Econometric Theory, **1**, 97–112. Printed in the United States of America.

THE ESTIMATION OF HIGHER-ORDER CONTINUOUS TIME AUTOREGRESSIVE MODELS

A. C. HARVEY and JAMES H. STOCK

London School of Economics and Harvard University

A method is presented for computing maximum likelihood, or Gaussian, estimators of the structural parameters in a continuous time system of higher-order stochastic differential equations. It is argued that it is computationally efficient in the standard case of exact observations made at equally spaced intervals. Furthermore it can be applied in situations where the observations are at unequally spaced intervals, some observations are missing and/or the endogenous variables are subject to measurement error. The method is based on a state space representation and the use of the Kalman–Bucy filter. It is shown how the Kalman–Bucy filter can be modified to deal with flows as well as stocks.

1. INTRODUCTION

A good deal of attention has been paid to the estimation of continuous time models in econometrics; see, for example, Wymer [14], Robinson [12] and the various papers in the book edited by Bergstrom [1]. In a more recent contribution, Bergstrom [2] presents a method for carrying out maximum likelihood (ML), or Gaussian, estimation of a closed higher-order system in the time domain. This article presents an alternative approach to ML estimation of such models which is more general and more attractive computationally.

The model contains N variables contained in a vector $y(t)$. These variables are assumed to be generated by a pth-order stochastic differential equation of the form

$$D^p y(t) = A_1 D^{p-1} y(t) + \cdots + A_{p-1} D y(t) + A_p y(t) + \zeta(t), \qquad (1)$$

This paper was written when the authors were, respectively, visiting professor and doctoral candidate in the Department of Economics, University of California, Berkeley. We would like to thank Richard Jones and Rex Bergstrom for their comments on a preliminary draft, though we remain responsible for errors.

where $A_1, \ldots, A_p$ are $N \times N$ matrices, D is the mean square differential operator, and $\zeta(t)$ is a multivariate continuous white noise process with mean zero and covariance matrix Σ, i.e., if we let

$$\zeta^*(r, s) = \int_r^s \zeta(t)\, dt$$

then

$$E[\zeta^*(r, s)] = 0$$
$$E[\zeta^*(t_1, t_2)\zeta^*(t_1, t_2)'] = (t_2 - t_1)\Sigma,$$

and

$$E[\zeta^*(t_1, t_2)\zeta^*(t_3, t_4)'] = 0 \qquad \text{for } t_1 < t_2 < t_3 < t_4.$$

The elements in the matrices $A_1, \ldots, A_n$ and Σ all depend on an $n \times 1$ vector of unknown parameters ψ.

Further discussion concerning the precise interpretation of (1) can be found in Bergstrom [1, pp. 188–120] and Wymer [14].

Suppose that the variables in $y(t)$ are observed at T equally spaced points in time. If the process is stationary, it can be shown that these discrete observations follow a multivariate ARMA $(p, p - 1)$ process if the variables are all stocks and a multivariate ARMA (p, p) process if some are flows. However, in both cases, the ARMA parameters are complicated functions of the original parameters in $A_1, \ldots, A_p$ and this makes it difficult to impose the a priori restrictions which economic theory places on these matrices. Bergstrom [2] suggests that a better way to parameterize the discrete model is in terms of the AR coefficient matrices and the autocovariance matrices of the MA part of the model. He then proposes computing an estimator of ψ by minimizing

$$\hat{L}(\psi) = \log|V| + y'V^{-1}y \tag{2}$$

where V is an $NT \times NT$ covariance matrix which depends on ψ, and y is the $NT \times 1$ vector of observations. If the integral of $\zeta(t)$ is a multivariate normal process, the resulting estimators are maximum likelihood. If this assumption is not made, the estimators are referred to as Gaussian and, in fact, it is under these weaker conditions that Bergstrom [2, Section 7] derives their properties.

The V matrix depends on ψ, and since $\hat{L}$ must be minimized numerically, the computational burden involved in constructing and inverting V a large number of times may be quite heavy; see Bergstrom [2, p. 134]. This difficulty can be avoided by approaching time domain estimation in a completely different way. The pth-order system in (1) can be reduced to a first-order system by defining

$$\alpha^*(t) = \begin{bmatrix} y(t) \\ Dy(t) \\ \vdots \\ D^{p-1}y(t) \end{bmatrix}, \qquad R = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ I \end{bmatrix}, \qquad \text{and } A = \left[\begin{array}{c|ccc} 0 & & & \\ 0 & & I & \\ \vdots & & & \\ \hline A_p & A_{p-1} & \cdots & A_1 \end{array}\right] \tag{3}$$

and writing

$$\frac{d}{dt}[\alpha^*(t)] = A\alpha^*(t) + R\zeta(t). \tag{4}$$

Estimating a genuine first-order model is, in principle, straightforward, but simply writing (1) in the form (3) does not provide an immediate solution to the problem at hand since, as Bergstrom [2, p. 120] points out, the derivatives in $\alpha^*(t)$ are not observable. However, if (4) is recognized as the transition equation in a continuous time state space model, as in Kalman and Bucy [9], a solution is available. For a stock variable, the state space formulation is completed by a measurement equation

$$y(t) = Z\alpha^*(t), \qquad t = 1, \ldots, T, \tag{5}$$

where Z is the $N \times Np$ matrix

$$Z = [I_N \quad 0 \quad \cdots \quad 0]. \tag{6}$$

Application of the Kalman–Bucy filter then offers the possibility of constructing the likelihood function via the prediction error decomposition. This approach has been adopted by Jones [8, 9]. However, Jones does not consider the problem of estimating unknown initial conditions in a nonstationary model, and his analysis does not cover flow variables. Both of these problems are of considerable importance in econometrics.

The attractions of a state space approach in the present context can be summarized as follows:

1. As already noted, the repeated construction and inversion of the $NT \times NT$ matrix, V is avoided.
2. The method can be extended very easily to handle situations where the data are measured at unequally spaced intervals.
3. It is not necessary to have observations on all the variables at any particular measurement point, i.e., there may be missing observations.
4. The model can be extended to allow for measurement error.
5. The estimation of nonstationary models and models with exogenous variables, (i.e., open systms) can be carried out by an appropriate treatment of the initial conditions.

In order to allow for the possibility of unequally spaced observations, the points at which observations are made will be indexed by τ. Thus the τth set of observations will be regarded as arising at time t_τ, $\tau = 1, \ldots, T$, while the time interval between the τth and $(\tau - 1)$th observations will be denoted by $\delta_\tau = t_\tau - t_{\tau-1}$. Missing observations can be handled by defining the **Z** matrix in the measurement equation (6) in such a way that the rows corresponding to missing observations are deleted. Note that with a flow variable, an alternative situation can arise in which the value of a missing observation is included in a future observation. This is known as temporal aggregation in a discrete model. While this problem can be handled in a continuous time model, to do so would complicate the exposition and so it will not be covered explicitly.

Measurement error has not usually been incorporated into continuous time models in econometrics. However, it can be allowed for by adding a disturbance term to the measurement equation. Whether it is worth introducing this additional source of error into the econometric model in practice is debatable, but it is certainly worth including it in the formulae presented below for the sake of generality.

