Econometric Theory, **25**, 2009, 1120–1137. Printed in the United States of America.
doi:10.1017/S0266466608090452

ESTIMATING CONTINUOUS-TIME MODELS ON THE BASIS OF DISCRETE DATA VIA AN EXACT DISCRETE ANALOG

J. RODERICK MCCRORIE
University of St. Andrews
and
CORE, Université catholique de Louvain

This paper offers a perspective on A.R. Bergstrom's contribution to continuous-time modeling, focusing on his preferred method of estimating the parameters of a structural continuous-time model using an exact discrete-time analog. Some inherent difficulties in this approach are discussed, which help to explain why, in spite of his prescience, the methods around his time were not universally adopted as he had hoped. Even so, it is argued that Bergstrom's contribution and legacy is secure and retains some relevance today for the analysis of macroeconomic and financial time series.

1. INTRODUCTION

Probably more than any other econometrician of his generation, Rex Bergstrom presaged the growth and development of the use of continuous-time models in econometrics. While concern is today mostly focused on single equation or small system nonlinear models in finance, his own work to a large extent was directed toward the estimation of large-scale macroeconometric systems, and proceeded in parallel with developments in linear-in-the-variables structural time series modeling. The continuous-time paradigm is often motivated as a basis for approximating processes that generate data—and no more has this been so than recently with high-frequency financial data. Rex's primary concern, however, was to use continuous-time models as the basis for taking account of the interaction among economic variables during the unit observation interval, whether implied by economic theory or other a priori information about the causal ordering of economic variables. Recognizing explicitly that economic data are generated in finer time intervals than the observation interval and realizing that naïvely-specified discrete-time models can suffer from a lack of time invariance,[1] he felt that only through a continuous-time model could accurate restrictions be placed on the

I thank Peter Phillips (the editor), three anonymous referees, and participants at the A.R. Bergstrom Memorial Conference held at the University of Essex, 24–25 May 2006, for comments on an earlier version of the paper. Address correspondence to J.R. McCrorie, School of Economics and Finance, University of St Andrews, Castlecliffe, The Scores, St Andrews KY16 9AL, U.K.; e-mail: mccrorie@st-andrews.ac.uk.

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

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

probability distribution of observed discrete data and estimation be properly directed toward the parameters of agents' objective functions. He believed, given the moderately sized macroeconomic data sets that were available, that these restrictions were important as a means of obtaining efficient estimates of the underlying structural parameters.

Rex's preferred method was to formulate a continuous-time model such that discrete data generated by the model satisfy an *exact* discrete time analog, called the exact discrete model. Given today's emphasis on financial time series, formulating what amounts to a model that has a closed-form transition density would be unduly restrictive, but at the time, Rex's recursive approach to modeling was innovative and forward-looking compared with the prevailing, nonrecursive method of simultaneous equations. While Bergstrom (1966) pointed toward the use of the exact discrete model as the basis of estimating the parameters of the continuous-time model, it was Phillips (1972) who proposed and demonstrated the feasibility of the method and illustrated the gains in efficiency in taking account of the exact restrictions implied by the continuous-time model. Subsequent developments for first-order models (Phillips, 1973, 1974, 1978; Sargan, 1974; Wymer, 1972) and for higher-order models (Bergstrom, 1983, 1985, 1986, 1997), and developments in computer technology, led to this method becoming his primary approach to estimation and, in particular, formed the basis of estimating the macroeconometric models (Bergstrom and Wymer, 1976; Bergstrom, Nowman, and Wymer, 1992 (BNW); and Bergstrom and Nowman, 2007 (BN)).[2,3]

The purpose of this paper is to offer a perspective on Bergstrom's contribution to the development of estimating continuous-time models on the basis of discrete data, focusing on two aspects of his work that have perhaps been underemphasized. We offer reasons why, in spite of Rex's prescience, the proposed continuous-time methods were not routinely adopted around the time they were conceived. And we discuss implications of his work for the burgeoning continuous-time literature, notably in empirical finance, that Rex began to see in the last decade of his life. Our focus will be on estimating continuous-time models based on an exact discrete model and what were ultimately found to be weaknesses and lacunae in both method and scope against the demands of the profession as they unfolded and evolved. Section 2 contains a brief outline of the method, Section 3 discusses some of the difficulties, and Section 4 considers wider aspects of Bergstrom's contribution. Section 5 concludes.

2. CONTINUOUS-TIME MODELS ESTIMATED USING AN EXACT DISCRETE MODEL

Most continuous-time models used today for financial time series are written in terms of a nonlinear stochastic differential equation (see Aït-Sahalia, 2007). The problem, should one wish to estimate the model via a likelihood function derived from the transition probability density of the discretely sampled data, is that only in exceptionally few cases can the transition density and log-likelihood

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

function be written in closed form. This restricted class includes models based on continuous-time vector autoregressive (VAR) models (including the Ornstein-Uhlenbeck process or Vasicek model); geometric Brownian motion; the square root (or Cox-Ingersoll-Ross) model; and the inverse square root model (see Phillips and Yu, 2009, for a general discussion). Contributing to an era where linear-in-the-variables models were predominant, Rex and other naturally used the framework of continuous-time vector autoregression, and for illustrative purposes we consider the closed first-order model with stock variables[4]

$$Dx(t) = A(\theta)x(t)\,dt + \zeta(dt), \tag{1}$$

subject to the initial condition

$$x(0) = y_0, \tag{2}$$

where D is the mean-square differential operator, $\{x(t), t > 0\}$ is an n-dimensional continuous-time stochastic process of finite variance, A is an $n \times n$ matrix whose elements are known functions of a p-dimensional vector θ of unknown parameters ($p \leq n^2$), y_0 is a non-random $n \times 1$ vector, and $\zeta(dt)$ is heuristically a white noise disturbance vector with covariance matrix $\Sigma(\mu)\,dt$, the elements of Σ being known functions of a q-dimensional vector μ of unknown parameters ($q \leq n(n+1)/2$).[5]

The exact discrete model is obtained from the solution of the continuous-time model (1) subject to (2): as shown by Bergstrom (1984, Thm. 3, p. 1167) among others, a sequence of equidistant discrete data $x(0), x(1), \ldots, x(T)$ generated by (1) satisfies the model

$$x(t) = F(\theta)\,x(t-1) + \varepsilon_t \qquad (t = 1, \ldots, T), \tag{3}$$

where

$$F(\theta) = \exp(A(\theta)) = \textstyle\sum_{i=0}^{\infty} A(\theta)^i / i', \tag{4}$$

$$\mathrm{E}(\varepsilon_t) = 0, \tag{5}$$

$$\mathrm{E}(\varepsilon_t \varepsilon_t') = \int_0^1 \exp(rA(\theta))\,\Sigma(\mu)\exp(rA(\theta))'\,dr = \Omega(\theta, \mu), \tag{6}$$

$$\mathrm{E}(\varepsilon_t \varepsilon_s') = 0 \qquad (s \neq t). \tag{7}$$

The exact discrete model (3)–(7) differs from conventional discrete-time *VAR* models in that the coefficient matrix and the covariance matrix are functions of the exponential of a matrix. This leads to several problems that we now briefly discuss.

