Temporal Aggregation and Ordinary Least Squares Estimation of Cointegrating Regressions

Gabriel Pons Rotger[a b] Andreu Sanso Rossello[a]
Jordi Surinyach Caralt[a]
[a]Econometrics Department
University of Barcelona
[b]Visiting Researcher
European University Institute

January 31, 2000

Abstract

The paper derives the asymptotic distribution of the ordinary least squares estimator of cointegrating vectors with temporally aggregated time series. It is shown, that temporal aggregation reduces the bias and variance of the estimator for average sampling (temporal aggregation of flow series) and does not affect the limiting distribution for systematic sampling (temporal aggregation of stock series). A Monte Carlo experiment shows the consistency of the finite sample results with the asymptotic theory.

1 Introduction[1]

Temporal aggregation effects on cointegrating vector estimation has not been analyzed in the literature, and as it is shown in this paper this transformation of the data can affect the statistical properties of some estimators. Some previous considerations must be made on the basis of theoretical results and empirical practice. First of all, while short-run dynamics changes with temporal aggregation (see for example Brewer, 1973), the cointegrating space is invariant (see Granger, 1990; Granger and Siklos, 1995; and Marcellino, 1996). Then, although economic time series are observed at a longer interval than are generated, the cointegrating vector of the generating mechanism can be estimated with lower frequency data.

Temporal aggregation is not taken into account for the estimation of cointegrating relationships, and the habitual thing is to assume that high frequency data provide better results than low frequency data because its bigger sample size. Moreover, in the literature of temporal aggregation and unit roots, a considerable amount of works exists that study the effects of temporal aggregation on unit root (see for example Shiller and Perron, 1985; Perron, 1989, 1991; Choi, 1992; and Choi and Chung, 1995), and cointegration testing (see Hooker, 1993; or Lahiri and Mamingi, 1995). These papers, although there is no an agreement on the magnitude of the effect, show that temporal aggregation reduces the power of the tests, so inference on unit roots and cointegrating relationships has better properties with high frequency data. On the other hand, sometimes in empirical modelling two samples of the same variables are available, a sample with a long span and temporally aggregated data, and another sample with shorter span but with high frequency data possibly with a bigger sample size. For example Gregory and Hansen (1996) use an annual sample covering 1901-1985 and a quarterly one for the period 1960-1991. So in order to decide the most appropiate sampling interval for estimate cointegrating vectors, it is necessary to study the temporal aggregation effects on the estimator.

The paper analyzes temporal aggregation effects on the behaviour of ordinary least squares (OLS) estimator when the generating mechanism of the disaggregated time series is an n-dimensional cointegrated system with one cointegrating vector. The limiting distribution of the estimator is derived for different types of temporal aggregation. It is shown that for average sampling and mixed sampling, the distribution depends on the temporal aggregation level of the variables, and it is possible for average sampling and some cases of mixed sampling to reduce the bias and variance of the estimator, by estimating the cointegrating vector with temporally aggregated time series. On the other hand, systematic sampling does not affect the limiting distribution of the OLS estimator.

The plan of the paper is first to establish in Section 2, the generating mechanism of the disaggregated time series, a CI(1,1) process, cointegrated process of order

[1] I am grateful to Jesus Gonzalo and Soren Johansen for many helpful comments on an earlier version of the paper.

(1,1). Section 3 provides the representation for different types of temporally aggregated time series. In Section 4, a multivariate invariance principle for temporally aggregated process is provided, and the asymptotic distribution of OLS estimator for different types of aggregation is derived. In Section 5 some simulation experiments are conducted. Section 6 concludes.

Throughout the paper, $\stackrel{L}{\to}$ denotes convergence in distribution and $\stackrel{p}{\to}$ stands for convergence in probability. The Brownian motion $B(r)$ and the standarized Brownian motion $W(r)$ on [0,1] are written as B and W, respectively, to achieve notational economy. Similarly, we write integrals with respect to Lebesgue measure such $\int_0^1 W(r)dr$ more simply as $\int W$. Vector Brownian motion with convariance matrix Ω is written $BM(\Omega)$. y_t denotes an n-dimensional time series measured at basic time units t, and Y_τ denotes an n-dimensional temporally aggregated time series measured at τ, where $\tau = [t/m]$ and $[x]$ is the integer part of x.

2 Generating mechanism of y_t: A cointegrated system

Consider the n-dimensional time series y_t; $t = 1;\ldots;T$; partitioned as $y_t = (y'_{1;t}; y'_{2;t})'$; where $y_{1;t}$ is r-dimensional and $y_{2;t}$ is g-dimensional ($g = n - r$), generated at discrete time t, by the system

$$y_{1;t} = \beta y_{2;t} + u_t; \quad (1)$$

$$\Delta y_{2;t} = v_t; \quad (2)$$

where the $r \times n$ matrix $(1, -\beta)$ is the matrix of cointegrating vectors, Δ is the difference operator in t-time, $\Delta = 1 - L$, where L is the lag operator in t-time, $L^j y_t = y_{t-j}$: u_t, the disequilibrium error of the cointegrating relationships (1) is called the temporary component, and v_t, the shocks that drive the stochastic trends, the permanent component. Define $\xi_t = (u'_t; v'_t)'$; and introduce the following assumptions:

1. $y_{1;t}$ and $y_{2;t}$ are not cointegrated processes,
2. ξ_t is a covariance stationary process with zero mean,
3. that satisfies the invariance principle (IP) (Phillips and Durlauf, 1986):

$$T^{-1/2} \sum_{t=1}^{[Tr]} \xi_t \stackrel{L}{\to} B_\xi \equiv BM(\Omega_{\xi\xi}); \quad (3)$$

for $r \in [0;1]$; where B_ξ is an n-dimensional vector Wiener process with covariance matrix $\Omega_{\xi\xi}$, the long-run covariance matrix of ξ_t:

$$\Omega_{\xi\xi} = \lim_{T\to\infty} T^{-1} E\left[\left(\sum_{t=1}^{T} \xi_t\right)\left(\sum_{t=1}^{T} \xi_t\right)'\right]: \quad (4)$$

The assumptions 1 and 2 imply that y_t is a CI(1,1) process. In this model, the short-run dynamics are absorbed in the long-run variance matrix of ε_t. The main drawback of this representation is that the consistency of the results depends on the normalization of the cointegrating regression being correct.

In order to derive the asymptotic distribution (a.d.) of the OLS estimator two decompositions of the long-run variance $\Omega_{\varepsilon\varepsilon}$ will be very useful. First, let express $\Omega_{\varepsilon\varepsilon}$ as

$$\Omega_{\varepsilon\varepsilon} = \Gamma_{\varepsilon\varepsilon}(0) + \Lambda_{\varepsilon\varepsilon} + \Lambda'_{\varepsilon\varepsilon};$$

where

$$\Lambda_{\varepsilon\varepsilon} = \sum_{j=1}^{\infty} \Gamma_{\varepsilon\varepsilon}(-j);$$

and

$$\Gamma_{\varepsilon\varepsilon}(j) = E(\varepsilon_t \varepsilon'_{t-j}).$$

Another useful decomposition of $\Omega_{\varepsilon\varepsilon}$ is the Cholesky factorization:

$$\Omega_{\varepsilon\varepsilon} = P_{\varepsilon\varepsilon} P'_{\varepsilon\varepsilon};$$

where $P_{\varepsilon\varepsilon}$, a triangular matrix, is the square root of $\Omega_{\varepsilon\varepsilon}$. With the Cholesky factorization we can express the Wiener process B_ε as the product

$$B_\varepsilon = P_{\varepsilon\varepsilon} \cdot W_\varepsilon;$$

where W_ε is a standard n-dimensional vector Wiener process.

Let partition B_ε, W_ε, $\Omega_{\varepsilon\varepsilon}$, $\Gamma_{\varepsilon\varepsilon}(j)$, $\Lambda_{\varepsilon\varepsilon}$, and $P_{\varepsilon\varepsilon}$ according with the dimensions of u_t and v_t:

$$B_\varepsilon = \begin{bmatrix} B_u \\ B_v \end{bmatrix};$$

$$W_\varepsilon = \begin{bmatrix} W_u \\ W_v \end{bmatrix};$$

$$\Omega_{\varepsilon\varepsilon} = \begin{bmatrix} \Omega_{uu} & \Omega_{uv} \\ \Omega_{vu} & \Omega_{vv} \end{bmatrix};$$

$$\Gamma_{\varepsilon\varepsilon}(j) = \begin{bmatrix} \Gamma_{uu}(j) & \Gamma_{uv}(j) \\ \Gamma_{vu}(j) & \Gamma_{vv}(j) \end{bmatrix};$$

$$\Lambda_{\varepsilon\varepsilon} = \begin{bmatrix} \Lambda_{uu} & \Lambda_{uv} \\ \Lambda_{vu} & \Lambda_{vv} \end{bmatrix};$$

and

$$P_{\varepsilon\varepsilon} = \begin{bmatrix} P_{uu} & P_{uv} \\ 0 & P_{vv} \end{bmatrix};$$

with the $r \times r$ matrix P_{uu}, the $r \times g$ matrix P_{uv}; and the $g \times g$ matrix P_{vv} given by

$$P_{uu} = \left(\Omega_{uu} - \Omega_{uv} \Omega_{vv}^{-1} \Omega_{vu} \right)^{1/2};$$
$$P_{uv} = \Omega_{uv} \Omega_{vv}^{-1/2};$$
$$P_{vv} = \Omega_{vv}^{1/2}.$$

Denoting the $r \times n$ matrix $P_u = [P_{uu}; P_{uv}]$ and the $g \times n$ matrix $P_v = [0; P_{vv}]$; B_u and B_v can be written as

$$B_u = P_{uu} \cdot W_u + P_{uv} \cdot W_v;$$
$$B_v = P_{vv} \cdot W_v.$$

3 Representations for temporally aggregated time series

Temporal aggregation does affect some properties of the cointegrating system (1) and (2), the short-run dynamics, but does not change the zero unit roots nor the cointegrating space (see Granger, 1990; Phillips, 1991b; Marcellino, 1996, 1999; or Li, 1998) because the linearlity of the transformation. In the literature, studies on temporal aggregation effects on time series models specifications have centered on ARIMA (see Brewer, 1973; Wei, 1981; and Weiss, 1984), ARMAX (see Brewer, 1973; and Weiss, 1984), GARCH (see Drost and Nijman, 1993), or VARMA (see Lutkepohl, 1987; and Marcellino, 1996, 1999) classes of models. This literature derives the aggregated model (representation of the temporally aggregated time series) and the relation of AR and MA polynomial orders of aggregated model with the orders of the disaggregated model (representation of the disaggregated time series), as well as the function relating the parameters of the aggregated model with the parameters of the disaggregated one.

Sometimes the links between the aggregated and disaggregated model are difficult to establish. For example, the derivation of the MA coefficients of a temporally aggregated VARMA (p,q) model can be a complicated task. For this reason the triangular representation is used in the paper. With this model we avoid the complications associated with the short-run dynamics specifications. In triangular cointegrated systems these dynamics are summarized by the long-run variance Ω_{33}; and as it is shown in section 4 the effects of temporal aggregation on this matrix can be obtained easily.

Three temporal aggregation schemes will be considered, depending on the components of the vector y_t being stock or flow variables. When all the elements of y_t are stock variables, temporally aggregated time series are obtained through systematic sampling. Purchasing power parity analysis or the Fisher equation theory are examples of long-run relationships between stocks. On the other hand, when all the elements of y_t are flow series (i.e., outcome, consumption and investment) average sampling is applied to y_t to obtain the temporally aggregated time series. The mixed sampling, that is when some variables are averaged sampled and others are systematically sampled (i.e., long-run money demand studies) is considered too.

