Journal of Econometrics 42 (1989) 319–336. North-Holland

ESTIMATING INTEGRATED HIGHER-ORDER CONTINUOUS TIME AUTOREGRESSIONS WITH AN APPLICATION TO MONEY-INCOME CAUSALITY*

A.C. HARVEY

London School of Economics, London WC2A 2AE, Great Britain

James H. STOCK

Harvard University, Cambridge, MA 02138, USA

Received July 1986, final version received January 1989

An algorithm is presented for computing the Gaussian maximum likelihood estimator of the parameters of a multivariate continuous time autoregressive process with multiple roots that equal zero. Some of the variables might be observed at a point in time ('stocks') and some might be observed as an integral over a unit interval ('flows'). This algorithm, based on the Kalman–Bucy filter, is used to investigate the possibility that previous findings that money Granger-causes industrial production spuriously arose because time-averaged variables were analyzed using discrete time methods.

1. Introduction

Economic time series are rarely stationary, but often can be approximated by stationary autoregressive processes after they have been differenced; see, for example, Granger and Newbold (1977, pp. 176–179). The treatment of discrete time models with integrated processes has received enormous attention and the issues are in general well understood. However, handling the analogous continuous time models is more difficult, particularly if a variable corresponds to a flow or a time-averaged measurement of a stock. A number of papers have considered maximum likelihood estimation of stationary continuous time autoregressions; see, for example, Bergstrom (1976, 1983), Phillips (1959), Phadke and Wu (1974), Jones (1981), Harvey and Stock (1985), and Zadrozny (1988). In addition, Hansen and Sargent (1981) and Christiano, Eichenbaum, and Marshall (1987) discuss frequency domain approaches to

*The authors are grateful to Robert Hall and Thomas Sargent for helpful discussions, and to an anonymous referee, an associate editor, and especially Peter Zadrozny for thoughtful suggestions on an earlier draft. Harvey's research was supported in part by the ESRC grant, 'Multivariate Aspects of Structural Time Series Models', B00232155. Stock's research was supported in part by NSF grants SES-84-08797 and SES-86-18984.

handling rational expectations models with integrated processes in continuous time.

This paper sets out a general procedure for computing the exact maximum likelihood estimators of the parameters of a higher-order continuous time autoregression with stock and flow variables, when some of the variables are integrated of positive integer orders. Specifically, we present an algorithm, based on the Kalman–Bucy (1961) filter, for computing the likelihood in the case that the continuous time autoregression has multiple zero roots, corresponding to multiple unit roots in a discrete time autoregression. This algorithm can also be used to estimate systems with cointegrated variables if the cointegrating vectors are known *a priori*, as in the models considered by Hansen and Sargent (1981) and Christiano, Eichenbaum, and Marshall (1987), or if they are estimated in a previous stage. In addition, the algorithm can be extended to irregularly spaced data or missing observations using the techniques discussed in, for example, Harvey and Stock (1985). To simplify exposition, however, it is presented here for the case of equally spaced observations.

The algorithm for a multivariate process containing both stock and flow variables with orders of integration zero and one is presented in section 2. The key step in deriving the algorithm is expressing the observable variables in terms of unobservable unit integrals of the mean-square derivatives of the state variables that enter into the stochastic differential equation system. These integrals can be expressed in terms of lower-order integrals of the state variables, which are in turn included in an augmented state vector. The observable variables can then be expressed as linear combinations of the elements of this state vector; this permits the application of the Kalman filter. An alternative approach is to work in levels using the strategy suggested by Zadrozny (1988) with the initial conditions handled as in, for example, Ansley and Kohn (1985). An advantage of the approach presented here is that, when it applies, it offers substantial computational simplifications.

In section 3, we briefly describe how this approach can be extended to variables with higher orders of integration, illustrating this extension for a univariate doubly integrated process observed as a flow.

In section 4, we examine the possibility that previous findings that money helps to predict output might spuriously have arisen by estimating discrete approximations to continuous time models. Eichenbaum and Singleton (1986), Christiano and Ljungqvist (1988), and Stock and Watson (1989) provided evidence that the proper handling of the apparent unit roots and time trends in discrete time models of money and output is important in determining the significance of conventional Granger-causality test statistics. In addition, Christiano and Eichenbaum (1987) showed that even if money does not Granger-cause output at a fine time interval, after the series are aggregated to a coarser frequency money will nevertheless contain information that is useful in predicting future output. This raises the possibility that money might not help to predict output in a continuous time model, but might spuriously do so in discrete time models, either because temporally aggregated data are used or because the restrictions in the discrete time model differ from those in the underlying continuous time model, or both. This is investigated in section 4 using monthly U.S. data from 1960 to 1985. The results provide some support for the view that the predictive role of money in discrete time models might in part be a spurious result of using temporally aggregated data.

2. Multivariate models with integrated mixed stock/flows data

This section presents an algorithm for evaluating the exact Gaussian likelihood for multivariate data generated by a system of stochastic differential equations in which the continuous time system contains roots that equal zero, and these zero root conditions are imposed.[1] The variables can be observed either as stocks or as flows, and some or all of the variables might have a zero root in their continuous time autoregressive representations. An important special case of this algorithm is the computation of the likelihood for a univariate flow process with a single zero root, such as might be appropriate for a univariate model of GNP. However, the multivariate formulation is of additional practical importance in at least three contexts.

First, suppose that one of the variables is stationary while another of the variables is integrated. Then it might be desirable to impose this zero root so that the usual asymptotic theory (predicated upon stationary processes) applies. An example of this situation might be a bivariate model relating the unemployment rate to changes in GNP, with unemployment modeled as a stationary variable observed at a point in time (say, the last survey date before the end of the quarter) and GNP modeled as a flow with a single continuous time zero root.

Second, these variables might individually be integrated but jointly be cointegrated as defined by Engle and Granger (1987), so that the number of zero roots in the multivariate specification is less than the number of zero roots in the individual univariate specifications. This situation has been extensively studied in discrete time models [e.g., Engle and Granger (1987) and Dickey (1986)] and has been examined in continuous time rational expectations models by Hansen and Sargent (1981) and Christiano, Eichenbaum, and Marshall (1987). In both cases, the reduced number of zero roots in the multivariate model can be imposed by reformulating the system in terms of stationary linear combinations of contemporaneous variables and first differ-

[1]The procedures in Harvey and Stock (1985) and Bergstrom (1985) can handle special cases with roots that equal zero, although even then the zero root restrictions are difficult to impose.

ences of the remaining nonstationary variables, a specification referred to as a vector error correction model.

Third, two variables might individually be integrated processes and not be cointegrated, so that their bivariate representation contains two roots that equal zero. We will argue in section 4 that postwar U.S. monthly Industrial Production and M1 are two such variables.

