Econometric Theory, **18**, 2002, 387–419. Printed in the United States of America.
DOI: 10.1017.S0266466602182077

MODELING CYCLICAL BEHAVIOR WITH DIFFERENTIAL-DIFFERENCE EQUATIONS IN AN UNOBSERVED COMPONENTS FRAMEWORK

MARCUS J. CHAMBERS
University of Essex
JOANNE S. MCGARRY
Loughborough University

This paper considers a continuous time unobserved components model in which the cyclical component follows a differential-difference equation whereas the trend and seasonal components follow more standard differential equations. Estimation of the parameters of the model with either a stock or a flow variable is analyzed using a frequency domain Gaussian estimator whose asymptotic properties are derived paying particular attention to the role of a truncation parameter that arises in the practical computation of the spectral density function. The results of a simulation exercise, which assesses the finite sample performance of the estimator, are also provided.

1. INTRODUCTION

This paper considers a continuous time unobserved components model in which the cyclical component follows a differential-difference equation whereas the trend and seasonal components follow more familiar differential equations. In economics, the use of differential-difference equations to characterize cyclical fluctuations can be traced back to Kalecki (1935). We believe that our formulation of the unobserved components model provides, at least in certain applications, more theoretical support for the specification of the cyclical component, as opposed to a pure time series representation. At the same time the model maintains the flexibility of the time series approach for the trend and seasonal, about which economic theory typically has rather less to say. The proposed method of estimation is based in the frequency domain and utilizes the discrete form of the Whittle likelihood. Furthermore, the techniques are flexible enough

An earlier version of this paper was prepared for the Cowles Foundation conference "New Developments in Time Series Econometrics," held at Yale University, October 23–24, 1999. We are grateful to Roy Bailey, Rex Bergstrom, Andrew Harvey, Peter Phillips, Neil Shephard, and four anonymous referees for helpful comments, without implicating them in any of the possible shortcomings of the paper. Address correspondence to: Marcus J. Chambers, Department of Economics, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, England; e-mail: mchamb@essex.ac.uk.

© 2002 Cambridge University Press 0266-4666/02 $9.50

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

to accommodate the point-sampling of a stock variable and the discrete time averaging inherent in the sampling of a flow variable.

Recently there has been a resurgence of interest in differential-difference equations in economics,[1] although so far there have been few empirical applications of such models.[2] Although we focus on cyclical behavior with economic applications in mind, it should be noted that differential-difference equations have been used extensively in other disciplines, notably physics, engineering, and biology. The estimation of unobserved (possibly noninteger) lag parameters, such as the one appearing in the differential-difference equation, has, in fact, been an important topic of research in the signal processing literature. The estimation of time delay, and the associated estimation of coherence and phase, by frequency domain methods was established in a sequence of articles by Hannan and Thomson (1971, 1973, 1981, 1988). In a regression context Hannan and Robinson (1973) propose a frequency domain method of estimating an unknown lag parameter and apply the method to some economic and oceanographic data, whereas a more general model with unknown lags is considered by Robinson (1974). The recent contributions of Chambers (1998) and McGarry (2000), allied to the results in the present paper, should help to encourage empirical investigations using differential-difference equations in economics.

The paper is organized as follows. Section 2 provides some relevant background on differential-difference equations and also some specific results concerning the particular equation that we use for modeling the cyclical component. These results concern the required conditions on the parameters that ensure stationarity and also conditions that yield a business cycle (which we take to be a cycle of length greater than twice the lag appearing in the differential-difference equation). The precise form of the spectral density functions corresponding to stock and flow variables is also derived. Section 3 concentrates on the unobserved components model in continuous time and provides results for stock and flow variables. The model is flexible enough to allow simultaneously for nonstationarity, path dependence, and the presence of unobservables in the process of interest.[3] In the case of the trend and seasonal being generated by differential equations it is possible to obtain exact autoregressive representations for the discrete time components. These exact representations are used to derive the spectral density functions of the discretely sampled processes in Section 4. In the case of flow variables the results we derive for the seasonal and trend are new and could even be used in a discrete time Kalman filtering approach to estimation of the continuous time components. We are unable to follow this route in the present context, however, because it is not possible to derive exact discrete time representations corresponding to an underlying differential-difference equation. Hence our focus is on spectral densities and frequency domain methods of estimation. Section 3 also gives by way of comparison some details of an alternative continuous time representation of the cycle in the form of a differential equation, an approach suggested by Harvey and Stock (1993). Section 4 derives the spectral densities of interest and considers the problem of estimating

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

the parameters of the model from a discrete time sample of observations. For reasons that will become apparent it is necessary to truncate a doubly infinite series in the computation of the spectral density function. We therefore investigate the asymptotic properties of the feasible frequency domain Gaussian estimator (based on the truncated series) concentrating specifically on the role played by the truncation parameter. Section 5 provides some simulation results for a model involving trend, cycle, and irregular components, and Section 6 concludes. All proofs of propositions and theorems are provided in the Appendix.

2. DIFFERENTIAL-DIFFERENCE EQUATIONS AND CYCLES

The use of differential-difference equations in this paper is for the purposes of modeling the cyclical component in an unobserved components framework in continuous time. Such models are termed *structural time series models* by Harvey (1989). There are two main requirements that the cyclical component should satisfy in this type of framework. The first, and most obvious, is that the specification used is capable of generating cycles. The second is that the specification should be relatively parsimonious, particularly when compared to the differential equation specification suggested by Harvey and Stock (1993) for modeling cycles in continuous time. The differential-difference equation proposed here satisfies these two requirements but also has the advantage that it often arises quite naturally from economic theory, a prominent early example being Kalecki (1935). We shall focus on the equation

$$dy(t) = [a_0 y(t) + a_1 y(t-p)]dt + \epsilon(dt), \quad t > 0, \tag{1}$$

where $y(t)$ is the scalar continuous time random process of interest, $\epsilon(dt)$ is a mean-zero random measure with variance $\sigma_\epsilon^2 dt$, a_0 and a_1 are unknown scalar parameters, and p denotes the unknown lag parameter that is real and positive. The process $y(t)$ is stationary provided that the roots of the equation $a(z) \equiv z - a_0 - a_1 e^{-pz} = 0$ have negative real parts. The equation $a(z) = 0$ typically has an infinite number of roots, and checking the required condition would appear to be an impossible task.[4] However, McGarry (2000) provides the following result (with proof) for the equation defined in (1).

PROPOSITION 1. *Let* $a(z) \equiv z - a_0 - a_1 e^{-pz}$, *where* a_0, a_1, *and* p *are real. Then all the roots of the equation* $a(z) = 0$ *have negative real parts if and only if*

(i) $a_0 < 1/p$, *and*
(ii) $a_0 < -a_1 < \sqrt{a_0^2 + x_1^2}$,

where x_1 *is the root of the equation* $x = a_0 \tan px$ *such that* $0 < x_1 < \pi/p$.

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

Proposition 1 is important because it provides conditions that are easily verified in practice for determining whether a process satisfying (1) is stationary or not.

With regard to cyclical behavior, it is rather important to be clear about precisely what is meant by the term *cycle*. In the case of differential-difference equations it is possible to obtain an infinite number of cycles whose lengths are equal to $2\pi p/r_j$, where the r_j $(j = 1,2,\ldots)$ are the solutions to a particular transcendental equation. James and Belz (1936) show that there is at most one such solution within each interval $((j-1)\pi, j\pi)$ $(j = 1,2,\ldots)$, and so all cycles have period less than $2p$ except for the one corresponding to the root r_1 in the interval $(0,\pi)$. It is this cycle that has been referred to as the major cycle by Frisch and Holme (1935), and it is the one that Kalecki (1935) regards as the business cycle. It is also the one that will be of most concern to us here and the one that will be implicitly meant by the unqualified use of the term *cycle*. Proposition 2 provides the requirement on the parameters of (1) for the model to produce cycles.

PROPOSITION 2. *The stochastic differential-difference equation (1) produces cycles provided that* $|-a_1 p| > e^{a_0 p - 1}$. *The period of the major, or business, cycle is given by* $2\pi p/r_1$ *where* $0 < r_1 < \pi$ *is the smallest positive root (when it exists) of the equation*

$$r\cot r + \ln\left|\frac{\sin r}{r}\right| = a_0 p - \ln|-a_1 p|. \tag{2}$$

The proof of Proposition 2 appears in the Appendix and is based on results of Kalecki (1935) and James and Belz (1936). A necessary, but not sufficient, condition for the existence of cycles is that $|-a_1| > a_0$, which arises because $e^{x-1} > x$ and hence $|-a_1 p| > e^{a_0 p - 1} > a_0 p$. Note that this condition is closely related to condition (ii) of Proposition 1, the stationarity requirement. The function $r\cot r + \ln|\sin r/r|$ is graphed in James and Belz (1936). The expression given previously for determining r_1, which in turn determines the length of the cycle, is more complicated than the corresponding expression in Kalecki (1935). This is because Kalecki's model implies $a_0 > 0$ and $a_1 < 0$ and further imposes the condition that the amplitude of the cycle is constant and equal to (in our notation) $a_0 p$. The model in this paper imposes no such constraints on a_0 and a_1, and the derivation of the root r_1 does not constrain the amplitude to be constant (although it does not exclude this possibility). The approach is therefore closer to James and Belz (1936) in this respect.

Unlike the case with ordinary (stochastic) differential equations, the representations of the solution to a differential-difference equation system do not lend themselves readily to the derivation of a convenient time domain representation for the observed process. Although the solution to a stochastic differential equation system of order p satisfies a discrete time ARMA$(p, p-1)$

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

equation for stock variables and an ARMA(p,p) equation for flow variables, the same is not true for variables satisfying an underlying differential-difference equation system. The spectral density function corresponding to a differential-difference equation is, however, easily derived, and hence the frequency domain approach is followed in this paper. Noting that (1) has the representation $a(D)y(t)dt = \epsilon(dt)$, where D denotes the mean square differential operator, a straightforward result from filtering theory[5] yields the spectral density function of the continuous time process $y(t)$ as

$$f_y^c(\lambda) = \frac{\sigma_\epsilon^2}{2\pi} |a(i\lambda)|^{-2}, \quad -\infty < \lambda < \infty, \tag{3}$$

where λ denotes frequency. From the definition of $a(z)$ the spectral density in (3) may be written more explicitly as

$$f_y^c(\lambda) = \frac{\sigma_\epsilon^2}{2\pi} [\lambda^2 + a_0^2 + a_1^2 + 2a_1(a_0 \cos p\lambda + \lambda \sin p\lambda)]^{-1}, \quad -\infty < \lambda < \infty.$$

The spectral density function therefore behaves like λ^{-2} as $\lambda \to \infty$ and is integrable over $(-\infty,\infty)$, so that $y(t)$ has finite variance.

The spectral density function of the observed process depends on whether it is a stock variable or a flow variable. For a stock variable, we shall assume that observations are equispaced and made at integer points in time, so that the sequence of observations is $\{y_t = y(t);\ t = 0,\pm1,\pm2,\ldots\}$. In this case, the spectral density function of y_t is

$$f_y(\lambda) = \sum_{j=-\infty}^{\infty} f_y^c(\lambda + 2\pi j) = \frac{\sigma_\epsilon^2}{2\pi} \sum_{j=-\infty}^{\infty} |a(i(\lambda + 2\pi j))|^{-2}, \quad -\pi < \lambda \le \pi, \tag{4}$$

which is the familiar "folding equation" that folds back the aliases of each $\lambda \in (-\pi,\pi]$ into that interval. Flow variables, on the other hand, are observed as the integral of the underlying rate of flow over the unit interval. The sequence of observations on a flow variable is therefore $\{Y_t = \int_{t-1}^{t} y(r)dr = \int_0^1 y(t-r)dr;\ t = 0,\pm1,\pm2,\ldots\}$. The frequency response function of the integrating filter, which transforms the continuous time process $y(t)$ into the observable process Y_t, is given by the function $g(z) = \int_0^1 e^{-rz}dr = (1 - e^{-z})/z$. The spectral density function of Y_t is then

$$\begin{aligned} f_Y(\lambda) &= \sum_{j=-\infty}^{\infty} |g(i(\lambda + 2\pi j))|^2 f_y^c(\lambda + 2\pi j) \\ &= \frac{\sigma_\epsilon^2}{2\pi} \sum_{j=-\infty}^{\infty} \frac{4\sin^2(\lambda/2)}{(\lambda + 2\pi j)^2} |a(i(\lambda + 2\pi j))|^{-2}, \quad -\pi < \lambda \le \pi, \end{aligned} \tag{5}$$

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

because $|g(i(\lambda + 2\pi j))|^2 = 4\sin^2((\lambda + 2\pi j)/2)/(\lambda + 2\pi j)^2 = 4\sin^2(\lambda/2)/(\lambda + 2\pi j)^2$. In the case where the lag parameter p is a positive integer, Chambers (1998) shows how to compute these spectral densities accurately using the method of residues, which transforms the doubly infinite series into a sum over a finite number of residues.[6] This method is not applicable, however, when p is noninteger, but McGarry (2000) has demonstrated that accurate approximations to the spectral density can be obtained by using an appropriate truncation of the series. Details of the computational procedures used in this paper are provided in subsequent sections.

