Some Consequences of Temporal Aggregation in Empirical Analysis

Massimiliano Marcellino
I.G.I.E.R., Bocconi University
Via Salasco 5, 20136 Milan, Italy.
E-mail: massimiliano.marcellino@uni-bocconi.it
and
European University Institute

Abstract

We derive the generating mechanism of a temporally aggregated process when the disaggregated one belongs to the VARIMA class. We then study the effects of temporal aggregation on a set of characteristics of usual interest such as exogeneity, causality, cointegration and common features. An empirical example with Canadian interest rates illustrates the main issues.

Key words: ARMA Process, Aggregated Process, Invariance.
JEL Classification: C32, C43, C5.

1. INTRODUCTION

Temporal aggregation arises when the frequency of data generation is lower than that of data collection, so that not all the realizations of the original stochastic process $x = \{x_t\}_{t=0}^{\infty}$ are observable. The available data can be thought of as realizations of the aggregated process $\mathsf{x} = \{\mathsf{x}_\tau\}_{\tau=0}^{\infty}$, where τ is the new temporal frequency and the elements of x are particular functions of those of x, the exact relationships depending on the temporal aggregation scheme.

If the elements of x are stock variables, the usual aggregated process is $\mathsf{x} = \{\mathsf{x}_\tau\}_{\tau=0}^{\infty} = \{x_{tk}\}_{t=0}^{\infty}$, that is, only the k^{th} elements of the original process are retained, where k is the frequency of aggregation. We refer to this situation as <u>point-in-time</u> <u>sampling</u>. When the elements of x are flow variables, the elements of the aggregated process are often partial sums of the disaggregated ones, namely, $\mathsf{x} = \{\mathsf{x}_\tau\}_{\tau=0}^{\infty} = \{\sum_{i=0}^{k-1} x_{tk+i}\}_{t=0}^{\infty}$. We refer to this case as <u>average</u> <u>sampling</u>.

Most of the theoretical literature has focused on the univariate ARIMA case; see in particular Brewer (1973), Wei (1981), and Weiss (1984). That said, temporal aggregation of multivariate processes seems to be the most interesting situation in economics. Economic models often imply that the variables under analysis follow vector ARIMA processes (e.g. Hansen and Sargent (1980), King, Plosser, and Rebelo (1988), Long and Plosser (1983), Taylor (1980)); and many properties of interest such as exogeneity, causality, and cointegration can be only defined in a multivariate context.

Section 2 develops a method to determine the effects of point-in-time and average sampling from a vector ARIMA process. This method can be easily modified to consider the more general case where the aggregated process is $\mathsf{x} = \{\mathsf{x}_\tau\}_{\tau=0}^{\infty} = \{\omega(L)x_{tk+k-1}\}_{t=0}^{\infty}$, with $\omega(L) = \sum_{i=0}^{k-1} \omega_i L^i$. Marcellino (1996a), deals also with <u>mixed</u> <u>sampling</u> (where different aggregation schemes are applied to the variables under analysis), discusses the ARMAX case, and reports results for continuous time processes.

Section 3 analyzes the effects of temporal aggregation on particular properties of interest, such as different notions of exogeneity, Granger causality, structural invariance, integration, cointegration, common features, responses to impulses, and measures of persistence. Our framework shows that these properties (other than long-run ones) are generally lost, and the <u>exact</u> effects of temporal aggregation can be derived. This suggests that economic propositions that have particular implications in terms of these properties should not be tested with temporally aggregated data. As an alternative, given the original process (or property of interest) and the particular temporal aggregation scheme which has generated the available data, the theoretical temporally aggregated process (or modification in the property) could be derived. Its compatibility with the data would then provide an indirect check of the appropriateness of the original hypothesis.

Section 4 presents an illustrative empirical example, using Canadian long-term and short-term interest rates. The data are analyzed at different temporal frequencies, and the theoretical predictions on the effects of temporal aggregation are checked. Section 5 concludes. The Appendix contains the proofs of the propositions in the text.

2. TEMPORAL AGGREGATION OF A VARIMA PROCESS

Let us assume that the disaggregated n dimensional process, $x = \{x_t\}_{t=0}^{\infty}$, evolves according to the system of difference equations

$$G(L)x_t = S(L)\varepsilon_t, \tag{1}$$

where $G(L) = I - G_1L - G_2L^2 - ... - G_gL^g$, $S(L) = I - S_1L - S_2L^2 - ... - S_sL^s$, the $\{G_i\}$ and $\{S_i\}$ are $n \times n$ matrices of coefficients, and ε_t is a multivariate white noise (WN) error, $\varepsilon_t \sim WN(0, \Sigma)$. We determine the generating mechanism of $\mathsf{x} = \{\mathsf{x}_\tau\}_{\tau=0}^{\infty}$, when x is obtained by point-in-time sampling from x. Having done so, we examine the consequences of that aggregation.

Lütkepohl (1987, chap. 6.5) derives upper bounds for the lag lengths of the AR and MA components in the final form representation for x_τ (see e.g. Zellner and Palm (1974) and Prothero and Wallis (1976)), which are gn and $gn + s$ respectively. Our method has two main advantages relative to Lütkepohl's procedure. First, ours yields a more parsimonious aggregated representation. Second, a proper VARIMA (and not a final form) representation is obtained, which is much more useful to evaluate the effects of temporal aggregation on particular properties of interest.

Let us define

$$\underset{1\times gk}{G^v} = (\underset{n\times n}{G_1}, G_2, ..., G_g, 0, ..., 0),$$

$$\underset{1\times gk-g}{B^v} = (\underset{n\times n}{B_1}, B_2, ..., B_{gk-g}),$$

and

$$\underset{gk-g\times gk}{G^m} = \begin{pmatrix} -I & G_1 & ... & G_g & 0 & 0 & ... & 0 \\ 0 & -I & ... & G_{g-1} & G_g & 0 & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & 0 & 0 & 0 & ... & G_g \end{pmatrix},$$

where the $\{B_i\}$ are the coefficients of the polynomial matrix $B(L)$. We also denote by G^v_{-k} and G^m_{-k} the $1 \times (gk-g)$ vector and $(gk-g) \times (gk-g)$ matrix of matrices which are obtained by deleting the k^{th} columns of G^v and G^m.