2.1. The model and notation

Suppose that n observed variables are arranged in the vector Y_t so that $Y_t = [Y_t^{s0\prime}\ \Delta Y_t^{s1\prime}\ Y_t^{f0\prime}\ \Delta Y_t^{f1\prime}]'$, where Y_t^{sj} is a n_{sj} vector of stock variables that are integrated of order j, Y_t^{fj} is a n_{fj} vector of flow variables integrated of order j, $\Delta x_t \equiv x_t - x_{t-1}$, and $n = n_{s0} + n_{s1} + n_{f0} + n_{f1}$. We adopt the convention that Y_t refers to the observable discrete time variables and $y(t)$ refers to their continuous time counterpart. Let $\xi(t) = [y^{s0}(t)'\ \mathrm{D}y^{s1}(t)'\ y^{f0}(t)'\ \mathrm{D}y^{f1}(t)']'$. Then $\xi(t)$ is assumed to obey the system of pth-order stochastic differential equations,

$$\mathrm{dD}^{p-1}\xi(t) = \left[A_1\mathrm{D}^{p-1}\xi(t) + \cdots + A_p\xi(t) + \theta\right]\mathrm{d}t + \mathrm{d}\zeta^*(t), \quad (2.1)$$

where $A_1, \dots, A_p$ are $n \times n$ matrices, θ is a $n \times 1$ constant vector, and $\zeta^*(t)$ is a n-dimensional Wiener process with covariance matrix Σ_ζ. The unknown parameters of the system are contained in $A_1, \dots, A_p$, Σ_ζ, and θ. See Bergstrom (1983) for a discussion of higher-order stochastic differential equations in the case when there are no zero roots (i.e., $n_{s1} = n_{f1} = 0$).

If there are integrated variables in the system, the stochastic differential equation for $y(t)$ will contain zero roots and will be unstable. Attempting to estimate such a system directly using the approach of Jones (1981) and Harvey and Stock (1985) presents a practical problem since it requires the diagonalization of a state transition matrix with multiple zero roots. Moreover, the levels specification does not readily lend itself to imposing the zero roots. We address these difficulties by treating (2.1) as a stable system in the unobservable variables $\xi(t)$. For flow variables with a single zero root – the case considered in this section – the observable process will be a double integral of $\mathrm{D}y(t)$, plus correction terms given below.

The key idea is to formulate the discrete time state space model so that (i) the continuous time roots that equal zero are imposed in such a way that the discrete time state space model has no unit roots, and (ii) a linear combination of the state vector can be observed as a stationary transformation of the basic variables. The formulation (2.1) achieves the first objective. The second is accomplished by writing the measurement equation of the state space model in terms of first differences of the nonstationary variables. To this end, note that

$$\Delta Y_t^{\mathrm{fl}} = \int_{t-1}^{t} y^{\mathrm{fl}}(\tau)\,\mathrm{d}\tau - \int_{t-2}^{t-1} y^{\mathrm{fl}}(\tau)\,\mathrm{d}\tau, \tag{2.2a}$$

$$\Delta Y_t^{\mathrm{sl}} = y^{\mathrm{sl}}(t) - y^{\mathrm{sl}}(t-1). \tag{2.2b}$$

It is shown below that the variables $y^{\mathrm{sl}}(t)$ and $\int y^{\mathrm{fl}}(\tau)\,\mathrm{d}\tau$ in (2.2) can be expressed in terms of integrals of the elements of $\xi(t)$ and its lags.

2.2. State space formulation

It is convenient to rewrite (2.1) in first-order form,

$$\mathrm{d}\xi^*(t) = [A\xi^*(t) + R\theta]\,\mathrm{d}t + R\,\mathrm{d}\zeta^*(t), \tag{2.3}$$

where

$$\xi^*(t) = \begin{bmatrix} \xi(t) \\ \mathrm{D}\xi(t) \\ \vdots \\ \mathrm{D}^{p-1}\xi(t) \end{bmatrix}, \quad R = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I_p \end{bmatrix}, \quad A = \left[\begin{array}{c|ccc} 0 & & & \\ \vdots & & I & \\ 0 & & & \\ \hline A_p & A_{p-1} & \cdots & A_1 \end{array}\right].$$

Following Jones (1981) and Harvey and Stock (1985), we diagonalize the state transition matrix A. As is discussed below, this facilitates the explicit evaluation of integrals that appear in the discrete time transition matrix for the system (2.3) and reduces the number of computations in the Kalman filter recursions. Accordingly, assume A to have the eigenvalue decomposition $A = G\Lambda G^{-1}$, where the eigenvalues λ_i, $i = 1,\dots,np$, are distinct and have negative real parts. Let $\mu = G^{-1}R\theta$, $\mathrm{d}\zeta(t) = G^{-1}R\,\mathrm{d}\zeta^*(t)$, and $\alpha(t) = G^{-1}\xi^*(t)$. Then (2.3) can be rewritten as

$$\mathrm{d}\alpha(t) = [\Lambda\alpha(t) + \mu]\,\mathrm{d}t + \mathrm{d}\zeta(t). \tag{2.4}$$

For $\tau > t-1$, $\alpha(t)$ satisfies

$$\alpha(\tau) = \mathrm{e}^{\Lambda(\tau-(t-1))}\alpha(t-1) + \int_{t-1}^{\tau} \mathrm{e}^{\Lambda(\tau-s)}\mu\,\mathrm{d}s + \int_{t-1}^{\tau} \mathrm{e}^{\Lambda(\tau-s)}\,\mathrm{d}\zeta(s). \tag{2.5}$$

Set $\tau = t$ in (2.5), so that the evolution of the transformed state vector at unit intervals of (discrete) time is given by

$$\alpha_t = T\alpha_{t-1} + \gamma + \eta_t, \tag{2.6}$$

where $T = \mathrm{e}^{A}$, $\gamma = \int_{t-1}^{t} \mathrm{e}^{A(t-s)}\mu \, \mathrm{d}s = W^{[1]}\mu$ with $W^{[1]} = A^{-1}(T - I)$, and where $\eta_t = \int_{t-1}^{t} \mathrm{e}^{A(t-s)} \mathrm{d}\zeta(s)$.