For every type of temporal aggregation, it is assumed that an n-dimensional vector time series, y_t, $t = 1; :::; T$; is generated at time interval, 1, indexed by t (i.e., a month), and it can be observed at a longer interval, $m > 1$, indexed by τ; where $\tau = [t/m]$ (i.e., if $m = 12$ then τ denotes years) and the sample size of aggregated time series is $T^{(m)} = [T/m]$. Note that, T is the sample size at disaggregated level and the span for all temporal aggregation order measured at t-units. In order to adopt a notation valid for all m, $T^{(m)}$ will denote the sample size and T the span of the sample.

The systematically sampled time series, Y_τ^{ss}; is defined as

$$Y_\tau^{ss} \equiv y_{m\tau}; \tag{5}$$

and the averaged sampled time series, Y_τ^{as} as

$$Y_\tau^{as} \equiv s(L)y_{m\tau}; \tag{6}$$

where $s(L)$ is the $(m-1)$-order lag polynomial

$$s(L) = \sum_{j=0}^{m-1} L^j. \tag{7}$$

Then the averaged sampled time series is the partial sum of the m nonoverlapping observations

$$Y_\tau^{as} = y_{m\tau} + y_{m\tau-1} + ::: + y_{m(\tau-1)+1};$$

Weiss (1984) observed that average sampling (6) is the result of two transformations. First, the lag polynomial (7) is applied to y_t; obtaining the partial sum of the m overlapping observations:

$$Y_t = s(L)y_t = y_t + y_{t-1} + ::: + y_{t-m+1}. \tag{8}$$

Then, systematic sampling of order m (5) is applied to (8) obtaining the averaged sampled time series

$$Y_\tau^{as} = Y_{m\tau}. \tag{9}$$

Note, that the difference between average sampling (6) and systematic sampling (5), is that in the first case sampling is applied to the partial sum of y_t (8), and in the second case is applied to y_t.

3.1 Model of systematically sampled time series

Let derive the representation for the systematically sampled time series, Y_τ^{ss}. In this case, systematic sampling is applied to both equations (1)-(2) obtaining:

$$
\begin{aligned}
Y_{1,\tau}^{ss} &= \beta' Y_{2,\tau}^{ss} + U_\tau^{ss}, \\
\Delta Y_{2,\tau}^{ss} &= V_\tau^{ss}.
\end{aligned} \tag{10}
$$

Note that in equation (10), Y_τ^{ss} is differenced in t-time, However, Y_τ^{ss} is measured in τ-time, so the first difference in the τ-time, $1 - L^m$, must be obtained. Then, define the difference operator in τ-time as

$$\Delta_m \equiv 1 - L^m,$$

where L^m is the lag operator in τ-time, such that $L^m Y_\tau^i = Y_{\tau-1}^i$ ($i =$ ss, as, ms). Given the equality

$$\Delta_m = s(L)\Delta,$$

the polynomial $s(L)$ must be applied to (10) to obtain Δ_m. Then, the common stochastic trend in the aggregated model is[2]

$$\Delta_m Y_{2,\tau}^{ss} = V_\tau^{as}. \tag{11}$$

Proposition 1 *Let y_t be an n-dimensional vector time series generated by the CI(1,1) process (1)-(2), then the systematically sampled time series Y_τ^{ss} follows a CI(1,1) process represented by*

$$
\begin{aligned}
Y_{1,\tau}^{ss} &= \beta' Y_{2,\tau}^{ss} + U_\tau^{ss}, \\
\Delta_m Y_{2,\tau}^{ss} &= V_\tau^{as},
\end{aligned}
$$

where

$$\varepsilon_\tau^{ss} = (U_\tau^{ss\prime}, V_\tau^{as\prime})' = S_{ss}(L)\varepsilon_{m\tau},$$

and the polynomial matrix $S_{ss}(L)$ is given by

$$S_{ss}(L) = \begin{bmatrix} I_r & 0 \\ 0 & s(L)\Delta I_g \end{bmatrix},$$

with $s(L)$ given by (7).

[2] Note that the permanent component of the triangular model for systematically sampled time series, V_τ^{as} is the averaged sampled permanent component of the disaggregated model, v_t.

3.2 Model of averaged sampled time series

In order to obtain the averaged sampled time series, Y_τ^{as}; we will follow the transformations observed by Weiss. So the polynomial $s(L)$; is applied to equations (1) and (2):

$$
\begin{aligned}
s(L)y_{1;t} &= s(L)\beta' y_{2;t} + s(L)u_t;\\
s(L)\Delta y_{2;t} &= s(L)v_t;
\end{aligned}
$$

such that

$$Y_{1;t} = \beta' Y_{2;t} + U_t; \qquad (12)$$

$$\Delta Y_{2;t} = V_t; \qquad (13)$$

where $Y_{1;t}$; $Y_{2;t}$; U_t, and V_t are the overlapping partial sums of $y_{1;t}$, $y_{2;t}$, u_t, and v_t. Next, systematic sampling is applied to equations (12) and (13) such that

$$Y_{1;\tau}^{as} = \beta' Y_{2;\tau}^{as} + U_\tau^{as};$$

$$\Delta Y_{2;\tau}^{as} = V_\tau^{as}: \qquad (14)$$

Like the preceding case, it is necessary to multiply the equation (14) by $s(L)$ to obtain the first-difference in τ-time; so finally the stochastic common trend is given by

$$\Delta_m Y_{2;\tau}^{as} = V_\tau^{as*};$$

where

$$V_\tau^{as*} = (s(L))^2 v_{m\tau}:$$

Proposition 2 *Let y_t be an n-dimensional vector time series generated by the CI(1,1) process (1)-(2), then the averaged sampled time series Y_τ^{as} follows a CI(1,1) process represented by*

$$
\begin{aligned}
Y_{1;\tau}^{as} &= \beta' Y_{2;\tau}^{as} + U_\tau^{as};\\
\Delta_m Y_{2;\tau}^{as} &= V_\tau^{as*};
\end{aligned}
$$

where

$$\xi_\tau^{as} = (U_\tau^{as\prime}, V_\tau^{as*\prime})' = S_{as}(L)\xi_{m\tau};$$

with the polynomial matrix $S_{as}(L)$ given by

$$S_{as}(L) = \begin{bmatrix} s(L)\cdot I_r & 0 \\ 0 & s^2(L)\cdot I_g \end{bmatrix};$$

and

$$s^2(L) = (s(L))^2 = 1 + 2L + \dots + mL^{m-1} + (m-1)L^m + \dots + L^{2m-2}:$$

3.3 Models of mixed sampled time series

The mixed sampling situation is common in empirical modelling of macroeconomic time series. The mixed sampled time series is denoted Y_τ^{ms}; and four cases of mixed sampling are considered: $y_{1;t}$ is systematically sampled and $y_{2;t}$ is averaged sampled, $y_{1;t}$ is averaged sampled and $y_{2;t}$ is systematically sampled $y_{1;t}$ is systematically sampled and $y_{2;t}$ is mixed sampled, and $y_{1;t}$ is averaged sampled and $y_{2;t}$ is mixed sampled. For the last two cases we consider the partition of the g-dimensional vector $y_{2;t}$ in two subvectors $y^1_{2;t}$ and $y^1_{2;t}$ of dimension g_1 and g_2, respectively, such that $y^1_{2;t}$ is averaged sampled and $y^2_{2;t}$ is systematically sampled.

3.3.1 $y_{1;t}$ is systematically sampled and $y_{2;t}$ is averaged sampled

To derive the representation of the temporally aggregated time series apply the following transformation on the cointegrating regression of the disaggregated model (1)

$$y_{1;t} = \left(s(L)\right)^{-1} \beta' s(L) y_{2;t} + u_t;$$

and multiply the polynomial $s(L)$ to (2)

$$s(L)\Delta y_{2;t} = s(L) v_t;$$

such that

$$y_{1;t} = m^{-1}\beta' Y_{2;t} + u_t. \qquad (15)$$

and

$$\Delta Y_{2;t} = V_t. \qquad (16)$$

Now, we apply sampling to equations (15) and (16) such that

$$Y^{ss}_{1;\tau} = m^{-1}\beta' Y^{as}_{2;\tau} + U^{ss}_\tau;$$
$$\Delta Y^{as}_{2;\tau} = V^{as}_\tau; \qquad (17)$$

and then apply $s(L)$ to (17) obtaining

$$\Delta_m Y^{as}_{2;\tau} = V^{as*}_\tau.$$

Denote $\xi^{ms1}_\tau = (U^{ss\prime}_\tau, V^{as*\prime}_\tau)'$, then $\xi^{ms1}_\tau = S_{ms1}(L)\xi_{m\tau}$; with

$$S_{ms1}(L) = \begin{bmatrix} I_r & 0 \\ 0 & s^2(L)\Delta I_g \end{bmatrix}.$$

3.3.2 $y_{1,t}$ is averaged sampled and $y_{2,t}$ is systematically sampled

To derive the temporal aggregated model apply $s(L)$ to (1)

$$s(L)y_{1,t} = s(L)\beta' y_{2,t} + s(L)u_t, \qquad (18)$$

and then systematic sampling to (18) obtaining

$$Y^{as}_{1,\tau} = m\beta' Y^{ss}_{2,\tau} + U^{as}_{\tau}.$$

On other hand, apply systematic sampling to equation (2) and then $s(L)$:

$$\Delta_m Y^{ss}_{2,\tau} = V^{as}_{\tau}.$$

The process ε^{ms2}_{τ} is

$$\varepsilon^{ms2}_{\tau} = (U^{as\prime}_{\tau}, V^{as\prime}_{\tau})' = S_{ms2}(L)\varepsilon_{m\tau},$$

with

$$S_{ms2}(L) = \begin{bmatrix} s(L)\cdot I_r & 0 \\ 0 & s(L)\cdot I_g \end{bmatrix}.$$

3.3.3 $y_{1,t}$ is systematically sampled and $y_{2,t}$ is mixed sampled

In this case, let write the DGP conformabily with the partition of $y_{2,t}$

$$\begin{aligned} y_{1,t} &= \beta_1' y^1_{2,t} + \beta_2' y^2_{2,t} + u_t, \\ \Delta y^1_{2,t} &= v^1_t, \\ \Delta y^2_{2,t} &= v^2_t, \end{aligned}$$

where $\beta = (\beta_1, \beta_2)$; $v_t = (v^{1\prime}_t, v^{2\prime}_t)'$, v^1_t is $g_1 \times 1$, and v^2_t is $g_2 \times 1$. Now partition ε_t, the n-dimensional standarized Wiener process, and the $\Sigma_{\varepsilon\varepsilon}$ matrix conformabily with $y_t = (y'_{1,t}, y^{1\prime}_{2,t}, y^{2\prime}_{2,t})'$:

$$\varepsilon_t = (u'_t, v^{1\prime}_t, v^{2\prime}_t)'$$

$$W_{\varepsilon} = [W_u', W_{v^1}', W_{v^2}']',$$

and

$$\Sigma_{\varepsilon\varepsilon} = \begin{bmatrix} \Sigma_{uu} & \Sigma_{uv^1} & \Sigma_{uv^2} \\ 0 & \Sigma_{v^1v^1} & \Sigma_{v^1v^2} \\ 0 & 0 & \Sigma_{v^2v^2} \end{bmatrix},$$

respectively.

