Modelling Cyclical Behaviour with Differential-Difference Equations in an Unobserved Components Framework

by

Marcus J. Chambers
University of Essex

and

Joanne S. McGarry
Loughborough University

September 1999

Abstract

This paper considers a continuous time unobserved components model in which the cyclical component follows a differential-difference equation while 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 analysed 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, as well as of an empirical application using data on the US gross national product, are also provided.

Address for correspondence: Professor Marcus J. Chambers, Department of Economics, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, England. Tel: +44 1206 872756; fax: +44 1206 872724; e-mail: mchamb@essex.ac.uk.

1. INTRODUCTION

The analysis of cycles in economic behaviour has been an important pursuit of economists since at least the mid-nineteenth century, and much research effort is still expended in an attempt to adequately describe the nature of business cycle fluctuations. The application of time series methodology plays a leading role in the empirical analysis of business cycles and a diverse range of techniques can be brought to bear in the quest for establishing empirical regularities.

The present contribution aims to integrate a particular type of model that has been proposed for the analysis of economic cycles within a flexible time series modelling framework. The model under consideration is characterised by a differential-difference equation and its presence in economics can be traced back to the 'macrodynamic theory of the business cycle' of Kalecki (1935). The source of the lag parameter in Kalecki's differential-difference equation is due to what has become known more recently as a time-to-build effect, which is characterised by a discrete time lag between the placing of investment orders and their realisation. Differential-difference equations have also appeared in other branches of economics[1] as well as in other disciplines, such as physics, engineering and biology. There has, in fact, recently been a resurgence of interest in differential-difference equations in economics,[2] although so far there appears to have been little (or no) empirical application of such models. The recent econometric implementations of differential-difference equations by Chambers (1998) and McGarry (1998), allied to the results in the present paper, should help to encourage empirical investigations using differential-difference equations.

The framework within which we incorporate a differential-difference equation in this paper is that of an unobserved components model in continuous time, in which the time series of interest in its most general form is comprised of trend, seasonal, cyclical and irregular components. The cyclical component is assumed to be governed by a differential-difference equation, while the trend and seasonal evolve according to standard stochastic differential equations (i.e. equations without a lag parameter). We believe that the formulation of the model in this way provides, at least in certain applications, more theoretical support for the specification of the cyclical component, as opposed to a pure time series representation, while maintaining 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 in this paper is based on the spectral density function of the discrete time process which is consistent with the underlying continuous time model. Furthermore, the techniques are also able to distinguish between the point-sampling of a stock variable and the discrete time averaging inherent in the sampling of a flow variable.

The paper is organised as follows. Section 2 provides some relevant background on differential-difference equations, as well as some specific results concerning the particular equation that we use for modelling the cyclical component. These results concern the required conditions on the parameters that ensure stationarity as well as conditions that yield

[1] Chambers (1998) provides some examples.

[2] See, for example, Boucekkine, Licandro and Paul (1997) and Asea and Zak (1999).

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. In the case of the trend and seasonal being generated by differential equations it is possible to derive exact autoregressive representations for the discrete time components which form a basis for the derivation of the spectral density functions 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 due to exact discrete time representations corresponding to an underlying differential-difference equation being unavailable. Hence our focus is on spectral densities and frequency domain methods of estimation. Section 3 also gives 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 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 and hence we investigate the asymptotic properties of the feasible frequency domain Gaussian estimator (based on the truncated series) concentrating specifically on the role of the truncation parameter. Section 5 provides some simulation results for a model involving trend, cycle and irregular components, while section 6 contains the results of an empirical illustration of the methods using data on the US gross national product. Section 7 concludes, and all proofs of propositions and theorems are provided in the appendix.

2. DIFFERENTIAL-DIFFERENCE EQUATIONS AND CYCLES

Recent work in econometrics concerning differential-difference equations has concentrated mainly on developing appropriate methods of estimation of the unknown parameters in complete systems of equations. Chambers (1998) deals with a random $n \times 1$ vector $y(t)$ satisfying the differential-difference equation system given by

$$dy(t) = \sum_{k=0}^{p} A_k(\theta)y(t-k)dt + \sum_{k=1}^{q} B_k(\theta)dy(t-k) + \zeta(dt), \quad -\infty < t < \infty, \tag{1}$$

where p and q are positive finite integers, A_k $(k = 0, 1, \ldots, p)$ and B_k $(k = 1, \ldots, q)$ are $n \times n$ matrices whose elements are known functions of an unknown $m \times 1$ parameter vector θ (where $m \leq n^2(p+q+1)$), and $\zeta(dt)$ is an $n \times 1$ vector of random measures satisfying $E[\zeta(dt)] = 0$, $E[\zeta(dt)\zeta(dt)'] = \Omega dt$ and $E[\zeta(\Delta_1)\zeta(\Delta_2)'] = 0$ for $\Delta_1 \cap \Delta_2 = \emptyset$. Chambers (1998) derives a spectral representation for the unique solution to (1) and shows that it will be stationary provided that the roots to the equation $\det[\alpha(z)] = 0$ have negative real parts, where[3]

$$\alpha(z) = z\left[I_n - \sum_{k=1}^{q} B_k(\theta)e^{-kz}\right] - \sum_{k=0}^{p} A_k(\theta)e^{-kz}.$$

In the case where $y(t)$ is comprised of a mixture of stock and flow variables, Chambers (1998) derives the spectral density function of the discretely sampled series and utilises it in a frequency domain Gaussian likelihood estimation procedure. The methods are further extended by McGarry (1998) who allows for the lags to assume non-integer values by considering the system

$$dy(t) = \sum_{k=0}^{r} A_k(\theta)y(t-\nu_k)dt + \sum_{k=1}^{s} B_k(\theta)dy(t-\mu_k) + \zeta(dt), \quad -\infty < t < \infty, \tag{2}$$

where ν_k $(k = 0, 1, \ldots, r)$ and μ_k $(k = 1, \ldots, s)$ denote the (unknown) non-integer lags[4] and r and s are finite integers. Allowing for non-integer lags causes additional computational difficulties which are addressed in McGarry (1998) and whose results are drawn upon later in this paper.