To construct the augmented state vector, note that

$$Y_t^{\mathrm{f}0} = \int_{\tau=t-1}^{t} y^{\mathrm{f}0}(\tau)\,\mathrm{d}\tau, \tag{2.7a}$$

$$\Delta Y_t^{\mathrm{s}1} = y^{\mathrm{s}1}(t) - y^{\mathrm{s}1}(t-1) = \int_{\tau=t-1}^{t} \mathrm{D}y^{\mathrm{s}1}(\tau)\,\mathrm{d}\tau, \tag{2.7b}$$

$$y^{\mathrm{f}1}(t) - y^{\mathrm{f}1}(t-1) = \int_{\tau=t-1}^{t} \mathrm{D}y^{\mathrm{f}1}(\tau)\,\mathrm{d}\tau, \tag{2.7c}$$

$$\int_{\tau=t-1}^{t}\int_{s=t-1}^{\tau} \mathrm{D}y^{\mathrm{f}1}(s)\,\mathrm{d}s\,\mathrm{d}\tau = \int_{\tau=t-1}^{t} y^{\mathrm{f}1}(\tau)\,\mathrm{d}\tau - y^{\mathrm{f}1}(t-1). \tag{2.8}$$

Combining (2.2), (2.7c), and (2.8), one obtains

$$\begin{aligned}
\Delta Y_t^{\mathrm{f}1} &= \int_{\tau=t-1}^{t}\int_{s=t-1}^{\tau} \mathrm{D}y^{\mathrm{f}1}(s)\,\mathrm{d}s\,\mathrm{d}\tau + y^{\mathrm{f}1}(t-1) \\
&\quad - \int_{\tau=t-2}^{t-1}\int_{s=t-2}^{\tau} \mathrm{D}y^{\mathrm{f}1}(s)\,\mathrm{d}s\,\mathrm{d}\tau - y^{\mathrm{f}1}(t-2) \\
&= \int_{\tau=t-1}^{t}\int_{s=t-1}^{\tau} \mathrm{D}y^{\mathrm{f}1}(s)\,\mathrm{d}s\,\mathrm{d}\tau - \int_{\tau=t-2}^{t-1}\int_{s=t-2}^{\tau} \mathrm{D}y^{\mathrm{f}1}(s)\,\mathrm{d}s\,\mathrm{d}\tau \\
&\quad + \int_{\tau=t-2}^{t-1} \mathrm{D}y^{\mathrm{f}1}(\tau)\,\mathrm{d}\tau.
\end{aligned} \tag{2.9}$$

Thus (2.7a), (2.7b), and (2.9) express stationary transformations of the observable discrete time variables in terms of integrals of elements of $\xi(t)$.

The integrals in (2.7) and (2.9) can be related to $\alpha(t)$ by defining two selection matrices: a $(n_{\mathrm{f}0} + n_{\mathrm{s}1} + n_{\mathrm{f}1}) \times np$ matrix

$$Z^{[1]} = \begin{bmatrix} 0_{n_{\mathrm{s}0}} & I_{n_{\mathrm{s}1}+n_{\mathrm{f}0}+n_{\mathrm{f}1}} & 0_{n(p-1)} \end{bmatrix} G,$$

where 0_m is a conformable matrix of zeros with m columns and I_m denotes the

$m \times m$ identity matrix, and a $n_{\mathrm{f1}} \times (n_{\mathrm{s1}} + n_{\mathrm{f0}} + n_{\mathrm{f1}})$ matrix

$$Z^{[2]} = \begin{bmatrix} 0_{n_{\mathrm{s1}} + n_{\mathrm{f0}}} & I_{n_{\mathrm{f1}}} \end{bmatrix}.$$

Thus

$$Z^{[1]}\alpha(t) = \begin{bmatrix} \mathrm{D}y^{\mathrm{s1}}(t)' & y^{\mathrm{f0}}(t)' & \mathrm{D}y^{\mathrm{f1}}(t)' \end{bmatrix}'$$

and $Z^{[2]}Z^{[1]}\alpha(t) = \mathrm{D}y^{\mathrm{f1}}(t)$. Also define the first and second integrals of the integrated linear combinations of $\alpha(t)$ to be

$$\alpha^{[1]}(\tau) = Z^{[1]}\int_{s=t-1}^{\tau} \alpha(s)\,\mathrm{d}s, \tag{2.10a}$$

$$\begin{aligned} \alpha_t^{[2]} &= Z^{[2]}Z^{[1]}\int_{\tau=t-1}^{t}\int_{s=t-1}^{\tau} \alpha(s)\,\mathrm{d}s\,\mathrm{d}\tau \\ &= Z^{[2]}\int_{\tau=t-1}^{t} \alpha^{[1]}(\tau)\,\mathrm{d}\tau, \end{aligned} \tag{2.10b}$$

and let $\alpha_t^{[1]} \equiv \alpha^{[1]}(t)$. The definitions (2.10) and the eqs. (2.7)–(2.9) permit rewriting the stationary transformations of the integrated variables in terms of $\alpha(t)$, $\alpha_t^{[1]}$, $\alpha_t^{[2]}$, and their first lags:

$$\Delta Y_t^{\mathrm{s1}} = \begin{pmatrix} I_{n_{\mathrm{s1}}} & 0 & 0 \end{pmatrix}\alpha_t^{[1]},$$

$$Y_t^{\mathrm{f0}} = \begin{pmatrix} 0 & I_{n_{\mathrm{f0}}} & 0 \end{pmatrix}\alpha_t^{[1]},$$

$$\Delta Y_t^{\mathrm{f1}} = \Delta\alpha_t^{[2]} + Z^{[2]}\alpha_{t-1}^{[1]}.$$

The augmented state vector of the discrete time system is $\alpha_t^{\dagger} = [\alpha_t' \; \alpha_t^{[1]\prime} \; \alpha_t^{[2]\prime} \; \alpha_{t-1}^{[1]\prime} \; \alpha_{t-1}^{[2]\prime}]'$. The transition equation for α_t is given by (2.6). To obtain the transition equation for $\alpha_t^{[1]}$, substitute (2.5) into (2.10a):

$$\begin{aligned} \alpha^{[1]}(\tau) = Z^{[1]}&\int_{s=t-1}^{\tau} \mathrm{e}^{A(s-(t-1))}\alpha(t-1)\,\mathrm{d}s \\ &+ Z^{[1]}\int_{s=t-1}^{\tau}\int_{r=t-1}^{s} \mathrm{e}^{A(s-r)}\mu\,\mathrm{d}r\,\mathrm{d}s \\ &+ Z^{[1]}\int_{s=t-1}^{\tau}\int_{r=t-1}^{s} \mathrm{e}^{A(s-r)}\mathrm{d}\zeta(r)\,\mathrm{d}s. \end{aligned} \tag{2.11}$$

Thus, letting $\tau = t$, one obtains

$$\alpha_t^{[1]} = Z^{[1]} W^{[1]} \alpha_{t-1} + \gamma^{[1]} + \eta_t^{[1]}, \tag{2.12}$$

where $W^{[1]} = \Lambda^{-1}(T - I)$, $\gamma^{[1]} = Z^{[1]} W^{[2]} \mu$ with $W^{[2]} = \Lambda^{-1}(W^{[1]} - I)$, and where $\eta_t^{[1]} = Z^{[1]} \int_{s=t-1}^{t} \int_{r=t-1}^{s} e^{\Lambda(s-r)} d\zeta(r)\, ds$.

