

Temporal Aggregation and Ordinary Least Squares Estimation of Cointegrating Regressions

Gabriel Pons-Rotger^{a,b} Andreu Sanso-Rossello^a

Jordi Surinyach-Caralt^a

^aEconometrics Department

University of Barcelona

^bVisiting 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 Introduction¹

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 Hocker, 1993; or Lahiri and Maringi, 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 appropriate 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 ...rst to establish in Section 2, the generating mechanism of the disaggregated time series, a $CI(1,1)$ process, cointegrated process of order

¹I 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, $\overset{L}{\rightarrow}$ denotes convergence in distribution and $\overset{P}{\rightarrow}$ stands for convergence in probability. The Brownian motion $B(r)$ and the standardized 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 as $\int_0^1 W(r)dr$ more simply as $\int W$. Vector Brownian motion with covariance matrix Σ is written $BM(\Sigma)$. y_t denotes an n -dimensional time series measured at basic time units t , and $Y_{\tilde{t}}$ denotes an n -dimensional temporally aggregated time series measured at \tilde{t} , where $\tilde{t} = [tm]$ 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 , $t = 1, \dots, T$; partitioned as $y_t = (y_{1;t}^0, y_{2;t}^0)^0$, 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

$$y_{1;t} = \Gamma y_{1;t-1} + u_t \quad (1)$$

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

where the $r \times n$ matrix (Γ, Γ') is the matrix of cointegrating vectors, Φ is the difference operator in t time, $\Phi = 1 - L$, 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 $\Sigma_t = (u_t^0, v_t^0)^0$, and introduce the following assumptions:

1. $y_{1;t}$ and $y_{2;t}$ are not cointegrated processes,
2. Σ_t is a covariance stationary process with zero mean,
3. that satisfies the invariance principle (IP) (Phillips and Dunsmuir, 1988):

$$T^{-1/2} \sum_{t=1}^{[T\lambda]} \Sigma_t \overset{L}{\rightarrow} B_{\Sigma} \sim BM(\Sigma_{\infty}); \quad (3)$$

for $\lambda \in [0, 1]$; where B_{Σ} is an n -dimensional vector Wiener process with covariance matrix Σ_{∞} , the long run covariance matrix of Σ_t :

$$\Sigma_{\infty} = \lim_{T \rightarrow \infty} T^{-1} E \sum_{t=1}^{[T\lambda]} \Sigma_t \Sigma_t^0 \sum_{t=1}^{[T\lambda]} \Sigma_t^0 \Sigma_t \quad (4)$$

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^3 . 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- Σ_{33} will be very useful. First, let express Σ_{33} as

$$\Sigma_{33} = \Sigma_{33}(0) + \alpha_{33} + \alpha_{33}^0;$$

where

$$\alpha_{33} = \sum_{j=1}^{\infty} \Sigma_{33}(j);$$

and

$$\Sigma_{33}(j) = E(\epsilon_t^3 \epsilon_{t-j}^{30});$$

Another useful decomposition of Σ_{33} is the Cholesky factorization:

$$\Sigma_{33} = P_{33} P_{33}^0;$$

where P_{33} ; a triangular matrix, is the square root of Σ_{33} . With the Cholesky factorization we can express the Wiener process B_3 as the product

$$B_3 = P_{33} W_3;$$

where W_3 is a standard n -dimensional vector Wiener process.

Let partition B_3 ; W_3 ; Σ_{33} ; $\Sigma_{33}(j)$; α_{33} ; and P_{33} according with the dimensions of u_t and v_t :

$$\begin{aligned} B_3 &= \begin{pmatrix} B_u \\ B_v \end{pmatrix}; \\ W_3 &= \begin{pmatrix} W_u \\ W_v \end{pmatrix}; \\ \Sigma_{33} &= \begin{pmatrix} \Sigma_{uu} & \Sigma_{uv} \\ \Sigma_{vu} & \Sigma_{vv} \end{pmatrix}; \\ \Sigma_{33}(j) &= \begin{pmatrix} \Sigma_{uu}(j) & \Sigma_{uv}(j) \\ \Sigma_{vu}(j) & \Sigma_{vv}(j) \end{pmatrix}; \\ \alpha_{33} &= \begin{pmatrix} \alpha_{uu} & \alpha_{uv} \\ \alpha_{vu} & \alpha_{vv} \end{pmatrix}; \end{aligned}$$

and

$$P_{33} = \begin{pmatrix} P_{uu} & P_{uv} \\ 0 & P_{vv} \end{pmatrix};$$

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

$$\begin{aligned} P_{uu} &= \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}; \\ P_{uv} &= \begin{bmatrix} \alpha_{13} & \alpha_{14} \\ \alpha_{23} & \alpha_{24} \end{bmatrix}; \\ P_{vv} &= \begin{bmatrix} \alpha_{33} & \alpha_{34} \\ \alpha_{43} & \alpha_{44} \end{bmatrix}; \end{aligned}$$

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

$$\begin{aligned} B_u &= P_{uu} \Phi W_u + P_{uv} \Phi W_v; \\ B_v &= P_{vv} \Phi W_v; \end{aligned}$$

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 linearity of the transformation. In the literature, studies on temporal aggregation effects on time series models specifications have centered on ARIMA (see Beveridge, 1973; Wei, 1981; and Weiss, 1984), ARMAX (see Beveridge, 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 matrix; 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 \downarrow ow 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 timeseries, $y_t, t = 1; \dots; T$; 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 i ; where $i = [t/m]$ (i.e., if $m = 12$ then i 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_i^{SS} ; is defined as

$$Y_i^{SS} = y_{m \cdot i}; \quad (5)$$

and the averaged sampled time series, Y_i^{AS} as

$$Y_i^{AS} = s(L)y_{m \cdot i}; \quad (6)$$

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

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

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

$$Y_i^{AS} = y_{m \cdot i} + y_{m \cdot (i-1)+1} + \dots + y_{m \cdot (i-1)+m}$$

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:

$$Y_{t=1} = s(L)y_t = y_t + y_{t-1} + \dots + y_{t-m+1}; \quad (8)$$

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

$$Y_i^{AS} = Y_{m \cdot i}; \quad (9)$$

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, $Y_{i,t}^{ss}$: In this case, systematic sampling is applied to both equations (1)-(2) obtaining

$$\begin{aligned} Y_{1;t}^{ss} &= -\alpha Y_{2;t}^{ss} + U_{i,t}^{ss}, \\ \Phi Y_{2;t}^{ss} &= V_{i,t}^{ss}; \end{aligned} \quad (10)$$

