

# COINTEGRATION ANALYSIS WITH MIXED-FREQUENCY DATA

BYEONGCHAN SEONG  
SUNG K. AHN  
PETER A. ZADROZNY

CESIFO WORKING PAPER NO. 1939  
CATEGORY 10: EMPIRICAL AND THEORETICAL METHODS  
MARCH 2007

*An electronic version of the paper may be downloaded*

- *from the SSRN website:* [www.SSRN.com](http://www.SSRN.com)
- *from the RePEc website:* [www.RePEc.org](http://www.RePEc.org)
- *from the CESifo website:* [www.CESifo-group.de](http://www.CESifo-group.de)

# COINTEGRATION ANALYSIS WITH MIXED-FREQUENCY DATA

## Abstract

We develop a method for directly modeling cointegrated multivariate time series that are observed in mixed frequencies. We regard lower-frequency data as regularly (or irregularly) missing and treat them with higher-frequency data by adopting a state-space model. This utilizes the structure of multivariate data as well as the available sample information more fully than the methods of transformation to a single frequency, and enables us to estimate parameters including cointegrating vectors and the missing observations of low-frequency data and to construct forecasts for future values. For the maximum likelihood estimation of the parameters in the model, we use an expectation maximization algorithm based on the state-space representation of the error correction model. The statistical efficiency of the developed method is investigated through a Monte Carlo study. We apply the method to a mixed-frequency data set that consists of the quarterly real gross domestic product and the monthly consumer price index.

JEL Code: C13, C22, C32.

Keywords: missing data, Kalman filter, expectation maximization algorithm, forecasting, error correction model, smoothing, maximum likelihood estimation.

*Byeongchan Seong*  
*Department of Statistics*  
*Chung-Ang University*  
*Seoul 156-756*  
*Korea*  
*bcseong@cau.ac.kr*

*Sung K. Ahn*  
*Department of Management and*  
*Operations*  
*Washington State University*  
*USA – Pullman, WA 99164*  
*ahn@wsu.edu*

*Peter A. Zadrozny*  
*Bureau of Labor Statistics*  
*2 Massachusetts Ave., NE*  
*USA – Washington, DC 20212*  
*zadrozny.peter@bls.gov*

February 27, 2007

B. Seong's research was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD, Basic Research Promotion Fund) (KRF-2006-003-C00052). S.K. Ahn's work was supported by the Korea Research Foundation Grant (KRF-2005-070-C00022) funded by the Korean Government (MOEHRD). P.A. Zadrozny is affiliated as Research Fellow with the Center for Economic Studies and Ifo Institute for Economic Research (CESifo), Munich, Germany, and with the Center for Financial Studies (CFS), Goethe University, Frankfurt, Germany. The paper represents the authors' views and does not necessarily represent any official positions of the Bureau of Labor Statistics.

## 1. Introduction

Multivariate time series that arise in economics and business are often observed in mixed frequencies. For example, data available from the database of the Bureau of Economic Analysis, the Bureau of Labor Statistics, and the Bureau of the Census are often in mixed frequencies, mostly with quarterly, monthly or weekly sampling intervals. Although cointegration, which represents a long-run equilibrium among the components of nonstationary multivariate time series, has been one of the most extensively investigated research topics, especially, in economics and business during the past two decades since Engle and Granger (1987), studies of cointegration have been limited to the case where all the components of a multivariate series are observed at the same frequency.

Data observed in mixed frequencies are usually transformed to a single frequency by temporally aggregating higher-frequency data to lower frequencies, or by interpolating lower-frequency data to higher frequencies. However, temporal aggregation destroys sample information (Zadrozny, 1990). Granger and Siklos (1995) examined the misinterpretation of the long-run component of a time series constructed by temporal aggregation. Marcellino (1999) theoretically derived the effects of temporal aggregation on cointegration such as the asymptotic invariance of cointegrating (CI) rank and vectors. However, he illustrated some possibilities of the loss of power of cointegration tests due to a decline in the number of available observations. Haug (2002) showed in Monte Carlo experiments with various cointegration tests and data generating processes (DGPs) that these power losses indeed occur and assessed their extents in samples of typical size used in empirical work.

For the case of interpolation of mixed-frequency data, Chow and Lin (1971, 1976) used the conventional regression approach. Other researchers, for example, Bernanke *et al.* (1997), Cuche and Hess (2000), and Liu and Hall (2001), used the state-space framework suggested by Harvey and Pierse (1984) where interpolation is based on univariate regression. When the goal is to estimate a multivariate model for forecasting or other purposes, this kind of interpolation is at best an intermediate nuisance and at worst a source of distortion in the data to be used for estimation (Zadrozny, 1990, p. 2).

Recently, Mariano and Murasawa (2003, 2004) considered multivariate models for constructing a new index of economic indicators using mixed-frequency data, which overcome the drawbacks of a univariate approach and exploit the cross-frequency sample information. However, they used differenced data instead of levels data. This causes a loss of information on the long-run dynamics among the variables.

In this paper we develop a method for “directly” modeling the cointegrated multivariate time series with mixed-frequency data, which is based on the state-space representation of the error correction model (ECM). We use the state-space formulation for mixed-frequency data in our development by fully utilizing the

structure of multivariate data as well as the available sample information. For the analysis, we exploit an expectation maximization (EM) algorithm. The method is applied to estimate the parameters and the missing observations of the low-frequency variables and to construct forecasts of their future values.

We introduce notation which is used in this paper:  $I_m$  denotes an  $m \times m$  identity matrix;  $O_m$ ,  $O_{m \times n}$ , and  $0_m$  denote an  $m \times m$  zero matrix, an  $m \times n$  zero matrix, and an  $m \times 1$  zero vector;  $\text{vec}(\cdot)$  vectorizes a matrix columnwise from left to right, and  $\text{vech}(\cdot)$  vectorizes the lower triangular part of a matrix columnwise; the symbol  $\otimes$  denotes the Kronecker product.

The structure of the paper is as follows. In Section 2, we describe preliminary concepts for mixed-frequency data and set up the state-space representation for the ECM of cointegration. In Section 3, we develop the procedures for estimating the parameters of the cointegrated model using the EM algorithm and discuss initial conditions for the procedures. In Section 4, we consider models with different types of deterministic terms. In Sections 5, we comment on smoothing, forecasting, and logarithmic transformation. In Section 6, we conduct Monte Carlo experiments for the investigation of the performance of the developed method. In Section 7, we consider a numerical example to illustrate the method and we conclude the paper in Section 8.

## 2. Model with mixed-frequency data

In this section, we define mixed-frequency data and the ECM for multivariate cointegrated time series.

### 2.1 Mixed-frequency data

We define high-frequency variables as those observed at the fundamental shortest sampling interval and define low-frequency variables as those observed at longer sampling intervals, either as temporal aggregates or as skip samples of their high-frequency values. A variable is skip sampled when, for example, it is generated every month but is sampled every third month, say, in the last month of every quarter. See, for example, Zdrozny (1988). We assume that the underlying data generating process of a multivariate time series of mixed-frequency data, composed of both the high and the low-frequency variables, operates at the highest frequency, as in Zdrozny (1990), Mitnik and Zdrozny (2004) and Mariano and Murasawa (2003, 2004). All variables are assumed to be produced at the highest frequency, but some variables are not observed at the highest frequency. For example, we consider a bivariate time series of the consumer price index (CPI) observed monthly and the gross domestic product (GDP) observed quarterly. CPI is the high-frequency variable and GDP is the low-frequency variable. The fundamental interval or highest frequency is monthly and GDP is in

principle ‘produced’ monthly but is observed only quarterly. Variables like GDP observed as temporal aggregates are often called flows, while variables observed at the fundamental frequency are often called stocks.

## 2.2 Error correction model and state-space representation

Let  $u_t$  be an  $n$ -dimensional vector autoregressive process of order  $p$ ,  $\text{VAR}(p)$ , which operates at the highest frequency with CI rank  $h$ , and consider the corresponding error correction form

$$\Delta u_t = ab'u_{t-1} + \sum_{j=1}^{p-1} \Psi_j \Delta u_{t-j} + \varepsilon_t, \quad (1)$$

where  $a$  and  $b$  are  $n \times h$  matrices with  $0 < h < n$  and  $h = n - d$ ,  $\Psi_j$ , for  $j = 1, \dots, p-1$ , are  $n \times n$  matrices, and  $\varepsilon_t$  is an independent  $N_n(0, \Omega)$  random vector. We assume that the elements of  $u_t$  are ordered such that the last  $d = n - h$  elements are not cointegrated. This assumption permits the normalization  $b = [I_h \ \beta_0]'$  for identification, as in Ahn and Reinsel (1990), where  $\beta_0$  is a  $d \times h$  matrix. The characteristic equation of model (1) has exactly  $d$  roots equal to one and all other roots are assumed to be outside the unit circle, so that  $u_t$  is cointegrated of order (1,1) (Engle and Granger, 1987).

We reorder the elements of  $u_t$  to form  $z_t = (z'_{1t}, z'_{2t})'$  such that an  $n_1 \times 1$  vector  $z_{1t}$  corresponds to high-frequency variables and an  $n_2 \times 1$  vector  $z_{2t}$  corresponds to low-frequency variables, where  $n = n_1 + n_2$ . For brevity, we assume that the low-frequency variables are observed as temporal aggregates because the method can be easily modified to accommodate the alternative case of skip-sampled data. The reordering implies  $z_t = Vu_t$ , where  $V$  is an  $n \times n$  permutation matrix. Substituting  $u_t$  with  $V^{-1}z_t$  and pre-multiplying both sides of equation (1) by  $V$ , we can rewrite model (1) as

$$\Delta z_t = \alpha \beta' z_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta z_{t-j} + e_t, \quad (2)$$

where  $\alpha = Va$ ,  $\beta = Vb$ ,  $\Gamma_j = V\Psi_j V'$ ,  $e_t = V\varepsilon_t$ ,  $e_t \sim N(0, \Sigma)$ , and  $\Sigma = V\Omega V'$ . Also, note that  $V^{-1} = V'$ . For convenience of estimation, we write  $\beta = V_1 I_h + V_2 \beta_0$ , where  $V_1$  and  $V_2$  are  $n \times h$  and  $n \times d$  matrices, such that  $V = [V_1 \ V_2]$ . The  $\text{VAR}(p)$  representation of model (2) is

$$z_t = \sum_{j=1}^p \Phi_j z_{t-j} + e_t, \quad (3)$$

where  $\Phi_1 = I_n + \alpha \beta' + \Gamma_1$ ,  $\Phi_j = \Gamma_j - \Gamma_{j-1}$ , for  $j = 2, \dots, p-1$ , and  $\Phi_p = -\Gamma_{p-1}$ .

In practice, as mentioned above, the low-frequency variables,  $z_{2t}$ , are not observed directly at the highest frequency, but either as temporally-aggregated flows or as skip-sampled stocks. Flows can be expressed as  $y_{2t} = C_0 z_{2t} + \xi_t$ , where  $\xi_t = \sum_{j=1}^v C_j z_{2t-j}$  is the lagged part in a temporal aggregation. Here,  $v$  denotes the maximum degree of aggregation and the  $C_j$ 's are  $n_2 \times n_2$  diagonal indicator matrices with zeros and ones on the principal diagonal. By adjusting the diagonal elements of  $C_j$ , Zadrozny (1990) suggested a way to treat a variable which is observed directly as a stock, but with a delay.