Proposition 1. If $|G^m_{-k}| \neq 0$, x is generated by (1), and $\mathsf{x} = \{\mathsf{x}_\tau\}_{\tau=0}^{\infty} = \{x_{tk}\}_{t=0}^{\infty}$, then $C(Z)\mathsf{x}_\tau = H(Z)\epsilon_\tau$, where $\epsilon_\tau \sim WN(0, \Sigma_\epsilon)$. The coefficients of $C(Z)$ are the k^{th} (matrix) columns of $G^v_{-k}(G^m_{-k})^{-1}G^m - G^v$, while those of $H(Z) = (I - H_1 Z - ... - H_h Z^h)$ and Σ_ϵ are the solutions of the nonlinear system

$$\begin{cases} \sum_{i=0}^{h} H_i \Sigma_\epsilon H_i' = \sum_{i=0}^{gk-g+s} N_i \Sigma N_i', \\ -H_j \Sigma_\epsilon + \sum_{i=1}^{h-j} H_{i+j} \Sigma_\epsilon H_i' = -N_{jk}\Sigma + \sum_{i=1}^{gk-g+s-jk} N_{i+jk} \Sigma N_i', \quad j = 1, ..., h, \end{cases} \tag{2}$$

where $N(L) = B(L)S(L)$ and $B^v = G^v_{-k}(G^m_{-k})^{-1}$. Upper bounds for h are reported in Table 1. ■

Proposition 1 provides a complete characterization of the aggregated process. It applies to stationary, integrated, and even explosive processes because it makes no assumptions about the roots of the disaggregated AR component. In the univariate case, the results by Brewer (1973), Wei (1981), and Weiss (1984) are obtained. The condition $|G^m_{-k}| \neq 0$ can be substituted for the weaker requirement that the rows of G^v_{-k} lie in the space spanned by the rows of G^m_{-k}. That condition is always satisfied, but it complicates the procedure; see Marcellino (1996a) for an example. The Kalman filter can be exploited for the calculation of the MA coefficients; see Hamilton (1994, p. 391). Our procedure is similar to that suggested by Ericsson, Hendry, and Tran (1994) to determine the relationship between the generating mechanism of seasonally adjusted and unadjusted variables. In particular, they also premultiply the generating mechanism of the original variables by a proper polynomial matrix, whose aim in their case is to annihilate the seasonal component.

We now turn to average sampling. Brewer, Wei, and Weiss, among others, repeat their analysis for this case, which they consider to be rather different from point-in-time sampling and to yield different results. By contrast, in our framework, average sampling is point-in-time sampling from a slightly different disaggregated process. As Campos, Ericsson, and Hendry (1990) note, average sampling can be thought of as a two-step procedure. In the first step, the process $x^* = \{x^*_t\}_{t=0}^{\infty} = \{\sum_{i=0}^{k-1} x_{t+i}\}_{t=0}^{\infty}$ is constructed. In the second step, point-in-time sampling is applied to x^* to obtain the desired aggregated process $\mathsf{x} = \{\mathsf{x}_\tau\}_{\tau=0}^{\infty} = \{x^*_{tk}\}_{t=0}^{\infty}$. Hence, the aggregated generating mechanism is obtained by applying Proposition 1 with x^* as the disaggregated process, where

$$G(L)x^*_t = S(L)(1 + L + ... + L^{k-1})\varepsilon_t.$$

The AR component is in general equal to that for x, which implies that the aggregated process will have the <u>same</u> AR component both for point-in-time and for average sampling.

Both the order and the coefficients of the MA component are different, so the aggregated MA components generally are different for the two aggregation schemes. In particular, the order of the aggregated MA component for average sampling is in general higher than that for point-in-time sampling, see Table 1.

A bivariate VAR(1) with 2-observations point-in-time sampling illustrates the substantial modifications in the coefficients induced by temporal aggregation. This example also will be used throughout the next section to highlight the changes in properties of interest. Hendry (1992) and Campos et al. (1990) examine a similar process, extending work by Telser (1967), Tiao (1972) and Campos, Ericsson, and Hendry (1988).

Suppose y and z are generated by the simultaneous equations model

$$\begin{array}{c} y_t = \alpha z_t + \beta z_{t-1} + \eta_{yt} \\ z_t = \gamma y_{t-1} + \eta_{zt} \end{array}, \quad \begin{pmatrix} \eta_{yt} \\ \eta_{zt} \end{pmatrix} \sim i.i.d.N(0, \begin{pmatrix} \sigma_{yy} & \sigma_{yz} \\ \sigma_{yz} & \sigma_{zz} \end{pmatrix}). \tag{3}$$

The first equation might represent the effects of a policy variable on a target variable, with the second equation as a control rule for the policy variable. The corresponding reduced form is a VAR(1)

$$\begin{array}{c} y_t = \alpha\gamma y_{t-1} + \beta z_{t-1} + \varepsilon_{yt} \\ z_t = \gamma y_{t-1} + \varepsilon_{zt} \end{array}, \quad \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix} = \begin{pmatrix} \eta_{yt} + \alpha\eta_{zt} \\ \eta_{zt} \end{pmatrix}. \tag{4}$$

If we indicate with y_τ and z_τ the temporally aggregated variables, Proposition 1 implies that

$$\begin{array}{c} \mathsf{y}_\tau = (\alpha^2\gamma^2 + \beta\gamma)\mathsf{y}_{\tau-1} + \alpha\beta\gamma\mathsf{z}_{\tau-1} + \varepsilon_{y\tau} \\ \mathsf{z}_\tau = \alpha\gamma^2\mathsf{y}_{\tau-1} + \beta\gamma\mathsf{z}_{\tau-1} + \varepsilon_{z\tau} \end{array}, \quad \begin{pmatrix} \varepsilon_{y\tau} \\ \varepsilon_{z\tau} \end{pmatrix} \sim i.i.d.N(0, \begin{pmatrix} \psi_{yy} & \psi_{yz} \\ \psi_{yz} & \psi_{zz} \end{pmatrix}), \tag{5}$$

with $\psi_{yy} = (1+\alpha^2\gamma^2)\sigma_{yy} + (\alpha^2 + (\beta+\alpha^2\gamma)^2)\sigma_{zz} + 2(\alpha+\alpha\gamma(\beta+\alpha^2\gamma))\sigma_{yz}$, $\psi_{zz} = (1+\alpha^2\gamma^2)\sigma_{zz} + \gamma^2\sigma_{yy} + 2\alpha\gamma^2\sigma_{yz}$, and $\psi_{yz} = \alpha\gamma^2\sigma_{yy} + (\alpha + \alpha\gamma(\beta+\alpha^2\gamma))\sigma_{zz} + (1+\alpha^2\gamma^2 + \gamma(\beta+\alpha^2\gamma))\sigma_{yz}$. Temporal aggregation confounds parameters across variables, time periods, and equations.

