Econometric Theory, **19**, 2003, 49–77. Printed in the United States of America.
DOI: 10.1017/S0266466603191037

THE ASYMPTOTIC EFFICIENCY OF COINTEGRATION ESTIMATORS UNDER TEMPORAL AGGREGATION

MARCUS J. CHAMBERS
University of Essex

This paper examines the effects of temporal aggregation on the asymptotic variances of estimators in cointegrated systems. Two important findings are obtained. First, estimators based on flow data alone are more efficient than when the data are all stocks or a mixture of stocks and flows. Second, estimators based on flow data are as efficient as when the data are recorded continuously. A method of improving efficiency with stock variables is also proposed, and an empirical illustration of the method is provided in the context of long-run money demand regressions.

1. INTRODUCTION

Asymptotic efficiency is a desirable feature for any estimator to possess, not least cointegration estimators. Many, though by no means all, of the estimators of cointegration relationships that have been proposed can be shown to be efficient within a certain class of estimators. Saikkonen (1991) characterizes asymptotic efficiency in the context of cointegration estimators and also proposes an efficient regression estimator of cointegrating parameters, whereas Phillips (1991b) uses the term *optimal estimator* to describe certain maximum likelihood estimators that are efficient within the class of estimators considered by Saikkonen (1991).

The analysis of the asymptotic efficiency of estimators in general, including the case of cointegration, takes the observed data series as given. In practice, however, the way in which measurements of the values of a variable are made depends on the nature of the variable itself. To give a simple example, the measurement of consumers' expenditure (a flow variable) requires observing the accumulated flow of expenditure over a particular time interval, and hence the observations have a time dimension associated with them (e.g., dollars *per quarter* or *per annum*). On the other hand, the measurement of the price of a com-

I thank Roy Bailey, Rex Bergstrom, Roderick McCrorie, a co-editor, and two anonymous referees for helpful comments. I also thank Katsumi Shimotsu for help with some data issues. None of these individuals are implicated, however, in any possible shortcomings of this paper. The financial support provided by the ESRC under grant R000221818 is gratefully acknowledged. Address correspondence to: Marcus J. Chambers, Department of Economics, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, England, e-mail: mchamb@essex.ac.uk.

© 2003 Cambridge University Press 0266-4666/03 $9.50

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

modity, or of an interest rate, can in principle be conducted almost at any point in time, these being examples of stock variables. As a result, the information content of these two types of observation schemes is rather different. Observations on flow variables contain information concerning the unobservable rate of flow during the interval over which the observation is made (effectively, an integral of the underlying rate of flow is being observed). Observations on stock variables, however, do not contain such information, being just point observations of an underlying process at given points in time. Although both types of data can be used to estimate the parameters of a model that describes the underlying (possibly continuous time) evolution of the variable concerned, so far no attempt has been made to investigate whether these different sampling schemes have any effect on the asymptotic efficiency of the resulting estimators.

This paper investigates a new avenue of research related to the asymptotic efficiency of cointegration estimators by considering the effects of temporal aggregation. In this paper, the term *temporal aggregation* is taken to mean the discrete time sampling of an underlying process, discrete or continuous, whose dynamics are determined at a different frequency than that with which the process is observed. Included within this definition, therefore, is the sampling of both stock and flow variables.[1] The paper also constitutes a new line of research in the temporal aggregation literature, which until now has not explored the effects of different sampling schemes on the efficiency of estimators.[2] Although earlier results concerning the asymptotic efficiency of cointegration estimators and the discrete time representation of an underlying continuous time process are utilized, the focus is on deriving the asymptotic variance matrices of a certain class of cointegration estimators, thereby allowing asymptotic efficiency comparisons to be made based on different discrete time sampling schemes. Although the form of discrete time sampling of a variable is determined by whether the variable is a stock or a flow, these results are used to suggest a method of improving efficiency with stock variables. An empirical example assesses the extent to which finite sample efficiency gains can be made using the proposed method.

There are three main theoretical contributions made in this paper. The first is that estimators based on flow data alone are shown to be more efficient than when the data are all stocks or a mixture of stocks and flows. Furthermore, it is also demonstrated that estimators based on flow data are as efficient as when the data are recorded continuously. Both of these results are obtained by a direct comparison of the asymptotic covariance matrices of the estimators under the relevant types of sampling scheme. Finally, a method of improving efficiency with stock variables is also proposed. This is based on the averaging of the stock variables in an attempt to mimic the filter inherent in the discrete time sampling of flow variables.

The paper is organized as follows. Section 2 defines the cointegrated continuous time model, which is based on Phillips (1991a) and provides the exact representation of the disturbance process in the corresponding cointegrated sys-

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

tem satisfied by the discrete time observations. This is achieved for each of the four possible combinations of stock and flow variables for the vectors of normalized (left-hand side) and explanatory (right-hand side) variables in the cointegrating relationships. This step is vital, because the asymptotic variances of the cointegration estimators depend on the value of the spectral density matrix of this disturbance process at the zero frequency. This section also discusses the concept of efficiency of cointegration estimators, following Saikkonen (1991), and provides conditions on the continuous time random disturbances such that the discrete time disturbance vectors satisfy a functional central limit theorem. Section 3 is organized into three sections. The first derives the precise form of the four time invariant linear filters that connect the discrete time disturbance vector with the continuous time disturbance vector. These filters are then used in the next section to derive expressions for the long-run variance matrix of the discrete time disturbance process, submatrices of which determine the asymptotic variance of the cointegration estimators. The final section of Section 3 derives the asymptotic variance of the (infeasible) continuous record estimator. The results in this section, and indeed in the paper as a whole, are quite general in the sense that the continuous time disturbance process is only assumed to be a stationary mixing process and to have a (Hermitian) positive definite spectral density matrix. It is not required to obey any particular law of motion. Section 4 investigates ways in which the relative inefficiencies associated with stock variables can be improved, and it incorporates an empirical example to assess the finite sample efficiency gains resulting from the proposed techniques. Section 5 concludes the paper, and Appendix A contains some supplementary lemmas required for the proofs of the principal theorems, lemmas, and propositions contained in the paper, the outline proofs of which appear in Appendix B. Proofs of the lemmas in Appendix A, and also more detailed proofs than appear in Appendix B, can be obtained from the author on request.[3]

2. THE MODEL AND ITS EFFICIENT ESTIMATION

2.1. The Model

The model of cointegration is given by the system of equations

$$y_1(t) = By_2(t) + u_1(t), \qquad t > 0, \tag{1}$$

$$dy_2(t) = u_2(t)dt, \qquad t > 0, \tag{2}$$

where $y_1(t)$ and $y_2(t)$ are continuous time $I(1)$ processes of dimensions $m_1 \times 1$ and $m_2 \times 1$, respectively, B is an $m_1 \times m_2$ matrix of unknown coefficients, and the $m_1 \times 1$ vector $u_1(t)$ and the $m_2 \times 1$ vector $u_2(t)$ are stationary continuous time random disturbances. The initial conditions $y_1(0)$ and $y_2(0)$ are not constrained to be fixed but may be drawn from a certain specified distribution. The long-run equilibrium relations between $y_1(t)$ and $y_2(t)$ are captured by (1),

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

whereas (2) represents the evolution of $y_2(t)$ as a continuous time $I(1)$ process.[4] Although the disturbance vector $u(t) \equiv [u_1(t)', u_2(t)']'$ is stationary, it is not necessary to specify a particular (parametric) model for this process. Indeed, one of the strengths of this paper is that the results do not require u_1 and u_2 to satisfy a particular law of motion.[5]

Defining the $m \times 1$ vector $y(t) \equiv [y_1(t)', y_2(t)']'$, where $m = m_1 + m_2$, it is possible to write[6] (1) and (2) as a continuous time triangular error correction model (ECM) of the form

$$dy(t) = -JAy(t)dt + w(t)dt, \qquad t > 0, \tag{3}$$

where $A = [I_{m_1}, -B]$, $J = [I_{m_1}, 0]'$ (in which 0 is a null matrix of dimension $m_1 \times m_2$), and

$$w(t)dt \equiv \begin{bmatrix} w_1(t)dt \\ w_2(t)dt \end{bmatrix} = \begin{bmatrix} [u_1(t) + Bu_2(t)]dt + du_1(t) \\ u_2(t)dt \end{bmatrix}.$$

The stationary vector $w(t)dt$ reflects the short-run dynamics of the system that are determined by the properties of u_1 and u_2.

Phillips (1991a) shows that the vector $y(t)$ satisfies the discrete time triangular ECM

$$\Delta y(t) = -JAy(t-1) + x(t), \qquad t = 1,2,\ldots, \tag{4}$$

$$x(t) = \int_0^1 e^{-sJA} w(t-s)ds + e^{-1}JAy(t-1) \tag{5}$$

and obtains two important results. The first is that the elements of the matrix B are identified directly from the discretely sampled data $y(1), y(2), \ldots, y(T)$, where T denotes sample size. This implies that there is no aliasing problem involved with the identification of cointegrating parameters in cointegrated continuous time systems. The second result is that it is not necessary to derive discrete models that explicitly account for the different ways in which stock and flow variables are observed, at least as far as the estimation of cointegrating parameters is concerned. These findings are in stark contrast to the situation in stationary systems of stochastic differential equations in which aliasing is a concern and in which it is vital to account for the differences between stocks and flows. Phillips derives these results in the cointegrated model by examining the behavior of the spectral density matrix of the observed series at the zero frequency. Certain filtered series, such as flows, do not affect the cointegrating relationships between the variables, because the value of the filter at the zero frequency is unity. This means that the discrete time ECM (4) may be used to estimate the matrix B irrespective of whether the data are stocks, flows, or a mixture of both. It is shown subsequently, however, that although it is possible to estimate the matrix B efficiently for a *given* data combination of stocks and

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

flows, the way in which the variables are observed *does* affect the asymptotic variance matrix of the efficient estimators when the different possible data combinations are considered.