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

$$\Phi_m = 1 - L^m;$$

where L^m is the lag operator in i -time, such that $L^m Y_{i,t}^i = Y_{i,t-1}^i$ ($i = ss; as; m$ s); Given the equality

$$\Phi_m = s(L) \Phi;$$

the polynomial $s(L)$ must be applied to (10) to obtain Φ_m : Then, the common stochastic trend in the aggregated model is²

$$\Phi_m Y_{2;t}^{ss} = V_{i,t}^{as}; \quad (11)$$

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 $Y_{i,t}^{ss}$ follows a CI(1,1) process represented by

$$\begin{aligned} Y_{1;t}^{ss} &= -\alpha Y_{2;t}^{ss} + U_{i,t}^{ss}, \\ \Phi_m Y_{2;t}^{ss} &= V_{i,t}^{as}; \end{aligned}$$

where

$$\Phi_m^{ss} = (U_{i,t}^{ss}, V_{i,t}^{as})' = S_{ss}(L) \Phi_m;$$

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

$$S_{ss}(L) = \begin{pmatrix} I_r & 0 \\ 0 & s(L) \Phi_g \end{pmatrix};$$

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

²Note that the permanent component of the triangular model for systematically sampled time series, $V_{i,t}^{as}$ 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, $Y_{i,t}^{as}$; we will follow the transformations observed by Weiss. So the polynomial $s(L)$; is applied to equations (1) and (2):

$$\begin{aligned} s(L)y_{1;t} &= s(L)^{-1}y_{2;t} + s(L)u_t; \\ s(L)\Phi y_{2;t} &= s(L)v_t; \end{aligned}$$

such that

$$Y_{1;t} = -Q_{2;t} + U_t; \quad (12)$$

$$\Phi Y_{2;t} = V_t; \quad (13)$$

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

$$\begin{aligned} Y_{1;i}^{as} &= -Q_{2;i}^{as} + U_i^{as}; \\ \Phi Y_{2;i}^{as} &= V_i^{as}; \end{aligned} \quad (14)$$

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

$$\Phi_m Y_{2;i}^{as} = V_i^{as\alpha},$$

where

$$V_i^{as\alpha} = (s(L))^2 v_{m,i};$$

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_i^{as} follows a CI (1,1) process represented by

$$\begin{aligned} Y_{1;i}^{as} &= -Q_{2;i}^{as} + U_i^{as}, \\ \Phi_m Y_{2;i}^{as} &= V_i^{as\alpha}, \end{aligned}$$

where

$$Y_i^{as\alpha} = (U_i^{as\alpha}, V_i^{as\alpha})' = S_{as}(L) Y_{m,i}^{as};$$

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

$$S_{as}(L) = \begin{pmatrix} s(L)\Phi_r & 0 \\ 0 & s^2(L)\Phi_g \end{pmatrix};$$

and

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

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 $Y_{\dot{t}}^{ms}$; 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 $y_{2;t}^1$ and $y_{2;t}^2$ of dimension g_1 and g_2 , respectively, such that $y_{2;t}^1$ is averaged sampled and $y_{2;t}^2$ 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)

$$y_{1;t} = s(L)^{i-1} \hat{0} s(L) y_{2;t} + u_t$$

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

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

such that

$$y_{1;t} = m^{i-1} \hat{0} y_{2;t} + u_t \quad (15)$$

and

$$\Phi Y_{2;t} = V_t \quad (16)$$

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

$$\begin{aligned} Y_{1;\dot{t}}^{ss} &= m^{i-1} \hat{0} Y_{2;\dot{t}}^{as} + U_{\dot{t}}^{ss}, \\ \Phi Y_{2;\dot{t}}^{as} &= V_{\dot{t}}^{as}, \end{aligned} \quad (17)$$

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

$$\Phi_m Y_{2;\dot{t}}^{as} = V_{\dot{t}}^{as\alpha}.$$

Denote $S_{\dot{t}}^{ms1} = (U_{\dot{t}}^{ss0}, V_{\dot{t}}^{as\alpha0})'$, then $S_{\dot{t}}^{ms1} = S_{ms1}(L) S_{\dot{t}}^{\#}$; with

$$S_{ms1}(L) = \begin{pmatrix} I_r & 0 \\ 0 & s^2(L) \Phi_g \end{pmatrix} :$$

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)

$$s(L)y_{1;t} = s(L)^{-1}y_{2;t} + s(L)u_t \quad (18)$$

and then systematic sampling to (18) obtaining

$$Y_{1;\zeta}^{as} = m^{-1}Y_{2;\zeta}^{ss} + U_{\zeta}^{as}$$

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

$$\Phi_m Y_{2;\zeta}^{ss} = V_{\zeta}^{as}$$

The process z_{ζ}^{ms} is

$$z_{\zeta}^{ms} = (U_{\zeta}^{as}, V_{\zeta}^{as})^0 = S_{ms}(L)z_{\zeta}^{\#}$$

with

$$S_{ms}(L) = \begin{pmatrix} s(L)\Phi_r & 0 \\ 0 & s(L)\Phi_g \end{pmatrix} \quad \#$$

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

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

$$\begin{aligned} y_{1;t} &= {}^{-1}y_{2;t} + {}^{-2}y_{2;t} + u_t \\ \Phi y_{2;t}^1 &= v_t^1 \\ \Phi y_{2;t}^2 &= v_t^2 \end{aligned}$$

where $\bar{v} = (\bar{v}_1; \bar{v}_2)$; $v_t = (v_t^1, v_t^2)^0$, v_t^1 is $g_1 \in 1$, and v_t^2 is $g_2 \in 1$. Now partition z_t ; then-dimensional standardized Wiener process, and the P_{33} matrix conformably with $y_t = (y_{1;t}; y_{2;t}^1; y_{2;t}^2)^0$:

$$\begin{aligned} z_t &= (u_t; v_t^1; v_t^2)^0 \\ W_3 &= [W_u^0; W_{v^1}^0; W_{v^2}^0] \end{aligned}$$

and

$$P_{33} = \begin{pmatrix} P_{uu} & P_{uv^1} & P_{uv^2} \\ 0 & P_{v^1v^1} & P_{v^1v^2} \\ 0 & 0 & P_{v^2v^2} \end{pmatrix}$$

respectively.

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

$$\begin{aligned} Y_{1;\zeta}^{ss} &= m^{-1}Y_{2;\zeta}^{1as} + Y_{2;\zeta}^{2ss} + U_{\zeta}^{ss}, \\ \Phi_m Y_{2;\zeta}^{1as} &= V_{\zeta}^{1as}, \\ \Phi_m Y_{2;\zeta}^{2ss} &= V_{\zeta}^{2as}, \end{aligned}$$

and

$$z_{i,t}^{ms3} = (U_{i,t}^{ss0}, V_{i,t}^{1\text{ as}0}, V_{i,t}^{2\text{ as}0})' = S_{ms3}(L) z_{i,t}^3$$

with

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

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 $y_{2;t}^1$ are averaged sampled, and $y_{2;t}^2$ is systematically sampled. Then, the aggregated model is

$$\begin{aligned} Y_{1;t}^{\text{as}} &= -\alpha Y_{2;t}^1 + m^{-1} Y_{2;t}^2 + U_{i,t}^{\text{as}}, \\ \Phi_m Y_{2;t}^1 &= V_{i,t}^1, \\ \Phi_m Y_{2;t}^2 &= V_{i,t}^2. \end{aligned}$$

and

$$z_{i,t}^{ms4} = (U_{i,t}^{\text{as}0}, V_{i,t}^{1\text{ as}0}, V_{i,t}^{2\text{ as}0})' = S_{ms4}(L) z_{i,t}^3$$

with

$$S_{ms4}(L) = \begin{bmatrix} s(L) \Phi_r & 0 & 0 \\ 0 & s^2(L) \Phi_{g_1} & 0 \\ 0 & 0 & s(L) \Phi_{g_2} \end{bmatrix}$$