$y_{1,t}$ and $y^2_{2,t}$ are systematically sampled, and $y^1_{2,t}$ is averaged sampled, so the representation for the temporally aggregated time series is given by:

$$\begin{aligned} Y^{ss}_{1,\tau} &= m^{-1}\beta_1' Y^{1\,as}_{2,\tau} + \beta_2' Y^{2\,ss}_{2,\tau} + U^{ss}_{\tau}, \\ \Delta_m Y^{1\,as}_{2,\tau} &= V^{1as*}_{\tau}, \\ \Delta_m Y^{2\,ss}_{2,\tau} &= V^{2\,as}_{\tau}, \end{aligned}$$

and

$$\xi^{ms3}_{\tau} = (U^{ss\prime}_{\tau}, V^{1\,as*\prime}_{\tau}, V^{2\,as\prime}_{\tau})' = S_{ms3}(L)\,\xi_{m\tau}$$

with

$$S_{ms3}(L) = \begin{bmatrix} I_r & 0 & 0 \\ 0 & s^2(L) \otimes I_{g_1} & 0 \\ 0 & 0 & s(L) \otimes I_{g_2} \end{bmatrix}.$$

3.3.4 $y_{1,t}$ is averaged sampled and $y_{2,t}$ is mixed sampled

Now, it is assumed that $y_{1,t}$ and $y^1_{2,t}$ are averaged sampled, and $y^2_{2,t}$ is systematically sampled. Then, the aggregated model is

$$\begin{aligned} Y^{as}_{1,\tau} &= \beta_1' Y^{1\,as}_{2,\tau} + m\,\beta_2' Y^{2\,ss}_{2,\tau} + U^{as}_{\tau}, \\ \Delta_m Y^{1\,as}_{2,\tau} &= V^{1\,as*}_{\tau}, \\ \Delta_m Y^{2\,ss}_{2,\tau} &= V^{2\,as}_{\tau}, \end{aligned}$$

and

$$\xi^{ms4}_{\tau} = (U^{as\prime}_{\tau}, V^{1\,as*\prime}_{\tau}, V^{2\,as\prime}_{\tau})' = S_{ms4}(L)\,\xi_{m\tau}$$

with

$$S_{ms4}(L) = \begin{bmatrix} s(L) \otimes I_r & 0 & 0 \\ 0 & s^2(L) \otimes I_{g_1} & 0 \\ 0 & 0 & s(L) \otimes I_{g_2} \end{bmatrix}.$$

4 Asymptotic distribution of OLS estimator for all temporal aggregation order of the variables

In this section we derive the a.d. of the OLS estimator for the different cases of temporal aggregation described above. The literature of temporal aggregation and unit roots is centered in unit root testing (see for example Shiller and Perron, 1985) and cointegration testing (see for example Lahiri and Mamingi, 1995). This literature analyzes the aggregation effects through Monte Carlo simulation, that although can provide some information on the direction of the effects, the results are always model-specific. The paper, as a first attempt studies the a.d. of OLS estimator when the vector time series y_t is temporally aggregated and y_t is the CI(1,1) process described in section 2.

The OLS estimator (Engle and Granger, 1987), consists in a simple static regression of the levels of the variables y_t. This simple method has the desirable property that is super-consistent (the estimator converging to the true value at rate T) (Stock, 1987). However the presence of serial correlation and/or cross-correlation in the variables, implies that this estimator is not asymptotically efficient and mixed normal distributed (AEMN) (Saikkonen, 1991), so an optimal inference theory is not applicable (see Phillips, 1991a). Moreover, the presence of non zero median terms in its distribution produces serious biases in finite samples (Banerjee et al., 1986). Thus, the OLS estimator is a sub-optimal method relative to AEMN estimators, like for example the full information maximum likelihood (see Johansen, 1988, 1991; and Ahn and Reinsel, 1990), the fully modified ordinary least squares (see Phillips and Hansen, 1990) or dynamic ordinary least squares (see Saikkonen, 1991; and Stock and Watson, 1993).

The reasons to study this no AEMN estimator are the following: first, the theoretical interest in temporal aggregation effects on a cointegrating estimator. Second, the OLS estimator is the basis for the construction of alternative procedures that try to deal with the correlation problems, like the FMLS, the canonical cointegrating regression of Park (1992), and the three step estimator of Engle and Yoo (1991). Third, the OLS residuals, $\hat{u}_t$, are used as the basis for some robuts cointegration tests (see Gonzalo and Lee, 1998) like the cointegration test of Phillips and Outliaris (1990), so the effects of temporal aggregation on the OLS estimator could affect the size and power of these tests.

To derive the distribution of OLS, it is assumed that $r = 1$, so there is one cointegrating vector and $g = n - 1$ common stochastic trends in the system:

$$\begin{aligned} y_{1,t} &= \beta' y_{2,t} + u_t, \\ \Delta y_{2,t} &= v_t, \end{aligned}$$

where $y_{1,t}$ is a scalar time series, β is a $g \times 1$ vector. The OLS estimator of β, $\hat{\beta}^{(m)}$, can be written (after centering and scaling) as:

$$T^{(m)}\left(\hat{\beta}^{(m)} - \beta\right) = \left[T^{(m)-2}\sum_{\tau=1}^{T^{(m)}} Y_{2,\tau}Y_{2,\tau}'\right]^{-1}\left[T^{(m)-1}\sum_{\tau=1}^{T^{(m)}} Y_{2,\tau-1}U_{\tau} + T^{(m)-1}\sum_{\tau=1}^{T^{(m)}} V_{\tau}U_{\tau}\right].$$

4.1 Multivariate invariance principle for temporally aggregated process

Because temporal aggregation is a linear transformation on ε_t, if ε_t satisfies an IP, then also ε_τ^i $(i = ss, as, ms)$ will satisfy an IP (see Li, 1998). The IP for temporally aggregated time series is established in the following proposition:

Proposition 3 (Invariance Principle for temporally aggregated process)
Suppose that ξ_t follows an n-dimensional random walk without drift

$$\xi_t = \xi_{t-1} + \varepsilon_t,$$

where »$_0$ ´ 0 and ³$_t$ is an n-dimensional covariance stationary process that satis..es the multivariate IP (3), then

$$T^{(m)\ i\ 1=2}\sum_{¿=1}^{[T^{(m)}¢]} ³^{i}_{¿} \ !^{L}\ B_{³ i} = P_{³ i ³ i}\ ¢W_{³}\ ´\ BM\ (-_{³ i ³ i});$$

where ³$^{i}_{¿}$= S$_i$(L)³$_{m¿}$, and -$_{³ i ³ i}$ is the long-run covariance matrix of ³$^{i}_{¿}$:

$$-_{³ i ³ i} = P_{³ i ³ i} P^{0}_{³ i ³ i};$$

where

$$P_{³ i ³ i} = m^{\ i\ 1=2}\ S_i(1)\ ¢P_{³ ³};$$

for i = as; ss; m s:

Proof. Let ³$_t$ be an n-dimensional covariance stationary process that satis..es the IP

$$T^{\ i\ 1=2}\sum_{t=1}^{[T¢]} ³_t \ !^{L}\ B_{³}; \qquad (19)$$

and ³$^{i}_{¿}$ = S$_i$(L)³$_{m¿}$;[3] a linear transformation of ³$_t$, then ³$^{i}_{¿}$ satis..es the IP

$$T^{(m)\ i\ 1=2}\sum_{¿=1}^{[T^{(m)}¢]} ³^{i}_{¿} \ !^{L}\ B_{³ i}: \qquad (20)$$

The relationship between the Wiener processes B$_³$ and B$_{³ i}$ is obtained after some substitutions in the left-hand side of (20):

$$\begin{aligned} T^{(m)\ i\ 1=2}\sum_{¿=1}^{[T^{(m)}¢]} ³^{i}_{¿} &= \mu \frac{T}{m} ¶^{\ i\ 1=2} \sum_{¿=1}^{[T=m\ ¢]} S_i(L)³_{m¿} \\ &= T^{\ i\ 1=2} m^{\ 1=2} m^{\ i\ 1} \sum_{t=1}^{[T¢]} S_i(L)³_t \\ &= m^{\ i\ 1=2} S_i(1) T^{\ i\ 1=2} \sum_{t=1}^{[T¢]} ³_t: \end{aligned}$$

This expression converges in distribution to

$$P_{³ i ³ i}\ ¢W;$$

where

$$P_{³ i ³ i} = m^{\ i\ 1=2}\ S_i(1)\ P_{³ ³};$$

[3]The proof is valid for any temporal aggregation scheme.

so the long-run covariance matrix of η_τ^{-i} is

$$\Omega_{\eta^{-i}\eta^{-i}} = \Pi_{\eta^{-i}\eta^{-i}}\Pi'_{\eta^{-i}\eta^{-i}}.$$

■

Now, define

$$\xi_{i\tau}^{(m)} = \xi_{i,\tau-1}^{(m)} + \eta_{i\tau}^{(m)}; \quad i = 1; 2; m\,1;$$

then the following asymptotic results are necessary to derive the asymptotic distributions of OLS estimators:

1. $T^{(m)-1}\sum_{\tau=1}^{T^{(m)}} \eta_\tau^{-i}\eta_\tau^{-i\prime} \xrightarrow{p} \Gamma_{\eta^{-i}\eta^{-i}}(0)$;
2. $T^{(m)-2}\sum_{\tau=1}^{T^{(m)}} \xi_{\tau-1}^{-i}\xi_{\tau-1}^{-i\prime} \xrightarrow{L} \Pi_{\eta^{-i}}\left(\int W_\eta W_\eta'\right)\Pi'_{\eta^{-i}}$;
3. $T^{(m)-1}\sum_{\tau=1}^{T^{(m)}} \xi_{\tau-1}^{-i}\eta_\tau^{-i\prime} \xrightarrow{L} \Pi_{\eta^{-i}}\left(\int W_\eta dW_\eta'\right)\Pi'_{\eta^{-i}} + \Lambda_{\eta^{-i}\eta^{-i}}$.

4.2 Systematic sampling

From proposition 3, the $\Pi_{\eta^1\eta^1}$ matrix and the long-run variance matrix of η_τ^1 are given by

$$\Pi_{\eta^1\eta^1} = \begin{bmatrix} m^{-1/2}\cdot\Pi_u \\ m^{1/2}\cdot\Pi_v \end{bmatrix};$$

$$\Omega_{\eta^1\eta^1} = \begin{bmatrix} m^{-1}\Omega_{uu} & \Omega_{uv} \\ \Omega_{vu} & m\,\Omega_{vv} \end{bmatrix};$$

respectively. Then we have

$$T^{(m)-2}\sum_{\tau=1}^{T^{(m)}} y_{2,\tau}^{1}y_{2,\tau}^{1\prime} \xrightarrow{L} m\,\Pi_{vv}\left(\int W_vW_v'\right)\Pi_{vv}';$$

$$T^{(m)-1}\sum_{\tau=1}^{T^{(m)}} y_{2,\tau-1}^{1}u_\tau^{1} \xrightarrow{L} \Pi_v\left(\int W\,dW'\right)\Pi_u' + \Lambda_{v^{as}u^{1}};$$

where $\Lambda_{v^{as}u^{1}} = \sum_{j=1}^{\infty}\Gamma_{v^{as}u^{1}}(-j)$ and assume that

$$T^{(m)-1}\sum_{\tau=1}^{T^{(m)}} v_\tau^{as}u_\tau^{1} \xrightarrow{p} \Gamma_{v^{as}u^{1}}(0).$$