The emphasis on frequency domain methods in the articles cited above arises because, 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,$p-1$) 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 frequency domain approach is therefore also followed in this paper.

The use of differential-difference equations in this paper is for the purposes of modelling 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 modelling 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 prominant early example being Kalecki (1935). We shall focus on the equation

$$dy(t) = [a_0 y(t) + a_1 y(t-p)]dt + \epsilon(dt), \quad t > 0, \tag{3}$$

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

[3] Note that (1) may be written $\alpha(D)y(t)dt = \zeta(dt)$ where D denotes the (mean square) differential operator.
[4] The lag parameter ν_0 would typically be set equal to zero.

the unknown lag parameter which is real and positive. An equation of this type was proposed by Kalecki (1935) for the study of business cycles and its properties were further investigated by James and Belz (1936) and Hayes (1950). 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.[5] However, McGarry (1998) provides the following result for the equation defined in (3).

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* $0 < x_1 < \pi/p$*.*

Proposition 1 is important because it provides conditions that are easily verified in practice for determining whether a process satisfying (3) is stationary or not. The proof of Proposition 1 is given in McGarry (1998).

With regard to cyclical behaviour, 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 (3) in order for the model to produce cycles.

PROPOSITION 2. *The stochastic differential-difference equation (3) produces cycles provided that* $a_1 < -p^{-1}e^{a_0 p - 1}$*. 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*

$$r \cot r + \ln \left| \frac{\sin r}{r} \right| = a_0 p - \ln|-a_1 p|. \tag{4}$$

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

[5] For certain (more complicated) differential-difference equation systems, such as those depicted in (1) and (2), verifying the required stationarity conditions could prove to be intractable.

of cycles is that $a_1 < -a_0$, which arises since the function $e^{x-1} > x$ and hence $a_1 p < -e^{a_0 p-1} < -a_0 p$. Note that this condition, in the form $a_0 < -a_1$, constitutes 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 above 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 does not exclude this possibility). The approach is therefore closer to James and Belz (1936) in this respect.

The spectral density function corresponding to (3) is given by, for the continuous time process $y(t)$, the function

$$f_y^c(\lambda) = \frac{\sigma_\epsilon^2}{2\pi} |a(i\lambda)|^{-2}, \quad -\infty < \lambda < \infty, \tag{5}$$

where λ denotes frequency. From the definition of $a(z)$ the spectral density in (5) may be written more explicitly as

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

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 $\{y_t = y(t);\ t = 0, \pm 1, \pm 2, \ldots\}$. In this case, the spectral density function of y_t is

$$f_y(\lambda) = \sum_{j=-\infty}^{\infty} f_y^c(\lambda + 2\pi j) = \frac{\sigma_\epsilon^2}{2\pi} \sum_{j=-\infty}^{\infty} |a(i(\lambda + 2\pi j))|^{-2}, \quad -\pi < \lambda \leq \pi, \tag{6}$$

which is the familiar 'folding equation' which 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 $\{Y_t = \int_{t-1}^{t} y(r)dr = \int_0^1 y(t-r)dr;\ t = 0, \pm 1, \pm 2, \ldots\}$. 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

$$\begin{aligned} f_Y(\lambda) &= \sum_{j=-\infty}^{\infty} |g(i(\lambda + 2\pi j))|^2 f_y^c(\lambda + 2\pi j) \\ &= \frac{\sigma_\epsilon^2}{2\pi} \sum_{j=-\infty}^{\infty} \frac{4 \sin^2(\lambda/2)}{(\lambda + 2\pi j)^2} |a(i(\lambda + 2\pi j))|^{-2}, \quad -\pi < \lambda \leq \pi, \end{aligned} \tag{7}$$

since $|g(i(\lambda+2\pi j))|^2 = 4\sin^2((\lambda+2\pi j)/2)/(\lambda+2\pi j)^2 = 4\sin^2(\lambda/2)/(\lambda+2\pi j)^2$. 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 non-integer, but McGarry (1998) 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 above results suggest that differential-difference equations in the form of (3) 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 which has been suggested by economic theory for certain types of variables. Equation (3) 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), while the cycle is governed by a differential-difference equation. The precise definitions of each of these components are given below.[7]

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

$$d\begin{bmatrix}\mu(t)\\ \beta(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}\mu(t)\\ \beta(t)\end{bmatrix}dt + \begin{bmatrix}\eta(dt)\\ \zeta(dt)\end{bmatrix}, \quad t>0, \tag{8}$$

where $\eta(dt)$ and $\zeta(dt)$ are mutually and serially uncorrelated random measures with zero

[6] For the model defined in (3), for a stock variable there are two terms in the resulting summation, while for a flow there are three terms. Equation (22) of Chambers (1998) provides the spectral density function for a stock variable generated by (3) 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.

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 $\Delta_1 \cap \Delta_2 = \emptyset$ 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.

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

$$d\begin{bmatrix} \gamma_j(t) \\ \gamma_j^*(t) \end{bmatrix} = \begin{bmatrix} 0 & \lambda_j \\ -\lambda_j & 0 \end{bmatrix}\begin{bmatrix} \gamma_j(t) \\ \gamma_j^*(t) \end{bmatrix} dt + \begin{bmatrix} \omega_j(dt) \\ \omega_j^*(dt) \end{bmatrix}, \quad j = 1, \ldots, s/2, \quad t > 0, \tag{9}$$

where $\lambda_j = 2\pi j/s$ ($j = 1, \ldots, s/2$) 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 (3) and is given by

$$d\psi(t) = [a_0\psi(t) + a_1\psi(t-p)]dt + \epsilon(dt), \quad t > 0, \tag{10}$$

where $\epsilon(dt)$ is a zero mean random measure with variance $\sigma_\epsilon^2 dt$, and 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

$$y_t = \mu_t + \gamma_t + \psi_t + \xi_t, \quad t = 1, \ldots, T, \tag{11}$$

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 realisation of a genuine irregular component in continuous time that affects $y(t)$, or as a

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

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 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

$$x_t = e^A x_{t-1} + \nu_t, \quad t = 1, \dots, T, \tag{12}$$

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 $\text{var}(\nu_t) = \int_0^1 e^{As}\Sigma_\nu e^{A's}ds$, 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, and are given in Harvey and Stock (1993).

