*Econometric Theory*, **25**, 2009, 985–994. Printed in the United States of America.
doi:10.1017/S0266466608090373

# ON DISCRETE SAMPLING OF TIME-VARYING CONTINUOUS-TIME SYSTEMS

PETER M. ROBINSON
*London School of Economics*

We consider a multivariate continuous-time process, generated by a system of linear stochastic differential equations, driven by white noise, and involving coefficients that possibly vary over time. The process is observable only at discrete, but not necessarily equally-spaced, time points (though equal spacing significantly simplifies matters). Such settings represent partial extensions of ones studied extensively by A.R. Bergstrom. A model for the observed time series is deduced. Initially we focus on a first-order model, but higher-order models are discussed in the case of equally-spaced observations. Some discussion of issues of statistical inference is included.

## 1. INTRODUCTION

Continuous-time modeling in econometrics was pioneered by Rex Bergstrom. At least if calendar effects can be ignored, econometric time series data are usually observed, perhaps after temporal aggregation, at discrete, regular intervals of time. On the other hand, the interval of measurement may have no economic significance, in which case an underlying continuous-time process can be imagined. If so, modeling based on the continuous-time process may be more natural and elegant than modeling directly the discrete-time observations. In his 1966 *Econometrica* paper and 1967 book, Bergstrom (1966, 1967) initiated continuous-time modeling of discrete-time macroeconometric data as a topic of econometric research. Consequently, interest expanded in the 1970s, as illustrated by the collection of articles by a number of researchers that Bergstrom (1976) edited. On a personal note, I was then a young researcher whose interest in the topic was considerably stimulated by reading Rex's early work. He kindly published a chapter of my Ph.D. thesis in Bergstrom (1976) and was always immensely kind and helpful to me in my career.

Building on his early development of the subject, Rex Bergstrom continued to contribute to it very significantly and uniquely for the rest of his life. His work included several advances in econometric theory, developing exact and approximate

This research was supported by ESRC Grant RES-062-23-0036. I am grateful for the helpful comments of Peter Phillips and three referees. Address correspondence to Peter M. Robinson, Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom; email: p.m.robinson@lse.ac.uk.

© 2009 Cambridge University Press 0266-4666/09 $15.00 

https://doi.org/10.1017/S0266466608090373 Published online by Cambridge University Press

discrete-time models and justifying rules of large-sample statistical inference on parameters in the underlying continuous-time model (e.g., Bergstrom, 1983); this work was marked by considerable rigor and care, and matched the challenge of the topic. Perhaps even more distinctively, Bergstrom successfully built and empirically estimated multiple equation models of the economy (e.g., Bergstrom, 1984); this work was considerably informed by economic theory, and his approach has been very influential, having been adopted by a number of empirical macroeconometric researchers internationally. Remarkably, he continued to be productive long into retirement, as illustrated by his final, posthumously-published book (Bergstrom and Nowman, 2007).

Developing statistical inference on continuous-time models using discrete-time data is an ambitious undertaking, as Rex Bergstrom was well aware. Intuitively, this is obvious from the fact that any sequence of discrete observations can be interpolated by uncountably many continuous paths. This leads to the classical "aliasing" phenomenon (most frequently discussed in the context of stationary processes, where the spectral density of the equally-spaced discrete sequence is generated from that of the continuous one by the "folding formula"), and thence to a lack of identifiability. The folding formula produces an upper bound for the continuous-time spectral density (because it is everywhere nonnegative), but uncountably many versions of it are consistent with a given discrete-time spectral density.