2.2. The Sampling Schemes and the Discrete Time Representations

The effects of temporal aggregation are most transparently analyzed by considering, in turn, the four possible combinations[7] of stock and flow variables for the vectors y_1 and y_2. It is convenient to define the stock variables in continuous time as $y_1^S(t)$ and $y_2^S(t)$ and to define the flow variables as $y_1^F(t)$ and $y_2^F(t)$ $(t \geq 0)$. The observed values of these variables, however, are given by $y_{1t}^S = y_1^S(t)$ and $y_{2t}^S = y_2^S(t)$ for the stock variables and by

$$y_{1t}^F = \int_0^1 y_1^F(t-s)ds \quad \text{and} \quad y_{2t}^F = \int_0^1 y_2^F(t-s)ds$$

for the flow variables, where in each case $t = 1,2,\ldots,T$ and T denotes sample size. Using y_t to denote the $m \times 1$ vector of observed values, the four sampling schemes of interest are defined in Table 1.

Despite the differences arising from the nature of stocks and flows detailed previously, the discrete time ECMs in each of the four stock-flow combinations can nevertheless be written in a common discrete time triangular ECM form, each case differing in the specification of the observable vector y_t (defined previously for each sampling scheme) and in the specification of the unobservable disturbance vector. The precise form of each discrete time ECM is given in Lemma 1, which follows.

TABLE 1. The four sampling schemes for the observed vector $y_t = [y_{1t}', y_{2t}']'$ $(t = 1,\ldots,T)$

Scheme	y_1	y_2	y_{1t}	y_{2t}
I	Stocks	Stocks	$y_1^S(t)$	$y_2^S(t)$
II	Stocks	Flows	$y_1^S(t)$	$\int_0^1 y_2^F(t-s)ds$
III	Flows	Stocks	$\int_0^1 y_1^F(t-s)ds$	$y_2^S(t)$
IV	Flows	Flows	$\int_0^1 y_1^F(t-s)ds$	$\int_0^1 y_2^F(t-s)ds$

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

LEMMA 1. *Let $y(t)$ satisfy the continuous time cointegrated system defined by (1) and (2). Then the observed vector y_t evolves according to the discrete time triangular* ECM

$$\Delta y_t = -JAy_{t-1} + \xi_t, \qquad t = 1, \ldots, T, \tag{6}$$

in which the components of the disturbance vector $\xi_t \equiv [\xi_{1t}', \xi_{2t}']'$ are defined, for each of the four sampling schemes, as follows.

Scheme I.

$$\xi_{1t} = u_1(t) + B\int_0^1 u_2(t-s)ds;$$

$$\xi_{2t} = \int_0^1 u_2(t-s)ds.$$

Scheme II.

$$\xi_{1t} = u_1(t) + B\left[\int_0^1 u_2(t-s)ds + \int_0^1 (1-s)u_2(t-1-s)ds\right];$$

$$\xi_{2t} = \int_0^1\int_0^1 u_2(t-r-s)drds.$$

Scheme III.

$$\xi_{1t} = \int_0^1 u_1(t-s)ds + B\int_0^1\int_0^1 u_2(t-r-s)drds - B\int_0^1 (1-s)u_2(t-1-s)ds;$$

$$\xi_{2t} = \int_0^1 u_2(t-s)ds.$$

Scheme IV.

$$\xi_{1t} = \int_0^1 u_1(t-s)ds + B\int_0^1\int_0^1 u_2(t-r-s)drds;$$

$$\xi_{2t} = \int_0^1\int_0^1 u_2(t-r-s)drds.$$

The first step in establishing the validity of Lemma 1 is to relate $x(t)$ in (4) directly to $u(t)$, a result that is given in Lemma A1 in Appendix A. Schemes I and IV in Lemma 1, in which y_1 and y_2 are either stocks or flows, are derived directly from (4) and by integrating (4), respectively, the latter case yielding

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

$$\int_0^1 \Delta y(t-s)ds = -JA\int_0^1 y(t-1-s)ds + \int_0^1 x(t-s)ds, \qquad t = 1,\ldots,T. \tag{7}$$

In the remaining two cases, the joint presence of both stocks and flows requires that the appropriate equations from the systems (4) and (7) are selected and transformed accordingly. Full details of the manipulations are given in the proof of Lemma 1 in Appendix B.

2.3. An Invariance Principle and Asymptotically Efficient Estimation

The derivation of the asymptotic properties of an estimator (efficient or otherwise) of cointegrating parameters is typically based on an invariance principle for ξ_t. Lemma 1 plays an important role in establishing such an invariance principle, because it relates ξ_t to the underlying continuous time disturbance process $u(t)$, about which the following assumptions are made. Let $(\Omega_u, \mathcal{F}, P)$ denote the probability space on which $u(t)$ is defined and let $\mathcal{F}_a^b$ $(a < b)$ denote a σ-subfield of $\mathcal{F}$. Then the uniform and strong mixing coefficients, ϕ_j and α_j, respectively, are defined by $\phi_j = \sup_t \phi(\mathcal{F}_{-\infty}^t, \mathcal{F}_{t+j}^\infty)$ and $\alpha_j = \sup_t \alpha(\mathcal{F}_{-\infty}^t, \mathcal{F}_{t+j}^\infty)$, where

$$\phi(\mathcal{F}_{-\infty}^t, \mathcal{F}_{t+j}^\infty) = \sup_{G\in\mathcal{F}_{-\infty}^t, H\in\mathcal{F}_{t+j}^\infty} |P(H|G) - P(H)|,$$

$$\alpha(\mathcal{F}_{-\infty}^t, \mathcal{F}_{t+j}^\infty) = \sup_{G\in\mathcal{F}_{-\infty}^t, H\in\mathcal{F}_{t+j}^\infty} |P(G\cap H) - P(G)P(H)|.$$

Assumption 1. The continuous time disturbance vector $u(t)$ is a stationary random process with Hermitian positive definite spectral density matrix $f(\lambda)$ $(-\infty < \lambda < \infty)$ and satisfies

(a) $E[u_i(t)] = 0$ $(i = 1,\ldots,m)$;
(b) $E|u_i(t)|^\beta < \infty$ $(i = 1,\ldots,m)$ for some $2 \le \beta < \infty$;
(c) either (i) $\sum_{j=1}^\infty \phi_j^{1-1/\beta} < \infty$ or (ii) $\beta > 2$ and $\sum_{j=1}^\infty \alpha_j^{1-2/\beta} < \infty$.

Assumption 2. The coefficients of the cointegrating matrix B in (1) satisfy $|B_{ij}| < \infty$ $(i = 1,\ldots,m_1; j = 1,\ldots,m_2)$.

Assumption 1 specifies $u(t)$ to be a zero mean stationary mixing process in which the mixing coefficients decay sufficiently rapidly to satisfy the stated summability conditions. Furthermore, absolute moments of a sufficiently high order (β) are also required to exist. Mixing conditions are particularly convenient to use in the present context, because in all four sampling schemes ξ_t is a measurable function of $u(t)$ over a finite interval and hence ξ_t inherits the same mixing properties.[8] Assumption 2 is a very weak requirement ensuring that the cointegrating coefficients are finite.

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

LEMMA 2. *Under Assumptions 1 and 2, the disturbance vector ξ_t in the discrete time ECM (6) is a stationary random process with Hermitian positive definite spectral density matrix $f_{\xi\xi}(\lambda)$ $(-\pi < \lambda \leq \pi)$ and satisfies the invariance principle*

$$T^{-1/2}\sum_{j=1}^{[Tr]}\xi_j \Rightarrow S(r) \quad \text{as } T\to\infty,$$

where the symbol $\Rightarrow$ denotes weak convergence of the relevant probability measures, $[Tr]$ denotes the integer part of Tr, and $S(r)$ is an $m\times 1$ Brownian motion process on $r\in[0,1]$ with covariance matrix $\Omega = 2\pi f_{\xi\xi}(0)$.

The invariance principle for ξ_t in Lemma 2 enables the limiting distributions of estimators of B to be expressed in terms of the Brownian motion process $S(r)$. The investigation of efficiency will be carried out with respect to estimators belonging to a well-defined class of estimators. It is convenient to partition $S(r)$ and Ω conformably with $y(t)$ and to define $S_{1.2}(r) = S_1(r) - \Omega_{12}\Omega_{22}^{-1}S_2(r)$. Following Saikkonen (1991),[9] an asymptotically efficient estimator, $\hat{B}$, of the cointegrating matrix B will be defined as one for which

$$T\,\mathrm{vec}(\hat{B}-B) \Rightarrow \left[\left(\int_0^1 S_2S_2'\right)^{-1}\otimes I_{m_1}\right]\mathrm{vec}\left(\int_0^1 dS_{1.2}S_2'\right) \quad \text{as } T\to\infty. \qquad \textbf{(8)}$$

Estimators satisfying the limiting distribution in (8) have alternatively been termed *optimal* estimators by Phillips (1991b). Because the Brownian motions $S_{1.2}(r)$ and $S_2(r)$ are independent, the limiting distribution in (8) can also be represented in the now familiar mixed normal form

$$\int_{G>0} N(0, G\otimes\Omega_{11.2})\,dP_G(G), \qquad \textbf{(9)}$$

in which $G \equiv (\int_0^1 S_2S_2')^{-1}$, $P_G(G)$ is the associated probability measure, and the matrix $\Omega_{11.2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$ denotes the covariance matrix of $S_{1.2}(r)$.

The importance of the mixed normal representation of the limiting distribution given in (9) was recognized by Phillips (1989). A number of cointegration estimators are known to have this limiting distribution and, hence, to be asymptotically efficient. Such efficient estimators typically fall into one of two categories. The first type treat the system dynamics either nonparametrically or make nonparametric adjustments to a first-stage estimator. Examples of this type of estimator include the spectral regression estimators suggested by Phillips (1991a), the fully modified estimators of Phillips and Hansen (1990), and the regression estimator of Saikkonen (1991). The second category of estimator is characterized by a fully parametric specification of the system dynamics, typically in vector autoregressive form. Estimators falling within this second category include the maximum likelihood estimators of Johansen (1991) and Phillips (1991b) and the reduced rank estimator of Reinsel (1990).