The distribution of $T^{(m)}(\hat\beta_1^{(m)} - \beta)$ is

$$T^{(m)}\left(\hat\beta_1^{(m)} - \beta\right) \xrightarrow{L} \left[\Pi_{vv}\left(\int W_vW_v'\right)\Pi_{vv}'\right]^{-1}\left[m^{-1}\Pi_v\left(\int W\,dW'\right)\Pi_u' + m^{-1}\cdot\Lambda_{v^{as}u^{1}}\right];$$

where $\Delta_{v_{ss}u_{ss}} = \Gamma_{v_{ss}u_{ss}}(0) + \Lambda_{v_{ss}u_{ss}}$. The nuisance parameter $\Delta_{v_{ss}u_{ss}}$ does not depend on temporal aggregation order because

$$\Gamma_{v_{ss}u_{ss}}(-j) = \sum_{i=jm}^{(j+1)m-1} \Gamma_{vu}(-i)$$

so

$$\Delta_{v_{ss}u_{ss}} = \sum_{j=0}^{\infty} \Gamma_{vu}(-j) = \Delta_{vu}. \tag{21}$$

and this implis that the a.d. of $T^{(m)}(\hat{\beta}^{(m)}_{ss} - \beta)$ is

$$T^{(m)}\left(\hat{\beta}^{(m)}_{ss} - \beta\right) \xrightarrow{L} \left[P_{vv}\left(\int W_v W_v'\right) P_{vv}'\right]^{-1} \left[m^{-1} P_v \left(\int W dW'\right) P_u' + m^{-1}\Delta_{vu}\right]; \tag{22}$$

The distribution (22) shows the joint effect of temporal aggregation and the span of the sample in the estimator, because it is expressed in terms of the sample size $T^{(m)}$, so for a proper analysis of the effects of temporal aggregation on the distribution of $\hat{\beta}^{(m)}_{ss}$; we must use the distribution of $T(\hat{\beta}^{(m)}_{ss} - \beta)$ instead of (22); that is in terms of the same span of the sample. So we must multiply (22) by m; and then the a.d. of $T(\hat{\beta}^{(m)}_{ss} - \beta)$ is given by

$$T\left(\hat{\beta}^{(m)}_{ss} - \beta\right) \xrightarrow{L} \left[P_{vv}\left(\int W_v W_v'\right) P_{vv}'\right]^{-1} \left[P_v \left(\int W dW'\right) P_u' + \Delta_{v_{ss}u_{ss}}\right]. \tag{23}$$

As it is seen in (23), the a.d. of OLS for systematically sampled time series does not depend on the sampling interval m.

Proposition 4 Let y_t be an n-dimesional time series generated by the CI(1,1) process (1)-(2), asymptotically the OLS estimator of the cointegrating vector β is equivalent for all finite temporal aggregation order of the systematically sampled time series Y^{ss}_{τ}.

4.3 Average sampling

The $P_{\xi_{as}\xi_{as}}$ matrix is given by

$$P_{\xi_{as}\xi_{as}} = \begin{bmatrix} m^{1/2}\,\Delta P_u \\ m^{3/2}\,\Delta P_v \end{bmatrix};$$

and the long-run variance matrix is

$$\Omega_{\xi_{as}\xi_{as}} = \begin{bmatrix} m\,\Omega_{uu} & m^{2}\Omega_{uv} \\ m^{2}\Omega_{vu} & m^{3}\Omega_{vv} \end{bmatrix}.$$

From the asymptotic results 1-3 of subsection 4.1

$$T^{(m)-1}\sum_{\tau=1}^{T^{(m)}} V_{\tau}^{as*}U_{\tau}^{as} \to_p \Gamma_{V^{as*}U^{as}}(0);$$

$$T^{(m)-2}\sum_{\tau=1}^{T^{(m)}} Y_{2,\tau}^{as}Y_{2,\tau}^{as\prime} \Rightarrow m^{3}P_v\left(\int WW'\right)P_v';$$

$$T^{(m)-1}\sum_{\tau=1}^{T^{(m)}} Y_{2,\tau-1}^{as}U_{\tau}^{as} \Rightarrow m^{2}P_v\left(\int WdW'\right)P_u' + \Lambda_{V^{as*}U^{as}};$$

where $\Lambda_{V^{as*}U^{as}} = \sum_{j=1}^{\infty}\Gamma_{V^{as*}U^{as}}(-j)$. Then,

$$T^{(m)}\left(\hat{\beta}_{as}^{(m)} - \beta\right) \Rightarrow \left[P_{vv}\left(\int W_vW_v'\right)P_{vv}'\right]^{-1}\left[m^{-1}P_v\left(\int WdW'\right)P_u' + m^{-3}\Delta_{V^{as*}U^{as}}\right]; \tag{24}$$

where $\Delta_{V^{as*}U^{as}} = \Gamma_{V^{as*}U^{as}}(0) + \Lambda_{V^{as*}U^{as}}$; and in terms of the covariances $\Gamma_{vu}(j)$ can be written as

$$\begin{aligned}\Delta_{V^{as*}U^{as}} = \; & \Gamma_{vu}(m-1) + (1+2)\Gamma_{vu}(m-2) + \ldots + (1+2+\ldots+m)\Gamma_{vu}(0)\\ & + \ldots + (1+2+\ldots+m+(m-1)+\ldots+2)\Gamma_{vu}(-m+2)\\ & + (1+2+\ldots+m+(m-1)+\ldots+1)\sum_{j=0}^{\infty}\Gamma_{vu}(-m+1-j).\end{aligned} \tag{25}$$

In terms of the same span of the sample (T),

$$T\left(\hat{\beta}_{as}^{(m)} - \beta\right) \Rightarrow \left[P_{vv}\left(\int W_vW_v'\right)P_{vv}'\right]^{-1}\left[P_v\left(\int WdW'\right)P_u' + m^{-2}\Delta_{V^{as*}U^{as}}\right]. \tag{26}$$

Unlike the previous case, now for average sampling, temporal aggregation order m affect the impact of the endogeneity in the distribution through the nuisance parameter $m^{-2}\Delta_{V^{as*}U^{as}}$ included in the 'simultaneous equation bias' (s.e.b) term:

$$m^{-2}\left[P_{vv}\left(\int W_vW_v'\right)P_{vv}'\right]^{-1}\Delta_{V^{as*}U^{as}};$$

a non zero mean distribution. Therefore, temporal aggregation can affect not only the variance but the bias of the OLS estimator. Concretely, average sampling can reduce the impact of the s.e.b. term in the bias and variance of the OLS if the next proposition is fulfilled, however average sampling will not eliminate the term and does not affect the unit root term, so although average sampling can improve the finite sample properties of OLS estimator, the OLS estimation with temporally aggregated data will not be an AEMN estimator.

Note that when $\Omega_{uv} = 0$; and so $P_{uv} = 0$; the OLS estimator is an AEMN estimator, like the full information maximum likelihood in a Cointegrated VAR system of Johansen (1988, 1991) and Ahn and Reinsel (1990). In this case temporal aggregation has no effects on the estimator.

Proposition 5 *Let y_t be an n-dimesional time series generated by the CI(1,1) system (1)-(2), then asymptotically the OLS estimator of the cointegrating vector β is less biased and more efficient when is estimated with averaged sampled data, if and only if*

$$\Delta_{V^{as*}U^{as}} < m^2 \Delta_{vu}.$$

As it is shown in the distribution (26), the OLS estimator with averaged sampled time series will depend on temporal aggregation through the nuisance parameter $m^{-2}\Delta_{V^{as*}U^{as}}$, all covariances between the present permanent component, V^{as*}_{τ}; and all leads of the temporary component, $U^{as}_{\tau+j}$ for $j = 0; 1; \ldots$. This term depends on the covariances $\Gamma_{vu}(-i)$ and $\Gamma_{vu}(j)$ for $i = 0; 1; 2; \ldots;$ and $j = 1; 2; \ldots; m-1$;[4] Comparing the expression (25) with the s.e.b. term of the a.d. for the disaggregated case Δ_{vu}, the following remarks establish the situations where the OLS estimation is not improved with average sampling:

Remark 6 The bias and variance of the OLS estimator will increase with the temporal aggregation order of the variables if and only if: i) $\Gamma_{vu}(-i) = 0$ $(i = 0; 1; 2; \ldots;)$ and ii) there exists some $j = 1; 2; \ldots; m-1$ such that $\Gamma_{vu}(j) \neq 0$:

Remark 7 The bias and variance of the OLS estimator will not depend on the temporal aggregation order if and only if: i i) $\Gamma_{vu}(-i) = 0$ $(i = 0; 1; 2; \ldots)$ and $\Gamma_{vu}(j) = 0$ $(j = 1; 2; \ldots; m-1)$ or ii) $\Delta_{V^{as*}U^{as}} = m^2 \Delta_{vu}$:

Remarks 6 and 7 are mathematically possible, but empirically not very probable, anyway a test on the significance of these covariances can be used to discard these remarks. Because the nuisance parameter $\Delta_{V^{as*}U^{as}}$ is a matrix, a deeper analysis is necessary.

Example 8 Consider the trivariate model

$$\begin{aligned} y_{1;t} &= \beta y_{2;t} + \alpha y_{3;t} + u_t; \\ \Delta y_{2;t} &= v_{1;t}; \\ \Delta y_{3;t} &= v_{2;t}; \end{aligned}$$

where $\zeta_t = (u_t; v_{1;t}; v_{2;t})$ is a covariance stationary process, the simulatenous equation bias can be expressed as

[4] Note the difference between $\Gamma_{vu}(-1) = E(v_t u_{t+1}) = E(u_t v_{t-1})$ and $\Gamma_{vu}(1) = E(v_t u_{t-1})$:

$$\begin{bmatrix} a_{11}\Delta_{v_1 u} + a_{12}\Delta_{v_2 u} \\ a_{21}\Delta_{v_1 u} + a_{22}\Delta_{v_2 u} \end{bmatrix}$$

where we assume that

$$\left[P_v\left(\int W_v W_v'\right)P_v'\right]^{-1} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix};$$

and for averaged sampled time series of order m is

$$\begin{bmatrix} a_{11}\frac{1}{m^2}\Delta_{v_1^{as}u^{as}} + a_{12}\frac{1}{m^2}\Delta_{v_2^{as}u^{as}} \\ a_{21}\frac{1}{m^2}\Delta_{v_1^{as}u^{as}} + a_{22}\frac{1}{m^2}\Delta_{v_2^{as}u^{as}} \end{bmatrix}.$$

Then, if we compare the OLS estimator of β, we must compare the nuisance parameter

$$a_{11}\Delta_{v_1 u} + a_{12}\Delta_{v_2 u}$$

with

$$a_{11}\frac{1}{m^2}\Delta_{v_1^{as}u^{as}} + a_{12}\frac{1}{m^2}\Delta_{v_2^{as}u^{as}},$$

and, so the following situations can occur (term 1: $\frac{1}{m^2}\Delta_{v_1^{as}u^{as}}$, term 2: $\frac{1}{m^2}\Delta_{v_2^{as}u^{as}}$) : i) If the two terms satisfy proposition 5, temporal aggregation improves the OLS estimation of β; ii) if only term 1 satisfies proposition 5 (without loss of generality), and remark 7 holds for term 2, temporal aggregation improves the OLS estimation of β; iii) if only one term satisfies proposition 5, and the other holds remark 6, it can occur that although term 2 increases the bias, the reduction of the bias because the term 2 is bigger and temporal aggregation improves the estimation; iv) the reduction of the bias because term 2 is not enough to compensate the increase in the bias implied through term 1, and so temporal aggregation implies a worst estimation; and v) if the proposition 5 is not satisfied by any term then temporal aggregation leads to a worst estimation. In practice the most probable situation is the first case whatever the dimension of the system.