We now construct a state-space representation of model (3). Let  $\tilde{z}_t = (z'_t, \xi'_t)' = (z'_{1t}, z'_{2t}, \xi'_t)'$  be an  $\tilde{n} \times 1$  vector with  $\tilde{n} = n_1 + 2n_2$ . We define the  $s \times 1$  state vector  $x_t = (\tilde{z}'_t, \dots, \tilde{z}'_{t-r+1})'$ , where  $r = \max(p, v)$  and  $s = \tilde{n}r$  for  $t = 1, \dots, T$ . Then, we define the state equation, as in Zadrozny (1990), by

$$x_t = Fx_{t-1} + Ge_t, \quad (4)$$

where the initial state,  $x_0$ , is assumed to be a normal random vector with mean vector  $\lambda$  and  $s \times s$  covariance matrix  $\Lambda$ . Here,  $F$  and  $G$  denote  $s \times s$  and  $s \times n$  matrices defined by

$$F = \begin{bmatrix} \Phi_1^* & \Phi_2^* & \dots & \Phi_r^* \\ I_{\tilde{n}} & O_{\tilde{n}} & \dots & O_{\tilde{n}} \\ O_{\tilde{n}} & I_{\tilde{n}} & \dots & O_{\tilde{n}} \\ \vdots & \vdots & \dots & \vdots \\ O_{\tilde{n}} & O_{\tilde{n}} & \dots & O_{\tilde{n}} \end{bmatrix}, \quad G = \begin{bmatrix} I_n \\ O_{(s-n) \times n} \end{bmatrix}, \quad \Phi_j^* = \begin{bmatrix} \Phi_j & O_{n \times n_2} \\ C_j^* & O_{n_2 \times n_2} \end{bmatrix}, \quad C_j^* = \begin{bmatrix} O_{n_2 \times n_1} & C_j \end{bmatrix},$$

$\Phi_j = O_n$ , if  $j > p$ , and  $C_j = O_{n_2}$  if  $j > v$ , so that  $\Phi_j^*$  and  $C_j^*$  are  $\tilde{n} \times \tilde{n}$  and  $n_2 \times n$  matrices.

Next, assuming no observation errors, we define the measurement equation

$$y_t \equiv \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} H_1 \\ H_2 \end{pmatrix} x_t, \quad (5)$$

where  $y_{1t} = z_{1t}$ ,  $H_1 \equiv [I_{n_1} \ O_{n_1 \times (s-n_1)}]$ , and  $H_2 \equiv [O_{n_2 \times n_1} \ C_0 \ I_{n_2} \ O_{n_2 \times (s-\tilde{n})}]$ . The matrix  $H_1$  picks out the high-frequency variables from state vector  $x_t$  and the matrix  $H_2$  picks out the low-frequency variables from the state vector. We introduce a new series,  $y_{2t}^+$ , which is observed only at lower frequencies, in order to deal with the missing observations in  $y_{2t}$ . As in Brockwell and Davis (1991), we fill in missing observations of  $y_{2t}$  with random vectors which are independent of  $y_t$  and are distributed independently of the parameters in model (2). Accordingly, we modify measurement equation (5) as

$$y_t^+ \equiv \begin{pmatrix} y_{1t} \\ y_{2t}^+ \end{pmatrix} = \begin{pmatrix} H_1 \\ H_{2t} \end{pmatrix} x_t + \begin{pmatrix} O_{n_1 \times n_2} \\ Q_{2t} \end{pmatrix} w_t, \quad (6)$$

$$H_{2t} = \begin{cases} H_2 & \text{if } y_{2t} \text{ is observable} \\ O_{n_2 \times s} & \text{otherwise} \end{cases}, \quad Q_{2t} = \begin{cases} O_{n_2} & \text{if } y_{2t} \text{ is observable} \\ I_{n_2} & \text{otherwise} \end{cases},$$

and  $w_t$  is an independent random vector distributed  $N_{n_2}(0, I_{n_2})$ . We also define  $H_t = [H_1', H_{2t}']'$  and  $Q_t = [O_{n_1 \times n_2}', Q_{2t}']'$ , which will appear in equations (16) and (20) and in Appendix B. In the implementation, because the realization of  $w_t$  is independent of  $y_t$ , setting  $w_t = 0$  is the preferred simple choice (Brockwell and Davis, 1991; Mariano and Murasawa, 2003). Instead of using  $w_t$ , a selection matrix may be used for constructing a measurement equation in order to control the mixed-frequency data, as in Zdrozny (1990).

We note that when the low-frequency variables are stocks, with missing data attributable to skip-sampling, we do not need to define  $y_{2t}$ ,  $\xi_t$ ,  $C_j$ , and  $\tilde{z}_t$ . Instead, we redefine  $\tilde{n} = n = n_1 + n_2$ ,  $r = \max(p, 2)$ ,  $\Phi_j^* = \Phi_j$ , and  $H_2 \equiv [O_{n_2 \times n_1} \quad I_{n_2} \quad O_{n_2 \times (s-n)}]$ . We need  $r = 2$  in order to construct the state equation even though the autoregressive order is one. Similar adjustments can be applied to more complicated cases, where the low-frequency variables are observed as both stocks and flows.

### 3. Maximum likelihood estimation of parameters

In this section, we consider maximum likelihood estimation (MLE) of parameters in error correction model (2) in the state-space form (4) and (6).

#### 3.1 EM algorithm

Dempster *et al.* (1977), Shumway and Stoffer (1982), and Watson and Engle (1983) developed and illustrated the EM algorithm for estimating a model in state-space form, when some variables are partly or completely unobserved (latent).

Let  $X_s = (x_t; 0 \leq t \leq s)$  and  $Y_s^+ = (y_t^+; 1 \leq t \leq s)$  be information sets. In order to develop an EM algorithm for estimating the parameters of the state-space model in (4) and (6), we consider several transformations for forming the likelihood function with respect to the complete data  $X_T$  and  $Y_T^+$ . Let

$$\Gamma = [\alpha \quad \Gamma_1 \quad \cdots \quad \Gamma_{p-1}], \quad \Gamma^* = [\alpha V_1' \quad \Gamma_1 \quad \cdots \quad \Gamma_{p-1}],$$

$$A = [I_n \quad O_{n \times n_2} \quad -I_n \quad O_{n \times (s-2n-n_2)}], \quad \text{and} \quad D = [V_2' \quad O_{d \times (s-n)}],$$

where  $Dx_{t-1} = V_2' z_{t-1}$ . Define

$$B = \begin{bmatrix} \beta' & O_{h \times n_2} & O_{h \times n} & O_{h \times n_2} & O_{h \times n} & \cdots \\ I_n & O_{n \times n_2} & -I_n & O_{n \times n_2} & O_n & \cdots \\ O_n & O_{n \times n_2} & I_n & O_{n \times n_2} & -I_n & \cdots \\ O_n & O_{n \times n_2} & O_n & O_{n \times n_2} & I_n & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \end{bmatrix},$$

$$B^* = \begin{bmatrix} I_n & O_{n \times n_2} & O_n & O_{n \times n_2} & O_n & \cdots \\ I_n & O_{n \times n_2} & -I_n & O_{n \times n_2} & O_n & \cdots \\ O_n & O_{n \times n_2} & I_n & O_{n \times n_2} & -I_n & \cdots \\ O_n & O_{n \times n_2} & O_n & O_{n \times n_2} & I_n & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \end{bmatrix},$$

where  $Bx_{t-1} = (z_{t-1}'\beta, \Delta z_{t-1}', \dots, \Delta z_{t-p+1}')'$ ,  $B^* x_{t-1} = (z_{t-1}', \Delta z_{t-1}', \dots, \Delta z_{t-p+1}')'$ , and  $B$  and  $B^*$  are  $\{h + (p-1)n\} \times s$  and  $np \times s$  matrices. We consider two transformations for forming the likelihood function and estimating  $\Gamma$ ,  $\Sigma$ , and  $\beta_0$ .

First,

$$Ax_t = \Delta z_t = \alpha \beta' z_{t-1} + \Gamma_1 \Delta z_{t-1} + \cdots + \Gamma_{p-1} \Delta z_{t-p+1} + e_t = \Gamma Bx_{t-1} + e_t, \quad (7)$$

$$\begin{aligned} Ax_t &= \alpha \beta_0' V_2' z_{t-1} + \alpha V_1' z_{t-1} + \Gamma_1 \Delta z_{t-1} + \cdots + \Gamma_{p-1} \Delta z_{t-p+1} + e_t \\ &= \alpha \beta_0' Dx_{t-1} + \Gamma^* B^* x_{t-1} + e_t \\ &= \{\alpha \otimes (Dx_{t-1})'\} \text{vec}(\beta_0) + \Gamma^* B^* x_{t-1} + e_t, \end{aligned} \quad (8)$$

because  $\alpha \beta_0' Dx_{t-1} = \text{vec}\{\alpha \beta_0' Dx_{t-1}\}$  and  $\text{vec}\{(Dx_{t-1})' \beta_0 \alpha'\} = \{\alpha \otimes (Dx_{t-1})'\} \text{vec}(\beta_0)$ , which are obtained using the vectorization rule  $\text{vec}(ABC) = (C' \otimes A) \text{vec}(B)$  (Magnus and Neudecker, 1988).

Then, we express the log-likelihood function as

$$\begin{aligned} \log L(\theta; X_T, Y_T^+) &= -\frac{1}{2} \log |\Lambda| - \frac{1}{2} (x_0 - \lambda)' \Lambda^{-1} (x_0 - \lambda) \\ &\quad - \frac{T}{2} \log |\Sigma| - \frac{1}{2} \sum_{t=1}^T (Ax_t - \Gamma Bx_{t-1})' \Sigma^{-1} (Ax_t - \Gamma Bx_{t-1}), \end{aligned} \quad (9)$$

or as

$$\begin{aligned} \log L(\theta; X_T, Y_T^+) &= -\frac{1}{2} \log |\Lambda| - \frac{1}{2} (x_0 - \lambda)' \Lambda^{-1} (x_0 - \lambda) \\ &\quad - \frac{T}{2} \log |\Sigma| - \frac{1}{2} \sum_{t=1}^T [Ax_t - \Gamma^* B^* x_{t-1} - \{\alpha \otimes (Dx_{t-1})'\} \text{vec}(\beta_0)]' \times \end{aligned}$$



$$\Sigma^{-1}[Ax_t - \Gamma^* B^* x_{t-1} - \{\alpha \otimes (Dx_{t-1})'\} \text{vec}(\beta_0)], \quad (10)$$

where terms that do not contain the parameters,  $\theta \equiv (\text{vec}(\beta_0)', \text{vec}(\Gamma)', \text{vech}(\Sigma)')$ , are omitted. Version (9) of the log-likelihood function is used to estimate the ‘‘stationary’’ parameters in  $\Gamma$  and  $\Sigma$ . Version (10) of the log-likelihood function is used to estimate the remaining ‘‘nonstationary’’ parameters in  $\beta_0$ . The distribution of  $w_t$  in the measurement equation does not have any effect on equations (9) and (10). Because the log likelihood function depends on the unobserved information,  $X_T$ , the EM algorithm is applicable, conditional on the observed information,  $Y_T^+$ . Specifically, we define the estimated parameters at iteration  $l+1$  as the value of  $\theta$  which maximizes