Trend component. The discrete time trend component μ_t satisfies

$$\begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \end{bmatrix} + \begin{bmatrix} \eta_t \\ \zeta_t \end{bmatrix}, \quad t = 1, \dots, T, \tag{13}$$

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

$$\text{var}\begin{bmatrix} \eta_t \\ \zeta_t \end{bmatrix} = \begin{bmatrix} \sigma_\eta^2 + \sigma_\zeta^2/3 & \sigma_\zeta^2/2 \\ \sigma_\zeta^2/2 & \sigma_\zeta^2 \end{bmatrix}. \tag{14}$$

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

$$\begin{bmatrix} \gamma_{jt} \\ \gamma_{jt}^* \end{bmatrix} = \begin{bmatrix} \cos\lambda_j & \sin\lambda_j \\ -\sin\lambda_j & \cos\lambda_j \end{bmatrix}\begin{bmatrix} \gamma_{j,t-1} \\ \gamma_{j,t-1}^* \end{bmatrix} + \begin{bmatrix} \omega_{jt} \\ \omega_{jt}^* \end{bmatrix}, \quad j = 1, \dots, s/2, \quad t = 1, \dots, T, \tag{15}$$

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

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

Although the discrete time representations for trend and seasonal given above 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 above representations to derive the spectral density functions of these components, which are detailed in section 4.

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 $\left\{Y_t = \int_{t-1}^t y(r)dr;\ t = 1, \dots, T\right\}$. The measurement equation for flow variables is

$$Y_t = A_t + \Gamma_t + \Psi_t + \Xi_t, \quad t = 1, \dots, T, \tag{16}$$

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^2_Ξ, 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, and the approach followed here is to integrate equation (12) 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 was 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. While 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 which provides the form of the discrete time model satisfied by the integrals $X_t = \int_{t-1}^t x(r)dr$ and which holds for each of the unobserved trend and seasonal components.

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

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

$$dx(t) = Ax(t)dt + \nu(dt), \quad t > 0, \tag{17}$$

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 $\Lambda(r) = \int_0^r \Phi(s)\Sigma_\nu \Phi(s)' ds$.

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 below.

THEOREM 2. *Let the unobserved components for trend and seasonal evolve in continuous time according to (8) and (9) 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*

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

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

$$V_{0j} = \frac{4\sigma_\omega^2}{\lambda_j^3} \begin{bmatrix} (\lambda_j - \sin\lambda_j) & 0 \\ 0 & (\lambda_j - \sin\lambda_j) \end{bmatrix}, \quad j = 1, \ldots, s/2,$$

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

The autoregressive matrices in the discrete time ARMA(1,1) representations for the unobserved components in the flow case are identical to those in 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) proposed 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 (9) and is given by

$$d\begin{bmatrix} \psi(t) \\ \psi^*(t) \end{bmatrix} = \begin{bmatrix} \ln\rho & \lambda_c \\ -\lambda_c & \ln\rho \end{bmatrix} \begin{bmatrix} \psi(t) \\ \psi^*(t) \end{bmatrix} dt + \begin{bmatrix} \kappa(dt) \\ \kappa^*(dt) \end{bmatrix}, \quad t > 0, \tag{18}$$

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 (18) is more parsimonious than the differential-difference equation (10), 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 (10) in some circumstances (e.g. time-to-build investment models), but can provide an alternative to (10) 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 (12) for stock variables and from Theorem 1 for flow variables. For stock variables the appropriate equations were provided by Harvey and Stock (1993) and are given by

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}, \quad t = 1, \ldots, T, \tag{19}$$

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 (18). Then the discrete time law of motion in the case of a flow variable Ψ_t is given by the ARMA(1,1) process*

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

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

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

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

where

$$c_1(\lambda) = \frac{-2\rho}{\delta(\lambda)^2}\left[\ln\rho\cos\lambda + \lambda\sin\lambda\right] + \frac{1}{\delta(\lambda)}\left[1 + \frac{\rho^2-1}{2\ln\rho} + \frac{2\ln\rho}{\delta(\lambda)}\right],$$

$$c_2(\lambda) = \frac{1}{\delta(\lambda)}\left[1+\rho^2-2\rho\cos\lambda\right], \quad c_3(\lambda) = \frac{\rho^2\ln\rho}{\delta(\lambda)^2}, \quad c_4(\lambda) = \frac{2\rho^2\lambda(\ln\rho)^2}{\delta(\lambda)^3},$$

$$c_5(\lambda) = \frac{\rho^2\lambda\left[(\ln\rho)^2-\lambda^2\right]}{\delta(\lambda)^3}, \quad \delta(\lambda) = (\ln\rho)^2+\lambda^2, \quad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

$$\begin{aligned} H(\lambda) &= \frac{\rho\ln\rho}{\delta(\lambda)}C_1(\lambda)C_4(\lambda)' + \frac{\rho\lambda}{\delta(\lambda)}C_2(\lambda)C_4(\lambda)' \\ &\quad -C_3(\lambda)\left[C_4(\lambda)' + \frac{\rho\left[(\ln\rho)^2-\lambda^2\right]}{\delta(\lambda)^2}C_1(\lambda)' + \frac{2\rho\lambda\ln\rho}{\delta(\lambda)^2}C_2(\lambda)'\right], \end{aligned}$$

$$C_1(\lambda) = \begin{bmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{bmatrix}, \quad C_2(\lambda) = \begin{bmatrix} \sin\lambda & -\cos\lambda \\ \cos\lambda & \sin\lambda \end{bmatrix}, \quad C_3(\lambda) = \frac{1}{\delta(\lambda)} \begin{bmatrix} \ln\rho & -\lambda \\ \lambda & \ln\rho \end{bmatrix},$$

$$C_4(\lambda) = \frac{1}{\delta(\lambda)^2} \begin{bmatrix} \lambda^2 - (\ln\rho)^2 - \delta(\lambda)\ln\rho & 2\lambda\ln\rho + \lambda\delta(\lambda) \\ -2\lambda\ln\rho - \lambda\delta(\lambda) & \lambda^2 - (\ln\rho)^2 - \delta(\lambda)\ln\rho \end{bmatrix}.$$

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 which utilises the spectral density function of the differential-difference equation component, 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 which 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 secondly, 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. In order 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.[11] 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

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

[11] 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.