2. ESTIMATION FROM A SAMPLE OF OBSERVATIONS AT DISCRETE POINTS IN TIME

Suppose that all the variables in the model are stocks, where a stock is used in the general sense to denote any variable which can be measured at a particular point in time. Examples include the capital stock, the rate of interest and the temperature in Rio de Janeiro. In order to allow for the possibility of measurement error, the vector of observations at time t_τ will be denoted by y_τ. This is related to the variables in the stochastic differential equation (1) by the measurement equation

$$y_\tau = Z_\tau \alpha^*(t_\tau) + \xi_\tau, \qquad t = 1, \ldots, T \tag{7}$$

where $\alpha^*(t)$ is as defined in (3) and the $N \times 1$ vector ξ_τ is a multivariate white noise disturbance term with mean zero and covariance matrix H_τ, which is distributed independently of $\zeta(t)$. If some of the observations are missing, the dimensions of y_τ, Z_τ, ξ_τ, and H_τ must be reduced accordingly. In fact the subscripts on Z and H have only been attached in order to make this possibility explicit. If there are no missing observations both these matrices are time invariant with the former defined by (6).

If A matrix in (3) has distinct characteristic roots it can be diagonalized, i.e.,

$$\mathrm{A} = \mathrm{G}\Lambda\mathrm{G}^{-1}, \tag{8}$$

where Λ is a diagonal matrix containing the roots, λ_i, $i = 1, \ldots, N_p$, of A, and G is an $Np \times Np$ matrix, the columns of which are the right characteristic vectors of A. Making the linear transformation

$$\alpha(t) = G^{-1}\alpha^*(t) \tag{9}$$

enables (4) to be written in the more manageable form

$$\frac{d}{dt}[\alpha(t)] = \Lambda\alpha(t) + \eta(t), \tag{10}$$

where $\eta(t)$ is a multivariate continuous white noise process, defined by $\eta(t) = G^{-1}R\zeta(t)$, which has mean zero and covariance matrix

$$Q = G^{-1}R\Sigma R'(\bar{G}^{-1})'. \tag{11}$$

Writing a "bar" over a matrix indicates the matrix of complex conjugates.

Suppose that at time $t_{\tau-1}$ the optimal estimator, $a_{\tau-1}$, of $\alpha(t_{\tau-1})$ is available, together with $P_{\tau-1}$, the covariance matrix of its estimation error. An optimal estimator in this context means a minimum mean square estimator (MMSE); see Duncan and Horn [4] or Harvey [7, Ch. 4]. It follows from (4) that

$$\alpha_\tau = \alpha(t_\tau) = e^{\Lambda\delta}\alpha_{\tau-1} + \int_0^\delta e^{\Lambda(\delta-s)}\eta(t_{\tau-1}+s)\,ds, \tag{12}$$

where $\delta = \delta_\tau$, but the subscript has been dropped for notational convenience. The optimal estimator of α_τ at time $t_{\tau-1}$ is, therefore,

$$a_{\tau/\tau-1} = e^{\Lambda\delta}a_{\tau-1} = T_\tau a_{\tau-1}, \tag{13a}$$

where T_τ corresponds to the transition matrix in the Kalman filter for a discrete time state space model. The covariance matrix of $\alpha_\tau - a_{\tau/\tau-1}$ is

$$P_{\tau/\tau-1} = T_\tau P_{\tau-1}\bar{T}'_\tau + Q_\tau, \tag{13b}$$

where the ijth element of Q_τ is

$$(Q_\tau)_{ij} = q_{ij}W(\lambda_i + \bar{\lambda}_j;\, \delta), \tag{13c}$$

the function $W(x; \delta)$ being defined as

$$W(x;\delta) = \begin{cases} \dfrac{\exp(\delta x) - 1}{x}, & x \neq 0 \\ \delta & x = 0 \end{cases} \tag{14}$$

and q_{ij} being the ijth element of the matrix Q in (11). Note that both T_τ and Q_τ depend on the interval between observations, δ_τ.

The MMSE of y_τ at time $t_{\tau-1}$ is

$$\tilde{y}_{\tau/\tau-1} = Z_\tau G a_{\tau/\tau-1}. \tag{15}$$

The corresponding vector of one-step ahead prediction errors, $v_\tau = y_\tau - \tilde{y}_{\tau/\tau-1}$, has mean zero and covariance matrix

$$F_\tau = Z_\tau G P_{\tau/\tau-1} \bar{G}' Z_\tau' + H_\tau. \tag{16}$$

Finally the transformed state vector is updated by

$$a_\tau = a_{\tau/\tau-1} + P_{\tau/\tau-1} \bar{G}' Z_\tau' F_\tau^{-1} v_\tau \tag{17a}$$

and

$$P_\tau = P_{\tau/\tau-1} - P_{\tau/\tau-1} \bar{G}' Z_\tau' F_\tau^{-1} Z_\tau G P_{\tau/\tau-1}. \tag{17b}$$

The equations in (13) and (17) are, respectively, the prediction and updating equations in the Kalman–Bucy filter for the continuous time state space model, (4) and (7). Note that $P_{\tau/\tau-1}$ and P_τ are Hermitian.

Evaluation of the Likelihood Function for Stationary Models

The process generated by (1) is stationary if the roots of A have negative real parts. In this case starting values for the Kalman–Bucy filter are given by the unconditional mean and covariance matrix of $\alpha^*(t)$. The unconditional mean is a vector of zeros and so the MMSE of α_1 before y_1 is observed as $a_{1/0} = 0$. Since

$$\alpha(t) = \int_{-\infty}^{t} e^{\Lambda(t-s)} \eta(s)\, ds$$

the ijth element of the covariance matrix of $\alpha_1 - a_{1/0}$ is

$$(P_{1/0})_{ij} = -q_{ij}/(\lambda_i + \bar{\lambda}_j) \tag{18}$$

The ML estimators are obtained by minimizing

$$\hat{L}(\psi) = \sum_{\tau=1}^{T} \log|F_\tau| + \sum_{\tau=1}^{T} v_\tau' F_\tau^{-1} v_\tau, \tag{19}$$

where the quantities v_τ and F_τ are obtained directly from the Kalman–Bucy filter. In the special case when the observations are all available, are equally spaced and are measured without error, (19) is identical to (2).

Evaluation of the Likelihood Function with Fixed Initial Conditions

If we do not wish to evaluate the initial conditions using (18) or if we are unable to use (18) because the model is nonstationary, approximate ML estimators can be obtained by starting the Kalman–Bucy filter with $a_{2/1} = G^{-1}a^*_{2/1}$, where $a^*_{2/1} = (y'_1\ 0', \ldots, 0')'$ and letting $P_{2/1} = 0$. This effectively assumes that the derivatives in $\alpha^*(t_1)$ are fixed and equal to zero. An alternative approach would be to treat these unknown derivatives as additional parameters to be estimated nonlinearly along with those in ψ. This approach can also be used when some of the observations in $y(t_1)$ are missing, and if all the unobserved elements in $y(t_1)$ are fixed it yields the exact ML estimator of ψ.