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

To compare the asymptotic efficiency of efficient estimators under different combinations of stock and flow data, it is convenient to rewrite the mixed normal distribution in (9) as[10]

$$\int_{\gamma>0} N(0, \gamma\Omega_{22}^{-1} \otimes \Omega_{11.2})\, dP_\gamma(\gamma), \tag{10}$$

where $\gamma = e_2'(\int_0^1 W_2 W_2')^{-1} e_2$, e_2 is any unit $m_2 \times 1$ vector, $P_\gamma(\gamma)$ is the probability measure associated with γ, and W_2 is an $m_2 \times 1$ vector of standard Brownian motions or Wiener processes (i.e., W_2 is Brownian motion with covariance matrix I_{m_2}). The representation of the distribution in (10) is particularly useful for the purposes of comparison because it is expressed as an integral with respect to the scalar random mixing variate γ that is common to each distribution, regardless of the underlying discrete time sampling scheme. Hence the efficiency comparison focuses on the matrices Ω_{22} and $\Omega_{11.2}$, both of which are obtained from Ω, the long-run variance matrix of the process ξ_t. Because the precise form of ξ_t has been shown in Lemma 1 to depend on whether the variables are stocks or flows, it suggests that the relative efficiency of efficient estimators across sampling schemes depends on the nature of the variables themselves. It should be emphasized once more that, for a *given* sampling scheme, an estimator with the limiting distribution in (8) or (10) is asymptotically efficient. The aim here is to ascertain the extent to which different sampling schemes affect the relative asymptotic efficiency of such estimators *across* the different sampling schemes. The next section, therefore, derives expressions for Ω_{22} and $\Omega_{11.2}$ in each of the four sampling schemes to facilitate such a comparison.

3. ASYMPTOTIC EFFICIENCY UNDER TEMPORAL AGGREGATION

3.1. The Spectral Density Matrix $f_{\xi\xi}(\lambda)$

The asymptotic variances of efficient estimators have been shown to depend on the submatrices Ω_{22} and $\Omega_{11.2}$ that are derived from the long-run variance matrix $\Omega = 2\pi f_{\xi\xi}(0)$. Lemma 1 has already provided the relationship between ξ_t and $u(t)$ for each of the four sampling schemes, and this information can be used to derive $f_{\xi\xi}(\lambda)$ from $f(\lambda)$, the spectral density matrix of $u(t)$. This task is greatly facilitated by the use of filters because ξ_t can be regarded as a filtered version of $u(t)$. The filters corresponding to the time domain representations given in Lemma 1 are provided in Lemma 3.

LEMMA 3. *The stationary process ξ_t is related to the continuous time disturbance process $u(t)$ by the filtering equation $\xi_t = M(D)u(t)$, where D denotes the mean square differential operator and the matrix filter function $M(z)$ is defined for each sampling scheme as follows.*

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

Scheme I.

$$M(z) = \begin{bmatrix} I_{m_1} & g(z)B \\ 0_{m_2,m_1} & g(z)I_{m_2} \end{bmatrix};$$

Scheme II.

$$M(z) = \begin{bmatrix} I_{m_1} & [g(z)+h(z)]B \\ 0_{m_2,m_1} & g(z)^2 I_{m_2} \end{bmatrix};$$

Scheme III.

$$M(z) = \begin{bmatrix} g(z)I_{m_1} & [g(z)^2 - h(z)]B \\ 0_{m_2,m_1} & g(z)I_{m_2} \end{bmatrix};$$

Scheme IV.

$$M(z) = \begin{bmatrix} g(z)I_{m_1} & g(z)^2 B \\ 0_{m_2,m_1} & g(z)^2 I_{m_2} \end{bmatrix}.$$

In the preceding definitions, $g(z) = (1 - e^{-z})/z$ *and* $h(z) = e^{-z}[1 - g(z)]/z$.

The filter relating ξ_t and $u(t)$, denoted $M(D)$, is a function of the (mean square) differential operator D because the filter operates on the continuous time process $u(t)$. This does *not* require, however, that $u(t)$ itself be mean square differentiable, because in none of the sampling schemes does the filter $M(D)$ involve the operation of applying D directly to $u(t)$, so that in no case is $Du(t)$ required to exist. The filter $g(D)$, for example, is the filter associated with the process of integration over the unit interval, so that $g(D)u(t) = \int_0^1 u(t-s)ds$. Although this requires that $u(t)$ be integrable in the wide sense[11] it does not require mean square differentiability. Similar comments apply to the filter $h(D)$ that arises in the mixed stock-flow cases.

The filter function $M(z)$ is utilized in deriving the spectral density function $f_{\xi\xi}(\lambda)$ of ξ_t. Because ξ_t is a discrete time process and $u(t)$ is a continuous time process, it is necessary to fold back the frequencies in the range $(-\infty,+\infty)$ into the interval $(-\pi,+\pi]$. The spectral density matrix of ξ_t, regarded as a continuous time process, is given by

$$f^c_{\xi\xi}(\lambda) = M(i\lambda)f(\lambda)M(i\lambda)^*, \qquad -\infty < \lambda < +\infty, \tag{11}$$

where $M(i\lambda)^*$ denotes the transpose of the complex conjugate of $M(i\lambda)$. This representation follows because $i\lambda$ is the frequency response function of the operator D and hence $M(i\lambda)$ is the frequency response function of the filter $M(D)$. The spectral density matrix of the discrete time process ξ_t is then related to $f^c_{\xi\xi}(\lambda)$ by the folding formula:[12]

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

$$f_{\xi\xi}(\lambda) = \sum_{j=-\infty}^{+\infty} f^{c}_{\xi\xi}(\lambda + 2\pi j)$$
$$= \sum_{j=-\infty}^{+\infty} M(i(\lambda + 2\pi j)) f(\lambda + 2\pi j) M(i(\lambda + 2\pi j))^{*},$$
$$-\pi < \lambda \leq +\pi. \qquad \textbf{(12)}$$

The properties of the matrix $\Omega = 2\pi f_{\xi\xi}(0)$ can be explored from the preceding representation for $f_{\xi\xi}(\lambda)$, and some important and general results are derived in the next section.

3.2. The Asymptotic Variances

The matrix of interest is $\Omega = 2\pi f_{\xi\xi}(0) = 2\pi \sum_{j=-\infty}^{+\infty} M(2\pi ij) f(2\pi j) M(2\pi ij)^{*}$, which is obtained from (12) by setting $\lambda = 0$. It is convenient, for the presentation of the results, to define the matrix filter function

$$H(z, B) = [I_{m_1}, h(z)B]$$

and also the matrix $f_{11.2}(0) = f_{11}(0) - f_{12}(0) f_{22}(0)^{-1} f_{21}(0)$, which is obtained from partitions of the spectral matrix $f(\lambda)$ at the zero frequency in the same way that $\Omega_{11.2}$ is obtained from Ω. The required expressions for the matrices Ω_{22} and $\Omega_{11.2}$ are presented in Theorem 1 for each of Schemes I–IV.

THEOREM 1. *The matrices Ω_{22} and $\Omega_{11.2}$ appearing in the limiting mixed normal distribution of an efficient estimator of the cointegrating matrix B are given by the following expressions for each of the four sampling schemes.*

Scheme I.

$$\Omega_{22} = 2\pi f_{22}(0); \qquad \Omega_{11.2} = 2\pi \left[\sum_{j \neq 0} f_{11}(2\pi j) + f_{11.2}(0) \right];$$

Scheme II.

$$\Omega_{22} = 2\pi f_{22}(0); \qquad \Omega_{11.2} = 2\pi \left[\sum_{j \neq 0} H(2\pi ij, B) f(2\pi j) H(2\pi ij, B)^{*} + f_{11.2}(0) \right];$$

Scheme III.

$$\Omega_{22} = 2\pi f_{22}(0); \qquad \Omega_{11.2} = 2\pi \left[B \sum_{j \neq 0} |h(2\pi ij)|^2 f_{22}(2\pi j) B' + f_{11.2}(0) \right];$$

Scheme IV.

$$\Omega_{22} = 2\pi f_{22}(0); \qquad \Omega_{11.2} = 2\pi f_{11.2}(0).$$

The expressions for the Ω_{22} and $\Omega_{11.2}$ matrices given in Theorem 1 are quite general in that they hold for any stationary $u(t)$ process with spectral density

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

matrix $f(\lambda)$ satisfying Assumption 1. The expressions are convenient representations of these matrices that utilize the precise form of the filter function $M(z)$ in each sampling scheme. Note, however, that the Ω_{22} matrices are the same in each scheme. This enables the comparisons of estimator efficiency to proceed on the basis of a comparison of the $\Omega_{11.2}$ matrix in each case. Such a comparison yields the following immediate consequence of Theorem 1.

PROPOSITION 1. *The most efficient estimators of cointegrating parameters are obtained when all the variables are flows (Scheme IV).*

The validity of this proposition is demonstrated by considering the three matrix differences $\Omega^{j}_{11.2} - \Omega^{IV}_{11.2}$ for $j = I, II, III$ and showing that each of these differences is positive semi-definite. This is the sense in which the term *efficient* is used in Proposition 1, even though the estimators for each sampling scheme are efficient for that particular scheme. Notwithstanding this fact, these results show that, when different sampling schemes are considered, it is flow data that yield the most efficient estimates of cointegrating parameters. If it were possible to choose the way in which one's variables were sampled, then the choice should unambiguously be as flows. Although such a choice is not available to econometricians working with given variables, it does suggest that estimates obtained with flow data are more precise (in the sense of having smaller variances) than those obtained with mixtures of stocks and flows or with stocks alone. Whether it is possible to obtain any improvements at all in the efficiency of estimators based on stock variables is explored in Section 4.