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

In this section we derive the ad. 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 Ahiri and Manning, 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 ad. 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 (AEM) (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 AEMM 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 non-AEMM 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 FIMLS, 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 robust cointegration tests (see Gonzalo and Lee, 1998) like the cointegration test of Phillips and Ouliaris (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:

$$\begin{aligned} y_{i;t} &= \beta_i y_{i;t-1} + u_{it} \\ \Phi y_{i;t} &= v_{it} \end{aligned}$$

where $y_{i;t}$ 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:

$$\begin{aligned} & T^{(m)} \hat{\beta}^{(m)} = \left(\sum_{i=1}^g \sum_{t=1}^m X_{it} X_{it}' \right)^{-1} \sum_{i=1}^g \sum_{t=1}^m X_{it} y_{it} \\ &= \left(\sum_{i=1}^g T^{(m)} X_{it} X_{it}' \right)^{-1} \sum_{i=1}^g T^{(m)} X_{it} y_{it} \\ &= \left(\sum_{i=1}^g \sum_{t=1}^m X_{it} X_{it}' \right)^{-1} \sum_{i=1}^g \sum_{t=1}^m X_{it} y_{it} \end{aligned}$$

4.1 Multivariate invariance principle for temporally aggregated process

Because temporal aggregation is a linear transformation on 3_t if 3_t satisfies an IP, then also $^3_t^i$ ($i = ss; as; ms$) 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 \mathbb{X}_t follows an n -dimensional random walk without drift

$$\mathbb{X}_t = \mathbb{X}_{t-1} + \mathbb{Z}_t$$

where $\mu_0 \neq 0$ and \mathbf{z}_t^3 is an n -dimensional covariance stationary process that satisfies the multivariate IP (3), then

$$\prod_{i=1}^m \sum_{t=1}^{i-1} [\mathbf{X}_t^{\otimes m}] \mathbf{z}_t^3 \mathbf{L} \mathbf{B}_{3i} = \mathbf{P}_{3i3i} \Phi \mathbf{W}_3 \sim \text{BM}(-\mathbf{z}_{3i3i});$$

where $\mathbf{z}_t^3 = \mathbf{S}(L) \mathbf{z}_t^3$, and \mathbf{z}_{3i3i} is the long run covariance matrix of \mathbf{z}_t^3 :

$$\mathbf{z}_{3i3i} = \mathbf{P}_{3i3i} \mathbf{P}_{3i3i}^0;$$

where

$$\mathbf{P}_{3i3i} = m^{i-1} \mathbf{S}(1) \Phi \mathbf{P}_{33};$$

for $i = as; ss; ms$:

Proof. Let \mathbf{z}_t^3 be an n -dimensional covariance stationary process that satisfies the IP

$$\prod_{t=1}^{i-1} \sum_{t=1}^{i-1} [\mathbf{X}_t^{\otimes m}] \mathbf{z}_t^3 \mathbf{L} \mathbf{B}_{3i}; \quad (19)$$

and $\mathbf{z}_t^3 = \mathbf{S}(L) \mathbf{z}_t^3$; ³ a linear transformation of \mathbf{z}_t^3 , then \mathbf{z}_t^3 satisfies the IP

$$\prod_{i=1}^m \sum_{t=1}^{i-1} [\mathbf{X}_t^{\otimes m}] \mathbf{z}_t^3 \mathbf{L} \mathbf{B}_{3i}; \quad (20)$$

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

$$\begin{aligned} \prod_{i=1}^m \sum_{t=1}^{i-1} [\mathbf{X}_t^{\otimes m}] \mathbf{z}_t^3 &= \frac{\mu_T}{m} \prod_{i=1}^m \sum_{t=1}^{i-1} [\mathbf{X}_t^{\otimes m}] \mathbf{S}(L) \mathbf{z}_t^3 \\ &= \prod_{i=1}^m m^{i-1} m^{i-1} \sum_{t=1}^{i-1} [\mathbf{X}_t^{\otimes m}] \mathbf{S}(L) \mathbf{z}_t^3 \\ &= m^{i-1} \mathbf{S}(1) \prod_{t=1}^{i-1} \sum_{t=1}^{i-1} [\mathbf{X}_t^{\otimes m}] \mathbf{z}_t^3 \end{aligned}$$

This expression converges in distribution to

$$\mathbf{P}_{3i3i} \Phi \mathbf{W};$$

where

$$\mathbf{P}_{3i3i} = m^{i-1} \mathbf{S}(1) \mathbf{P}_{33};$$

³The proof is valid for any temporal aggregation scheme

so the long run covariance matrix of $\bar{z}^{(m)}$ is

$$- \Sigma_{\bar{z}} = P_{\bar{z}} P_{\bar{z}}^0;$$

■

Now define

$$\bar{z}_i^{(m)} = \bar{z}_{i-1}^{(m)} + z_i^{(m)}, \quad i = 1, \dots, m;$$

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

1. $T^{(m)} \bar{z}^{(m)} \xrightarrow{D} N(\bar{z}, P_{\bar{z}}^0);$
2. $T^{(m)} \bar{z}^{(m)} \xrightarrow{D} N(\bar{z}, P_{\bar{z}}^0 + \alpha_{\bar{z}});$
3. $T^{(m)} \bar{z}^{(m)} \xrightarrow{D} N(\bar{z}, P_{\bar{z}}^0 + \alpha_{\bar{z}});$

4.2 Systematic sampling

From proposition 3, the $P_{\bar{z}^{SS}}$ matrix and the long run variance matrix of \bar{z}^{SS} are given by

$$P_{\bar{z}^{SS}} = \begin{pmatrix} m^{-1} \Sigma_U \\ m^{-1} \Sigma_V \end{pmatrix};$$

$$- \Sigma_{\bar{z}^{SS}} = \begin{pmatrix} m^{-1} \Sigma_{UU} & -\Sigma_{UV} \\ -\Sigma_{VU} & m^{-1} \Sigma_{VV} \end{pmatrix};$$

respectively. Then we have

$$T^{(m)} \bar{z}^{(m)} \xrightarrow{D} N(\bar{z}, P_{\bar{z}}^0 + \alpha_{\bar{z}});$$

$$T^{(m)} \bar{z}^{(m)} \xrightarrow{D} N(\bar{z}, P_{\bar{z}}^0 + \alpha_{\bar{z}});$$

where $\alpha_{\bar{z}} = \sum_{j=1}^1 \Sigma_{V \text{ as } U^{SS}}(i, j)$ and assume that

$$T^{(m)} \bar{z}^{(m)} \xrightarrow{D} N(\bar{z}, P_{\bar{z}}^0 + \alpha_{\bar{z}});$$