$$Q(\theta | \theta^{(l)}) = E_l \left\{ \log L(\theta; X_T, Y_T^+) | Y_T^+ \right\}, \quad (11)$$

where  $\theta^{(l)}$  denotes estimated  $\theta$  after  $l$  iterations and  $E_l\{\cdot | Y_T^+\}$  denotes the conditional expectation with respect to a density containing  $\theta^{(l)}$ , given  $Y_T^+$ . Then, using the derivatives of (11) with respect to  $\theta$  stated in Appendix A, we obtain the following equations for updating  $\theta^{(l+1)}$  at the end of  $l$  iterations,

$$\begin{aligned} \text{vec}(\beta_0^{(l+1)}) &= \left\{ \alpha^{(l)'} \Sigma^{(l)-1} \alpha^{(l)} \otimes (DM_{11}D') \right\}^{-1} \times \\ &\quad \text{vec} \left[ D \left\{ M_{10}A' - M_{11}(\Gamma^{*(l)}B^*)' \right\} \Sigma^{(l)-1} \alpha^{(l)} \right], \end{aligned} \quad (12)$$

$$\Gamma^{(l+1)} = (AM_{01}B'^{(l+1)}) \left( B^{(l+1)}M_{11}B'^{(l+1)} \right)^{-1}, \quad (13)$$

$$\Sigma^{(l+1)} = T^{-1} \left( AM_{00}A' - \Gamma^{(l+1)}B^{(l+1)}M_{10}A' \right), \quad (14)$$

where, in equation (12),  $\alpha^{(l)}$ ,  $\Sigma^{(l)}$ , and  $\Gamma^{*(l)}$  are given by the previous iteration; in equation (13),  $B^{(l+1)}$  is given by  $\beta_0^{(l+1)}$ , according to equation (12); and, in equation (14),  $\Gamma^{(l+1)}$  is given by equation (13). In equations (12) to (14),  $M_{00}$ ,  $M_{01}$ ,  $M_{10}$ , and  $M_{11}$  are given by

$$M_{jk} = \sum_{t=1}^T E_l(x_{t-j}x'_{t-k} | Y_T^+) = \sum_{t=1}^T (P_{t-j,t-k}^T + x_{t-j}^T x_{t-k}^T), \quad (15)$$

for  $j, k = 0, 1$ , where  $P_{t-j,t-k}^T$  and  $x_{t-j}^T$  are produced by the prediction and updating recursions of the Kalman filter. In Appendix B, we adopt the fixed-interval smoothing algorithm of De Jong (1989), which avoids inversion of large matrices, and, hence, is computationally more efficient than the classical smoothing equations (Durbin and Koopman, 2001). We note that the mean vector,  $\lambda$ , and the covariance matrix,  $\Lambda$ , of the initial state vector,  $x_0$ , cannot be estimated simultaneously. Following Shumway and Stoffer (1982, 2000), we

preset the covariance matrix and estimate the mean vector as  $\lambda^{(l+1)} = x_0^T$  by maximizing (11).

We summarize the iterative EM procedure as follows:

- (1) Calculate  $M_{jk}$  for  $j, k = 0, 1$  using equations (B1) to (B6) in Appendix B, with the initial values  $\lambda^{(0)}$ ,  $\Lambda$ , and  $\theta^{(0)}$ .
- (2) Estimate  $\lambda^{(1)} = x_0^T$  and calculate  $\theta^{(1)}$  using equations (12) to (14).
- (3) Iterate on steps (1) and (2) above until the parameter estimates or the likelihood values converge. At each iteration, we calculate the innovations form of the log-likelihood function (Schweppe, 1965),

$$\begin{aligned} \log L(\theta; Y_T^+) &= -\frac{1}{2} \sum_{t=1}^T \log |H_t P_t^{t-1} H_t' + Q_t Q_t'| \\ &\quad - \frac{1}{2} \sum_{t=1}^T (y_t - H_t x_t^{t-1})' (H_t P_t^{t-1} H_t' + Q_t Q_t')^{-1} (y_t - H_t x_t^{t-1}), \end{aligned} \quad (16)$$

and stop when the difference between  $\log L(\theta^{(l+1)}; Y_T^+)$  and  $\log L(\theta^{(l)}; Y_T^+)$  is less than a predetermined small value.

As an alternative to the EM algorithm, the Newton-Raphson (NR) method can be applied to maximize the likelihood. However, the NR method is more likely to fail because it is very sensitive to initial values; see Shumway and Stoffer (2000) for more comparisons between the NR and EM algorithms. Thus, we recommend using the EM algorithm especially because in cointegration with mixed-frequency data initial values of parameters are difficult to obtain.

### 3.2 Initialization

To start the EM algorithm with the Kalman filter, we need to specify the initial values,  $\lambda^{(0)}$ ,  $\Lambda$ , and  $\theta^{(0)}$ . When the state equation is nonstationary, the unconditional distribution of the state vector is not defined. Usually, the initial distribution of  $x_0$  must be specified by a diffuse or noninformative prior because genuine prior information is generally not available (Harvey, 1991). Therefore, we set  $\lambda^{(0)} = 0$  and  $\Lambda = \delta I$ , where  $\delta$  is a large value, for example,  $\delta = 10^8$ .

Regarding initial values of parameters, we first obtain an estimate of the nonstationary parameters by using single-frequency data that are usually obtained by transforming the high-frequency variables,  $y_{1t}$ , to match the low frequency of  $y_{2t}$ . Then, we use the relationship between the nonstationary parameters in mixed-frequency non-temporally-aggregated and single-frequency temporally-aggregated models, estimated using mixed- and single-frequency data, in order to

obtain an initial estimate of the nonstationary parameters in  $\beta_0^{(0)}$ . For example, see Marcellino (1999) and Pons and Sansó (2005) for the explicit formulas for the relationship in several cases. For the remaining stationary parameters, however, it is not easy to obtain an explicit formula to describe the relationship between the parameters of models of mixed- and single-frequency data. Therefore, treating the initial estimate  $\beta_0^{(0)}$  as known, and, thus, fixed in equation (2), we obtain initial estimates of the stationary parameters,  $\Gamma^{(0)}$  and  $\Sigma^{(0)}$ , using the iterative EM algorithm. Henceforth, we call this the “initial EM” algorithm. We use the output from the initial EM algorithm to produce initial values for the “main EM” algorithm which we are advocating. Accordingly, we obtain a better starting value  $\lambda^{(0)}$  for the main EM algorithm than the starting value  $\lambda^{(0)} = 0$  used for the initial EM algorithm.

In order to prespecify the CI rank in the analysis of cointegration with mixed-frequency data, we use the fact stated by Marcellino (1999) that the CI rank is invariant to temporal aggregation. Then, we can use the CI rank, obtained by applying the CI rank test to the temporally-aggregated single-frequency data.

We select as “best” the VAR( $p$ ) model whose MLE yields the lowest values of Akaike’s information criterion (AIC) and Schwartz’s Bayesian information criterion (SBC),

$$\text{AIC} = T^{-1} \{-2 \log L(\hat{\theta}; Y_T^+) + 2 \dim(\theta)\},$$

$$\text{SBC} = T^{-1} \{-2 \log L(\hat{\theta}; Y_T^+) + \dim(\theta) \log T\},$$

where  $\log L(\hat{\theta}; Y_T^+)$  is given by equation (16),  $\hat{\theta}$  is the MLE of  $\theta$ , obtained using the main EM algorithm, and  $\dim(\theta)$  is the dimension of  $\theta$ .

#### 4. Model with deterministic terms

In this section, we describe a state-space representation that can accommodate constant and linear trend terms in model (1). To this end, we consider the model

$$\Delta u_t = \gamma_1 + \gamma_2(t-1) + ab'u_{t-1} + \sum_{j=1}^{p-1} \Psi_j \Delta u_{t-j} + \varepsilon_t, \quad (17)$$

where  $\gamma_1$  and  $\gamma_2$  are  $n \times 1$  vectors. If  $\gamma_1$  and  $\gamma_2$  are restricted, they can be reparameterized as  $\gamma_j = a\rho_j$  and incorporated into the cointegration relationship as

$$a\rho_1' + a\rho_2'(t-1) + ab'u_{t-1} = a \begin{bmatrix} \rho_1' & \rho_2' & b' \end{bmatrix} \begin{pmatrix} 1 \\ t-1 \\ u_{t-1} \end{pmatrix}.$$

For brevity of exposition, we assume  $\gamma_1$  and  $\gamma_2$  are all unrestricted in this section. For cases with restrictions, see Appendix C, where we consider state-space representations for various combinations of deterministic terms, as in Johansen (1996). Then, model (2) is rewritten as

$$\Delta z_t = \mu_1 + \mu_2(t-1) + \alpha\beta'z_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta z_{t-j} + e_t, \quad (18)$$

where  $\mu_1 = V\gamma_1$  and  $\mu_2 = V\gamma_2$ . For model (18), we obtain a state-space representation by replacing state vector  $x_t$  with  $\tilde{x}_t = (1, t, x_t)'$ , so that state equation (4) becomes

$$\tilde{x}_t = \tilde{F}\tilde{x}_{t-1} + \tilde{G}e_t, \quad (19)$$

where

$$\tilde{F} = \begin{bmatrix} 1 & 0 & \vdots & O_{2 \times s} \\ 1 & 1 & \vdots & \\ \mu_1 & \mu_2 & \vdots & F \\ 0_{s-n} & 0_{s-n} & \vdots & \end{bmatrix}, \quad \tilde{G} = \begin{bmatrix} O_{2 \times n} \\ G \end{bmatrix},$$

and, correspondingly, measurement equation (6) becomes

$$y_t^+ = \tilde{H}_t \tilde{x}_t + Q_t w_t, \quad (20)$$

where  $\tilde{H}_t = [O_{n \times 2} \quad H_t]$ .