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

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

$$f_z(\lambda) = \left|\phi(e^{-i\lambda})\right|^2 [f_\mu(\lambda) + f_\gamma(\lambda) + f_\psi(\lambda) + f_\xi(\lambda)], \quad -\pi < \lambda \leq \pi. \tag{20}$$

Precise expressions for each of the constituents of (20) are given below.

Trend component.

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

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.

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

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

Cyclical component.

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

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

Irregular component.

$$f_\xi(\lambda) = \frac{\sigma_\xi^2}{2\pi}.$$

The functions $f_\mu(\lambda)$ and $f_\gamma(\lambda)$ are perhaps best described as being pseudo-spectral 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 functions above are derived directly from the discrete time representations in section 3.2, while $f_\psi(\lambda)$ corresponds directly to (6) 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 the previous subsection. The spectral density function in this case, $f_Z(\lambda)$, is given by

$$f_Z(\lambda) = \left|\phi(e^{-i\lambda})\right|^2 \left[f_A(\lambda) + f_\Gamma(\lambda) + f_\Psi(\lambda) + f_\Xi(\lambda)\right], \quad -\pi < \lambda \leq \pi. \tag{21}$$

Precise expressions for each of these components are given below.

Trend component.

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

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.

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

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.

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

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

Irregular component.

$$f_\Xi(\lambda) = \frac{\sigma_\Xi^2}{2\pi}.$$

As in the stock variable case, the functions $f_A(\lambda)$ and $f_\Gamma(\lambda)$ represent pseudo-spectra. All of the functions for the processes satisfying differential equations are based on the discrete time representations given in Theorem 2, while $f_\Psi(\lambda)$ corresponds to (7) in section 2. The term involving $\sin^2(\lambda/2)$ in $f_\Psi(\lambda)$ arises due to 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 (18). When the variable is a stock, the relevant expression is

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

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

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

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 (18), the above spectral densities can nevertheless be used in the frequency domain approach detailed below.

4.4. Maximum Likelihood Estimation

Under the assumption that the variable $y(t)$ is a Gaussian process, maximisation of the Gaussian likelihood function yields maximum likelihood estimators of the unknown parameters. For notational convenience, we shall not distinguish in this subsection 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:

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

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) $f^c(\lambda)$ as follows:

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

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 observed process. If the truncation point is fixed i.e. 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, non-truncated 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 Whittle likelihood) is given by (apart from a constant term)

$$L(\theta)=\frac{1}{T}\sum_{j\in J_T}\left[\ln f(\lambda_j;\theta)+\frac{I(\lambda_j)}{f(\lambda_j;\theta)}\right], \tag{23}$$

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

$$f^M(\lambda;\theta)=\left|\phi(e^{-i\lambda})\right|^2\sum_{k=-M}^{M} f^c(\lambda+2\pi k;\theta),\quad -\pi<\lambda\leq\pi, \tag{24}$$

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

$$r^M(\lambda;\theta)=\left|\phi(e^{-i\lambda})\right|^2\sum_{k=M+1}^{\infty}\left[f^c(\lambda+2\pi k;\theta)+f^c(\lambda-2\pi k;\theta)\right] \tag{25}$$

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

$$L^M(\theta)=\frac{1}{T}\sum_{j\in J_T}\left[\ln f^M(\lambda_j;\theta)+\frac{I(\lambda_j)}{f^M(\lambda_j;\theta)}\right]. \tag{26}$$

Our aim is to examine the properties of (23) and (26) and to compare

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

where Θ denotes the parameter space. In order 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 below.

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 $\sum_{s=0}^{\infty} |cov(z_t, z_{t+s})| < \infty$.

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 is, in some senses, redundant given 2 and 3, for it can be used to verify these conditions. However, we use it principally to stress that z_t is assumed to be a stationary Gaussian process.[12] 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 below; 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. Finally, Assumption 5 specifies an appropriate rate of increase of M with T.

THEOREM 3. *Under Assumptions 1 to 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 to 3.

In order 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 $T \to \infty$.

[12] Note that, for the model specified for z_t, it is implicit that $f(\lambda;\theta)$ is a continuous function of λ and θ and that $0 < f(\lambda;\theta) < \infty$.

A mean value expansion of the score vector $s(\hat{\theta}_T) = \partial L(\theta)/\partial\theta|_{\theta=\hat{\theta}_T}$ 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 $\|\bar{\theta} - \theta_0\| \leq \|\hat{\theta}_T - \theta_0\|$. This yields

$$\sqrt{T}(\hat{\theta}_T - \theta_0) = -S(\bar{\theta})^{-1}\sqrt{T}s(\theta_0). \tag{27}$$

We make the following high level assumptions based on the above expression.

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 $T \to \infty$.

Under Assumptions 6 and 7, and the previously established consistency of $\hat{\theta}_T$ (under Assumptions 1 to 3), we obtain the required limiting distribution for $\sqrt{T}(\hat{\theta}_T - \theta_0)$ with $V = S(\theta_0)^{-1}$. In fact,

$$S(\theta_0) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{\partial \ln f(\lambda;\theta_0)}{\partial\theta}\frac{\partial \ln f(\lambda;\theta_0)}{\partial\theta'}d\lambda, \tag{28}$$

where $\partial \ln f(\lambda;\theta_0)/\partial\theta$ denotes the vector $\partial \ln f(\lambda;\theta)/\partial\theta|_{\theta=\theta_0}$.

In order 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 obtain

$$\sqrt{T}(\hat{\theta}_T^M - \theta_0) = -S^M(\tilde{\theta})^{-1}\sqrt{T}s^M(\theta_0), \tag{29}$$

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 (27) with (29) yields

$$\begin{aligned}\sqrt{T}(\hat{\theta}_T - \hat{\theta}_T^M) &= S^M(\tilde{\theta})^{-1}\sqrt{T}s^M(\theta_0) - S(\bar{\theta})^{-1}\sqrt{T}s(\theta_0) \\ &= \left[S^M(\tilde{\theta})^{-1} - S(\bar{\theta})^{-1}\right]\sqrt{T}s^M(\theta_0) + S(\bar{\theta})^{-1}\sqrt{T}\left[s^M(\theta_0) - s(\theta_0)\right]. \end{aligned} \tag{30}$$

The analysis of this expression yields the following result.