The preceding results suggest that differential-difference equations in the form of (1) satisfy the requirements for the cyclical component in the unobserved components model in continuous time. The equation is capable of producing cycles; it is relatively parsimonious; and, furthermore, it is of a form that has been suggested by economic theory for certain types of variables. Equation (1) will therefore constitute the basis for the cyclical component in this paper.

3. AN UNOBSERVED COMPONENTS FRAMEWORK FOR CONTINUOUS TIME PROCESSES

3.1. The Continuous Time Components

In this section we define the continuous time unobserved components model in which the cyclical component satisfies a differential-difference equation. For simplicity we focus on the univariate case, but the methods extend straightforwardly to the multivariate case. The basic idea is that the scalar continuous time random process $y(t)$ consists of a trend component $\mu(t)$, a seasonal component $\gamma(t)$, and a cyclical component $\psi(t)$, so that $y(t) = \mu(t) + \gamma(t) + \psi(t)$. The trend and seasonal components are assumed to be governed by a standard first-order stochastic differential equation, as in Harvey and Stock (1993), whereas the cycle is governed by a differential-difference equation. The precise definitions of each of these components are given subsequently.[7]

Trend component. The trend component, in its most general form, evolves according to the bivariate system

$$d\begin{bmatrix}\mu(t)\\ \beta(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}\mu(t)\\ \beta(t)\end{bmatrix}dt + \begin{bmatrix}\eta(dt)\\ \zeta(dt)\end{bmatrix}, \quad t > 0, \tag{6}$$

where $\eta(dt)$ and $\zeta(dt)$ are mutually and serially uncorrelated random measures with zero means and variances $\sigma_\eta^2 dt$ and $\sigma_\zeta^2 dt$, respectively. The random measures are serially uncorrelated in the sense that $E[\eta(\Delta_1)\eta(\Delta_2)] = E[\zeta(\Delta_1)\zeta(\Delta_2)] = 0$ for any two sets (intervals) Δ_1 and Δ_2 on the real line such that $\Delta_1 \cap \Delta_2 = \varnothing$, and they are mutually uncorrelated in the sense that $E[\eta(\Delta_1)\zeta(\Delta_2)] = 0$ for any two intervals Δ_1 and Δ_2. The components $\mu(t)$ and $\beta(t)$ are often referred to as the level and slope, respectively.

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

Seasonal component. The seasonal component is defined as $\gamma(t) = \sum_{j=1}^{s/2} \gamma_j(t)$, in which s denotes the number of seasons in a year and the $\gamma_j(t)$ satisfy

$$d\begin{bmatrix}\gamma_j(t)\\ \gamma_j^*(t)\end{bmatrix} = \begin{bmatrix}0 & \lambda_j^s\\ -\lambda_j^s & 0\end{bmatrix}\begin{bmatrix}\gamma_j(t)\\ \gamma_j^*(t)\end{bmatrix}dt + \begin{bmatrix}\omega_j(dt)\\ \omega_j^*(dt)\end{bmatrix}, \quad j = 1,\ldots,s/2, \quad t > 0, \qquad \textbf{(7)}$$

where $\lambda_j^s = 2\pi j/s$ $(j = 1,\ldots,s/2)$ denote the seasonal frequencies and $\omega_j(dt)$ and $\omega_j^*(dt)$ are zero mean mutually and serially uncorrelated random measures with common variance $\sigma_\omega^2 dt$. It is also assumed that $\omega_j(dt)$ and $\omega_j^*(dt)$ are uncorrelated with $\eta(dt)$ and $\zeta(dt)$ for all j and at all leads and lags.

Cyclical component. The form of the cyclical component follows the specification in (1) and is given by

$$d\psi(t) = [a_0\psi(t) + a_1\psi(t-p)]dt + \epsilon(dt), \quad t > 0, \qquad \textbf{(8)}$$

where $\epsilon(dt)$ is a zero mean random measure with variance $\sigma_\epsilon^2 dt$, which is uncorrelated with $\eta(dt)$, $\zeta(dt)$, $\omega_j(dt)$, and $\omega_j^*(dt)$ (for all j) at all leads and lags.

The specification of the model is completed with an equation that relates the observed process to the underlying continuous time components. The next two sections provide the relevant details for the discrete time sampling of stock variables and flow variables, respectively.

3.2. Discrete Time Sampling: Stock Variables

Stock variables are assumed to be observed at integer points in time, so that the observed sequence is given by $\{y_t = y(t);\ t = 1,2,\ldots,T\}$, where T denotes sample size. We assume throughout that observations are evenly spaced at unit intervals of time, but it is possible to extend the analysis to situations where $y(t)$ is observed at irregular intervals.[8] The observations satisfy the measurement equation

$$y_t = \mu_t + \gamma_t + \psi_t + \xi_t, \quad t = 1,\ldots,T, \qquad \textbf{(9)}$$

where $\mu_t = \mu(t)$, $\gamma_t = \gamma(t)$, $\psi_t = \psi(t)$, and the irregular component ξ_t is white noise with variance σ_ξ^2. The random disturbance ξ_t can be interpreted either as the discrete time realization of a genuine irregular component in continuous time that affects $y(t)$ or as a measurement error that is associated with the discrete time sampling of the continuous time process.

For the trend and seasonal components, which are generated by differential equations, it is possible to obtain exact time domain representations for μ_t and γ_t, which can be used to derive the spectral density functions for use in estimation. However, for the reasons outlined in Section 2, such time domain representations are not possible for the cyclical component ψ_t, although its spectral density function can still be obtained as in Section 2. Each of the trend and

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

seasonal components can be written as $dx(t) = Ax(t)dt + \nu(dt)$, where $x(t)$ is a 2×1 vector and A is a 2×2 coefficient matrix. The uncorrelatedness of the random measures in each component, and the absence of any interactions between the individual components, means that it is legitimate to treat the components separately in this way. The solution of the continuous time representation for $x(t)$ yields the first-order vector autoregressive representation for $x_t = x(t)$ given by

$$x_t = e^A x_{t-1} + \nu_t, \quad t = 1, \ldots, T, \tag{10}$$

where the matrix exponential is defined by $e^A = \sum_{j=0}^{\infty} A^j/(j!)$ and the discrete time disturbance vector is given by[9]

$$\nu_t = \int_{t-1}^{t} e^{A(t-s)} \nu(ds).$$

The vector ν_t is white noise with variance matrix $\mathrm{var}(\nu_t) = \int_0^1 e^{As} \Sigma_\nu e^{A's} ds$, where $\Sigma_\nu dt$ is the covariance matrix of $\nu(dt)$. Applying these results to the trend and seasonal components yields the following discrete time representations for μ_t and γ_t, which are given in Harvey and Stock (1993).

Trend component. The discrete time trend component μ_t satisfies

$$\begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \end{bmatrix} + \begin{bmatrix} \eta_t \\ \zeta_t \end{bmatrix}, \quad t = 1, \ldots, T, \tag{11}$$

where the disturbance vector $[\eta_t, \zeta_t]'$ is white noise with variance matrix

$$\mathrm{var} \begin{bmatrix} \eta_t \\ \zeta_t \end{bmatrix} = \begin{bmatrix} \sigma_\eta^2 + \sigma_\zeta^2/3 & \sigma_\zeta^2/2 \\ \sigma_\zeta^2/2 & \sigma_\zeta^2 \end{bmatrix}. \tag{12}$$

Note that the discrete time disturbances η_t and ζ_t are correlated even though the underlying continuous time random measures are not.

Seasonal component. The discrete time seasonal component $\gamma_t = \sum_{j=1}^{s/2} \gamma_{jt}$, where each γ_{jt} satisfies

$$\begin{bmatrix} \gamma_{jt} \\ \gamma_{jt}^* \end{bmatrix} = \begin{bmatrix} \cos \lambda_j^s & \sin \lambda_j^s \\ -\sin \lambda_j^s & \cos \lambda_j^s \end{bmatrix} \begin{bmatrix} \gamma_{j,t-1} \\ \gamma_{j,t-1}^* \end{bmatrix} + \begin{bmatrix} \omega_{jt} \\ \omega_{jt}^* \end{bmatrix}, \quad j = 1, \ldots, s/2, \quad t = 1, \ldots, T, \tag{13}$$

where the vectors $[\omega_{jt}, \omega_{jt}^*]'$ are white noise with common variance matrix $\sigma_\omega^2 I_2$ and are uncorrelated across j.

Although the discrete time representations for trend and seasonal given previously lend themselves readily to the application of time domain maximum likelihood estimation via the prediction error decomposition of the Gaussian likelihood function and application of the Kalman filter (as in Harvey and Stock, 1993), the presence of the cyclical component ψ_t rules out such an approach. Instead, we use the preceding representations to derive the spectral density functions of these components, which are detailed in Section 4.

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

3.3. Discrete Time Sampling: Flow Variables

Flow variables are observed as the integral of the underlying rate of flow over the unit time interval, so that the observed sequence is $\{Y_t = \int_{t-1}^{t} y(r)dr;\ t = 1,\ldots,T\}$. The measurement equation for flow variables is

$$Y_t = A_t + \Gamma_t + \Psi_t + \Xi_t, \quad t = 1,\ldots,T, \tag{14}$$

where $A_t = \int_{t-1}^{t} \mu(r)dr$, $\Gamma_t = \int_{t-1}^{t} \gamma(r)dr$, $\Psi_t = \int_{t-1}^{t} \psi(r)dr$, Ξ_t is white noise with variance σ_Ξ^2, and the integrals are defined in the wide sense.[10] The random disturbance Ξ_t has the same interpretation as in the case of stock variables.

Following the approach for stock variables, the next task is to derive discrete time representations for the trend and seasonal components, based on the generic continuous time representation $dx(t) = Ax(t)dt + \nu(dt)$. It is well known that flow variables generated by a first-order stochastic differential equation satisfy an ARMA(1,1) process in discrete time,[11] and the approach followed here is to integrate equation (10) over the interval $(t-1, t)$ to obtain such a representation. The disturbance term in this representation has the form of a double integral of the underlying random measures, giving rise to the MA(1) autocorrelations. This approach is followed here because it is convenient for the frequency domain methods employed later. An alternative approach is used by Harvey and Stock (1993) and has the advantage that the resulting disturbance term is white noise, although the lagged vector in the autoregressive component is not $\int_{t-1}^{t} x(r)dr$ but $x(t-1)$ instead. Although the Kalman filter is ideally suited to take advantage of such an unobservable component, this discrete time state space representation does not lend itself well to the frequency domain approach that we must adopt.

We present, first of all, a general theorem that provides the form of the discrete time model satisfied by the integrals $X_t = \int_{t-1}^{t} x(r)dr$ and that holds for each of the unobserved trend and seasonal components.

THEOREM 1. *Let $x(t)$ be a $q \times 1$ continuous time random vector satisfying*

$$dx(t) = Ax(t)dt + \nu(dt), \quad t > 0, \tag{15}$$

where A is a fixed $q \times q$ matrix of coefficients and $\nu(dt)$ is a $q \times 1$ vector of serially uncorrelated random measures with mean vector zero and variance matix $\Sigma_\nu dt$. Then $X_t = \int_{t-1}^{t} x(r)dr$ $(t = 1,\ldots,T)$ satisfies

$$X_t = e^A X_{t-1} + N_t, \quad t = 1,\ldots,T,$$

where $N_t = \int_{t-1}^{t} \int_{s-1}^{s} e^{A(s-r)}\nu(dr)ds$ is an MA(1) disturbance vector with

$$E(N_t N_t') \equiv V_0 = 2\Lambda(1) + \Phi(1)\Sigma_\nu\Phi(1)' - \Phi(1)\Sigma_\nu\Upsilon(1)' - \Upsilon(1)\Sigma_\nu\Phi(1)',$$

$$E(N_t N_{t-1}') \equiv V_1 = \Phi(1)\Sigma_\nu\Upsilon(1)' - \Lambda(1),$$

in which $\Phi(r) = \int_0^r e^{As}ds$, $\Upsilon(r) = \int_0^r \Phi(s)ds$, and $\Lambda(r) = \int_0^r \Phi(s)\Sigma_\nu\Phi(s)'ds$.

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

The discrete time representation given in Theorem 1, which holds for both stationary and nonstationary (integrated) processes, does not appear to have been given in this precise form in the literature to date. The representation encompasses, for example, the model for flow variables given by Bergstrom (1984, Theorem 8), which is derived for (asymptotically) stationary processes for which the coefficient matrix in the continuous time model is nonsingular. The representation in Theorem 1 holds for both unobserved components of interest, including the trend in which the coefficient matrix is singular. Theorem 1 can be applied to each component separately, in view of the uncorrelatedness of the underlying random measures between components and the absence of any other interactions between components. The results are presented in Theorem 2, which follows.