3.3. Continuous Record Asymptotics

It is of interest to compare the efficiency of the estimators based on discretely sampled data with an efficient estimator that is obtained using a continuous record of data. In such circumstances, the classification of variables as stocks or flows is somewhat irrelevant, because the variables are observed on a continuous basis. Although, in practice, continuous time observations are not available to the econometrician, the analysis of the estimator based on a continuous record does serve to illuminate the results obtained earlier.

Assuming that a continuous record of observations is available on $y(t)$ and $dy(t)$ for $0 \leq t \leq T$, estimation is based on the continuous time triangular system (3). Analogous to the discrete time case, the asymptotic properties of the continuous record estimator $\tilde{B}$ require that, as $T \to \infty$,

$$T^{-1/2} \int_0^{[Tr]} w(s)ds \Rightarrow \tilde{S}(r), \tag{13}$$

where $\tilde{S}(r)$ is an $m \times 1$ Brownian motion process with covariance matrix $\tilde{\Omega} = 2\pi f_{ww}(0)$ and $f_{ww}(\lambda)$ $(-\infty < \lambda < \infty)$ denotes the spectral density matrix of $w(t)$. It then follows that the asymptotic distribution of $\tilde{B}$ will have the representation (9) or, equivalently, (10), but in which the matrices Ω_{22} and $\Omega_{11.2}$ will

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

be submatrices of $\tilde{\Omega} = 2\pi f_{ww}(0)$ rather than of $\Omega = 2\pi f_{\xi\xi}(0)$. This is because the spectral matrix of the disturbance process in the model at frequency zero determines the long-run variance matrix and in the continuous record case it is $w(t)$, rather than ξ_t, that constitutes the disturbance vector. These observations lead immediately to the following result.

THEOREM 2. *The asymptotic distribution of* $T\,\mathrm{vec}(\tilde{B} - B)$ *is given by*

$$T\,\mathrm{vec}(\tilde{B} - B) \Rightarrow \int_{\gamma>0} N(0, \gamma\tilde{\Omega}_{22}^{-1} \otimes \tilde{\Omega}_{11.2})\, dP_\gamma(\gamma) \quad \text{as } T \to \infty,$$

where $\tilde{\Omega}_{22} = 2\pi f_{22}(0)$, $\tilde{\Omega}_{11.2} = 2\pi f_{11.2}(0)$, *and the random mixing variate* γ *is defined following (10).*

Theorem 2 characterizes the asymptotic distribution of an efficient estimator in the idealized case where observations on the variables are recorded continuously. Note that the matrices $\tilde{\Omega}_{22}$ and $\tilde{\Omega}_{11.2}$ are defined in terms of $f(\lambda)$, the spectral density matrix of $u(t)$, which enables a direct comparison with the expressions in Theorem 1. Such a comparison yields the following result.

PROPOSITION 2. *An efficient estimator obtained using flow data alone (Scheme IV) is asymptotically as efficient as the estimator obtained using a continuous record of data.*

This remarkable result indicates that efficient estimators of cointegrating parameters are asymptotically as efficient with discretely sampled flow data as with a continuous record of observations. The observed values of flow variables, by their nature, contain information about the sample path of the variable over each observation interval, whereas stock variables do not contain the same type of information, being observed only at a single point in each interval. Such a difference lies at the heart of the result reported previously.

4. IMPROVING EFFICIENCY WITH STOCK VARIABLES

4.1. The Theory

The result that cointegration estimators are more efficient when obtained using flow data is of little comfort to a researcher working with stock variables. The question arises, therefore, as to whether the results contained in this paper can be used to suggest ways of improving the efficiency of estimators obtained using stock variables. Effectively this reduces to the issue of whether it is possible to approximate the filter function $g(z)$, associated with the operation of integration, which relates the unobservable flow $y^F(t)$ to its discretely observed value, $y_t^F = \int_0^1 y^F(t-s)ds$, so that $y_t^F = g(D)y^F(t)$. Recall that the filter function $g(z)$ appears in the matrix filter function $M(z)$ relating ξ_t to $u(t)$ and that $M(2\pi ij)$ (for integer j) helps to determine the matrix $f_{\xi\xi}(0)$, upon which $\Omega_{11.2}$ is based. The particular values of $g(z)$ that are of interest are,

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

therefore, $g(0) = 1$ and $g(2\pi ij) = 0$ for integer $j \neq 0$. Given observations on a stock variable at a fixed frequency, there seems little scope for improving efficiency. Attempting to approximate a flow by taking the average of p successive observations using $\tilde{y}_t^S = (1/p)\sum_{k=0}^{p-1} y^S(t-k)$ involves the filter function $a(z) = (1/p)\sum_{k=0}^{p-1} e^{-kz}$. This function satisfies $a(0) = 1$ and $a(2\pi ij) = 1$ for integer $j \neq 0$, which has no effect whatsoever on asymptotic efficiency.

One situation in which efficiency can be improved is where observations on the stock variable are available at a higher frequency than that used for estimation. In models with mixtures of stocks and flows, this situation might arise if it were possible to sample the stocks more frequently than the flows. For example, some stock variables, such as interest rates, exchange rates, and certain prices, are observable at much higher frequencies than flow variables such as consumption, income, and exports. In these cases it would be advantageous to average the stock variables over the (unit) observation interval associated with the sampling of the flows. Suppose that the stock variables are recorded p times each unit observation period. In terms of unit time intervals, the sequence of adjusted stock data is of the form

$$y_{t,p}^S = \frac{1}{p}\sum_{k=0}^{p-1} y^S\left(t - \frac{k}{p}\right), \qquad t = 1, \ldots, T.$$

In terms of filters, the averaged stock variable $y_{t,p}^S$ is related to $y^S(t)$ by the expression $y_{t,p}^S = \sigma_p(D)y^S(t)$, where the filter function $\sigma_p(z)$ is given by

$$\sigma_p(z) = \frac{1}{p}\sum_{k=0}^{p-1} e^{-(k/p)z},$$

which satisfies $\sigma_p(0) = 1$ and

$$\sigma_p(2\pi ij) = \begin{cases} 1, & j = 0, \pm p, \pm 2p, \ldots; \\ 0, & \text{otherwise.} \end{cases}$$

As $p \to \infty$, the discrete sum determining the averaged variable $y_{t,p}^S$ converges to an integral, and so $\lim_{p\to\infty} y_{t,p}^S = \int_0^1 y^S(t-s)ds$. In practice, although the stock variable $y^S(t)$ may be observed very frequently relative to y_t^F, this limiting case will never be achieved, and so there will always be some loss of efficiency associated with stock variables as compared to flows.

To formalize the arguments, the three cases in which stock variables occur need to be modified to account for the averaging of the stocks over the finer grid of time points comprising each unit time interval. Details of these sampling schemes are given in Table 2, in which $y_{t,p}$ refers to the $m \times 1$ vector of observations.

For the vector $y_{t,p}$, the form of the triangular discrete time ECM representation given in (6) is still valid, although the form of the disturbance vector ξ_t is, of course, changed. Lemma 4, which follows, therefore provides the exact dis-

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

TABLE 2. The sampling schemes with averaged stock data in the observed vector $y_{t,p} = [y'_{1t,p}, y'_{2t,p}]'$ $(t = 1, \ldots, T)$

Scheme	y_1	y_2	$y_{1t,p}$	$y_{2t,p}$
I(p)	Stocks	Stocks	$\frac{1}{p}\sum_{k=0}^{p-1} y_1^S\left(t-\frac{k}{p}\right)$	$\frac{1}{p}\sum_{k=0}^{p-1} y_2^S\left(t-\frac{k}{p}\right)$
II(p)	Stocks	Flows	$\frac{1}{p}\sum_{k=0}^{p-1} y_1^S\left(t-\frac{k}{p}\right)$	$\int_0^1 y_2^F(t-s)\,ds$
III(p)	Flows	Stocks	$\int_0^1 y_1^F(t-s)\,ds$	$\frac{1}{p}\sum_{k=0}^{p-1} y_2^S\left(t-\frac{k}{p}\right)$

crete time ECM representation when the stock variables are averaged over the unit observation period.

LEMMA 4. *Let $y(t)$ satisfy the continuous time cointegrated system defined by (1) and (2). Then the observed vector $y_{t,p}$ evolves according to the discrete time triangular* ECM

$$\Delta y_{t,p} = -JAy_{t-1,p} + \xi_{t,p}, \qquad t = 1, \ldots, T, \tag{14}$$

in which the components of the disturbance vector $\xi_{t,p} \equiv [\xi'_{1t,p}, \xi'_{2t,p}]'$ are defined, for each of the three sampling schemes involving averaged stock variables, as follows.

Scheme I(p).

$$\xi_{1t,p} = \frac{1}{p}\sum_{k=0}^{p-1}\left[u_1\left(t-\frac{k}{p}\right) + B\int_0^1 u_2\left(t-\frac{k}{p}-s\right)ds\right];$$

$$\xi_{2t,p} = \frac{1}{p}\sum_{k=0}^{p-1}\left[\int_0^1 u_2\left(t-\frac{k}{p}-s\right)ds\right].$$

Scheme II(p).

$$\xi_{1t,p} = \frac{1}{p}\sum_{k=0}^{p-1}\left[u_1\left(t-\frac{k}{p}\right) + B\int_0^1 u_2\left(t-\frac{k}{p}-s\right)ds\right] + B\left[\frac{1}{p}\sum_{k=0}^{p-1}\int_{k/p}^1 u_2(t-1-s)ds - \int_0^1 su_2(t-1-s)ds\right];$$

$$\xi_{2t,p} = \int_0^1 u_2(t-s)ds.$$

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

Scheme III(p).

$$\xi_{1t,p} = \int_0^1 u_1(t-s)ds + B\int_0^1\int_0^1 u_2(t-r-s)drds + B\left[\int_0^1 su_2(t-1-s)ds - \frac{1}{p}\sum_{k=0}^{p-1}\int_{k/p}^1 u_2(t-1-s)ds\right];$$

$$\xi_{2t,p} = \frac{1}{p}\sum_{k=0}^{p-1}\left[\int_0^1 u_2\left(t-\frac{k}{p}-s\right)ds\right].$$

The results in Lemma 4 generalize those in Lemma 1, which may be obtained as a special case of Lemma 4 by setting $p = 1$. These definitions relate $\xi_{t,p}$ to the continuous time disturbance process $u(t)$ and, hence, implicitly determine the matrix filter relating these two processes. The precise form of these filters is given in Lemma 5, which is analogous to the results in Lemma 3 for the unadjusted stock variables.