Example 9 For example, the nuisance parameter when the time series is averaged sampled two periods is $-\frac{1}{4}\gamma_{vu}(1)+\frac{3}{4}\gamma_{vu}(0)+\sum_{i=1}^{\infty}\gamma_{vu}(-i)$; and three periods $-\frac{1}{9}\gamma_{vu}(2)-\frac{1}{3}\gamma_{vu}(1)+\frac{2}{3}\gamma_{vu}(0)+\frac{8}{9}\gamma_{vu}(-1)+\sum_{i=2}^{\infty}\gamma_{vu}(-i)$. So, the most probable situation is that temporal aggregation improves the behaviour of the OLS estimator. As it is shown in the next example, the s.e.b. term is the main source of bias and variance in the estimator so average sampling may have an important impact on the precision of the estimator.

Example 10 Consider the bivariate cointegrated system used by many authors in their simulations and theoretical analysis (Banerjee et al., 1986; Engle and Granger, 1987; or Gonzalo, 1994):

$$y_{1,t} = \beta y_{2,t} + u_t,$$

$$\Delta y_{2,t} = v_t.$$

where $\zeta_t = (u_t, v_t)'$ is an i.i.d. sequence with covariance matrix

$$\Gamma_{\zeta\zeta}(0) = \begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix};$$

so the long-run variance matrix $\Omega_{\zeta\zeta} = \Gamma_{\zeta\zeta}(0)$. The OLS estimator of β is given by

$$\hat{\beta}^{(m)} = \frac{\sum_{\tau=1}^{T^{(m)}} y^{as}_{2,\tau} y^{as}_{1,\tau}}{\sum_{\tau=1}^{T^{(m)}} y^{as2}_{2,\tau}}; \qquad (27)$$

and the limiting distribution of (27) for all temporal aggregation order of the variables when the temporal aggregation is obtained through average sampling in terms of $T^{(m)}(\hat{\beta}^{(m)} - \beta)$ is given by:

$$T^{(m)}(\hat{\beta}^{(m)} - \beta) \xrightarrow{L} \frac{1}{m}\frac{(1-\rho^2)^{1/2}}{\sigma}\frac{Q_1}{Q_2} + \frac{1}{m}\frac{\gamma}{\sigma}\frac{Q_3}{Q_2} + \frac{m+1}{2m^2}\frac{\gamma}{\sigma}\frac{1}{Q_2}; \qquad (28)$$

and in terms of $T(\hat{\beta}^{(m)} - \beta)$ by:

$$T(\hat{\beta}^{(m)} - \beta) \xrightarrow{L} \frac{(1-\rho^2)^{1/2}}{\sigma}\frac{Q_1}{Q_2} + \frac{\gamma}{\sigma}\frac{Q_3}{Q_2} + \frac{m+1}{2m}\frac{\gamma}{\sigma}\frac{1}{Q_2}; \qquad (29)$$

where $Q_1 = \int W_v dW_u$, $Q_2 = \int W_v dW_v$, $Q_3 = \int W_v^2$, W_u and W_v are two independent scalar standarized Wiener processes. The distribution (28) shows the jointly effect of the span and temporal aggregation on the estimator, and the distribution (29) shows the effect of temporal aggregation when the span of the sample is fixed. The second case is more interesting, and shows that for a fixed span of the sample temporal aggregation will reduce the bias and variance of the OLS estimator through the term:

$$\frac{m+1}{2m}\frac{\gamma}{\sigma}\frac{1}{Q_2}.$$

In fact, this improve in the precision of the estimator is obtained reducing the sample size. It is interesting to compare the exact mean and variance of (29).[5] These moments are:

$$TE(\hat{\beta}^{(m)} - \beta) = \frac{\gamma}{\sigma}\left(-1.78143 + \frac{m+1}{2m}5.56286\right);$$

[5] For this purpose the exact moments of the ratios Q_1/Q_2, Q_1/Q_2^2; $(Q_1/Q_2)^2$; Q_3/Q_2; $1/Q_2$; $(Q_3/Q_2)^2$; $(1/Q_2)^2$; (Q_3/Q_2^2); and Q_1Q_3/Q_2 are needed. $E(Q_1/Q_2) = 0$ because the mixed gaussian distribution is centered at zero, $E(Q_1/Q_2^2) = 0$; and $E(Q_1Q_3/Q_2) = 0$. For $E(Q_3/Q_2)$; $E(1/Q_2)$; $E(Q_3/Q_2)^2$; $E(1/Q_2)^2$ and $E(Q_3/Q_2^2)$ see Gonzalo and Pitarakis (1998, Lemma 3.1); and for $E(Q_1/Q_2)^2$ see Abadir and Paruolo (1994, Theorem 1).

$$
\begin{aligned}
TVar(\hat{\beta}^{(m)} - \beta) = {} & 5.196 + \left(8.090 - 45.580\frac{m+1}{2m} + 67.831\left(\frac{m+1}{2m}\right)^2\right)\alpha \\
& + \left(-6.347 + 39.639\frac{m+1}{2m} - 61.891\left(\frac{m+1}{2m}\right)^2\right)\frac{\alpha}{\sigma} \\
& + \left(3.173 + 11.126\frac{m+1}{2m}\right)\frac{\alpha}{\sigma^2}.
\end{aligned}
$$

Letting $\gamma = 0.5$, and $\sigma = 0.5$, figures 4.1 and 4.2 represents $TE(\hat{\beta}^{(m)} - \beta)$ and $TVar(\hat{\beta}^{(m)} - \beta)$, respectively. As it is shown in both figures average sampling reduces the bias and variance of the OLS estimator, where the most important improvement in the estimator is produced when we move from $m = 1$ to $m = 2$.

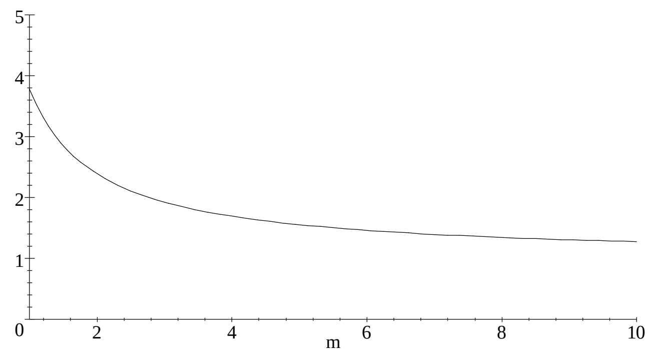

Figure 4.1 $TE(\hat{\beta}^{(m)} - \beta)$

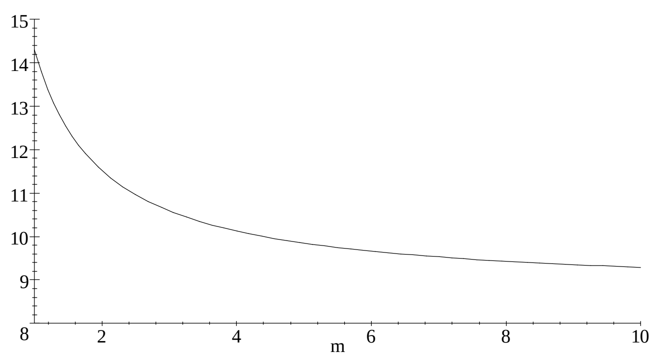

Figure 4.2 $TVar(\hat{\beta}^{(m)} - \beta)$

The aggregation effects of averaged sampling on the OLS estimator imply that in practice, can be better to estimate the cointegrating vector with averaged sampling data than with disaggregated data. A direct implication of this result is that the historical data, the sample with longer span and temporally aggregated data will provide more accurate estimation of the long-run relationship than the shorter span and higher frequency sample, because the span and the temporal aggregation have a positive impact on the precision of the OLS estimator of cointegrating vectors.

4.4 Mixed sampling

4.4.1 $y_{1,t}$ is systematically sampled and $y_{2,t}$ is averaged sampled

The matrix $P_{\xi^{ms1}\xi^{ms1}}$ is

$$P_{\xi^{ms1}\xi^{ms1}} = \begin{bmatrix} m^{-1/2} \cdot P_u \\ m^{3/2} \cdot P_v \end{bmatrix};$$

and the long-run variance of ξ_{τ}^{ms1} is

$$\Omega_{\xi^{ms1}\xi^{ms1}} = \begin{bmatrix} m^{-1}\Omega_{uu} & m\Omega_{uv} \\ m\Omega_{vu} & m^{3}\Omega_{vv} \end{bmatrix}.$$

This implies that the following quantities converge in distribution to

$$T^{(m)-2}\sum_{\tau=1}^{T^{(m)}} Y_{2,\tau}^{as} Y_{2,\tau}^{as\prime} \xrightarrow{L} m^{3} P_v \left(\int W W'\right) P_v',$$

$$T^{(m)-1}\sum_{\tau=1}^{T^{(m)}} Y_{2,\tau-1}^{as} U_{\tau}^{ss\prime} \xrightarrow{L} m P_v \left(\int W dW'\right) P_u' + \Delta_{V^{as}U^{ss}};$$

so the a.d. for $T^{(m)}(\hat{\beta}^{*(m)}_{ms1} - m^{-1}\beta)$ is

$$T^{(m)}\left(\hat{\beta}^{*(m)}_{ms1} - m^{-1}\beta\right) \xrightarrow{L} \left[P_{vv}\left(\int W_v W_v'\right) P_{vv}'\right]^{-1} \left[m^{-2} P_v \left(\int W dW'\right) P_u' + m^{-3}\Delta_{V^{as}U^{ss}}\right];$$

and for $T(\hat{\beta}^{(m)}_{ms1} - \beta)$ is

$$T\left(\hat{\beta}^{(m)}_{ms1} - \beta\right) \xrightarrow{L} \left[P_{vv}\left(\int W_v W_v'\right) P_{vv}'\right]^{-1} \left[P_v \left(\int W dW'\right) P_u' + m^{-1}\Delta_{V^{as}U^{ss}}\right].$$

where

$$\begin{aligned} &\Delta_{V^{as}U^{ss}} \\ &= \Gamma_{vu}(0) + 2\Gamma_{vu}(-1) + \dots + (m-1)\Gamma_{vu}(-(m-2)) + m\sum_{i=0}^{\infty}\Gamma_{vu}(-(m-1+i)); \end{aligned}$$

so the term $m^{-1}\Delta_{V^{as}U^{ss}}$ is smaller than Δ_{vu}.