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

$$T^{(m)}(\hat{\beta}_{SS} - \beta) \xrightarrow{D} N(0, A^{-1} P_{\bar{z}}^0 A^{-1} + m^{-1} P_V + m^{-1} \Sigma_{V \text{ as } U^{SS}});$$

where $\Phi_{V asUss} = \sum_{j=0}^{m-1} \Phi_{V asUss}(j) + \alpha_{V asUss}$. The nuisance parameter $\Phi_{V asUss}$ does not depend on temporal aggregation order because

$$\Phi_{V asUss}(j) = \sum_{i=jm}^{(j+1)m-1} \Phi_{VU}(i-i)$$

so

$$\Phi_{V asUss} = \sum_{j=0}^{m-1} \Phi_{VU}(j) = \Phi_{VU} \quad (21)$$

and this implies that the a.d. of $T^{(m)}(\hat{\beta}_{SS}^{(m)} | \beta)$ is

$$T^{(m)} @ \hat{\beta}_{SS}^{(m)} | \beta \sim N \left(\beta, \frac{1}{m} \left[P_W^{-1} + \sum_{j=0}^{m-1} W_j^{-1} \right]^{-1} \right) \quad (22)$$

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 $\hat{\beta}_{SS}^{(m)}$; we must use the distribution of $T(\hat{\beta}_{SS} | \beta)$ 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 $T(\hat{\beta}_{SS} | \beta)$ is given by

$$T @ \hat{\beta}_{SS} | \beta \sim N \left(\beta, P_W^{-1} + \sum_{j=0}^{m-1} W_j^{-1} \right) = \Phi_{V asUss} \quad (23)$$

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 -dimensional 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_t^{SS} .

4.3 Average sampling

The P_{3as3as} matrix is given by

$$P_{3as3as} = \begin{pmatrix} m^{-1} & \Phi_U \\ m^{-3/2} & \Phi_V \end{pmatrix} ;$$

and the long run variance matrix is

$$\Sigma_{3as3as} = \begin{pmatrix} m^{-1} & \Sigma_{UV} \\ m^{-2} & \Sigma_{VV} \end{pmatrix} ;$$

From the asymptotic results 1-3 of subsection 4.1

$$T^{(m)i-1} \sum_{\zeta=1}^{T^{(m)}} V_{\zeta}^{as} U_{\zeta}^{as} \rightarrow i_{V^{as}U^{as}}(0);$$

$$T^{(m)i-2} \sum_{\zeta=1}^{T^{(m)}} Y_{2;\zeta}^{as} Y_{2;\zeta}^{as} \rightarrow m^3 P_V \mu_Z W W^0 P_V^0;$$

$$T^{(m)i-1} \sum_{\zeta=1}^{T^{(m)}} Y_{2;\zeta}^{as} U_{\zeta}^{as} \rightarrow m^2 P_V \mu_Z W W^0 P_U^0 + \alpha_{V^{as}U^{as}};$$

where $\alpha_{V^{as}U^{as}} = \sum_{j=1}^m i_{V^{as}U^{as}}(j)$. Then

$$T^{(m)} \hat{\alpha}_{as}^{(m)} \rightarrow i_{V^{as}U^{as}}^{-1} A \rightarrow P_{VV} \mu_Z W_V W_V^0 P_{VV}^0 \rightarrow i_{V^{as}U^{as}}^{-1} P_V \mu_Z W W^0 P_U^0 + m^i \zeta_{V^{as}U^{as}}; \quad (24)$$

where $\zeta_{V^{as}U^{as}} = i_{V^{as}U^{as}}(0) + \alpha_{V^{as}U^{as}}$; and in terms of the covariances $i_{VU}(j)$ can be written as

$$\begin{aligned} \zeta_{V^{as}U^{as}} &= i_{VU}(m-i-1) + (1+2)i_{VU}(m-i-2) + \dots + (1+2+\dots+m)i_{VU}(0) \\ &+ \dots + (1+2+\dots+m+(m-i-1) + \dots + 2)i_{VU}(i-m+2) \\ &+ (1+2+\dots+m+(m-i-1) + \dots + 1) \sum_{j=0}^{i-1} i_{VU}(i-m+1+j); \end{aligned} \quad (25)$$

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

$$T \hat{\alpha}_{as}^{(m)} \rightarrow i_{V^{as}U^{as}}^{-1} A \rightarrow P_{VV} \mu_Z W_V W_V^0 P_{VV}^0 \rightarrow i_{V^{as}U^{as}}^{-1} P_V \mu_Z W W^0 P_U^0 + m^i \zeta_{V^{as}U^{as}}; \quad (26)$$

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 $m^i \zeta_{V^{as}U^{as}}$ included in the 'simultaneous equation bias' (s.e.b) term:

$$m^i \zeta_{V^{as}U^{as}} \rightarrow P_{VV} \mu_Z W_V W_V^0 P_{VV}^0 \rightarrow i_{V^{as}U^{as}}^{-1} \zeta_{V^{as}U^{as}};$$

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

Proposition 5 Let y_t be an n -dimensional 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

$$\Phi_{V^{as}U^{as}} < m^2 \Phi_{VU};$$

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 $m^{-1} \Phi_{V^{as}U^{as}}$, all covariances between the present permanent component, V_t^{as} ; and all leads of the temporary component, U_{t+j}^{as} for $j = 0; 1; \dots$. This term depends on the covariances $\gamma_{vu}(j-i)$ and $\gamma_{vu}(j)$ for $i = 0; 1; 2; \dots$; and $j = 1; 2; \dots; m-i-1$.⁴ Comparing the expression (25) with the s.e. term of the a.d. for the disaggregated case Φ_{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) $\gamma_{vu}(j-i) = 0$ ($i = 0; 1; 2; \dots$;) and ii) there exists some $j = 1; 2; \dots; m-i-1$ 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) $\gamma_{vu}(j-i) = 0$ ($i = 0; 1; 2; \dots$;) and $\gamma_{vu}(j) = 0$ ($j = 1; 2; \dots; m-i-1$) or ii) $\Phi_{V^{as}U^{as}} = m^2 \Phi_{VU}$:

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 $\Phi_{V^{as}U^{as}}$ is a matrix, a deeper analysis is necessary.