LEMMA 5. *The stationary process $\xi_{t,p}$ is related to the continuous time disturbance process $u(t)$ by the filtering equation $\xi_{t,p} = M_p(D)u(t)$, where the matrix filter function $M_p(z)$ is defined in each sampling scheme involving averaged stock variables as follows.*

Scheme I(p).

$$M_p(z) = \begin{bmatrix} \sigma_p(z)I_{m_1} & \sigma_p(z)g(z)B \\ 0_{m_2,m_1} & \sigma_p(z)g(z)I_{m_2} \end{bmatrix};$$

Scheme II(p).

$$M_p(z) = \begin{bmatrix} \sigma_p(z)I_{m_1} & [\sigma_p(z)g(z) + h_p(z)]B \\ 0_{m_2,m_1} & g(z)^2 I_{m_2} \end{bmatrix};$$

Scheme III(p).

$$M_p(z) = \begin{bmatrix} g(z)I_{m_1} & [g(z)^2 - h_p(z)]B \\ 0_{m_2,m_1} & \sigma_p(z)g(z)I_{m_2} \end{bmatrix}.$$

The function $h_p(z)$ is defined by $h_p(z) = e^{-z}[\sigma_p(z) - g(z)]/z$.

The dependence of these filter functions on p has been emphasized by the notation $M_p(z)$. In the same way that the filter functions $M(z)$ in Lemma 3 were an essential part of deriving the matrices determining the asymptotic variances of optimal estimators in Section 3, the filter functions in Lemma 5 enable the asymptotic variance matrices to be determined for efficient estimators based on the averaged stock variables. These asymptotic variance matrices are presented in Theorem 3.

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

THEOREM 3. *The matrices Ω_{22} and $\Omega_{11.2}$ appearing in the limiting mixed normal distribution of an efficient estimator of the cointegrating matrix B are given by the following expressions under the three sampling schemes involving averaged stock variables.*

Scheme I(p).

$$\Omega_{22} = 2\pi f_{22}(0); \qquad \Omega_{11.2} = 2\pi\left[\sum_{j\neq 0} f_{11}(2\pi jp) + f_{11.2}(0)\right];$$

Scheme II(p).

$$\Omega_{22} = 2\pi f_{22}(0); \qquad \Omega_{11.2} = 2\pi\left[\sum_{j\neq 0} H(2\pi ijp, B) f(2\pi jp) H(2\pi ijp, B)^* + f_{11.2}(0)\right];$$

Scheme III(p).

$$\Omega_{22} = 2\pi f_{22}(0); \qquad \Omega_{11.2} = 2\pi\left[B\sum_{j\neq 0} |h(2\pi ijp)|^2 f_{22}(2\pi jp)B' + f_{11.2}(0)\right].$$

In each case, note that the only nonzero terms in the expression for $\Omega_{11.2}$ are multiples of p due to the filter $\sigma_p(z)$. The expressions in Theorem 3 can be compared directly to those in Theorem 1 to yield the following qualitative result.

PROPOSITION 3. *In models containing stock variables, estimators of cointegrating parameters are more efficient when based on the averaged stock variable $y^S_{t,p}$ than when based on the unadjusted stock variable y^S_t.*

The validity of Proposition 3 is established by comparing the $\Omega_{11.2}$ matrices in Theorems 1 and 3 and showing that the differences are positive definite. The implication for applied work is clear. If it is possible to construct the averaged variables $y^S_{t,p}$ for stock data, then doing so unambiguously improves the asymptotic efficiency of the cointegration estimators. This could prove to be important in models containing mixed stock and flow variables, because in many cases it will be possible to observe the stock variables at higher frequencies than the flow variables. In such circumstances, the efficiency of the estimators can be improved by applying the filter $\sigma_p(D)$ to the stock variables.

Some further comments concerning the proposed averaging are in order.[13] The first is that, although the averaging of stock variables may be a good idea theoretically, it may not always be sensible to do so in practice, particularly in circumstances in which the stock variable provides a measure of a quantity that lies at the heart of certain testable propositions that would be adversely affected by the averaging. The second concerns the issue of seasonality. Although the underlying continuous time model in (1) and (2) does not explicitly account for seasonality, it must be recognized that the observed data may be

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

seasonal and hence some accounting for seasonal effects may need to be made. Further care may also be needed when averaging stock variables observed at a higher frequency so as not to introduce further seasonal distortions. Finally, it is interesting to briefly reflect on the form in which stock variables are typically made available to researchers in published national accounts and other databases. Variables such as the money stock and price indices are usually available at a monthly frequency but often reflect a certain degree of averaging in their construction. Other variables, such as stock prices, interest rates, and exchange rates, are available on a daily basis in databases such as *Datastream* and are, hence, amenable to the proposed averaging method. Other data sources, such as *Citibase/DRI* and the *International Financial Statistics* published by the International Monetary Fund, typically report monthly or quarterly averages, although the latter (at least) also contains mid-quarter exchange rates in addition to quarterly averages. There is, therefore, scope for using the proposed methods in empirical research in economics.

4.2. An Empirical Illustration

It is of interest to ascertain the extent to which the filter $\sigma_p(D)$ can yield efficiency gains in finite samples, and the long-run money demand relationship provides a convenient model within which to explore such issues empirically. Defining $m(t)$ as the real money stock, $y(t)$ as the level of real income, and $r(t)$ as the interest rate and assuming the variables are individually $I(1)$ processes, the long-run money demand relationship can be cast as the cointegrated system

$$\ln m(t) = \beta_1 \ln y(t) + \beta_2 r(t) + u_1(t), \qquad t > 0, \tag{15}$$

$$d \ln y(t) = u_2(t)dt, \qquad dr(t) = u_3(t)dt, \qquad t > 0, \tag{16}$$

where the vector $u(t) \equiv [u_1(t), u_2(t), u_3(t)]'$ is a stationary random disturbance vector. This model is in the form of the cointegrated system (1) and (2), and hence the preceeding results apply to the estimation of the parameters β_1 and β_2.

The data used in this empirical illustration refer to the United States and cover the period January 1, 1959, to December 31, 1999. The money stock series (real M1 money balances) is observed monthly, the interest rate series (Treasury bill rate) is observed daily, and the income series (real GDP) is observed quarterly. The variables $m(t)$ and $r(t)$ are examples of stock variables, whereas $y(t)$ is clearly a flow, although there may, however, be some notion of averaging implicit in the construction of the money stock series. For this reason, quarterly averages of the money stock data were used, although different versions of the interest rate variable were considered. In particular, mid-quarter, end-quarter, and mean-quarter values of the interest rate variable were computed from the daily observations. The preceding theory suggests that the esti-

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

mates obtained using the latter averages are the most efficient asymptotically, and this empirical application enables the finite sample variances to be compared across the three different sampling schemes for the interest rate.

Two types of efficient estimator are considered. The first is the fully modified least squares (FM-OLS) estimator of Phillips and Hansen (1990), based on the triangular ECM representation, and the second is the Gaussian (maximum likelihood) estimator of Reinsel (1990) and Johansen (1991), based on the vector ECM (VECM) representation. The FM-OLS estimators were computed using the automatic bandwidth selection technique of Andrews (1991), based on an AR(1) approximation for the disturbances, to estimate the long-run variance matrix, for three different kernel functions, namely, Parzen, Tukey–Hamming, and Bartlett. These estimators are denoted FM-OLS(P), FM-OLS(TH), and FM-OLS(B) to distinguish the different underlying kernel functions. Using the VECM approach, two estimators were considered. The first uses the Schwartz information criterion to determine the autoregressive order of the system, which turns out to be 6. This estimator is denoted VECM(6). Inspection of the residuals in this system, however, suggests that they are serially correlated, and so the second estimator is based on a system with 12 lags and is denoted VECM(12). Although the second system is less parsimonious in modeling the dynamics, the residuals are closer to being white noise. The estimates of β_1 and β_2 obtained using each estimator are given in Table 3, which also contains the OLS estimates for comparison.

TABLE 3. Estimates of the long-run money demand relationship

	Mean-quarter		End-quarter		Mid-quarter	
Estimator	β_1	β_2	β_1	β_2	β_1	β_2
OLS	0.8934	−0.0236	0.8921	−0.0232	0.8920	−0.0231
	(0.0135)	(0.0052)	(0.0136)	(0.0052)	(0.0133)	(0.0051)
FM-OLS(P)	0.8452	−0.0219	0.8279	−0.0212	0.8293	−0.0216
	(0.0291)	(0.0112)	(0.0371)	(0.0143)	(0.0343)	(0.0131)
FM-OLS(TH)	0.8427	−0.0219	0.8445	−0.0208	0.8283	−0.0214
	(0.0293)	(0.0112)	(0.0399)	(0.0154)	(0.0347)	(0.0133)
FM-OLS(B)	0.8837	−0.0229	0.8814	−0.0227	0.8829	−0.0233
	(0.0572)	(0.0219)	(0.0600)	(0.0231)	(0.0594)	(0.0227)
VECM(6)	0.6484	−0.0193	0.5541	0.0079	0.6379	−0.0145
	(0.0610)	(0.0232)	(0.0719)	(0.0275)	(0.0629)	(0.0239)
VECM(12)	0.7147	−0.0308	0.7331	−0.0561	0.6993	−0.0333
	(0.0384)	(0.0146)	(0.0513)	(0.0197)	(0.0472)	(0.0180)

Note: Figures in parentheses are standard errors.