Using state-space representation (19), we can similarly get transformation (7) or (8). First, we define matrices

$$\tilde{A} = [O_{n \times 2} \quad A], \quad \tilde{\Gamma} = [\mu_1 \quad \mu_2 \quad \Gamma], \quad \tilde{\Gamma}^* = [\mu_1 \quad \mu_2 \quad \Gamma^*],$$

$$B_\rho = \begin{bmatrix} 0_h & 0_h & \vdots \\ \vdots & \vdots & \vdots \\ O_{(p-1)n \times 2} & \vdots & B \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} I_2 & O_{2 \times s} \\ \vdots & \vdots \\ B_\rho \end{bmatrix}.$$

Then, we obtain

$$\tilde{A}\tilde{x}_t = \mu_1 + \mu_2(t-1) + \Gamma B_\rho \tilde{x}_{t-1} + e_t = \tilde{\Gamma} \tilde{B} \tilde{x}_{t-1} + e_t, \quad (21)$$

$$\tilde{A}\tilde{x}_t = \mu_1 + \mu_2(t-1) + \alpha\beta'_0 D x_{t-1} + \Gamma^* B^* x_{t-1} + e_t = \alpha\beta'_0 \tilde{D} \tilde{x}_{t-1} + \tilde{\Gamma}^* \tilde{B}^* \tilde{x}_{t-1} + e_t, \quad (22)$$

where

$$\tilde{B}^* = \begin{bmatrix} I_2 & O_{2 \times s} \\ \vdots & \vdots \\ O_{np \times 2} & B^* \end{bmatrix}, \quad \tilde{D} = [O_{d \times 2} \quad D].$$

Similarly to the case with no deterministic terms, by using equations (21) or (22), we can obtain estimated parameters at iteration  $l+1$  of the EM algorithm by using derivatives of conditional expectation (11), namely

$$\text{vec}(\beta_0^{(l+1)}) = \left\{ \alpha'^{(l)} \Sigma^{(l-1)} \alpha^{(l)} \right\} \otimes \left( \tilde{D} \tilde{M}_{11} \tilde{D}' \right)^{-1} \times \text{vec} \left[ \tilde{D} \left\{ \tilde{M}_{10} \tilde{A}' - \tilde{M}_{11} \left( \tilde{\Gamma}^{*(l)} \tilde{B}^* \right)' \right\} \Sigma^{(l-1)} \alpha^{(l)} \right], \quad (23)$$

$$\tilde{\Gamma}^{(l+1)} = \left( \tilde{A} \tilde{M}_{01} \tilde{B}'^{(l+1)} \right) \left( \tilde{B}^{(l+1)} \tilde{M}_{11} \tilde{B}'^{(l+1)} \right)^{-1}, \quad (24)$$

$$\Sigma^{(l+1)} = T^{-1} \left( \tilde{A} \tilde{M}_{00} \tilde{A}' - \tilde{\Gamma}^{(l+1)} \tilde{B}^{(l+1)} \tilde{M}_{10} \tilde{A}' \right), \quad (25)$$

where for  $j, k = 0, 1$ ,

$$\tilde{M}_{jk} = \sum_{t=1}^T \text{E}_t \left( \tilde{x}_{t-j} \tilde{x}_{t-k}' \mid Y_T^+ \right) = \sum_{t=1}^T \left( \tilde{P}_{t-j, t-k}^T + \tilde{x}_{t-j}^T \tilde{x}_{t-k}^T \right), \quad (26)$$

and  $\tilde{P}_{t-j, t-k}^T$  and  $\tilde{x}_{t-j}^T$  are obtained using the Kalman filter, as in the case with no deterministic terms.

## 5. Comments

One of the important purposes of the paper is to estimate missing or unobserved low-frequency variables,  $y_{2t}$ , which satisfy the structure of the cointegrated multivariate time-series model. This can be done by estimating  $x_t^T$ , together with its covariance matrix  $P_t^T$ , using smoothing equations (B5) and (B6) in Appendix B. One of the advantages of the proposed method is that we can use it to forecast low-frequency variables which are generated jointly with variables observed at higher frequencies. For example, with quarterly GDP and other monthly variables, we can forecast monthly GDP even if GDP is observed only quarterly.

Usually, logarithms of variables are taken before fitting a multivariate time-series model, especially to stabilize the variances of series. This creates no difficulties for high-frequency series (Harvey and Pierce, 1984). However, for a temporally-aggregated low-frequency series, although the sum of the original variables is observed, the logarithm of a sum is not equal to the sum of the logarithms. In such a case, we assume that the logarithms of the data are integrated of order 1. Furthermore, we assume that temporally-aggregated variables are geometric means of unobserved high-frequency variables. This is not a common accounting identity that links high- and low-frequency variables. Usually, low-frequency values are arithmetic means of high-frequency values. In this case, the low-frequency values are geometric means of the high-frequency values. Otherwise, we would have to work with a nonlinear state-space model,

using the extended Kalman filter, as in Anderson and Moore (1979, pp. 193-195). See also Harvey and Pierce (1984) and Mariano and Murasawa (2003).

## 6. Monte Carlo experiments

Monte Carlo experiments are conducted to investigate the performance of the proposed method. The data generating process we consider is similar to the one in Ahn and Reinsel (1990), except that we consider a 3-dimensional process, specifically,

$$\Delta u_t = (\Delta u_{1t}, \Delta u_{2t}, \Delta u_{3t})' = \gamma_1 + ab'u_{t-1} + \varepsilon_t,$$

where  $\varepsilon_t \sim \text{i.i.d. } N_3(0, \Omega)$ , for  $t = 1, \dots, T$ , and  $\gamma_1$  denotes an unrestricted constant term. The parameters are set at the following values:

$$\begin{aligned} \gamma_1 = (\gamma_{11}, \gamma_{21}, \gamma_{31})' &= (-0.2, 0.1, 0.3)', \quad a = (a_1, a_2, a_3)' = (0.6, 1, 0.4)', \\ b = (b_1, b_2, b_3)' &= (1, -2, 3)' \quad \text{and} \quad \Omega = \begin{pmatrix} 25 & 7.5 & 2.5 \\ 7.5 & 9 & 1.5 \\ 2.5 & 1.5 & 1 \end{pmatrix}, \end{aligned}$$

so that  $u_t$  is cointegrated with rank one.

For convenience, we assume that permutation matrix  $V$  is an identity matrix, that is,  $u_{2t}$  and  $u_{3t}$  are not cointegrated. Then,  $u_t = z_t$  and the parameters in models (17) and (18) are identical. After generating  $u_t$ , we set  $y_{1t} = u_{1t}$  and  $y_{2t} = u_{2t}$ , making them high-frequency variables, and, then, generate the low-frequency variable,  $y_{3t}$ , in the following two ways. First (CASE I), we set  $y_{3t} = u_{3t} + u_{3t-1} + u_{3t-2}$ , for  $t = 3, 6, 9, \dots$ , which makes  $y_{3t}$  a temporally-aggregated flow variable. Second (CASE II), we set  $y_{3t} = u_{3t}$ , for  $t = 3, 6, 9, \dots$ , which makes  $y_{3t}$  a skip-sampled stock variable. The remaining values of  $y_{3t}$ , for  $t \neq 3, 6, 9, \dots$ , are assumed to be unobserved.

We generate 1,000 replications of the series for sample sizes  $T = 120$  and  $T = 240$ , which represent 10 and 20 years of monthly data, such that the first 50 values are discarded in order to reduce dependence on starting values. We estimate a VAR(1) with an unrestricted constant.

Because, to our best knowledge, there is no other method available to analyze mixed-frequency data with cointegration, we evaluate the performance of the proposed method against the case in which all variables are observed at the highest frequency, that is, there are no missing data. We expect our method to perform nearly as well with intermittently-missing mixed-frequency data as with complete high-frequency data. The same issue arose in Chen and Zadrozny (1998), where a similar Monte Carlo experiment was conducted to evaluate the

performance of their method for estimating a stationary VAR model using mixed-frequency data, relative to using complete high-frequency data.

The method of Ahn and Reinsel (1990) is used for estimating with the complete high-frequency data. As mentioned before, the initial estimate of the cointegrating vector for the proposed method is obtained using low-frequency data obtained by temporally-aggregated high-frequency data. The remaining initial estimates of the stationary parameters are obtained by treating the initial estimate of the cointegrating vector as if it were known.

Tables 1 and 2 contain the results of the simulations for CASEs I and II. We observe that the differences in the tables are generally fairly small and that the main EM method performs well with the mixed-frequency data, compared with the complete data, in terms of biases and root mean-squared errors (RMSEs) of estimated parameters. As expected, neither the initial nor the main EM algorithms perform better using mixed-frequency data than using complete high-frequency data.

It is interesting to compare tables 1 and 2 with tables 2 to 7 in Chen and Zadrozny (1998), which show how MLE and extended Yule-Walker (XYW) parameter estimates deteriorate, in terms of RMSE, when going from complete to mixed-frequency data. In their table 5, the RMSEs of MLE increase about 44%, whereas the best RMSEs of XYW increase about 77% or more. Consider how the RMSEs change in table 1, for  $T = 240$ , as we move from using complete data to using mixed-frequency data and the main EM algorithm. The RMSEs increase from 5.7% to 65.1%, on average by 40.3%. We conclude that MLEs of parameters in both stationary and nonstationary (cointegrated) VAR processes lose a similar amount of RMSE accuracy when going from complete to mixed-frequency data. However, we should be cautious in drawing this as a general conclusion because it surely also depends on whether the VAR process is stationary or not, on the dimension of the process, and on the sample size.

In all cases, regardless of the value of  $T$ , the main EM algorithm performs better than the initial EM, in terms of bias and RMSE. For longer series, biases and RMSEs of both the initial and main algorithm are smaller. In table 3, we report decline rates of RMSEs for estimated parameters in CASEs I and II, when  $T$  doubles from 120 to 240. For stationary parameters,  $a$  and  $\Omega_{ij}$ , and nonstationary parameters,  $b$ , RMSEs in tables 1 and 2 are consistent with the respective convergence rates of  $O_p(T^{-1/2})$  and  $O_p(T^{-1})$ . RMSEs of stationary parameters generally decline by 29% ( $=1 - \sqrt{120/240}$ ) or more when  $T$  doubles from 120 to 240 and RMSEs of nonstationary parameters generally decline by 50% ( $=1 - 120/240$ ) or more when  $T$  doubles. However, changes in the RMSEs of stationary parameters,  $\gamma$ , are ambiguous, because their RMSEs decline faster than is predicted for stationary parameters but slower than is predicted for nonstationary parameters. Most importantly, this also occurs when we apply the method of Johansen (1988) or Ahn and Reinsel (1990) to the analysis of complete data and the RMSEs of the stationary parameters,  $\gamma$ , do not decline more than

those of the nonstationary parameters.

Figure 1 shows RMSEs of estimates of missing values of  $y_{3t}$ , for  $t \neq 3, 6, 9, \dots$ , for CASEs I and II. The figure shows all except two starting values for CASE I and one starting value for CASE II. Two final outlying values (0.588 and 0.693, for  $t=118$  and  $119$ , when  $T=120$ , and 0.556 and 0.626, for  $t=238$  and  $239$ , when  $T=240$ ) are omitted from Figure 1, because they are spuriously large due to endpoint effects, which arise as a result of the backward smoothing recursions (B3) to (B6), with  $r_{T+1}=0$  and  $R_{T+1}=0$  being a sort of initialization. In Figure 2, this phenomenon is weaker because the aggregation order of the variable is regarded as one. From Figures 1 and 2, we see that as  $T$  doubles from 120 to 240, the RMSEs decline by 8.87% and 9.05% in CASEs I and II. Therefore, for stationary parameters, RMSEs of estimates of missing values decline more slowly than the expected convergence rate  $O_p(T^{-1/2})$ .