Example 8 Consider the trivariate model

$$\begin{aligned} y_{1;t} &= \beta_1 y_{1;t-1} + \beta_2 y_{2;t-1} + \beta_3 y_{3;t-1} + u_{1;t} \\ \Phi y_{1;t} &= v_{1;t} \\ \Phi y_{3;t} &= v_{2;t} \end{aligned}$$

where $^3_t = (u_{1;t}; v_{1;t}; v_{2;t})$ is a covariance stationary process, the simultaneous equation bias can be expressed as

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

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} v_{1u} \\ v_{2u} \end{pmatrix}$$

where we assume that

$$P_V^{-1} W_V W_V' P_V^{-1} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix};$$

and for averaged sampled time series of order m is

$$\begin{pmatrix} a_{11} \frac{1}{m^2} & a_{12} \frac{1}{m^2} \\ a_{21} \frac{1}{m^2} & a_{22} \frac{1}{m^2} \end{pmatrix} \begin{pmatrix} v_{1as} \\ v_{2as} \end{pmatrix};$$

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

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} v_{1u} \\ v_{2u} \end{pmatrix}$$

with

$$\begin{pmatrix} a_{11} \frac{1}{m^2} & a_{12} \frac{1}{m^2} \\ a_{21} \frac{1}{m^2} & a_{22} \frac{1}{m^2} \end{pmatrix} \begin{pmatrix} v_{1as} \\ v_{2as} \end{pmatrix};$$

and so the following situations can occur (term 1: $\frac{1}{m^2} \sum v_{1as}$; term 2: $\frac{1}{m^2} \sum v_{2as}$):
 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 $\frac{1}{4} \sum v_{1u}(1) + \frac{3}{4} \sum v_{1u}(0) + \frac{1}{4} \sum v_{2u}(j-1)$; and three periods $\frac{1}{9} \sum v_{1u}(2) + \frac{1}{3} \sum v_{1u}(1) + \frac{2}{3} \sum v_{1u}(0) + \frac{8}{9} \sum v_{2u}(j-1) + \frac{1}{9} \sum v_{2u}(j-i)$: 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):

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

$$\zeta y_{i,t} = v_t$$

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

$$\tilde{\Sigma} = \begin{pmatrix} 1 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{pmatrix};$$

so the long run variance matrix $\tilde{\Sigma} = \tilde{\Sigma}(0)$: The OLS estimator of β is given by

$$\hat{\beta}^{(m)} = \frac{\sum_{i=1}^m \sum_{t=1}^m y_{i,t} \zeta_t}{\sum_{i=1}^m \sum_{t=1}^m \zeta_t^2}; \quad (27)$$

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:

$$T^{(m)}(\hat{\beta}^{(m)} | \beta) \xrightarrow{d} \frac{1}{m} \frac{(1 - \frac{1}{4})^{1/2}}{\frac{1}{4}} Q_1 + \frac{1}{m} \frac{Q_3}{\frac{1}{4}} + \frac{m+1}{2m^2} \frac{1}{\frac{1}{4}}; \quad (28)$$

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

$$T(\hat{\beta}^{(m)} | \beta) \xrightarrow{d} \frac{(1 - \frac{1}{4})^{1/2}}{\frac{1}{4}} Q_1 + \frac{Q_3}{\frac{1}{4}} + \frac{m+1}{2m} \frac{1}{\frac{1}{4}}; \quad (29)$$

where $Q_1 = \int W_u dW_u$, $Q_2 = \int W_v dW_v$, $Q_3 = \int W_v^2$, W_u and W_v are two independent scalar standardized Wiener processes. The distribution (28) shows the joint 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:

$$\frac{m+1}{2m} \frac{1}{\frac{1}{4}};$$

In fact, this improvement 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:

$$TE(\hat{\beta}^{(m)} | \beta) = \frac{1}{4} \beta + \frac{m+1}{2m} \frac{1}{4} \beta;$$

⁵For 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 $Q_1 Q_3=Q_2$ 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 $E(Q_1 Q_3=Q_2) = 0$; 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 González and Pitarakis (1998, Lemma 3.1); and for $E(Q_1=Q_2)^2$ see Albadir and Paolucci (1994, Theorem 1).

$$\begin{aligned}
 \text{TE}(\hat{\beta}_i^{(m)}) &= 5.196 + 8.090 \frac{\bar{A}}{m+1} + 45.580 \frac{\mu}{2m} + 67.831 \frac{\mu^2}{2m} \\
 &+ 6.317 + 39.639 \frac{\bar{A}}{m+1} + 61.891 \frac{\mu}{2m} + \frac{\mu^2}{3/4} \\
 &+ 3.173 + 11.126 \frac{\mu}{2m} + \frac{\mu^2}{3/4}
 \end{aligned}$$

Letting $\mu = 0.5$, and $\bar{A} = 0.5$, Figures 4.1 and 4.2 represent $\text{TE}(\hat{\beta}_i^{(m)})$ and $\text{TVar}(\hat{\beta}_i^{(m)})$, 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 $m = 2$.

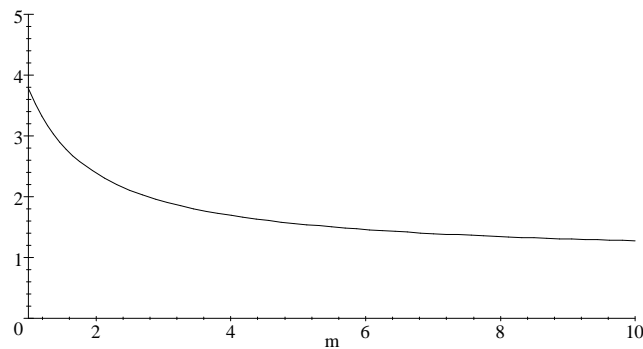


Figure 4.1 $\text{TE}(\hat{\beta}_i^{(m)})$

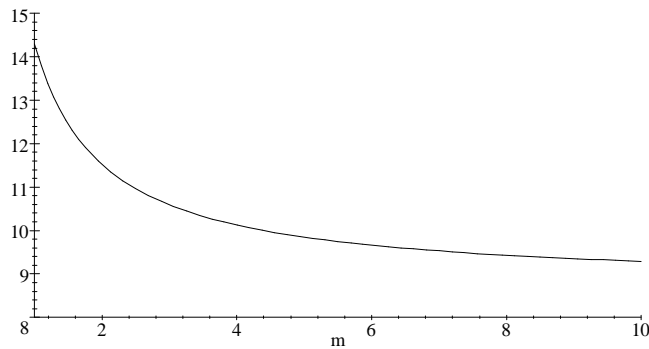


Figure 4.2 $\text{TVar}(\hat{\beta}_i^{(m)})$

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 $P_{3m \times 3m \times s1}$ is

$$P_{3m \times 3m \times s1} = \begin{pmatrix} m^{i-2} \Phi_U \\ m^{3-2} \Phi_V \end{pmatrix};$$

and the long run variance of $\hat{\beta}^{3m \times s1}$ is

$$\hat{\beta}^{3m \times s1} = \begin{pmatrix} m^{i-1} - u_u & m - u_w \\ m - v_u & m^{3-} - v_v \end{pmatrix};$$

This implies that the following quantities converge in distribution to

$$\begin{aligned} T^{(m) i 2} X^{(m)} &= Y_{2;\zeta}^{as} Y_{2;\zeta}^{as0} !^L m^3 P_V W W^0 P_V^0; \\ T^{(m) i 1} X^{(m)} &= Y_{2;\zeta i 1}^{as} U_{\zeta}^{ss} !^L m P_V W d W^0 P_U^0 + \alpha_V^{as} U_{ss}; \end{aligned}$$

so the ad. for $T^{(m)}(\hat{\beta}^{(m)})$ is

$$T^{(m)} @_{m \times s1} \hat{\beta}^{(m)} - A !^L \cdot P_{VV} W_V W_V^0 P_{VV}^0 \rightarrow i 1 \cdot m^{i 2} P_V W d W^0 P_U^0 + m^{i 3} \zeta_V^{as} U_{ss};$$