However, a parsimonious parametric model for macroeconomic processes is usually necessary due to a shortage of data, especially in a multivariate setting, and such a model can considerably alleviate the identification problem. Typically, the model has been a system of constant-coefficient stochastic differential equations. Then it is possible (see Phillips, 1973) that the parameters in the model can be at least locally identified, though this can be difficult to demonstrate analytically. Unfortunately, though an exact discrete-time model can be written down, and may belong to the familiar autoregressive moving average (ARMA) class, the parameters of the discrete-time model can be extremely complicated functions of the underlying parameters; their analytic form depends, indeed, on whether any roots of the continuous-time autoregressive operator are repeated, and, if so, on their multiplicities. Consequently, interest in simpler, discrete-time "approximations" developed, for example, ones generated by replacing derivatives by differences (see, e.g., Bergstrom, 1966). However, there is an element of arbitrariness in the choice of approximation, and for a given model a continuum of "discrete approximations" can be justified, in which the utilization of each leads to different parameter estimates and possibly different conclusions about economic behavior (see Robinson, 1977a). One solution is to employ the basic discrete ARMA structure that is derived, and to estimate this unconstrainedly, ignoring the restrictions. However, the ARMA typically has more coefficients than the initial continuous-time model, so inefficiency results. As a further pitfall, the linear stochastic differential equations are typically assumed to be driven by white noise, or more generally, in the stationary case, the process has a rational spectral density. This

https://doi.org/10.1017/S0266466608090373 Published online by Cambridge University Press

assumption cannot be taken for granted; hence, specifying the dynamics of the unobservable input process is a hazardous undertaking, and misspecification leads to a misspecified "exact" discrete model.

Nevertheless, a considerable body of work has developed, in large part due to Rex Bergstrom and his students, treating the modeling of systems with or without predetermined variables, developing the linear model to allow for forms of nonlinearity, and extending stationary models to nonstationary ones, to reflect the growing econometric interest in nonlinear and nonstationary dynamics in purely discrete time. The literature is described in the historical review article Bergstrom (1988). In fact, this literature, beginning more than 40 years ago, with its focus on macroeconomic time series, is reflected in considerable recent activity with continuous-time modeling in finance. Here the main difficulties facing Bergstrom, discussed above, are reduced due to the very finely spaced sampling of many financial time series and their considerable length. Nevertheless, high frequency financial data and microstructure noise affect model formulations in a way that introduces considerable analytic complexity and calls for new inferential approaches. Phillips and Yu (2005a, 2005b, 2006) emphasize the magnitude of finite sample bias due to discrete sampling, and its economic implications.

The bulk of research on continuous-time modeling has assumed constancy of parameters across time. On the other hand, time-varying models have over the years become a focus of study within mainstream discrete-time series econometrics and statistics. In large part this is due to a need to describe one or more changes in structure, caused by economic, political, or natural events that have a sudden effect on the economy. There has also been considerable work with parameters that can change over time, possibly smoothly and possibly with respect to each observation point, or perhaps with discontinuous jumps. Here we will discuss an extension of a famous result for time-invariant continuous-time models that was alluded to above: a discrete-time equally-spaced sequence from a continuous-time ARMA process has a certain discrete ARMA representation, as found by Bartlett (1946) and subsequent researchers. We attempt to extend this sort of result to a setting that allows general variation across time in coefficients. We also allow for multivariate data, and additionally consider the possibility of irregularly-spaced sampling. Finally, we discuss issues of modeling statistical inference.

## 2. TIME-VARYING CONTINUOUS-TIME MODEL AND DISCRETE SAMPLING

On the real interval $T = [t_0, .)$, define $X(t)$, $t \in T$, to be a mean-zero, real-valued, $p$-dimensional vector homogeneous random process with orthogonal increments, that is,

$$\begin{aligned} \mathrm{E}(X(t_1) - X(t_2))(X(t_3) - X(t_4))' &= (t_1 - t_2) I_p, \qquad t_1 = t_3 > t_2 = t_4, \\ &= 0, \qquad\qquad t_1 > t_2 \geq t_3 > t_4, \end{aligned} \tag{1}$$

https://doi.org/10.1017/S0266466608090373 Published online by Cambridge University Press

$I_p$ being the $p$-rowed identity matrix and the prime denoting transposition.