Neither of the preceding methods is entirely satisfactory. The first because it introduces an approximation into the likelihood function and the second because it increases the uncertainty associated with the numerical optimization procedure. However, the second procedure can be modified by observing that the unknown parameters in $\alpha^*(t_1)$ are linear in the observations, and hence can be concentrated out of the likelihood function. This makes it considerably more attractive.

An algorithm enabling unknown elements in an initial state vector to be concentrated out of the likelihood function was first given by Rosenberg [13]. This algorithm can be modified for use in the present problem as follows. Let there be $M \leqslant Np$ unknown elements in $\alpha^*(t_1)$ and let these be denoted by an $M \times 1$ vector θ. The Kalman–Bucy filter is initialized at time t_1 with the unknown elements in $\alpha^*(t_1)$ set equal to zero. Premultiplying by G^{-1} yields the appropriate initial state vector estimator, $\tilde{\mathrm{a}}_1$, while $\tilde{\mathrm{P}}_1 = 0$. As the Kalman-Bucy filter proceeds, a set of $Np \times M$ matrices J_τ, $\tau = 1, \ldots, T-1$, are produced by the following auxiliary recursions:

$$J_\tau = T_{\tau+1}(I - K_\tau Z_\tau G)J_{\tau-1}, \qquad \tau = 2, \ldots, T-1, \tag{20}$$

where T_τ is defined in (13a), K_τ is the Kalman gain, which in this case is given by $K_\tau = P_{\tau/\tau-1}\bar{G}Z'_\tau F_\tau^{-1}$, and J_1 consists of the M columns of T_2G^{-1} corresponding to unknown elements in $\alpha^*(t_1)$. (When all the elements in $y(t_1)$ are observed, this means the last $N(p-1)$ columns.) When all the observations have been processed, the ML estimators of θ, conditional on ψ, is

$$\tilde{\theta} = \left[\sum_{\tau=2}^{T} \bar{J}'_{\tau-1}Z'_\tau F_\tau^{-1}Z_\tau J_{\tau-1}\right]^{-1} \sum_{\tau=2}^{T} \bar{J}'_{\tau-1}Z'_\tau F_\tau^{-1}\tilde{v}_\tau, \tag{21}$$

where $\tilde{v}_2, \ldots, \tilde{v}_T$, are the vectors of one-step ahead prediction errors produced by the Kalman–Bucy filter. Therefore, the elements of θ can be concentrated out of the likelihood function, leaving the following function to be minimized with respect to ψ:

$$L = \sum_{\tau=2}^{T} \log|F_\tau| + \sum_{\tau=2}^{T} \tilde{v}'_\tau F_\tau^{-1}\tilde{v}_\tau - \tilde{\theta}' \sum_{\tau=2}^{T} \bar{J}'_{\tau-1}Z'_\tau F_\tau^{-1}\tilde{v}_\tau. \tag{22}$$

Further details on the derivation of this algorithm are found in Appendix A.

Miscellaneous Comments

1. When the observations are equally spaced and none are missing, δ_τ can be set equal to unity for all τ, and the Kalman–Bucy filter will converge to a steady state under fairly mild conditions; cf. Chan *et al.* [3]. The P_τ matrix is time invariant in the steady state, and so once close to a steady state, a considerable reduction in computing time can be achieved by recognizing this fact and dropping (13b) and (17b); a similar device was used in [5] for computing exact ML estimators in univariate ARMA models.
2. Once the parameters ψ, have been estimated, (13a) can be used to make prediction for any lead time. Mean square errors, conditional on ψ, can be computed from (13b).
3. Stationarity can be imposed on the model by a suitable parameterization; cf. Jones [8, pp. 656–7].
4. A nonzero mean can be introduced into the model by adding an $N \times 1$ vector β to the right-hand side of (1). The mean is then $\mu = -A_p^{-1}\beta$, and this can be estimated efficiently from the sample mean of the observations. The parameters in ψ are then estimated as before with the observations in deviation from the mean form. If the actual ML estimator of μ (or β) is required, it can be obtained by using an appropriate modification of Rosenberg's algorithm to concentrate the likelihood function.
5. If there is no measurement error on y_τ, all the formulae remain valid with $H_\tau = 0$.

3. ESTIMATION FROM A SAMPLE OF INTEGRAL OBSERVATIONS

Flow variables can only be measured with respect to a particular interval of time. Examples include national income, sales and rainfall. If all the variables in $y(t)$ are flows and are measured without error at time t_ι, the vector of observations is

$$y_\tau = \int_0^\delta y(t_{\tau-1} + s)\, ds. \tag{23}$$

More generally, the measurement equation appropriate to the transformed first-order system, (4), is

$$y_\tau = \int_0^\delta Z_\tau G\alpha(t_{\tau-1} + s)\, ds + \xi_\tau, \qquad \tau = 1, \ldots, T, \tag{24}$$

where ξ_T is a multivariate white noise disturbance term with mean zero and covariance matrix $\mathbf{H}_\tau$ and the subscripts on Z_τ and $\mathbf{H}_\tau$ indicate the possibility of missing observations. As before, δ can vary with τ, but the subscript has been omitted for notational convenience. Note that it is possible to conceive of a variation of (24) in which measurement error accumulates continuously, that is,

$$y_\tau = \int_0^\delta \{y(t_{\tau-1} + s) + \xi(t_{\tau-1} + s)\}\, ds.$$

This can be written in the form (24), but with ξ_τ having a covariance matrix $\delta_\tau H_\tau$. In what follows it will be assumed that the measurement error is as originally given in (24), but clearly the second model can be handled almost as easily. Of course if δ_τ is constant there is effectively no difference between the two formulations.

In view of (24) it is necessary to consider predictions of the state vector at all points between $t_{\tau-1}$ and t_τ. If $a(t + s/t)$ denotes the MMSE of $\alpha(t + s)$ at time t, then

$$a(t_{\tau-1} + s/t_{\tau-1}) = e^{\Lambda s} a_{\tau-1}. \tag{25}$$

(Note that in the special case when $s = \delta$, this can be written as $a_{\tau/\tau-1}$.) Given $a_{\tau-1}$, the MMSE of y_τ is

$$\begin{aligned} \tilde{y}_{\tau/\tau-1} &= \int_0^\delta Z_\tau G a(t_{\tau-1} + s/t_{\tau-1})\, ds \\ &= Z_\tau G \left[\int_0^\delta e^{\Lambda s}\, ds \right] a_{\tau-1} \\ &= Z_\tau G W_\tau a_{\tau-1}, \end{aligned} \tag{26}$$

where W_τ is a diagonal matrix with ith diagonal element $W(\lambda_i; \delta)$. The covariance matrix of the prediction error, $v_\tau = y_\tau - \tilde{y}_{\tau/\tau-1}$, is

$$F_\tau = Z_\tau G P^{ff}_{\tau/\tau-1} \bar{G}' Z'_\tau + H_\tau. \tag{27}$$

The difference between (27) and (16) is that $P_{\tau/\tau-1}$ is replaced by a new matrix $P^{ff}_{\tau/\tau-1}$ such that

$$P^{ff}_{\tau/\tau-1} = W_\tau P_{\tau-1} \bar{W}_\tau + Q^{ff}_\tau \tag{28}$$

where the (i, j) element of Q^{ff}_τ is given by

$$(Q^{ff}_\tau)_{ij} = \begin{cases} \dfrac{q_{ij}}{\lambda_i + \bar{\lambda}_j}\left[W(\lambda_i; \delta) W(\bar{\lambda}_j; \delta) - \dfrac{\{W(\lambda_i; \delta) - \delta\}}{\lambda_i} - \dfrac{\{W(\bar{\lambda}_j; \delta) - \delta\}}{\bar{\lambda}_j} \right], & \lambda_i + \bar{\lambda}_j \neq 0 \\ \dfrac{2q_{ij}}{\lambda_i^2}\left[\dfrac{\sin h(\lambda_i \delta)}{\lambda_i} - \delta \right], & \lambda_i + \bar{\lambda}_j = 0. \end{cases} \tag{29}$$