3. PROBLEMS WITH THE APPROACH

3.1. The Aliasing Problem

While not a problem that Rex addressed in his career, at least in print, the aliasing problem is fundamental to continuous-time modeling for the reason that

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

identification is prior to estimation. The problem is multivariate in character and relates to the fact that there are, in principle, many different matrices that share the same exponential F in (3) and (4). These matrices are aliases of A in the sense that through taking the place of A in (1), they could generate the same equidistant discrete data. The identification problem is therefore more severe for continuous-time models compared with discrete-time models but, on the other hand, discrete-time models suffer from a lack of time invariance. If Gaussianity is assumed, the general problem is to find a necessary and sufficient condition such that the pair (A, Σ) is identifiable in (F, Ω). In any particular application, of course, the matrices depend on the parameter vectors θ and μ and it could be that fewer or further restrictions are required.

McCrorie (2003) offered a framework for the problem by considering the following matrix M that allows the pair (A, Σ) to be treated together: if

$$M = \begin{bmatrix} -A & \Sigma \\ 0 & A' \end{bmatrix}, \tag{8}$$

then, as an application of Van Loan (1978, Thm. 1),

$$\exp(M) = \begin{bmatrix} F^{-1} & F^{-1}\Omega \\ 0 & F' \end{bmatrix}. \tag{9}$$

The following theorem, which is a consequence of Theorem 2 of Culver (1966), contains the basic result on identification in terms of when the matrix exponential mapping is bijective in general.[6]

THEOREM 1 (McCrorie, 2003). (A, Σ) *is identifiable in* (F, Ω) *if the eigenvalues of* M *are strictly real and no Jordan block of* M *belonging to any eigenvalue appears more than once.*

Note that the eigenvalues of M are simply the eigenvalues and reverse eigenvalues of A, and so if A has no complex eigenvalues and there is no confluence in its eigenvalues, the aliasing problem reduces essentially to a univariate problem involving the exponential function that, when viewed as real-valued, is bijective. Unfortunately, both restrictions are inapt for economic time series: they rule out plausible cyclical behavior resulting from complex eigenvalues and plausible trend behavior resulting from multiple unit roots (multiple zero eigenvalues of A). In the complex eigenvalue case, several authors achieve identification through additional restrictions—Phillips (1973) uses Cowles Commission type restrictions; Hansen and Sargent (1983) restrictions inherent in the requirement that Ω is positive semidefinite; and Hansen and Sargent (1991) cross-equation restrictions implied by the rational expectations hypothesis.[7] Bergstrom et al. (1992) and Bergstrom and Nowman (2007) use prior bounds on the parameters as a means of achieving identification.[8] The results by Hansen and Sargent (1983) show that without importing restrictions from beyond the problem in hand, identification can only be local. The problem of finding a necessary and sufficient condition,

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

however, remains open even in the most basic model. It is not surprising, therefore, when Rex was starting to conceive of the use of continuous-time models, at the same time as identification issues were to the fore in the simultaneous equations methodology, that the aliasing problem presented an impediment to their use. It was then much easier computationally to obtain estimates from discrete-time models, and problems in interpreting estimates that arise from their lack of time invariance were largely ignored.

From today's perspective, the aliasing problem is not as important as it once was because the continuous-time models being used in empirical finance tend to be small system models that are nonlinear in the parameters. The structure of these models often necessitates using a discretization based on an *approximation* of its solution, and so estimation proceeds on the basis of a time-varying but identified discrete-time model whose estimates are interpreted against the parameters of the continuous-time model (see Aït-Sahalia, 2007; Phillips and Yu, 2009). An example illustrating the identification problem is given in the Appendix.

3.2. The Controllability Problem

Erickson (1971) noted that Ω is positive semidefinite but not necessarily positive definite, and Phillips (1974) among others showed that if

$$H(A, \Sigma) = [\, \Sigma \quad A\Sigma \quad A^2\Sigma \quad \dots \quad A^{n-1}\Sigma \,], \tag{10}$$

then Ω is positive definite if and only if rank $H(A, \Sigma) = n$. This is equivalent to solving the controllability problem in linear systems theory.[9] In the case that Ω is not positive definite, it is possible to find a nonsingular matrix to transform the system into a canonical form where a subsystem is completely controllable (see Kwakernaak and Sivan, 1972, Sect. 1.6). This aspect is untreated in the time domain methods proposed by Bergstrom (1983) and Harvey and Stock (1985), in the representation theory of Chambers (1999) and McCrorie (2000), and in the formulation of the BNW and BN macroeconometric models.

3.3. Multimodal Likelihood Functions and the Sampling Properties of Gaussian Estimators

When the exact discrete model is used as the basis of estimation, the objective function may be multimodal to the extent that it may be difficult numerically to find its global optimum (Robinson, 1980). Belcher, Hampton, and Tunnicliffe (1994) have argued that this is a prima facie reason not to use a structural continuous-time autoregressive model and its exact discrete model. Results from the BNW and BN models demonstrate that such models and estimation procedures *can* produce plausible parameter estimates and offer a satisfactory post-sample forecasting performance. Nevertheless, the multimodality problem remains a severe drawback in approaches based on an exact discrete model, even if it did not manifest itself at the time Rex began working on continuous-time methods, when computer technology was still in its infancy.

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

The problem of finite sample estimation bias has the potential to beset *all* maximum likelihood (ML) methods, including those based on the exact discrete model, and is particularly relevant when estimating the mean reversion parameter in the common interest rate models (see Phillips and Yu, 2009).[10] In a sampling experiment, Phillips and Yu show that this bias, which can be substantial in samples as large as 1,000, is more important than the discretization bias arising from estimation being based on only an approximation to the solution of the continuous-time model. Phillips and Yu (2005, 2009) successfully propose jackknife techniques and a simulation-based method based on the method of indirect inference as bias reduction methods. The latter method, although computationally more costly, was otherwise superior and retained the good asymptotic properties of ML estimation.

Finally, the asymptotic sampling properties of estimators from continuous-time models are technically more difficult to derive than for discrete-time models, and they have not yet been formally established for either the BNW or BN models. The emergence of excess differencing in the exact discrete model of the latter (Bergstrom, 1997) is discussed by Bergstrom (2009) in a posthumous contribution to this volume, and is explained by Chambers (2006, 2009) in work that represents an important step toward establishing the asymptotic properties of Gaussian estimators.