Table 1. Performance of EM algorithm in complete versus mixed-frequency data, for CASE I of temporally-aggregated mixed-frequency data

	$T = 120$			$T = 240$		
	Complete data	Mixed-freq. data		Complete data	Mixed-freq. data	
		Init. EM	Main EM		Init. EM	Main EM
$a_1 = 0.6$						
Mean	0.585	0.566	0.578	0.593	0.583	0.588
RMSE	0.049	0.069	0.056	0.035	0.042	0.037
$a_2 = 1$						
Mean	1.007	0.987	0.995	1.002	0.987	0.993
RMSE	0.032	0.064	0.040	0.020	0.036	0.025
$a_3 = 0.4$						
Mean	0.399	0.391	0.396	0.399	0.394	0.397
RMSE	0.011	0.028	0.016	0.007	0.015	0.010
$b_1 = -2$						
Mean	-1.999	-1.949	-1.990	-2.001	-2.003	-2.001
RMSE	0.030	0.171	0.050	0.011	0.058	0.015
$b_2 = 3$						
Mean	2.997	2.896	2.980	3.002	3.009	3.003
RMSE	0.066	0.380	0.112	0.023	0.126	0.033
$\gamma_{11} = -0.2$						
Mean	-0.148	0.013	-0.199	-0.203	-0.473	-0.276
RMSE	0.860	3.568	1.650	0.466	1.669	0.665
$\gamma_{21} = 0.1$						
Mean	0.114	0.362	0.019	0.071	-0.386	-0.052
RMSE	1.229	6.186	2.652	0.590	2.743	0.953
$\gamma_{31} = 0.3$						
Mean	0.314	0.418	0.271	0.294	0.107	0.240
RMSE	0.480	2.457	1.065	0.232	1.103	0.383

Table 1. (Continued)

	$T = 120$			$T = 240$		
	Complete data	Mixed-freq. data		Complete data	Mixed-freq. data	
		Init. EM	Main EM		Init. EM	Main EM
$\Omega_{11} = 25$						
Mean	24.470	27.767	25.891	24.843	26.186	25.483
RMSE	3.114	5.163	3.874	2.240	2.894	2.537
$\Omega_{12} = 7.5$						
Mean	7.374	12.365	9.816	7.469	9.644	8.642
RMSE	1.508	6.390	3.999	1.074	3.066	2.226
$\Omega_{13} = 2.5$						
Mean	2.461	5.209	3.773	2.486	3.690	3.090
RMSE	0.511	3.504	2.175	0.352	1.676	1.173
$\Omega_{22} = 9$						
Mean	8.806	16.121	12.950	8.931	12.434	11.063
RMSE	1.177	8.899	5.957	0.863	4.500	3.352
$\Omega_{23} = 1.5$						
Mean	1.461	5.266	3.590	1.474	3.345	2.572
RMSE	0.307	4.575	3.034	0.216	2.343	1.692
$\Omega_{33} = 1$						
Mean	0.976	2.782	1.931	0.984	1.853	1.461
RMSE	0.132	2.188	1.400	0.092	1.095	0.758

Table 2. Performance of EM algorithm in complete versus mixed-frequency data, for CASE II of skip-sampled mixed-frequency data

	$T = 120$			$T = 240$		
	Complete data	Mixed-freq. data		Complete data	Mixed-freq. data	
		Init. EM	Main EM		Init. EM	Main EM
$a_1 = 0.6$						
Mean	0.585	0.576	0.576	0.593	0.587	0.587
RMSE	0.049	0.060	0.058	0.035	0.039	0.038
$a_2 = 1$						
Mean	1.007	0.989	0.989	1.002	0.990	0.991
RMSE	0.032	0.054	0.047	0.020	0.032	0.029
$a_3 = 0.4$						
Mean	0.399	0.394	0.394	0.399	0.396	0.396
RMSE	0.011	0.022	0.020	0.007	0.013	0.011
$b_1 = -2$						
Mean	-1.999	-1.988	-1.993	-2.001	-2.000	-2.001
RMSE	0.030	0.087	0.055	0.011	0.034	0.019
$b_2 = 3$						
Mean	2.997	2.972	2.984	3.002	3.001	3.002
RMSE	0.066	0.194	0.121	0.023	0.074	0.041
$\gamma_{11} = -0.2$						
Mean	-0.148	0.026	-0.104	-0.203	-0.206	-0.238
RMSE	0.860	2.140	1.631	0.466	1.014	0.704
$\gamma_{21} = 0.1$						
Mean	0.114	0.420	0.196	0.071	0.070	0.014
RMSE	1.229	3.566	2.672	0.590	1.626	1.069
$\gamma_{31} = 0.3$						
Mean	0.314	0.431	0.341	0.294	0.289	0.267
RMSE	0.480	1.417	1.066	0.232	0.654	0.429

Table 2. (Continued)

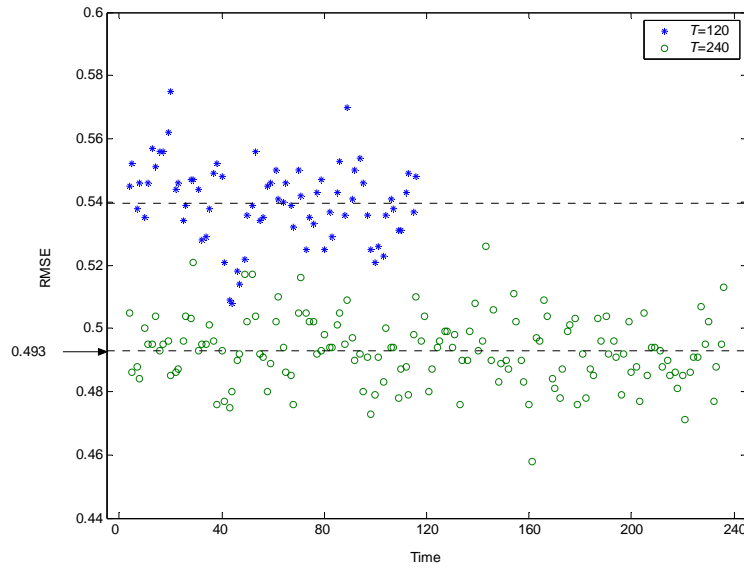
	$T = 120$			$T = 240$		
	Complete data	Mixed-freq. data		Complete data	Mixed-freq. data	
		Init. EM	Main EM		Init. EM	Main EM
$\Omega_{11} = 25$						
Mean	24.470	26.146	25.893	24.843	25.591	25.523
RMSE	3.114	3.932	3.855	2.240	2.585	2.570
$\Omega_{12} = 7.5$						
Mean	7.374	10.341	9.928	7.469	8.893	8.743
RMSE	1.508	4.402	4.134	1.074	2.441	2.350
$\Omega_{13} = 2.5$						
Mean	2.461	3.952	3.769	2.486	3.187	3.124
RMSE	0.511	2.272	2.159	0.352	1.229	1.189
$\Omega_{22} = 9$						
Mean	8.806	14.003	13.338	8.931	11.599	11.291
RMSE	1.177	6.894	6.422	0.863	3.832	3.625
$\Omega_{23} = 1.5$						
Mean	1.461	4.039	3.689	1.474	2.799	2.633
RMSE	0.307	3.388	3.148	0.216	1.862	1.756
$\Omega_{33} = 1$						
Mean	0.976	2.104	1.938	0.984	1.543	1.466
RMSE	0.132	1.521	1.411	0.092	0.811	0.764

Table 3. Decline rates of RMSE in CASEs I and II as  $T$  goes from 120 to 240

Parameter	Complete data	CASE I: temp. agg.		Complete data	CASE II: skip sampled	
		Init. EM	Main EM		Init. EM	Main EM
$a_1$	29%	39%	34%	29%	35%	34%
$a_2$	38%	44%	38%	38%	41%	38%
$a_3$	36%	46%	38%	36%	41%	45%
$b_1$	63%	66%	70%	63%	61%	65%
$b_2$	65%	67%	71%	65%	62%	66%
$\gamma_{11}$	46%	53%	60%	46%	53%	57%
$\gamma_{21}$	52%	56%	64%	52%	54%	60%
$\gamma_{31}$	52%	55%	64%	52%	54%	60%
$\Omega_{11}$	28%	44%	35%	28%	34%	33%
$\Omega_{12}$	29%	52%	44%	29%	45%	43%
$\Omega_{13}$	31%	52%	46%	31%	46%	45%
$\Omega_{22}$	27%	49%	44%	27%	44%	44%
$\Omega_{23}$	30%	49%	44%	30%	45%	44%
$\Omega_{33}$	30%	50%	46%	30%	47%	46%

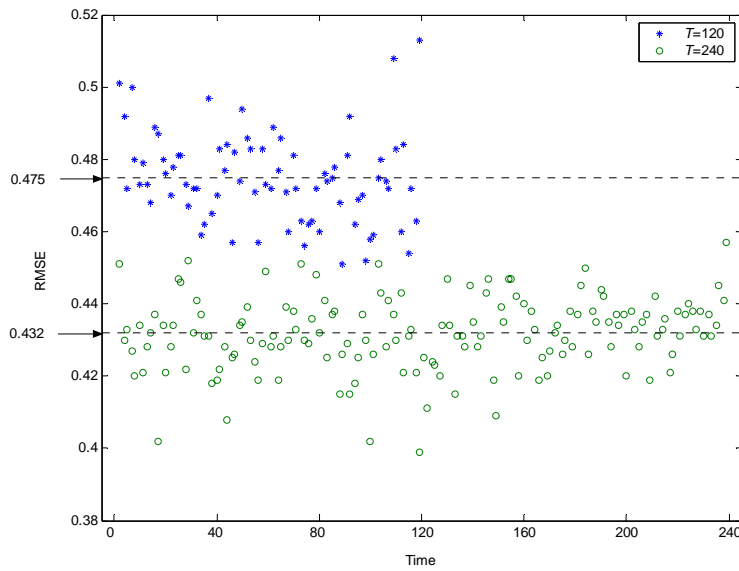
Note: Decline rate =  $1 - (\text{RMSE for } T=240) / (\text{RMSE for } T=120)$

Figure 1. RMSE of monthly estimated  $y_{3t}$  for CASE I



Note: The dotted lines denote average values of the corresponding RMSE.

Figure 2. RMSE of monthly estimated  $y_{3t}$  for CASE II



Note: Same as in Figure.

## 7. Examples

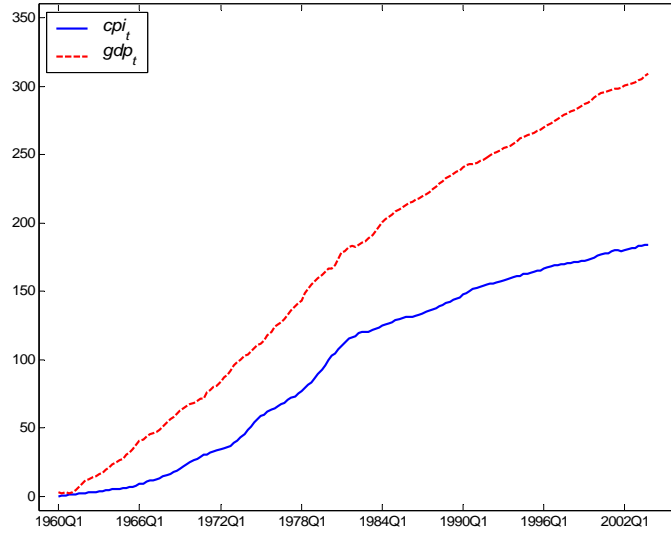
We illustrate the proposed main EM algorithm using mixed-frequency U.S. data of monthly CPI and quarterly GDP, from December 1959 to December 2003, which comprises 176 quarters or 529 months of observations. The CPI and GDP data are from the Bureau of Labor Statistics ([www.bls.gov](http://www.bls.gov)) and from the Bureau of Economic Analysis ([www.bea.gov](http://www.bea.gov)), are seasonally adjusted, have a base CPI value of 100 in 1982-1984, and have a unit of GDP in billions of current dollars. The original data, denoted  $CPI_t$  and  $GDP_t$  in month  $t$ , were transformed to  $cpi_t$  and  $gdp_t$  by taking natural logarithms and subtracting from them the natural logs at the starting dates, as

$$y_t = \begin{pmatrix} cpi_t \\ gdp_t \end{pmatrix} = \begin{pmatrix} \log(CPI_t) - \log(CPI_{1959:12}) \\ \log(GDP_t) - \log(GDP_{1959:12}) \end{pmatrix}.$$

Henceforth, the lower-case variable names,  $cpi$  and  $gdp$ , will refer to these transformations of  $CPI$  and  $GDP$ .