The transition equation for $\alpha_t^{[2]}$ is found by integrating (2.11) from $t-1$ to t:

$$\begin{aligned}\alpha_t^{[2]} = {} & Z^{[2]} Z^{[1]} \int_{\tau=t-1}^{t} \int_{s=t-1}^{\tau} e^{\Lambda(s-(t-1))} \alpha(t-1)\, ds\, d\tau \\ & + Z^{[2]} Z^{[1]} \int_{\tau=t-1}^{t} \int_{s=t-1}^{\tau} \int_{r=t-1}^{s} e^{\Lambda(s-r)} \mu\, dr\, ds\, d\tau \\ & + Z^{[2]} Z^{[1]} \int_{\tau=t-1}^{t} \int_{s=t-1}^{\tau} \int_{r=t-1}^{s} e^{\Lambda(s-r)} d\zeta(r)\, ds\, d\tau. \end{aligned} \tag{2.13}$$

Rewriting (2.13) and evaluating the first two sets of integrals, we have

$$\alpha_t^{[2]} = Z^{[2]} Z^{[1]} W^{[2]} \alpha_{t-1} + \gamma^{[2]} + \eta_t^{[2]}, \tag{2.14}$$

where $\gamma^{[2]} = Z^{[2]} Z^{[1]} W^{[3]} \mu$ with $W^{[3]} = \Lambda^{-1}(W^{[2]} - \frac{1}{2} I)$, and where $\eta_t^{[2]} = Z^{[2]} Z^{[1]} \int_{\tau=t-1}^{t} \int_{s=t-1}^{\tau} \int_{r=t-1}^{s} e^{\Lambda(s-r)} d\zeta(r)\, ds\, d\tau$. Note that $\eta_t^{[1]}$ and $\eta_t^{[2]}$ involve continuous time disturbances occurring only between $t-1$ and t.

Combining (2.6), (2.12), and (2.14), one obtains the augmented state vector and its discrete time transition equation:

$$\alpha_t^\dagger = T^\dagger \alpha_{t-1}^\dagger + R^\dagger \gamma^\dagger + R^\dagger \eta_t^\dagger, \tag{2.15}$$

where

$$\alpha_t^\dagger = \begin{bmatrix} \alpha_t \\ \alpha_t^{[1]} \\ \alpha_t^{[2]} \\ \alpha_{t-1}^{[1]} \\ \alpha_{t-1}^{[2]} \end{bmatrix}, \quad T^\dagger = \begin{bmatrix} T & 0 & 0 & 0 & 0 \\ Z^{[1]} W^{[1]} & 0 & 0 & 0 & 0 \\ Z^{[2]} Z^{[1]} W^{[2]} & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \end{bmatrix},$$

$$R^\dagger = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \gamma^\dagger = \begin{bmatrix} \gamma \\ \gamma^{[1]} \\ \gamma^{[2]} \end{bmatrix}, \quad \eta_t^\dagger = \begin{bmatrix} \eta_t \\ \eta_t^{[1]} \\ \eta_t^{[2]} \end{bmatrix}.$$

The measurement equation, which relates the stationary transformations of the observed variables to the elements of the augmented state vector, is obtained by combining (2.2) and the expressions that follow (2.10b). Thus:

$$\begin{bmatrix} Y_t^{s0} \\ \Delta Y_t^{s1} \\ Y_t^{f0} \\ \Delta Y_t^{f1} \end{bmatrix} = \begin{bmatrix} (I_{n_{s0}}\ 0)G & 0 & 0 & 0 & 0 \\ 0 & (I_{n_{s1}}\ 0\ 0) & 0 & 0 & 0 \\ 0 & (0\ I_{n_{f0}}\ 0) & 0 & 0 & 0 \\ 0 & 0 & I_{n_{f1}} & Z^{[2]} & -I_{n_{f1}} \end{bmatrix} \begin{bmatrix} \alpha_t \\ \alpha_t^{[1]} \\ \alpha_t^{[2]} \\ \alpha_{t-1}^{[1]} \\ \alpha_{t-1}^{[2]} \end{bmatrix}, \tag{2.16}$$

where the selection matrix is partitioned conformably with $\alpha_t^\dagger$. The measurement equation (2.16) can be written more compactly as

$$Y_t = Z^\dagger \alpha_t^\dagger. \tag{2.17}$$

Note that ΔY_t^{s1} and Y_t^{f0} are treated symmetrically in the measurement equation.[2]

2.3. *The Kalman filter*

The Kalman filter is given by the usual set of equations for a complex-valued state vector. Let $Q = \mathrm{E}\eta_t^\dagger \bar{\eta}_t^{\dagger\prime}$, where $\bar{x}'$ denotes the complex-conjugate transpose of x. Also let $a_{t|t-1}^\dagger$ and $P_{t|t-1}^\dagger$, respectively, be the prediction of the state vector made at $t-1$ and its covariance matrix, and let $a_t^\dagger$ and $P_t^\dagger$, respectively, denote the filtered or predicted values of α_t at time t and its covariance matrix. With this notation, the prediction equations of the Kalman filter are

$$a_{t|t-1}^\dagger = T^\dagger a_{t-1}^\dagger + R^\dagger \gamma^\dagger,$$

$$P_{t|t-1}^\dagger = T^\dagger P_{t-1}^\dagger \bar{T}^{\dagger\prime} + R^\dagger Q R^{\dagger\prime}.$$

The forecast errors are $\nu_t = Y_t - Z^\dagger a_{t|t-1}^\dagger$, and the updating equations are

$$a_t^\dagger = a_{t|t-1}^\dagger + P_{t|t-1}^\dagger \bar{Z}^{\dagger\prime} F_t^{-1} \nu_t,$$

$$P_t^\dagger = P_{t|t-1}^\dagger - P_{t|t-1}^\dagger \bar{Z}^{\dagger\prime} F_t^{-1} Z^\dagger P_{t|t-1}^\dagger,$$

[2]A Gaussian measurement error could be added to the measurement equations, (2.16). Such a formulation would, however, imply that the measurement errors are introduced differently depending on the order of integration of the variable.

where F_t, the covariance matrix of ν_t, is given by $F_t = Z^{\dagger}P^{\dagger}_{t|t-1}\bar{Z}^{\dagger\prime}$. Estimation proceeds by maximizing the Gaussian likelihood function as discussed in Jones (1981) and Harvey and Stock (1985).