3.4. Higher-Order Models

The problems involved in estimating higher-order models are more complicated and involve the problem of defining observable vectors with mixed stock and flow data; the need for a rank condition to eliminate unobservable variables; and the difficulty in expressing this rank condition in a way that traces restrictions back to the original coefficient matrix and not just to its exponential. These issues are discussed by Harvey and Stock (1985, 1993), Chambers (1999), McCrorie (2000), and McGarry (2003). The frequency domain methods described by Robinson (1993) are much more amenable to higher-order models, but they have yet to be used in applied work. Unlike many discrete-time approaches, the continuous-time methods developed for higher-order models incorporate the aspect that data are observed as both stocks and flows. The cost, however, is to introduce other minor problems and lacunae that entail other difficulties that need to be overcome when formulating and estimating continuous-time models via an exact discrete model.

4. WIDER ASPECTS OF BERGSTROM'S CONTRIBUTION

Rex will be remembered as a pioneer: he was the first to apply the method of limited information maximum likelihood; with Basmann he opened up the field of finite-sample theory in econometrics; he contributed to the formulation and estimation of cyclical growth models; and, of course, he will be remembered for his

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

contribution to continuous-time methods, both as a proponent and in the context of large-scale model building (see the paper by Nowman, 2009, in this volume).

One aspect that has perhaps been neglected is his motivation to pursue the development of continuous-time methods because of what he felt were deficiencies in discrete-time modeling. This was not just in terms of a question of whether to use a recursive or nonrecursive model, but also in terms of the lack of time invariance of discrete-time models. Estimates of parameters of naïvely specified discrete-time models, in consequence, are only true with respect to the particular rate at which the data are sampled and therefore do not readily admit an economic interpretation in the absence of a treatment of temporal aggregation bias. The implications of this are still to feed through fully to applied work. That being said, the recent availability of ultra-high frequency data has changed the nature of the problems the econometrician has to face. Rex foresaw and indeed exploited the developments in computer technology, but he did not imagine the astounding advances we have seen in data collection, particularly in financial time series, that have led in some cases to data today being available on a tick-by-tick basis. This makes fitting nonparametric models directly to data all the more practical (see, e.g., Aït-Sahalia, 1996; Bandi and Phillips, 2003) and has moved the continuous-time econometric literature away from the strongly parametric, linear, multivariate approach that Rex adopted for the macreoconometric models. It has also led to research problems of a different nature, such as how optimally to sample continuous-time processes in the presence of microstructure noise.[11] There are, however, potentially some insights to be gained from his way of thinking. High-frequency data offer the possibility of comparing estimates of models based on different sampling schemes, and the efficacy of a discretization method could be measured against a rival method via the sensitivity of its parameter estimates to temporal aggregation bias. The possibility not available to Rex of randomly sampling economic processes is now feasible, not just for observed data that are generated in random intervals in time, which motivated the study by Aït-Sahalia and Mykland (2003), but as a potential solution to the aliasing problem, especially for models formulated for and estimated using low frequency data.[12] The status of such models, however, must now be open to question; for when viewed against the data that are now available, an approach that seeks to identify the parameters of one fixed specification neglects the gains that could be made from sampling at significantly higher frequencies in allowing for more structure to be uncovered. It may be pragmatic to consider radically different types of models for the same time series sampled at different rates, and to an extent this already happens. In the exchange rate literature, for example, we see that accounting for microstructure noise is important for models tested against high-frequency data (Lyons, 2001; Evans and Lyons, 2002), yet at the same time Engel, Mark, and West (2007) show that traditional models based on an asset market approach, when augmented by natural assumptions about exogenous variables and the rate of time preference of economic agents, can be corroborated by low frequency data of the same series. There is also the question of the role of economic theory

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

in statistical models and Rex's early vision of relating parameter estimates to the deep parameters of economic agents' behavior being superseded today by concerns relating to microstructure noise. The results by Aït-Sahalia, Mykland, and Zhang (2005) show that, although the presence of microstructure noise makes it optimal to sample less often than otherwise would be the case, modeling the noise and using all the data is preferable to first order, even if the noise distribution is misspecified. It could be that, costs permitting, the move toward high-frequency data considerations becomes inexorable should the data be available. Be that as it may, Rex will always be a significant part in the history of econometric model building and the school of thought that attempts to explain data through the imposition of restrictions from economic theory.

The macroeconometric literature still retains a strong interest in low-frequency data modeling, primarily because of the preponderance of trends in economic time series observable at such frequencies, and the relationships observed among them. Additionally, high-frequency macro time series data are often expensive to obtain. One aspect that is often overlooked is the invariance of the property of co-integration to temporal aggregation (see Stock, 1987, and especially Phillips, 1991). This implies that long-run information could be embedded in a *continuous-time model* and for a given discrete-time model the property of co-integration should manifest itself in the same way if economic variables are sampled equidistantly, regardless of the rate at which they are sampled. Franchi (2007) has recently characterized the sense in which we can "count" the number of unit roots in a system. The results by Phillips (1991) imply that this could be achieved equivalently by characterizing the singularities in a stochastic differential system. This approach, currently being pursued by McCrorie (2009), has the advantage that it works with the full system (together with a generalized inverse), and so offers the promise of a uniform approach. Inferentially speaking, the *essential* properties of econometric models are given by their singularities, and one connects to a discrete-time model through the solution of a model formulated in continuous time, just as envisioned by Rex and his former student Peter C.B. Phillips (see Bergstrom, 1984; Phillips, 1972, 1991; Yu and Phillips, 2001). In any case, the principle of working with the *solution* of the continuous-time model offers a benchmark for continuous-time estimation methods, and is an aspect of model building to which Rex and Phillips will be historically associated, even if today, in practice, the nature of the formulated model means that its solution needs to be approximated.

5. CONCLUDING COMMENTS

In the last decade of his life, Rex Bergstrom finally witnessed a burgeoning literature on methods of estimating parameters of continuous-time models on the basis of discrete data, as he had foreseen more than 40 years earlier. The methods he saw, however, did not entirely embrace those he had developed. Rex's thinking on continuous time had been developed in an era of linear-time-series modeling, with

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

his own focus toward estimation and testing of the parameters of agents' objective functions in the context of a large-scale dynamic macroeconometric model. His efforts were therefore directed, by necessity, toward confronting some of the problems that have been discussed in this article. The methods he proposed were never "off the shelf" in a way some applied workers like: they required both the conceptual development of a model based on causal restrictions and computing power that was at the time on the edge of technology. Estimation of a model as late as Bergstrom et al. (1992) required the use of a CRAY X-MP/48 supercomputer, even if today estimation of such a model would be standard. We have pointed to some problems in econometrics today that might benefit from the continuous-time perspective that he and some others, notably Peter C.B. Phillips, have offered. Rex will, of course, be remembered in this area for his progressive and visionary role in the development of the subject and for his indefatigable promotion of continuous-time methods over four decades. The book by Bergstrom and Nowman (2007) allows for a retrospective view of what was actually achieved.