As mentioned in section 3.2, we first estimate the CI rank and the CI vector by using a quarterly single-frequency sample, obtained by picking one monthly value of  $cpi$  per quarter (we call this “skip-sampling”) and keeping quarterly  $gdp$  as is. In Figure 3, the quarterly single-frequency data indicate that the estimated model may have unrestricted, constant and linear, deterministic terms. The model selection criteria, minimum AIC and SBC, both indicated choosing a VAR(4) model. Then, applying Johansen’s trace test to the VAR(4) model, estimated using the single-frequency data, we obtained a p-value of .0437, which indicated a CI rank of one. The corresponding estimate of the CI vector was  $(1, -1.041)'$ . For the mixed-frequency observations on  $y_t = (cpi_t, gdp_t)'$ , we obtained  $(1, -3.123)'$  as the estimate of the cointegrating vector by using the relationship in Pons and Sansó (2005).

Figure 3. Quarterly  $cpi_t$  and  $gdp_t$  from 1960:1 to 2003:12



Note: Quarterly  $cpi_t$  are skip samplings of observed monthly  $cpi_t$  in a quarter.

Table 4 reports minimum AIC and SBC, used to select the best monthly (highest frequency) model and associated log-likelihood values. Both AIC and SBC select the VAR(2) model,

$$\Delta y_t = \begin{pmatrix} -0.0149 \\ 0.1599 \end{pmatrix} + \begin{pmatrix} -0.0040 \\ -0.0007 \end{pmatrix} t + \begin{pmatrix} -0.0161 \\ -0.0022 \end{pmatrix} (1, -3.1289) y_{t-1} \\ + \begin{pmatrix} 0.5156 & 0.0046 \\ 0.1250 & -0.0042 \end{pmatrix} \Delta y_{t-1} + \varepsilon_t,$$

$$\hat{\text{var}}(\varepsilon_t) = \begin{pmatrix} 0.0463 & -0.0036 \\ -0.0036 & 0.0388 \end{pmatrix},$$

as the best VAR( $p$ ) model. For this model, we computed monthly smoothed estimates of  $gdp_t$  and monthly forecasts of  $cpi_t$  and  $gdp_t$ .

Table 5 shows monthly smoothed estimates of  $gdp_t$ , from 2003:1 to 2003:12, as examples of monthly smoothed estimates of  $gdp_t$  in the sample period. Two types of estimates are computed, temporally aggregated and not temporally aggregated. The temporally-aggregated estimates reflect low-frequency quarterly aggregated  $gdp_t$ , i.e., sums of three consecutive monthly  $gdp_t$ . The not-temporally-aggregated estimates reflect high-frequency monthly disaggregated  $gdp_t$ , i.e., monthly  $gdp_t$  by itself. We note that in the table the not-temporally aggregated estimates are multiplied by 3 in order to be in quarterly



form comparable to observed quarterly  $gdp_t$ .

Tables 6 and 7 state out-of-sample (that is, out of model estimation sample) forecasts of  $cpi_t$  and  $gdp_t$  for 2004 and compare these with true values. The tables show smaller forecast errors when using mixed-frequency data than when using single-frequency data. Specifically, the forecast errors for mixed-frequency data are smaller on average by about 52% in Table 6 and about 32% in Table 7.

Table 4. Model selection criteria of optimal VAR( $p$ ) for mixed-frequency data

$P$	Log likelihood function ( $\times 10^6$ )	AIC ( $\times 10^3$ )	SBC ( $\times 10^3$ )
1	-1.7246	6.5326	6.5327
2	-1.7245	6.5323	6.5324
3	-1.7245	6.5323	6.5324
4	-1.7245	6.5323	6.5325
5	-1.7245	6.5324	6.5326
6	-1.7249	6.5337	6.5340
7	-1.7251	6.5346	6.5349
8	-1.7254	6.5356	6.5359
9	-1.7250	6.5341	6.5344
10	-1.7245	6.5325	6.5329

Table 5. Monthly smoothed estimates of in-sample  $gdp_t$ , 2003:1 to 2003:12

Year: month	Observed	Temp. agg.	Not temp. agg.
2003:1		303.33	303.75
2003:2		303.74	304.14
2003:3	304.15	304.15	304.53
2003:4		304.56	305.01
2003:5		304.96	305.34
2003:6	305.44	305.44	305.97
2003:7		306.06	306.87
2003:8		306.82	307.62
2003:9	307.55	307.55	308.16
2003:10		308.14	308.64
2003:11		308.59	308.97
2003:12	308.93	308.93	309.18

Note: “Temporally aggregated” and “not temporally aggregated” estimates in columns 3 and 4, respectively, reflect quarterly sums of monthly values ending in the indicated month and monthly values for that month multiplied by 3, in order to be in quarterly form comparable to observed quarterly  $gdp$  in column 2.

Table 6. Monthly out-of-sample forecasts of  $cpi_t$ 

Year: month	Observed	Single-frequency		Mixed-frequency	
		Forecast	Error	Forecast	Error
2004:1	184.39			184.11	0.28
2004:2	184.71			184.32	0.39
2004:3	185.14	183.99	1.14	184.53	0.61
2004:4	185.35			184.73	0.62
2004:5	185.94			184.93	1.01
2004:6	186.20	184.39	1.81	185.13	1.07
2004:7	186.15			185.32	0.83
2004:8	186.20			185.50	0.70
2004:9	186.36	184.61	1.75	185.69	0.67
2004:10	186.94			185.87	1.07
2004:11	187.20			186.04	1.16
2004:12	187.20	184.75	2.45	186.22	0.98

Note: “Out-of-sample” means for months beyond earlier months, from 1959:12 to 2003:12, used to estimate the model which was used to produce the forecasts.

Table 7. Monthly out-of-sample forecasts of  $gdp_t$ 

Year: month	Observed	Single-frequency		Mixed-frequency		
		Forecast	Error	Forecast Low Freq.	Error	Forecast High Freq.
2004:1				309.22		309.51
2004:2				309.51		309.84
2004:3	310.71	309.47	1.24	309.83	0.88	310.14
2004:4				310.15		310.47
2004:5				310.46		310.77
2004:6	312.30	310.23	2.07	310.78	1.52	311.10
2004:7				311.09		311.40
2004:8				311.41		311.73
2004:9	313.65	310.74	2.91	311.72	1.93	312.03
2004:10				312.04		312.36
2004:11				312.35		312.66
2004:12	315.11	311.22	3.89	312.66	2.45	312.96

Note: Same as in Table 6.

## 8. Conclusion

We have developed and illustrated a method, for estimating a multivariate cointegrated VAR model with mixed-frequency time-series data, by using a state-space representation of an error correction model. The method allows us not only to estimate such a model using mixed-frequency data, but also to estimate missing or unobserved high-frequency values of the low-frequency variables. Monte Carlo experiments, applied to mixed-frequency data, with missing observations, and to single-frequency data, with complete observations, indicate that the proposed method performs well with missing data due to mixed frequencies.

## Appendix A

To compute estimated parameters  $\Gamma^{(l+1)}$  and  $\Sigma^{(l+1)}$  at iteration  $l+1$  of the EM algorithm, we differentiate equation (11) with respect to  $\Gamma$  and  $\Sigma^{-1}$ , using equation (9), and obtain

$$\frac{\partial Q(\theta | \theta^{(l)})}{\partial \Gamma} = -2\Sigma^{-1} \left\{ A \sum_{t=1}^T E_t(x_t x_{t-1}' | Y_T^+) B' - \Gamma B \sum_{t=1}^T E_t(x_{t-1} x_{t-1}' | Y_T^+) B' \right\}, \quad (\text{A1})$$

$$\begin{aligned} \frac{\partial Q(\theta | \theta^{(l)})}{\partial \Sigma^{-1}} = & \frac{T}{2} \{2\Sigma - \text{diag}(\Sigma)\} - \frac{1}{2} \left[ 2 \sum_{t=1}^T E_t \left\{ (Ax_t - \Gamma Bx_{t-1})(Ax_t - \Gamma Bx_{t-1})' | Y_T^+ \right\} \right. \\ & \left. - \text{diag} \left( \sum_{t=1}^T E_t \left\{ (Ax_t - \Gamma Bx_{t-1})(Ax_t - \Gamma Bx_{t-1})' | Y_T^+ \right\} \right) \right], \quad (\text{A2}) \end{aligned}$$

where  $\text{diag}(\cdot)$  denotes the diagonal matrix formed from the diagonal elements of the argument matrix. To find the estimated parameters  $\tilde{\beta}^{(l+1)}$ , we differentiate equation (11) with respect to  $\beta_0$ , using equation (10), and obtain

$$\begin{aligned} \frac{\partial Q(\theta | \theta^{(l)})}{\partial \text{vec}(\beta_0)} = & -2 \sum_{t=1}^T E_t \left[ \left\{ \alpha' \otimes (Dx_{t-1}) \right\} \Sigma^{-1} (Ax_t - \Gamma^* B^* x_{t-1}) | Y_T^+ \right] \\ & + 2 \sum_{t=1}^T E_t \left[ \left\{ \alpha' \otimes (Dx_{t-1}) \right\} \Sigma^{-1} \left\{ \alpha \otimes (Dx_{t-1})' \right\} | Y_T^+ \right] \text{vec}(\beta_0). \quad (\text{A3}) \end{aligned}$$

By noting that  $\left\{ \alpha' \otimes (Dx_{t-1}) \right\} \Sigma^{-1} Ax_t = \text{vec} \left\{ Dx_{t-1} (x_t' A' \Sigma^{-1}) \alpha \right\}$ , we can rewrite equation (A3) as

$$\frac{\partial Q(\theta | \theta^{(l)})}{\partial \text{vec}(\beta_0)} = -2 \text{vec} \left\{ D \sum_{t=1}^T E_t(x_{t-1} x_t' | Y_T^+) A' \Sigma^{-1} \alpha \right\}$$

$$\begin{aligned}
& + 2\text{vec}\left\{D\sum_{t=1}^T E_l(x_{t-1}x'_{t-1} | Y_T^+)(\Gamma^* B^*)'\Sigma^{-1}\alpha\right\} \\
& + 2\left[(\alpha'\Sigma^{-1}\alpha)\otimes\left\{D\sum_{t=1}^T E_l(x_{t-1}x'_{t-1} | Y_T^+)D'\right\}\right]\text{vec}(\beta_0).
\end{aligned}$$

## Appendix B

Let  $x_t^s = E_l(x_t | Y_s^+)$ ,  $P_t^s = \text{cov}_l(x_t | Y_s^+)$  and  $P_{t,t-1}^s = \text{cov}_l(x_t, x_{t-1} | Y_s^+)$ , where  $E_l(\cdot | Y_s^+)$  and  $\text{cov}_l(\cdot | Y_s^+)$  denote the conditional expectation and conditional covariance with respect to the density based on  $\theta^{(l)}$ .