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

TABLE 4. Eigenvalues of covariance matrix differences

Estimator	End-quarter	Mid-quarter
FM-OLS(P)	5.95, 0.11	3.73, 0.07
FM-OLS(TH)	8.27, 0.15	3.84, 0.08
FM-OLS(B)	3.77, 0.07	2.80, 0.07
VECM(6)	16.42, 0.29	2.58, 0.07
VECM(12)	13.12, 0.24	8.50, 0.17

Note: All entries have been multiplied by 10^4.

The FM-OLS estimates in Table 3 are remarkably stable across the different treatments of the interest rate data. The estimates using different kernels are also very similar, particularly in the case of the Parzen and Tukey–Hamming kernels. The VECM estimates show more variation both across different treatments of the interest rate data and between different system orders. An interesting feature of the results in Table 3 is that in no case (with the exception of the inefficient OLS estimator) are the standard errors using end-quarter or mid-quarter interest rate data less than those using the mean-quarter data. The theoretical results, however, concern the positive semidefiniteness of the difference in covariance matrices, not a comparison of individual elements of the covariance matrices. With this in mind, we have constructed Table 4, which contains the eigenvalues of the differences in the covariance matrices of the estimators based on unaveraged interest rate data with respect to the averaged data (which the theory suggests should be positive semidefinite).

In all cases in Table 4, the relevant covariance matrix differences are positive semidefinite, as predicted by the asymptotic theory. This suggests that averaging stock data can lead to more precise estimates of cointegrating parameters in finite samples. These empirical results are only indicative, however, and should not be interpreted as meaning that such an outcome will always be achieved in finite samples, as the theory derived only relates to the asymptotic distribution of the estimators and not their finite sample distributions.

5. CONCLUDING COMMENTS

This paper has investigated a new avenue of research in the field of temporal aggregation in econometrics by considering the effects that the discrete time sampling of stock and flow variables has on the asymptotic efficiency of estimators in cointegrated systems. It is shown that estimators based on flow variables alone are more efficient than estimators based on stocks alone or on mixtures of stocks and flows. Furthermore, efficient estimators obtained with

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

flow variables are shown to be as efficient as the estimators that would be obtained using a continuous record of data, if such an estimator were feasible.

Although such findings are of academic interest, the situation faced by a researcher is typically one in which the method of sampling a variable is determined by its very nature. For example, it is not possible to sample the money stock as a flow variable even if it is desirable for improved estimator efficiency to do so. The paper therefore provides a practical method for improving efficiency with stock variables that can be implemented in certain cases, these being where the stock can be observed at a higher frequency than the other variables. The method involves discrete averaging of the higher frequency data over the unit time interval and is shown to improve the efficiency of the estimators. An empirical example using long-run money demand regressions is carried out to assess the extent to which this aggregation of stock variables can improve efficiency, and it is found that for a number of estimators the averaging of monthly money stock data and daily interest rate data in the regressions using quarterly data can improve efficiency. Monte Carlo evidence on this and other predictions of the theory is obtained and assessed in Chambers (2001).

NOTES

1. This is the sense in which the term *temporal aggregation* is used by Bergstrom (1984), Phillips (1991a), Drost and Nijman (1993), and Chambers (1998). Some authors, however, among them Weiss (1984) and Granger and Siklos (1995), use the term *temporal aggregation* to refer more explicitly to the sampling of flow data alone, and they employ the term *systematic sampling* for the case of stock variables.

2. There have been two main avenues of research in the temporal aggregation literature. The first examines the implications of sampling at discrete time intervals for the specification of econometric and time series models, and the second concentrates on methods of estimating the parameters of the underlying data generation process, based on a sample of observations at discrete time intervals. Weiss (1984) and Bergstrom (1984) are examples of the first line of research, whereas Bergstrom (1990), Harvey and Stock (1985), and Robinson (1993) cover various time and frequency domain methods of estimation in the second.

3. These proofs are also available electronically at http://privatewww.essex.ac.uk/~mchamb.

4. Phillips (1991a) uses the representation $Dy_2(t) = u_2(t)$ in place of (2), where D denotes the mean square differential operator, which requires the sample paths of $y_2(t)$ to satisfy certain smoothness conditions. The specification in (2), however, does not require $y_2(t)$ to be mean square differentiable (although not ruling out the possibility) but simply requires increments in $y_2(t)$ to be stationary.

5. The specification of the model is, of course, consistent with u_2 (and also u_1) being generated by a stochastic differential equation system. Chambers (2000) deals with such a fully parametric continuous time cointegrated system.

6. See Phillips (1991a) for details. The matrix J in (3) is denoted E by Phillips (1991a), but J is used to denote this matrix here so as to avoid any confusion with the expectation operator.

7. The situation where y_1 and y_2 are each combinations of stocks and flows can, of course, also be analyzed using the methods in this paper. See Chambers (2000) for an example.

8. This follows from, for example, Theorem 14.1 of Davidson (1994) or Theorem 3.49 of White (1984).

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

9. The following representations are based on Definition 2.1 of Saikkonen (1991), to which the vec operator has been applied.

10. Details of the equivalence of the representations (8)–(10) can be found in Phillips (1989).

11. See Bergstrom (1984) for a definition of wide sense integration of continuous time random processes.

12. See, for example, Hannan (1970, p. 45) for details.

13. I am grateful to a co-editor and a referee for emphasizing these points.

REFERENCES

Ahn, S.K. & G.C. Reinsel (1990) Estimation for partially nonstationary multivariate autoregressive models. *Journal of the American Statistical Association* 85, 813–823.

Andrews, D.W.K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–858.

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. 1145–1212. Amsterdam: North-Holland.

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

Chambers, M.J. (1998) Long memory and aggregation in macroeconomic time series. *International Economic Review* 39, 1053–1072.

Chambers, M.J. (2000) Frequency Domain Gaussian Estimation of Temporally Aggregated Cointegrated Systems. Manuscript, University of Essex.

Chambers, M.J. (2001) Temporal aggregation and the finite sample performance of spectral regression estimators in cointegrated systems: A simulation study. *Economic Theory* 17, 591–607.

Davidson, J.E.H. (1994) *Stochastic Limit Theory.* Oxford: Oxford University Press.

Drost, F.C. and T.E. Nijman (1993) Temporal aggregation of GARCH processes. *Econometrica* 61, 909–927.

Granger, C.W.J. & P.L. Siklos. (1995) Systematic sampling, temporal aggregation, seasonal adjustment, and cointegration: Theory and evidence. *Journal of Econometrics* 66, 357–369.

Hannan, E.J. (1970) *Multiple Time Series*. New York: Wiley.

Harvey, A.C. & J.H. Stock (1985) The estimation of higher-order continuous time autoregressive models. *Econometric Theory* 1, 97–118.

Johansen, S. (1991) Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* 59, 1551–1580.

Phillips, P.C.B. (1989) Partially identified econometric models. *Econometric Theory* 5, 181–240.

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

Phillips, P.C.B. (1991b) Optimal inference in cointegrated systems. *Econometrica* 59, 283–306.

Phillips, P.C.B. & S.N. Durlauf (1986) Multiple time series regression with integrated processes. *Review of Economic Studies* 53, 473–495.

Phillips, P.C.B. & B.E. Hansen (1990) Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* 57, 99–125.

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: Essays in Honour of A.R. Bergstrom*, pp. 71–90. Oxford: Blackwell.

Saikkonen, P. (1991) Asymptotically efficient estimation of cointegration regressions. *Econometric Theory* 7, 1–21.

Weiss, A.A. (1984) Systematic sampling and temporal aggregation in time series models. *Journal of Econometrics* 26, 271–281.

White, H. (1984) *Asymptotic Theory for Econometricians*. New York: Academic Press.

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

APPENDIX A

The following lemmas, A1–A7, are used in the proofs of the lemmas, theorems, and propositions in the main text. Their proofs are contained in a supplementary document that may be obtained from the author on request.

LEMMA A1. *The disturbance vector* $x(t) \equiv [x_1(t)', x_2(t)']'$ *defined in (5) is related to the continuous time disturbance process* $u(t)$ *by*

$$x_1(t) = u_1(t) + B\int_0^1 u_2(t-s)ds, \qquad x_2(t) = \int_0^1 u_2(t-s)ds.$$

LEMMA A2. *Let* $y_2(t)$ *be a continuous time random variable that evolves according to the law of motion* $dy_2(t) = u_2(t)dt$, *where* $u_2(t)$ *is a stationary continuous time random process. Then*

$$y_2(t) - \int_0^1 y_2(t-s)ds = \int_0^1 (1-s)u_2(t-s)ds.$$

LEMMA A3. *Let* $f(s)$ *be a continuous function on* $s \in [0,1]$ *and let* $u(t)$ *be a stationary continuous time scalar random process. Then*

$$E\left|\int_0^1 f(s)u(t-s)ds\right|^\beta \le E|u(t)|^\beta\left[\int_0^1 |f(s)|ds\right]^\beta,$$

where β *is a finite positive constant.*

LEMMA A4. *Let* $u(t)$ *denote a wide sense integrable continuous time random process. Then*

$$\int_0^1 u(t-s)ds = g(D)u(t) \quad \textit{and} \quad \int_0^1 (1-s)u(t-s)ds = k(D)u(t),$$

where $g(z) = (1-e^{-z})/z$ *and* $k(z) = [1-g(z)]/z$.

LEMMA A5. *The functions* $g(z) = (1-e^{-z})/z$ *and* $h(z) = e^{-z}[1-g(z)]/z$ *satisfy*

$$\lim_{z\to 0} g(z) = 1 \quad \textit{and} \quad \lim_{z\to 0} h(z) = \frac{1}{2},$$

whereas for integer $j \neq 0$, $g(2\pi ij) = 0$ *and* $h(2\pi ij) = (2\pi ij)^{-1}$.