THEOREM 4. *Under Assumptions 1 to 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 (28).*

The proof of Theorem 4 proceeds by demonstrating that $\sqrt{T}\left[s^M(\theta_0) - s(\theta_0)\right] = o_p(1)$ and that $S^M(\tilde{\theta})^{-1} - S(\bar{\theta})^{-1} = o_p(1)$. Combined with Assumptions 6 and 7, this yields the desired result. The elements of the matrix $S(\theta_0)$ are typically estimated using $\widehat{S}(\hat{\theta}_T^M)$ which is defined by the computationally more convenient expression

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

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 it is required that the truncation parameter $M = O(T^\delta)$ for some $\delta > 0$, but it remains a matter of some interest 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. First, and foremost, 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 versus flow); and the presence of a stochastic trend. In its most general form, the underlying continuous time model used in the simulations consists of a trend and cycle, defined as follows:

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

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 $2/3$ and first order autocovariance $1/6$. 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

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

therefore seven stationary combinations of the cyclical parameters, which shall be referred to as Experiments I to VII and which are summarized in Table 1.

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

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 1 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 upwards if T^δ is not already an integer. Values of $\delta = \{0.25, 0.50\}$ were considered, for three sample sizes, $T = \{64, 128, 256\}$, which gives values of M equal to $\{3, 8\}$ ($T = 64$), $\{4, 12\}$ ($T = 128$), and $\{4, 16\}$ ($T = 256$). A total of 1000 replications of each experiment were conducted, resulting in $28 \times 2 \times 3 \times 1000 = 168{,}000$ optimizations of the likelihood function. Details concerning the generation of data are provided in the Appendix, while the results of the simulations are presented in Tables 2 to 5. Each Table contains the mean square errors (multiplied by 10^4) for the seven experiments corresponding, in order, to stock and flow variables with no trend, and stock and flow variables with a stochastic trend.

In attempting to summarise the results contained in Tables 2 to 5, 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 (MSEs) of the following:

Truncation parameter M: there are noticeable differences in the MSEs 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 MSEs 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 4 than when it was of length 10.

Table 2. Mean square errors: stock variable, no trend

Experiment	δ	T	a_0	a_1	p	σ_ϵ	σ_ξ
I	0.25	64	217	292	1419	887	1532
		128	31	77	452	277	578
		256	13	27	84	22	95
	0.50	64	205	275	1529	878	1537
		128	31	77	451	268	574
		256	13	27	89	26	91
II	0.25	64	762	957	1764	1090	1922
		128	161	246	513	342	582
		256	38	69	92	36	87
	0.50	64	891	1066	1716	1124	1926
		128	136	229	468	328	597
		256	40	70	103	37	88
III	0.25	64	2489	11091	2482	1879	1459
		128	603	1915	433	439	266
		256	95	204	11	63	42
	0.50	64	5020	15096	2663	1987	1521
		128	560	1738	404	410	265
		256	93	194	9	62	42
IV	0.25	64	5122	17182	3840	2153	1725
		128	892	4043	713	432	200
		256	104	368	64	41	34
	0.50	64	3320	15135	3510	2094	1694
		128	1176	4929	772	470	201
		256	102	360	66	42	35
V	0.25	64	38	53	462	710	1152
		128	13	22	197	309	423
		256	5	9	68	105	131
	0.50	64	44	58	466	706	1173
		128	13	22	197	306	429
		256	5	9	68	106	134
VI	0.25	64	5392	10948	1911	2136	1052
		128	685	2419	521	802	324
		256	188	280	23	177	75
	0.50	64	2241	6966	1886	2056	1049
		128	667	2317	536	776	321
		256	189	274	23	168	74
VII	0.25	64	6167	13095	2788	2396	1312
		128	499	2488	948	660	368
		256	120	340	194	109	60
	0.50	64	3001	10337	2348	2128	1310
		128	491	2437	940	631	373
		256	111	344	186	98	61

All entries have been multiplied by 10^4.

Table 3. Mean square errors: flow variable, no trend

Experiment	δ	T	a_0	a_1	p	σ_ϵ	σ_ξ
I	0.25	64	188	483	1241	1921	1300
		128	34	170	598	685	496
		256	12	32	84	55	100
	0.50	64	189	479	1246	1897	1306
		128	34	171	599	686	494
		256	12	32	84	55	100
II	0.25	64	10429	7146	1446	3819	1820
		128	103	227	666	649	782
		256	30	56	102	66	123
	0.50	64	8850	6166	1442	3491	1823
		128	104	228	695	649	780
		256	30	56	101	64	123
III	0.25	64	1160	3686	2337	2576	885
		128	465	926	832	856	214
		256	105	146	11	151	40
	0.50	64	1139	3726	2366	2565	861
		128	457	923	489	854	213
		256	105	145	11	151	40
IV	0.25	64	10002	7341	3336	11934	728
		128	475	1237	788	858	204
		256	70	112	8	103	32
	0.50	64	11498	7716	3305	13486	733
		128	478	1263	782	861	200
		256	70	112	8	103	32
V	0.25	64	33	50	465	889	623
		128	12	20	223	377	234
		256	5	9	81	125	84
	0.50	64	33	50	465	889	624
		128	12	20	223	377	234
		256	5	9	81	125	84
VI	0.25	64	1509	5776	2006	2759	779
		128	469	1667	403	1046	290
		256	223	307	16	319	75
	0.50	64	1518	5843	1998	2737	782
		128	459	1633	393	1044	297
		256	186	293	16	274	69
VII	0.25	64	1710	6802	2627	3265	862
		128	578	2531	931	1162	319
		256	133	429	129	192	56
	0.50	64	1723	6720	2631	3314	850
		128	582	2529	920	1165	314
		256	133	429	129	192	56

All entries have been multiplied by 10^4.