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

$$T @_{m \times s1} \hat{\beta}^{(m)} - A !^L \cdot P_{VV} W_V W_V^0 P_{VV}^0 \rightarrow i 1 \cdot P_V W d W^0 P_U^0 + m^{i 1} \zeta_V^{as} U_{ss};$$

where

$$\begin{aligned} \zeta_V^{as} U_{ss} &= i_{vu}(0) + 2 i_{vu}(j 1) + \dots + (m - i 1) i_{vu}(j (m - i 2)) + m \sum_{i=0}^{\infty} i_{vu}(j (m - i 1 + i)); \end{aligned}$$

so the term $m^{i 1} \zeta_V^{as} U_{ss}$ is smaller than ζ_{vu} :

Proposition 11 Let y_t be an n -dimensional 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 $Y_{\zeta}^{m \times s1}$ 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

$P_{3m \times 3m}$ is

$$P_{3m \times 3m} = \begin{pmatrix} \mu & \mu \\ \mu & \mu \end{pmatrix} \begin{pmatrix} P_U & \\ & P_V \end{pmatrix};$$

and the long run variance matrix is

$$- \Sigma_{3m \times 3m} = \begin{pmatrix} m - \omega & m - \omega \\ m - \nu_U & m - \nu_V \end{pmatrix};$$

so

$$T^{(m)} = \begin{pmatrix} Y_{2;\zeta}^{ss} \\ Y_{2;\zeta}^{ss} \\ \vdots \\ Y_{2;\zeta}^{ss} \end{pmatrix} \begin{pmatrix} X^{(m)} \\ X^{(m)} \\ \vdots \\ X^{(m)} \end{pmatrix} = \begin{pmatrix} \mu Z \\ \mu Z \\ \vdots \\ \mu Z \end{pmatrix} \begin{pmatrix} W \\ W \\ \vdots \\ W \end{pmatrix} \begin{pmatrix} P_U^0 \\ P_V^0 \\ \vdots \\ P_V^0 \end{pmatrix}$$

The head for $T^{(m)}(\hat{\alpha}^{(m)})$ is

$$T^{(m)} \hat{\alpha}^{(m)} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} \mu Z \\ \mu Z \\ \vdots \\ \mu Z \end{pmatrix} \begin{pmatrix} W \\ W \\ \vdots \\ W \end{pmatrix} \begin{pmatrix} P_U^0 \\ P_V^0 \\ \vdots \\ P_V^0 \end{pmatrix} + m^{-1} \zeta_{V \text{ as } U \text{ as}};$$

and for $T^{(m)}(\hat{\alpha}^{(m)})$

$$T^{(m)} \hat{\alpha}^{(m)} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} \mu Z \\ \mu Z \\ \vdots \\ \mu Z \end{pmatrix} \begin{pmatrix} W \\ W \\ \vdots \\ W \end{pmatrix} \begin{pmatrix} P_U^0 \\ P_V^0 \\ \vdots \\ P_V^0 \end{pmatrix} + m^{-1} \zeta_{V \text{ as } U \text{ as}};$$

where

$$\zeta_{V \text{ as } U \text{ as}} = i_{\nu_U}(m-1) + 2i_{\nu_U}(m-2) + \dots + (m-1)i_{\nu_U}(1) + m \zeta_{\nu_U};$$

Unlike the preceding case of mixed sampling in this case the nuisance parameter $m^{-1} \zeta_{V \text{ as } U \text{ as}}$ is bigger than ζ_{ν_U} , because the presence of the terms

$$\frac{1}{m} i_{\nu_U}(m-1) + \frac{2}{m} i_{\nu_U}(m-2) + \dots + \frac{(m-1)}{m} i_{\nu_U}(1);$$

Proposition 12 Let y_t be an n-dimensional 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 $Y_{\zeta}^{m \times 2}$ 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 $P_{3m \times 3m \times 3}$ is now expressed partitioned conformably with $y_{1;t}$, $y_{2;t}^1$ and $y_{2;t}^2$

$$P_{3m \times 3m \times 3} = \begin{bmatrix} \Phi_U & 0 & 0 \\ \Phi_{V^1} & \Phi_U & 0 \\ \Phi_{V^2} & 0 & \Phi_U \end{bmatrix};$$

and the long run variance matrix is

$$- \Sigma_{3m \times 3m \times 3} = \begin{bmatrix} \omega & 0 & 0 \\ \omega & \omega^2 & 0 \\ \omega & 0 & \omega \end{bmatrix}$$

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

$$T^{(m)} = \begin{bmatrix} \Phi_U & 0 & 0 \\ \Phi_{V^1} & \Phi_U & 0 \\ \Phi_{V^2} & 0 & \Phi_U \end{bmatrix} \begin{bmatrix} \omega & 0 & 0 \\ \omega & \omega^2 & 0 \\ \omega & 0 & \omega \end{bmatrix}$$

then the ad. of $T^{(m)}$ ($\hat{A}^{(m)}$) is

$$\hat{A}^{(m)} = \begin{bmatrix} \omega & 0 & 0 \\ \omega & \omega^2 & 0 \\ \omega & 0 & \omega \end{bmatrix}$$

By introducing the habitual substitutions we obtain the ad. of $T^{(m)}$ ($\hat{A}^{(m)}$)

$$\hat{A}^{(m)} = \begin{bmatrix} \omega & 0 & 0 \\ \omega & \omega^2 & 0 \\ \omega & 0 & \omega \end{bmatrix}$$

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

$$T @_{m \text{ s3 i}}^{\wedge(m)} - A !^L \cdot P_{VV} \quad W_V W_V^0 P_{VV}^0 \quad P_{VV}^0 \quad P_V \quad W d W^0 P_U^0 + \quad m i^1 \zeta_{V1 \text{ asU ss}} \quad \zeta_{V2 \text{ asU ss}} \quad \# \# :$$

According with the preceding results the coefficients of the cointegrating vector associate with the flow regressors $y_{2;t}^1$ are estimated more precise at temporal aggregated level, and the coefficients associated with the stock regressors $y_{2;t}^2$ 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 $P_{3m \text{ s3 m s4}}$

$$P_{3m \text{ s3 m s4}} = \begin{matrix} 2 & m^{1=2} & \Phi_{U^2} & 3 \\ 6 & m^{3=2} & \Phi_{V1} & 7 \\ & m^{1=2} & \Phi_{V2} & 5 \end{matrix} ;$$

implies the following convergences in distribution

$$T^{(m)i^2} X^{(m)} \begin{matrix} Y_{2;\zeta}^{1 \text{ as}} \\ Y_{2;\zeta}^{2 \text{ ss}} \end{matrix} \# \begin{matrix} Y_{2;\zeta}^{1 \text{ as0}} \\ Y_{2;\zeta}^{2 \text{ ss0}} \end{matrix} !^L \quad m^{1=2} \quad m \quad 0 \quad \# \quad \mu Z \quad W W^0 P_V^0 \quad m \quad 0 \quad \# \\ 0 \quad 1 \quad P_V \quad W W^0 P_V^0 \quad 0 \quad 1 \quad m^{1=2} ;$$