Initially, let $A(t)$ and $B(t)$ be, respectively, $q \times q$ and $q \times p$ matrices of real-valued, bounded, measurable functions, and consider the system

$$dY(t) = A(t)Y(t)\,dt + B(t)\,dX(t), \qquad t \in T;$$

$$Y(t_0) = Y_0. \tag{2}$$

Many processes $Y(t)$ that arise, not only in economics but also in the natural sciences and engineering, are so modeled, but often with $A(t)$ and $B(t)$ assumed constant across $t$, as in Rex Bergstrom's work. Note that changing $A(t)$ with $t$ entails time-varying dynamics, whereas changing $B(t)$ with $t$ entails time-varying conditional heteroskedasticity.

It is assumed that observations can be made only at countably many points in $T$, $t_1 < t_2 < \cdots$. Indeed the property (1) may limit the accuracy with which $Y(t)$ can be continuously recorded. We deduce a model for $Y(t_j)$, $j \geq 1$.

Let there exist, for all $t$, a nonsingular fundamental matrix $\Phi(t)$ for $A(t)$ (Coddington and Levinson, 1955, pp. 69–71), satisfying

$$\frac{d}{dt}\Phi(t) = -\Phi(t)A(t), \qquad t > t_0,$$

$$\Phi(t_0) = I_q. \tag{3}$$

Suppose, in addition to (1), that

$$\begin{aligned} \mathrm{E}Y_0 &= 0, \\ \mathrm{E}Y_0 Y_0' &= \Omega < \infty, \\ \mathrm{E}Y_0 (X(t_2) - X(t_1))' &= 0, \qquad t_0 < t_1 \leq t_2. \end{aligned} \tag{4}$$

A solution will be described as unique if it is unique in the mean square sense considered by Friedman (1975, Thm. 1.1). The following theorem demonstrates equivalence in second-order moment properties of continuous-time and discrete-time solutions.

THEOREM. *Let $Y_c(t)$, $t \in T$, be the unique solution of (2). For any given sequence $t_1, t_2, \ldots \in T$, where $t_0 < t_1 < t_2, \ldots$, let $Y_d(t_j)$, $j \geq 1$, be the unique solution of*

$$\begin{aligned} Y(t_j) - \Phi(t_j)^{-1}\Phi(t_{j-1})Y(t_{j-1}) &= Z(t_j), \qquad j \geq 1, \\ Y(t_0) &= Y_0, \end{aligned} \tag{5}$$

*where*

$$\begin{aligned} \mathrm{E}Z(t_j) &= 0, \\ \mathrm{E}Z(t_j)Z(t_j)' &= \Phi(t_j)^{-1}\Sigma_j\Phi(t_j)^{-1\prime}, \end{aligned}$$

https://doi.org/10.1017/S0266466608090373 Published online by Cambridge University Press

$$\mathrm{E}Z(t_j)Z(t_k)' = 0, \qquad j \neq k,$$

$$\Sigma_j = \int_{t_{j-1}}^{t_j} \Phi(t)B(t)B(t)'\Phi(t)'dt. \tag{6}$$

*Then*

$$\mathrm{E}Y_c(t_j)Y_c(t_k)' = \mathrm{E}Y_d(t_j)Y_d(t_k)'$$

*for all* $j, k$.

**Proof.** Premultiplication of (2) by $\Phi(t)$ and use of (3) produce

$$d\left(\Phi(t)Y(t)\right) = \Phi(t)B(t)\,dX(t),$$

whence integration over $(t_{j-1}, t_j)$, $j \geq 1$, produces

$$\Phi(t_j)Y(t_j) - \Phi(t_{j-1})Y(t_{j-1}) = \int_{t_{j-1}}^{t_j} \Phi(t)B(t)\,dX(t),$$

which may be written as (5), with

$$Z(t_j) = \Phi(t_j)^{-1} \int_{t_{j-1}}^{t_j} \Phi(t)B(t)\,dX(t).$$

It is readily verified that (1) implies (6). Now (5) has the unique solution

$$Y_d(t_j) = \Phi(t_j)^{-1}\left(Y_0 + \sum_{i=1}^{j} \Phi(t_i)Z(t_i)\right), \qquad j \geq 1,$$

so that, from (4), (6)

$$\mathrm{E}Y_d(t_j)Y_d(t_k)' = \Phi(t_j)^{-1}\left(\Omega - \Sigma_1 + \cdots + \Sigma_{\min(j,k)}\right)\Phi(t_k)^{-1\prime}.$$