LEMMA A6. *Let* $y_2(t)$ *be a continuous time random variable that evolves according to the law of motion* $dy_2(t) = u_2(t)dt$, *where* $u_2(t)$ *is a stationary continuous time random process. Then, for some positive integer* p,

$$\frac{1}{p}\sum_{k=0}^{p-1} y_2\left(t-\frac{k}{p}\right) - \int_0^1 y_2(t-s)ds = \frac{1}{p}\sum_{k=0}^{p-1}\int_{k/p}^1 u_2(t-s)ds - \int_0^1 su_2(t-s)ds.$$

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

In terms of filters, this expression has the representation

$$\frac{1}{p}\sum_{k=0}^{p-1} y_2\left(t-\frac{k}{p}\right)-\int_0^1 y_2(t-s)ds = m_p(D)u_2(t),$$

where $m_p(z) = [\sigma_p(z) - g(z)]/z$.

LEMMA A7. *The function* $h_p(z) = e^{-z}[\sigma_p(z) - g(z)]/z$ *satisfies* $\lim_{z\to 0} h_p(z) = 1/(2p)$, *while for integer* $j \neq 0$,

$$h_p(2\pi ij) = \frac{\sigma_p(2\pi ij)}{2\pi ij} = \begin{cases} \dfrac{1}{2\pi ij}, & j = 0, \pm p, \pm 2p, \ldots, \\ 0, & \text{otherwise.} \end{cases}$$

APPENDIX B

Outline proofs of the lemmas, theorems, and propositions given in the main text are provided here. In many cases the outline is given only for sampling scheme I, with the proofs for the remaining schemes following in a similar fashion. More detailed proofs, covering all the sampling schemes, are available from the author on request.

Proof of Lemma 1. We discuss each sampling scheme in turn.

Scheme I. The expressions are obtained directly from partitioning (4), which gives $\xi_{1t} = x_1(t)$ and $\xi_{2t} = x_2(t)$, and then applying Lemma A1.

Scheme II. Taking the first m_1 equations of (4) gives

$$\Delta y_{1t}^S = -y_{1,t-1}^S + By_2^F(t-1) + x_1(t) = -y_{1,t-1}^S + By_{2,t-1}^F + \xi_{1t},$$

with $\xi_{1t} = x_1(t) + B[y_2^F(t-1) - y_{2,t-1}^F]$. The expression for ξ_{1t} in Lemma 1 results from the form of $x_1(t)$ given in Lemma A1 and using Lemma A2 for the expression

$$y_2^F(t-1) - y_{2,t-1}^F = y_2^F(t-1) - \int_0^1 y_2^F(t-1-s)ds.$$

The expression for ξ_{2t} is obtained from the last m_2 equations of (7), so that $\xi_{2t} = \int_0^1 x_2(t-s)ds$. Lemma A1 is again applied to relate ξ_{2t} to $u_2(t)$.

Scheme III. The first m_1 equations of (7) yield

$$\Delta y_{1t}^F = -y_{1,t-1}^F + B\int_0^1 y_2^S(t-1-s)ds + \int_0^1 x_1(t-s)ds = -y_{1,t-1}^F + By_{2,t-1}^S + \xi_{1t},$$

where $\xi_{1t} = \int_0^1 x_1(t-s)ds + B[\int_0^1 y_2^S(t-1-s)ds - y_{2,t-1}^S]$. Application of Lemmas A1 and A2 yields the desired expression for ξ_{1t}, and the expression for ξ_{2t} is obtained directly from the last m_2 equations of (4) as in Scheme I.

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

Scheme IV. The expressions for ξ_{1t} and ξ_{2t} come directly from partitioning (7), so that $\xi_{1t} = \int_0^1 x_1(t-s)ds$ and $\xi_{2t} = \int_0^1 x_2(t-s)ds$. Lemma A1 once more yields the result. ■

Proof of Lemma 2. The stationarity of ξ_t follows directly from the definitions of ξ_t in terms of $u(t)$ in Lemma 1. The invariance principle is justified by verifying that the conditions of Corollary 2.2 of Phillips and Durlauf (1986) are satisfied. The required conditions are that (i) $E(\xi_t) = 0$; (ii) $E|\xi_{it}|^\beta < \infty$ for some $2 \leq \beta < \infty$; and (iii) the mixing conditions in Assumption 1(c) also hold for ξ_t. Condition (i) is satisfied by Assumption 1(a) and Lemma 1. Condition (iii) holds as a result of the mixing properties in Assumption 1(c) and the fact that ξ_t is a measurable function of $u(t)$ involving finite lags, which follows from the definitions in Lemma 1 and application of Theorem 14.1 of Davidson (1994) or Theorem 3.49 of White (1984). It remains to demonstrate that (ii) holds for each of the four sampling schemes. Details are provided here for Scheme I, and the results for the remaining sampling schemes follow in a similar fashion.

Scheme I. From Lemma 1 it follows that

$$
\begin{aligned}
E|\xi_{1i,t}|^\beta &= E\left| u_{1i}(t) + \sum_{j=1}^{m_2} B_{ij} \int_0^1 u_{2j}(t-s)ds \right|^\beta \\
&\leq (m_2+1)^{\beta-1}\left[E|u_{1i}(t)|^\beta + \sum_{j=1}^{m_2} E\left| B_{ij}\int_0^1 u_{2j}(t-s)ds \right|^\beta \right] \\
&\leq (m_2+1)^{\beta-1}\left[E|u_{1i}(t)|^\beta + \sum_{j=1}^{m_2} |B_{ij}|^\beta E\left| \int_0^1 u_{2j}(t-s)ds \right|^\beta \right] \\
&\leq (m_2+1)^{\beta-1}\left[E|u_{1i}(t)|^\beta + \sum_{j=1}^{m_2} |B_{ij}|^\beta E|u_{2j}(t)|^\beta \right] < \infty
\end{aligned}
$$

under Assumptions 1(b) and 2 and where the first inequality arises from the c_r-inequality and the third from application of Lemma A3. Also,

$$
E|\xi_{2i,t}|^\beta = E\left| \int_0^1 u_{2i}(t-s)ds \right|^\beta \leq E|u_{2i}(t)|^\beta < \infty
$$

by Assumption 1(b) and where Lemma A3 has been applied. ■

Proof of Lemma 3. The filters are derived directly from the expressions in Lemma 1, which relate ξ_t to $u(t)$, along with the results in Lemma A4. Details are provided here for Scheme I, and the results for the remaining sampling schemes follow in a similar fashion.

Scheme I. From Lemma 1 and application of Lemma A4,

$$
\xi_{1t} = u_1(t) + B\int_0^1 u_2(t-s)ds = u_1(t) + Bg(D)u_2(t),
$$

$$
\xi_{2t} = \int_0^1 u_2(t-s)ds = g(D)u_2(t),
$$

resulting in the form of $M(z)$ in the lemma. ■

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

Proof of Theorem 1. To simplify the presentation of the required summations, the notation

$$S(X_1, f, X_2) = \sum_{j=-\infty}^{+\infty} X_1(2\pi ij) f(2\pi j) X_2(-2\pi ij)$$

is utilized, where $X_1(\cdot)$ (of dimension $l \times m$) and $X_2(\cdot)$ (of dimension $m \times n$) are matrix functions and $f(\cdot)$ (of dimension $m \times m$) represents a spectral density matrix or a submatrix thereof. In each of the four sampling schemes, the long-run variance matrix of interest is given by $\Omega = 2\pi f_{\xi\xi}(0) = 2\pi S(M, f, M')$, where $M(z)$ is defined in Lemma 3. The submatrices of Ω yield $\Omega_{11.2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$. Some of the summations involving the functions $g(z)$ and $h(z)$ simplify in view of Lemma A5. For example, $S(g, f, g) = f(0) = S(g^2, f, g^2)$, $S(g, f, h) = f(0)h(0) = (\frac{1}{2})f(0)$, and $S(h, f, h) = (\frac{1}{4})f(0) + \sum_{j\neq 0}|h(2\pi ij)|^2 f(2\pi ij)$. Details are provided here for Scheme I, and the results for the remaining sampling schemes follow in a similar fashion.

Scheme I. From the definition of $M(z)$, the following submatrices of Ω are obtained:

$$\begin{aligned}\Omega_{11} &= 2\pi[S(I_{m_1}, f_{11}, I_{m_1}) + BS(g, f_{21}, I_{m_1}) + S(I_{m_1}, f_{12}, g)B' + BS(g, f_{22}, g)B'] \\ &= 2\pi\left[\sum_{j=-\infty}^{\infty} f_{11}(2\pi j) + Bf_{21}(0) + f_{21}(0)'B' + Bf_{22}(0)B'\right];\end{aligned}$$

$$\Omega_{12} = 2\pi[S(I_{m_1}, f_{12}, g) + BS(g, f_{22}, g)] = 2\pi[\, f_{12}(0) + Bf_{22}(0)];$$

$$\Omega_{22} = 2\pi S(g, f_{22}, g) = 2\pi f_{22}(0).$$

Combining these expressions gives

$$\Omega_{12}\Omega_{22}^{-1}\Omega_{21} = 2\pi[\, f_{12}(0) f_{22}(0)^{-1} f_{21}(0) + Bf_{21}(0) + f_{21}(0)'B' + Bf_{22}(0)B'],$$

and hence, because $\Omega_{11.2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$, it follows that

$$\begin{aligned}\Omega_{11.2} &= 2\pi\left[\sum_{j\neq 0} f_{11}(2\pi j) + f_{11}(0) + Bf_{21}(0) + f_{21}(0)'B' + Bf_{22}(0)B'\right] \\ &\quad - 2\pi[\, f_{12}(0) f_{22}(0)^{-1} f_{21}(0) + Bf_{21}(0) + f_{21}(0)'B' + Bf_{22}(0)B'] \\ &= 2\pi\left[\sum_{j\neq 0} f_{11}(2\pi j) + f_{11.2}(0)\right],\end{aligned}$$

as given in the theorem. ■

Proof of Proposition 1. The validity of Proposition 1 is established by considering, in turn, the three matrix differences $\Omega_{11.2}^{j} - \Omega_{11.2}^{IV}$ for $j = I, II, III$ and demonstrating that each one is positive semidefinite. The first of these differences is

$$\Omega_{11.2}^{I} - \Omega_{11.2}^{IV} = 2\pi \sum_{j\neq 0} f(2\pi j),$$

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

which is clearly (Hermitian) positive definite in view of $f(\lambda)$ being a spectral density matrix. Next,

$$\Omega_{11.2}^{II} - \Omega_{11.2}^{IV} = 2\pi \sum_{j\neq 0} H(2\pi ij, B) f(2\pi j) H(2\pi ij, B)^*.$$