If, say, $\lambda_i = 0$, then $\{W(\lambda_i; \delta) - \delta\}/\lambda_i$ is replaced by $\delta^2/2$. The updating equations are

$$a_\tau = a_{\tau/\tau-1} + P^f_{\tau/\tau-1}\bar{G}'Z'_\tau F_\tau^{-1} v_\tau \tag{30a}$$

and

$$P_\tau = P_{\tau/\tau-1} - P^f_{\tau/\tau-1}\bar{G}'Z'_\tau F_\tau^{-1} Z_\tau G\overline{P^f_{\tau/\tau-1}}', \tag{30b}$$

where $a_{\tau/\tau-1}$ and $P_{\tau/\tau-1}$ are defined as in (13), but F_τ is defined by (27) and the matrix $P^f_{\tau/\tau-1}$ is

$$P^f_{\tau/\tau-1} = e^{\Lambda\delta}P_{\tau-1}\bar{W}_\tau + Q^f_\tau \tag{31a}$$

where

$$(Q^f_\tau)_{ij} = \begin{cases} \dfrac{q_{ij}}{\lambda_i + \bar{\lambda}_j}[W(\bar{\lambda}_j; \delta) - W(-\lambda_i; \delta)], & \lambda_i + \bar{\lambda}_j \neq 0 \\ \dfrac{q_{ij}(W(\lambda_i; \delta) - \delta)}{\lambda_i}, & \lambda_i + \bar{\lambda}_j = 0. \end{cases} \tag{31b}$$

The derivation of (27) and (29) is given in the next two subsections. These may be omitted without any loss in continuity.

The modified Kalman–Bucy filter thus consists of the prediction equations (13), together with the updating equations (30). When the model is stationary, starting values can be computed from (18), exactly as before. The likelihood function is as in (19), except that F_τ is defined as in (27).

Computing the likelihood function with fixed initial conditions is best carried out by treating all the elements in $\alpha^*(t_0)$ as fixed. Since no observations are available at time t_0, $\theta = \alpha^*(t_0)$ and $M = Np$. Rosenberg's algorithm operates essentially as before, but with F_τ defined as in (27), $P_{\tau/\tau-1}$ replaced by $P^f_{\tau/\tau-1}$, $Z_\tau G$ replaced by $Z_\tau G K_\tau$, and τ running from 1 to $T-1$. The starting value for (20) is the $Np \times Np$ matrix $J_0 = T_1 G^{-1}$.

Derivation of the Covariance Matrix of the Prediction Error Vector

Since

$$\alpha(t_{\tau-1} + s) = e^{\Lambda s}\alpha_{\tau-1} + \int_0^s e^{\Lambda(s-u)}\eta(t_{\tau-1} + u)\, du \tag{32}$$

it follows from (24) and (26) that the prediction error vector is

$$\begin{aligned} v_\tau &= y_\tau - \tilde{y}_{\tau/\tau-1} \\ &= Z_\tau G\left[W_\tau(\alpha_{\tau-1} - a_{\tau-1}) + \int_0^\delta \int_0^s e^{\Lambda(s-u)}\eta(t_{\tau-1} + u)\,du\,ds \right] + \xi_\tau. \end{aligned} \tag{33}$$

The matrix we need to evaluate is $P^{ff}_{\tau/\tau-1}$ in (27), which is the expectation of the matrix in square brackets in (33) multiplied by its complex conjugate transpose. It is apparent from (33) that $P^{ff}_{\tau/\tau-1}$ is made up of the sum of two matrices, the first of which is simply $W_\tau P_{\tau-1} \bar{W}_\tau$. When $\lambda_i + \bar{\lambda}_j \neq 0$, the ijth element of the second matrix Q^{ff}_τ, is

$$\begin{aligned} (Q^{ff}_t)_{ij} &= q_{ij} \int_0^\delta \int_0^\delta \int_0^{\min(s,r)} e^{\lambda_i s + \bar{\lambda}_j r} \cdot e^{-(\lambda_i + \bar{\lambda}_j)u}\,du\,dr\,ds \\ &= \frac{q_{ij}}{\lambda_i + \bar{\lambda}_j} \int_0^\delta \int_0^\delta e^{\lambda_i r + \bar{\lambda}_j s} - e^{\phi(s-r)}\,dr\,ds, \end{aligned} \tag{34}$$

where

$$\phi(s-r) = \begin{cases} \lambda_i(s-r), & s > r \\ -\bar{\lambda}_j(s-r), & s < r \end{cases}.$$

Now

$$\int_0^\delta \int_0^\delta e^{\lambda_i r + \lambda_j s}\,dr\,ds = W(\lambda_i; \delta) W(\bar{\lambda}_j; \delta) \tag{35}$$

while

$$\begin{aligned} \int_0^\delta \int_0^\delta e^{\phi(s-r)}\,dr\,ds &= \int_0^\delta \int_0^s e^{\lambda_i(s-r)}\,dr\,ds + \int_0^\delta \int_0^r e^{-\bar{\lambda}_j(s-r)}\,ds\,dr \\ &= \left[\frac{e^{\lambda_i \delta}}{\lambda_i^2} - \frac{\delta}{\lambda_i} - \frac{1}{\lambda_i^2} \right] + \left[\frac{e^{\bar{\lambda}_j \delta}}{\bar{\lambda}_j^2} - \frac{\delta}{\bar{\lambda}_j} - \frac{1}{\bar{\lambda}_j^2} \right]. \end{aligned} \tag{36}$$

Substituting (35) and (36) into (34) gives (29). In the case $\lambda_i + \bar{\lambda}_j = 0$, expression (34) becomes

$$\begin{aligned} &\int_0^\delta \int_0^\delta e^{\lambda_i s + \bar{\lambda}_j r} \cdot \min(r,s)\,dr\,ds \\ &= \left[\frac{\delta^2}{2\lambda_i} - \frac{\delta}{\lambda_i^2} + \frac{e^{\bar{\lambda}_j \delta} - 1}{\lambda_i^2 \bar{\lambda}_j} \right] + \left[\frac{\delta^2}{2\bar{\lambda}_j} - \frac{\delta}{\bar{\lambda}_j^2} + \frac{e^{\lambda_i \delta} - 1}{\bar{\lambda}_j^2 \lambda_i} \right]. \end{aligned} \tag{37}$$

which can be rewritten as in (29).