$$T^{(m)i^1} X^{(m)} \begin{matrix} Y_{2;\zeta i^1}^{1 \text{ as}} \\ Y_{2;\zeta i^1}^{2 \text{ ss}} \end{matrix} \# \quad U_{\zeta}^{\text{as}} !^L \quad m^{1=2} \quad m \quad 0 \quad \# \quad \mu Z \quad W d W^0 P_U^0 \quad m^{1=2} + \quad \alpha_{V1 \text{ asU as}} \quad \alpha_{V2 \text{ asU as}} \quad \# \\ 0 \quad 1 \quad P_U \quad m^{1=2} + \quad \alpha_{V1 \text{ asU as}} \quad \alpha_{V2 \text{ asU as}} \quad \# \\ 0 \quad \wedge(m) \quad 1$$

that after some transformations and substitutions the limiting distribution of $T @_{m \text{ s4 i}}^{\wedge(m)} - A$ is

$$T @_{m \text{ s4 i}}^{\wedge(m)} - A !^L \cdot P_{VV} \quad W_V W_V^0 P_{VV}^0 \quad P_{VV}^0 \quad P_V \quad W d W^0 P_U^0 + \quad m i^2 \zeta_{V1 \text{ asU as}} \quad m i^1 \zeta_{V2 \text{ asU as}} \quad \# \# :$$

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

5 Monte Carlo study

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

$$y_{1;t} = \beta y_{2;t} + u_t; \quad t= 1;2; \dots; T$$

$$\zeta y_{2;t} = v_t;$$

where $u_t = \frac{1}{2}u_{t-1} + \varepsilon_t$; $|\rho| < 1$, ε_t is i.i.d. $(0,1)$; v_t is i.i.d. $(0,1)$ and $E(\varepsilon_t v_t) = 0$; so the parameter space is $\rho \in [-1, 1]$; $\theta \in [0, 1]$. The values for these parameters are the following

$$\rho = 0.2; \theta = 0.6; 0.9; \sigma = 0.75; 0.9;$$

such that four different DGPs will be considered: DGP 1 ($\rho = 0, \theta = 0.6$), DGP 2 ($\rho = 0, \theta = 0.9$); DGP 3 ($\rho = 0.75, \theta = 0.6$); and DGP 4 ($\rho = 0.75, \theta = 0.9$). The correlation ρ determines the optimality (or not) when $\rho = 0$ (or $\rho = 1$) of the OLS estimator. The autoregressive parameter θ 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 $u_0 = v_0 = 0$; 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 ad. of OLS is:

$$T(\hat{\beta}_i - \beta_i)' \rightarrow N\left(0, \frac{1}{1-\theta^2} \left(\frac{\sigma^2}{Q_2} + \frac{\sigma^2}{Q_2} \right)\right);$$

and for averaged sampled time series is:

$$T(\hat{\beta}_{as}^{(m)} - \beta_i)' \rightarrow N\left(0, \frac{1}{1-\theta^2} \left(\frac{\sigma^2}{Q_2} + \sum_{j=0}^{m-1} \frac{\sigma^2}{Q_2} \theta^{2j} \right)\right);$$

where

$$A = \frac{1}{1-\theta^2} \left(\frac{\sigma^2}{Q_2} + \frac{\sigma^2}{Q_2} \right);$$

The OLS estimator for DGP 1 and DGP 2 is BLUE 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 DGP 3 and DGP 4 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.M.E.), Root Mean Square Error (RMSE), and the concentration probability calculated by $Pr(|\hat{\beta}_i - \beta_i| < 0.05)$ (Pr.); that is the frequency that the estimation deviates from the population parameter by 0.05.

Tables 1 and 2 show the simulation results for systematic sampling when OLS is an BLUE 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 $\frac{1}{2} = 0.6$; but in the case $\frac{1}{2} = 0.9$; the precision of OLS is almost the same.

In tables 5 and 6 ($\rho = 0$) the results for average sampling and DG P1 and DG P2 (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 $\frac{1}{2}$ 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 (DG P3 and DG P4). 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 0.047 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 $\frac{1}{2}$ (table 8). As it is seen in table 8 the reduction in the bias for $\frac{1}{2} = 0.9$ is quite inferior to the reduction for $\frac{1}{2} = 0.6$. 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 0.037 and to 0.028 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 joint 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 theoretical rate T^{-1} .

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 allows 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 AEM estimators like FM-OLS or 3SLS, or as the basis of some robust cointegration tests like the cointegration test of Phillips and Perron.

7 References

Amisano, 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 Paradosi, 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.

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

Chi, 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 $I(1)$ 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 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 $I(1)$ 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* 66 357-8.

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

Haldrup, 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) Lowpass filtered 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 9639, Department of Economics, University of California, San Diego

- Marcello 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 cointegrated 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 Ouliaris, 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 cointegrating 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 (D & P 1)

m	1	2	3	4	5	6	7	8	9	10
[T ^(m)]	50	25	16	12	10	8	7	6	5	5
B.meas	-0.0006	-0.0007	0.0002	0.0003	0.0003	0.0019	-0.0022	0.0011	0.0019	0.0009
RM SE	0.0056	0.006	0.0077	0.0087	0.0093	0.0118	0.0137	0.0150	0.0209	0.0182
Pr.	0.674	0.664	0.5936	0.5874	0.573	0.5314	0.5222	0.504	0.4526	0.468
[T ^(m)]	100	50	33	25	20	16	14	12	11	10
B.meas	-0.0005	-0.0005	-0.0007	-0.0008	-0.0006	-0.0007	-0.0009	-0.0017	-0.0006	-0.0005
RM SE	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	0.784	0.751	0.7072	0.666	0.652	0.64	0.524
[T ^(m)]	150	75	50	37	30	25	21	18	16	15
B.meas	0.0002	0.0004	0.0002	0.0004	-0.0002	0.0006	0.0002	0.0002	0.0003	0.0005
RM SE	0.0010	0.0012	0.0012	0.0015	0.0015	0.0018	0.0020	0.0024	0.0026	0.0028
Pr.	0.892	0.8864	0.8784	0.856	0.8496	0.8288	0.8112	0.7846	0.7702	0.78

Table 2: Behaviour of OLS estimation of cointegrated vectors with systematic sampling (D & P 2)

m	1	2	3	4	5	6	7	8	9	10
[T ^(m)]	50	25	16	12	10	8	7	6	5	5
B.meas	-0.0029	-0.0026	-0.0026	-0.0027	-0.0036	-0.0022	-0.0024	-0.0027	-0.0034	-0.0047
RM SE	0.0550	0.0558	0.0598	0.067	0.063	0.062	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.meas	-0.0017	-0.0017	-0.0016	-0.0016	-0.0008	-0.0016	-0.0021	-0.0016	-0.0026	-0.0011
RM SE	0.0226	0.0227	0.0233	0.0234	0.0238	0.0256	0.0256	0.0272	0.0273	0.0284
Pr.	0.3644	0.366	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.meas	0.0018	0.0017	0.002	0.0018	0.002	0.0022	0.0022	0.0021	0.002	0.0026
RM SE	0.0123	0.0123	0.0125	0.0128	0.0127	0.0130	0.0137	0.0145	0.0147	0.0141
Pr.	0.4354	0.438	0.437	0.4394	0.437	0.4448	0.435	0.436	0.43	0.438