THEOREM 2. *Let the unobserved components for trend and seasonal evolve in continuous time according to (6) and (7), respectively. Then the discrete time flow variables follow ARMA(1,1) laws of motion as follows.*

Trend component. The discrete time trend component A_t satisfies

$$\begin{bmatrix} A_t \\ B_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_{t-1} \\ B_{t-1} \end{bmatrix} + \begin{bmatrix} H_t \\ Z_t \end{bmatrix}, \quad t = 1, \ldots, T,$$

where the MA(1) disturbance vector $[H_t, Z_t]'$ has variance matrix V_0 and first-order autocovariance matrix V_1 given by

$$V_0 = \begin{bmatrix} \frac{2}{3}\sigma_\eta^2 + \frac{11}{60}\sigma_\zeta^2 & \frac{1}{3}\sigma_\zeta^2 \\ \frac{1}{3}\sigma_\zeta^2 & \frac{2}{3}\sigma_\zeta^2 \end{bmatrix}, \quad V_1 = \begin{bmatrix} \frac{1}{6}\sigma_\eta^2 - \frac{1}{30}\sigma_\zeta^2 & \frac{1}{8}\sigma_\zeta^2 \\ \frac{1}{24}\sigma_\zeta^2 & \frac{1}{6}\sigma_\zeta^2 \end{bmatrix}.$$

Seasonal component. The discrete time seasonal component is given by $\Gamma_t = \sum_{j=1}^{s/2} \Gamma_{jt}$, with each Γ_{jt} satisfying

$$\begin{bmatrix} \Gamma_{jt} \\ \Gamma_{jt}^* \end{bmatrix} = \begin{bmatrix} \cos\lambda_j^s & \sin\lambda_j^s \\ -\sin\lambda_j^s & \cos\lambda_j^s \end{bmatrix} \begin{bmatrix} \Gamma_{j,t-1} \\ \Gamma_{j,t-1}^* \end{bmatrix} + \begin{bmatrix} \Omega_{jt} \\ \Omega_{jt}^* \end{bmatrix}, \quad j = 1, \ldots, s/2, \quad t = 1, \ldots, T,$$

where the MA(1) disturbance vector $[\Omega_{jt}, \Omega_{jt}^*]'$ has variance matrix V_{0j} and first-order autocovariance matrix V_{1j} given by (for $j = 1, \ldots, s/2$)

$$V_{0j} = \frac{4\sigma_\omega^2}{(\lambda_j^s)^3} \begin{bmatrix} (\lambda_j^s - \sin\lambda_j^s) & 0 \\ 0 & (\lambda_j^s - \sin\lambda_j^s) \end{bmatrix},$$

$$V_{1j} = \frac{\sigma_\omega^2}{(\lambda_j^s)^3} \begin{bmatrix} \lambda_j^s(1 - \cos\lambda_j^s) - 2(\lambda_j^s - \sin\lambda_j^s) & 2(1 - \cos\lambda_j^s) - \lambda_j^s \sin\lambda_j^s \\ \lambda_j^s \sin\lambda_j^s - 2(1 - \cos\lambda_j^s) & \lambda_j^s(1 - \cos\lambda_j^s) - 2(\lambda_j^s - \sin\lambda_j^s) \end{bmatrix}.$$

The autoregressive matrices in the discrete time ARMA(1,1) representations for the unobserved components in the flow case are identical to those in

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

the case of stocks. The properties of the disturbances, however, are rather different. In particular the disturbances are MA(1) processes for flows but white noise vectors for stocks. Note that, for the seasonal component, the discrete time disturbances remain contemporaneously uncorrelated but they display cross-correlation at the one period lag.

The ARMA(1,1) representations in Theorem 2 provide the basis for deriving the spectral density functions of the unobserved components and, hence, of the observed flow variable Y_t. These expressions will be provided in detail in Section 4, in which issues of estimation are addressed.

3.4. An Alternative Representation for the Cyclical Component

Harvey and Stock (1993) propose a differential equation to capture the cyclical component in a continuous time unobserved components model. Its form is similar to the differential equation for the seasonal component in (7) and is given by

$$d\begin{bmatrix}\psi(t)\\ \psi^*(t)\end{bmatrix} = \begin{bmatrix}\ln\rho & \lambda_c\\ -\lambda_c & \ln\rho\end{bmatrix}\begin{bmatrix}\psi(t)\\ \psi^*(t)\end{bmatrix}dt + \begin{bmatrix}\kappa(dt)\\ \kappa^*(dt)\end{bmatrix}, \quad t > 0, \tag{16}$$

where $\kappa(dt)$ and $\kappa^*(dt)$ are zero mean mutually and serially uncorrelated random measures with common variance $\sigma_\kappa^2\,dt$, λ_c denotes the frequency of the cycle, and $0 < \rho < 1$ for the process to be stationary. In this specification the frequency of the cycle is estimated directly, and the length of the cycle is given by $2\pi/\lambda_c$. The representation for the cycle in (16) is more parsimonious than the differential-difference equation (8), involving three rather than four unknown parameters, and also lends itself to a convenient time domain representation. Its motivation on economic grounds is, perhaps, weaker than for (8) in some circumstances (e.g., time-to-build investment models), but it can provide an alternative to (8) for estimating the features of cycles in continuous time models.

In terms of the discrete time sampling of stock and flow variables, the discrete time representations follow from (10) for stock variables and from Theorem 1 for flow variables. For stock variables the appropriate equations are provided by Harvey and Stock (1993) and are given by

$$\begin{bmatrix}\psi_t\\ \psi_t^*\end{bmatrix} = \rho\begin{bmatrix}\cos\lambda_c & \sin\lambda_c\\ -\sin\lambda_c & \cos\lambda_c\end{bmatrix}\begin{bmatrix}\psi_{t-1}\\ \psi_{t-1}^*\end{bmatrix} + \begin{bmatrix}\kappa_t\\ \kappa_t^*\end{bmatrix}, \quad t = 1,\ldots,T, \tag{17}$$

where the vector $[\kappa_t, \kappa_t^*]'$ is white noise with variance matrix $-\sigma_\kappa^2(1-\rho^2)I_2/(2\ln\rho)$. For flow variables, however, we present the result as a corollary to Theorem 2.

COROLLARY TO THEOREM 2. *Let the unobserved cyclical component evolve in continuous time according to (16). Then the discrete time law of motion in the case of a flow variable Ψ_t is given by the ARMA(1,1) process*

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

$$\begin{bmatrix} \Psi_t \\ \Psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \Psi_{t-1} \\ \Psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} K_t \\ K_t^* \end{bmatrix},$$

where the MA(1) disturbance vector $[K_t, K_t^*]'$ *has variance matrix* V_{0K} *and first-order autocovariance matrix* V_{1K} *given by*

$$V_{0K} = \sigma_\kappa^2\{[2c_1(\lambda_c) + c_2(\lambda_c) - 2c_3(\lambda_c)]I_2 - H(\lambda_c) - H(\lambda_c)'\},$$
$$V_{1K} = \sigma_\kappa^2\{[c_3(\lambda_c) - c_1(\lambda_c)]I_2 + [c_4(\lambda_c)J + c_5(\lambda_c)J' + H(\lambda_c)]\},$$

where

$$c_1(\lambda) = \frac{-2\rho}{\delta(\lambda)^2}[\ln\rho\cos\lambda + \lambda\sin\lambda] + \frac{1}{\delta(\lambda)}\left[1 + \frac{\rho^2 - 1}{2\ln\rho} + \frac{2\ln\rho}{\delta(\lambda)}\right],$$
$$c_2(\lambda) = \frac{1}{\delta(\lambda)}[1 + \rho^2 - 2\rho\cos\lambda], \quad c_3(\lambda) = \frac{\rho^2\ln\rho}{\delta(\lambda)^2}, \quad c_4(\lambda) = \frac{2\rho^2\lambda(\ln\rho)^2}{\delta(\lambda)^3},$$
$$c_5(\lambda) = \frac{\rho^2\lambda[(\ln\rho)^2 - \lambda^2]}{\delta(\lambda)^3}, \quad \delta(\lambda) = (\ln\rho)^2 + \lambda^2, \quad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$
$$H(\lambda) = \frac{\rho\ln\rho}{\delta(\lambda)} C_1(\lambda)C_4(\lambda)' + \frac{\rho\lambda}{\delta(\lambda)} C_2(\lambda)C_4(\lambda)'$$
$$- C_3(\lambda)\left[C_4(\lambda)' + \frac{\rho[(\ln\rho)^2 - \lambda^2]}{\delta(\lambda)^2} C_1(\lambda)' + \frac{2\rho\lambda\ln\rho}{\delta(\lambda)^2} C_2(\lambda)'\right],$$
$$C_1(\lambda) = \begin{bmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{bmatrix}, \quad C_2(\lambda) = \begin{bmatrix} \sin\lambda & -\cos\lambda \\ \cos\lambda & \sin\lambda \end{bmatrix},$$
$$C_3(\lambda) = \frac{1}{\delta(\lambda)}\begin{bmatrix} \ln\rho & -\lambda \\ \lambda & \ln\rho \end{bmatrix},$$
$$C_4(\lambda) = \frac{1}{\delta(\lambda)^2}\begin{bmatrix} \lambda^2 - (\ln\rho)^2 - \delta(\lambda)\ln\rho & 2\lambda\ln\rho + \lambda\delta(\lambda) \\ -2\lambda\ln\rho - \lambda\delta(\lambda) & \lambda^2 - (\ln\rho)^2 - \delta(\lambda)\ln\rho \end{bmatrix}.$$

As can be seen from the equations in the corollary, the passage from continuous time to discrete time in the case of the cyclical component of a flow variable imposes complicated restrictions on the moving average disturbance process. The discrete time representation in this corollary can be used to derive the spectral density function of Ψ_t, which can be used in the estimation of the continuous time model.

4. FREQUENCY DOMAIN APPROACHES TO ESTIMATION

The difficulties associated with deriving discrete time representations for processes generated by differential-difference equations have already been outlined. We therefore follow a frequency domain estimation procedure that utilizes the spectral density function of the differential-difference equation component,

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

which is readily derived. We note, however, that in the situation where all of the unobserved components are generated by stochastic differential equations rather than differential-difference equations, it is possible to use time domain methods of estimation. For example, Harvey and Stock (1993) propose Kalman filtering techniques that can be used in the construction of the Gaussian likelihood function. Our treatment of the model based on stochastic differential equations therefore complements the approach of Harvey and Stock (1993) and differs from their approach in two main respects. First, we use the discrete time representations for flows given in Theorem 2, and second, we use frequency domain methods of estimation rather than time domain methods.

In what follows we shall assume that the observed variable (either y_t or Y_t) is comprised of all four unobserved components, trend, seasonal, cycle, and irregular. The necessary modifications that need to be made in circumstances in which one or more of these components is absent are straightforward. As defined earlier, the trend and seasonal components are nonstationary, and, hence, so too is the observed variable. To carry out analysis in the frequency domain, it is necessary to transform the observations so that they form a stationary sequence. In the case of the trend component, the operator that needs to be applied to induce stationarity is the second difference operator $\Delta^2 = (1 - L)^2$, where L denotes the lag operator.[12] In the case of the seasonal component, the relevant operator is $S(L) = 1 + \sum_{j=1}^{s-1} L^j$. Hence in the most general case, where y_t or Y_t contain all four unobserved components, it is necessary to work with the transformed variables

$$z_t = \phi(L)y_t \quad \text{and} \quad Z_t = \phi(L)Y_t, \quad \text{where } \phi(L) = \Delta^2 S(L).$$

The spectral density functions for each of these two transformed variables are derived subsequently.

4.1. Spectral Density Function: Stock Variables

The uncorrelatedness of the individual unobserved components ensures that the spectral density function of z_t can be expressed as the sum of the spectral densities of the individual components. Let the spectral density function of a variable x_t be denoted $f_x(\lambda)$, where $-\pi < \lambda \leq \pi$ denotes frequency. Then the spectral density function of z_t, denoted $f_z(\lambda)$, is given by

$$f_z(\lambda) = |\phi(e^{-i\lambda})|^2 [f_\mu(\lambda) + f_\gamma(\lambda) + f_\psi(\lambda) + f_\xi(\lambda)], \quad -\pi < \lambda \leq \pi. \tag{18}$$

Precise expressions for each of the constituents of (18) are given subsequently.