2.4. Calculation of disturbance covariances

The Kalman filter equations depend on the covariance matrix Q, which consists of nonstochastic integrals over $t-1$ to t. Let

$$Q = \begin{bmatrix} Q^{00} & Q^{01} & Q^{02} \\ \overline{Q^{01\prime}} & Q^{11} & Q^{12} \\ \overline{Q^{02\prime}} & \overline{Q^{12\prime}} & Q^{22} \end{bmatrix},$$

where $Q^{00} = \mathrm{E}\eta_t\bar{\eta}_t'$, $Q^{0j} = \mathrm{E}\eta_t\bar{\eta}_t^{[j]\prime}$ for $j = 1,2$, and $Q^{ij} = \mathrm{E}\eta_t^{[i]}\bar{\eta}_t^{[j]\prime}$ for $i, j = 1,2$. To calculate the integrals comprising Q^{ij}, we adopt the heuristic device of first rewriting the innovations $\eta_t^{[i]}$ as single weighted integrals of $\mathrm{d}\zeta(s)$ from $t-1$ to t. In the case of $\eta_t^{[1]}$,

$$\begin{aligned} \eta_t^{[1]} &= Z^{[1]}\int_{s=t-1}^{t}\int_{r=t-1}^{s} \mathrm{e}^{A(s-r)}\mathrm{d}\zeta(r)\,\mathrm{d}s \\ &= Z^{[1]}\int_{r=t-1}^{t}\int_{s=r}^{t} \mathrm{e}^{A(s-r)}\mathrm{d}s\,\mathrm{d}\zeta(r) \\ &= Z^{[1]}\int_{r=t-1}^{t} A^{-1}[\mathrm{e}^{A(t-r)} - I]\,\mathrm{d}\zeta(r). \end{aligned}$$

A similar calculation applies to $\eta_t^{[2]}$. Thus:

$$\eta_t^{[1]} = Z^{[1]}A^{-1}v_t^{[0]} - Z^{[1]}A^{-1}v_t^{[1]}, \tag{2.18a}$$

$$\eta_t^{[2]} = Z^{[2]}Z^{[1]}A^{-2}v_t^{[0]} - Z^{[2]}Z^{[1]}A^{-2}v_t^{[1]} - Z^{[2]}Z^{[1]}A^{-1}v_t^{[2]}, \tag{2.18b}$$

where $v_t^{[0]} = \eta_t = \int_{t-1}^{t}\mathrm{e}^{A(t-s)}\mathrm{d}\zeta(s)$, $v_t^{[1]} = \int_{t-1}^{t}\mathrm{d}\zeta(s)$ and $v_t^{[2]} = \int_{t-1}^{t}(t-s)\,\mathrm{d}\zeta(s)$. Thus $\eta_t^{\dagger} = S^{\dagger}v_t^{\dagger}$, where $v_t^{\dagger} = [v_t^{[0]\prime}\ v_t^{[1]\prime}\ v_t^{[2]\prime}]'$ and where

$$S^{\dagger} = \begin{bmatrix} I & 0 & 0 \\ Z^{[1]}A^{-1} & -Z^{[1]}A^{-1} & 0 \\ Z^{[2]}Z^{[1]}A^{-2} & -Z^{[2]}Z^{[1]}A^{-2} & -Z^{[2]}Z^{[1]}A^{-1} \end{bmatrix}. \tag{2.19}$$

It follows that $Q = S^{\dagger}(\mathrm{E}v_t^{\dagger}\bar{v}_t^{\dagger\prime})\bar{S}^{\dagger\prime}$, where

$$\mathrm{E}v_t^{\dagger}\bar{v}_t^{\dagger\prime} = \begin{bmatrix} Q^{00} & W^{[1]}\Omega & \Lambda^{-1}(T - W^{[1]})\Omega \\ \overline{(W^{[1]}\Omega)'} & \Omega & \Omega/2 \\ \overline{[\Lambda^{-1}(T - W^{[1]})\Omega]'} & \bar{\Omega}'/2 & \Omega/3 \end{bmatrix}, \tag{2.20}$$

where $\Omega = G^{-1}R\Sigma_{\zeta}R'\bar{G}^{-1}$ and where the (i, j) element of the $p \times p$ covariance matrix Q^{00} is given by $Q_{ij}^{00} = \Omega_{ij}[\exp(\lambda_i + \bar{\lambda}_j) - 1]/(\lambda_i + \bar{\lambda}_j)$.

2.5. *Initial conditions*

Because the system (2.1) is assumed to be stable, the initial conditions for the Kalman filter are the unconditional expectation of $\alpha_t^{\dagger}$ and its unconditional covariance matrix. These are given by

$$\alpha_{1|0}^{\dagger} = \sum_{k=0}^{\infty} (T^{\dagger})^k R^{\dagger}\gamma^{\dagger} = (I - T^{\dagger})^{-1} R^{\dagger}\gamma^{\dagger}, \tag{2.21a}$$

$$P_{1|0}^{\dagger} = \sum_{k=0}^{\infty} (T^{\dagger})^k R^{\dagger} Q R^{\dagger\prime} (\bar{T}^{\dagger\prime})^k. \tag{2.21b}$$

Because the roots of $T^{\dagger}$ have modulus less than one, $I - T^{\dagger}$ is invertible and the sum in (2.21b) is finite. The sum in (2.21b) (and the corresponding Lyapunov equation) can be simplified considerably because $T^{\dagger}$ is sparse. For example, the block of $P_{1|0}^{\dagger}$ corresponding to α_t has the (i, j) element, $(P_{1|0}^{\dagger 00})_{ij} = -\Omega_{ij}/(\lambda_i + \bar{\lambda}_j)$, $i, j = 1, \dots, np$.

2.6. *Extensions and alternatives*

This approach is readily extended to include additional deterministic or exogenous variables by replacing θ in (2.1) with a time-varying function of a finite number of unknown parameters, say $X(t; \theta)$, where θ is a parameter vector. In this case integrals of these exogenous or deterministic variables [such as appear starting with (2.5) for the special case that $G^{-1}RX(t; \theta) = G^{-1}R\theta = \mu$] enter the discrete time state transition equation. When $X(t; \theta)$ is a deterministic function of time, these integrals are typically straightforward (if tedious) to calculate. Bergstrom (1976, 1983) and the references therein discuss various approximations that can be made when $X(t; \theta)$ is an exogenous random process.

The Kalman filter for the augmented state vector has been developed here using the eigenvector/eigenvalue decomposition of the continuous time state transition matrix. This facilitates obtaining explicit representations for e^A as $G\,e^{\Lambda}G^{-1}$, for the discrete time innovation covariance matrix Q, and for $P_{1|0}^{\dagger 00}$. An alternative approach is to develop the filter for the system (2.3) without a preliminary linear transformation, in which case all matrices and variables are real-valued. The derivation parallels the one given here. This formulation makes it possible to adopt alternative methods for computing the matrix exponentials (and their integrals) that do not involve eigenvalue decompositions; a variety of techniques for this purpose are examined by Van Loan (1978). For further discussions and linear filters formulated along these lines, see Melino (1985), Harvey and Stock (1988), and Zadrozny (1988).