Table 4. Mean square errors: stock variable, stochastic trend

Experiment	δ	T	a_0	a_1	p	σ_ϵ	σ_ξ	σ_η
I	0.25	64	1287	1033	1922	1519	1413	1502
		128	317	311	727	482	478	797
		256	66	85	139	36	120	200
	0.50	64	991	760	1914	1474	1444	1507
		128	311	315	737	510	464	788
		256	65	85	141	37	121	204
II	0.25	64	1296	1370	1315	1470	1617	1016
		128	372	455	487	562	606	568
		256	79	127	134	69	132	197
	0.50	64	1128	1285	1325	1448	1594	1025
		128	378	460	521	564	601	568
		256	79	127	135	69	130	197
III	0.25	64	1310	3948	1099	1454	1220	934
		128	180	190	62	232	323	376
		256	5	3	1	19	98	137
	0.50	64	1329	3978	1081	1385	1196	922
		128	184	184	62	225	328	368
		256	4	3	1	19	99	137
IV	0.25	64	2340	9093	1370	1361	1411	867
		128	319	922	59	225	344	311
		256	1	2	1	10	90	125
	0.50	64	2541	9284	1310	1317	1387	889
		128	299	831	65	202	316	309
		256	1	2	1	10	89	125
V	0.25	64	313	485	782	1568	1602	2267
		128	58	43	396	763	787	1284
		256	17	17	117	198	255	387
	0.50	64	318	477	786	1543	1623	2246
		128	61	43	390	748	796	1277
		256	17	17	117	198	258	386
VI	0.25	64	2132	5896	996	3093	1102	1151
		128	599	1048	293	937	393	433
		256	21	10	3	38	120	140
	0.50	64	2201	5437	946	3115	1148	1169
		128	605	1029	287	923	400	438
		256	22	11	3	38	120	144
VII	0.25	64	1374	5556	1846	1905	1366	719
		128	431	1737	754	591	505	323
		256	28	71	47	41	130	135
	0.50	64	1400	5666	1833	1858	1352	747
		128	419	1734	741	586	507	325
		256	26	59	46	39	124	135

All entries have been multiplied by 10^4.

Table 5. Mean square errors: flow variable, stochastic trend

Experiment	δ	T	a_0	a_1	p	σ_ϵ	σ_ξ	σ_η
I	0.25	64	950	897	1765	2546	1116	975
		128	290	320	594	604	400	490
		256	66	73	101	41	102	176
	0.50	64	956	877	1767	2546	1110	975
		128	290	319	594	601	404	490
		256	64	72	100	38	101	176
II	0.25	64	766	845	1496	1930	1415	761
		128	351	467	435	789	646	386
		256	77	120	73	41	124	129
	0.50	64	759	848	1527	1988	1420	761
		128	360	471	435	798	641	386
		256	77	121	74	41	124	129
III	0.25	64	1326	3329	1519	2900	813	632
		128	273	258	188	727	241	186
		256	3	1	2	10	67	76
	0.50	64	1349	3395	1514	2892	810	636
		128	272	256	188	726	238	186
		256	3	1	2	10	67	76
IV	0.25	64	1291	3333	1675	2463	787	490
		128	286	351	240	570	207	177
		256	1	1	1	3	58	78
	0.50	64	1305	3368	1677	2516	805	490
		128	286	350	239	568	208	177
		256	1	1	1	3	58	78
V	0.25	64	314	226	942	2323	775	1526
		128	87	55	455	998	288	882
		256	18	16	132	243	113	361
	0.50	64	315	226	928	2288	764	1526
		128	87	55	454	996	288	882
		256	18	16	132	243	113	361
VI	0.25	64	1929	6135	963	3309	636	728
		128	656	953	167	1521	248	279
		256	24	6	2	54	65	91
	0.50	64	1885	6094	972	3335	639	729
		128	663	955	165	1529	248	279
		256	24	6	2	54	65	91
VII	0.25	64	1421	5345	1769	3251	984	561
		128	573	2002	604	1257	367	214
		256	33	37	32	70	74	82
	0.50	64	1391	5307	1834	3196	992	560
		128	566	1979	603	1259	372	214
		256	33	37	32	70	74	82

All entries have been multiplied by 10^4.

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.

The above comments are derived from a visual inspection of the entries in the tables 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 summarise 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.[14] We have therefore made an attempt at carrying out a test of the hypothesis that the means of the distribution of mean square errors 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.[15] The relevant statistics are presented in Table 6.

Table 6. 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	−1.4494	−1.7852	−1.6062	0.4361	0.2401	*na*
Flow	−1.7323	−1.3885	−1.3289	−0.9061	−0.2509	*na*
Stock versus flow						
No trend	0.6061	−1.3947	−0.1204	2.1208*	−1.3544	*na*
Trend	−0.5892	−0.7453	0.2448	2.0280*	−2.0767*	−2.1490*

* denotes significance at the 5% level; *na* denotes test not applicable.

The first set of results tests the hypothesis that, for a given sampling scheme (stock or flow), the means of the MSEs 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. None 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 MSEs. The second

[14] 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 due to 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 optimisations of the likelihood function involving, in each case, at least five unknown parameters.

[15] The test relies on normality in finite samples and on the equivalance 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.

set of statistics tests the hypothesis that, for a given trend scenario, the means of the MSEs are the same regardless of whether the variable is a stock or a flow. A negative statistic occurs if the mean with a flow is smaller than the mean with a stock. The significant entries here occur, interestingly, for the variance parameters, particularly in the situation in which there is a stochastic trend. This is somewhat surprising, because the form of the spectral densities, and hence the likelihood itself, should be compensating for these variations. 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 therefore conclude that there do not exist significant differences in the means of the distributions of the MSEs caused either by the sampling scheme or by the presence of a stochastic trend.

6. AN EMPIRICAL ILLUSTRATION

In order to illustrate the proposed differential-difference equation methods in action with real data, we have chosen a time series for which there are already published estimates available from a discrete time unobserved components model. The variable itself is the logarithm of US gross national product (GNP) covering the period 1910 to 1970, a total of 61 annual observations. Harvey (1989, p.92) reports the results of fitting a discrete time trend-plus-cycle model to these data. We also fit a continuous time version of this model to the data as well as the continuous time model containing a differential-difference equation cyclical component. The results are contained in Tables 7 and 8.