The focus above has been on Rex's research contribution, but it should never be forgotten that Rex was also an outstanding teacher. Mention has been made by others, for example Phillips (2005) and Nowman (2009), of his characteristic appearance in a charcoal black suit—indeed there were rumors he had a wardrobe collection of them—but in spite of his diminutive size, he had tremendous presence. Many students at the University of Essex will remember entering his room to be greeted by Rex engulfed by his black executive chair and backed by a library that was like a wall of learning, yet at all times it was his personality, his enthusiasm, and, occasionally, his raucous laugh that filled the room. I was fortunate to take lectures in his last M.Sc. course before his retirement. He had moved on from the series of lectures based on Malinvaud's *Statistical Methods of Econometrics* described by Phillips (1988) to give a course on time series based on Rozanov's *Stationary Random Processes*; Anderson's *The Statistical Analysis of Time Series*; Hannan and Deistler's *The Statistical Theory of Linear Systems*; and his own work (Bergstrom 1984, 1990) on continuous-time modeling. This course has been considerably helpful to a number of people currently contributing to this volume. He began with the immortal line; "I'll assume you're all familiar with the notion of the Lebesgue integral," and moved immediately to a discussion of random measure. The lectures were characterized by their formality and total rigor, and were delivered without notes in blocks of one-and-a-half hours without a break. He was Rex *tremendae majestatis*.

NOTES

1. McCrorie and Chambers (2006) illustrate the consequences of the lack of time invariance of naïvely-specified discrete-time models in testing for money-income causality and discuss related literature.

2. This aspect in the development of methods of estimating continuous-time models was one of the topics outside of the scope of the recent, important survey by Aït-Sahalia (2007). Phillips (1972)

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

was based on his M.A. dissertation supervised by Rex at the University of Auckland in 1969. See Bergstrom (1988) and Nowman (2008) for further discussion.

3. Other contributions to the method of using an exact discrete model include Sims (1971); Geweke (1978); Hansen and Sargent (1983, 1991); Marcet (1991); Harvey and Stock (1985, 1988, 1989, 1993); Zadrozny (1988); Robinson (1976, 1977, 1993, 2009); Phillips (1991); Simos (1996); Nowman (1997); Chambers (1999, 2003, 2004, 2009); McCrorie (2000); Yu and Phillips (2001); Chen and Zadrozny (2001); McGarry (2003); Campbell, Chacko, Rodriguez, and Viceira (2004); McCrorie and Chambers (2006); and Chambers and McCrorie (2006, 2007).

4. We shall not consider open systems, which throw up their own difficulties owing to the need to interpolate the exogenous variables. See, e.g., Bergstrom (1986), Robinson (1992), and McCrorie (2001). We also suppress a constant term that can be treated exactly in the context of an open system.

5. The stochastic differential equation system (1) is expressed in a way that is, in fact, individual to Bergstrom and reflects his own solution to the problem of defining white noise in a way that explicitly recognizes that there is no wide sense stationary process whose integrals over every pair of disjoint intervals are uncorrelated. Stochastic differential equations were not commonplace then as they are today and, motivated to work under an assumption weaker than the disturbance being Brownian motion, Rex constructed his own L^2-theory that used orthogonal increments based on the concept of random measure. The theory was provided by Bergstrom (1983, 1984), and McCrorie (2000, 2002) offers some additional results and comment.

6. See also Kessler and Rahbek (2004).

7. Other examples of cross-equation restrictions being used to identify continuous-time models include Christiano (2005), Christiano and Eichenbaum (1987), and Heaton (1993). Hansen and Scheinkman (1985) and Hansen, Scheinkman, and Touzi (1998) discuss the identification of nonlinear continuous-time Markov processes from discrete data.

8. Identification of the macroeconometric models was not, however, formally established and remains open. This aspect does not *appear* to have interfered with parameter estimation: in the BN model, for example, there were only four complex pairs of eigenvalues, two of which failed (and only just failed) the stationarity condition of having negative real parts.

9. A time-invariant linear system is controllable if the state of the system can be transferred from an initial value to an arbitrary value within a finite time. See the reference cited in the main text.

10. The source of the bias is exactly the same as the estimation bias associated with local-to-unity discrete-time autoregressive models. See also Yu (2008).

11. See Aït-Sahalia et al. (2005) and Zhang, Mykland, and Aït-Sahalia (2005).

12. See Robinson (1980), Lii and Masry (1992, 1994), and Duffie and Glynn (2004) for a discussion of random sampling with parameter identification in mind. Alongside the work by Aït-Sahalia and Mykland (2003), there is a simulation experiment by Jordà and Marcellino (2003) that showed that if data are generated in short, random intervals of time but are sampled equidistantly, the stylized facts of financial time series—nonnormality, conditional heteroskedasticity, and leptokurtosis—can all be generated purely as an artifact of the sampling scheme. It is therefore important in high-frequency data contexts, just as it was for Rex and others dealing with low-frequency macroeconomic data, that there is an awareness of the potential of the sampling scheme to distort the true data generation process.

13. Strictly, Σ_0 should be parametrized not just as a symmetric matrix but also as a positive definite matrix. This could be achieved, in principle, using the fact that it must have a Hermitian logarithm (see Zwanziger, 1964), although we do not require this result in the illustrations that follow.

14. The model (A.3)–(A.4) with a boundary condition is validly defined in terms of the mean-square differential operator as a special case of (1) and (2), but in the empirical finance literature it is usual to use the (stronger) assumption of Gaussianity to write the stochastic differential equation in terms of the ordinary differential operator, with solution a random process whose sample functions are almost all continuous.

15. This result is often attributed to Hille (1958) but is in fact due to Morinaga and Nono (1954). See Bourgeois (2007) for further comment.

16. I am grateful to a referee for this suggestion.

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

REFERENCES

Aït-Sahalia, Y. (1996) Nonparametric pricing of interest rate derivative securities. *Econometrica* 64, 527–560.

Aït-Sahalia, Y. (2007) Estimating continuous-time models with discretely sampled data. In R. Blundell, T. Persson, & W. Newey (eds.), *Advances in Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, vol. 3, pp. 261–327. Cambridge University Press.

Aït-Sahalia, Y. & P.A. Mykland (2003) The effects of random sampling and discrete sampling when estimating continuous-time diffusions. *Econometrica* 71, 483–549.

Aït-Sahalia, Y., P.A. Mykland, & L. Zhang (2005) How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies* 18, 351–416.

Babbs, S.H. & K.B. Nowman (1999) Kalman filtering of generalized Vasicek term structure models. *Journal of Financial and Quantitative Analysis* 34, 115–130.

Bandi, F.M. & P.C.B. Phillips (2003) Fully nonparametric estimation of scalar diffusion models. *Econometrica* 71, 241–283.