3. Extension to higher orders of integration

The approach of section 2 can be extended to the case that one or more of the variables is integrated of integer order d. This entails (i) defining the multiple integrals of the state vector $\alpha_t^{[d+1]}$ (for flows) and $\alpha_t^{[d]}$ (for stocks) and (ii) writing the measurement equation in terms of the stationary discrete time process $\Delta^d Y_t$.

We illustrate the approach for a univariate doubly integrated flow variable, so that the second difference of Y_t is stationary and $y(t)$ satisfies

$$\mathrm{d}D^{p-1+d}y(t)=\left[a_1\mathrm{D}^{p-1+d}y(t)+\cdots+a_p\mathrm{D}^{d}y(t)+\theta\right]\mathrm{d}t+\mathrm{d}\zeta^*(t), \tag{3.1}$$

with $d=2$. Let $\xi^*(t)=[\mathrm{D}^2y(t)\ldots\mathrm{D}^{p+1}y(t)]'$ and let A denote the continuous time transition matrix for the first-order system based on $\xi^*(t)$ as follows (2.3). As in section 2, suppose that A has the eigenvalue decomposition $A=G\Lambda G^{-1}$ (with no zero roots by assumption) and define $\alpha(t)=G^{-1}\xi^*(t)$. To derive the measurement equation, let $Z^*=[1\,0\ldots0]G$ and define $\alpha_t^{[1]}=Z^*\int_{t-1}^{t}\alpha(\tau)\,\mathrm{d}\tau$, $\alpha_t^{[2]}=Z^*\int_{\tau=t-1}^{t}\int_{s=t-1}^{\tau}\alpha(s)\,\mathrm{d}s\,\mathrm{d}\tau$, and $\alpha_t^{[3]}=Z^*\int_{\tau=t-1}^{t}\int_{s=t-1}^{\tau}\int_{r=t-1}^{s}\alpha(r)\,\mathrm{d}r\,\mathrm{d}s\,\mathrm{d}\tau$. Because $Z^*\alpha(\tau)=\mathrm{D}^2y(\tau)$, these integrals reduce to

$$\alpha_t^{[1]}=\mathrm{D}y(t)-\mathrm{D}y(t-1), \tag{3.2a}$$

$$\alpha_t^{[2]}=y(t)-y(t-1)-\mathrm{D}y(t-1), \tag{3.2b}$$

$$\alpha_t^{[3]}=\int_{t-1}^{t}y(\tau)\,\mathrm{d}\tau-y(t-1)-\tfrac{1}{2}\mathrm{D}y(t-1). \tag{3.2c}$$

Next solve (3.2) for the stationary transformation of the observed flow variable:

$$\Delta^2 Y_t = \Delta^2 \int_{t-1}^{t} y(\tau)\,\mathrm{d}\tau = \Delta^2 \alpha_t^{[3]} + \Delta \alpha_{t-1}^{[2]} + \tfrac{1}{2}\left(\alpha_{t-1}^{[1]} + \alpha_{t-2}^{[1]}\right). \tag{3.3}$$

The augmented state vector $\alpha_t^{\dagger}$ here consists of α_t and of values at times t, $t-1$, and $t-2$ of $\alpha_t^{[1]}$, $\alpha_t^{[2]}$, and $\alpha_t^{[3]}$. The measurement equation relates these elements of the state vector to the (stationary) second difference of the observed flow variable according to (3.3).

The discrete time state space model for $\alpha_t^{\dagger}$ is obtained by augmenting the transition equations for α_t, $\alpha_t^{[1]}$, and $\alpha_t^{[2]}$ in section 2 (modified to use Z^* rather than $Z^{[1]}$ and $Z^{[2]}Z^{[1]}$) with the similarly obtained transition equation for $\alpha_t^{[3]}$:

$$\alpha_t^{[3]} = Z^* W^{[3]} \alpha_{t-1} + \gamma^{[3]} + \eta_t^{[3]}, \tag{3.4}$$

where $\gamma^{[3]} = Z^* W^{[4]} \mu$ and $\eta_t^{[3]} = Z^* \int_{t-1}^{t} \int_{t-1}^{\tau} \int_{t-1}^{s} \int_{t-1}^{r} \mathrm{e}^{\Lambda(r-v)}\,\mathrm{d}\zeta(v)\,\mathrm{d}r\,\mathrm{d}s\,\mathrm{d}\tau$. The matrices $W^{[3]}$ and $W^{[4]}$ are obtained from the recursion $W^{[j]} = \Lambda^{-1}(W^{[j-1]} - I/(j-1)!)$ for $j \geq 2$.[3] Finally, the additional elements of the error covariance matrix can be computed by expressing $\eta_t^{[3]}$ as a sum of single weighted integrals over $\mathrm{d}\zeta(\tau)$, as was done in (2.18) for $\eta_t^{[1]}$ and $\eta_t^{[2]}$. The initial conditions are obtained using (2.21).

[3] To prove the stated recursion for $W^{[j]}$, let

$$W^{[j]}(\tau) = \int_{u_j=0}^{\tau} \int_{u_{j-1}=0}^{u_j} \cdots \int_{u_1=0}^{u_2} \mathrm{e}^{\Lambda u_1}\,\mathrm{d}u_1 \cdots \mathrm{d}u_j,$$

and note that $W^{[j]} = W^{[j]}(1)$ and that $W^{[j]}(\tau) = \int_{u=0}^{\tau} W^{[j-1]}(u)\,\mathrm{d}u$. Suppose that $W^{[j]}(\tau) = \Lambda^{-1}[W^{[j-1]}(\tau) - I\tau^{j-1}/(j-1)!]$. Then

$$\begin{aligned} W^{[j+1]}(\tau) &= \int_0^{\tau} W^{[j]}(u)\,\mathrm{d}u \\ &= \int_0^{\tau} \Lambda^{-1}\left[W^{[j-1]}(u) - Iu^{j-1}/(j-1)!\right]\mathrm{d}u \\ &= \Lambda^{-1}\left[\int_0^{\tau} W^{[j-1]}(u)\,\mathrm{d}u - I\int_0^{\tau} u^{j-1}\,\mathrm{d}u/(j-1)!\right] \\ &= \Lambda^{-1}\left[W^{[j]}(\tau) - I\tau^j/j!\right]. \end{aligned}$$

The recursion for $W^{[j+1]}(\tau)$ (and thus for $W^{[j+1]}$) therefore holds for all $j \geq 2$ if it can be shown for $j = 2$, which is readily shown by direct calculation.