Proposition 11 *Let y_t be an n-dimesional time series generated by the CI(1,1) process (1)-(2), then asymptotically the OLS estimator of the cointegrating vector β is more precise when it is estimated with mixed sampled (case 1) time series Y_{τ}^{ms1} than when is estimated with temporally disaggregated time series.*

4.4.2 $y_{1,t}$ is averaged sampled and $y_{2,t}$ is systematically sampled

$P_{3m\,s2\,3m\,s2}$ is

$$P_{3m\,s2\,3m\,s2} = \begin{bmatrix} m^{1/2} \cdot P_u \\ m^{1/2} \cdot P_v \end{bmatrix};$$

and the long-run variace matrix is

$$\Omega_{3m\,s2\,3m\,s2} = \begin{bmatrix} m\,\Omega_{uu} & m\,\Omega_{uv} \\ m\,\Omega_{vu} & m\,\Omega_{vv} \end{bmatrix};$$

so

$$T^{(m)-2}\sum_{\tau=1}^{T^{(m)}} Y^{ss}_{2,\tau} Y^{ss\prime}_{2,\tau} \to^{L} m\, P_v \left(\int W W'\right) P_v';$$

$$T^{(m)-1}\sum_{\tau=1}^{T^{(m)}} Y^{ss}_{2,\tau-1} U^{as}_{\tau} \to^{L} m\, P_v \left(\int W\, dW'\right) P_u' + \ast_{v^{as}u^{as}}:$$

The a.d. for $T^{(m)}(\hat{\beta}^{*(m)}_{m\,s2} - m\beta)$ is

$$T^{(m)}\left(\hat{\beta}^{*(m)}_{m\,s2} - m\beta\right) \to^{L} \left[P_{vv}\left(\int W_v W_v'\right) P_{vv}'\right]^{-1}\left[P_v\left(\int W\, dW'\right) P_u' + m^{-1}\Delta_{v^{as}u^{as}}\right];$$

and for $T(\hat{\beta}^{(m)}_{m\,s2} - \beta)$

$$T\left(\hat{\beta}^{(m)}_{m\,s2} - \beta\right) \to^{L} \left[P_{vv}\left(\int W_v W_v'\right) P_{vv}'\right]^{-1}\left[P_v\left(\int W\, dW'\right) P_u' + m^{-1}\Delta_{v^{as}u^{as}}\right];$$

where

$$\Delta_{v^{as}u^{as}} = \gamma_{vu}(m-1) + 2\gamma_{vu}(m-2) + \ldots + (m-1)\gamma_{vu}(1) + m\,\Delta_{vu}:$$

Unlike the preceding case of mixed sampling, in this case the nuisance parameter $m^{-1}\Delta_{v^{as}u^{as}}$ is bigger than Δ_{vu}, because the presence of the terms

$$\frac{1}{m}\gamma_{vu}(m-1) + \frac{2}{m}\gamma_{vu}(m-2) + \ldots + \frac{(m-1)}{m}\gamma_{vu}(1):$$

Proposition 12 *Let y_t be an n-dimesional time series generated by the CI(1,1) process (1)-(2), then asymptotically the OLS estimator of the cointegrating vector β is more precise when is estimated with mixed sampled (case 2) time series $Y^{m\,s2}_{\tau}$ than when is estimated with temporally disaggregated time series.*

4.4.3 $y_{1,t}$ is systematically sampled and $y_{2,t}$ is mixed sampled

The matrix $P_{3ms3\,3ms3}$ is now expressed partitioned conformabily with $y_{1,t}$, $y^1_{2,t}$, and $y^2_{2,t}$

$$P_{3ms3\,3ms3} = \begin{bmatrix} m^{-1/2} \cdot P_u \\ m^{3/2} \cdot P_{v^1} \\ m^{1/2} \cdot P_{v^2} \end{bmatrix};$$

and the long-run variance matrix is

$$\Omega_{3ms3\,3ms3} = \begin{bmatrix} m^{-1}\Omega_{uu} & m\,\Omega_{uv^1} & \Omega_{uv^2} \\ m\,\Omega_{v^1u} & m^{3}\Omega_{v^1v^1} & m^{2}\Omega_{v^1v^2} \\ \Omega_{v^2u} & m^{2}\Omega_{v^2v^1} & m\,\Omega_{v^1v^1} \end{bmatrix}$$

where $P_{v^1} = [0; P_{v^1v^1}; P_{v^1v^2}]$ and $P_{v^2} = [0; 0; P_{v^2v^2}]$, so

$$T^{(m)-2}\sum_{\tau=1}^{T^{(m)}} \begin{bmatrix} y^{1(m)}_{as\,2,\tau} \\ y^{2(m)}_{ss\,2,\tau} \end{bmatrix} \begin{bmatrix} y^{1(m)\prime}_{as\,2,\tau} & y^{2(m)\prime}_{ss\,2,\tau} \end{bmatrix} \xrightarrow{L} m^{1/2}\begin{bmatrix} m & 0 \\ 0 & 1 \end{bmatrix} P_v \left(\int W W'\right) P_v' \begin{bmatrix} m & 0 \\ 0 & 1 \end{bmatrix} m^{1/2};$$

$$T^{(m)-1}\sum_{\tau=1}^{T^{(m)}} \begin{bmatrix} Y^{1as}_{2,\tau-1} \\ Y^{2ss}_{2,\tau-1} \end{bmatrix} U^{ss}_{\tau} \xrightarrow{L} m^{1/2}\begin{bmatrix} m & 0 \\ 0 & 1 \end{bmatrix} P_v \left(\int W\,dW'\right) P_u' m^{-1/2} + \begin{bmatrix} \Lambda_{V^{1as}*U^{ss}} \\ \Lambda_{V^{2as}U^{ss}} \end{bmatrix};$$

then the a.d. of $T^{(m)}(\hat{\beta}^{*(m)}_{ms3} - \beta^{(m)})$ is

$$T^{(m)}\left(\hat{\beta}^{*(m)}_{ms3} - \beta^{(m)}\right) \xrightarrow{L}$$

$$\left[m^{1/2}\begin{bmatrix} m & 0 \\ 0 & 1 \end{bmatrix} P_v \left(\int W W'\right) P_v' \begin{bmatrix} m & 0 \\ 0 & 1 \end{bmatrix} m^{1/2} \right]^{-1}$$

$$\times \left[m^{1/2}\begin{bmatrix} m & 0 \\ 0 & 1 \end{bmatrix} P_v \left(\int W W'\right) P_u' m^{-1/2} + \begin{bmatrix} \Delta_{V^{1as}*U^{ss}} \\ \Delta_{V^{2as}U^{ss}} \end{bmatrix} \right].$$

By introducing the habitual substitutions we obtain the a.d. of $T(\hat{\beta}^{(m)}_{ms3} - \beta)$

$$T\left(\hat{\beta}^{(m)}_{ms3} - \beta\right) \xrightarrow{L}$$

$$\left[\begin{bmatrix} m & 0 \\ 0 & 1 \end{bmatrix} \left[\begin{bmatrix} m & 0 \\ 0 & 1 \end{bmatrix} P_v \left(\int W W'\right) P_v' \begin{bmatrix} m & 0 \\ 0 & 1 \end{bmatrix} \right] \right]^{-1}$$

$$\times \left[\begin{bmatrix} m & 0 \\ 0 & 1 \end{bmatrix} P_v \left(\int W\,dW'\right) P_u' + \begin{bmatrix} \Delta_{V^{1as}*U^{ss}} \\ \Delta_{V^{2as}U^{ss}} \end{bmatrix} \right].$$

and then the limiting distribution depending on the temporal aggregation order is

$$T\left(\hat{\beta}^{(m)}_{ms3} - \beta\right) \xrightarrow{L} \left[\int P_{vv} W_v W_v' P_{vv}'\right]^{-1}\left[\int P_v W\, dW' P_u' + \begin{pmatrix} m^{-1}\Delta_{v_1^{as} u^{ss}} \\ \Delta_{v_2^{as} u^{ss}} \end{pmatrix}\right].$$

According with the preceding results the coefficients of the cointegrating vector associate with the flow regressors $y^1_{2,t}$, β_1, are estimated more precise at temporal aggregated level, and the coefficients associated with the stock regressors $y^2_{2,t}$, β_2, are estimated equally well with disaggregated than with temporally aggregated data, so the whole vector β must be estimated with temporal aggregated time series.

4.4.4 $y_{1,t}$ is averaged sampled and $y_{2,t}$ is mixed sampled

The matrix $P_{3ms4\,3ms4}$

$$P_{3ms4\,3ms4} = \begin{bmatrix} m^{1/2}\,\Delta P_u \\ m^{3/2}\,\Delta P_{v_1} \\ m^{1/2}\,\Delta P_{v_2} \end{bmatrix},$$

implies the following convergences in distribution

$$T^{(m)-2}\sum_{\tau=1}^{T^{(m)}}\begin{bmatrix} Y^{1as}_{2,\tau} \\ Y^{2ss}_{2,\tau} \end{bmatrix}\begin{bmatrix} Y^{1as\prime}_{2,\tau} & Y^{2ss\prime}_{2,\tau} \end{bmatrix} \xrightarrow{L} m^{1/2}\begin{bmatrix} m & 0 \\ 0 & 1 \end{bmatrix} P_v \left(\int W W'\right) P_v' \begin{bmatrix} m & 0 \\ 0 & 1 \end{bmatrix} m^{1/2},$$

$$T^{(m)-1}\sum_{\tau=1}^{T^{(m)}}\begin{bmatrix} Y^{1as}_{2,\tau-1} \\ Y^{2ss}_{2,\tau-1} \end{bmatrix} U^{as}_\tau \xrightarrow{L} m^{1/2}\begin{bmatrix} m & 0 \\ 0 & 1 \end{bmatrix} P_v \left(\int W\, dW'\right) P_u' m^{1/2} + \begin{bmatrix} \Lambda_{v_1^{as} u^{as}} \\ \Lambda_{v_2^{as} u^{as}} \end{bmatrix},$$

that after some transformations and substitutions the limiting distribution of $T\left(\hat{\beta}^{(m)}_{ms4} - \beta\right)$ is

$$T\left(\hat{\beta}^{(m)}_{ms4} - \beta\right) \xrightarrow{L} \left[\int P_{vv} W_v W_v' P_{vv}'\right]^{-1}\left[\int P_v W\, dW' P_u' + \begin{pmatrix} m^{-2}\Delta_{v_1^{as} u^{as}} \\ m^{-1}\Delta_{v_2^{as} u^{as}} \end{pmatrix}\right].$$

Then, the coefficients of the flow regressors $y^1_{2,t}$, β_1, are estimated with more precision at temporal aggregated level, and the coefficients of the stock regressors $y^2_{2,t}$, β_2, are estimated more precise at disaggregated level.

5 A Monte Carlo study

The model considered in our Monte Carlo simulation is the bivariate cointegrated system:

$$\begin{aligned} y_{1,t} &= \beta y_{2,t} + u_t, \qquad t = 1, 2, \ldots, T \\ \Delta y_{2,t} &= v_t, \end{aligned}$$

where $u_t = \rho u_{t-1} + \varepsilon_t$; $|\rho| < 1$, ε_t is i.i.d. $N(0;1)$; v_t is i.i.d. $N(0,1)$ and $E(\varepsilon_t; v_t) = \gamma$; so the parameter space is $\beta \times \rho \times \gamma$. The values for these parameters are the following:

$$\beta = 2;\ \rho = \{0.6; 0.9\};\ \gamma = \{0; 0.75\};$$

such that four different DGPs will be considered: DGP1 ($\gamma = 0$, $\rho = 0.6$), DGP2 ($\gamma = 0$, $\rho = 0.9$); DGP3 ($\gamma = 0.75$, $\rho = 0.6$); and DGP4 ($\gamma = 0.75$, $\rho = 0.9$). The correlation γ determines the optimality (or not) when $\gamma = 0$ (or $\gamma = 1$) of the OLS estimator. The autoregressive parameter ρ determines the velocity of the adjustment to the long-run relationship. The spans of the sample considered are $T = \{50; 100; 150\}$ and the temporal aggregation orders $m = \{1; 2; 3; 4; 5; 6; 7; 8; 9; 10\}$.

The results for the systematic sampling case are shown in tables 1 to 4, and for average sampling in tables 5 to 8.[6] In all the simulations we have generated 5000 series of length $T + 50$, starting with $\varepsilon_0 = v_0 = 0$; and then discarding the initial 50 observations.[7]

The asymptotic distributions for the different temporal aggregation schemes are given by next expressions. For disaggregated (and systematically) sampled time series the a.d. of OLS is:

$$T(\hat\beta - \beta) \xrightarrow{L} A + \left(\frac{1}{1-\rho}\right)\gamma\frac{Q_3}{Q_2} + \left(\frac{1}{1-\rho}\right)\gamma\frac{1}{Q_2};$$

and for averaged sampled time series is:

$$T(\hat\beta^{(m)}_{as} - \beta) \xrightarrow{L} A + \left[\frac{m+1}{2m} + \dots + \frac{m^2-1}{m^2}\rho^{m-2} + \sum_{j=0}\rho^{(m-1+j)}\right]\gamma\frac{1}{Q_2}.$$

where

$$A = \left(\frac{1}{1-\rho}\right)\left(1-\alpha^2\right)^{1/2}\frac{Q_1}{Q_2} + \left(\frac{1}{1-\rho}\right)\gamma\frac{Q_3}{Q_2}.$$

The OLS estimator for DGP1 and DGP2 is AEMN for all temporal aggregation order of the variables and for any type of aggregation. So theoretically temporal aggregation will not affect the estimator. On the other hand, for DGP3 and DGP4 average sampling will improve the estimator, and systematic sampling will not affect the estimator. The estimators are compared in terms of the bias in mean (B.mea.), Root Mean Square Error (RMSE), and the concentration probability calculated by $Pr(|\hat\beta^{(m)}_i - \beta| < 0.05)$ (Pr.); that is the frequency that the estimation deviates from the population parameter by 0.05.