A given aggregated VARIMA process might result from temporal aggregation of a non-VARIMA process or any of many different VARIMA processes. This problem parallels aliasing in the frequency domain analysis of time series, see, e.g., Koopmans (1974). This issue has important implications for the related topic of temporal disaggregation, i.e., the construction of disaggregated series from aggregated values, see e.g. Marcellino (in press).

3. TEMPORAL AGGREGATION AND PARTICULAR PROPERTIES

Having determined the relationship between the disaggregated and aggregated data, we can now systematically analyze the effects of temporal aggregation on particular properties of interest. Table 2 provides main references to former studies, where available. Applied researchers sometime interpret results concerning these properties as corroborating or falsifying particular economic propositions. Yet, for this interpretation to be robust, these properties must be invariant to temporal aggregation. Otherwise, spurious conclusions can be drawn. We determine whether these properties are invariant or not, and also derive the exact effects of temporal aggregation. See Marcellino (1996a) for numerous analytical examples and illustrations.

3.1. Stationary and Unit Roots

A general result follows immediately for the effects of temporal aggregation on the roots of the AR component.

Proposition 2. If $|G(l)| = 0$ has roots $\{l_j = \lambda_j,\ j = 1, ..., gn\}$, then $|C(z)| = 0$ has roots $\{z_j = \lambda_j^k,\ j = 1, ..., gn\}$, and the $gn(k-1)$ roots of $|B(l)| = 0$ satisfy the equation $\prod_{j=1}^{gn}(\sum_{i=0}^{k-1} \lambda_j^{k-1-i} l^i) = 0$. ■

Thus, in the stationary case, the roots of $|C(z)| = 0$ are smaller in absolute value than those of $|G(l)| = 0$: λ_j^k vs. λ_j. They decrease with the sampling frequency, and their signs can change for even values of k.

Positive unit roots are not affected by temporal aggregation. If x is $I(d)$, then x is still $I(d)$ — a well known result; see e.g. Granger and Siklos (1995). Thus, unit-root tests should be asymptotically invariant to different temporal aggregation schemes and sampling frequency, with the decrease in absolute value of the non-unit roots balancing the decrease in the number of available observations. Pierse and Snell (1995) have shown that the asymptotic local power of a one-sided unit-root test is independent of the frequency of aggregation, as long as the same data span is adopted.

Proposition 2 also provides a rationale for Tiao's (1972) result that an ARIMA(g,d,s) subject to point-in-time sampling becomes an IMA(d,d-1) process as k increases; and the aggregated MA component is the same regardless of the disaggregated AR and MA coefficients.

3.2. Cointegration

If there exist some linear combinations of $I(d)$ variables that are $I(b)$ with $b < d$, the variables are said to be cointegrated of order d, b $[CI(d, b)]$ and the coefficients of the $I(b)$

linear combinations are the cointegration vectors. See, e.g., Engle and Granger (1987), Johansen (1988), and Stock and Watson (1988). The $I(b)$ linear combinations remain $I(b)$ after temporal aggregation, as noted by Granger (1990). But one might wonder whether other linear cointegrating combinations are created, or whether some of the existing ones become collinear. We have

Proposition 3. If x is $CI(1,0)$, both the number and the composition of the cointegration vectors are invariant to temporal aggregation. The aggregated loadings of the error correction terms are equal to the disaggregated ones premultiplied by $B(1)$. ■

Marcellino (1996b) proves this result for the $I(2)$ case and indicates how it can be demonstrated for the $I(d)$ case. In the analytical example above, (y,z) are $CI(1,0)$ for $(\alpha+\beta)\gamma = 1$, with cointegration vector $(\gamma,-1)$ and loadings $(-\beta,1)'$. For the aggregated variables, it is $rank(C(1)) = 1$, the cointegration vector remains $(\gamma,-1)$, and the loadings become $(-\alpha\gamma\beta, \alpha\gamma)'$.

3.3 Negative and Complex Unit Roots

Solutions of $|G(l)| = 0$ may include negative or complex unit roots, as, for example, when the variables follow a stochastic seasonal pattern. The existence of such roots is not invariant to temporal aggregation because they can become positive unit roots, as follows from Proposition 2. Thus, temporal aggregation may alter the frequency at which integration and cointegration occur. When $(\beta-\alpha)\gamma = 1$ in the analytical example, the variables are $C(1,0)$ at frequency π. After point-in-time sampling, they become $C(1,0)$ at frequency 0, i.e., spurious long-run cointegration is created.

Granger and Siklos (1995) make a similar point. However, they state (p. 361) that this non-invariance is related only to point-in-time aggregation and not to average sampling (temporal aggregation in their vocabulary). This appears correct only if there is a total or partial coincidence of the stochastic seasonal component with the weighting scheme, in which case the stochastic seasonal component is simply lost with average sampling.

3.4 Common Cycles

Linear combinations of $I(1)$ variables can be stationary. Other linear combinations can be such that their first differences are innovations, or at least admit an MA representation of finite order, say h: see Vahid and Engle (1993a, 1993b). In the former case, fewer than n shocks drive the short run evolution of the variables. In the latter case, Vahid and Engle prove the existence of non-synchronous common cycles ($NSCC(h)$) among the variables. Assume s such so-called cofeature combinations, group their coefficients into the $n \times s$ matrix of cofeature vectors γ (so that $\gamma'\Delta x_t = \varepsilon_t$ where ε_t is MA(h)), and indicate with γ_k its aggregated counterpart. Then we have:

Proposition 4. $\gamma_k = \gamma$, but $NSCC(h)$ become $NSCC(h_k)$, where h_k can be determined from Table 1 with $g = 1$ and $s = h$. ■

When $(\alpha + \beta)\gamma = 1$ in the analytical example, the cofeature vector is $(1, \beta)'$, which is orthogonal to the loadings of the error correction term; and $\Delta y_t + \beta \Delta z_t$ is white noise. The variable $\Delta \mathsf{y}_\tau + \beta \Delta \mathsf{z}_\tau$ is still white noise after point-in-time sampling, but is MA(1) after average sampling. Temporal aggregation may also create additional cofeature vectors.