Belcher, J., J.S. Hampton, & G. Tunnicliffe Wilson (1994) Parametrization of continuous time autoregressive models for irregularly sampled time series data. *Journal of the Royal Statistical Society, Series B* 56, 141–155.

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

Bergstrom, A.R. (1983) Gaussian estimation of structural parameters in higher-order continuous time dynamic models. *Econometrica* 51, 117–152.

Bergstrom, A.R. (1984) Continuous time stochastic models and issues of aggregation over time. In Z. Griliches & M.D. Intriligator (eds.), *Handbook of Econometrics*, vol. 2, pp. 1145–1221. North-Holland.

Bergstrom, A.R. (1985) The estimation of parameters in nonstationary higher-order continuous time dynamic models. *Econometric Theory* 1, 369–385.

Bergstrom, A.R. (1986) The estimation of open higher-order continuous time dynamic models with mixed stock and flow data. *Econometric Theory* 2, 350–373.

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

Bergstrom, A.R. (1990) *Continuous Time Econometric Modelling*. Oxford University Press.

Bergstrom, A.R. (1997) Gaussian estimation of mixed order continuous time dynamic models with unobservable stochastic trends from mixed stock and flow data. *Econometric Theory* 13, 467–505.

Bergstrom, A.R. (2009) The effects of differencing on the Gaussian likelihood of models with unobservable stochastic trends: A simple example. *Econometric Theory* 25 (this issue), 903–913.

Bergstrom, A.R. & K.B. Nowman (1999) Gaussian estimation of a two-factor continuous time model of the short-term interest rate. *Economic Notes* 28, 25–41.

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

Bergstrom, A.R., K.B. Nowman, & C.R. Wymer (1992) Gaussian estimation of a second order continuous time macroeconometric model of the United Kingdom. *Economic Modelling* 9, 313–351.

Bergstrom, A.R. & C.R. Wymer (1976) A model of disequilibrium neoclassical growth and its application to the United Kingdom. In A.R. Bergstrom (ed.), *Statistical Inference in Continuous Time Economic Models*, pp. 267–327. North-Holland.

Bernstein, D.S. & W. So (1993) Some explicit formulas for the matrix exponential. *IEEE Transactions on Automatic Control* 38, 1228–1232.

Bourgeois, G. (2007) On commuting exponentials in low dimensions. *Linear Algebra and its Applications* 423, 277–286.

Campbell, J.Y., G. Chacko, J. Rodriguez, & L.M. Viceira (2004) Strategic asset allocation in a continuous-time VAR model. *Journal of Economic Dynamics and Control* 28, 2195–2214.

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

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

Chambers, M.J. (2003) The asymptotic efficiency of cointegration estimators under temporal aggregation. *Econometric Theory* 19, 49–77.

Chambers, M.J. (2004) Testing for unit roots with flow data and varying sampling frequency. *Journal of Econometrics* 119, 1–18. (Corrigendum: *Journal of Econometrics* 144 (2008), 524–525)

Chambers, M.J. (2006) On Excess Differencing in Discrete Time Representations of Cointegrated Continuous Time Models with Mixed Sample. Mimeo, University of Essex.

Chambers, M.J. (2009) Discrete time representations of cointegrated continuous time models with mixed sample data. *Econometric Theory* 25 (this issue), 1030–1049.

Chambers, M.J. & J.R. McCrorie (2006) Identification and estimation of exchange rate models with unobservable fundamentals. *International Economic Review* 47, 573–582.

Chambers, M.J. & J.R. McCrorie (2007) Frequency domain estimation of temporally aggregated Gaussian cointegrated systems. *Journal of Econometrics* 136, 1–29.

Chen, B. & P. Zadrozny (2001) Analytic derivatives for the matrix exponential for estimation of linear continuous-time models. *Journal of Economic Dynamics and Control* 25, 1867–1879.

Chen, R. & L. Scott (1992) Pricing interest rate options in a two-factor Cox-Ingersoll-Ross model of the term structure. *Review of Financial Studies* 5, 613–636.

Christiano, L.J. (2005) A method for estimating the timing interval in a linear econometric model with an application to Taylor's model of staggered contracts. *Journal of Economic Dynamics and Control* 9, 363–404.

Christiano, L.J. & M. Eichenbaum (1987) Temporal aggregation and structural inference in macroeconomics. *Carnegie-Rochester Conference Series on Public Policy* 26, 63–130.

Culver, W.J. (1966) On the existence and uniqueness of the real logarithm of a matrix. *Proceedings of the American Mathematical Society* 17, 1146–1151.

Duffie, D. & P. Glynn (2004) Estimation of continuous-time Markov processes sampled at random time intervals. *Econometrica* 7, 1773–1808.

Engel, C, N.C. Mark, & K.D. West (2007) Exchange models are not as bad as you think. *NBER Macroeconomics Annual* 22, 381–441.

Ercolani, J.S. (2009) Cyclical trends in continuous time models. *Econometric Theory* 25 (this issue), 1112–1119.

Erickson, R.V. (1971) Constant coefficient linear differential equations driven by white noise. *Annals of Mathematical Statistics* 42, 820–823.

Evans, M.D. & R.K. Lyons (2002) Order flow and exchange rate dynamics. *Journal of Political Economy* 110, 170–180.

Franchi, M. (2007) The integration order of vector autoregressive processes. *Econometric Theory* 23, 546–553.

Geweke, J.B. (1978) Temporal aggregation in the multivariate regression model. *Econometrica* 46, 643–662.

Hansen, L.P. & T.J. Sargent (1981) The Dimensionality of the Aliasing Problem in Models with Rational Spectral Densities. Research Department Staff Report 72, Federal Reserve Bank of Minneapolis.

Hansen L.P. & T.J. Sargent (1983) The dimensionality of the aliasing problem in models with rational spectral densities. *Econometrica* 51, 377–387.

Hansen, L.P. & T.J. Sargent (1991) Identification of continuous time rational expectations from discrete data. In L.P. Hansen and T.J. Sargent (eds.), *Rational Expectations Econometrics*, pp. 219–235. Westview Press.

Hansen, L.P. & J.A. Scheinkman (1995) Back to the future: generating moment implications for continuous-time Markov processes. *Econometrica* 63, 767–804.

Hansen, L.P., J.A. Scheinkman, & N. Touzi (1998) Spectral methods for identifying scalar diffusions. *Journal of Econometrics* 86, 1–32.

Harvey, A.C. & J.H. Stock (1985) The estimation of higher-order continuous time autoregressive models. *Econometric Theory* 1, 97–117.

Harvey A.C. & J.H. Stock (1988) Continuous time autoregressive models with common stochastic trends. *Journal of Economic Dynamics and Control* 12, 365–384.