On the other hand, from Soong (1973, p.154) a solution of (2) is

$$Y_c(t) = \Phi(t)^{-1}\left(Y_0 + \int_{t_0}^{t} \Phi(u)B(u)\,dX(u)\right),$$

and this is unique (Friedman, 1975, Thm. 1.1). Thus

$$\begin{aligned}\mathrm{E}Y_c(t_j)Y_c(t_k)' &= \Phi(t_j)^{-1}\left(\Omega + \int_0^{\min(t_j,t_k)} \Phi(u)B(u)B(u)'\Phi(u)'du\right)\Phi(t_k)^{-1\prime} \\ &= \mathrm{E}Y_d(t_j)Y_d(t_k)'.\end{aligned}$$

This completes the proof. ■

This "skip -sampling" result could be extended to cover temporally-aggregated data. It would be more difficult to deal with a situation in which different elements of $Y(t)$ are observed at different time points.

https://doi.org/10.1017/S0266466608090373 Published online by Cambridge University Press

In the usual constant-coefficient case, $A(t) \equiv A$, we have, with $t_0 = 0$, $\Phi(t) = e^{-At}$, where for a square matrix $Z$, $e^Z = 1 + \sum_{j=1}^{\infty} Z^j/j!$ when the series converges. Then we deduce that the AR coefficient matrix $\Phi\left(t_j\right)^{-1}\Phi\left(t_{j-1}\right) = e^{A\left(t_j - t_{j-1}\right)}$. In the regularly-spaced case where $t_j - t_{j-1} = \Delta$, the latter matrix becomes $e^{A\Delta}$, as in Rex Bergstrom's work. With irregular spacing we have a multivariate extension of a result obtained by Robinson (1977b).

More generally, for a scalar absolutely continuous function $g(t)$, suppose $A(t) = Af(t)$, where $A$ is a $p \times p$ constant matrix and $f(t) = (d/dt)g(t)$. Thus $\Phi(t) = e^{-Ag(t)}$, and thence $\Phi(t_j)^{-1}\Phi\left(t_{j-1}\right) = e^{A\{g(t_j) - g(t_{j-1})\}}$, in which we may write $g(t_j) - g(t_{j-1}) = \int_{t_{j-1}}^{t_j} f(u)\,du$. For example, in the periodic case $f(t) = \cos(bt)$, $b \neq 0$, we have $g(t) = \sin(bt)/b$ and thence $\Phi(t_j)^{-1}\Phi\left(t_{j-1}\right) = e^{A\{\sin(bt_j) - \sin(bt_{j-1})\}/b} = e^{2A\sin\{b(t_j - t_{j-1})/2\}\cos\{b(t_j + t_{j-1})/2\}}$. With regular spacing, $t_j - t_{j-1} \equiv \Delta$, this becomes $e^{2A\sin(b\Delta/2)\cos(b(t - \Delta/2))}$ at $t_j = t$. Thus the discrete-time model also has a periodic structure.

We have focused on closed systems, containing no predetermined or exogenous variables. An open system, which contains a vector $U(t)$ of such observable variables, is

$$dY(t) = A(t)Y(t)\,dt + C(t)U(t) + B(t)\,dX(t), \quad t \in T.$$

We can infer from the Theorem proof, by simply substituting $C(t)U(t) + B(t)\,dX(t)$ for $B(t)\,dX(t)$, a model that includes $Y(t)$ only discretely but depends on the continuous $U(t)$ as well as the continuous $X(t)$. However, it is likely that $U(t)$ will only be observed discretely. This mirrors the problem encountered with constant-coefficient models, where a full discrete-time model has been deduced either "exactly," by an implicit assumption about the continuous-time dynamics of $U(t)$, which is inevitably a source of misspecification if incurred, or a discrete "approximation," which again really has the status of a good approximation for suitable continuous-time dynamics (see, e.g., Phillips, 1974; Sargan, 1974; Robinson, 1976; Bergstrom, 1983). Of course an allowance for irregular spacing raises the possibility that $U(t)$ could be observed at different times from $Y(t)$, which would make for further difficulties.