3.5 Exogeneity and Causality

Engle, Hendry, and Richard (1983) define weak exogeneity, structural invariance, and super exogeneity. These properties generally are affected by temporal aggregation. Weak exogeneity basically requires that all the information on a set of parameters of interest be contained in a conditional model, so that the marginal model for the conditioning variables can be neglected when inference on those parameters of interest is the goal of the analysis. Structural invariance requires that the parameters of the conditional model be invariant to changes in those of the marginal model. That implies immunity to the Lucas Critique, see Lucas (1976) and, e.g., Ericsson and Irons (1995). Super exogeneity holds when the first two properties are satisfied. From Proposition 1, the parameters of the aggregated conditional model are a complicated function of those of the disaggregated conditional and marginal models, thereby usually destroying the exogeneity and invariance properties.

Re-examining the analytical example, assume that the parameter of interest in (3) is α. The variable z is weakly exogenous for α if α can be recovered without loss of information from the parameters of the conditional model alone:

$$y_t = a z_t + b y_{t-1} + c z_{t-1} + u_{yt}, \tag{6}$$

where $a = \alpha + \frac{\sigma_{yz}}{\sigma_{zz}}$; $b = -\gamma \frac{\sigma_{yz}}{\sigma_{zz}}$; $c = \beta$; $var(u_{yt}) = \sigma_{yy} - \frac{\sigma_{yz}^2}{\sigma_{zz}}$; and a, b, c, and $var(u_{yt})$ are variation free from the parameters of the marginal model for z, γ and σ_{zz}. Hence, when $\sigma_{yz} = 0$, z is weakly exogenous for α. It is also super exogenous if the conditional model is structurally invariant. For point-in-time sampling with $k = 2$, the conditional model becomes:

$$\mathsf{y}_\tau = a_k \mathsf{z}_\tau + b_k \mathsf{y}_{\tau-1} + c_k \mathsf{z}_{\tau-1} + \mathsf{u}_\tau, \tag{7}$$

where $a_k = \xi$, $b_k = \alpha^2\gamma^2 + \beta\gamma - \xi\alpha\gamma^2$, $c_k = \alpha\beta\gamma - \xi\beta\gamma$, $var(\mathsf{u}_\tau) = \psi_{yy} - \xi\psi_{yz}$, and $\xi = \psi_{yz}/\psi_{zz}$. Hence, even if $\sigma_{yz} = 0$, z is not weakly exogenous for α. Moreover, even if the coefficients in (6) were invariant to changes in σ_{zz}, this is not generally true for those in (7). Such a lack of invariance due to temporal aggregation could be erroneously interpreted as evidence in favor of the Lucas Critique.

As Sims (1971) and Wei (1982) show, Granger non-causality is generally not invariant to aggregation. So, strict and strong exogeneity are also not invariant. While Granger non causality is usually lost after temporal aggregation, it may also be spuriously created.

3.6 Other Properties

The approach above can also determine the effects of temporal aggregation on the Wold representation of the process. If x is generated by (1) (with $s = 0$ for simplicity), then $\Delta x_t = T(L)\varepsilon_t$, with $T(L)G(L) = (1-L)I$. If x is stationary, then $T(L) = (1-L)G^{-1}(L)$. If x is $C(1,0)$, the exact relationship between $G(L)$ and $T(L)$ can be determined from Mellander, Vredin and Warne (1992, p. 374). Point-in-time sampling obtains $\Delta \mathsf{x}_\tau = U(Z)H(Z)\epsilon_\tau$ with $U(Z)C(Z) = (1-Z)I$. Several results follow.

First, the impulse response function and variance decomposition for the temporally aggregated process differ from those of the original process. This follows from the change in the contemporaneous correlation structure of the residuals, as noted in Granger (1988) and Granger and Swanson (1992), and from having different coefficients in the Wold representation.

Second, trend-cycle decompositions of a series and measures of persistence of the shocks (as in Beveridge and Nelson (1981), and Pesaran, Pierse, and Lee (1991)) are generally not invariant to temporal aggregation. See Lippi and Reichlin (1991) and Rossana and Seater (1995) for details and examples for the univariate case.

Third, temporal aggregation affects the quality of forecasts. Forecasts of aggregated values can be obtained either directly from an aggregated model or by aggregating disaggregated forecasts. Lütkepohl (1987, chap. 7,8) compares the two possibilities for the multivariate ARMA case, and shows that the latter leads to a smaller mean squared forecast error, a part from a few special cases. See also Campos (1992).

In summary, temporal aggregation usually alters most properties existing at the disaggregated frequency, excepting long-run properties. Thus some care should be taken when interpreting empirical results on these properties, particularly if they are used to corroborate or falsify an economic proposition.

4. AN EMPIRICAL EXAMPLE

Temporal aggregation has clear empirical consequences. Table 3 lists papers that examine its empirical effects.

To illustrate some of the theoretical results derived, we now discuss a simple empirical example on the term structure of interest rates. The disaggregated data are monthly observations on the Canadian 10-year Government bond yield (RL) and 90-day deposit rate (RS). The series are from the $OECD$ database through TSM (codes: $7mh2camn$ and $7mc2camn$).. The sample period is 1961:1-1993:11, with 395 available monthly observations, and the data appear in Figure 1a.

We first analyze these data under the assumption that their generating process can be approximated by a $VAR(g)$, where g is chosen by a recursive $F-$test for the significance in both equations of the g^{th} lag of at least one of the two variables, starting with $g = 24$. Then, we study the effects of different temporal aggregation schemes on exogeneity, Granger-non causality, the presence of common trends, and common cycles, having approximated the aggregated process by a VAR model. The main references for the tests employed are Johansen (1988,1991,1992) for cointegration and weak exogeneity, and Vahid and Engle (1993a, 1993b) and Engle and Kozicki (1993) for cofeatures. Table 4 summarizes the results, asterisks * and ** indicate rejection of the null hypothesis at the 5% and 1% levels. Full details are available upon request. All the calculations were performed in PcGive 8.1 and PcFiml 8.1 (Doornik and Hendry (1994a, 1994b)).