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

Harvey, A.C. & J.H. Stock (1989) Estimating integrated higher-order continuous time autoregressions with an application to money-income causality. *Journal of Econometrics* 42, 319–336.

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

Heaton, J. (1983) The interaction between time-nonseparable preferences and time aggregation. *Econometrica* 61, 353–386.

Hille, E. (1958) On roots and logarithms of elements of a complex Banach algebra. *Mathematische Annalen* 136, 46–57.

Jordà, O. & M. Marcellino (2003) Modeling high-frequency foreign exchange data dynamics. *Macroeconomic Dynam*ics 7, 618–635.

Kessler M. & A. Rahbek (2004) Identification and inference for multivariate cointegrated and ergodic Gaussian diffusions. *Statistical Inference for Stochastic Processes* 7, 137–151.

Kwakernaak, H. & R. Sivan (1972) *Linear Optimal Control Systems*. Wiley.

Lii, K.-S. & E. Masry (1992) Model fitting for continuous-time processes from discrete time data. *Journal of Multivariate Analysis* 30, 56–79.

Lii, K.-S. & E. Masry (1994) Spectral estimation of continuous-time stationary processes from random sampling. *Stochastic Processes and their Applications* 52, 39–64.

Lyons, R. (2001) *The Microstructure Approach to Exchange Rates*. MIT Press.

Marcet, A. (1991) Temporal aggregation of economic time series. In L.P. Hansen and T.J. Sargent (eds.), *Rational Expectations Econometrics*, pp. 237–282. Westview Press.

McCrorie, J.R. (2000) Deriving the exact discrete analog of a continuous time system. *Econometric Theory* 16, 998–1015.

McCrorie, J.R. (2001) Interpolating exogenous variables in open continuous time dynamic models. *Journal of Economic Dynamics and Control* 25, 1399–1427.

McCrorie, J.R. (2002) The likelihood of the parameters of a continuous time vector autoregressive model. *Statistical Inference for Stochastic Processes* 5, 273–286.

McCrorie, J.R. (2003) The problem of aliasing in identifying finite parameter continuous time models. *Acta Applicandae Mathematicae* 79, 9–16.

McCrorie, J.R. (2009) The Integration Order of Vector Autoregressive Processes. Mimeo, University of St. Andrews.

McCrorie, J.R. & M.J. Chambers (2006) Granger causality and the sampling of economic processes. *Journal of Econometrics* 132, 311–326.

McGarry, J.S. (2003) The exact discrete time representation of a system of fourth-order differential equations. *Computers and Mathematics with Applications* 46, 213–230.

Morinaga, K. & T. Nono (1954) On the non-commutative solution of the exponential equation $e^x e^y = e^{x+y}$, II. *Journal of Science of the Hiroshima University, Series A* 18, 137–178.

Nowman, K.B. (1997) Gaussian estimation of single-factor continuous time models of the term structure of interest rates. *Journal of Finance* 52, 1695–1703.

Nowman, K.B. (2009) Rex Bergstrom's contributions to continuous-time macroeconometric modeling. *Econometric Theory* 25 (this issue), 1087–1098.

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

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, 803–824.

Phillips, P.C.B. (1978) The treatment of flow data in the estimation of continuous time systems. In A.R. Bergstrom, A.J.L. Catt, M.H. Peston, & B.D.J. Silverstone (eds.), *Stability and Inflation*, pp. 257–274. Wiley.

Phillips, P.C.B. (1988) The ET interview: Professor A.R. Bergstrom. *Econometric Theory* 4, 301–327.

Phillips, P.C.B. (1991) Error correction and long-run equilibrium in continuous time. *Econometrica* 59, 967–980.

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

Phillips, P.C.B. (2005) Albert Rex Bergstrom 1925–2005. *New Zealand Economic Papers* 39, 129–152.

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

Phillips, P.C.B. & J. Yu (2009) Maximum likelihood and Gaussian estimation of continuous time models in finance. In T.G. Anderson, R.A. Davis, J.-P. Kreiss, & T. Mikosch (eds.), *Handbook of Financial Time Series*. Springer, forthcoming.

Robinson, P.M. (1976) Fourier estimation of continuous time models. In A.R. Bergstrom (ed.), *Statistical Inference in Continuous Time Economic Models*, pp. 215–266. North-Holland.

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

Robinson, P.M. (1980) Continuous model fitting from discrete data. In D.R. Brillinger and G.C. Tiao (eds.), *Directions in Time Series*, pp. 263–278. Institute of Mathematical Statistics.

Robinson, P.M. (1992) Review of A.R. Bergstrom's "Continuous Time Econometric Modelling." *Econometric Theory* 8, 571–579.

Robinson, P.M. (1993) Continuous time models in econometrics: Closed and open systems, stocks and flows. In P.C.B. Phillips (ed.), *Models, Methods and Applications of Econometrics*, pp. 71–90. Blackwell.

Robinson, P.M. (2009) On discrete sampling of time-varying continuous-time systems. *Econometric Theory* 25 (this issue), 985–994.

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

Sims, C.A. (1971) Discrete approximations to continuous time lag distributions in econometrics. *Econometrica* 39, 545–564.

Stock, J.H. (1987) Temporal aggregation and structural inference in macroeconomics: A comment. *Carnegie-Rochester Conference Series in Public Policy* 26, 131–140.

Van Loan, C.F. (1978) Computing integrals involving the matrix exponential. *IEEE Transactions on Automatic Control* AC-23, 395–404.

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

Yu, J. (2008) Bias in the Estimation of Mean Reversion Parameter in Continuous Time Models. Mimeo, Singapore Management University.

Yu, J. & P.C.B. Phillips (2001) A Gaussian approach for continuous time models of the short-term interest rate. *Econometrics Journal* 4, 221–225.

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

Zhang, L, P.A. Mykland, & Y. Aït-Sahalia (2005) A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* 100, 1394–1411.

Zwanziger, D. (1964) Simple theorem on Hermitian matrices and an application to the polarization of vector particles. *Physical Review* 136, B558–B562.