Derivation of the Updating Equations

Consider the state vector defined at a finite number of equally spaced points between $t_{\tau-1}$ and t_τ, i.e., $\alpha(t_{\tau-1} + \{i\delta/k\})$, $i = 1, \ldots, k$. The observation y_τ is

$$
\begin{aligned}
y_\tau &\simeq Z_\tau G\left[\sum_{i=1}^{k} \alpha(t_{\tau-1} + \{i\delta/k\}) \times \delta/k\right] + \xi_\tau \\
&= Z_\tau G S^{(k)} \alpha_\tau^{(k)} + \xi_\tau
\end{aligned}
\tag{38}
$$

where $\alpha_\tau^{(k)}$ is an $Npk \times 1$ vector containing the stacked state vectors at $t_{\tau-1} + i\delta/k$, $i = 1, \ldots, k$, and $S^{(k)}$ is an $Np \times Npk$ summation matrix

$$S^{(k)} = (I_{Np} I_{Np} \ldots . I_{Np}) \times \delta/k. \tag{39}$$

Let the MMSE of $\alpha(t_{\tau-1} + \{i\delta/k\})$ be $a(t_{\tau-1} + \{i\delta/k\}/t_{\tau-1})$, for $i = 1, \ldots, k$, and let $a_{\tau-1}^{(k)}$ be the vector corresponding to $\alpha_\tau^{(k)}$. The Kalman filter updating equations combine the information in $a_{\tau-1}^{(k)}$ with that in y_τ to produce the MMSE of $\alpha_\tau^{(k)}$. Following Duncan and Horn [4] or Harvey [7, p. 108–9], we may write

$$
\begin{bmatrix} a_{\tau-1}^{(k)} \\ y_\tau \end{bmatrix} = \begin{bmatrix} I \\ Z_\tau G S^{(k)} \end{bmatrix} \alpha_\tau^{(k)} + \begin{bmatrix} a_{\tau-1}^{(k)} - \alpha_\tau^{(k)} \\ \xi_\tau \end{bmatrix},
\tag{40}
$$

from which the updating equations are

$$a_\tau^{(k)} = a_{\tau-1}^{(k)} + P_{\tau/\tau-1}^{(k)} S^{(k)\prime} \bar{G}' Z_\tau' \{F_\tau^{(k)}\}^{-1} v_\tau \tag{41}$$

and

$$P_\tau^{(k)} = P_{\tau/\tau-1}^{(k)} - P_{\tau/\tau-1}^{(k)} S^{(k)\prime} \bar{G}' Z_\tau' \{F_\tau^{(k)}\}^{-1} Z_\tau G S^{(k)} P_{\tau/\tau-1}^{(k)}, \tag{42}$$

where

$$F_\tau^{(k)} = Z_\tau G S^{(k)} P_{\tau/\tau-1}^{(k)} S^{(k)\prime} \bar{G}' Z_\tau' + H_\tau \tag{43}$$

and

$$P_{\tau/\tau-1}^{(k)} = E[\{\alpha_\tau^{(k)} - a_{\tau-1}^{(k)}\}\{\bar{\alpha}_\tau^{(k)} - \bar{a}_{\tau-1}^{(k)}\}'].$$

If we now let $k \to \infty$, it can be seen that $S^{(k)} P_{\tau/\tau-1} S^{(k)\prime} \to P_{\tau/\tau-1}^{ff}$, where $P_{\tau/\tau-1}^{ff}$ is defined in (28). Hence $F_\tau^{(k)} \to F_\tau$ in (27). Turning to (41), it will be observed that we only need consider the last Np rows since we are only interested in obtaining an updated estimator for the state vector at time t_τ.

Thus, on letting $k \to \infty$, the updating equations for α_τ become (30a) and (30b) with

$$P^f_{\tau/\tau-1} = \int_0^\delta E[\{\alpha(t_{\tau-1}+\delta) - a(t_{\tau-1}+\delta/t_{\tau-1})\}\{\overline{\alpha(t_{\tau-1}+s)} \\ -\overline{a(t_{\tau-1}+s/t_{\tau-1})}\}']\,ds. \tag{44}$$

Substituting from (25) and (32) gives

$$P^f_{\tau/\tau-1} = \int_0^\delta e^{\Lambda\delta} P_{\tau-1} e^{\overline{\Lambda}s}\,ds + \int_0^\delta \int_0^s e^{\Lambda(\delta-u)} Q e^{\overline{\Lambda}(s-u)}\,du\,ds, \tag{45}$$

and evaluating the integrals gives (31).

4. MIXTURES OF STOCKS AND FLOWS

In many models, $y(t)$ will consist of both stocks and flows. In order to develop the appropriate formulas for this case it will be assumed that there are N^s stock variables and N^f flow variables and that the $N = N^s + N^f$ elements of $y(t)$ have been ordered so that the stocks come first.

Using obvious notation the measurement equations can be written as

$$y^s_\tau = Z^s_\tau G\alpha_\tau + \xi^s_\tau \tag{46a}$$

$$y^f_\tau = Z^f G \int_0^\delta \alpha(t_{\tau-1}+s)\,ds + \xi^f_\tau \tag{46b}$$

where Z^s and Z^f are $N^s \times Np$ and $N^f \times Np$ matrices, respectively. The covariance matrix of $(\xi^{s\prime}\xi^{f\prime})'$ will be denoted by H_τ. The elements of $\tilde{y}_{\tau/\tau-1}$ are obtained by replacing the Z_τ matrices in (15) and (26) by Z^s_τ and Z^f_τ, respectively. The covariance matrix of the prediction error, v_τ, is

$$F_\tau = \left[\begin{array}{c|c} Z^s_\tau G P_{\tau/\tau-1} \overline{G}' Z^{s\prime}_\tau & Z^s_\tau G P^f_{\tau/\tau-1} \overline{G}' Z^{f\prime}_\tau \\ \hline Z^f_\tau G P^f_{\tau/\tau-1}{}' \overline{G}' Z^{s\prime}_\tau & Z^f_\tau G P^{ff}_{\tau/\tau-1} \overline{G}' Z^{f\prime}_\tau \end{array}\right] + H_\tau \tag{47}$$

where $P^f_{\tau/\tau-1}$ and $P^{ff}_{\tau/\tau-1}$ are as defined by (31) and (28) respectively. This expression for F_τ can be derived by using the approach set out in Section 3. Similarly, it can be shown that the updating equations are

$$a_\tau = a_{\tau/\tau-1} + P^{sf}_{\tau/\tau-1} F^{-1}_\tau v_\tau \tag{48a}$$

and

$$P_\tau = P_{\tau/\tau-1} - P^{sf}_{\tau/\tau-1} F^{-1}_\tau P^{sf}_{\tau/\tau-1}{}' \tag{48b}$$

where

$$P^{sf}_{\tau/\tau-1} = [P_{\tau/\tau-1}\bar{G}'Z^{s'}_{\tau} \;\vdots\; P^{f}_{\tau/\tau-1}\bar{G}'Z^{f'}_{\tau}]. \tag{48c}$$