Table 7 contains the results of estimating the discrete time and continuous time trend-plus-cycle models. The discrete time estimates are taken directly from Harvey (1989, p.92) who does not report standard errors for these particular estimates. The main difference between the estimated discrete time and continuous time models is that, in the latter, the variance of $\zeta(dt)$ is zero, and the estimated values of ρ and λ_c are smaller than in the discrete time model. The estimated length of the cycle is, as a result, correspondingly larger in the continuous time model, at 18.3 years compared to 7 years.[16] The estimates in Table 8, for the differential-difference equation cycle, are based on three choices of truncation parameter for the spectrum. Using the same method as in the simulations in the previous section, the truncation parameter M is determined by $M = [T^\delta] + 1(T^\delta \notin \mathcal{N})$ for $\delta = \{0.25, 0.50, 0.75\}$ (the corresponding values of M are 3, 8 and 22 respectively). The best models were those for which $\sigma^2_\Xi = \sigma^2_\zeta = 0$ and the estimates display little variation across the three columns. Each of the sets of estimates of the parameters that characterise the cyclical component satisfy the stationarity condition in Proposition 1, and the implied cycle lengths are all approximately equal to 7.3 years.[17] This is very close to the estimated cycle length obtained by Harvey (1989) for the trend-plus-cycle model formulated directly in discrete time. A disappointing feature, however, of the estimated values in Tables 7 and 8 is the relatively large size of the standard errors compared to the estimates themselves. In fact, none of the estimates has a t-ratio greater than 1.76. One possible explanation for this is that the models are

[16] The cycle length is given by $2\pi/\lambda_c$.
[17] The cycle length is equal to $2\pi p/r_1$; see Proposition 2.

Table 7. Estimates of Unobserved Components Models: Discrete Time and Continuous Time

Discrete time model

$$y_t = \mu_t + \psi_t + \Xi_t,$$

$$\begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ \beta_{t-1} \end{pmatrix} + \begin{pmatrix} \eta_t \\ \zeta_t \end{pmatrix},$$

$$\begin{pmatrix} \psi_t \\ \psi_t^* \end{pmatrix} = \rho \begin{pmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{pmatrix} \begin{pmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{pmatrix} + \begin{pmatrix} \kappa_t \\ \kappa_t^* \end{pmatrix},$$

Continuous time model

$$y_t = A_t + \Psi_t + \Xi_t,$$

$$A_t = \int_{t-1}^{t} \mu(r)dr, \quad \Psi_t = \int_{t-1}^{t} \psi(r)dr,$$

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

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

Variable

y_t: logarithm of US GNP, 1910–1970

Estimates

	Discrete time[1]	Continuous time
ρ	0.97	0.7116 (0.5730)
λ_c	0.90	0.3433 (1.3373)
σ^2_κ	3.3×10^{-4}	80.66×10^{-4} (144.28×10^{-4})
σ^2_η	23.7×10^{-4}	20.29×10^{-4} (76.08×10^{-4})
σ^2_ζ	6.1×10^{-4}	0.00
σ^2_Ξ	0.0	0.00
Cycle length	7.0	18.30

Figures in parentheses are standard errors.

[1] Estimates taken from Harvey (1989, p.92).

Table 8. Estimates of Continuous Time Unobserved Components Model with Differential-Difference Equation Cycle

Model

$$y_t = A_t + \Psi_t + \Xi_t,$$

$$A_t = \int_{t-1}^{t} \mu(r)dr, \quad \Psi_t = \int_{t-1}^{t} \psi(r)dr,$$

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

$$d\psi(t) = [a_0\psi(t) + a_1\psi(t-p)]dt + \epsilon(dt),$$

Variable

y_t: logarithm of US GNP, 1910–1970

Estimates

$\{\delta, M\}$	$\{0.25, 3\}$	$\{0.50, 8\}$	$\{0.75, 22\}$
a_0	0.2369	0.2370	0.2362
	(0.5788)	(0.4169)	(0.5304)
a_1	−0.8617	−0.8607	−0.8631
	(1.4165)	(1.0049)	(1.5041)
p	1.4723	1.4717	1.4733
	(1.0412)	(0.8368)	(1.1834)
σ_ϵ^2	0.59×10^{-4}	0.60×10^{-4}	0.58×10^{-4}
	(10.35×10^{-4})	(9.63×10^{-4})	(11.46×10^{-4})
σ_η^2	92.06×10^{-4}	92.04×10^{-4}	92.07×10^{-4}
	(70.74×10^{-4})	(71.98×10^{-4})	(71.75×10^{-4})
σ_ζ^2	0.00	0.00	0.00
σ_Ξ^2	0.00	0.00	0.00
r_1	1.2688	1.2679	1.2708
Cycle length	7.2908	7.2931	7.2843

Figures in parentheses are standard errors.

misspecified in some way, and Harvey (1989) does indeed find that the discrete time model in Table 7 is inferior to a cyclical trend model in which μ_t also depends on ψ_{t-1} and $y_t = \mu_t + \Xi_t$. Further investigations with continuous time cyclical trend models may be fruitful, but are beyond the scope of this simple illustration.

7. 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 which 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 in order 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 some simulation results are reported that investigate the role of the truncation parameter in finite samples. An empirical illustration using US GNP data is also provided.

There are a number of natural routes by which the results in this paper could be extended. Multivariate extensions would be relatively straightforward and the results of Chambers (1998) and McGarry (1998) could be utilised 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. As mentioned in the previous section, it may be of some worth to consider continuous time cyclical trend models, in which the cyclical component influences the trend directly, rather than appearing as a separate component in the measurement equation. This could be achieved for the case of a differential-difference equation cyclical component although there would be some increase in the complexity of the resulting spectral density function. Perhaps allied with this is the important task of undertaking more extensive empirical investigations with the proposed models, possibly in the estimation of business cycles with variables such as GNP and investment. Indeed, it was in this area that differential-difference equations were introduced into economics by Kalecki (1935). 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 generated by a differential-difference equation. Such issues of cycle length under misspecification can be investigated using appropriately designed sampling experiments. Some of these issues are currently under investigation by the authors.