APPENDIX: Illustration of the Identification Problem

Consider model (1) and (2) with $n = 2$, and let

$$A_0 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{and} \quad \Sigma_0 = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \tag{A.1}$$

be real matrices where A_0 represents a coefficient matrix and Σ_0 represents a covariance matrix.[13] In practice, establishing that the model is identified involves both the aliasing

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

problem and the more traditional problem of ensuring that there are not too many free parameters, meaning in practice the matrices will be further restricted by exclusion restrictions and a priori causal information to depend on a smaller number of parameters. For example, consider the following two-factor interest rate model, where the spot interest rate r is given by

$$r(t) = \upsilon - x_1(t) - x_2(t), \tag{A.2}$$

where υ is the long-run average rate, and the factors x_1 and x_2 represent the current effect of streams of short- and long-term economic "news", generated by

$$Dx_1(t) = ax_1(t)\,dt + \sigma_1 W_1(dt), \tag{A.3}$$

$$Dx_2(t) = dx_2(t)\,dt + \sigma_2 W_2(dt), \tag{A.4}$$

where a, $d < 0$ are mean reversion parameters, $\sigma_1, \sigma_2 > 0$ are diffusion coefficients, and $W_1(dt)$ and $W_2(dt)$ represent independent standard Brownian motion processes (Babbs and Nowman, 1997; Babbs and Nowman, 1999).[14] Using the exact discrete model, Bergstrom and Nowman show that if the Brownian motion processes were correlated, the model even with $b = c = 0$ is unidentified, and make the zero correlation restriction a priori. The lack of correlation between the disturbance processes, although restrictive, is also a feature of the two-factor Cox-Ingersoll-Ross model considered by Chen and Scott (1992) and the cyclical models introduced by Ercolani (2009) in this volume.

Now we show using (A.1) that aliasing alone impacts on identification in different ways depending on the values of the parameters. If A_0 has complex eigenvalues λ_1, $\lambda_2 = \frac{1}{2}\mathrm{tr}(A_0) \pm i\delta$, where $\delta = \frac{1}{2}\sqrt{\left|\mathrm{tr}(A_0)^2 - 4\det(A_0)\right|}$, $\mathrm{tr}(A_0)^2 < 4\det(A_0)$, $\mathrm{tr}(A_0) = a+d$, and $\det(A_0) = ad - bc$, then it is diagonalizable, and an invertible matrix T such that $T^{-1}A_0T = \mathrm{diag}(\lambda_1, \lambda_2)$ is the matrix $T = \begin{bmatrix} b & b \\ \lambda_1 - a & \lambda_2 - a \end{bmatrix}$. Then if $M_0 = \begin{bmatrix} -A_0 & \Sigma_0 \\ 0 & A_0' \end{bmatrix}$, $\exp(M_0) = \begin{bmatrix} F_0^{-1} & F_{12} \\ 0 & F_0' \end{bmatrix}$, where (cf. Bernstein and So, 1993)

$$\begin{aligned} F_0 = \exp(A_0) &= T[\mathrm{diag}(\exp(\lambda_1), \exp(\lambda_2))]T^{-1} \\ &= e^{tr(A_0)/2}\begin{bmatrix} \cos\delta + \frac{a-d}{2}\frac{\sin\delta}{\delta} & b\frac{\sin\delta}{\delta} \\ c\frac{\sin\delta}{\delta} & \cos\delta - \frac{a-d}{2}\frac{\sin\delta}{\delta} \end{bmatrix}, \end{aligned} \tag{A.5}$$

and $F_{12} = F_0^{-1}\Omega_0$. Since the characteristic equation of M_0 is a quartic and its eigenvalues are distinct, its exponential can be expressed directly in terms of the elements of A_0 and Σ_0 via Lagrange interpolation. Notice, however, that if X solves the Lyapunov equation $A_0X + XA_0' = -\Sigma_0$, then

$$\begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}^{-1}\begin{bmatrix} -A_0 & \Sigma_0 \\ 0 & A_0' \end{bmatrix}\begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} = \begin{bmatrix} -A_0 & 0 \\ 0 & A_0' \end{bmatrix}, \tag{A.6}$$

and so $F_{12} = \exp(-A_0)\Omega_0 = \exp(-A_0)X - X\exp(A_0')$, implying that

$$\Omega_0 = X - \exp(A_0)X\exp(A_0'). \tag{A.7}$$

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

Phillips (1973, p. 354) alternately defines Ω_0 as the solution of a Lyapunov equation in $\exp(A_0)$ and Σ_0. If we assume, as Phillips did, that the eigenvalues of A_0 do not differ by an integer multiple of $2\pi i$, then the framework of McCrorie (2003) that uses (8) and (9) can be used to discuss identification without the need to deal with Ω_0 directly. The following matrices are constructed to be aliases of A_0:

$$A_p = T[\mathrm{diag}(\lambda_1, \lambda_2) + \mathrm{diag}(2\pi ip, -2\pi ip)]T^{-1}, \tag{A.8}$$

where p runs through the integers. Consider $M^* = \begin{bmatrix} -A_p & \Sigma^* \\ 0 & A_p' \end{bmatrix}$. Then $\exp(M^*) = \begin{bmatrix} F_0^{-1} & F_0^{-1}\Omega^* \\ 0 & F_0' \end{bmatrix}$, where $\Omega^* = \int_0^1 \exp(rA_p)\Sigma^* \exp(rA_p)' dr$. If $\Omega^* = \Omega_0$, then $\exp(M^*) = \exp(M_0)$, and since by assumption M_0 is incongruent mod $2\pi i$, $M_0 M^* = M^* M_0$ (Morinaga and Nono, 1954, Thm. 4).[15] Equating the upper right-hand entries in this expression gives

$$A_0\Sigma^* + \Sigma^* A_0' = A_p\Sigma_0 + \Sigma_0 A_p', \tag{A.9}$$

and diagonalizing A_0 and A_p, premultiplying by T^{-1}, and postmultiplying by its conjugate transpose $(\bar{T}^{-1})'$ gives the Lyapunov equation

$$\Lambda_0(T^{-1}\Sigma^*(\bar{T}^{-1})') + (T^{-1}\Sigma^*(\bar{T}^{-1})')\Lambda_0 = \Lambda^* T^{-1}\Sigma_0(\bar{T}^{-1})' + T^{-1}\Sigma_0(\bar{T}^{-1})'\Lambda^*, \tag{A.10}$$

where $\Lambda_0 = \mathrm{diag}(\lambda_1, \lambda_2)$ and $\Lambda^* = \Lambda_0 + \mathrm{diag}(2\pi ip, -2\pi ip)$. Using vec to denote the vector formed by concatenating the columns of a matrix, $\otimes$ to denote the related Kronecker product, and I_n the $n \times n$ identity matrix, we obtain

$$\begin{aligned} \mathrm{vec}((T^{-1}\Sigma^*(\bar{T}^{-1})') &= (\Lambda_0 \otimes I_2 + I_2 \otimes \Lambda_0)^{-1}(\Lambda^* \otimes I_2 + I_2 \otimes \Lambda^*)\mathrm{vec}(T^{-1}\Sigma_0(\bar{T}^{-1})') \\ &= [I_4 + 2\pi ip\ \mathrm{diag}(1/\lambda_1, 0, 0, -1/\lambda_2)]\,\mathrm{vec}(T^{-1}\Sigma_0(\bar{T}^{-1})'), \end{aligned} \tag{A.11}$$

and so

$$\mathrm{vec}(\Sigma^*) = [I_4 + 2\pi ip(T \otimes T)\,\mathrm{diag}(1/\lambda_1, 0, 0, -1/\lambda_2)(T^{-1} \otimes T^{-1})]\,\mathrm{vec}(\Sigma_0),$$

$$\begin{aligned} \Sigma^* = \Sigma_0 &+ \frac{\pi p}{\delta}\begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \\ &- \frac{\pi p}{\delta(ad-bc)}\begin{bmatrix} (d\sigma_1 - b\sigma_2)^2 + 2bd(1-\rho)\sigma_1\sigma_2 & -cd\sigma_1^2 + 2bc\rho\sigma_1\sigma_2 - ab\sigma_2^2 \\ -cd\sigma_1^2 + 2bc\rho\sigma_1\sigma_2 - ab\sigma_2^2 & (c\sigma_1 - a\sigma_2)^2 + 2ac(1-\rho)\sigma_1\sigma_2 \end{bmatrix}. \end{aligned} \tag{A.12}$$