The Kalman–Bucy filter therefore consists of the predicton equations, (13), together with the updating equations (48). The ML estimators are again obtained by minimizing a function of the form (19). Rosenberg's algorithm can be applied by fixing $y(t)$ at t_0, as in the pure flow case. Since $\tilde{y}_{\tau/\tau-1}$ can be written as

$$\tilde{y}_{\tau/\tau-1} = \begin{bmatrix} Z^s_\tau G \\ Z_\tau G K_\tau \end{bmatrix} a_{\tau-1}, \tag{49}$$

the matrix $Z_\tau G$ in (20) to (22) is replaced by the matrix in square brackets in (49). F_τ is defined as in (47), while the Kalman gain is

$$K_\tau = P^{sf}_{\tau/\tau-1}F_\tau^{-1}.$$

5. EXOGENOUS VARIABLES

In an open system, the model in (1) becomes

$$D^p y(t) = \sum_{j=1}^{p} A_j D^{p-j} y(t) + Bx(t) + \xi(t),$$

where $x(t)$ is a $K \times 1$ vector of exogenous variables and B is an $N \times K$ matrix of parameters. The system can again be written in first-order form by defining an $Np \times K$ matrix $B^* = [0 \quad \cdots \quad 0 \quad B']'$ and adding $B^*x(t)$ to the right-hand side of (4). The transition equation governing the change in the transformed state vector from $t_{\tau-1}$ to t_τ is

$$\alpha_\tau = e^{\Lambda\delta}\alpha_{\tau-1} + c_\tau + \int_0^\delta e^{\Lambda(\delta-s)}\eta(t_{\tau-1} + s)\,ds,$$

where

$$c_\tau = \int_0^\delta e^{\Lambda(\delta-s)}G^{-1}B^*x(t_{\tau-1} + s)\,ds;$$

cf. (12). If c_τ can be evaluated, the procedures set out in the previous sections will yield the exact likelihood function. The only change necessary is the addition of c_τ to the right-hand side of (13a). The starting values are obtained using Rosenberg's algorithm and since nonstationarity in the elements of $y(t)$

will normally be regarded as coming from nonstationarity in the elements of $x(t)$, it will usually be reasonable to constrain the roots of A to have negative real parts.

The difficulty in applying ML to an open system, even a first-order one, is that the exogenous variables will not normally be analytic functions of time. Hence c_τ cannot be evaluated exactly. However, various approximations can be made to c_τ and if the elements of $x(t)$ are reasonably smooth, the ML estimators will tend to have satisfactory properties; see, for example, References 1 (Chs. 4, 7, and 8), 11, and 14.

6. ASYMPTOTIC PROPERTIES OF ESTIMATOR

For a stationary model with observations at equally spaced intervals, the asymptotic properties of the estimator described in this paper have been established by Bergstrom [2, pp. 147–9] without the assumption that the observations are normally distributed. Under fairly general conditions he shows that the Gaussian estimator of ψ is asymptotically normal with mean vector zero and a covariance matrix equal to the inverse of the large sample information matrix for normally distributed observations. In terms of the quantities used in (19), the finite sample approximation to the large sample information matrix has as its ijth element

$$I_{ij} = \frac{1}{2}\sum_t \operatorname{tr}[F_t^{-1}(\partial F_t/\partial\psi_i)F_t^{-1}(\partial F_t/\partial\psi_j)] + \sum_t (\partial v_t/\partial\psi_i)'F_t^{-1}\partial v_t/\partial\psi_j,$$
$$i, j = 1, \ldots, n; \qquad \textbf{(50)}$$

cf. Engle and Watson [5]. This expression depends on first derivatives only and these derivatives can be evaluated recursively as shown in Appendix B. Note that the derivation of these recursions is not altogether straightforward here because of the transformation needed to diagonalize the **A** matrix.

The corresponding analytic expression for the score vector is

$$\frac{\partial \log L(\psi)}{\partial\psi_i} = -\frac{1}{2}\sum_t \operatorname{tr}(F_t^{-1}\partial F_t/\partial\psi_i)(I - F_t^{-1}v_t v_t') - \sum_t (\partial v_t/\partial\psi_i)'F_t^{-1}v_t,$$
$$i = 1, \ldots, n. \qquad \textbf{(52)}$$

Expression (52) can be used in the numerical optimization procedure, as can (51), although the complexity of the recursions suggests that this may not be worthwhile in practice. Nevertheless, (51) can be expected to yield a more acceptable asymptotic covariance matrix than an estimate based on numerical derivatives.

REFERENCES

1. Bergstrom, A. R. (Ed.) (1976). *Statistical inference in continuous time economic models.* Amsterdam: North Holland, 1976.
2. ———. (1983). Gaussian estimation of structural parameters in higher order continuous time dynamic models. *Econometrica, 51,* 117–152.
3. Chan, S. W., Goodwin, G. C., & Sin, K. W., (1984). Convergence properties of the solution of the Riccati difference equations in optimal filtering of nonstabilizable systems. *IEEE Transactions on Automatic Control, AC-29,* 110–118.
4. Duncan, D. B., & Horn, S. D. (1972). Linear dynamic regression from the viewpoint of regression analysis. *Journal of the American Statistical Association, 67,* 815–821.
5. Engle, R. F., & Watson, M. (1981). A one-factor multivariate time series model of metropolitan wage rates. *Journal of the American Statistical Association 76,* 774–781.
6. Gardner, G., Harvey, A. C., & Phillips, G.D.A. An algorithm for exact maximum likelihood estimation of autoregressive–moving average models by means of Kalman filtering. *Applied Statistics, 29,* 311–322.
7. Harvey, A. C. (1981). *Time Series Models.* Deddington, Oxford: Phillip Allan, and New York: Humanities Press.
8. Jones, R. H. (1981). Fitting a continuous time autoregression to discrete data. In D. F. Findley, (Ed.) *Applied Time Series Analysis 11.* New York: Academic Press, pp. 651–682.
9. ———. (1983). Fitting multivariate models to unequally spaced data. In E. Parzen (Ed.), *Time Series Analysis of Irregular Observations.* New York: Springer-Verlag.
10. Kalman, R. E., & Bucy, R. S. (1961). New results in linear filtering and prediction theory. *Transactions of the ASME, Series D, Journal of Basic Engineering, 83,* 95–108.
11. Phillips, P.C.B. (1974). The estimation of some continuous time models. *Econometrica, 42,* 803–824.
12. Robinson, P. M. (1976). The estimation of linear differential equations with constant coefficients. *Econometrica, 44,* 751–764.
13. Rosenberg B. (1973). Random coefficient models: The analysis of a cross section of time series by stochastically convergent parameter regression. *Annals of Economic and Social Measurement, 2,* 399–428.
14. Wymer, C. R. (1972). Econometric estimation of stochastic differential equation system. *Econometrica, 40,* 565–577.