## 3. HIGHER-ORDER EXTENSION

Successive application of the step from (1) to (5) yields a sampled version of the $r^{th}$-order vector continuous-time model

$$dD^{r-1}Y(t) = \sum_{i=1}^{r} A_i(t)D^{i-1}Y(t)\,dt + B(t)\,dX(t),$$

$$Y(t_0) = Y_0, \tag{7}$$

where $D^j = d^j/dt^j$. First write (7) in the form (2) with $X(t)$, $Y(t)$, $A(t)$, and $B(t)$ replaced by

https://doi.org/10.1017/S0266466608090373 Published online by Cambridge University Press

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ X(t) \end{bmatrix}, \begin{bmatrix} Y(t) \\ DY(t) \\ \vdots \\ D^{r-2}Y(t) \\ D^{r-1}Y(t) \end{bmatrix}, \begin{bmatrix} 0 & I_q & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & I_q \\ A_1(t) & A_2(t) & \cdots & \cdots & A_r(t) \end{bmatrix}, \begin{bmatrix} \cdots & 0 \\ \cdots & 0 \\ \cdots & \vdots \\ \cdots & 0 \\ \cdots & B(t) \end{bmatrix},$$

respectively.

For simplicity consider only the familiar case of equally-spaced observations, $t_{j+1} - t_j \equiv 1$. Then a discrete-time model like (5) is obtained, with $t_j$ and $t_{j-1}$ replaced by $t$ and $t-1$. (Although we call it "discrete," it is of course valid continuously over $t$.) Now the first subvector of the latter model has the form

$$Y(t) = \sum_{i=1}^{r} F_i(t) D^{i-1} Y(t-1) + \int_{t-1}^{t} G(t)\, dX(t).$$

If $F_r(t)$ is nonsingular, an expression for $D^{r-1}Y(t-1)$ is obtained. By means of this, $D^{r-1}Y(t-1)$ is eliminated from the remainder of (5), and a new (2) is formed, in which $Y(t)$ is replaced by

$$\left(Y(t)', Y(t-1)', DY(t)', DY(t-1)', \ldots, D^{r-2}Y(t)', D^{r-2}Y(t-1)'\right)'.$$

A new (5) is derived and from the first two subvectors of this, expressions for $D^{r-2}Y(t-1)$, $D^{r-2}Y(t-2)$, deduced, much as before. A further (2) is constructed and so on, until the final (2) has $Y(t)'$ replaced by $\left(Y(t)', \ldots, Y(t-r+1)'\right)$. From the first subvector of the associated (5) model we deduce

$$Y(t) - \sum_{j=1}^{r} R_j(t) Y(t-j) = \sum_{j=0}^{r-1} S_j(t) Z(t-j) \qquad \textbf{(8)}$$

for suitable matrices $R_1(t), \ldots, R_r(t)$, $S_1(t), \ldots, S_{r-1}(t)$, and where for each $t$ elements of the sequence $Z(t), Z(t-1), \ldots$ are uncorrelated but possibly heteroskedastic vectors, and the right side of (8) results from the $r-1$ integrations that have been performed. The fact that (7) and (8) have solutions with identical autocovariance properties may be demonstrated as in the Theorem.

Similar results were earlier established for the special case that coefficients and sampling intervals are constant over time, early references being Bartlett (1946), Walker (1950), and Phillips (1959). Then if the initial conditions are at $-\infty$, and, say, $Y(t)$ is covariance stationary, (5) and (8) become special cases of ARMA models, about which a great deal is known. An open question is the extension of such results to spatial and spatiotemporal data.

## 4. MODELING AND INFERENCE ISSUES

The properties we have developed demonstrate some equivalence between continuous-time models with time-varying coefficients and discrete-time models,

https://doi.org/10.1017/S0266466608090373 Published online by Cambridge University Press

but they fall far short of providing a basis for statistical inference, given discrete-time data.