The expression for Σ^* denotes the set of all matrices such that $\Omega^* = \Omega_0$; however, some of the Σ^* are not positive semidefinite and so are not admissible as covariance matrices. Here, we use (A.12) to illustrate the aliasing problem using the example in Hansen and Sargent (1981).

Let $A_0 = \begin{bmatrix} -1 & 1 \\ -1 & -2 \end{bmatrix}$, $\quad \Sigma_0 = \begin{bmatrix} 2 & 2\rho \\ 2\rho & 2 \end{bmatrix}$, $\quad (|\rho| < 1)$.

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

Then $\Lambda_0 = \text{diag}\left(-\frac{3}{2}+\frac{\sqrt{3}}{2}i, -\frac{3}{2}-\frac{\sqrt{3}}{2}i\right)$, $\quad T = \begin{bmatrix} 1 & 1 \\ -\frac{1}{2}+\frac{\sqrt{3}}{2}i & -\frac{1}{2}-\frac{\sqrt{3}}{2}i \end{bmatrix}$.

Using (A.8), the aliases of A_0 are

$$A_p = \begin{bmatrix} -1 & 1 \\ -1 & -2 \end{bmatrix} + \frac{2\pi p}{\sqrt{3}} \begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix}, \tag{A.13}$$

and from (A.12) the candidate covariance matrices are

$$\Sigma^* = \begin{bmatrix} 2 & 2\rho \\ 2\rho & 2 \end{bmatrix} + \frac{4\pi p(1+2\rho)}{3\sqrt{3}} \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}, \tag{A.14}$$

with associated eigenvalues

$$(2-\gamma) \pm 2\sqrt{(\rho+\gamma)^2 + \left(\frac{3}{2}\gamma\right)^2}, \tag{A.15}$$

where $\gamma = 2\pi p(1+2\rho)/(3\sqrt{3})$. From (A.14), we see that if $\rho = -1/2$, then $\Sigma^* = \Sigma_0$, meaning that each of the A_p matrices is associated with the *same* covariance matrix Σ_0. There is therefore a countable infinity of continuous-time models consistent with the data. Notice, however, that as indicated in Table 1 of Hansen and Sargent (1981), if ρ is allowed to move away from $-1/2$ in both directions, there are fewer and fewer observationally equivalent models and indeed the model is identified in a region of the parameter space above $\rho = -1$ and in a larger region below $\rho = 1$. For example, with $\rho = -0.6$, we can see from (A.15) that only $p = -2$, -1, 0, and 1 are admissible, and with $\rho = -0.8$ (and smaller permissible ρ), only $p = 0$ is admissible, i.e., the model is identified. For $\rho = -0.3$, only $p = -1$, 0, and 1 are admissible and for $\rho = -0.2$ (and larger permissible ρ), only $p = 0$ is admissible.

The above approach can also be used to shed light on the aliasing problem in a second-order univariate model.[16] Consider

$$d[Dx(t)] = a\,Dx(t) + b\,x(t) + \zeta(dt), \tag{A.16}$$

which represents the integral equation

$$Dx(t) - Dx(0) = \int_0^t [a\,Dx(r) + b\,x(r)]\,dr + \int_0^t \zeta(dr) \tag{A.17}$$

for all t. By assumption, $\{Dx(t),\ t > 0\}$ is mean-square continuous, and so

$$x(t) - x(0) = \int_0^t Dx(r)\,dr. \tag{A.18}$$

Combining (A.17) and (A.18) gives a *first-order* system in $\begin{bmatrix} Dx(t) \\ x(t) \end{bmatrix}$:

$$\begin{bmatrix} Dx(t) \\ x(t) \end{bmatrix} - \begin{bmatrix} Dx(0) \\ x(0) \end{bmatrix} = \int_0^t \left\{ \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix} \begin{bmatrix} Dx(r) \\ x(r) \end{bmatrix} \right\} dr + \begin{bmatrix} \int_0^t \zeta(dr) \\ 0 \end{bmatrix}. \tag{A.19}$$

We can therefore discuss the aliasing problem using the analysis above with $c = 1, d = 0$, and $\sigma_2 = 0$ in (A.1). Theorem 1 tells us that for $n = 2$, as here, aliasing is possible only

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press

with complex eigenvalues, and so we consider $b < -a^2/4$. Then $\lambda_1, \lambda_2 = \frac{1}{2}a \pm i\delta$, where $\delta = \frac{1}{2}\sqrt{|a^2+4b|}$. Using (A.12),

$$\Sigma^* = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & 0 \end{bmatrix} + \frac{\pi p}{\delta}\begin{bmatrix} \sigma_1^2 & 0 \\ 0 & 0 \end{bmatrix} + \frac{\pi p}{\delta b}\begin{bmatrix} 0 & 0 \\ 0 & \sigma_1^2 \end{bmatrix} = \sigma_1^2 \begin{bmatrix} 1+\pi p\delta^{-1} & 0 \\ 0 & \pi p(b\delta)^{-1} \end{bmatrix}. \quad \textbf{(A.20)}$$

Since $b < 0, \quad p = 1, 2, \ldots$ are all inadmissible because $\pi p(b\delta)^{-1} < 0$. For negative p, $\pi p(b\delta)^{-1} > 0$. If, however, for $p = -1$, $1 + \pi p\delta^{-1} < 0$, the model is identified; for then $1 + \pi p\delta^{-1} < 0$ for all negative p. Take $\delta < \pi$. Then $|a^2 + 4b| < 4\pi^2$. Hence the model is identified, provided $b > -\pi^2 - a^2/4$, indicating that, for given a, the model is identified across a wider range of values of b than is offered through considering the bijectivity of the mapping between the coefficient matrix and its exponential alone.

Both examples given in the Appendix illustrate the fundamental conclusion of Hansen and Sargent (1983) and others that continuous-time models are only locally identifiable on the basis of discrete data, and the potential impact of aliasing cannot be ignored.

https://doi.org/10.1017/S0266466608090452 Published online by Cambridge University Press