APPENDIX A

Rosenberg [13] gives an algorithm for obtaining starting values in a discrete time Kalman filter, but provides very little detail on how the algorithm was derived. This appendix provides a derivation.

Consider the discrete time state space model

$$\alpha_t = T_t \alpha_{t-1} + \eta_t, \qquad \eta_t \sim \text{NID}(0, Q_t) \tag{A.1a}$$

$$y_t = Z_t \alpha_t + \xi_t, \qquad \xi_t \sim \text{NID}(0, H_t) \tag{A.1b}$$

for $t = 1, \ldots, T$, with $E(\eta_t \xi_s') = 0$ for all t, s. The Kalman filter for this model is essentially the same as the Kalman–Bucy filter for the continuous time model in Section 2. The only difference lies in the presence of the transformation matrix G and the fact that this results in some of the matrices being complex.

Suppose that the initial-state vector, α_0 is unknown but fixed, so that its elements can be treated as additional parameters. Rosenberg sets up a Kalman filter in which the starting values are $\tilde{a}_0 = 0$ and $\tilde{P}_0 = 0$ and a set of prediction error vectors, $\tilde{v}_1, \ldots, \tilde{v}_T$ are produced. If α_0 were known, the Kalman–Bucy filter would be run with $a_0 = \alpha_0$ and $P_0 = 0$ to give the prediction error vectors, $v_1, \ldots, v_T$. This second set of prediction errors are needed to form the likelihood function. Bearing in mind that the P_t matrices in the two filters are the same, the idea behind Rosenberg's algorithm is to obtain $v_1, \ldots, v_T$ from the output from the first filter.

Write the initial state vector as

$$\boldsymbol{\alpha}_0 = \tilde{a}_0 + \alpha_0$$

where $\tilde{a}_0 = 0$. The prediction and updating equations for the state vector, (13) and (17) can be written together as

$$a_{t/t-1} = T_t(I - K_{t-1}Z_{t-1})a_{t-1/t-2} + T_t K_{t-1} y_{t-1}, \qquad t = 1, \ldots, T \tag{A.2}$$

where k_{t-1} is the Kalman gain. The vector $a_{t/t-1}$ is conditional on α_0, but it can be written as

$$a_{t/t-1} = \tilde{a}_{t/t-1} + J_{t-1}\alpha_0, \qquad t = 1, \ldots, T \tag{A.3}$$

where $\tilde{a}_{t/t-1}$ is output from the first filter and J_t is computed recursively from

$$J_t = T_{t+1}(I - K_t Z_t) J_{t-1}, \qquad t = 1, \ldots, T-1 \tag{A.4}$$

with $J_0 = T_1$.

The prediction error vector can similarly be split up into two parts, i.e.,

$$\begin{aligned} v_t &= y_t - Z_t a_{t/t-1} = y_t - Z_t \tilde{a}_{t/t-1} - Z_t J_{t-1}\alpha_0 \\ &= \tilde{v}_t - Z_t J_{t-1}\alpha_0, \qquad t = 1, \ldots, T. \end{aligned} \tag{A.5}$$

Substituting (A.5) into a likelihood function of the form (19) and differentiating with respect to α_0 gives

$$\tilde{\alpha}_0 = \left[\sum_{t=1}^{T} J_{t-1}{}' Z_t' F_t^{-1} Z_t J_{t-1} \right]^{-1} \sum_{t=1}^{T} J_{t-1}' Z_t F_t^{-1} \tilde{v}_t \tag{A.6}$$

This expression is the ML estimator of α_0, conditional on any other parameters in the model. It can therefore be used to concentrate α_0 out of the likelihood function. This gives

$$\begin{aligned} L(y_1, \ldots, y_T) &= -\frac{1}{2} \sum_{t=1}^{T} \log|F_t| - \frac{1}{2} \sum_{t=1}^{T} (\tilde{v}_t - Z_t J_{t-1} \tilde{\alpha}_0)' F_t^{-1} (\tilde{v}_t - Z_t J_{t-1} \tilde{\alpha}_0) \\ &= -\frac{1}{2} \sum_{t=1}^{T} \log|F_t| - \frac{1}{2} \sum_{t=1}^{T} \tilde{v}_t{}' F_t^{-1} \tilde{v}_t + \frac{1}{2} \tilde{\alpha}_0' \sum_{t=1}^{T} J_{t-1}' Z_t F_t^{-1} \tilde{v}_t. \end{aligned} \tag{A.7}$$

APPENDIX B

In this appendix, computable recursive relations for the derivatives $\partial v_\tau/\partial\psi$ and $\partial F_\tau/\partial\psi$ are presented, first for the case that all variables are stocks, and then for the case that all variables are flows. It is assumed that the parameter values are such that the process is stationary. The extension to the case of both stock and flow variables is straightforward.

First we adopt some notation. Recall that $\alpha_\tau^* = G\alpha_\tau$. Similarly, let $W_\tau^* = GW_\tau G^{-1}$ and $T_\tau^* = GT_\tau G^{-1}$. Let $P_\tau^* = GP_\tau\bar{G}'$, $Q_\tau^* = GQ_\tau\bar{G}'$, and so forth for P_τ^{f*}, Q_τ^{f*}, etc. Let C_n be generated by the recursion $C_n = C_0\Lambda^n + \Lambda C_{n-1}$, where $C_0 = G^{-1}(\partial A/\partial\psi)G$, and let $C_n^* = GC_nG^{-1}$. Also, let

$$\Phi_n(x) = e^{\Lambda x}(-\Lambda)^{-(n+1)} \sum_{k=n+1}^{\infty} (-\Lambda x)^k/k!$$

$$\Phi_n^f(x) = (-\Lambda)^{-(n+1)} \sum_{k=n+1}^{\infty} (k+1)(-\Lambda x)^k/(k+2)!$$

$$\Pi_{n\tau} = (W_\tau - \delta I)\Phi_{n+2}(\delta)e^{-\Lambda\delta} + \Phi_{n+1}^f(\delta).$$

Note that $\Phi_n(x)$ and $\Phi_n^f(x)$ are diagonal; furthermore, they can be generated by relatively simple recursion relations. Finally, let $\Gamma = G^{-1}(\partial(R\Sigma\bar{R}')/\partial\psi)(\bar{G}')^{-1}$, and let the (i, j) elements of Γ_τ, Γ_τ^f, and Γ_τ^{ff} be defined respectively by

$$(\Gamma_\tau)_{ij} = (\Gamma)_{ij}W(\lambda_i + \bar{\lambda}_j; \delta)$$

$$(\Gamma_\tau^f)_{ij} = (\Gamma)_{ij}e^{\lambda_i\delta}[W(\bar{\lambda}_j; \delta) - W(-\lambda_i; \delta)]/(\lambda_i + \bar{\lambda}_j)$$

$$(\Gamma_\tau^{ff})_{ij} = \frac{(\Gamma)_{ij}}{\lambda_i + \bar{\lambda}_j}\left[W(\lambda_i; \delta)W(\bar{\lambda}_j; \delta) - \frac{W(\lambda_i; \delta) - \delta}{\lambda_i} - \frac{W(\bar{\lambda}_j; \delta) - \delta}{\bar{\lambda}_j}\right].$$