The first column of Table 4 indicates that there is one cointegration relationship, which can be written as $RS - \beta_1 RL + \beta_2$. Moreover, $\beta_1 = 1$; and because RL tends to be higher than RS, $\beta_2 > 0$. This result agrees with previous empirical analyses on other interest rates data, see e.g. Campbel and Shiller (1987), Engle and Granger (1987), and Hall, Anderson and Granger (1992). RL is weakly exogenous for the parameters of the cointegration vector. The non-significance of the lags of ΔRS in the error correction model for ΔRL is rejected, so RL cannot be considered as a strongly exogenous regressor. The presence of common cycles among the two interest rates is rejected, as well as that of non-synchronous common cycles. (Only the result for $NSCC(1)$ is reported in Table 4.)

We construct quarterly data by point-in-time (QP) and average (QA) sampling from the original observations; see Figures 1b and 1c. In both cases, there is still one cointegration vector, and the hypothesis that its coefficients are equal to those for the monthly series is accepted (second and third columns of Table 4). Weak exogeneity of RL still seems valid. However, the lags of ΔRS in the error correction model for ΔRL are insignificant, so RL is not Granger caused by RS, and RL appears strongly exogenous for the long-run parameters. Moreover, some non-synchronous common cycles are now detected.

For half-yearly data, denoted HP and HA, there is still one cointegration vector, and the hypothesis of equality to that for monthly variables cannot be rejected (last two columns of Table 4). See also the plots of the data, which appear in Figures 1d and 1e. However, RL is no longer weakly exogenous for the coefficients of this cointegration vector, so that RL could not be strongly exogenous, even if Granger non-causality held. The number of cofeature vectors does not decrease, even with a change in the order of the obtainable MA linear combination, which agrees with Proposition 4.

In summary, the empirical example matches and illustrates the theoretical results. That said, the empirical results could differ from the theoretical ones in small samples or in mis-specified models. For example, if only 8 lags are used in the analysis of the monthly data, no cointegration is found, giving the appearance that temporal aggregation increased the cointegration rank. Another possible explanation for a mismatch between empirical and theoretical results could be incorrect initial assumptions. For example, the original variables could be generated by a nonlinear model, or the aggregation scheme could be more general than these we consider (see e.g. Jorda and Marcellino (1998)). Even in the models that we have adopted there are some signs of mis-specification, including non normality and heteroskedasticity. However, the simulations by Eithreim (1994) and Gonzalo (1994) indicate that Johansen's tests are quite robust to this type of mis-specification..

5. CONCLUSIONS

This paper identifies and reviews several consequences of temporal aggregation. First, even linear models in the ARMA class become much more complicated after temporal aggregation, with substantial modifications in the order and coefficients of their AR and MA components. Second, most empirical properties are not invariant to aggregation. Third, even for invariant properties (e.g. integration and cointegration) the finite sample power of testing procedures may fall when temporally aggregated data are used, due to the smaller number of available observations. Fourth, temporal aggregation results generally in an higher mean squared forecast error.

The methods in this paper can be used to identify and alleviate some of these problems. We have provided formulae to determine the aggregated generating mechanism and also the effects of temporal aggregation on particular properties, when the disaggregated generating mechanism and aggregation scheme are specified. Unfortunately, the latter condition is seldom verified in practice. Theoretical economic models rarely specify the temporal frequency: a serious drawback, particularly when these models aim at characterizing the temporal evolution of certain variables or phenomena.. Hence, care should be taken when economic theory is used to support the results of empirical analysis with temporally aggregated data, or viceversa.

ACKNOWLEDGMENT

I wish to thank an Associate Editor, a referee, Rob Engle, Giampiero Gallo, Clive Granger, Jim Hamilton, Sren Johansen, Marco Lippi, Grayham Mizon, Mark Salmon, and Pravin Trivedi for helpful comments. I alone am responsible for remaining errors.

APPENDIX

Proof of Proposition 1. We want to determine a polynomial matrix, $B(L)$, of degree $gk-g$ in L, such that the coefficients not associated with a multiple of L^k in $B(L)G(L)$ are equal to zero. These coefficients can be grouped in the (matrix) columns of $B^v G^m_{-k} - G^v_{-k}$, given that the (matrix) coefficient of L^i in the product $B(L)G(L)$ coincides with the i^{th} (matrix) column of $B^v G^m - G^v$. Thus, the elements of $B(L)$ must satisfy the linear system

$$B^v G^m_{-k} - G^v_{-k} = 0.$$

If $|G^m_{-k}| \neq 0$, the former system admits the unique solution $B^v = G^v_{-k}(G^m_{-k})^{-1}$. The coefficients of the aggregated AR component, $C(Z)$, are those associated with a multiple of L^k in $B(L)G(L)$, and therefore they are those in the k^{th} (matrix) columns of $G^v_{-k}(G^m_{-k})^{-1}G^m - G^v$.

Premultiplication of both sides of (1) by $B(L)$ implies that the autocovariance function of the MA component is:

$$\Gamma_k(j) = \begin{cases} \sum_{i=0}^{gk-g+s} N_i \Sigma_\varepsilon N_i', & j = 0 \\ -N_{jk}\Sigma_\varepsilon + \sum_{i=1}^{\alpha-jk} N_{i+jk}\Sigma_\varepsilon N_i', & j \in N : \alpha \geq jk \\ 0 & j \in N : \alpha < jk \end{cases}$$

$$= \begin{cases} \sum_{i=0}^{h} H_i \Sigma_\epsilon H_i', & j = 0 \\ -H_j\Sigma_\epsilon + \sum_{i=1}^{h-j} H_{i+j}\Sigma_\epsilon H_i', & j = 1, ..., h \\ 0 & j \geq h+1, \end{cases}$$

with $\alpha = gk - g + s$. Therefore $H(Z)$ and Σ_ϵ must solve the system in (2), with the identifying invertibility condition $|H(z)| \neq 0$, $z \leq 1$, and $H(0) = I$. Upper bounds for h depend on g, s, and k; and the different possibilities are summarized in Table 1. ■