Table 3: Behaviour of OLS estimation of cointegrated vectors with systematic sampling (D G P 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.meas	0.055	0.0546	0.0535	0.0537	0.0524	0.0492	0.0494	0.0483	0.0502	0.0456
RM SE	0.0558	0.0554	0.0546	0.0549	0.0538	0.0511	0.0515	0.0511	0.0548	0.0497
Pr.	0.504	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.meas	0.0384	0.0379	0.0384	0.0374	0.0354	0.0375	0.036	0.036	0.0351	0.0323
RM SE	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	0.688	0.634	0.696	0.616	0.694	0.618	0.616
[T ^(m)]	150	75	50	37	30	25	21	18	16	15
B.meas	0.0289	0.0287	0.0284	0.0284	0.0284	0.0275	0.0282	0.0275	0.0273	0.0259
RM SE	0.0290	0.0288	0.0285	0.0285	0.0285	0.0276	0.0283	0.0277	0.0275	0.026
Pr.	0.817	0.8114	0.8054	0.8024	0.791	0.781	0.768	0.7442	0.741	0.7384

Table 4: Behaviour of OLS estimation of cointegrated vectors with systematic sampling (D G P 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.meas	0.1724	0.1707	0.173	0.162	0.166	0.161	0.166	0.169	0.168	0.1539
RM SE	0.1892	0.1876	0.1913	0.1875	0.1834	0.1873	0.186	0.1855	0.1922	0.176
Pr.	0.16	0.16	0.162	0.175	0.1728	0.162	0.1756	0.1754	0.166	0.17
[T ^(m)]	100	50	33	25	20	16	14	12	11	10
B.meas	0.1274	0.126	0.1272	0.1254	0.1253	0.1264	0.1242	0.1247	0.1233	0.122
RM SE	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	0.246	0.246	0.2436	0.2406	0.256	0.2592
[T ^(m)]	150	75	50	37	30	25	21	18	16	15
B.meas	0.1051	0.1048	0.1041	0.1046	0.1032	0.1026	0.1035	0.1039	0.1032	0.1018
RM SE	0.1076	0.1073	0.1066	0.1072	0.1057	0.1051	0.106	0.106	0.106	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 (D & P 1)

m	1	2	3	4	5	6	7	8	9	10
[T ^(m)]	50	25	16	12	10	8	7	6	5	5
B.meas	0.002	0.002	0.002	0.0018	0.0019	0.0018	0.0019	0.0021	0.0024	0.002
RM SE	0.0064	0.006	0.0076	0.0077	0.0074	0.0084	0.0082	0.0089	0.01087	0.0091
Pr.	0.672	0.628	0.696	0.632	0.656	0.604	0.607	0.5996	0.5842	0.604
[T ^(m)]	100	50	33	25	20	16	14	12	11	10
B.meas	0.0011	0.0011	0.0011	0.0011	0.0012	0.001	0.001	0.0011	0.0012	0.0013
RM SE	0.0022	0.0022	0.0023	0.0023	0.0023	0.0024	0.0024	0.0025	0.0025	0.0026
Pr.	0.816	0.8124	0.8102	0.8098	0.8106	0.7994	0.804	0.792	0.8016	0.806
[T ^(m)]	150	75	50	37	30	25	21	18	16	15
B.meas	0.0001	0.0001	0.0001	0.0001	0.0001	0	0	0	0	0.0001
RM SE	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 (D & P 2)

m	1	2	3	4	5	6	7	8	9	10
[T ^(m)]	50	25	16	12	10	8	7	6	5	5
B.meas	0.0044	0.0045	0.0052	0.0053	0.0042	0.006	0.0051	0.006	0.0051	0.0055
RM SE	0.0584	0.061	0.063	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	0.216
[T ^(m)]	100	50	33	25	20	16	14	12	11	10
B.meas	0.0014	0.0014	0.0014	0.0012	0.0013	0.0017	0.0014	0.0015	0.0013	0.0012
RM SE	0.0239	0.0247	0.0257	0.026	0.0266	0.0287	0.0285	0.0298	0.0291	0.0296
Pr.	0.3518	0.3504	0.3442	0.3492	0.348	0.336	0.3364	0.3358	0.345	0.3438
[T ^(m)]	150	75	50	37	30	25	21	18	16	15
B.meas	0.0005	0.0005	0.0005	0.0005	0.0005	0.0006	0.0006	0.0005	0.0005	0.0006
RM SE	0.0125	0.0129	0.0131	0.0136	0.0135	0.0137	0.0145	0.0153	0.0155	0.0145
Pr.	0.468	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 (D G P 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.meas	0.056	0.046	0.0415	0.036	0.0331	0.0306	0.0283	0.026	0.0252	0.0233
RM SE	0.056	0.0474	0.0422	0.0376	0.0337	0.0313	0.0290	0.0271	0.026	0.0241
Pr.	0.556	0.582	0.5948	0.636	0.638	0.608	0.674	0.626	0.66	0.604
[T ^(m)]	100	50	33	25	20	16	14	12	11	10
B.meas	0.0387	0.0325	0.029	0.0259	0.023	0.0225	0.0211	0.0198	0.0187	0.0176
RM SE	0.0389	0.0327	0.0291	0.026	0.0240	0.0226	0.0212	0.0200	0.0188	0.0177
Pr.	0.712	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.meas	0.0286	0.0242	0.0215	0.0197	0.018	0.018	0.0158	0.0151	0.0142	0.0135
RM SE	0.0287	0.0243	0.0215	0.0197	0.0180	0.018	0.0158	0.0151	0.0142	0.0135
Pr.	0.8154	0.8452	0.866	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 (D G P 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.meas	0.166	0.1574	0.1546	0.1502	0.1452	0.1424	0.1385	0.1348	0.1353	0.1281
RM SE	0.1787	0.1726	0.1707	0.166	0.1610	0.1596	0.1555	0.1534	0.1572	0.146
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.meas	0.1295	0.1249	0.1221	0.1185	0.1157	0.1155	0.1117	0.1104	0.106	0.1042
RM SE	0.1348	0.1300	0.1270	0.1233	0.1204	0.1205	0.116	0.1153	0.1111	0.1086
Pr.	0.234	0.2432	0.2456	0.2514	0.2552	0.2506	0.2596	0.2578	0.266	0.272
[T ^(m)]	150	75	50	37	30	25	21	18	16	15
B.meas	0.0987	0.0951	0.0924	0.0908	0.0879	0.0859	0.0849	0.084	0.0822	0.0789
RM SE	0.1008	0.0971	0.0943	0.0927	0.0897	0.0876	0.086	0.086	0.0840	0.0805
Pr.	0.3244	0.333	0.3382	0.3426	0.349	0.3556	0.356	0.3594	0.364	0.3716