Recursions for Stocks

$$\frac{\partial v_\tau}{\partial\psi} = -Z_\tau\frac{\partial T_\tau^*}{\partial\psi}a_{\tau-1}^* - Z_\tau T_\tau^*\frac{\partial a_{\tau-1}^*}{\partial\psi}$$

$$\frac{\partial F_\tau}{\partial\psi} = Z_\tau\frac{\partial P_{\tau/\tau-1}^*}{\partial\psi}Z_\tau' + \frac{\partial H_\tau}{\partial\psi},$$

where

$$\frac{\partial T_\tau^*}{\partial\psi} = \sum_{n=0}^{\infty} \frac{1}{(n+1)!} \delta_\tau^{n+1} C_n^*; \qquad \frac{\partial a_{\tau/\tau-1}^*}{\partial\psi} = \frac{\partial T_\tau^*}{\partial\psi} a_{\tau-1}^* + T_\tau^* \frac{\partial a_{\tau-1}^*}{\partial\psi}$$

$$\frac{\partial a_\tau^*}{\partial\psi} = \frac{\partial a_{\tau/\tau-1}^*}{\partial\psi} + \frac{\partial P_{\tau/\tau-1}^*}{\partial\psi} Z_\tau F_\tau^{-1} v_\tau + P_{\tau/\tau-1}^* Z_\tau F_\tau^{-1} \frac{\partial F_\tau}{\partial\psi} F_\tau^{-1} v_\tau + P_{\tau/\tau-1}^* Z_\tau' F_\tau^{-1} \frac{\partial v_\tau}{\partial\psi}$$

$$\frac{\partial P_{\tau/\tau-1}^*}{\partial\psi} = \frac{\partial T_\tau^*}{\partial\psi} P_{\tau-1}^* \bar{T}_\tau^{*\prime} + T_\tau^* P_{\tau-1}^* \left(\frac{\overline{\partial T_\tau^*}}{\partial\psi}\right)' + T_\tau^* \frac{\partial P_{\tau-1}^*}{\partial\psi} \bar{T}_\tau^{*\prime} + \frac{\partial Q_\tau^*}{\partial\psi}$$

$$\frac{\partial P_\tau^*}{\partial\psi} = \frac{\partial P_{\tau/\tau-1}^*}{\partial\psi} [I - Z_\tau' F_\tau^{-1} Z_\tau P_{\tau/\tau-1}^*] + P_{\tau/\tau-1}^* Z_\tau' F_\tau^{-1} \frac{\partial F_\tau}{\partial\psi} F_\tau^{-1} Z_\tau P_{\tau/\tau-1}^*$$

$$- P_{\tau/\tau-1}^* Z_\tau' F_\tau^{-1} Z_\tau \frac{\partial P_{\tau/\tau-1}^*}{\partial\psi}$$

$$\frac{\partial Q_\tau^*}{\partial\psi} = G\left[\Gamma_\tau + \sum_{n=0}^{\infty} (C_n Q \bar{\Phi}_{n+1}(\delta) + \Phi_{n+1}(\delta) Q \bar{C}_n')\right]\bar{G}'.$$

The derivatives with respect to the initial conditions are $\partial a_{1/0}^*/\partial\psi = 0$ and

$$\frac{\partial P_{1/0}^*}{\partial\psi} = G\left[\Gamma_{1/0} - \sum_{n=0}^{\infty} (C_n Q \bar{\Phi}_{n+1}(-\infty) + \Phi_{n+1}(-\infty) Q \bar{C}_n')\right]\bar{G}'$$

where $(\Gamma_{1/0})_{ij} = (\Gamma)_{ij}/(\lambda_i + \bar{\lambda}_j)$.

Recursions for Flows

$$\frac{\partial v_\tau}{\partial\psi} = -Z_\tau \frac{\partial W_\tau^*}{\partial\psi} a_{\tau-1}^* - Z_\tau W_\tau^* \frac{\partial a_{\tau-1}^*}{\partial\psi}$$

$$\frac{\partial F_\tau}{\partial\psi} = Z_\tau \frac{\partial P_{\tau/\tau-1}^{ff*}}{\partial\psi} Z_\tau' + \frac{\partial H_\tau}{\partial\psi}$$

where

$$\frac{\partial W_\tau^*}{\partial\psi} = \sum_{n=0}^{\infty} \delta_\tau^{n+2} C_n^*/(n+2)!$$

$$\frac{\partial a_\tau^*}{\partial\psi} = \frac{\partial a_{\tau/\tau-1}^*}{\partial\psi} + \frac{\partial P_{\tau/\tau-1}^{f*}}{\partial\psi} Z_\tau' F_\tau^{-1} v_\tau + P_{\tau/\tau-1}^{f*} Z_\tau' F_\tau^{-1} \frac{\partial F_\tau}{\partial\psi} F_\tau^{-1} v_\tau + P_{\tau/\tau-1}^{f*} Z_\tau' F_\tau^{-1} \frac{\partial v_\tau}{\partial\psi}$$

$$\frac{\partial P_{\tau/\tau-1}^{f*}}{\partial\psi} = \frac{\partial T_\tau^*}{\partial\psi} P_{\tau-1}^* \bar{W}_\tau^{*\prime} + T_\tau^* \frac{\partial P_{\tau-1}^*}{\partial\psi} \bar{W}_\tau^{*\prime} + T_\tau^* P_{\tau-1}^* \left(\frac{\overline{\partial W_\tau^*}}{\partial\psi}\right)' + \frac{\partial Q_\tau^{f*}}{\partial\psi}$$

$$\frac{\partial Q_\tau^{f*}}{\partial\psi} = G\left[\Gamma_\tau^f + \sum_{n=0}^{\infty} (C_n Q(\bar{\Phi}_{n+1}(\delta)\bar{W}_\tau - \bar{\Phi}_{n+2}(\delta))e^{-\bar{\Lambda}\delta} + \Phi_{n+2}(\delta) Q \bar{C}_n')\right]\bar{G}'$$

$$\frac{\partial P^{ff*}_{\tau/\tau-1}}{\partial \psi} = \frac{\partial W^*_\tau}{\partial \psi} P^*_{\tau-1} \bar{W}^{*\prime}_\tau + W^*_\tau \frac{\partial P^*_{\tau-1}}{\partial \psi} \bar{W}^{*\prime}_\tau + W^*_\tau P^*_{\tau-1} \left(\frac{\overline{\partial W^*_\tau}}{\partial \psi} \right)' + \frac{\partial Q^{ff*}_\tau}{\partial \psi}$$

$$\frac{\partial Q^{ff*}_\tau}{\partial \psi} = G \left[\Gamma^{ff}_\tau + \sum_{n=0}^{\infty} (C_n Q \bar{\Pi}_{n\tau} + \Pi_{n\tau} Q \bar{C}'_n) \right] \bar{G}'$$

and where T^*_τ, $\partial T^*_\tau / \partial \psi$, P^*_τ, and $\partial P^*_\tau / \partial \psi$ are given by the same expressions as in the case of stocks.