Trend component.

$$f_\mu(\lambda) = \frac{1}{2\pi} w'(I_2 - Ce^{-i\lambda})^{-1} V_0 (I_2 - C'e^{i\lambda})^{-1} w, \quad -\pi < \lambda \leq \pi,$$

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

where

$$w = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad V_0 = \begin{bmatrix} \sigma_\eta^2 + \sigma_\zeta^2/3 & \sigma_\zeta^2/2 \\ \sigma_\zeta^2/2 & \sigma_\zeta^2 \end{bmatrix}.$$

Seasonal component.

$$f_\gamma(\lambda) = \frac{\sigma_\omega^2}{2\pi} \sum_{j=1}^{s/2} w'[I_2 - C_1(\lambda_j^s)e^{-i\lambda}]^{-1}[I_2 - C_1(\lambda_j^s)'e^{i\lambda}]^{-1}w, \quad -\pi < \lambda \leq \pi,$$

where w is as defined previously and $C_1(\lambda)$ is defined in the corollary to Theorem 2.

Cyclical component.

$$f_\psi(\lambda) = \frac{\sigma_\epsilon^2}{2\pi} \sum_{j=-\infty}^{\infty} |a(i(\lambda + 2\pi j))|^{-2}, \quad -\pi < \lambda \leq \pi,$$

where $a(z) = z - a_0 - a_1 e^{-pz}$.

Irregular component.

$$f_\xi(\lambda) = \frac{\sigma_\xi^2}{2\pi}, \quad -\pi < \lambda \leq \pi.$$

The functions $f_\mu(\lambda)$ and $f_\gamma(\lambda)$ are perhaps best described as being pseudospectral densities in view of the variables μ and γ being nonstationary. These functions contain poles at certain frequencies, these being the origin ($\lambda = 0$) and the seasonal frequencies, respectively. These poles are, however, annihilated by the operator $\phi(L)$, whose frequency response function is $\phi(e^{-i\lambda})$. For the processes generated by differential equations, the preceding functions are derived directly from the discrete time representations in Section 3.2, whereas $f_\psi(\lambda)$ corresponds directly to (4) in Section 2.

4.2. Spectral Density Function: Flow Variables

The spectral density function for the transformed flow variable Z_t can be derived in a similar form to the spectral density for the transformed stock variable in Section 4.1. The spectral density function in this case, $f_Z(\lambda)$, is given by

$$f_Z(\lambda) = |\phi(e^{-i\lambda})|^2[f_A(\lambda) + f_\Gamma(\lambda) + f_\Psi(\lambda) + f_\Xi(\lambda)], \quad -\pi < \lambda \leq \pi. \tag{19}$$

Precise expressions for each of these components are given subsequently.

Trend component.

$$f_A(\lambda) = \frac{1}{2\pi} w'(I_2 - Ce^{-i\lambda})^{-1}(V_0 + V_1 e^{-i\lambda} + V_1' e^{i\lambda})(I_2 - C'e^{i\lambda})^{-1}w,$$

$$-\pi < \lambda \leq \pi,$$

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

where w and C are the same as in the stock variable case and the matrices V_0 and V_1 are defined in Theorem 2.

Seasonal component.

$$f_\Gamma(\lambda) = \frac{1}{2\pi}\sum_{j=1}^{s/2} w'[I_2 - C_1(\lambda_j^s)e^{-i\lambda}]^{-1}(V_{0j} + V_{1j}e^{-i\lambda} + V_{1j}'e^{i\lambda}) \times [I_2 - C_1(\lambda_j^s)'e^{i\lambda}]^{-1}w$$

$(-\pi < \lambda \leq \pi)$, where w and $C_1(\lambda)$ are the same as in the stock variable case and the matrices V_{0j} and V_{1j} are defined in Theorem 2.

Cyclical component.

$$f_\Psi(\lambda) = \frac{\sigma_\epsilon^2}{2\pi}\sum_{j=-\infty}^{\infty} \frac{4\sin^2(\lambda/2)}{(\lambda + 2\pi j)^2}\,|a(i(\lambda + 2\pi j))|^{-2}, \quad -\pi < \lambda \leq \pi,$$

where $a(z) = z - a_0 - a_1 e^{-pz}$.

Irregular component.

$$f_\Xi(\lambda) = \frac{\sigma_\Xi^2}{2\pi}, \quad -\pi < \lambda \leq \pi.$$

As in the stock variable case, the functions $f_A(\lambda)$ and $f_\Gamma(\lambda)$ represent pseudospectra. All of the functions for the processes satisfying differential equations are based on the discrete time representations given in Theorem 2, whereas $f_\Psi(\lambda)$ corresponds to (5) in Section 2. The term involving $\sin^2(\lambda/2)$ in $f_\Psi(\lambda)$ arises as a result of the way in which flows are observed, as an integral. This term is the frequency response function of the integrating filter.

4.3. Spectral Densities with the Alternative Cyclical Component

For completeness we also provide the spectral densities for the cyclical component when it is generated by the stochastic differential equation (16). When the variable is a stock, the relevant expression is

$$f_\psi(\lambda) = -\frac{\sigma_\kappa^2(1-\rho^2)}{4\pi\ln\rho} w'[I_2 - \rho C_1(\lambda_c)e^{-i\lambda}]^{-1}[I_2 - \rho C_1(\lambda_c)'e^{i\lambda}]^{-1}w, \quad -\pi < \lambda \leq \pi,$$

where w is as defined before and $C_1(\lambda)$ is defined in the corollary to Theorem 2. When the variable of interest is a flow, the spectral density becomes (for $-\pi < \lambda \leq \pi$)

$$f_\Psi(\lambda) = \frac{1}{2\pi} w'[I_2 - \rho C_1(\lambda_c)e^{-i\lambda}]^{-1}(V_{0K} + V_{1K}e^{-i\lambda} + V_{1K}'e^{i\lambda}) \times [I_2 - \rho C_1(\lambda_c)'e^{i\lambda}]^{-1}w,$$

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

where w and $C_1(\lambda)$ are the same as in the stock variable case and the matrices V_{0K} and V_{1K} are defined in the corollary to Theorem 2. Although time domain methods of estimation are possible when the cycle satisfies (16), the preceding spectral densities can nevertheless be used in the frequency domain approach detailed subsequently.

4.4. Maximum Likelihood Estimation

Under the assumption that the variable $y(t)$ is a Gaussian process, maximization of the Gaussian likelihood function yields maximum likelihood estimators of the unknown parameters. For notational convenience, we shall not distinguish in this section between stocks and flows, so that the transformed variable will simply be denoted z_t. Furthermore, the $m \times 1$ vector of unknown parameters to be estimated will be denoted θ, again regardless of whether the variable is a stock or a flow, so that

$$\theta = (\sigma_\eta^2, \sigma_\zeta^2, \sigma_\omega^2, \sigma_\xi^2, a_0, a_1, p, \sigma_\epsilon^2)'$$

for a stock and

$$\theta = (\sigma_\eta^2, \sigma_\zeta^2, \sigma_\omega^2, \sigma_\Xi^2, a_0, a_1, p, \sigma_\epsilon^2)'$$

for a flow. The spectral density function corresponding to z_t will be denoted $f(\lambda;\theta)$. It is convenient to represent the spectral density function $f(\lambda;\theta)$ in terms of an underlying function $f^c(\lambda;\theta)$ as follows:

$$f(\lambda;\theta) = |\phi(e^{-i\lambda})|^2 \sum_{k=-\infty}^{\infty} f^c(\lambda + 2\pi k;\theta), \quad -\pi < \lambda \leq \pi. \tag{20}$$

Because we have provided exact expressions for the components of $f(\lambda;\theta)$ corresponding to the trend, seasonal, and irregular components of z_t, which do not depend on a doubly infinite series, we define (suppressing the dependence on θ for convenience) $f^c(\lambda)$ in the case of a stock variable as follows:

$$f^c(\lambda + 2\pi k) = \begin{cases} f_\mu(\lambda) + f_\gamma(\lambda) + f_\psi^c(\lambda) + f_\xi(\lambda), & k = 0, \\ f_\psi^c(\lambda + 2\pi k), & k \neq 0, \end{cases}$$

where $f_\psi^c(\lambda) = (\sigma_\epsilon^2/2\pi)|a(i\lambda)|^{-2}$ $(-\infty < \lambda < \infty)$. In the case of a flow, $f^c(\lambda)$ becomes

$$f^c(\lambda + 2\pi k) = \begin{cases} f_A(\lambda) + f_\Gamma(\lambda) + f_\Psi^c(\lambda) + f_\Xi(\lambda), & k = 0, \\ f_\Psi^c(\lambda + 2\pi k), & k \neq 0, \end{cases}.$$

where $f_\Psi^c(\lambda) = (\sigma_\epsilon^2/2\pi)(4\sin^2(\lambda/2)/\lambda^2)|a(i\lambda)|^{-2}$ $(-\infty < \lambda < \infty)$.

The estimation of the continuous time unobserved components model with a differential-difference equation representing the cyclical component means that, in practical terms, it is necessary to approximate, by truncating an infinite series, a component of the spectral density function of the discretely ob-

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

served process. If the truncation point is fixed, that is, independent of sample size T, then even asymptotically the likelihood function will contain a bias term that will potentially cause an inconsistency in the estimator. We therefore wish to investigate the asymptotic properties of the estimator based on the approximate likelihood (using the truncated series for the spectrum) and propose conditions under which this feasible estimator will be consistent and have the same limiting distribution as the full maximum likelihood estimator (using the exact, nontruncated spectrum). In view of our interest being on the effects of the truncation of the spectral density function on the properties of the estimator, we do not seek to provide a set of minimal conditions under which the maximum likelihood estimator (MLE) will have the desired properties. Rather, we impose high-level assumptions that ensure the consistency and asymptotic normality of the MLE, so as to focus more transparently on the properties that the truncation scheme must satisfy for the two estimators to have the same asymptotic properties.

A commonly used and computationally convenient version of $(-2/T)$ times the frequency domain Gaussian log-likelihood function (often called the discrete Whittle likelihood) is given by (apart from a constant term)

$$L(\theta) = \frac{1}{T} \sum_{j \in J_T} \left[\ln f(\lambda_j;\theta) + \frac{I(\lambda_j)}{f(\lambda_j;\theta)} \right], \tag{21}$$

where $J_T = \{j : -T/2 < j \leq [T/2]\}$, $\lambda_j = 2\pi j/T$ for $j \in J_T$, and $I(\lambda)$ denotes the periodogram defined by $I(\lambda) = w(\lambda)w(-\lambda)$, where $w(\lambda) = (2\pi T)^{-1/2} \sum_{t=1}^{T} z_t e^{it\lambda}$. In practice, $f(\lambda;\theta)$ is approximated using a truncated series, given by

$$f^M(\lambda;\theta) = |\phi(e^{-i\lambda})|^2 \sum_{k=-M}^{M} f^c(\lambda + 2\pi k;\theta), \quad -\pi < \lambda \leq \pi, \tag{22}$$

where M denotes the truncation point. It follows that $f(\lambda;\theta) = f^M(\lambda;\theta) + r^M(\lambda;\theta)$ where

$$r^M(\lambda;\theta) = |\phi(e^{-i\lambda})|^2 \sum_{k=M+1}^{\infty} [f^c(\lambda + 2\pi k;\theta) + f^c(\lambda - 2\pi k;\theta)] \tag{23}$$

denotes the error arising from the truncation. The approximate log-likelihood based on the truncated spectrum is therefore

$$L^M(\theta) = \frac{1}{T} \sum_{j \in J_T} \left[\ln f^M(\lambda_j;\theta) + \frac{I(\lambda_j)}{f^M(\lambda_j;\theta)} \right]. \tag{24}$$

Our aim is to examine the properties of (21) and (24) and to compare

$$\hat{\theta}_T = \arg\min_{\theta \in \Theta} L(\theta) \quad \text{with} \quad \hat{\theta}_T^M = \arg\min_{\theta \in \Theta} L^M(\theta),$$

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

where Θ denotes the parameter space. To focus on the properties of the approximation, in particular on the role of the truncation parameter M, we make high-level assumptions concerning the likelihood function $L(\theta)$. The assumptions are given here.

Assumption 1. $\theta_0 \in \Theta$, where Θ is a compact subset of R^m.

Assumption 2. $L(\theta) \xrightarrow{p} l(\theta)$ as $T \to \infty$.

Assumption 3. $l(\theta) > l(\theta_0)$ for $\theta \neq \theta_0$.

Assumption 4. z_t is a stationary Gaussian process and $\sum_{s=0}^{\infty} |\mathrm{cov}(z_t, z_{t+s})| < \infty$.

Assumption 5. $M = O(T^{\delta})$ for some $\delta > 0$.

Assumption 1 is a standard assumption in proofs of consistency for optimization estimators. Assumptions 2 and 3 ensure that, in the limit, the minimizer (over Θ) of $L(\theta)$ (and hence of $l(\theta)$) is unique and equal to θ_0 in probability. More primitive conditions could be used to establish such properties and typically require that $f(\lambda;\theta) \neq f(\lambda;\theta_0)$ for $\theta \neq \theta_0$, an assumption that is implicit here. Assumption 4 stresses that z_t is assumed to be a stationary Gaussian process. The condition on the autocovariances ensures, in particular, that $I(\lambda) = O_p(1)$, a property that is used in the proof of Theorem 3, which follows; see, for example, Brillinger (1975, Theorem 5.2.6). The Gaussian property implicitly requires that the continuous time random measure disturbances are increments of Brownian motion processes. Assumption 4 can, in fact, be used to verify Assumptions 2 and 3, but we state 2 and 3 explicitly as high-level assumptions for convenience. Finally, Assumption 5 specifies an appropriate rate of increase of M with T.