Proof of Proposition 2. We have $|C(z)| = |C(l^k)| = |B(l)G(l)| = |B(l)||G(l)|$. Thus, considering $C(L^k)$ as a function of L, $|C(l^k)| = 0$ has gkn roots, which are equal to the gn roots of $|G(l)| = 0$ plus the $gn(k-1)$ roots of $|B(l)| = 0$. If we consider $C(L^k)$ as a function of $L^k = Z$, then $|C(z)| = 0$ has only gn roots, and $|C(\lambda_j^k)| = 0$ for $j = 1, ..., gn$ implies that these roots are just λ_j^k, $j = 1, ..., gn$. If we consider again $C(L^k)$ as a function of L, it also follows that $\prod_{j=1}^{gn}(\lambda_j^k - l^k) = 0$. This admits the decomposition

$$\prod_{j=1}^{gn}(\lambda_j^k - l^k) = \prod_{j=1}^{gn}(\sum_{i=0}^{k-1} \lambda_j^{k-1-i} l^i) \prod_{j=1}^{gn}(\lambda_j - l) = 0.$$

Thus, the $gn(k-1)$ roots of $|B(l)| = 0$ satisfy the equation $\prod_{j=1}^{gn}(\sum_{i=0}^{k-1} \lambda_j^{k-1-i} l^i) = 0$. ■

Proof of Proposition 3. Suppose there exist p cointegration vectors for x, with $G(1) = \alpha\beta'$, where α and β are the $n \times p$ matrices of rank p whose columns contain the loadings and the coefficients of the cointegration vectors. If $B(1)$ is invertible, $C(1) = B(1)G(1)$ implies

$$p = r(G(1)) \leq min[r(B^{-1}(1), r(C(1))] = r(C(1)) = p_k,$$

where p_k is the number of cointegration vectors at the aggregate level. Because $p \geq p_k$ also, then $p_k = p$. Hence, it remains to show that $B(1)$ is invertible, i.e., that $|B(1)| \neq 0$. From Proposition 2, the roots of $|B(l)| = 0$ satisfy the equation $\prod_{j=1}^{gn}(\sum_{i=0}^{k-1} \lambda_j^{k-1-i} l^i) = 0$. Hence, if $B(1)$ were not invertible, then $\sum_{i=0}^{k-1} \lambda_j^{k-1-i} = 0$ for some j, which is not possible when $|\lambda_j| > 1$ or $\lambda_j = 1$ for all j. ■

Proof of Proposition 4. The proof follows immediately from Proposition 1, noting that the problem is to aggregate $\gamma' \Delta x_t = \varepsilon_t$, where ε_t is MA(h), which is a VARMA(1,h) with a diagonal AR component. ■

REFERENCES

Bergström, R., and Edin P.A. (1992), "Time Aggregation and the Distributional Shape of Unemployment Duration," Journal of Applied Econometrics, 7, pp. 5-30.

Beveridge, S., and Nelson C.R. (1981), "A New Approach to the Decomposition of Economic Time Series into Permanent and Transient Components with Particular Attention to Measurement of Business Cycle," Journal of Monetary Economics, 7, pp. 151-174.

Brewer, K.R.W. (1973), "Some Consequences of Temporal Aggregation and Systematic Sampling for ARMA and ARMAX models," Journal of Econometrics, 1, pp. 133-154.

Campbell, J.Y., and Shiller, R.J. (1987), "Cointegration and Tests of Present Value Models," Journal of Political Economy, 95, pp. 1062-1088.

Campos, J. (1992), "Confidence Intervals for Linear Combinations of Forecasts from Dynamic Econometric Models," Journal of Policy Modeling, 14, pp. 535-560.

Campos, J., Ericsson, N.R., and Hendry, D.F. (1988), "Comment on Telser," Journal of the American Statistical Association, 83, p. 581.

— (1990), "An Analogue Model of Phase-averaging Procedures," Journal of Econometrics, 43, pp. 275-292.

Christiano, L.J., and Eichenbaum M. (1987), "Temporal Aggregation and Structural Inference in Macroeconomics," Carnegie-Rochester Conference Series on Public Policy, 26, pp. 63-130.

Cunningham, T.J., and Hardouvelis G.A. (1992), "Money and Interest Rates: the Effects of Temporal Aggregation and Data Revision," Journal of Economics and Business, 44, pp. 19-30.

Doornik, J.A., and Hendry, D.F. (1994a), PcGive 8.0: An Interactive Econometric Modelling System, London: International Thomson.

— (1994b), PcFiml 8.0: Interactive Econometric Modelling of Dynamic Systems, London: International Thomson.

Eithreim, O. (1994), "Inference in Small Cointegrated Systems: Some Monte Carlo Results," W.P. 94-31, University of California at San Diego.

Engle, R.F., and Granger, C.W.J. (1987), "Cointegration and Error Correction: Representation, Estimation and Testing," Econometrica, 55, pp. 251-276.

Engle, R.F., Hendry, D.F., and Richard, J.F. (1983), "Exogeneity," Econometrica, 51, pp. 277-304.

Engle, R.F., and Kozicki, S. (1993), "Testing for Common Features," (with comments), Journal of Business and Economic Statistics, 11, pp. 369-395.

Ericsson, N.R., Hendry, D.F., and Tran H.A. (1994), "Cointegration, Seasonality, Encompassing, and the Demand for Money in the United Kingdom," in Hargreaves C.P. (ed.), Nonstationary Time Series Analysis and Cointegration, Oxford: Oxford University Press.

Ericsson, N.R., and Irons J.S. (1995), "The Lucas Critique in Practice: Theory without Measurement," in Hoover K.D. (ed.), Macroeconometrics: Developments, Tensions and Prospects, Boston: Kluwer Academic Publishers.

Ermini, L. (1988), "Temporal Aggregation and Hall's Model of Consumption Behaviour," Applied Economics, 20, pp. 1317-1320.

— (1989), "Some New Evidence on the Timing of Consumption Decisions and on Their Generating Process," Review of Economics and Statistics, 71, pp. 643-650.

Gamber, E.N., and Joutz, F.L. (1993), "An Application of Estimating Structural Vector Autoregression Models with Long Run Restrictions," Journal of Macroeconomics, 15, pp. 723-745.

Gonzalo, J. (1994), "Five Alternative Methods of Estimating Long-Run Equilibrium Relationships," Journal of Econometrics, 60, pp. 203-233.

Goodhart, CA.E., McMahon, P.C., and Ngama, Y.L. (1991), "Testing for Unit Roots with Very High Frequency Spot Exchange Rate Data," Journal of Macroeconomics, 13, pp. 423-438.