Tables 1 and 2 shown the simulation results for systematic sampling when OLS is an AEMN estimator. The estimation is not biased for all temporal aggregation

[6] The Monte Carlo results for mixed sampling are available upon request.

[7] The GAUSS programming language and its RNDN function were used to generate the pseudo-normal innovations.

order, but the precision of the estimator decreases with temporal aggregation when the size of sample is small, showing the sample size effect. When OLS is an inefficient estimator (tables 3 and 4), the bias is hardly reduced with temporal aggregation, and the precision decreases for $\rho = 0.6$; but in the case $\rho = 0.9$; the precision of OLS is almost the same.

In tables 5 and 6 ($\gamma = 0$) the results for average sampling and DGP1 and DGP2 (OLS is an AEMN estimator), respectively, show how the estimator is not biased for any aggregation level of the variables, according with the theoretical results and how the most accurate estimation is obtained with temporal disaggregated data, although the dispersion of the empirical distribution hardly decreases, showing the sample size effect. As it is observed in table 2 when the parameter ρ is closer to 1 although the bias of the estimator is similar the dispersion of the distribution increases with m. More interesting results are observed when the estimator is not AEMN (DGP3 and DGP4). The temporal aggregation effects for a fixed span of the sample are observed horizontally for every span. If we look at tables 7 and 8, it is observed that for a fixed span, temporal aggregation reduces the bias and the dispersion of the estimator. For example, the bias in mean in table 3 for case ($m = 1, T = 50, T^{(1)} = 50$) is 0.056 and goes to 0.0467 for ($m = 2, T = 50, T^{(2)} = 25$) and to 0.0415 for ($m = 3, T = 50, T^{(3)} = 16$): The reduction in the bias is almost the same for every span, but strongly depends on ρ (table 8). As it is seen in table 8 the reduction in the bias for $\rho = 0.9$ is quite inferior to the reduction for $\rho = 0.6$. Another interesting comparison can be done vertically in the table. When we move vertically we observe the effect of the span on the estimator, for fixed m. In fact, this is what is done in Monte Carlo analysis, ignoring any temporal aggregation effects. As it is seen, a bigger span reduces the bias and dispersion of the estimator, and the effect is more important than the effect of temporal aggregation according with some results in the literature of temporal aggregation and unit root and cointegration testing (see Shiller and Perron, 1985; or Lahiri and Mamingi, 1995). Following with the example (table 7), an increase at the same rate of the span, moving from ($m = 1, T = 50, T^{(1)} = 50$) to ($m = 1, T = 100, T^{(1)} = 100$) and ($m = 1, T = 150, T^{(1)} = 150$) implies a higher reduction of the bias, from 0.056 to 0.0387 and to 0.0286, respectively. This effect does not depend on the temporal aggregation order nor the value of the autoregressive parameter. Notice that, for a fixed span to obtain a reduction of bias similar to the obtained when moving from $T = 50$ to $T = 150$, we must estimate the cointegrating vector with averaged sampled data of order 7.

Two preceding cases show how it is possible to improve the estimator by reducing the sample size (for a fixed span and more aggregated data) or by increasing the sample size (for a fixed temporal aggregation order and a increasing span), because the temporal aggregation and the span have a positive effect on the statistical properties of the OLS estimator. Finally, it is possible to compare the behaviour of the OLS estimator for fixed a sample size, when the span and m are varying at the same rate, that is the join effect of the span and aggregation. In this case we can observe these effects on the diagonal lines of the tables, i.e., the line ($m=1, T=50, T^{(1)}=50$)-($m=2, T=100, T^{(2)}=50$)-($m=3, T=150, T^{(3)}=50$). In this case (table 7), the bias goes from 0.056 to 0.0325 and to 0.0215, showing the biggest improving of the estimator, almost at the theoretically rate T.

6 Conclusions

The objective of the paper was to analyze the OLS estimator of cointegrating vectors with temporally aggregated data. A triangular system representation of the different schemes of temporally aggregated time series allow us to derive the limiting distribution of the OLS estimator for all finite temporal aggregation order, and it is shown how temporal aggregation affects its limiting distribution, and how for some types of aggregation like average sampling, the OLS estimator with temporally aggregated data is better (in terms of bias and variance) than the estimator with disaggregated data, contrary to some results in the literature of temporal aggregation and unit roots. A Monte Carlo study is conducted and shows the consistency of the theoretical results for the average and systematic sampling. The main implications of this paper are: First, it is possible to improve the behaviour of OLS estimator of cointegrating vectors with average sampling. In fact, this transformation can be applied to flow variables as well to stock variables because it doesn't affect the cointegrating space. Second, OLS estimation with averaged sampled time series can be used as the basis of some AEMN estimators like FM-OLS or 3SLS, or as the basis of some robust cointegration tests like the cointegration test of Phillips and Outliaris.

7 References

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

Abadir, K.M., and Paruolo, P. (1994) The marginal density of bivariate cointegration estimators. Discussion paper in Economics 94/05, University of Exeter.

Banerjee, A., Dolado, J.J., Hendry, D.F., and Smith, G.W. (1986) Exploring equilibrium relationships in econometrics through static models: Some Monte Carlo evidence. Oxford Bulletin of Economics and Statistics 48, 253-77.

Brewer, K.R.W. (1973) Some consequences of temporal aggregation and systematic sampling for ARMA and ARMAX models. Journal of Econometrics 1, 133-154.

Choi, I. (1992) Effects of data aggregation on the power of tests for a unit root. Economics Letters 40, 397-401.

Choi, I., and Chung, B.S. (1995) Sampling frequency and the power of tests for a unit root: A simulation study. Economics Letters 49, 131-136.

Drost, F.C., and Nijman, T.E. (1993) Temporal aggregation of GARCH processes. Econometrica 61(4), 909-927.

Engle, R.F., and Granger, C.W.J. (1987) Cointegration and error correction: representation, estimation and testing. Econometrica 55, 251-76.

Engle, R.F., and Yoo, S. (1991) Cointegrated economic time series: an overview with new results. Ch.12, 237-266, in R.F. Engle, and Granger, C.W.J., (eds.), Long run economic relationships, readings in cointegration, Oxford: Oxford University Press.

Gonzalo, J. (1994) Five methods of estimating long-run relationships. Journal of Econometrics 60, 203-33.

Gonzalo, J., and Lee, T. (1998) Pitfalls in testing for long-run relationships. Journal of Econometrics 86, 129-154.

Gonzalo, J., and Pitarakis, J. (1998) On the exact moments of asymptotic distributions in a unstable AR(1) with dependent errors. International Economic Review 39-1, 71-88.

Granger, C.W.J. (1990) Aggregation of time-series variables: a survey. in Barker, T., and Pesaran, M.H. (eds.), Disaggregation in Econometric Modelling, London: Routledge.

Granger, C.W.J., and Siklos, P.L. (1995) Systematic sampling, temporal aggregation, seasonal adjustment, and cointegration. Theory and evidence. Journal of Econometrics 66, 357-69.

Gregory, A.W., and Hansen, B.E. (1996) Residual-based tests for cointegration in models with regime shifts. Journal of Econometrics 70, 99-126.

Hooker, M.A. (1993) Testing for cointegration. Power versus frequency of observation. Economics Letters 41, 359-362.

Johansen, S. (1988) Statistical analysis of cointegrating vectors. Journal of Economic Dynamics and Control 12, 231-54.

Johansen, S. (1991) Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. Econometrica 59, 1551-1581.

Lahiri, K., and Mamingi, N. (1995) Testing for cointegration: Power versus frequency of observation - another view. Economics Letters 49, 121-124.

Li, Y. (1998) Low-pass …ltered least squares estimators of cointegrating vectors. Journal of Econometrics 85, 289-316.

Lutkepohl, H. (1987) Forecasting aggregated vector ARMA process. Berlin: Springer-Verlag.

Marcellino, M. (1996) Some temporal aggregation issues in empirical analysis. Discussion paper 96-39, Department of Econmics, University of California, San Diego.

Marcellino, M. (1999) Some consequences of temporal aggregation in empirical analysis. Journal of Business and Economic Statistics 17, 129-136.

Park, J.Y. (1992) Canonical cointegrating regressions. Econometrica 60, 119-143.

Perron, P. (1989) Testing for the random walk: A simulation experiment when the sampling interval is varied. In: B. Raj (ed.), Advances in econometrics and modelling, Dordrecht: Kluwer Academic Publisher, 47-68.

Perron, P. (1991) Test consistency with varying sampling frequency. Econometric Theory 7, 341-368.

Phillips, P.C.B. (1991a) Optimal inference in co-integrated systems. Econometrica 59, 282-306.

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

Phillips, P.C.B., and Durlauf, S.N. (1986) Multiple time series regression with integrated processes. Review of Economic Studies 53, 473-95.

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

Phillips, P.C.B., and Outliaris, S. (1990) Asymptotic properties of residual based tests for cointegration. Econometrica 58, 165-193.

Saikkonen, P. (1991) Asymptotically efficient estimation of cointegrating regressions. Econometric Theory 7, 1-21.

Shiller, R.J., and Perron, P. (1985) Testing the random walk hypothesis: Power versus frequency of observation. Economics Letters 18, 381-386.

Stock, J.H. (1987) Asymptotic properties of least-squares estimators of co-integrating vectors. Econometrica 55, 1035-56.

Stock, J.H., and Watson, M. (1993) A simple estimator of cointegrating vectors in Higher Order Integrated Systems. Econometrica 61, 783-820.

Wei, W.W.S. (1981) The effects of systematic sampling on ARIMA models. Communications in Statistical-Theoretical Mathematics 10, 2389-2398.

Weiss, A. (1984) Systematic sampling and temporal aggregation in time series models. Journal of Econometrics 2, 1-16.