THEOREM 3. *Under Assumptions 1–5,* $\hat{\theta}_T^M \xrightarrow{p} \theta_0$ *as* $T \to \infty$.

The proof of Theorem 3 proceeds by demonstrating that $|L(\theta) - L^M(\theta)| = o_p(1)$ and much of it rests on bounding the error in approximating $f(\lambda;\theta)$, denoted $r^M(\lambda;\theta)$. We note in passing that the consistency of $\hat{\theta}_T$ can also be established under Assumptions 1–3.

To show that $\sqrt{T}(\hat{\theta}_T^M - \theta_0)$ has the same limiting distribution as $\sqrt{T}(\hat{\theta}_T - \theta_0)$, it is convenient to first specify conditions under which $\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{d} N(0, V)$ as $T \to \infty$. We therefore make the following high-level assumptions, in which $s(\theta) = \partial L(\theta)/\partial\theta$ denotes the score vector, $S(\theta) = \partial^2 L(\theta)/\partial\theta\partial\theta'$ denotes the Hessian matrix, and $\bar{\theta}$ satisfies $\|\bar{\theta} - \theta_0\| \leq \|\hat{\theta}_T - \theta_0\|$.

Assumption 6. $S(\bar{\theta}) \xrightarrow{p} S(\theta_0)$ as $T \to \infty$, where $S(\theta_0)$ is a finite positive definite matrix.

Assumption 7. $\sqrt{T}s(\theta_0) \xrightarrow{d} N(0, S(\theta_0))$ as $T \to \infty$.

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

Under Assumptions 6 and 7, and the previously established consistency of $\hat{\theta}_T$ (under Assumptions 1–3), it follows that $\sqrt{T}(\hat{\theta}_T - \theta_0)$ has the required limiting distribution with $V = S(\theta_0)^{-1}$, where

$$S(\theta_0) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\partial \ln f(\lambda;\theta_0)}{\partial \theta} \frac{\partial \ln f(\lambda;\theta_0)}{\partial \theta'} d\lambda \tag{25}$$

and $\partial \ln f(\lambda;\theta_0)/\partial\theta$ denotes the vector $\partial \ln f(\lambda;\theta)/\partial\theta|_{\theta=\theta_0}$. When considering the limiting distribution of the estimator of interest, $\hat{\theta}_T^M$, it is necessary to strengthen the condition concerning the bandwidth exponent δ. Assumption 5 is replaced with Assumption 5′.

Assumption 5′. $M = O(T^{\delta})$ for some $\delta > \frac{1}{2}$.

THEOREM 4. *Under Assumptions 1–4, 5′, 6, and 7,* $\sqrt{T}(\hat{\theta}_T - \hat{\theta}_T^M) = o_p(1)$*, and it follows that* $\sqrt{T}(\hat{\theta}_T^M - \theta_0) \xrightarrow{d} N(0, S(\theta_0)^{-1})$ *as* $T \to \infty$*, where the matrix* $S(\theta_0)$ *is defined in (25).*

Under Theorem 4, asymptotic normality of the estimator requires that the truncation parameter M grows at a sufficiently fast rate relative to sample size T. Although consistency merely requires $M = O(T^{\delta})$ for some $\delta > 0$, asymptotic normality requires $\delta > \frac{1}{2}$. Finally, note that the elements of the matrix $S(\theta_0)$ are typically estimated using $\hat{S}(\hat{\theta}_T^M)$, which is defined by the computationally more convenient expression

$$\hat{S}(\hat{\theta}_T^M) = \frac{1}{2T} \sum_{j \in J_T} \frac{\partial \ln f(\lambda_j;\hat{\theta}_T^M)}{\partial \theta} \frac{\partial \ln f(\lambda_j;\hat{\theta}_T^M)}{\partial \theta'},$$

where $\partial \ln f(\lambda;\hat{\theta}_T^M)/\partial\theta = \partial \ln f(\lambda;\theta)/\partial\theta|_{\theta=\hat{\theta}_T^M}$.

5. SOME SIMULATION EVIDENCE

The previous section established the consistency and asymptotic normality of the Gaussian estimator of θ based on the truncated spectral density. Under Assumption 5 consistency is assured if the truncation parameter $M = O(T^{\delta})$ for some $\delta > 0$, whereas asymptotic normality results if $M = O(T^{\delta})$ for some $\delta > \frac{1}{2}$ (see Assumption 5′). It remains a matter of some interest, however, to investigate the finite sample properties of the estimator based on different choices of δ (and, hence, of M). This section therefore reports the results of a simulation study in an attempt to obtain some evidence as to the properties of the estimator in finite samples.

The simulation experiments are designed to address a number of issues. The first concerns the effect of the choice of M on the properties of the estimator. Closely related to this is the effect of increasing sample size T, but in addition we also investigate the effects on the estimator of cycle length, sampling scheme (i.e., stock vs. flow), and the presence of a stochastic trend. In its most general

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

form, the underlying continuous time model used in the simulations consists of a trend and cycle, defined as follows:

$$y(t) = \mu(t) + \psi(t), \quad d\mu(t) = \eta(dt),$$
$$d\psi(t) = [a_0\psi(t) + a_1\psi(t-p)]dt + \epsilon(dt), \quad t > 0,$$

where $\eta(dt)$ and $\epsilon(dt)$ are mutually and serially uncorrelated random measures, each with variance dt (i.e., $\sigma_\eta^2 = \sigma_\epsilon^2 = 1$). The discrete time observations are given by $y_t = \mu_t + \psi_t + \xi_t$ for a stock variable and by $Y_t = A_t + \Psi_t + \Xi_t$ for a flow, where in each case the irregular component (ξ_t or Ξ_t) is white noise with unit variance. Note that $\mu(t)$ is a continuous time stochastic trend and that $\mu_t = \mu_{t-1} + \eta_t$ where η_t is white noise with unit variance. However, in the case of a flow, $A_t = \int_0^1 \mu(t-r)dr$ also satisfies a unit root process, given by $A_t = A_{t-1} + H_t$, but where $H_t = \int_{t-1}^{t}\int_{s-1}^{s} \eta(dr)ds$ is a first-order moving average process[13] with variance $\frac{2}{3}$ and first-order autocovariance $\frac{1}{6}$. The values of the parameters a_0, a_1, and p in the cyclical component were chosen as follows. Each of a_0 and p were allowed to take on two values, so that $a_0 = \{0.5, -1.0\}$ and $p = \{0.5, 1.0\}$, with a_1 chosen so that for each of the four combinations of a_0 and p there would be a cycle of length four and one of length ten. This results in eight parameter combinations, although for $a_0 = 0.5$ and $p = 1.0$ the value of a_1 (-2.5898) required to generate a cycle of length four violated the stationarity requirement (see Proposition 1) and hence was not used. There were therefore seven stationary combinations of the cyclical parameters, which shall be referred to as Experiments I–VII and which are summarized in Table 1.

When considering the four combinations of stock or flow variable, trend or no trend, this yields a total of 28 experiments. The variance parameters (σ_ϵ^2, σ_ξ^2, σ_Ξ^2, σ_η^2) used to simulate the data are all equal to unity. With regard to the choice of M, estimates were obtained using the rule $M = [T^\delta] + 1(T^\delta \notin \mathcal{N})$, where $[x]$ denotes the integer part of x, $1(x)$ denotes the indicator function (equal to one if x is true and zero otherwise), and $\mathcal{N}$ denotes the set of (positive) integers. Hence M is equal to the integer part of T^δ, rounded upward if T^δ is not

Table 1. Parameter values of cyclical component in simulation experiments

Experiment	a_0	a_1	p	Cycle length
I	0.5	−0.9928	0.5	10
II	0.5	−1.3005	0.5	4
III	−1.0	−0.4690	0.5	10
IV	−1.0	−0.6143	0.5	4
V	0.5	−0.7423	1.0	10
VI	−1.0	−0.1656	1.0	10
VII	−1.0	−0.5778	1.0	4

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

already an integer. Values of $\delta = \{0.25,0.60\}$ were considered, for three sample sizes, $T = \{64,128,256\}$, which gives values of M equal to $\{3,12\}$ ($T = 64$), $\{4,18\}$ ($T = 128$), and $\{4,27\}$ ($T = 256$). A total of 1,000 replications of each experiment were conducted, resulting in $28 \times 2 \times 3 \times 1{,}000 = 168{,}000$ optimizations of the likelihood function. Details concerning the generation of data are provided in the Appendix. To save on space, we report the simulation results corresponding to experiments in which the cycle length is equal to ten, namely, experiments I, III, V, and VI. These appear in Table 2, which contains the mean square errors (multiplied by 10^4) for these four experiments corresponding, in order, to stock and flow variables with no trend and stock and flow variables with a stochastic trend. As an additional space-saving measure, we only report the mean square errors for the parameters a_0, a_1, and p of the differential-difference equation cycle. The full set of results, for all seven experiments and for all parameters (including the variance parameters), is available from the authors on request.[14]

In attempting to summarize the results contained in Table 2, it is convenient to focus, in turn, on each of the issues raised at the start of this section, namely, on the effects on the mean square errors (MSE's) of the following variables.

Truncation parameter M: there are noticeable differences in the MSE's corresponding to different values of M at the smallest sample size (64), and it is interesting to note that it is not always the larger of the two values of M that produces the smallest MSE for $T = 64$. Although some differences remain between the MSE's for larger sample sizes, any such differences are negligible for $T = 256$.

Sample size T: as one would expect, the MSE falls as T increases in every case.

Cycle length: no uniform pattern emerges here, and the cycle length appears to affect the estimation of different parameters in different ways, although we did detect a slight tendency for the MSE to be larger when the cycle was of length four than when it was of length ten.

Stocks versus flows: although there are a few differences to be observed in the results as between stock and flow variables, the differences are fairly minor. Indeed, if the model and likelihood function are correct, we should not expect to find any significant differences.

Stochastic trend: again, some differences emerge, but there appears to be no uniform pattern across the entries in the tables, which is to be expected if the model and likelihood function are correct.

Another matter of interest, not revealed by the reported MSE's, is the extent to which the limiting normal distribution describes the distribution of the estimators in finite samples. Allied to this is the extent of departures from normality when the exponent determining the truncation parameter (δ) is less than one-half. Figure 1, therefore, contains the smoothed empirical densities of the normalized estimator of the lag parameter p, i.e., $\sqrt{T}(\hat{p}_T^M - p_0)$, obtained in Experiment V with $\delta = 0.25$ for a stock variable and $T = 64$, 128, and 256. Figure 1 also contains a normal curve with the same variance as the distribution for $T = 256$. As can be seen, the distributions are reasonably well described by the normal, although, not surprisingly, there is some evidence of

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

TABLE 2. Mean square errors for simulations with cycle length equal to ten

			Without stochastic trend						With stochastic trend					
			Stock variable			Flow variable			Stock variable			Flow variable		
Experiment	δ	T	a_0	a_1	p	a_0	a_1	p	a_0	a_1	p	a_0	a_1	p
I	0.25	64	243	394	1,393	177	448	1,601	1,662	1,431	2,279	830	780	1,466
		128	30	80	494	31	167	469	349	341	586	241	310	504
		256	13	24	101	11	30	67	79	93	127	68	92	118
	0.60	64	220	319	1,408	176	447	1,604	1,666	1,426	2,392	815	769	1,458
		128	30	80	503	32	167	470	352	346	592	242	310	503
		256	13	24	97	11	30	67	78	92	123	68	92	118
III	0.25	64	2,837	12,406	2,438	1,073	3,830	2,633	1,096	3,720	892	1,332	3,553	1,224
		128	456	1,432	377	514	1,107	548	115	130	24	440	534	215
		256	100	205	9	122	167	10	1	3	1	11	8	4
	0.60	64	2,818	12,367	2,644	1,468	4,680	3,630	1,153	3,945	856	1,324	3,559	1,228
		128	520	1,591	395	502	1,065	504	117	128	26	440	538	216
		256	102	208	9	122	167	11	1	2	1	11	8	4
V	0.25	64	31	40	432	44	68	522	130	100	730	224	136	1,019
		128	13	19	201	12	25	216	41	42	339	21	25	120
		256	5	9	78	5	10	92	12	15	103	11	13	105
	0.60	64	30	40	431	44	68	522	157	125	731	224	136	1,019
		128	13	19	201	12	25	216	43	41	334	21	25	120
		256	5	9	78	5	10	92	12	15	102	11	13	105
VI	0.25	64	1,987	9,609	2,236	1,542	4,842	1,621	1,970	5,962	814	1,786	5,334	766
		128	606	3,132	1,035	525	1,629	426	825	1,298	205	819	1,431	186
		256	118	435	186	165	278	11	29	9	4	67	9	4
	0.60	64	2,430	11,423	2,278	1,500	4,987	1,622	2,130	5,793	813	1,805	5,344	768
		128	563	3,029	1,088	523	1,616	425	793	1,268	185	821	1,452	187
		256	116	441	194	165	278	11	30	8	4	67	9	4

Note: All entries have been multiplied by 10^4.