Because $f(\lambda)$ is Hermitian positive definite, $H(2\pi ij) f(2\pi j) H(2\pi ij)^*$ is clearly positive semidefinite for each j, and hence so is the preceding sum over $j \neq 0$. Finally,

$$\Omega_{11.2}^{III} - \Omega_{11.2}^{IV} = 2\pi B \sum_{j\neq 0} |h(2\pi ij)|^2 f_{22}(2\pi j) B'.$$

In view of Lemma A5, $|h(2\pi ij)|^2 = (2\pi ij)^{-1}(-2\pi ij)^{-1} = (4\pi^2 j^2)^{-1} > 0$ for all $j \neq 0$, and because $f_{22}(\lambda)$ is Hermitian positive definite, the preceding sum is positive semidefinite. ■

Proof of Theorem 2. The proof follows straightforwardly from the definition of an optimal estimator by evaluating the long-run variance matrix $\tilde{\Omega} = 2\pi f_{ww}(0)$ and demonstrating that it is equal to $2\pi f(0)$. Because $w(t)dt = \Gamma(D)u(t)dt$, where

$$\Gamma(z) = \begin{bmatrix} (1+z)I_{m_1} & B \\ 0 & I_{m_2} \end{bmatrix},$$

it follows that $f_{ww}(\lambda) = \Gamma(i\lambda) f(\lambda) \Gamma(i\lambda)^*$ $(-\infty < \lambda < \infty)$. Evaluation of the submatrices of $f_{ww}(\lambda)$ yields

$$\tilde{\Omega}_{11} = 2\pi f_{ww,11}(0) = 2\pi[\, f_{11}(0) + B f_{21}(0) + f_{12}(0)B' + B f_{22}(0) B'\,],$$

$$\tilde{\Omega}_{12} = 2\pi f_{ww,12}(0) = 2\pi[\, f_{12}(0) + B f_{22}(0)],$$

$$\tilde{\Omega}_{22} = 2\pi f_{ww,22}(0) = 2\pi f_{22}(0).$$

The expression for $\tilde{\Omega}_{11.2}$ follows straightforwardly. ■

Proof of Proposition 2. This follows directly from the fact that the variance matrices in the limiting distributions of the optimal estimators in Scheme IV (Theorem 1) and in the continuous record case (Theorem 2) are identical. ■

Proof of Lemma 4. We consider each of the three sampling schemes in turn.

Scheme I(p). Here, $y_{t,p} = (1/p)\sum_{k=0}^{p-1} y(t - k/p)$, and so the discrete time ECM is obtained by applying the lag operator function $(1/p)\sum_{k=0}^{p-1} L^{(k/p)}$ to (4), yielding the ECM representation $\Delta y_{t,p} = -JAy_{t-1,p} + \xi_{t,p}$, where $\xi_{t,p} = (1/p)\sum_{k=0}^{p-1} x(t - k/p)$. Application of Lemma A1 for $x(t)$ yields the required expressions.

Scheme II(p). Taking the first m_1 equations of the equation in Scheme I(p) gives

$$\begin{aligned}\Delta y_{1t,p}^S &= -y_{1,t-1,p}^S + B\frac{1}{p}\sum_{k=0}^{p-1} y_2^F\left(t - 1 - \frac{k}{p}\right) + \frac{1}{p}\sum_{k=0}^{p-1} x_1\left(t - \frac{k}{p}\right) \\ &= -y_{1,t-1,p}^S + B y_{2,t-1}^F + \xi_{1t,p},\end{aligned}$$

where $\xi_{1t,p} = (1/p)\sum_{k=0}^{p-1} x_1(t - k/p) + B[(1/p)\sum_{k=0}^{p-1} y_2^F(t - 1 - k/p) - y_{2,t-1}^F]$. Applying Lemmas A1 and A6 yields the required expression. The vector $\xi_{2t,p}$ is obtained di-

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

rectly from the last m_2 equations of (7) and hence is the same as the expression in Scheme II of Lemma 1.

Scheme III(p). The first m_1 equations of (7) yield

$$\Delta y_{1t}^F = -y_{1,t-1}^F + B\int_0^1 y_2^S(t-1-s)ds + \int_0^1 x_1(t-s)ds$$

$$= -y_{1,t-1}^F + By_{2,t-1,p} + \xi_{1t,p},$$

where $\xi_{1t,p} = \int_0^1 x_1(t-s)ds + B[\int_0^1 y_2^S(t-1-s)ds - y_{2,t-1,p}^S]$ and the result follows by using Lemmas A1 and A6. The expression for $\xi_{2t,p}$ comes from the last m_2 equations of the expression in Scheme I(p), which describe the evolution of $y_{2t,p}^S$. ■

Proof of Lemma 5. We take each sampling scheme in turn.

Scheme I(p). Because $\xi_{t,p} = \sigma_p(D)\xi_t$, where ξ_t is from Scheme I in Lemma 1, it follows immediately that $\xi_{t,p} = \sigma_p(D)M(D)u(t)$, with $M(z)$ defined in Scheme I of Lemma 3, because $\xi_t = M(D)u(t)$.

Scheme II(p). From the definition of $\xi_{1t,p}$ it follows, using Lemma A6, that

$$\xi_{1t,p} = \sigma_p(D)u_1(t) + B\sigma_p(D)g(D)u_2(t) + Bm_p(D)e^{-D}u_2(t),$$

and because $h_p(z) = e^{-z}m_p(z)$, the filter on $u_2(t)$ is $\sigma_p(D)g(D) + h_p(D)$. The vector $\xi_{2t,p}$ is equal to ξ_{2t} in Scheme II, and so it follows immediately that $\xi_{2t,p} = g(D)^2u_2(t)$. The expression for $M_p(z)$ follows by combining these results.

Scheme III(p). Again from the definition of $\xi_{1t,p}$ and using Lemma A6,

$$\xi_{1t,p} = g(D)u_1(t) + Bg(D)^2u_2(t) - Be^{-D}m_p(D)u_2(t),$$

and the expression in the lemma follows by noting that $h_p(z) = e^{-z}m_p(z)$. Finally, the vector $\xi_{2t,p}$ is equivalent to $\xi_{2t,p}$ in Scheme I(p), and the expression for $M_p(z)$ then follows immediately. ■

Proof of Theorem 3. The proof follows the same approach as the proof of Theorem 1, with $M_p(z)$ replacing $M(z)$. Expressions involving the filter function $\sigma_p(z)$ simplify in view of $\sigma_p(2\pi ij) = 1$ for $j = 0, \pm p, \pm 2p, \ldots$ and $\sigma_p(2\pi ij) = 0$ otherwise. For example, $S(\sigma_p, f, \sigma_p) = \sum_{j=-\infty}^{\infty} f(2\pi jp)$. An outline of the derivations for Scheme I(p) is provided here. The results for the remaining sampling schemes follow in a similar fashion.

Scheme I(p). The expressions of interest are

$$\Omega_{11} = 2\pi[S(\sigma_p, f_{11}, \sigma_p) + Bf_{21}(0) + f_{21}(0)'B' + Bf_{22}(0)B'];$$

$$\Omega_{12} = 2\pi[S(\sigma_p g, f_{12}, \sigma_p g) + BS(\sigma_p g, f_{22}, \sigma_p g)] = 2\pi f_{12}(0) + Bf_{22}(0);$$

$$\Omega_{22} = 2\pi S(\sigma_p g, f_{22}, \sigma_p g) = 2\pi f_{22}(0).$$

Combining these expressions gives

$$\Omega_{12}\Omega_{22}^{-1}\Omega_{21} = 2\pi[\, f_{12}(0)f_{22}(0)^{-1}f_{21}(0) + Bf_{21}(0) + f_{21}(0)'B' + Bf_{22}(0)B']$$

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press

and hence

$$\Omega_{11.2} = 2\pi[S(\sigma_p, f_{11}, \sigma_p) - f_{12}(0)f_{22}(0)^{-1}f_{21}(0)] = 2\pi\left[\sum_{j\neq 0} f_{11}(2\pi jp) + f_{11.2}(0)\right],$$

because $S(\sigma_p, f_{11}, \sigma_p) = f_{11}(0) + \sum_{j\neq 0} f_{11}(2\pi jp)$. ■

Proof of Proposition 3. As the expressions for the Ω_{22} matrices are identical in Theorems 1 and 3, the proof proceeds by comparing the $\Omega_{11.2}$ matrices defined in these theorems for the cases with unadjusted and adjusted data. Details are given for Scheme I(p); those for Schemes II(p) and III(p) follow in a similar fashion.

Scheme I(p). The expressions in the two theorems give

$$\begin{aligned}\Omega^{I}_{11.2} - \Omega^{I(p)}_{11.2} &= 2\pi\left[\sum_{j\neq 0} f_{11}(2\pi j) + f_{11.2}(0)\right] - 2\pi\left[\sum_{j\neq 0} f_{11}(2\pi jp) + f_{11.2}(0)\right] \\ &= 2\pi \sum_{j\in J_p} f_{11}(2\pi j),\end{aligned}$$

where $J_p = \{j : j \in \mathcal{N}, j \neq 0, \pm p, \pm 2p, \ldots\}$ and $\mathcal{N} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ denotes the set of positive and negative integers, including zero. Because $f(\lambda)$ is Hermitian positive definite by assumption, it follows that this matrix difference is positive definite. ■

https://doi.org/10.1017/S0266466603191037 Published online by Cambridge University Press