Table 1: Behaviour of OLS estimation of cointegrated vectors with systematic sampling (DGP1)

m	1	2	3	4	5	6	7	8	9	10
$[T^{(m)}]$	50	25	16	12	10	8	7	6	5	5
B.mea.	-0.0006	-0.0007	0.0002	0.0003	0.0003	0.0019	-0.0022	0.0011	0.0019	0.0009
RMSE	0.0056	0.0060	0.0077	0.0087	0.0093	0.0118	0.0137	0.0150	0.0209	0.0182
Pr.	0.6474	0.6364	0.5936	0.5874	0.573	0.5314	0.5222	0.504	0.4526	0.4668
$[T^{(m)}]$	100	50	33	25	20	16	14	12	11	10
B.mea.	-0.0005	-0.0005	-0.0007	-0.0008	-0.0006	-0.0007	-0.0009	-0.0017	-0.0006	-0.0005
RMSE	0.0022	0.0024	0.0025	0.0029	0.0033	0.0041	0.0045	0.0051	0.0052	0.0057
Pr.	0.8188	0.8036	0.7864	0.7684	0.751	0.7072	0.6966	0.6752	0.649	0.6524
$[T^{(m)}]$	150	75	50	37	30	25	21	18	16	15
B.mea.	0.0002	0.0004	0.0002	0.0004	-0.0002	0.0006	0.0002	0.0002	0.0003	0.0005
RMSE	0.0010	0.0012	0.0012	0.0015	0.0015	0.0018	0.0020	0.0024	0.0026	0.0028
Pr.	0.8962	0.8864	0.8784	0.856	0.8496	0.8288	0.8112	0.7846	0.7702	0.767

Table 2: Behaviour of OLS estimation of cointegrated vectors with systematic sampling (DGP2)

m	1	2	3	4	5	6	7	8	9	10
$[T^{(m)}]$	50	25	16	12	10	8	7	6	5	5
B.mea.	-0.0029	-0.0026	-0.0026	-0.0027	-0.0036	-0.0022	-0.0024	-0.0027	-0.0034	-0.0047
RMSE	0.0550	0.0558	0.0598	0.0627	0.0613	0.0682	0.0706	0.0750	0.0938	0.0852
Pr.	0.2378	0.2332	0.2294	0.2342	0.2328	0.2178	0.2286	0.2174	0.2148	0.2176
$[T^{(m)}]$	100	50	33	25	20	16	14	12	11	10
B.mea.	-0.0017	-0.0017	-0.0016	-0.0016	-0.0008	-0.0016	-0.0021	-0.0016	-0.0026	-0.0011
RMSE	0.0226	0.0227	0.0233	0.0234	0.0238	0.0256	0.0256	0.0272	0.0273	0.0284
Pr.	0.3644	0.3606	0.3576	0.356	0.3634	0.3524	0.3512	0.3512	0.3446	0.357
$[T^{(m)}]$	150	75	50	37	30	25	21	18	16	15
B.mea.	0.0018	0.0017	0.002	0.0018	0.002	0.0022	0.0022	0.0021	0.002	0.0026
RMSE	0.0123	0.0123	0.0125	0.0128	0.0127	0.0130	0.0137	0.0145	0.0147	0.0141
Pr.	0.4354	0.4368	0.437	0.4394	0.437	0.4448	0.435	0.436	0.43	0.4368

Table 3: Behaviour of OLS estimation of cointegrated vectors with systematic sampling (DGP 3)

m	1	2	3	4	5	6	7	8	9	10
$[T^{(m)}]$	50	25	16	12	10	8	7	6	5	5
B.mea.	0.055	0.0546	0.0535	0.0537	0.0524	0.0492	0.0494	0.0483	0.0502	0.0456
RMSE	0.0558	0.0554	0.0546	0.0549	0.0538	0.0511	0.0515	0.0511	0.0548	0.0497
Pr.	0.5694	0.5582	0.5524	0.5244	0.5218	0.498	0.4822	0.4706	0.4326	0.4494
$[T^{(m)}]$	100	50	33	25	20	16	14	12	11	10
B.mea.	0.0384	0.0379	0.0384	0.0374	0.0354	0.0375	0.036	0.0367	0.0351	0.0323
RMSE	0.0386	0.0381	0.0386	0.0377	0.0357	0.0379	0.0364	0.0372	0.0357	0.0330
Pr.	0.7188	0.7156	0.709	0.6988	0.6934	0.6596	0.6516	0.6394	0.6318	0.6216
$[T^{(m)}]$	150	75	50	37	30	25	21	18	16	15
B.mea.	0.0289	0.0287	0.0284	0.0284	0.0284	0.0275	0.0282	0.0275	0.0273	0.0259
RMSE	0.0290	0.0288	0.0285	0.0285	0.0285	0.0276	0.0283	0.0277	0.0275	0.0261
Pr.	0.817	0.8114	0.8054	0.8024	0.791	0.781	0.7658	0.7442	0.741	0.7384

Table 4: Behaviour of OLS estimation of cointegrated vectors with systematic sampling (DGP 4)

m	1	2	3	4	5	6	7	8	9	10
$[T^{(m)}]$	50	25	16	12	10	8	7	6	5	5
B.mea.	0.1724	0.1707	0.173	0.1692	0.1656	0.1671	0.166	0.1629	0.1648	0.1539
RMSE	0.1892	0.1876	0.1913	0.1875	0.1834	0.1873	0.1860	0.1855	0.1922	0.1765
Pr.	0.168	0.167	0.1692	0.175	0.1728	0.1672	0.1756	0.1754	0.1676	0.17
$[T^{(m)}]$	100	50	33	25	20	16	14	12	11	10
B.mea.	0.1274	0.1267	0.1272	0.1254	0.1253	0.1264	0.1242	0.1247	0.1233	0.122
RMSE	0.1323	0.1316	0.1322	0.1304	0.1303	0.1319	0.1295	0.1304	0.1290	0.1277
Pr.	0.2388	0.243	0.2406	0.2522	0.2462	0.246	0.2436	0.2406	0.2568	0.2592
$[T^{(m)}]$	150	75	50	37	30	25	21	18	16	15
B.mea.	0.1051	0.1048	0.1041	0.1046	0.1032	0.1026	0.1035	0.1039	0.1032	0.1018
RMSE	0.1076	0.1073	0.1066	0.1072	0.1057	0.1051	0.1061	0.1067	0.1060	0.1045
Pr.	0.3124	0.3126	0.3138	0.3116	0.317	0.3138	0.3184	0.3136	0.3146	0.322

Table 5: Behaviour of OLS estimation of cointegrated vectors with averaged sampling (DGP1)

m	1	2	3	4	5	6	7	8	9	10
$[T^{(m)}]$	50	25	16	12	10	8	7	6	5	5
B.mea.	0.002	0.002	0.002	0.0018	0.0019	0.0018	0.0019	0.0021	0.0024	0.002
RMSE	0.0064	0.0068	0.0076	0.0077	0.0074	0.0084	0.0082	0.0089	0.01087	0.0091
Pr.	0.6272	0.6228	0.6096	0.6132	0.6156	0.6004	0.607	0.5996	0.5842	0.6064
$[T^{(m)}]$	100	50	33	25	20	16	14	12	11	10
B.mea.	0.0011	0.0011	0.0011	0.0011	0.0012	0.001	0.001	0.0011	0.0012	0.0013
RMSE	0.0022	0.0022	0.0023	0.0023	0.0023	0.0024	0.0024	0.0025	0.0025	0.0026
Pr.	0.8162	0.8124	0.8102	0.8098	0.8106	0.7994	0.804	0.792	0.8016	0.8062
$[T^{(m)}]$	150	75	50	37	30	25	21	18	16	15
B.mea.	0.0001	0.0001	0.0001	0.0001	0.0001	0	0	0	0	0.0001
RMSE	0.0009	0.0010	0.0010	0.0010	0.0010	0.001	0.0011	0.0011	0.0011	0.0011
Pr.	0.905	0.9034	0.901	0.8982	0.9	0.8992	0.892	0.8882	0.8908	0.8952

Table 6 Behaviour of OLS estimation of cointegrated vectors with averaged sampling (DGP2)

m	1	2	3	4	5	6	7	8	9	10
$[T^{(m)}]$	50	25	16	12	10	8	7	6	5	5
B.mea.	0.0044	0.0045	0.0052	0.0053	0.0042	0.0061	0.0051	0.0061	0.0051	0.0055
RMSE	0.0584	0.0621	0.0683	0.0714	0.0704	0.0772	0.0780	0.0848	0.1003	0.0850
Pr.	0.2302	0.229	0.2222	0.2218	0.2246	0.2186	0.2196	0.2142	0.208	0.2162
$[T^{(m)}]$	100	50	33	25	20	16	14	12	11	10
B.mea.	0.0014	0.0014	0.0014	0.0012	0.0013	0.0017	0.0014	0.0015	0.0013	0.0012
RMSE	0.0239	0.0247	0.0257	0.0260	0.0266	0.0287	0.0285	0.0298	0.0291	0.0296
Pr.	0.3518	0.3504	0.3442	0.3492	0.348	0.3362	0.3364	0.3358	0.345	0.3438
$[T^{(m)}]$	150	75	50	37	30	25	21	18	16	15
B.mea.	0.0005	0.0005	0.0005	0.0005	0.0005	0.0006	0.0006	0.0005	0.0005	0.0006
RMSE	0.0125	0.0129	0.0131	0.0136	0.0135	0.0137	0.0145	0.0153	0.0155	0.0145
Pr.	0.4608	0.46	0.4594	0.4536	0.4566	0.4546	0.4514	0.443	0.443	0.453

Table 7: Behaviour of OLS estimation of cointegrated vectors with averaged sampling (DGP3)

m	1	2	3	4	5	6	7	8	9	10
$[T^{(m)}]$	50	25	16	12	10	8	7	6	5	5
B.mea.	0.056	0.0467	0.0415	0.0369	0.0331	0.0306	0.0283	0.0263	0.0252	0.0233
RMSE	0.0560	0.0474	0.0422	0.0376	0.0337	0.0313	0.0290	0.0271	0.0262	0.0241
Pr.	0.5564	0.582	0.5948	0.6036	0.6238	0.6208	0.6274	0.6226	0.606	0.6404
$[T^{(m)}]$	100	50	33	25	20	16	14	12	11	10
B.mea.	0.0387	0.0325	0.029	0.0259	0.0239	0.0225	0.0211	0.0198	0.0187	0.0176
RMSE	0.0389	0.0327	0.0291	0.0260	0.0240	0.0226	0.0212	0.0200	0.0188	0.0177
Pr.	0.7162	0.755	0.7742	0.7884	0.8	0.7934	0.807	0.8054	0.8194	0.8202
$[T^{(m)}]$	150	75	50	37	30	25	21	18	16	15
B.mea.	0.0286	0.0242	0.0215	0.0197	0.018	0.0168	0.0158	0.0151	0.0142	0.0135
RMSE	0.0287	0.0243	0.0215	0.0197	0.0180	0.0168	0.0158	0.0151	0.0142	0.0135
Pr.	0.8154	0.8452	0.8686	0.8804	0.8946	0.9002	0.9026	0.9	0.9058	0.9174

Table 8: Behaviour of OLS estimation of cointegrated vectors with averaged sampling (DGP4)

m	1	2	3	4	5	6	7	8	9	10
$[T^{(m)}]$	50	25	16	12	10	8	7	6	5	5
B.mea.	0.1636	0.1574	0.1546	0.1502	0.1452	0.1424	0.1385	0.1348	0.1353	0.1281
RMSE	0.1787	0.1726	0.1707	0.1664	0.1610	0.1596	0.1555	0.1534	0.1572	0.1469
Pr.	0.1808	0.1818	0.182	0.185	0.1904	0.19	0.1904	0.1908	0.1844	0.1872
$[T^{(m)}]$	100	50	33	25	20	16	14	12	11	10
B.mea.	0.1295	0.1249	0.1221	0.1185	0.1157	0.1155	0.1117	0.1104	0.1065	0.1042
RMSE	0.1348	0.1300	0.1270	0.1233	0.1204	0.1205	0.1164	0.1153	0.1111	0.1086
Pr.	0.2394	0.2432	0.2456	0.2514	0.2552	0.2506	0.2596	0.2578	0.2668	0.272
$[T^{(m)}]$	150	75	50	37	30	25	21	18	16	15
B.mea.	0.0987	0.0951	0.0924	0.0908	0.0879	0.0859	0.0849	0.084	0.0822	0.0789
RMSE	0.1008	0.0971	0.0943	0.0927	0.0897	0.0876	0.0867	0.0860	0.0840	0.0805
Pr.	0.3244	0.333	0.3382	0.3426	0.349	0.3556	0.3564	0.3594	0.3634	0.3716