4. Temporal aggregation and money as a predictor of output

Work of Sims (1972, 1980a, 1980b) and others suggests that money helps to predict output in a bivariate relation, i.e., money enters significantly into the equation for output in a discrete time bivariate vector autoregression (VAR). Two methodological arguments have recently been raised to suggest that these results might be misleading. First, Bernanke (1986), Eichenbaum and Singleton (1986), and Runkle (1987) have pointed out that finding Granger-causality in discrete time models is sensitive to how the apparent unit roots in these series are handled and to whether deterministic time trends are included in the specification. This raises the possibility that the inference that money is a useful predictor of future output – which is based on conventional asymptotic distribution theory – might be overturned by a more careful examination of the 'unit root' properties of the series and the associated nonstandard asymptotic distribution theory. Second, Christiano and Eichenbaum (1987) have suggested that perceived money–output causality might be the spurious product of another econometric pitfall, the use of time-averaged data at a coarser frequency than the true economic relationship. To address the first of these arguments, Stock and Watson (1989) examined the unit root and time trend properties of money and output and concluded that the deviation of money growth from a small but statistically significant time trend Granger-causes the growth of industrial production. Christiano and Ljungqvist (1988) reached a similar conclusion using a different empirical approach. But the possibility remains that this result spuriously reflects the use of temporally aggregated variables in a discrete time approximation to an underlying continuous time relation in which money does not enter the reduced-form equation for output.

We examine this possibility by estimating a continuous time bivariate VAR(2) of the form (2.1), in which zero roots are imposed in both the money and output equations. The continuous time analog of the discrete time hypothesis of Granger-noncausality is examined by considering the restriction that derivatives of money do not appear in the output equation, that is, that the bivariate system (2.1) is recursive with output ordered first. The data are monthly observations from January 1959 to December 1985 on M1 and the Index of Industrial Production (IP), both taken from the Citibase data base. Since the M1 data are monthly averages of daily figures and the IP index is a measure of output produced during the month, we treat both series as flows.

Previous research on money–output causality has employed logarithms of these flow data. However, both frequency domain procedures and the Kalman filter algorithm developed above handle flows by specifying the variables in levels rather than logarithms. This poses a conflict, since a logarithmic transformation conveniently handles the heteroskedasticity and the deterministic components of the trend in aggregate variables. We take two approaches to

Table 1

Estimated continuous time bivariate M1 industrial production system; exponentially detrended data; 1960:1–1985:12.[a]

A. Unrestricted system

$$\mathrm{d}\begin{bmatrix} \mathrm{D}^2 y(t) \\ \mathrm{D}^2 m(t) \end{bmatrix} = \left\{ \begin{bmatrix} 1.07 \\ 0.01 \end{bmatrix} + \begin{bmatrix} -32.77 & -31.51 \\ -1.36 & -3.25 \end{bmatrix} \begin{bmatrix} \mathrm{D}^2 y(t) \\ \mathrm{D}^2 m(t) \end{bmatrix} + \begin{bmatrix} -108.25 & -14.67 \\ -3.93 & -7.22 \end{bmatrix} \begin{bmatrix} \mathrm{D} y(t) \\ \mathrm{D} m(t) \end{bmatrix} \right\} \mathrm{d}t + \begin{bmatrix} \mathrm{d}\zeta_1(t) \\ \mathrm{d}\zeta_2(t) \end{bmatrix}$$

B. Comparison of restricted and unrestricted systems

	Unrestricted	Restricted
Roots of discrete time transition matrix $\exp(A)$	$-0.281 \pm 0.238i$	$-0.254 \pm 0.231i$
	0.033	0.015
	0.000	0.000
SEE: Δy_t	0.807	0.813
Δm_t	0.381	0.381
Log likelihood	-249.20	-252.17

[a] The coefficients were estimated by maximum likelihood, with the likelihood function computed using the Kalman filter algorithm described in section 2. The continuous time transition matrix A is defined following (2.3). In the restricted estimation, the two coefficients on the derivatives of m entering the equation for y are set to zero. The method of exponential detrending is described in the text. The roots stated as 0.000 are positive but rounded to three decimals.

this problem. The first approach uses multiplicative exponential detrending. Specifically, the estimated parameters from the regression equation $\ln X_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + u_t$ are used to construct a detrended series, $x_t = X_t \exp(-\hat{\alpha}_1 t - \hat{\alpha}_2 t^2)$, where X_t refers to M1 and IP.[4] The continuous time model is then estimated using x_t as data. Since Stock and Watson (1989) found that money and output are well described as being integrated but not cointegrated, first differences were used in the measurement equation to impose the zero continuous time roots. Let m_t and y_t denote the detrended variables associated with x_t; then in the notation of section 2 the variables entering the measurement equation are Δy_t and Δm_t, with $n = n_{f1} = 2$. In the second approach, the continuous time VAR's are estimated using first differences of the logarithms of the data; these growth rates were linearly detrended (using OLS) prior to estimation in the continuous time model.

[4] The t^2 term was included because Stock and Watson (1989) found evidence of a deterministic time trend in M1 growth. An argument based on orders of magnitudes of the regressors and the error term shows that the OLS estimators of $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are consistent even if $\ln X_t$ contains a discrete time unit root. Christiano, Eichenbaum, and Marshall (1987) estimate these deterministic trends imposing the restriction that these coefficients are the same in both their series (consumption and income measures). However, there is no theoretical or empirical reason to believe that these deterministic components are the same in the case of M1 and IP.

Table 2

Estimated continuous time bivariate M1 industrial production system; linearly detrended growth rates; 1960:1–1985:12.[a]

A. Unrestricted system

$$\mathrm{d}\begin{bmatrix} \mathrm{D}^2 y(t) \\ \mathrm{D}^2 m(t) \end{bmatrix} = \left\{ \begin{bmatrix} 0.002 \\ -0.001 \end{bmatrix} + \begin{bmatrix} -17.30 & -12.86 \\ 0.04 & -2.97 \end{bmatrix} \begin{bmatrix} \mathrm{D}^2 y(t) \\ \mathrm{D}^2 m(t) \end{bmatrix} + \begin{bmatrix} -40.24 & 1.54 \\ 0.49 & -10.28 \end{bmatrix} \begin{bmatrix} \mathrm{D} y(t) \\ \mathrm{D} m(t) \end{bmatrix} \right\} \mathrm{d}t + \begin{bmatrix} \mathrm{d}\zeta_1(t) \\ \mathrm{d}\zeta_2(t) \end{bmatrix}$$

B. Comparison of restricted and unrestricted systems

	Unrestricted	Restricted
Roots of discrete time transition matrix $\exp(A)$	$-0.199 \pm 0.037i$ 0.081 0.000	$-0.053 \pm 0.046i$ $-0.008 \pm 0.008i$
SEE: Δy_t	0.0082	0.0083
Δm_t	0.0038	0.0039
Log likelihood	2615.29	2612.67

[a]Before estimation the data were transformed by taking the first differences of their logarithms, then computing the residuals from a regression of this series against a constant and a linear time trend. See the notes to table 1.