Granger, C.W.J. (1988), "Some Recent Developments in a Concept of Causality," Journal of Econometrics, 39, pp. 199-211.

— (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 Swanson, N.R. (1992), "Impulse Response Functions Based on a Causal Approach to Residual Orthogonalization in Vector Autoregressions," W.P. # 92-50, Univeristy of California at San Diego.

Granger, C.W.J., and Siklos, P.R. (1995), "Systematic Sampling, Temporal Aggregation, Seasonal Adjustment, and Cointegration. Theory and Evidence," Journal of Econometrics, 66, pp. 357-369.

Hall, D.A., Anderson, H.M., and Granger, C.W.J. (1992), "A Cointegration Analysis of Treasury Bill Yields," Review of Economics and Statistics, 75, pp. 116-126.

Hamilton, J.D. (1994), Time Series Analysis, Princeton: Princeton University Press.

Hansen, L., and Sargent, T. (1980), "Formulating and Estimating Linear Rational Expectations Models," Journal of Economic Dynamics and Control, 2, pp. 1-46.

Hendry, D.F. (1992), "An Econometric Analysis of TV Advertising Expenditure in the United Kingdom," Journal of Policy Modelling, 14, pp. 281-311.

Hendry, D.F., and Ericsson N.R. (1991), "An Econometric Analysis of UK Money Demand in Monetary Trends in the United States and the United Kingdom by Milton Friedman and Anna J. Schwartz," American Economic Review, 81, pp. 8-38.

Johansen, S. (1988), "Statistical Analysis of Cointegration Vectors," Journal of Economic Dynamics and Control, 12, pp. 231-254.

— (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," Econometrica, 59, pp. 1551-1580.

— (1992), "Cointegration in Partial Systems and the Efficiency of Single-Equation Analysis," Journal of Econometrics, 52, pp. 389-402.

Jorda, O., and Marcellino, M. (1998), "Random-Time Aggregation," mimeo, University of California at Davis and Bocconi University.

King R., Plosser, C.I., and Rebelo, S. (1988), "Production, Growth and Business Cycles, I," Journal of Monetary Economics, 21, pp309-342.

Koopmans, L.H. (1974), The Spectral Analysis of Time Series, New York: Academic Press.

Lippi, M., and Reichlin, L. (1991), "Trend-Cycle Decompositions and Measures of Persistence: Does Time Aggregation Matter?," Economic Journal, pp. 314-323.

Long, J.B., and Plosser, C.I. (1983), "Real Business Cycles," Journal of Political Economy, 91, pp 39-69.

Lucas, R.E. (1976), "Econometric Policy Evaluation: A Critique," Carnegie-Rochester Conference Series on Public Policy, 1, pp. 19-46.

Marcellino, M. (1996a), "Some Consequences of Temporal Aggregation in Empirical Analysis," W.P. 96/39, University of California at San Diego.

— (1996b), "Aggregation of an I(2) Process," (in Italian with English summary), Proceedings of the XXXVIII Meeting of the Italian Statistical Society, Vol. 2, pp 549-556. ^{t}

— (in press) "Temporal Disaggregation, Missing Observations, Outliers and Forecasting: A Unifying Non-Model Based Procedure," in T.B. Fomby and Hill, R.C. (eds), Advances in Econometrics: Messy Data - Missing Observations, Outliers and Mixed-frequency Data, Greenwich: JAI Press.

Mellander, E., Vredin, A., and Warne, A. (1992), "Stochastic Trends and Economic Fluctuations in a Small Open Economy," Journal of Applied Econometrics, 7, pp. 369-394.

Pesaran, M.H., Pierse, R.G., and Lee, K.C., (1991), "Persistence of Shocks and their Sources in a Multivariate Model of UK Output Growth," DAE W.P. # 9115, University of Cambridge.

Pierse, R.G., and Snell, A.J. (1995), "Temporal Aggregation and the Power of Tests for a Unit Root," Journal of Econometrics, 65, pp. 333-345.

Prothero, D.L., and Wallis, K.F. (1976). "Modelling Macroeconomic Time Series," Journal of the Royal Statistical Society, Ser. A, 139, pp. 468-500.

Rossana, R.J., and Seater, J.J. (1992), "Aggregation, Unit Roots and the Time Series Structure of Manifacturing Real Wages," International Economic Review, 33, pp. 159-179.

— (1995), "Temporal Aggregation and Economic Time Series," Journal of Business and Economic Statistics, 13, pp. 441-451.

Rowley, J.C.R., and Trivedi, P.K. (1975), Econometrics of Investment, Wiley: New York.

Sims, C.A. (1971), "Discrete Approximations to Continuous Time Distributed Lags in Econometrics," Econometrica, 39, pp. 545-563.

Stock, J.H., and Watson, M.W. (1988), "Testing for Common Trends," Journal of the American Statistical Association, 83, pp. 1097-1107.

Taylor J. (1980), "Aggregate Dynamics and Staggered Contracts," Journal of Political Economy, 88, pp. 1-24.

Telser, L.G. (1967), "Discrete Samples and Moving Sums in Stationary Stochastic Processes," Journal of the American Statistical Association, 62, pp. 484-499.

Tiao G.C. (1972), "Asymptotic Behaviour of Temporal Aggregates of Time Series," Biometrika, 59, pp. 525-531.

Vahid, F., and Engle, R.F. (1993a), "Common Trends and Common Cycles," Journal of Applied Econometrics, 8, pp. 341-360.

— (1993b), "Non-Synchronous Common Cycles," W.P. # 9355, University of California at San Diego.

Wei, W.W.S. (1981), "Effect of Systematic Sampling on ARIMA Models," Communications in Statistical-Theoretical Mathematics, 10, pp. 2389-2398.

— (1982), "The Effects of Systematic Sampling and Temporal Aggregation on Causality - A Cautionary Note," Journal of the American Statistical Association, 77, pp. 316-319.

Weiss, A.A. (1984), "Systematic Sampling and Temporal Aggregation in Time Series Models," Journal of Econometrics, pp. 271-281.

Zellner, A., and Palm, F.I. (1974), "Time Series Analysis and Simultaneous Equation Models," Journal of Econometrics, 2, pp. 17-54.