We calculate the prediction and updating recursions using the following equations (for example, Shumway and Stoffer, 2000). For  $t = 1, \dots, T$ ,

$$x_t^{t-1} = Fx_{t-1}^{t-1}, \quad P_t^{t-1} = FP_{t-1}^{t-1}F' + GG', \quad (\text{B1})$$

$$x_t^t = x_t^{t-1} + K_t(y_t^+ - H_t x_t^{t-1}), \quad P_t^t = P_t^{t-1} - K_t H_t P_t^{t-1}, \quad (\text{B2})$$

where  $K_t = P_t^{t-1}H_t'(H_t P_t^{t-1}H_t' + Q_t Q_t')^{-1}$ . We start iterations (B1) and (B2) by setting  $x_0^0 = \lambda$  and  $P_0^0 = \Lambda$ . In order to calculate  $x_t^T$ ,  $P_t^T$  and  $P_{t,t-1}^T$  using equations (B1) and (B2), for  $t = T, \dots, 1$ , we iterate over the backwards recursions

$$r_t = H_t'(H_t P_t^{t-1}H_t' + Q_t Q_t')^{-1}(y_t^+ - H_t x_t^{t-1}) + L_t' r_{t+1} \quad (\text{B3})$$

$$R_t = H_t'(H_t P_t^{t-1}H_t' + Q_t Q_t')^{-1}H_t + L_t' R_{t+1} L_t \quad (\text{B4})$$

where  $r_{T+1} = 0$ ,  $R_{T+1} = 0$ , and  $L_t = F(I_s - K_t H_t)$ . Following Durbin and Koopman (2001), we obtain the smoothing equations

$$x_t^T = x_t^{t-1} + P_t^{t-1} r_t, \quad P_t^T = P_t^{t-1} - P_t^{t-1} R_t P_t^{t-1}, \quad (\text{B5})$$

for  $t = 1, \dots, T$ , and

$$P_{t+1,t}^T = (I_s - P_{t+1}^t R_{t+1}) L_t P_t^{t-1}, \quad (\text{B6})$$

for  $t = 1, \dots, T-1$ .

## Appendix C

Here, we discuss the modifications of state-space equations (19) to (22) which are necessary in order for the state-space representation to incorporate the constant and linear deterministic terms of equation (17). We follow the classifications of the restrictions given by Johansen (1996).

TYPE 1 restrictions are  $\gamma_1 = \alpha\rho_1$  or  $\mu_1 = \alpha\rho_1$  and  $\gamma_2 = 0$  or  $\mu_2 = 0$ . In this case,  $\rho_1$  can be put into the cointegration relationship by replacing  $\beta_0$  with  $\tilde{\beta}_0 = (\rho_1', \beta_0')$ , so that  $\rho_1$  is estimated instead of  $\mu_1$ . TYPE 2 restrictions are an unrestricted  $\gamma_1$  and  $\gamma_2 = 0$ . The model is easily correspondingly changed by deleting the terms involving  $\gamma_2$  and  $\mu_2$ , which are now zero. TYPE 3 restrictions are an unrestricted  $\gamma_1$  and  $\gamma_2 = \alpha\rho_2$  or  $\mu_2 = \alpha\rho_2$ . In this case,  $\rho_2$  is put into the cointegration relationship and  $\beta_0$  is replaced by  $\tilde{\beta}_0 = (\rho_2', \beta_0')$ . Table 8 states these modifications explicitly in algebraic notation.

Table 8. Definitions of state-space components for Type I to Type III restrictions on deterministic terms

	Type 1	Type 2	Type 3
$\tilde{x}_t$	$(1, x_t)'$	$(1, x_t)'$	$(1, t, x_t)'$
$\tilde{F}$	$\begin{bmatrix} 1 & O_{1 \times s} \\ \alpha \rho_1 & F \\ 0_{s-n} & \end{bmatrix}$	$\begin{bmatrix} 1 & O_{1 \times s} \\ \mu_1 & F \\ 0_{s-n} & \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & O_{2 \times s} \\ 1 & 1 & \\ \mu_1 & \alpha \rho_2 & F \\ 0_{s-n} & 0_{s-n} & \end{bmatrix}$
$\tilde{G}$	$\begin{bmatrix} O_{1 \times n} \\ G \end{bmatrix}$	$\begin{bmatrix} O_{1 \times n} \\ G \end{bmatrix}$	$\begin{bmatrix} O_{2 \times n} \\ G \end{bmatrix}$
$\tilde{H}_t$	$[O_{n \times 1} \ H_t]$	$[O_{n \times 1} \ H_t]$	$[O_{n \times 2} \ H_t]$
$\tilde{A}$	$[O_{n \times 1} \ A]$	$[O_{n \times 1} \ A]$	$[O_{n \times 2} \ A]$
$\tilde{\Gamma}$	$\Gamma$	$[\mu_1 \ \Gamma]$	$[\mu_1 \ \Gamma]$
$\tilde{\Gamma}^*$	$[0_n \ \Gamma^*]$	$[\mu_1 \ \Gamma^*]$	$[\mu_1 \ 0_n \ \Gamma^*]$
$\beta_0$	$\tilde{\beta}_0 = (\rho_1', \beta_0)'$	$\beta_0$	$\tilde{\beta}_0 = (\rho_2', \beta_0)'$
$B_\rho$	$\begin{bmatrix} \rho_1' & B \\ O_{(p-1)n \times 1} & \end{bmatrix}$	$[O_{\{h+(p-1)n\} \times 1} \ B]$	$\begin{bmatrix} 0_h \ \rho_2' & B \\ O_{(p-1)n \times 2} & \end{bmatrix}$
$\tilde{B}$	$B_\rho$	$\begin{bmatrix} 1 & O_{1 \times s} \\ B_\rho \end{bmatrix}$	$\begin{bmatrix} 1 & O_{1 \times (s+1)} \\ B_\rho \end{bmatrix}$
$\tilde{B}^*$	$\begin{bmatrix} 1 & O_{1 \times s} \\ O_{np \times 1} & B^* \end{bmatrix}$	$\begin{bmatrix} 1 & O_{1 \times s} \\ O_{np \times 1} & B^* \end{bmatrix}$	$\begin{bmatrix} I_2 & O_{2 \times s} \\ O_{np \times 2} & B^* \end{bmatrix}$
$\tilde{D}$	$\begin{bmatrix} 1 & O_{1 \times s} \\ O_{d \times 1} & D \end{bmatrix}$	$[O_{d \times 1} \ D]$	$\begin{bmatrix} 0 & 1 & O_{1 \times s} \\ O_{d \times 2} & D \end{bmatrix}$

## References

1. Ahn, S. K. and G. C. Reinsel (1990), Estimation for Partially Nonstationary Multivariate Autoregressive Models, *Journal of the American Statistical Association*, 85, 813-823.
2. Anderson, B. D. A. and J. B. Moore (1979), *Optimal Filtering*, Englewood Cliffs, NJ: Prentice Hall.
3. Bernanke, B. S., M. Gertler and M. Watson (1997), Systematic Monetary Policy and the Effects of Oil Price Shocks, *Brookings Papers on Economic Activity*, 1, 91–157.
4. Brockwell, P. J. and R. A. Davis (1991), *Time Series: Theory and Methods*, 2<sup>nd</sup> ed., New York: Springer-Verlag.
5. Chen, B. and P. A. Zdrozny (1998), An Extended Yule-Walker Method for Estimating Vector Autoregressive Models with Mixed-Frequency Data, in *Advances in Econometrics: Messy Data–Missing Observations, Outliers, and Mixed-Frequency Data*, Vol. 13, pp. 47-73, T. B. Fomby and R. C. Hill (eds.), Greenwich, CT: JAI Press.
6. Chow, G. C. and A. Lin (1971), Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series, *Review of Economics and Statistics*, 53, 372–375.
7. Chow, G. C. and A. Lin (1976), Best Linear Unbiased Estimation of Missing Observations in an Economic Time Series, *Journal of the American Statistical Association*, 71, 719–721.
8. Cuche, N. A. and M. K. Hess (2000), Estimating Monthly GDP in a General Kalman Filter Framework: Evidence from Switzerland, *Economic & Financial Modelling*, 7, 153–194.
9. De Jong, P. (1989), Smoothing and Interpolation with the State-Space Model, *Journal of the American Statistical Association*, 84, 1085-1088.
10. De Jong, P. and M. J. MacKinnon (1988), Covariances for Smoothed Estimates in State Space Models, *Biometrika*, 75, 601-602.

11. Dempster, A. P., N. M. Laird and D. B. Rubin (1977), Maximum Likelihood from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society, Series B*, 39, 1-38.
12. Durbin, J. and S. J. Koopman (2001), *Time Series Analysis by State Space Methods*, Oxford: Oxford University Press.
13. Engle, R. F. and C. W. J. Granger (1987), Co-Integration and Error Correction: Representation, Estimation, and Testing, *Econometrica*, 55, 251-276.
14. Granger, C. W. J. and P. L. Siklos (1995), Systematic Sampling, Temporal Aggregation, Seasonal Adjustment, and Cointegration Theory and Evidence, *Journal of Econometrics*, 66, 357-369.
15. Harvey, A. C. and R. G. Pierce (1984), Estimating Missing Observations in Economic Time Series, *Journal of the American Statistical Association*, 79, 125-131.
16. Haug, A. A. (2002), Temporal Aggregation and the Power of Cointegration Tests: a Monte Carlo Study, *Oxford Bulletin of Economics and Statistics*, 64, 399-412.
17. Johansen, S. (1988), Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control*, 12, 231-254.
18. Johansen, S. (1996), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, 2nd ed., Oxford: Oxford University Press.
19. Liu, H. and S. G. Hall (2001), Creating High-Frequency National Accounts with State-Space Modeling: a Monte Carlo Experiment, *Journal of Forecasting*, 20, 441-449.
20. Magnus, J. R. and H. Neudecker (1988), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Chichester: John Wiley and Sons.
21. Marcellino, M. (1999), Some Consequences of Temporal Aggregation in Empirical Analysis, *Journal of Business and Economic Statistics*, 17, 129-



136.