The autoregressive coefficients of the unrestricted bivariate continuous time VAR(2) based on the exponentially detrended data are presented in table 1.[5] The table also contains summary statistics comparing the restricted and the unrestricted systems, where the restriction is that derivatives of m do not enter the equation for y. Both systems exhibit only modest serial dependence, with a pair of complex conjugate roots in their 4×4 discrete time transition matrix $\exp(A)$ and with two roots close to zero. The corresponding results for the linearly detrended growth rates data are given in table 2. Again, both systems exhibit small serial dependence.

Discrete and continuous time VAR's with money and output are compared in table 3. Using either the exponentially detrended data (panel A) or the linearly detrended growth rates (panel B), in all cases the hypothesis of Granger-noncausality is rejected at the 10% level. However, the marginal significance levels are greatest for the continuous time models, with neither continuous time specification rejecting noncausality at the 5% level. Using either data set, the in-sample predictive performance of the continuous and discrete models are comparable, with m being predicted somewhat more

[5]The likelihood function was maximized using a combination of the DFP and steepest descent maximization algorithms. The programs were written in the programming language MATLAB.

Table 3

Comparison of discrete and continuous time results; 1960:1–1985:12.[a]

A. Exponentially detrended levels specification

		Granger causality test: p-value	SEE ($\times 10^{-2}$)	
			Δy_t	Δm_t
Discrete time models	VAR(2)	0.0019	0.801	0.400
	VAR(4)	0.0001	0.787	0.391
	VAR(6)	0.0000	0.779	0.389
	VAR(12)	0.0016	0.785	0.382
Continuous time VAR(2)		0.0515	0.807	0.381

B. Linearly detrended log differences specification

		Granger causality test: p-value	SEE ($\times 10^{-2}$)	
			Δy_t	Δm_t
Discrete time models	VAR(2)	0.0027	0.812	0.402
	VAR(4)	0.0002	0.800	0.393
	VAR(6)	0.0001	0.793	0.390
	VAR(12)	0.0045	0.801	0.384
Continuous time VAR(2)		0.0728	0.823	0.384

[a]The discrete time systems are VAR's estimated using Δy_t and Δm_t, including an intercept term; observations prior to 1960 were used as initial conditions. The continuous time result in panel A is for the VAR(2)'s reported in table 1. The corresponding result in panel B is for the VAR(2)'s in table 2. The SEE's refer to the unrestricted systems. The p-values (i.e., marginal significance levels) for the Granger causality tests were computed using the usual F distribution for the discrete time VAR's and using the asymptotic χ^2 distribution of the likelihood ratio statistic in the continuous time models.

precisely and y somewhat less in the continuous time specification. Finally, comparing the results across models indicates that the use of the logarithmic or levels specification is less important than the choice of lag length and the use of a continuous or a discrete time specification.

In summary, our findings lend some empirical support for the view that tests of predictive content may be affected by temporal aggregation. In the continuous time models of money and output, however, the hypothesis that money is not on the margin a useful predictor of output is still rejected at the 10% level of significance.

References

Ansley, C.F. and R. Kohn, 1985, Estimation, filtering and smoothing in state space models with incompletely specified initial conditions, Annals of Statistics 13, 1286–1316.

Bergstrom, A.R., ed., 1976, Statistical inference in continuous time models (North-Holland, Amsterdam).

Bergstrom, A.R., 1983, Gaussian estimation of structural parameters in higher order continuous time dynamic models, Econometrica 51, 117–152.

Bergstrom, A.R., 1985, The estimation of parameters in nonstationary higher-order continuous time dynamic models, Econometric Theory 1, 369–385.

Bernanke, B.S., 1986, Alternative explanations of the money-income correlation, Carnegie–Rochester Conference on Public Policy 25, 49–100.

Christiano, L.J. and M. Eichenbaum, 1987, Temporal aggregation and structural inference in macroeconomics, Carnegie–Rochester Conference on Public Policy 26, 63–130.

Christiano, L.J., M. Eichenbaum, and D. Marshall, 1987, The permanent income hypothesis revisited, Manuscript (Federal Reserve Bank of Minneapolis, Minneapolis, MN).

Christiano, L.J. and L. Ljungqvist, 1988, Money does Granger-cause output in the bivariate money–output relation, Journal of Monetary Economics 22, 217–236.

Dickey, D.A., 1986, Stationarity transformations in vector autoregressions, Manuscript (Department of Statistics, North Carolina State University, Raleigh, NC).

Eichenbaum, M. and K.J. Singleton, 1986, Do equilibrium real business cycle theories explain postwar U.S. business cycles, NBER Macroeconomics Annual 1, 91–134.

Engle, R.F. and C.W.J. Granger, 1987, Co-integration and error correction: Representation, estimation and testing, Econometrica 55, 251–276.

Granger, C.W.J. and P. Newbold, 1977, Forecasting economic time series (Academic Press, New York).

Hansen, L.P. and T.J. Sargent, 1981, Formulating and estimating continuous time rational expectations models, Staff report no. 75 (Federal Reserve Bank of Minneapolis, Minneapolis, MN).

Harvey, A.C. and J.H. Stock, 1985, The estimation of higher order continuous time autoregressive models, Econometric Theory 1, 97–112.

Harvey, A.C. and J.H. Stock, 1988, Continuous time autoregressive models with common stochastic trends, Journal of Economic Dynamics and Control 12, 365–384.

Jones, R.H., 1981, Fitting a continuous time autoregression to discrete data, in: D.F. Findley, ed., Applied time series analysis II (Academic Press, New York).

Kalman, R.E. and R.S. Bucy, 1961, New results in linear filtering and prediction theory, Transactions of the AMSE (Journal of Basic Engineering) 83D, 95–108.

Melino, A., 1985, Estimation of unit average diffusion processes, Working paper (Institute for Policy Analysis, University of Toronto, Toronto).

Phadke, M.S. and S.M. Wu, 1974, Modeling of continuous stochastic processes from discrete observations with application to sunspots data, Journal of the American Statistical Association 69, no. 346, 325–329.

Phillips, A.W., 1959, The estimation of parameters in systems of stochastic differential equations, Biometrika, 67–76.

Runkle, D., 1987, Vector autoregressions and reality, Journal of Business and Economic Statistics 5, 437–442.

Sims, C.A., 1972, Money, income, and causality, American Economic Review 62, 540–552.

Sims, C.A., 1980a, Macroeconomics and reality, Econometrica 48, 1–48.

Sims, C.A., 1980b, Comparison of interwar and postwar business cycles: Monetarism reconsidered, American Economic Review 70, 250–257.

Stock, J.H. and M.W. Watson, 1989, Interpreting the evidence on money-income causality, Journal of Econometrics 40, 161–182.

Van Loan, C.F., 1978, Computing integrals involving the matrix exponential, IEEE Transactions on Automatic Control AC-23, 395–404.

Zadrozny, P., 1988, Gaussian likelihood of continuous-time ARMAX models when data are stocks and flows at different frequencies, Econometric Theory 4, 108–124.