Point-in-time Sampling	Average Sampling
$C(Z) = g, \quad H(Z) = g - 1 - q$ for $qk < g - s \leq (q+1)k$ $q = 0, 1, ..., g-1$	$C(Z) = g, \quad H(Z) = g - q$ for $qk < g - s + 1 \leq (q+1)k$ $q = 0, 1, ..., g$
$C(Z) = g, \quad H(Z) = g$ for $g = s$	$C(Z) = g, \quad H(Z) = g$ for $g = s$
$C(Z) = g, \quad H(Z) = g + q$ for $qk \leq s - g < (q+1)k$ $q = 0, 1, ...$	$C(Z) = g, \quad H(Z) = g + 1 + q$ for $qk \leq s - 1 - g < (q+1)k$ $q = 0, 1, ...$

Table 1: *Upper bounds for the order in Z of C(Z) and H(Z)*
Note: q is the lowest value such that the inequalities are satisfied.

Property	**Author/s**	**Invariance**
Unit roots	Pierse and Snell (1995)	Yes
Cointegration	Granger (1990)	Yes
Seasonal unit roots	Granger and Siklos (1995)	No
Exogeneity	Campos *et al.* (1990), Hendry (1992)	No
Causality	Sims (1971), Wei (1982)	No
Impulse response functions	Granger and Swanson (1992)	No
Trend-cycle decompositions	Lippi and Reichlin (1991)	No
Measures of persistence	Rossana and Seater (1995)	No
Forecasting	Lutkepohl (1987)	No

Table 2: *The effects of temporal aggregation on particular properties*

Author/s	**Subject**
Bergstrom and Edin (1992)	Unemployment
Christiano and Eichenbaum (1987)	Inventories
Cunningham and Hardouvelis (1992)	Effects of monetary shocks
Ermini (1988,1989)	Consumption
Gamber and Joutz (1993)	Effects of real shocks
Goodhart *et al.* (1991)	Exchange rates
Hendry (1992)	Demand analysis
Hendry and Ericsson (1991)	Money demand
Rossana e Seater (1992)	Real wages
Rowley e Trivedi (1975)	Investment

Table 3: *Empirical analyses on the effects of temporal aggregation*

Statistic	Monthly data	Quarterly data		Half-yearly data	
		QP	QA	HP	HA
$\lambda - \max$ test for $p = 0$	17.47*	18.59*	22.89*	28.81**	17.97*
$Trace$ test for $p = 0$	21.91*	21.92*	26.18*	33.02**	21.84*
$\lambda - \max / Trace$ test for $p \leq 1$	4.44	3.32	3.29	4.21	3.87
LR test for $\beta_1 = 1, \beta_2 = 1.322$	2.01	0.01	0.12	2.21	0.52
LR test for weak exogeneity	0.97	0.89	2.91	3.97*	4.31*
Wald test for no lags of $\Delta cansi$	30.58**	7.25	4.21	1.77	-
LR test for CC	95.11**	15.81*	23.74**	8.35	-
LR test for NSCC (order)	55.36** (1)	7.79 (1)	8.21 (2)	-	-
Number of lags	12	4	6	4	1

Table 4: *Properties at different temporal frequencies and for different aggregation schemes*

Figure 1. Graph of the variables at different temporal frequencies and for different aggregation schemes.

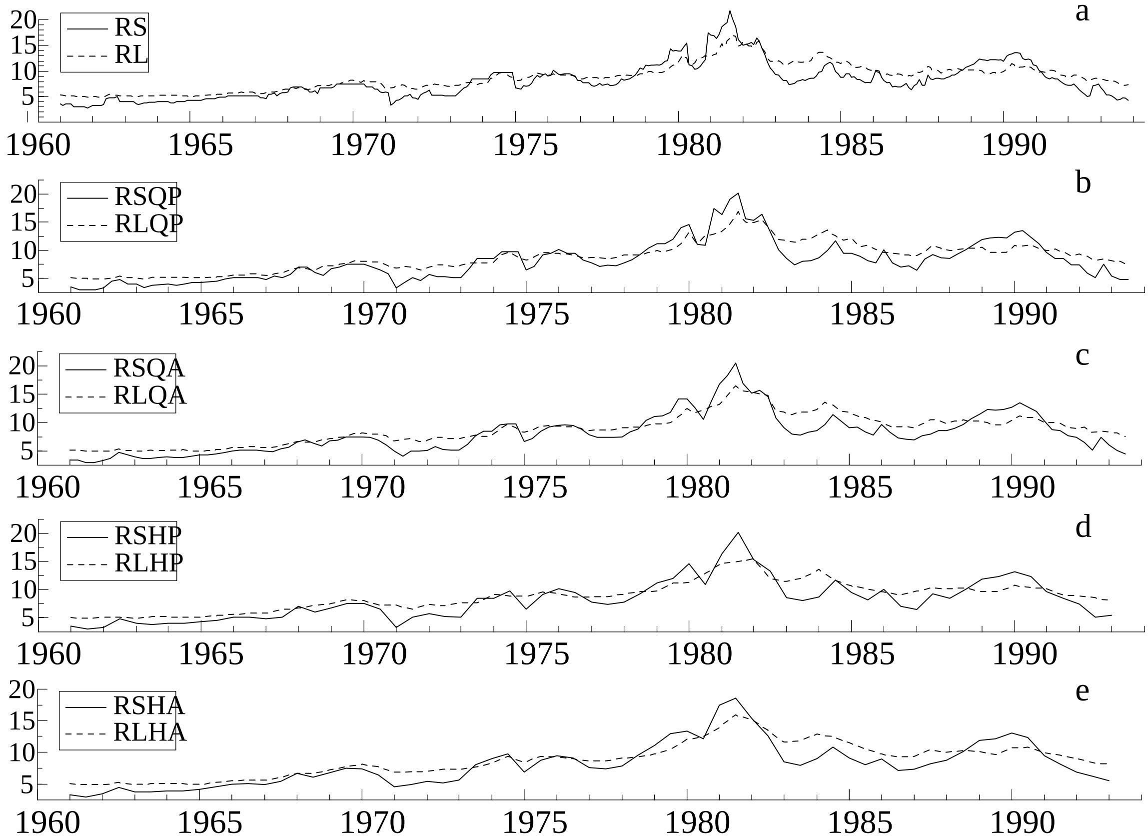
RS
RL
a
RSQP
RLQP
b
RSQA
RLQA
c
RSHP
RLHP
d
RSHA
RLHA
e
20
15
10
5
1960
1965
1970
1975
1980
1985
1990

Figure 1: