IDENTIFICATION AND ESTIMATION OF CONTINUOUS TIME DYNAMIC SYSTEMS WITH EXOGENOUS VARIABLES USING PANEL DATA

ALFRED HAMERLE AND HERMANN SINGER
University of Regensburg

WILLI NAGL
University of Konstanz

This paper deals with the identification and maximum likelihood estimation of the parameters of a stochastic differential equation from discrete time sampling. Score function and maximum likelihood equations are derived explicitly. The stochastic differential equation system is extended to allow for random effects and the analysis of panel data. In addition, we investigate the identifiability of the continuous time parameters, in particular the impact of the inclusion of exogenous variables.

1. INTRODUCTION

In econometrics, dynamic models are mainly based on discrete time. However, continuous time models, already propounded by Koopmans, were recommended by some authors, including Bergstrom [2-6,8], Phillips [22], Sargan [28], Gandolfo [11], Wymer [32], and others. Another approach was pioneered by Robinson [24-27] who handled the systems in the frequency domain.

This paper focuses on maximum likelihood estimation of the unknown model parameters from discrete time sampling. We estimate the "exact discrete model" which corresponds to the continuous time model in the sense *that observations at given points in time that are generated by the latter system also satisfy the former.* Score function and maximum likelihood equations for the continuous time parameters are derived explicitly. An explicit formula for the Fisher information is given by Singer [29]. As far as we know, only nonlinear optimization techniques with numerical derivatives have been used up to now for maximizing the likelihood function (Bergstrom [5,7,9], Harvey and Stock [13], Jones [15], Jones and Tryon [16], Jones

and Ackerson [17], Zadrozny [33]). However, in larger systems these algorithms may break down (see Dennis and Schnabel [10]).

In the present paper we discuss a fully parameterized model without restrictions. In some applications, restrictions on the system matrices, guided by economic theory, have to be accounted for. This can be achieved by letting the matrices depend on a lower dimensional parameter vector (see, e.g., Bergstrom [6], Zadrozny [33], Harvey and Stock [13]). Using the chain rule, the maximum likelihood equations can be derived analogously. The system can be extended to allow for mixed flow and stock data. Singer [30,31] introduces a continuous/discrete state space model, which allows the treatment of higher order models including errors of measurement, individual specific random effects, and flow data. ML estimates are obtained using EM and quasi-Newton algorithms with exact analytic derivatives.

Both simulation studies and empirical applications show that our estimation procedure is much more efficient and robust than the derivative-free optimization algorithms. In the present paper the stochastic differential equations system is extended to allow for the analysis of panel data involving short time series but many observations at any given point in time. For a discussion of the advantages of panel data see Hsiao [14]. In addition, we investigate the identifiability of the continuous time parameters, in particular the impact of the inclusion of exogenous variables.

In the next section we give a short review of the solution of stochastic differential equations, including additional random effects to account for time-invariant omitted variables. In Section 3 we consider in detail maximum likelihood estimation, and Section 4 contains the identification results.

2. MODEL SPECIFICATION

In the following we consider the system of linear stochastic differential equations

$$dy_n(t) = [Ay_n(t) + Bx_n(t) + \pi_n] \, dt + G \, dW_n(t),$$

$$t \in [t_o, t_T], \qquad n = 1, \dots, N \tag{2.1}$$

where $W_n(t)$ is the m-dimensional Wiener process, and formally $dW_n(t) = \xi_n(t) \, dt$ where $\xi_n(t)$ is "white noise" with zero mean and $E(\xi_n(t)\xi_n(s)') = I_m \delta(t - s)$. $W_n(t)$ and $W_{n'}(t)$ are assumed to be independent. $y_n(t)$ is a p-vector of endogenous variables, and $x_n(t)$ is a q-vector of exogenous or control variables. A, B, and G are $p \times p$, $p \times q$, and $p \times m$-matrices. In general, A, B, and G contain unknown parameters which must be estimated from the data. π_n are the individual specific components representing the effects of unobserved variables which remain constant over t for given n. Here we only deal with the random effect approach. We assume that the $\{\pi_n\}$ are i.i.d. according to $N(0; V_\pi)$. Moreover, we assume that π_n is independent of

the Wiener process $W_n(t)$. Note that the individual specific effects in (2.1) affect the rates of change and not the absolute levels of the dependent variables. The initial values $y_n(t_o)$ may vary from individual to individual. They may be constants or random variables.

Another approach is the fixed effects model. Then, the $\{\pi_n\}$ are unknown parameters which have to be estimated along with the regression coefficients and the other model parameters. This approach is already included in (2.1). The normality assumption is dropped and the parameters π_n can be absorbed in B by using appropriate dummy coding of the exogenous variables. However, as we shall see in the next section, the resulting discrete-time model for the observations includes lagged endogenous variables. Then, if the number T of observation times is fixed (and small) and N is large, we face an incidental parameter problem (see, e.g., Hsiao [14, Chapter 4], for details). Hence we use a random effect framework. On the other hand, time effects λ_t (such as economy wide shocks) can be easily included in the model as fixed effects, since T is small. We only have to incorporate appropriate dummy variables in the matrix of exogenous variables.

We note that, if random effects are included, the solution of the stochastic differential equation (2.1) is no longer a Markov process. But if we extend the system by including the vector $d\pi_n(t) = 0$ with the random initial value $\pi_n(t_o) = \pi_n$ where π_n is $N(0; V_\pi)$, then the solution of the complete dynamic system is a Markov process. The solution of the y_n-part in which we are interested is given by (Arnold [1]):

$$y_n(t) = \exp(A(t - t_o))y_n(t_o) + \int_{t_o}^{t} \exp(A(t - s))Bx_n(s)\, ds$$

$$+ A^{-1}(\exp(A(t - t_o)) - I)\pi_n + \int_{t_o}^{t} \exp(A(t - s))GdW_n(s).$$

$$(2.2)$$

Under the assumptions stated above, $y_n(t)$ is also a Gauss process.

3. MAXIMUM LIKELIHOOD ESTIMATION

3.1. The General Maximum Likelihood Estimator

Assuming equally spaced observation intervals $\Delta t = t_j - t_{j-1}$, we can write (Phillips [23, p. 137])

$$y_n(t_j) - \exp(A\,\Delta t)y_n(t_{j-1}) - \int_0^{\Delta t} \exp(A(\Delta t - s))Bx_n(t_{j-1} + s)\, ds$$

$$= \delta_n + \varepsilon_{nj} \qquad n = 1,\ldots,N, \quad j = 1,\ldots,T \qquad (3.1)$$

where

$$\delta_n = A^{-1}(\exp(A \Delta t) - I)\pi_n \quad \text{and}$$

$$\varepsilon_{nj} = \int_0^{\Delta t} \exp(A(\Delta t - s))G \, dW_n(t_{j-1} + s).$$

The covariance structure of the disturbances is given by

$$\text{cov}(\delta_n, \delta_{n'}) = \delta_{nn'} A^{-1}(\exp(A \Delta t) - I)V_\pi(\exp(A' \Delta t) - I)A^{-T}$$

$$:= \delta_{nn'} V_\delta \tag{3.2}$$

and

$$\text{cov}(\varepsilon_{nj}, \varepsilon_{n'k}) = \delta_{nn'} \delta_{jk} V_\varepsilon$$

where

$$V_\varepsilon = \int_0^{\Delta t} \exp(As)\Omega \exp(A's) \, ds, \qquad \Omega = GG'$$

and δ_{jk} denotes the Kronecker symbol.

For the covariance structure of $\delta_n + \varepsilon_{ni}$ we obtain

$$\text{cov}(\delta_n + \varepsilon_{ni}, \delta_m + \varepsilon_{mj}) = \delta_{nm}(V_\delta + \delta_{ij}V_\varepsilon),$$

$$i,j = 1, \ldots, T; \qquad n,m = 1, \ldots, N.$$

In matrix notation this reads

$$I_N \otimes (1_T 1_T' \otimes V_\delta + I_T \otimes V_\varepsilon) \tag{3.3}$$

where I_N is an identity matrix of order N, 1_T a $T \times 1$ vector of ones, and $\otimes$ denotes the Kronecker product.

We have used the assumption that the Wiener process $W_n(t)$ is independent of the individual specific effect π_n.

We define

$$z_n(t_j) = y_n(t_j) - \exp(A \Delta t)y_n(t_{j-1})$$

$$- \int_0^{\Delta t} \exp(A(\Delta t - s))Bx_n(t_{j-1} + s) \, ds$$

and the error term is denoted by $u_{nj} = \delta_n + \varepsilon_{nj}$. In addition, let us define

$$Z_n = [z_n(t_1), \ldots, z_n(t_T)]', \qquad Y_n = [y_n(t_1), \ldots, y_n(t_T)]',$$

$$Y_{-1,n} = [y_n(t_0), \ldots, y_n(t_{T-1})]' \qquad F_n(A,B,X) = [f_{n1}, \ldots, f_{nT}]',$$

where

$$f_{nj} = \int_0^{\Delta t} \exp(A(\Delta t - s))Bx_n(t_{j-1} + s)\, ds$$

$$U_n = [u_{n1}, \ldots, u_{nT}]' \qquad n = 1, \ldots, N, \qquad j = 1, \ldots, T. \tag{3.4}$$

The system (3.1) can be rewritten as

$$Z_n = Y_n - Y_{-1,n}A_* - F_n(A,B,X) = U_n$$

where $A_* = \exp(A' \Delta t)$.

For notational convenience we define:

$$Z = \begin{bmatrix} Z_1 \\ \cdots \\ Z_N \end{bmatrix}, \qquad Y = \begin{bmatrix} Y_1 \\ \cdots \\ Y_N \end{bmatrix}, \qquad Y_{-1} = \begin{bmatrix} Y_{-1,1} \\ \cdots \\ Y_{-1,N} \end{bmatrix},$$

$$F = \begin{bmatrix} F_1 \\ \cdots \\ F_N \end{bmatrix}, \qquad U = \begin{bmatrix} U_1 \\ \cdots \\ U_N \end{bmatrix}.$$

The covariance matrix of the disturbances U_n can be written compactly as

$$\Sigma = \text{cov}(\text{vec}(U_n)) = 1_T 1_T' \otimes V_\delta + I_T \otimes V_\varepsilon.$$

Here the vec-operation is defined by stacking the rows of a matrix into a column vector.

Since we have assumed that π_n is normally distributed, the log-likelihood of N individuals (conditional on the initial values) is obtained as

$$l(A,B,V_\varepsilon(A,\Omega), V_\delta(A,V_\pi); Y, Y_{-1}, X)$$

$$= \text{const} - N/2[\log(\det(\Sigma)) + \text{tr}(\Sigma^{-1}S)] \tag{3.5}$$

where $S = 1/N\sum\text{vec}(Z_n)\text{vec}'(Z_n)$.

Utilizing the McDonald–Swaminathan-calculus for symbolic matrix differentiation (see McDonald and Swaminathan [19] and also Magnus and Neudecker [18]) we obtain the following maximum likelihood equations:

(1) $\partial l/\partial\Omega = -\frac{1}{2}\partial\Sigma/\partial\Omega\, \text{vec}(\Sigma^{-1} - \Sigma^{-1}S\Sigma^{-1})$ $= 0$

(2) $\partial l/\partial V_\pi = -\frac{1}{2}\partial\Sigma/\partial V_\pi\, \text{vec}(\Sigma^{-1} - \Sigma^{-1}S\Sigma^{-1})$ $= 0$

(3) $\partial l/\partial A = -\frac{1}{2}\partial\Sigma/\partial A\, \text{vec}(\Sigma^{-1} - \Sigma^{-1}S\Sigma^{-1}) - \frac{1}{2}\partial S/\partial A\, \text{vec}(\Sigma^{-1}) = 0$

(4) $\partial l/\partial B = \qquad\qquad\qquad\qquad\qquad - \frac{1}{2}\partial S/\partial B\, \text{vec}(\Sigma^{-1}) = 0$

$$\tag{3.6}$$

If there are restrictions on the structural parameter matrices, they are assumed to depend on a lower dimensional parameter vector Θ. Then, using

the chain rule, the matrix derivatives $(\partial/\partial\Omega)$, $(\partial/\partial V_\pi)$, $(\partial/\partial A)$, and $(\partial/\partial B)$ have to be replaced by $(\partial\Omega/\partial\Theta)(\partial/\partial\Omega)$, $(\partial V_\pi/\partial\Theta)(\partial/\partial V_\pi)$, $(\partial A/\partial\Theta)(\partial/\partial A)$, and $(\partial B/\partial\Theta)(\partial/\partial B)$. Since we attempt to obtain an explicit solution, we only consider the case without restrictions.

In order to solve the ML-equations (3.6) we have to consider the block structure of the covariance matrix Σ and distinguish the following cases:

3.2. Model Without Individual Specific Effects

Here $\Sigma = I_T \otimes V_\varepsilon$, so that (see 3.6.1)

$$(\partial\Sigma/\partial\Omega) = [\partial V_\varepsilon/\partial\Omega, 0,0,\ldots,\partial V_\varepsilon/\partial\Omega, 0,\ldots,\partial V_\varepsilon/\partial\Omega] = \text{vec}'I_T \otimes \partial V_\varepsilon/\partial\Omega$$

where we used partitioned McDonald–Swaminathan-derivatives. Abbreviating the block matrix $\Sigma^{-1} - \Sigma^{-1}S\Sigma^{-1}$ by R and denoting the blocks by $R_{ij}, i,j = 1,\ldots,T$, we get from (3.6.1)

$$(\partial\Sigma/\partial\Omega)\text{vec}\, R = \partial V_\varepsilon/\partial\Omega \text{ vec} \sum_i R_{ii} = 0,$$

where $R_{ij} = \delta_{ij}V_\varepsilon^{-1} - V_\varepsilon^{-1}S_{ij}V_\varepsilon^{-1}$.

This implies

$$0 = \partial V_\varepsilon/\partial\Omega \text{ vec} \left(TV_\varepsilon^{-1} - V_\varepsilon^{-1}\sum_i S_{ii}V_\varepsilon^{-1}\right).$$

A sufficient condition is $\hat{V}_\varepsilon = 1/T\sum S_{ii}$.

Thus equation (3.6.1) can be fulfilled by the choice $\hat{\Sigma} = I_T \otimes \hat{V}_\varepsilon$. Since $\partial\Sigma/\partial A$ is of the same block structure as $\partial\Sigma/\partial\Omega$, the first term in equation (3.6.3) vanishes when $\hat{\Sigma}$ is inserted. Thus we are left with the equations

$$-\tfrac{1}{2}\partial S/\partial A \text{ vec}(\Sigma^{-1}) = 0$$

$$-\tfrac{1}{2}\partial S/\partial B \text{ vec}(\Sigma^{-1}) = 0. \tag{3.7}$$

3.3. Model with Individual Specific Effects

Here $\Sigma = I_T \otimes V_\varepsilon + 1_T 1'_T \otimes V_\delta$ and $\Sigma^{-1} = I_T \otimes V_\varepsilon^{-1} - 1_T 1'_T \otimes W^{-1}$, where $W = TV_\varepsilon + V_\varepsilon V_\delta^{-1} V_\varepsilon$. Equations (3.6.1) and (3.6.2) now yield

$$(\partial\Sigma/\partial\Omega)\text{vec}\, R = [\partial V_\varepsilon/\partial\Omega, 0,0,\ldots,\partial V_\varepsilon/\partial\Omega, 0,\ldots,\partial V_\varepsilon/\partial\Omega]\text{vec}\, R = 0$$

$$(\partial\Sigma/\partial V_\pi)\text{vec}\, R = [\partial V_\delta/\partial V_\pi,\ldots,\partial V_\delta/\partial V_\pi]\text{vec}\, R = 0 \tag{3.8}$$

where $R = \Sigma^{-1} - \Sigma^{-1}S\Sigma^{-1}$,

$$R_{il} = \delta_{il}V_\varepsilon^{-1} - W^{-1}$$

$$- (V_\varepsilon^{-1}S_{il}V_\varepsilon^{-1} - V_\varepsilon^{-1}S_{i.}W^{-1} - W^{-1}S_{.l}V_\varepsilon^{-1} + W^{-1}S_{..}W^{-1})$$

and $S_{i.} = \sum_j S_{ij}$, and so on.

Equations (3.8) imply

$$\partial V_\varepsilon / \partial \Omega \, \text{vec} \sum_i R_{ii} = 0$$

$$\partial V_\delta / \partial V_\pi \, \text{vec} \sum_{i,j} R_{ij} = 0. \tag{3.9}$$

An obvious solution is $\sum_i R_{ii} = 0$ and $\sum_{ij} R_{ij} = 0$. Defining $S_= = \sum_i S_{ii}$ and $S_{\neq} = \sum_{i \neq j} S_{ij}$, these equations read

$$0 = T(V_\varepsilon^{-1} - W^{-1}) - [V_\varepsilon^{-1} S_= V_\varepsilon^{-1} - V_\varepsilon^{-1} S_{..} W^{-1}$$
$$- W^{-1} S_{..} V_\varepsilon^{-1} - TW^{-1} S_{..} W^{-1}] \tag{3.10.1}$$

$$0 = TV_\varepsilon^{-1} - T^2 W^{-1} - [V_\varepsilon^{-1} S_{..} V_\varepsilon^{-1} - TV_\varepsilon^{-1} S_{..} W^{-1}$$
$$- TW^{-1} S_{..} V_\varepsilon^{-1} - T^2 W^{-1} S_{..} W^{-1}]. \tag{3.10.2}$$

After some manipulations (3.10.2) can be written as

$$T^2 V_\delta + TV_\varepsilon = S_{..}$$

and inserting (3.10.1) we obtain

$$\hat{V}_\varepsilon = \frac{1}{T} S_= - \frac{1}{T(T-1)} S_{\neq}$$

$$\hat{V}_\delta = \frac{1}{T(T-1)} S_{\neq}. \tag{3.11}$$

Again (3.6.1) and (3.6.2) can be fulfilled by the choice

$$\hat{\Sigma} = I_T \otimes \hat{V}_\varepsilon + 1_T 1_T' \otimes \hat{V}_\delta \tag{3.12}$$

and since the structure of $\partial \Sigma / \partial A$ is

$$[\partial V_\varepsilon / \partial A, 0, 0, \ldots, \partial V_\varepsilon / \partial A, 0, 0, \ldots, \partial V_\varepsilon / \partial A] + [\partial V_\delta / \partial A, \ldots, \partial V_\delta / \partial A]$$

we are led via (3.9) to the same set of equations (3.7).

3.4. ML-Equations for *A* and *B*

In order to solve (3.7), $\partial S / \partial A$ and $\partial S / \partial B$ can be expressed as

$$\partial S / \partial A = 1/N \sum \partial [\text{vec}(Z_n) \, \text{vec}'(Z_n)] / \partial A$$
$$= 1/N \sum \partial Z_n / \partial A (I_{pT} \otimes \text{vec}' Z_n + \text{vec} Z_n' \otimes I_{pT})$$
$$\partial S / \partial B = 1/N \sum \partial [\text{vec}(Z_n) \, \text{vec}'(Z_n)] / \partial B$$
$$= 1/N \sum \partial Z_n / \partial B (I_{pT} \otimes \text{vec}' Z_n + \text{vec} Z_n' \otimes I_{pT})$$

and we obtain from (3.7)

$$\sum (\partial Z_n/\partial A)\Sigma^{-1} \text{vec } Z_n = 0$$

$$\sum (\partial Z_n/\partial B)\Sigma^{-1} \text{vec } Z_n = 0.$$

Calculating $\partial Z_n/\partial A$ and $\partial Z_n/\partial B$ we can write [using vec(vec R) = vec(vec R')]

$$\sum [(\partial A_*/\partial A)(Y'_{-1,n} \otimes I_p) + \partial F_n/\partial A]\Sigma^{-1} \text{vec}(Y_n - Y_{-1,n}A_* - F_n) = 0$$

$$\sum (\partial F_n/\partial B)\Sigma^{-1} \text{vec}(Y_n - Y_{-1,n}A_* - F_n) = 0. \tag{3.13}$$

F_n is a functional of the exogenous variables and the parameters A and B. The values of the exogenous variables, however, are known only at fixed points in time. So, in general, the functional F_n can be calculated only approximately. Two simple approximations are discussed in the next section. Sometimes there is knowledge about the functional form of the time paths and the integrals f_{nj} can be calculated explicitly. Another approach is assuming that the vector $x_n(t)$ is itself generated by a differential equation.

4. ON THE IDENTIFICATION PROBLEM

The drift matrix A is, in general, nonsymmetric and may have complex eigenvalues. Then, the matrix equation $A^* = \exp(A \Delta t)$ will not have a unique solution and we run into a well-known identification problem (Phillips [21]). It turns out that the inclusion of exogenous variables may help to identify the continuous time model.

We discuss two simple approximation methods for the exogenous variables in order to calculate the functional

$$f_{nj} = \int_0^{\Delta t} \exp(A(\Delta t - s))Bx_n(t_{j-1} + s) \, ds \tag{4.1}$$

explicitly (more complicated approximations are discussed in Phillips [23]): (i) $x(t)$ are (or are approximated by) step functions. (ii) $x(t)$ change linearly between the observation times (polygonal lines). It is well known that the approximations introduce an asymptotic bias which depends on the smoothness of the true path $x_n(t)$ (see, e.g., Sargan [28], Phillips [23], Bergstrom [6]). According to the mean value theorem, the step function approximation can be shown to introduce a bias of order $O(\Delta t^2)$, whereas in the case of polygonal lines an approximation error of $O(\Delta t^3)$ arises. Our starting point is the exact discrete model (cf. (3.1),(3.4))

$$Y - Y_{-1}A_* - F(A,B,X) = U$$

where $A^* = \exp(A \Delta t) = (A_*)'$, and F is given in (3.4).

In the sequel we employ the following **assumptions:**

Assumption 1. All eigenvalues of A are distinct. Then, A can be diagonalized and represented as $A = P\Lambda P^{-1}$. Let A have r real and $2c$ complex eigenvalues ($p = r + 2c$). If A is a solution of $A^* = \exp(A \Delta t)$, then other solutions A_K can be represented as

$$A_K = A + (2\pi i/\Delta t)P\,\text{diag}(0,K,-K)P^{-1}$$

$$:= A + M_K \tag{4.2}$$

where 0 is an r-nullvector and K is a c-vector of integers (see Phillips [21]).
■

Assumption 2. The real parts of the eigenvalues Λ are negative. This implies that A and A_K are nonsingular.   ■

Assumption 3. The reduced form parameters A^*, B^*, C^*, V_ε, and V_δ are identified.   ■

Furthermore, the following **properties** are utilized (we use the notation $[X, Y] := XY - YX$ for the commutator of X and Y).

Property 1. A commutes with A^*, A^{-1}, A_K, and M_K.   ■

Property 2. A^* and A are diagonalized by the same matrix P.   ■

Property 3. $[A \otimes I, I \otimes A] = 0$. This implies $\exp(L \Delta t) = \exp((A \otimes I + I \otimes A)\Delta t) = (\exp(A \Delta t) \otimes I)(I \otimes \exp(A \Delta t)) = A^* \otimes A^*$.   ■

Property 4. $(A \otimes I + I \otimes A)$ is nonsingular. This is a consequence of Assumption 2.   ■

(i) *Exogenous variables approximated by step functions.* If the exogenous variables are piecewise constant between measurements we obtain from (4.1)

$$f_{nj} = B^* x_n(t_{j-1})$$

where $B^* = A^{-1}(A^* - I)B$. Then, the exact discrete model reads ($B^* = B'_*$)

$$Y - Y_{-1}A_* - X_{-1}B_* = U \tag{4.3}$$

where $X_{-1} = [x_1(t_0),\ldots,x_1(t_{T-1}),\ldots,x_N(t_0),\ldots,x_N(t_{T-1})]'$.

PROPOSITION 1. *If the exogenous variables are approximated by step functions and the reduced form (discrete time) parameters are identified, not all continuous time parameters can be uniquely inferred and we have:*

1. *If A^* has complex eigenvalues, A cannot be uniquely determined from A^*.*
2. *Ω is identified, iff the condition L_K vec $V_\varepsilon = L$ vec $V_\varepsilon(*)$ is fulfilled.*
3. *If ($*$) is not fulfilled, there are only finitely many solutions (A_K, Ω_K) with $\Omega_K \geq 0$ (Hansen and Sargent [12]).*
4. *B is identified iff B^* is in the null space of $M_K = (2\pi i/\Delta t)P\,\text{diag}(0,K,-K)P^{-1}$.*
5. *V_π is identified iff V_δ satisfies the condition $A_K V_\delta A'_K = A V_\delta A'$.*   ■

Proof. Suppose that $(A^*, B^*, V_\varepsilon, V_\delta)$ is given and that some of the eigenvalues of A^* are complex. We have to solve the equations

$$A^* = \exp(A \, \Delta t) \tag{4.4.1}$$

$$B^* = A^{-1}(A^* - I)B \tag{4.4.2}$$

$$V_\varepsilon = \int \exp(As)\Omega \exp(A's) \, ds \tag{4.4.3}$$

$$V_\delta = A^{-1}(A^* - I)V_\pi(A^* - I)'(A^{-1})' \tag{4.4.4}$$

for the continuous time parameters A, B, Ω, and V_π.

According to (4.2), equation (4.4.1) has, in general, infinitely many solutions A_K, where A can be obtained by taking the matrix logarithm of A^* (principle value). Then, (4.4.2) implies the solutions

$$B_K = (A^* - I)^{-1}A_K B^*. \tag{4.5}$$

However, it may happen that $A_K B^* = AB^*$. Inserting (4.1), we obtain

$$M_K B^* = 0_{p \times q} \tag{4.6}$$

where

$$M_K = (2\pi i/\Delta t)P \, \mathrm{diag}(0, K, -K)P^{-1}.$$

Now

$$\mathrm{rank}(\mathrm{diag}(0, K, -K)) \geq 2, \tag{4.7}$$

and thus $\mathrm{rank}(M_K) \geq 2$, if the trivial cases $p = 1$ or $r = p \geq 2$ (all eigenvalues are real; no aliasing) and $K = (0, \ldots, 0)$ are excluded. If B^* is such that (4.6) can be fulfilled (this is the case if the columns of B^* or B are in the null space of M_K), all $B_K = B$ and B is identified.

Let us now exploit (4.4.3). The integral

$$V_\varepsilon = \int_0^{\Delta t} \exp(As)\Omega \exp(A's) \, ds$$

is explicitly given as

$$\mathrm{vec} \, V_\varepsilon = L^{-1}(\exp(L \, \Delta t) - I)\mathrm{vec} \, \Omega \tag{4.8}$$

where $L = (A \otimes I + I \otimes A)$ (cf. Properties 3 and 4).

Equation (4.8) implies the solutions

$$\mathrm{vec} \, \Omega_K = (\exp(L_K \Delta t) - I)^{-1}L_K \, \mathrm{vec} \, V_\varepsilon$$

$$= (A^* \otimes A^* - I)^{-1}L_K \, \mathrm{vec} \, V_\varepsilon \tag{4.9}$$

where $L_K = A_K \otimes I + I \otimes A_K$.

Ω_K is symmetric but, as pointed out by Hansen and Sargent [12], not necessarily positive semidefinite with the consequence that the set of observationally equivalent matrices is finite. However, if

$$L_K \text{ vec } V_\varepsilon = L \text{ vec } V_\varepsilon \tag{4.10}$$

could be fulfilled, only one solution $\Omega_K = \Omega$ is obtained. (4.10) is equivalent to

$$[M_K \otimes I + I \otimes M_K] \text{vec } V_\varepsilon = 0$$

or

$$M_K V_\varepsilon + V_\varepsilon M_K' = 0. \tag{4.11}$$

According to Theorem 2 of Hansen and Sargent [12], if we have $V_\varepsilon(A_K, \Omega) = V_\varepsilon(A, \Omega)$ for some $A_K \neq A$, then infinitely many structures (A_K, Ω) are obtained. If (4.10) or (4.11) is fulfilled, their Theorem 2 can be applied.

Finally, let us solve (4.4.4). We obtain

$$V_{\pi K} = (A^* - I)^{-1} A_K V_\delta A_K' (A^{*'} - I)^{-1}. \tag{4.12}$$

All $V_{\pi K}$ are positive semidefinite, as required. However, if the condition $A_K V_\delta A_K' = A V_\delta A'$ holds, there is only one solution V_π. Inserting (4.2) it is seen that

$$[M_K \otimes A + A \otimes M_K + M_K \otimes M_K] \text{vec } V_\delta = 0$$

must be satisfied, which is possible for $p > 2$, since rank $[.] \geq (p - 2)^2 = p^2 - 4(p - 1)$. ∎

(ii) *Polygonal approximation.*   Here the exogenous variables are approximated by piecewise linear functions. Now we obtain the following (pseudo)-exact discrete model

$$Y - Y_{-1} A_* - X_{-1} B_* - \Delta X C_* = U \tag{4.13}$$

where $C^*(A, B) = -A^{-1} B + (A \Delta t)^{-1} B^*$, $\Delta X = [\Delta x_{11}, \ldots, \Delta x_{1T}, \ldots, \Delta x_{N1}, \ldots, \Delta x_{NT}]'$ and $\Delta x_{nj} = x_n(t_j) - x(t_{j-1})$.

In this case we obtain a further condition (the others are given in (4.4))

$$C^* = -A^{-1} B + (A \Delta t)^{-1} B^*$$

$$= -(A^* - I)^{-1} B^* + (A \Delta t)^{-1} B^*. \tag{4.14}$$

Now one of the solutions A_K of (4.4.1) must fulfill (4.14), say $A_{K'}$. The other A_K, $K \neq K'$, can satisfy (4.14) only if

$$A_{K'}^{-1} B^* = A_K^{-1} B^*.$$

From this we again get condition (4.6) (set $K - K' \to K$)

$$M_K B^* = 0_{p \times q}. \tag{4.15}$$

Since rank$(M_K) \geq 2$, there are at most $p - 2$ linear independent columns of B^* which are in the null space of M_K. Thus we have proved:

PROPOSITION 2. *Let B^* (and thus B) have column rank $\geq p - 1$. Then (4.15) cannot be fulfilled and only one A_K solves (4.14). In this case all continuous time parameters are identified from equally spaced observations.* ∎

Proposition 2 shows that the inclusion of additional exogenous variables may help to identify the continuous time parameters from discrete time observations. Note that in the case $p = 2$ equation (4.15) cannot hold and the model is identified. Furthermore, it should be noted that restrictions on the parameter matrices of the system may also be an aid to identification.

REFERENCES

1. Arnold, L. *Stochastic Differential Equations.* New York: Wiley, 1974.
2. Bergstrom, A.R. Non-recursive models as discrete approximations to systems of stochastic differential equations. *Econometrica* 34 (1966): 173–182.
3. Bergstrom, A.R. *The Construction and Use of Economic Models.* London: English Universities Press, 1967.
4. Bergstrom, A.R. (ed.), *Statistical Inference in Continuous Time Economic Models.* Amsterdam: North-Holland, 1976.
5. Bergstrom, A.R. Gaussian estimation of structural parameters in higher continuous time dynamic models. *Econometrica* 51 (1983): 117–152.
6. Bergstrom, A.R. Continuous time stochastic models and issues of aggregation over time. In Z. Griliches and M.D. Intrilligator (eds.), *Handbook of Econometrics*, Vol. II. Amsterdam: North-Holland, 1984.
7. Bergstrom, A.R. The estimation of parameters in nonstationary higher-order continuous-time dynamic models. *Econometric Theory* 1 (1985): 369–385.
8. Bergstrom, A.R. The history of continuous-time econometric models. *Econometric Theory* 4 (1988): 365–383.
9. Bergstrom, A.R. *Continuous Time Econometric Modelling.* Oxford: Oxford University Press, 1990.
10. Dennis, J.E. Jr. & R.B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* Englewood Cliffs; NJ: Prentice-Hall, 1983.
11. Gandolfo, G. *Qualitative Analysis and Econometric Estimation of Continuous Time Dynamic Models.* Amsterdam: North-Holland, 1981.
12. Hansen, L.P. & T.J. Sargent. The dimensionality of the aliasing problem in models with rational spectral densities. *Econometrica* 51 (1983): 377–387.
13. Harvey, A.C. & J.H. Stock. The estimation of higher-order continuous time autoregressive models. *Econometric Theory* 1 (1985): 97–117.
14. Hsiao, C. *Analysis of Panel Data.* Cambridge: Cambridge University Press, 1986.
15. Jones, R.H. Fitting multivariate models to unequally spaced data. In E. Parzen (ed.), *Time Series Analysis of Irregularly Observed Data.* New York: Springer, 1984.
16. Jones, R.H. & P.V. Tryon. Continuous time series models for unequally spaced data applied to modeling atomic clocks. *SIAM Journal of Scientific Statistical Computing* 8 (1987): 71–81.
17. Jones, R.H. & L.M. Ackerson. Serial correlation in unequally spaced longitudinal data. *Biometrika* 77 (1990): 721–731.
18. Magnus, J.R. & H. Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics.* Chichester: Wiley, 1988.
19. McDonald, R.P. & H. Swaminathan. A simple matrix calculus with applications to multivariate analysis. *General Systems* XVIII (1973): 37–54.

20. Phillips, P.C.B. The structural estimation of a stochastic differential equation system. *Econometrica* 40 (1972): 1021-1041.

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

22. Phillips, P.C.B. The estimation of some continuous time models. *Econometrica* 42 (1974): 803-824.

23. Phillips, P.C.B. The estimation of linear stochastic differential equations with exogenous variables. In A.R. Bergstrom (ed.), *Statistical Inference in Continuous Time Economic Models*, pp. 135-173. Amsterdam: North-Holland, 1976.

24. Robinson, P.M. The estimation of linear differential equations with constant coefficients. *Econometrica* 44 (1976): 751-764.

25. Robinson, P.M. Instrumental variables estimation of differential equations. *Econometrica* 44 (1976): 765-776.

26. Robinson, P.M. The construction and estimation of continuous time models and discrete approximations in econometrics. *Journal of Econometrics* 6 (1977): 173-198.

27. Robinson, P.M. Continuous model fitting from discrete data. In D.R. Brillinger and G.C. Tiao (eds.), *Directions in Time Series*. Ames: Iowa State University, 1978.

28. Sargan, J.D. Some discrete approximations to continuous time stochastic models. In A.R. Bergstrom (ed.), *Statistical Inference in Continuous Time Economic Models*, pp. 27-80. Amsterdam: North-Holland, 1976.

29. Singer, H. *Parameterschätzung in Zeitkontinuierlichen Dynamischen Systemen.* Konstanz: Hartung-Gorre, 1990.

30. Singer, H. *LSDE—A Program Package for the Simulation, Graphical Display, Optimal Filtering and Maximum Likelihood Estimation of Linear Stochastic Differential Equations, User's Guide.* Meersburg: Author, 1991.

31. Singer, H. Continuous-time dynamical systems with sampled data, errors of measurement and unobserved components. *Journal of Time Series Analysis* (1993), forthcoming.

32. Wymer, C.R. Econometric estimation of stochastic differential equation systems. *Econometrica* 40 (1972): 565-577.

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