22. Mariano, R. S. and Y. Murasawa (2003), A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series, *Journal of Applied Econometrics*, 18, 427-443.
23. Mariano, R. S. and Y. Murasawa (2004), Constructing a Coincident Index of Business Cycles without Assuming a One-Factor Model, Discussion Paper 2004-6, College of Economics, Osaka Prefecture University.
24. Mittnik, S. and P. A. Zadrozny (2004), Forecasting Quarterly German GDP at Monthly Intervals Using Monthly Ifo Business Conditions Data, in *Ifo Survey Data in Business Cycle and Monetary Policy Analysis*, J. E. Sturm and T. Wollmershäuser (eds.), Heidelberg: Physica Verlag.
25. Pons, G. and A. Sansó (2005), Estimation of Cointegrating Vectors with Time Series Measured at Different Periodicity, *Econometric Theory*, 21, 735-756.
26. Schweppe, F. C. (1965), Evaluation of Likelihood Functions for Gaussian Signals, *IEEE Transactions on Information Theory*, IT-4, 294-305.
27. Shumway, R. H. and D. S. Stoffer (1982), An Approach to Time Series Smoothing and Forecasting Using the EM Algorithm, *Journal of Time Series Analysis*, 3, 253-264.
28. Shumway, R. H. and D. S. Stoffer (2000), *Time Series Analysis and its Applications*, New York: Springer-Verlag.
29. Watson, M. W. and R. F. Engle (1983), Alternative Algorithms for the Estimation of Dynamic Factor, Mimic and Varying Coefficient Regression Models, *Journal of Econometrics*, 23, 385-400.
30. Zadrozny, P.A. (1988), Gaussian Likelihood of Continuous-Time ARMAX Models when Data Are Stocks and Flows at Different Frequencies, *Econometric Theory*, 4, 108-124.
31. Zadrozny, P.A. (1990), Estimating a Multivariate ARMA Model with Mixed-Frequency Data: an Application to Forecasting U.S. GNP at Monthly Intervals, Center for Economic Studies Discussion Paper 90-5, Bureau of the

Census, Washington, DC.

# CESifo Working Paper Series

(for full list see [www.cesifo-group.de](http://www.cesifo-group.de))

---

- 1876 Agnieszka Stażka, Sources of Real Exchange Rate Fluctuations in Central and Eastern Europe – Temporary or Permanent?, December 2006
- 1877 Xavier Calsamiglia, Teresa Garcia-Milà and Therese J. McGuire, Why do Differences in the Degree of Fiscal Decentralization Endure?, December 2006
- 1878 Natacha Gilson, How to be Well Shod to Absorb Shocks? Shock Synchronization and Joining the Euro Zone, December 2006
- 1879 Scott Alan Carson, Modern Health Standards for Peoples of the Past: Biological Conditions by Race in the American South, 1873 – 1919, December 2006
- 1880 Peter Huber, Michael Pfaffermayr and Yvonne Wolfmayr, Are there Border Effects in the EU Wage Function?, December 2006
- 1881 Harry Flam and Håkan Nordström, Euro Effects on the Intensive and Extensive Margins of Trade, December 2006
- 1882 Panu Poutvaara and Mikael Priks, Hooliganism in the Shadow of the 9/11 Terrorist Attack and the Tsunami: Do Police Reduce Group Violence?, December 2006
- 1883 Ruud A. de Mooij and Gaëtan Nicodème, Corporate Tax Policy, Entrepreneurship and Incorporation in the EU, December 2006
- 1884 Johannes Becker and Clemens Fuest, Corporate Tax Policy and International Mergers and Acquisitions – Is the Tax Exemption System Superior?, January 2007
- 1885 Momi Dahan and Udi Nisan, The Effect of Benefits Level on Take-up Rates: Evidence from a Natural Experiment, January 2007
- 1886 José García-Solanes, Francisco I. Sancho-Portero and Fernando Torrejón-Flores, Beyond the Salassa-Samuelsen Effect in some New Member States of the European Union, January 2007
- 1887 Peter Egger, Wolfgang Eggert and Hannes Winner, Saving Taxes Through Foreign Plant Ownership, January 2007
- 1888 Timothy J. Goodspeed and Andrew Haughwout, On the Optimal Design of Disaster Insurance in a Federation, January 2007
- 1889 Wim Groot, Henriëtte Maassen van den Brink and Bernard van Praag, The Compensating Income Variation of Social Capital, January 2007
- 1890 Bas Jacobs, Ruud A. de Mooij and Kees Folmer, Analyzing a Flat Income Tax in the Netherlands, January 2007

- 1891 Hans Jarle Kind, Guttorm Schjelderup and Frank Stähler, Newspapers and Advertising: The Effects of Ad-Valorem Taxation under Duopoly, January 2007
- 1892 Erkki Koskela and Rune Stenbacka, Equilibrium Unemployment with Outsourcing under Labour Market Imperfections, January 2007
- 1893 Maarten Bosker, Steven Brakman, Harry Garretsen, Herman de Jong and Marc Schramm, The Development of Cities in Italy 1300 – 1861, January 2007
- 1894 Michel Beine, Oscar Bernal, Jean-Yves Gnabo and Christelle Lecourt, Intervention Policy of the BoJ: A Unified Approach, January 2007
- 1895 Robert S. Chirinko and Daniel J. Wilson, State Investment Tax Incentives: A Zero-Sum Game?, January 2007
- 1896 Theo S. Eicher and Oliver Roehn, Sources of the German Productivity Demise – Tracing the Effects of Industry-Level ICT Investment, January 2007
- 1897 Helge Berger, Volker Nitsch and Tonny Lybek, Central Bank Boards around the World: Why does Membership Size Differ?, January 2007
- 1898 Gabriel Felbermayr and Wilhelm Kohler, Does WTO Membership Make a Difference at the Extensive Margin of World Trade?, January 2007
- 1899 Benno Torgler and Friedrich Schneider, The Impact of Tax Morale and Institutional Quality on the Shadow Economy, January 2007
- 1900 Tomer Blumkin and Efraim Sadka, On the Desirability of Taxing Charitable Contributions, January 2007
- 1901 Frederick van der Ploeg and Reinhilde Veugelers, Higher Education Reform and the Renewed Lisbon Strategy: Role of Member States and the European Commission, January 2007
- 1902 John Lewis, Hitting and Hoping? Meeting the Exchange Rate and Inflation Criteria during a Period of Nominal Convergence, January 2007
- 1903 Torben M. Andersen, The Scandinavian Model – Prospects and Challenges, January 2007
- 1904 Stephane Dees, Sean Holly, M. Hashem Pesaran and L. Vanessa Smith, Long Run Macroeconomic Relations in the Global Economy, January 2007
- 1905 Richard Jong-A-Pin and Jakob De Haan, Political Regime Change, Economic Reform and Growth Accelerations, January 2007
- 1906 Sascha O. Becker and Peter H. Egger, Endogenous Product versus Process Innovation and a Firm's Propensity to Export, February 2007

- 1907 Theo S. Eicher, Chris Papageorgiou and Oliver Roehn, Unraveling the Fortunate of the Fortunate: An Iterative Bayesian Model Averaging (IBMA) Approach, February 2007
- 1908 Liliana E. Pezzin, Robert A. Pollak and Barbara S. Schone, Efficiency in Family Bargaining: Living Arrangements and Caregiving Decisions of Adult Children and Disabled Elderly Parents, February 2007
- 1909 Christian Keuschnigg and Soren Bo Nielsen, Self-Selection and Advice in Venture Capital Finance, February 2007
- 1910 Rune Jansen Hagen and Gaute Torsvik, Irreversible Investments, Dynamic Inconsistency and Policy Convergence, February 2007
- 1911 Eric A. Hanushek and Ludger Woessmann, The Role of School Improvement in Economic Development, February 2007
- 1912 Bernard M. S. van Praag, Perspectives from the Happiness Literature and the Role of New Instruments for Policy Analysis, February 2007
- 1913 Volker Grossmann and Thomas M. Steger, Growth, Development, and Technological Change, February 2007
- 1914 Margarita Katsimi and Thomas Moutos, Human Capital and the Feldstein-Horioka Puzzle, February 2007
- 1915 Oliver Roehn, Theo S. Eicher and Thomas Strobel, The Ifo Industry Growth Accounting Database, February 2007
- 1916 Ian Babetskii, Aggregate Wage Flexibility in Selected New EU Member States, February 2007
- 1917 Burkhard Heer, Alfred Maussner and Paul D. McNelis, The Money-Age Distribution: Empirical Facts and Limited Monetary Models, February 2007
- 1918 Yin-Wong Cheung, Menzie D. Chinn and Eijii Fujii, The Overvaluation of Renminbi Undervaluation, February 2007
- 1919 Jim Malley, Apostolis Philippopoulos and Ulrich Woitek, To React or Not? Fiscal Policy, Volatility and Welfare in the EU-3, February 2007
- 1920 Mattias Polborn, Competing for Recognition through Public Good Provision, February 2007
- 1921 Lars P. Feld and Benno Torgler, Tax Morale after the Reunification of Germany: Results from a Quasi-Natural Experiment, February 2007
- 1922 Robert S. Chirinko and Huntley Schaller, Fundamentals, Misvaluation, and Investment: The Real Story, February 2007

- 1923 Benno Torgler and Friedrich Schneider, Shadow Economy, Tax Morale, Governance and Institutional Quality: A Panel Analysis, February 2007
- 1924 Adrian Pagan and M. Hashem Pesaran, On Econometric Analysis of Structural Systems with Permanent and Transitory Shocks and Exogenous Variables, February 2007
- 1925 Hans-Werner Sinn, The Welfare State and the Forces of Globalization, February 2007
- 1926 Michael Smart, Raising Taxes through Equalization, February 2007
- 1927 Øystein Foros, Kåre P. Hagen and Hans Jarle Kind, Price-Dependent Profit Sharing as an Escape from the Bertrand Paradox, February 2007
- 1928 Balázs Égert, Kirsten Lommatzsch and Amina Lahrière-Révil, Real Exchange Rates in Small Open OECD and Transition Economies: Comparing Apples with Oranges?, February 2007
- 1929 Aleksander Berentsen and Cyril Monnet, Monetary Policy in a Channel System, February 2007
- 1930 Wolfgang Ochel, The Free Movement of Inactive Citizens in the EU – A Challenge for the European Welfare State?, February 2007
- 1931 James K. Hammitt and Nicolas Treich, Statistical vs. Identified Lives in Benefit-Cost Analysis, February 2007
- 1932 Wilhelm Kohler, The Bazaar Effect, Unbundling of Comparative Advantage, and Migration, February 2007
- 1933 Karsten Staehr, Fiscal Policies and Business Cycles in an Enlarged Euro Area, February 2007
- 1934 Michele Bernasconi and Paola Profeta, Redistribution or Education? The Political Economy of the Social Race, March 2007
- 1935 Axel Dreher, Martin Gassebner and Lars-H. R. Siemers, Does Terror Threaten Human Rights? Evidence from Panel Data, March 2007
- 1936 Naércio Aquino Menezes Filho and Marc-Andreas Muendler, Labor Reallocation in Response to Trade Reform, March 2007
- 1937 Gebhard Flaig and Timo Wollmershaeuser, Does the Euro-zone Diverge? A Stress Indicator for Analyzing Trends and Cycles in Real GDP and Inflation, March 2007
- 1938 Michael Funke and Michael Paetz, Environmental Policy Under Model Uncertainty: A Robust Optimal Control Approach, March 2007
- 1939 Byeongchan Seong, Sung K. Ahn and Peter A. Zadrozny, Cointegration Analysis with Mixed-Frequency Data, March 2007