The approach taken to modeling $A(t)$ and $B(t)$ in (2) is a crucial initial consideration. They could be parametric, in the sense of depending on $t$ and finitely many unknown parameters, or they could be nonparametric, or else a semiparametric setting could be considered; for example $A(t)$ could be parametric but $B(t)$ nonparametric. Then the corresponding discrete model (5) inherits the same kind of properties (though the innovations covariance matrix $\Sigma_j$ depends on both $A(t)$ and $B(t)$). For parametric $A(t)$, however, the extent to which a parametric representation of the autoregressive coefficient matrix $\Phi(t_j)^{-1}\Phi(t_{j-1})$ can be deduced is far from clear. Even in the equally-spaced context, with $t_{j+1} - t_j$ constant, matters are more complicated even than in the time-invariant settings developed by Rex Bergstrom. Rather, a desire for efficiency gains from utilizing the restrictions required by the continuous-time model is liable to give way to a more pragmatic approach to specifying the discrete-time model, either parametrically or nonparametrically, with the continuous-time one essentially playing the role of motivating a loosely-defined discrete-time dynamic structure, via the Theorem. In higher-order models (7), choice of the order $r$ is as always an issue. Notice that $A(t)$ and $B(t)$ could also be regarded as stochastically generated but independent of $\{X(t)\}$ (see, e.g., Nicholls and Quinn 1980), in which case the Theorem is valid conditional on $\{A(t)\}, \{B(t)\}$.

Given a discrete-time specification, a variety of strategies for estimation are possible. In parametric models, these are likely to be based on a form of Gaussian likelihood or pseudo-likelihood, as in the time-invariant models considered by Rex Bergstrom. For the purpose of statistical inference, developing asymptotic distribution theory presents formidable difficulties. The irregular spacing permitted in the Theorem is a severe complication in itself, even for time-invariant models. This is due essentially to the loss of Toeplitz structure in the covariance matrix of the data, making it very difficult to develop regularity conditions that comprehensibly separate out assumptions on the process generating the observation times, be it deterministic or stochastic, and assumptions on $\{A(t)\}$, $\{B(t)\}, \{X(t)\}$. For a scalar, time-invariant version of (2), with irregular spacing, Robinson (1977b) developed Gaussian estimates and asymptotic theory. In a purely discrete-time setting, Dunsmuir (1983) developed asymptotic theory for estimates of time-invariant ARMA models when data are subject to missing observations from a regularly-spaced grid. For time-varying discrete-time models, asymptotic theory will rely inter alia on stability properties, or suitable knowledge of the nature of any instability. Some theoretical properties of time-varying discrete-time models, with regularly-spaced observations, have been developed by, e.g., Hallin (1978), Melard and Herteleer-de-Schutter (1989), and parameter estimation has been considered by, e.g., Subba Rao (1970) and Kitagawa and Gersch (1985). For nonparametric $A(t)$ and $B(t)$, a local kernel smoothing approach can be employed, as in the time-varying semiparametric regression and regularly-spaced locally stationary time series contexts of Robinson (1989) and

https://doi.org/10.1017/S0266466608090373 Published online by Cambridge University Press

Dahlhaus (2000), respectively; here, to enable a useful asymptotic theory, $A(t)$ would be replaced by $A(t/n)$, where $n$ is the number of observations and observations are drawn at $t = 1, \ldots, n$. The above discussion has had in mind the closed systems that are the focus of the paper. With open systems, briefly discussed in Section 2, after making the interpolation/approximation that creates a discrete-time model, there is potential for less dependence on strong assumptions on the input noise $X(t)$, as stressed by Robinson (1976) in the case of constant-coefficient models. Rex Bergstrom's pioneering investigations have not only solved important issues in time series modeling and inference, but opened the way for much further research.

## REFERENCES


Bartlett, M.S. (1946) On the theoretical specification of sampling properties of autocorrelated time series. *Journal of the Royal Statistical Society Supplement* 8, 27–41.

Bergstrom, A.R. (1966) Nonrecursive models as discrete approximations to systems of stochastic differential equations. *Econometrica* 34, 173–182.