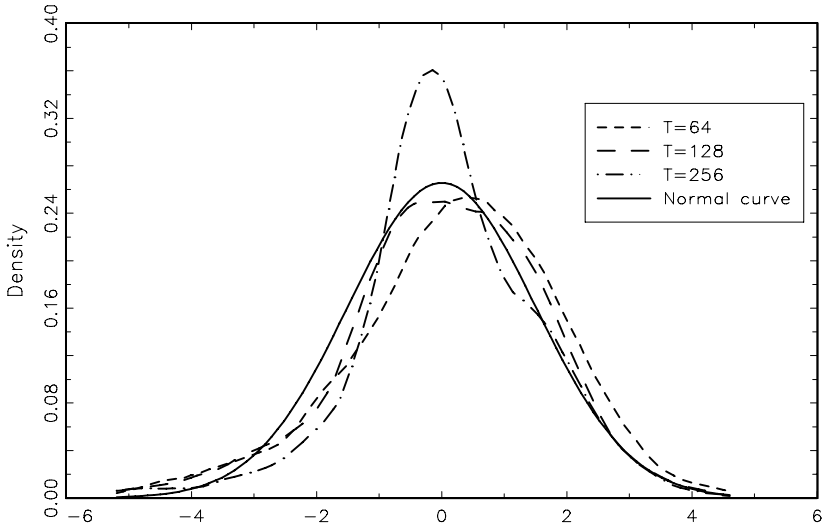


FIGURE 1. Estimated densities of p, Experiment V, $\delta = 0.25$

departures from normality. The curves depicted in Figure 1 are fairly representative of the distributions across experiments and are indistinguishable from the curves in Experiment V with $\delta = 0.6$.

The preceding comments are derived from a visual inspection of the full set of results but are not based on any statistical analysis of the results. A more detailed set of simulation results, derived from a more extensive exploration of the sample space, would enable response surfaces to be estimated in an attempt to summarize the findings. But with only two values of a_0 and p considered, and no variation across the variance parameters, such an exercise in the present circumstances would be of limited reliability.[15] We have therefore made an attempt at carrying out a test of the hypothesis that the means of the distribution of MSE's are equal in the trend–no trend comparison (for given stock or flow variable) and in the stock-flow comparison (for given trend or no trend). The test statistic is based on the difference of the relevant sample means and has an approximate t-distribution.[16] The relevant statistics are presented in Table 3.

The first set of results tests the hypothesis that, for a given sampling scheme (stock or flow), the means of the MSE's are the same regardless of whether a stochastic trend is present. A negative statistic indicates that the mean of the distribution with a stochastic trend is smaller than the mean of the distribution without a stochastic trend. Only one of these statistics is significant at the 5% level, supporting our conjecture from a visual inspection of the results that the presence of a stochastic trend does not significantly affect the MSE's. The second set of statistics tests the hypothesis that, for a given trend scenario, the means of the MSE's are the same regardless of whether the variable is a stock

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

TABLE 3. Test statistics for differences in means of mean square errors

	a_0	a_1	p	σ_ϵ	σ_ξ	σ_η
			Stochastic trend versus no trend			
Stock	−0.7943	−1.3391	−1.1821	−0.1700	0.1363	N.A.
Flow	0.8764	−0.3496	−1.8261*	0.1480	0.1732	N.A.
			Stock versus flow			
No trend	−1.8692*	−1.8783*	0.5543	1.7306	−1.5808	N.A.
Trend	−0.8706	−0.9217	−0.1976	1.9970*	−1.6913*	−1.5130

Note: * denotes significance at the 5% level; N.A. denotes test not applicable.

or a flow. A negative statistic occurs if the mean with a flow is smaller than the mean with a stock. There are four significant entries here, which is somewhat surprising, because the form of the spectral densities, and hence the likelihood itself, should be compensating for these variations in sampling scheme. It must be noted, however, that the significance of the statistics is not great in any of these cases and that the distributions may only be approximate anyway. We are therefore inclined to conclude that there do not appear to be significant differences in the means of the distributions of the MSE's caused either by the sampling scheme or by the presence of a stochastic trend.

6. CONCLUDING COMMENTS

This paper has proposed a class of continuous time unobserved components models in which the cyclical component evolves according to a differential-difference equation, thereby extending the range of models in the literature that until now have relied on a stochastic differential equation (with no lags) to model all unobserved components. Conditions on the parameters of the differential-difference equation are provided under which the process is stationary and under which it admits a business cycle. A frequency domain approach is adopted by necessity in view of the differential-difference equations not being particularly amenable to neat discrete time representations. Spectral density functions are derived when the variable is a stock or a flow, and exact discrete time representations are also derived for the stochastic differential equation based components, which are used in deriving the appropriate spectral densities of these components. In view of the spectrum of the cycle being defined by a doubly infinite series it is necessary to employ some form of truncation to calculate its value at any given frequency. Conditions are provided under which the frequency domain Gaussian estimator will have the same asymptotic properties as the estimator based on the exact spectral density, and

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

some simulation results are reported that investigate the role of the truncation parameter in finite samples.

There are a number of natural routes by which the results in this paper could be extended. Multivariate extensions and the allowance of multiple lags in the differential-difference equation would be relatively straightforward, and the results of Chambers (1998) and McGarry (2000) could be utilized in such a direction, although the derivations of the conditions for stationarity and for the existence of a business cycle would be rather more difficult. A particularly important task is that of undertaking empirical investigations with the proposed models, possibly in the estimation of business cycles with variables such as gross national product and investment. Indeed, it was in this area that differential-difference equations were introduced into economics by Kalecki (1935). It would also be possible to allow for long memory in the trend component by utilizing a fractional derivative of order $\frac{1}{2} < k < \frac{3}{2}$ along the lines developed in Chambers (1996). In this case it would also be necessary to truncate the doubly infinite series defining the spectral density function of the long memory trend component, and the analysis of Section 4 concerning such truncation would be pertinent. Of related interest is the investigation of the extent to which the differential-difference equation cyclical component can estimate cycle length in circumstances in which the cycle is not actually generated by a differential-difference equation. Such issues of cycle length estimation under misspecification can be investigated using appropriately designed sampling experiments. Some of these issues are currently under investigation by the authors.

NOTES

1. See, for example, Boucekkine, Licandro and Paul (1997) and Asea and Zak (1999). Chambers (1998) also provides some examples.

2. One of the few empirical applications can be found in McGarry (2000), for example.

3. We are grateful to a referee for emphasizing the path-dependence feature.

4. For certain (more complicated) differential-difference equation systems, verifying the required stationarity conditions could prove to be intractable.

5. See, for example, Hannan (1970, p. 68).

6. For the model defined in (1), for a stock variable there are two terms in the resulting summation, whereas for a flow there are three terms. Equation (22) of Chambers (1998) provides the spectral density function for a stock variable generated by (1) with $p = 1$.

7. Note that in our definitions we use random measures, such as $\eta(dt)$, to capture the disturbance terms, whereas Harvey and Stock (1993) use increments in random processes, such as $d\eta(t)$, for this purpose. These differences can best be described as differences in style rather than in substance, because the properties of the resulting series are not drastically affected by these alternative definitions.

8. See, for example, Harvey and Stock (1993).

9. See Rozanov (1967) and Bergstrom (1984) for details of integration with respect to a random measure.

10. See Bergstrom (1984) for a definition of wide sense integration of continuous time processes.

11. See, for example, Theorem 8 of Bergstrom (1984).

12. In the so-called local level model, where $\sigma_\zeta^2 = 0$, and $\beta(t) = \beta(0)$, $\mu(t)$ follows a random walk, and hence the appropriate operator is Δ rather than Δ^2.

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

13. See the trend component in Theorem 2 and set $B_t = 0$, $\sigma_\xi^2 = 0$, and $\sigma_\eta^2 = 1$. The result can also be derived directly from first principles by integrating $d\mu(t) = \eta(dt)$ twice.

14. The results can also be downloaded from URL http://privatewww.essex.ac.uk/~mchamb.

15. Simulation exercises carried out for models with perhaps only two free parameters and using least squares and related methods can afford to be more extensive in the exploration of the parameter space as a result of the ease with which simulations can be executed. We would remind the reader again that the results we have obtained have required 168,000 highly nonlinear optimizations of the likelihood function involving, in each case, at least five unknown parameters.

16. The test relies on normality in finite samples and on the equivalence of variances in the two distributions. Although there is evidence to suggest that these two conditions may not be met in all cases, we carry out the tests regarding the *t*-distribution as holding approximately.

REFERENCES

Asea, P.K. & P.J. Zak (1999) Time-to-build and cycles. *Journal of Economic Dynamics and Control* 23, 1155–1175.

Bergstrom, A.R. (1984) Continuous time stochastic models and issues of aggregation over time. In Z. Griliches & M.D. Intriligator (eds.), *Handbook of Econometrics*, vol. 2, pp. 1145–1212. Amsterdam: North-Holland.

Boucekkine, R., O. Licandro, & C. Paul (1997) Differential-difference equations in economics: On the numerical solution of vintage capital growth models. *Journal of Economic Dynamics and Control* 21, 347–362.

Brillinger, D.R. (1975) *Time Series: Data Analysis and Theory*. New York: Holt, Rinehart, and Winston.

Chambers, M.J. (1996) The estimation of continuous parameter long-memory time series models. *Econometric Theory* 12, 274–290.

Chambers, M.J. (1998). The estimation of systems of joint differential-difference equations. *Journal of Econometrics* 85, 1–31.

Chambers, M.J. (1999) Discrete time representation of stationary and nonstationary continuous time systems. *Journal of Economic Dynamics and Control* 23, 619–639.

Frisch, R. & H. Holme (1935) The characteristic solutions of a mixed difference and differential equation occurring in economic dynamics. *Econometrica* 3, 225–239.

Gradshteyn, I.S. & I.M. Ryzhik (1994) *Table of Integrals, Series, and Products*, 5th ed. New York: Academic Press.

Hannan, E.J. (1970) *Multiple Time Series*. New York: Wiley.

Hannan, E.J. & P.M. Robinson (1973) Lagged regression with unknown lags. *Journal of the Royal Statistical Society, Series B* 35, 252–267.

Hannan, E.J. & P.J. Thomson (1971) The estimation of coherence and group delay. *Biometrika* 58, 469–482.

Hannan, E.J. & P.J. Thomson (1973) Estimating group delay. *Biometrika* 60, 241–253.

Hannan, E.J. & P.J. Thomson (1981) Delay estimation and the estimation of coherence and phase. *IEEE Transactions on Acoustics, Speech, and Signal Processing* ASSP-29, 485–490.

Hannan, E.J. & P.J. Thomson (1988) Time delay estimation. *Journal of Time Series Analysis* 9, 21–33.

Harvey, A.C. (1989) *Forecasting, Structural Time Series Models, and the Kalman Filter*. Cambridge: Cambridge University Press.

Harvey, A.C. & J.H. Stock (1993) Estimation, smoothing, interpolation, and distribution for structural time series models in continuous time. In P.C.B. Phillips (ed.), *Models, Methods, and Applications of Econometrics: Essays in Honour of A. R. Bergstrom*, pp. 55–70. Oxford: Blackwell.

James, R.W. & M.H. Belz (1936) On a mixed difference and differential equation. *Econometrica* 4, 157–160.

Kalecki, M. (1935) A macrodynamic theory of business cycles. *Econometrica* 3, 327–344.

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

McGarry, J.S. (2000) Estimating Lag Parameters in Differential-Difference Equations. Economic research paper 00/14, Loughborough University.
Robinson, P.M. (1974) Stochastic difference equations with non-integral differences. *Advances in Applied Probability* 6, 524–545.
Robinson, P.M. (1988) The stochastic difference between econometric statistics. *Econometrica* 56, 531–548.
Rozanov, Y.A. (1967). *Stationary Random Processes*. New York: Holden Day.