REFERENCES

Asea, P. K. and Zak, P. J. (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 and M. D. Intriligator (eds.), *Handbook of Econometrics, Volume 2*, pp. 1145–1212. North-Holland, Amsterdam.

Boucekkine, R., Licandro, O. and Paul, C. (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*. Holt, Rinehart and Winston, New York.

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

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

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

Harvey, A. C. and Stock, J. H. (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. Blackwell, Oxford.

Hayes, N. D. (1950) Roots of the transcendental equation associated with a certain differential-difference equation. *Journal of the London Mathematical Society* 25, 226–232.

James, R. W. and Belz, M. H. (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.

McGarry, J. S. (1998) The estimation of systems of joint differential-difference equations with non-integer lags. University of Essex Discussion Paper No. 483.

Robinson, P. M. (1988) The stochastic difference between econometric statistics. *Econometrica* 56, 531–548.

Rozanov, Y. A. (1967) *Stationary Random Processes*. Holden Day, New York.

APPENDIX

Proof of Proposition 2. The result comes directly from Kalecki (1935) who deals with the deterministic model (using his notation for the parameters) given by

$$Dy(t) = \frac{m}{\theta}y(t) - \left(\frac{m}{\theta} + n\right)y(t-\theta), \tag{A1}$$

in which $D = d/dt$ denotes the usual (nostochastic) differential operator. Kalecki (1935, equation 29) shows that cycles are produced provided that the condition $m + \theta n > e^{m-1}$ is met. Since, in our notation, $p = \theta$, $a_0 = m/\theta$ and $a_1 = -[(m/\theta) + n]$, it follows that $m = a_0 p$ and $n = -(a_0 + a_1)$. Using these definitions the equivalent condition in our notation is that $a_0 p - p(a_0 + a_1) > e^{a_0 p - 1}$, which, upon rearrangement, gives the expression in the Proposition.

To derive the length of the cycle, we begin by noting that solutions to (A1) are of the form $Ke^{\alpha t}$, where K is some constant. Substituting into (A1) 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 above equation are, respectively,

$$q = \sigma e^q \cos r, \tag{A2}$$

$$r = \sigma e^q \sin r. \tag{A3}$$

From (A3) it follows that $\sigma e^q = r/\sin r$, and substituting this into (A2) yields

$$q = r \cot r.$$

Dividing (A3) 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

$$q = -\ln|\sigma| - \ln\left|\frac{\sin r}{r}\right|.$$

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. □

Proof of Theorem 1. The solution to (17) is given by

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

which yields the stochastic difference equation

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

see, for example, Bergstrom (1984, Theorem 3). Integrating the above equation over the interval $(t-1, t]$ yields

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

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

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

It follows that

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

which yields the expression given in the theorem, while

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

which yields the expression for V_1. □

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 (8), 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, since $\Lambda(1) = \int_0^1 \Phi(s)\Sigma_\nu\Phi(s)'ds$, 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_ζ^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 (9) and $\Sigma_\nu = \sigma_\omega^2 I_2$. It can be shown that

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

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

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

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

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

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 (18) 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, since $\Phi(s)\Phi(s)' = \{[1+\rho^{2s} - 2\rho^s\cos s\lambda_c]/\delta(\lambda_c)^2\}I_2$ 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 to 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. Since $f_j^M = f_j - r_j^M$, 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 form of which depend on the 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 below. 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 DDE and where $0 < |x_{1j}^M|, |x_{2j}^M| < 1$ for all j and large enough M. Since $\ln(1+x) = O(x)$ 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. Since $\arctan(x) = O(x)$ as $x \to 0$, 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

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

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 $|L(\theta) - L^M(\theta)| \leq \delta_1^M(\theta) + \delta_2^M(\theta)$, where

$$\delta_1^M(\theta) = \left| \frac{1}{T} \sum_{j \in J_T} \ln\left(1 + \frac{r_j^M}{f_j - r_j^M}\right)\right| = O(M^{-1}),$$

and

$$\delta_2^M(\theta) = \left| \frac{1}{T} \sum_{j \in J_T} I_j e_j^M \right|.$$

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

Proof of Theorem 4. First, note that under Assumption 6, $S(\bar{\theta})^{-1} = O_p(1)$, while under Assumption 7, $\sqrt{T}s(\theta_0) = O_p(1)$. Hence in order to show that $\sqrt{T}(\hat{\theta}_T - \hat{\theta}_T^M) = o_p(1)$, it is therefore sufficient to show that:
(i) $\sqrt{T}\left[s^M(\theta_0) - s(\theta_0)\right] = o_p(1)$, 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

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

$$-\frac{1}{\sqrt{T}}\sum_{j\in J_T} I_j e_j^M\left(\frac{\partial \ln f_{0j}}{\partial\theta}+\kappa_{1j}^M\right)$$

$$= \sqrt{T}s(\theta_0)+O_p(M^{-1}).$$

Therefore, $\sqrt{T}[s^M(\theta_0)-s(\theta_0)] = O_p(M^{-1}) = O_p(T^{-\delta})$ 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

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

where

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

and

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

$$+\frac{1}{T}\sum_{j\in J_T} I_j e_j^M \left(\frac{\partial \ln f_j}{\partial\theta}+\kappa_{1j}^M\right)\left(\frac{\partial \ln f_j}{\partial\theta'}+\kappa_{1j}^{M\prime}\right)$$

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

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 ensures 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 $\eta_t \sim NID(0,1)$, while 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 $\alpha = 2 \pm \sqrt{3}$. Taking the invertible root we have $\alpha = 2 - \sqrt{3}$ and this is the value used to generate H_t using the above MA representation. It is also possible to show that $\sigma_e^2 = \sigma_\eta^2/(6\alpha)$ and so we took $e_t \sim NID(0, 1/6\alpha)$ since $\sigma_\eta^2 = 1$. 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 discretisation 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 100'th observation is chosen, while 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.