Bergstrom, A.R. (1967) *The Construction and Use of Economic Models*. English University Press.

Bergstrom, A.R. (1976) *Statistical Inference in Continuous -Time Economic Models*. North-Holland.

Bergstrom, A.R. (1983) Gaussian estimates and structural parameters in higher-order continuous-time dynamic models. *Econometrica* 51, 117–162.

Bergstrom, A.R. (1984) Monetary fiscal and exchange rate policy in a continuous-time model of the United Kingdom. In P. Malgrange & P. Muel (eds.) *Contemporary Macroeconomic Modelling*, pp. 183–206. Blackwell.

Bergstrom, A.R. (1988) The history of continuous-time econometric models. *Econometric Theory* 4, 365–383.

Bergstrom, A.R. & B. Nowman (2007) *A Continuous Time Econometric Model of the United Kingdom with Stochastic Trends.* Cambridge University Press.

Coddington, E.A. & N. Levinson (1955) *Theory of Ordinary Differential Equations*. McGraw-Hill.

Dahlhaus, R. (2000) A likelihood approximation for locally stationary processes. *The Annals of Statistics* 28, 1762–1794.

Dunsmuir, W.T.M. (1983) A central limit theorem for estimation in Gaussian stationary time series observed at unequally spaced times. *Stochastic Processes and their Applications* 14, 279–295.

Friedman, A. (1975) *Stochastic Differential Equations and Applications.* Academic Press.

Hallin, M. (1978) Mixed autoregressive-moving average multivariate processes with time-dependent coefficients. *Journal of Multivariate Analysis* 8, 567–572.

Kitagawa, G. & W. Gersch (1985) A smoothness priors time-varying AR coefficient modeling of nonstationary covariance time series. *IEEE Transactions on Automatic Control* AC-30, 48–55.

Melard, G. & A. Herteleer-de-Schutter (1989) Contributions to evolutionary spectral theory. *Journal of Time Series Analysis* 10, 41–63.

Nicholls, D. & B.G. Quinn (1980) The estimation of random coefficient autoregressive models I. *Journal of Time Series Analysis* 1, 37–46.

Phillips, A.W. (1959) The estimation of parameters in systems of stochastic differential equations. *Biometrika* 46, 67–76.

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

Phillips, P.C.B. (1974) The estimation of some continuous time models. *Econometrica* 42, 803–823.

Phillips, P.C.B. & J. Yu (2005a) Jackknifing bond option prices. *Review of Financial Studies* 18, 707–742.


https://doi.org/10.1017/S0266466608090373 Published online by Cambridge University Press

Phillips, P.C.B. & J. Yu (2005b) Comments: A selective overview of nonparametric methods in financial econometrics. *Statistical Science* 20, 338–343.

Phillips, P.C.B. & J. Yu (2006) Maximum likelihood and Gaussian estimation of continuous time models in finance. Working paper, Cowles Foundation for Research in Economics.

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

Robinson, P.M. (1977a) The construction and estimation of continuous time models and discrete approximations in econometrics. *Journal of Econometrics* 6, 173–197.

Robinson, P.M. (1977b) Estimation of a time series model from unequally spaced data. *Stochastic Processes and their Applications* 6, 9–24.

Robinson, P.M. (1989) Nonparametric estimation of time-varying parameters. In P. Hackl (ed.), *Statistical Analysis and Forecasting of Economic Structural Change* pp. 253–264. Springer-Verlag.

Sargan, J.D. (1974) Some discrete approximations to continuous time stochastic models. *Journal of the Royal Statistical Society* B 36, 74–90.

Soong, T.T. (1973) *Random Differential Equations in Science and Engineering*. Academic Press.

Subba Rao, T. (1970) The fitting of non-stationary time-series models with time-dependent parameters. *Journal of the Royal Society* B 32, 312–322.

Walker, A.M. (1950) Note on a generalization of the large sample goodness of fit test for linear autoregressive schemes. *Journal of the Royal Statistical Society* B 12, 102–107.

https://doi.org/10.1017/S0266466608090373 Published online by Cambridge University Press