APPENDIX

Proof of Proposition 2. To derive the length of the cycle, consider the deterministic equation

$$Dy(t) = a_0 y(t) + a_1 y(t-p), \tag{A.1}$$

in which $D = d/dt$ denotes the usual (nonstochastic) differential operator. Note that solutions to (A.1) are of the form $Ke^{\alpha t}$, where K is some constant. Substituting into (A.1) and dividing through by $Ke^{\alpha t}$ yields the equation $\alpha = a_0 + a_1 e^{-\alpha p}$, which can be expressed as $-a_1 p e^{-a_0 p} e^{(a_0-\alpha)p} = (a_0 - \alpha)p$. Defining $z = (a_0 - \alpha)p = q + ir$ and $\sigma = -a_1 p e^{-a_0 p}$ this can be expressed more simply as $\sigma e^z = z$ or $q + ir = \sigma e^q(\cos r + i \sin r)$. When cycles exist, there will be an infinite number of solutions to this equation of the form $q_k + ir_k$. We are looking for the smallest r_k, call it r_1, in which case the period of the cycle is $2\pi p/r_1$ (see Kalecki, 1935, pp. 332–336). The real and imaginary parts of the preceding equation are, respectively,

$$q = \sigma e^q \cos r, \tag{A.2}$$

$$r = \sigma e^q \sin r. \tag{A.3}$$

From (A.3) it follows that $\sigma e^q = r/\sin r$, and substituting this into (A.2) yields $q = r \cot r$. Dividing (A.3) by r gives $1 = \sigma e^q \sin r/r = |\sigma| e^q |\sin r/r|$. Taking logs and rearranging yields a second expression for q, namely, $q = -\ln|\sigma| - \ln|\sin r/r|$. Equating these two expressions for q yields $r \cot r + \ln|\sin r/r| = -\ln|\sigma|$, which, when substituting for $\ln|\sigma|$, gives the required expression.

Regarding the existence of cycles, it is required that

$$r \cot r + \ln\left|\frac{\sin r}{r}\right| = a_0 p - \ln|-a_1 p| < 1$$

(see James and Belz, 1936). The condition stated in Proposition 1 follows immediately. ■

Proof of Theorem 1. The solution to (15) is given by

$$x(t) = e^{At}x(0) + \int_0^t e^{A(t-r)}\nu(dr), \quad t > 0,$$

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

which yields the stochastic difference equation

$$x(t) = e^A x(t-1) + \int_{t-1}^{t} e^{A(t-r)}\nu(dr), \quad t = 1,\ldots,T$$

(see, e.g., Bergstrom, 1984, Theorem 3). Integrating over $(t-1, t]$ yields

$$X_t = e^A X_{t-1} + N_t, \quad t = 1,\ldots,T,$$

where $N_t = \int_{t-1}^{t}\int_{s-1}^{s} e^{A(s-r)}\nu(dr)ds$. From Chambers (1999, p. 630) we obtain the equivalent representation of N_t as a pair of single integrals with respect to $\nu(dt)$, given by

$$N_t = \int_{t-1}^{t} \Phi(t-r)\nu(dr) + \int_{t-2}^{t-1} [\Phi(1) - \Phi(t-1-r)]\nu(dr).$$

It follows that

$$E(N_t N_t') = V_0 = \int_0^1 \Phi(s)\Sigma_\nu \Phi(s)'ds + \int_0^1 [\Phi(1) - \Phi(s)]\Sigma_\nu [\Phi(1) - \Phi(s)]'ds,$$

$$E(N_t N_{t-1}') = V_1 = \int_0^1 [\Phi(1) - \Phi(s)]\Sigma_\nu \Phi(s)'ds,$$

which yields the expressions given in the theorem. ■

Proof of Theorem 2. The proof proceeds by applying Theorem 1 to each of the unobserved components and deriving the matrices e^A, V_0, and V_1 in turn. The latter autocovariance matrices require evaluation of $\Phi(r)$, $\Phi(1)$, $\Upsilon(1)$, and $\Lambda(1)$. The proof for each component will be given in turn.

Trend component. The matrix A is the coefficient matrix in (6), and it follows that $A^j = 0$ for $j > 1$. Hence $e^{Ar} = I_2 + Ar$, and so $\Phi(s) = \int_0^s e^{Ar}dr = sI_2 + (s^2/2)A$ and $\Upsilon(1) = \int_0^1 \Phi(s)ds = (1/2)I_2 + (1/6)A$. It also follows, because $\Lambda(1) = \int_0^1 \Phi(s)\Sigma_\nu\Phi(s)'ds$, that

$$\Lambda(1) = \int_0^1 \left[sI_2 + \frac{s^2}{2}A\right]\Sigma_\nu\left[sI_2 + \frac{s^2}{2}A\right]'ds = \frac{1}{3}\Sigma_\nu + \frac{1}{8}A\Sigma_\nu + \frac{1}{8}\Sigma_\nu A' + \frac{1}{20}A\Sigma_\nu A',$$

where Σ_ν is the 2×2 diagonal matrix with σ_η^2 and σ_ξ^2 on the diagonal. Combining these terms yields the expressions for V_0 and V_1.

Seasonal component. The matrix A is the coefficient matrix in (7), and $\Sigma_\nu = \sigma_\omega^2 I_2$. It can be shown that

$$e^{As} = \begin{bmatrix} \cos s\lambda_j^s & \sin s\lambda_j^s \\ -\sin s\lambda_j^s & \cos s\lambda_j^s \end{bmatrix},$$

and so to evaluate $\Phi(s)$ it is necessary to evaluate the integrals $\int_0^s \cos r\lambda_j^s dr = (1/\lambda_j^s)\sin s\lambda_j^s$ and $\int_0^s \sin r\lambda_j^s dr = (1/\lambda_j^s)(1 - \cos s\lambda_j^s)$. Application of these formulae yields

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

$$\Phi(1) = \frac{1}{\lambda_j^s}\begin{bmatrix} \sin\lambda_j^s & 1-\cos\lambda_j^s \\ \cos\lambda_j^s - 1 & \sin\lambda_j^s \end{bmatrix}, \quad \Upsilon(1) = \frac{1}{(\lambda_j^s)^2}\begin{bmatrix} 1-\cos\lambda_j^s & \lambda_j^s - \sin\lambda_j^s \\ \sin\lambda_j^s - \lambda_j^s & 1-\cos\lambda_j^s \end{bmatrix}.$$

Furthermore, because $\Phi(s)\Phi(s)' = [2(1-\cos s\lambda_j^s)/(\lambda_j^s)^2]I_2$, we obtain

$$\Lambda(1) = \sigma_\omega^2 \int_0^1 \Phi(s)\Phi(s)'ds = \frac{2\sigma_\omega^2}{(\lambda_j^s)^3}\begin{bmatrix} \lambda_j^s - \sin\lambda_j^s & 0 \\ 0 & \lambda_j^s - \sin\lambda_j^s \end{bmatrix}.$$

Combining these terms yields the expressions for V_{0j} and V_{1j} in the theorem. ■

Proof of Corollary to Theorem 2. The matrix A is the coefficient matrix in (16), and $\Sigma_\nu = \sigma_\kappa I_2$. It can be shown that

$$e^{As} = \rho^s\begin{bmatrix} \cos s\lambda_c & \sin s\lambda_c \\ -\sin s\lambda_c & \cos s\lambda_c \end{bmatrix},$$

and so to evaluate $\Phi(s)$ it is necessary to evaluate $\int_0^s \rho^r \cos r\lambda_c dr$ and $\int_0^s \rho^r \sin r\lambda_c dr$. Noting that $\rho^r = e^{r\ln\rho}$, Gradshteyn and Ryzhik (1994, equations 2.663.1 and 2.663.2) yield, respectively,

$$\int_0^s \rho^r \sin r\lambda_c dr = \frac{\rho^s(\ln\rho\sin s\lambda_c - \lambda_c\cos s\lambda_c) + \lambda_c}{\delta(\lambda_c)},$$

$$\int_0^s \rho^r \cos r\lambda_c dr = \frac{\rho^s(\ln\rho\cos s\lambda_c + \lambda_c\sin s\lambda_c) - \ln\rho}{\delta(\lambda_c)}.$$

Application of these formulae yields

$$\Phi(1) = \frac{\rho\ln\rho}{\delta(\lambda_c)}C_1(\lambda_c) + \frac{\rho\lambda_c}{\delta(\lambda_c)}C_2(\lambda_c) - C_3(\lambda_c),$$

$$\Upsilon(1) = \frac{\rho[(\ln\rho)^2 - \lambda_c^2]}{\delta(\lambda_c)^2}C_1(\lambda_c) + \frac{2\rho\lambda_c\ln\rho}{\delta(\lambda_c)^2}C_2(\lambda_c) - C_4(\lambda_c).$$

Now, because $\Phi(s)\Phi(s)' = \{[1 + \rho^{2s} - 2\rho^s\cos s\lambda_c]/\delta(\lambda_c)^2\}I_2$ and $\Lambda(1) = \sigma_\kappa^2\int_0^1\Phi(s)\Phi(s)'ds$, we obtain $\Lambda(1) = \sigma_\kappa^2 c_1(\lambda_c)I_2$. Combining these results yields, after some manipulations, the sequence of formulae defining V_{0K} and V_{1K} in the theorem. ■

Proof of Theorem 3. In what follows, it is convenient notationally to define $f_j = f(\lambda_j;\theta)$ and to similarly define f_j^M, r_j^M, and I_j. Note that Assumptions 1–3 ensure that $\hat\theta_T \xrightarrow{p} \theta_0$. If we can show that $|L(\theta) - L^M(\theta)| = o_p(1)$, then $\hat\theta_T^M$ will also be consistent. Because $f_j^M = f_j - r_j^M$, it follows that

$$L^M(\theta) = \frac{1}{T}\sum_{j\in J_T}\left[\ln(f_j - r_j^M) + \frac{I_j}{f_j - r_j^M}\right].$$

Note that $0 \le |\phi(e^{-i\lambda})|^2 \le P_1 < \infty$ for some constant P_1 (depending on the nature of the filter needed to induce stationarity) and for all λ. The results of McGarry (1998, Proposition 2) enable bounds for r_j^M to be derived, the forms of which depend on the

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

relative magnitudes of $|a_0|$ and $|a_1|$. In what follows, bounds involving P_1 will be absorbed within the generic constants (e.g., k_1) used subsequently. For $|a_0| < 2|a_1|$,

$$\begin{aligned} 0 \le r_j^M &\le k_1 \ln\left|\frac{2\pi M + \lambda_j + k_2 + k_3}{2\pi M + \lambda_j + k_2 - k_3}\right| + k_1 \ln\left|\frac{2\pi M - \lambda_j + k_2 + k_3}{2\pi M - \lambda_j + k_2 - k_3}\right| \\ &= k_1 \ln\left|1 + \frac{2k_3}{2\pi M + \lambda_j + k_2 - k_3}\right| + k_1 \ln\left|1 + \frac{2k_3}{2\pi M - \lambda_j + k_2 - k_3}\right| \\ &= k_1 \ln|1 + x_{1j}^M| + k_1 \ln|1 + x_{2j}^M|, \end{aligned}$$

where k_1, k_2, and k_3 are constants depending on the parameters of the differential-difference equation (DDE) and where $0 < |x_{1j}^M|, |x_{2j}^M| < 1$ for all j and large enough M. Because $\ln(1 + x) = O(x)$ as $x \to 0$ we have, for fixed j, $r_j^M = O(x_{1j}^M) + O(x_{2j}^M) = O(M^{-1})$. For $|a_0| > 2|a_1|$,

$$0 \le r_j^M \le k_4\left[\pi - \arctan\left(\frac{2\pi M + \lambda_j + c_2}{k_5}\right) - \arctan\left(\frac{2\pi M - \lambda_j + c_2}{k_5}\right)\right],$$

where, once more, the constants k_4 and k_5 depend on the parameters of the DDE. Because $\arctan(x) = (\pi/2) - O(x^{-1})$ as $x \to \infty$, it follows in this case also that $r_j^M = O(M^{-1})$. The final case, when $|a_0| = 2|a_1|$, yields

$$0 \le r_j^M \le \frac{k_6}{2\pi M + \lambda_j + k_2} + \frac{k_6}{2\pi M - \lambda_j + k_2} = O(M^{-1}),$$

where the constant k_6 depends on the parameters of the model.

Now consider the difference

$$L(\theta) - L^M(\theta) = \frac{1}{T}\sum_{j \in J_T}\left[\ln\left(\frac{f_j}{f_j - r_j^M}\right) + \frac{I_j}{f_j} - \frac{I_j}{f_j - r_j^M}\right].$$

Note that

$$\frac{1}{f_j - r_j^M} = \frac{1}{f_j(1 - r_j^M/f_j)} = \frac{1}{f_j}\left(1 + \frac{r_j^M}{f_j} + \frac{r_j^{M2}}{f_j^2} + \dots\right) = \frac{1}{f_j} + e_j^M,$$

where $e_j^M = f_j^{-1}\sum_{k=1}^{\infty} r_j^{Mk}/f_j^k = O(M^{-1})$. Note that e_j^M is convergent in view of the fact that $r_j^M/f_j < 1$ for all j. Therefore $|L(\theta) - L^M(\theta)| \le \delta_1^M(\theta) + \delta_2^M(\theta)$, where

$$\delta_1^M(\theta) = \left|\frac{1}{T}\sum_{j \in J_T}\ln\left(1 + \frac{r_j^M}{f_j - r_j^M}\right)\right| = O(M^{-1}) \quad \text{and} \quad \delta_2^M(\theta) = \left|\frac{1}{T}\sum_{j \in J_T} I_j e_j^M\right|.$$

Because, under Assumption 4, $I_j = O_p(1)$, it follows that $\delta_2(\theta) = O_p(M^{-1})$. Hence, under Assumption 5, $|L(\theta) - L^M(\theta)| \le O(M^{-1}) + O_p(M^{-1}) = o_p(T^{-\delta+\epsilon})$ for some $0 < \epsilon < \delta$, which is $o_p(1)$ as required. ■

Proof of Theorem 4. A mean value expansion of $s(\hat{\theta}_T) = \partial L(\theta)/\partial\theta|_{\theta=\hat{\theta}_T}$ around θ_0 yields $s(\hat{\theta}_T) = 0 = s(\theta_0) + S(\bar{\theta})(\hat{\theta}_T - \theta_0)$, where $S(\bar{\theta}) = \partial^2 L(\theta)/\partial\theta\partial\theta'|_{\theta=\bar{\theta}}$ and $\|\bar{\theta} - \theta_0\| \le \|\hat{\theta}_T - \theta_0\|$. This results in

$$\sqrt{T}(\hat{\theta}_T - \theta_0) = -S(\bar{\theta})^{-1}\sqrt{T}s(\theta_0), \tag{A.4}$$

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

and the stated distribution follows immediately using Assumptions 6 and 7. For $\sqrt{T}(\hat{\theta}_T^M - \theta_0) \xrightarrow{d} N(0,V)$ also, it is sufficient to establish the property that $\sqrt{T}(\hat{\theta}_T - \hat{\theta}_T^M) = o_p(1)$. General results for establishing this kind of property can be found in Robinson (1988). We proceed, however, by using a mean value expansion of the score vector for determining $\hat{\theta}_T^M$, and we obtain

$$\sqrt{T}(\hat{\theta}_T^M - \theta_0) = -S^M(\tilde{\theta})^{-1}\sqrt{T}s^M(\theta_0), \tag{A.5}$$

where $\tilde{\theta}$ satisfies $\|\tilde{\theta} - \theta_0\| \leq \|\hat{\theta}_T - \theta_0\|$ and $S^M(\tilde{\theta})$ and $s^M(\theta_0)$ represent, respectively, the Hessian matrix and the score vector based on the approximate likelihood function $L^M(\theta)$. Comparing (A.4) with (A.5) yields

$$\begin{aligned}\sqrt{T}(\hat{\theta}_T - \hat{\theta}_T^M) &= S^M(\tilde{\theta})^{-1}\sqrt{T}s^M(\theta_0) - S(\bar{\theta})^{-1}\sqrt{T}s(\theta_0)\\ &= [S^M(\tilde{\theta})^{-1} - S(\bar{\theta})^{-1}]\sqrt{T}s^M(\theta_0) + S(\bar{\theta})^{-1}\sqrt{T}[s^M(\theta_0) - s(\theta_0)]. \end{aligned}\tag{A.6}$$

Under Assumption 6, $S(\bar{\theta})^{-1} = O_p(1)$, whereas under Assumption 7, $\sqrt{T}s(\theta_0) = O_p(1)$. Hence to show that $\sqrt{T}(\hat{\theta}_T - \hat{\theta}_T^M) = o_p(1)$, it is therefore sufficient to show that

(i) $\sqrt{T}[s^M(\theta_0) - s(\theta_0)] = o_p(1)$, and
(ii) $S^M(\tilde{\theta})^{-1} - S(\bar{\theta})^{-1} = o_p(1)$.

We shall treat each of these terms in turn.

(i) From the form of the likelihood functions we obtain

$$s(\theta) = \frac{1}{T}\sum_{j\in J_T}\left(1 - \frac{I_j}{f_j}\right)\frac{\partial \ln f_j}{\partial \theta}, \qquad s^M(\theta) = \frac{1}{T}\sum_{j\in J_T}\left(1 - \frac{I_j}{f_j^M}\right)\frac{\partial \ln f_j^M}{\partial \theta}.$$

We have already established that $(1/f_j^M) = (1/f_j) + e_j^M$ where $e_j^M = O(M^{-1})$, and it is easy to show that

$$\ln f_j^M = \ln(f_j - r_j^M) = \ln[f_j(1 - r_j^M/f_j)] = \ln f_j + \ln(1 - r_j^M/f_j) = \ln f_j + \kappa_{0j}^M,$$

where $\kappa_{0j}^M = O(M^{-1})$. It is also possible to show that

$$\frac{\partial \ln f_j^M}{\partial \theta} = \frac{\partial \ln f_j}{\partial \theta} + \kappa_{1j}^M,$$

where the elements of the $m \times 1$ vector κ_{1j}^M are $O(M^{-1})$. Defining $f_{0j} = f(\lambda_j;\theta_0)$ and letting $\partial \ln f_{0j}/\partial\theta$ denote the derivative vector evaluated at $\theta = \theta_0$, we obtain

$$\begin{aligned}\sqrt{T}s^M(\theta_0) &= \frac{1}{\sqrt{T}}\sum_{j\in J_T}\left(1 - \frac{I_j}{f_{0j}} - I_j e_j^M\right)\left(\frac{\partial \ln f_{0j}}{\partial\theta} + \kappa_{1j}^M\right)\\ &= \frac{1}{\sqrt{T}}\sum_{j\in J_T}\left(1 - \frac{I_j}{f_{0j}}\right)\frac{\partial \ln f_{0j}}{\partial\theta} + \frac{1}{\sqrt{T}}\sum_{j\in J_T}\left(1 - \frac{I_j}{f_{0j}}\right)\kappa_{1j}^M\\ &\quad - \frac{1}{\sqrt{T}}\sum_{j\in J_T} I_j e_j^M\left(\frac{\partial \ln f_{0j}}{\partial\theta} + \kappa_{1j}^M\right)\\ &= \sqrt{T}s(\theta_0) + O_p(T^{1/2}M^{-1}).\end{aligned}$$

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

Therefore, $\sqrt{T}[s^M(\theta_0) - s(\theta_0)] = O_p(T^{1/2}M^{-1}) = O_p(T^{1/2-\delta})$ under Assumption 5′, which is $o_p(1)$ as required.

(ii) From the definitions of the score vectors we obtain

$$S(\theta) = \frac{1}{T}\sum_{j\in J_T}\left(1 - \frac{I_j}{f_j}\right)\frac{\partial^2 \ln f_j}{\partial\theta\partial\theta'} + \frac{1}{T}\sum_{j\in J_T}\frac{I_j}{f_j}\frac{\partial \ln f_j}{\partial\theta}\frac{\partial \ln f_j}{\partial\theta'},$$

and similarly

$$S^M(\theta) = \frac{1}{T}\sum_{j\in J_T}\left(1 - \frac{I_j}{f_j^M}\right)\frac{\partial^2 \ln f_j^M}{\partial\theta\partial\theta'} + \frac{1}{T}\sum_{j\in J_T}\frac{I_j}{f_j^M}\frac{\partial \ln f_j^M}{\partial\theta}\frac{\partial \ln f_j^M}{\partial\theta'}.$$

In fact, the first term in $S(\theta)$ is $o_p(1)$. It can be established that

$$\frac{\partial^2 \ln f_j^M}{\partial\theta\partial\theta'} = \frac{\partial^2 \ln f_j}{\partial\theta\partial\theta'} + \kappa_{2j}^M,$$

where $\kappa_{2j}^M = O(M^{-1})$. It follows that

$$\begin{aligned} S^M(\theta) &= \frac{1}{T}\sum_{j\in J_T}\left(1 - \frac{I_j}{f_j} - I_j e_j^M\right)\left(\frac{\partial^2 \ln f_j}{\partial\theta\partial\theta'} + \kappa_{2j}^M\right) \\ &\quad + \frac{1}{T}\sum_{j\in J_T}\left(\frac{I_j}{f_j} + I_j e_j^M\right)\left(\frac{\partial \ln f_j}{\partial\theta} + \kappa_{1j}^M\right)\left(\frac{\partial \ln f_j}{\partial\theta'} + \kappa_{1j}^{M\prime}\right) \\ &= S_1^M(\theta) + S_2^M(\theta), \end{aligned}$$

where

$$\begin{aligned} S_1^M(\theta) &= \frac{1}{T}\sum_{j\in J_T}\left(1 - \frac{I_j}{f_j}\right)\frac{\partial^2 \ln f_j}{\partial\theta\partial\theta'} + \frac{1}{T}\sum_{j\in J_T}\left(1 - \frac{I_j}{f_j}\right)\kappa_{2j}^M \\ &\quad - \frac{1}{T}\sum_{j\in J_T} I_j e_j^M\left(\frac{\partial^2 \ln f_j}{\partial\theta\partial\theta'} + \kappa_{2j}^M\right) \\ &= \frac{1}{T}\sum_{j\in J_T}\left(1 - \frac{I_j}{f_j}\right)\frac{\partial^2 \ln f_j}{\partial\theta\partial\theta'} + O_p(M^{-1}), \\ S_2^M(\theta) &= \frac{1}{T}\sum_{j\in J_T}\frac{I_j}{f_j}\frac{\partial \ln f_j}{\partial\theta}\frac{\partial \ln f_j}{\partial\theta'} \\ &\quad + \frac{1}{T}\sum_{j\in J_T}\frac{I_j}{f_j}\left(\frac{\partial \ln f_j}{\partial\theta}\kappa_{1j}^{M\prime} + \kappa_{1j}^M\frac{\partial \ln f_j}{\partial\theta'} + \kappa_{1j}^M\kappa_{1j}^{M\prime}\right) \\ &\quad + \frac{1}{T}\sum_{j\in J_T} I_j e_j^M\left(\frac{\partial \ln f_j}{\partial\theta} + \kappa_{1j}^M\right)\left(\frac{\partial \ln f_j}{\partial\theta'} + \kappa_{1j}^{M\prime}\right) \\ &= \frac{1}{T}\sum_{j\in J_T}\frac{I_j}{f_j}\frac{\partial \ln f_j}{\partial\theta}\frac{\partial \ln f_j}{\partial\theta'} + O_p(M^{-1}). \end{aligned}$$

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press

It follows that $S^M(\theta) - S(\theta) = O_p(M^{-1}) = O_p(T^{-\delta})$ under Assumption 5′, and hence $S^M(\theta) = S(\theta) + o_p(1)$. Slutsky's theorem and the previously established consistency of the two estimators then ensure that $S^M(\tilde{\theta})^{-1} - S(\bar{\theta})^{-1} = o_p(1)$ as required. ■

Generation of Data. The data for the trend component were generated according to the difference equations provided in Section 3. For a stock variable, $\mu_t = \mu_{t-1} + \eta_t$ with $\eta_t \sim NID(0,1)$, whereas for a flow variable $A_t = A_{t-1} + H_t$ where H_t is required to be MA(1) with variance $V_0 = 2\sigma_\eta^2/3$ and first-order autocorrelation $V_1 = \sigma_\eta^2/6$. Suppose $H_t = e_t + \alpha e_{t-1}$ where $e_t \sim IID(0,\sigma_e^2)$. Then $V_0 = (1 + \alpha^2)\sigma_e^2$ and $V_1 = \alpha\sigma_e^2$. Combining the two expressions for V_0 and the two for V_1 yields the equation $(\alpha^2 - 4\alpha + 1) = 0$, which can be solved to give $\alpha = 2 \pm \sqrt{3}$. The invertible root is $\alpha = 2 - \sqrt{3}$, and this is the value used to generate H_t using the preceding MA representation. It is also possible to show that $\sigma_e^2 = \sigma_\eta^2/(6\alpha)$, and so we took $e_t \sim NID(0,1/6\alpha)$ because $\sigma_\eta^2 = 1$. The irregular component was straightforward, and we generated $\xi_t \sim NID(0,1)$ for stocks and $\Xi_t \sim NID(0,1)$ for flows.

Generating discrete observations on the cyclical component from a differential-difference equation is more difficult in view of the lack of an exact discrete time representation. We therefore used a discretization method, dividing each unit time interval into 100 subintervals of length $h = 0.01$ and generating

$$\psi(t) - \psi(t-h) = [a_0\psi(t-h) + a_1\psi(t-p-h)]h + \sigma_\epsilon\sqrt{h}\epsilon(t),$$

where $\epsilon(t) \sim NID(0,\sigma_\epsilon^2)$ with $\sigma_\epsilon^2 = 1$. For a stock variable, every 100th observation is chosen, whereas for a flow the average value over every successive 100 observations is recorded. Although there is an element of approximation inherent in this approach, its effects do not appear to have been too severe.

https://doi.org/10.1017/S0266466602182077 Published online by Cambridge University Press