

This article was downloaded by: [Silva, Eliud]

On: 6 February 2011

Access details: Access Details: [subscription number 933195018]

Publisher Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of Applied Statistics

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713428038>

Temporal disaggregation and restricted forecasting of multiple population time series

E. Silva^a; V. M. Guerrero^b; D. Peña^a

^a Departamento de Estadística, Universidad Carlos III de Madrid, Madrid, Spain ^b Departamento de Estadística, Instituto Tecnológico Autónomo de México (ITAM), D. F., México

First published on: 06 January 2011

To cite this Article Silva, E. , Guerrero, V. M. and Peña, D.(2011) 'Temporal disaggregation and restricted forecasting of multiple population time series', Journal of Applied Statistics, 38: 4, 799 – 815, First published on: 06 January 2011 (iFirst)

To link to this Article: DOI: 10.1080/02664761003692316

URL: <http://dx.doi.org/10.1080/02664761003692316>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Temporal disaggregation and restricted forecasting of multiple population time series

E. Silva^{a*}, V.M. Guerrero^b and D. Peña^a

^aDepartamento de Estadística, Universidad Carlos III de Madrid, 28903 Getafe, Madrid, Spain;

^bDepartamento de Estadística, Instituto Tecnológico Autónomo de México (ITAM), 01080 D. F., México

(Received 19 April 2009; final version received 27 January 2010)

This article presents some applications of time-series procedures to solve two typical problems that arise when analyzing demographic information in developing countries: (1) unavailability of annual time series of population growth rates (PGRs) and their corresponding population time series and (2) inappropriately defined population growth goals in official population programs. These problems are considered as situations that require combining information of population time series. Firstly, we suggest the use of temporal disaggregation techniques to combine census data with vital statistics information in order to estimate annual PGRs. Secondly, we apply multiple restricted forecasting to combine the official targets on future PGRs with the disaggregated series. Then, we propose a mechanism to evaluate the compatibility of the demographic goals with the annual data. We apply the aforementioned procedures to data of the Mexico City Metropolitan Zone divided by concentric rings and conclude that the targets established in the official program are not feasible. Hence, we derive future PGRs that are both in line with the official targets and with the historical demographic behavior. We conclude that growth population programs should be based on this kind of analysis to be supported empirically. So, through specialized multivariate time-series techniques, we propose to obtain first an optimal estimate of a disaggregate vector of population time series and then, produce restricted forecasts in agreement with some data-based population policies here derived.

Keywords: compatibility testing; demographic forecasting; measurement error; multivariate time series; preliminary series; VAR models

1. Introduction

Unavailability of annual population growth rates (PGRs) represents a problem for policy and decision-makers, particularly in developing countries. This problem occurs in México in spite of the fact that census data are generated regularly every 10 years and that annual vital statistics of births and deaths are also available. Another problem is that inappropriate targets of PGRs

*Corresponding author. Email: jsilvaurrutia@hotmail.com

are usually proposed in the official programs for political reasons. Demographers typically apply easy-to-use, but suboptimal, tools to solve those problems. Besides, there is no unique solution to those problems due to the subjectivity involved in its application. For instance, a demographer would solve the previous problems by interpolating the census data to obtain annual data and then he/she would use personal beliefs to describe the patterns of fertility, mortality, and migration, in order to build scenarios of the future population growth. It should be clear that in such a case, it is not possible to associate a confidence level or credibility to the scenarios. This is in contrast with our proposal, because we suggest solving those problems from a statistical point of view and using optimality criteria. Another point worth emphasizing is that demographers tend to rely on univariate procedures, while our proposal involves multivariate techniques.

Our proposal goes as follows, firstly we use a disaggregation technique to estimate time series of population growth, based primarily on census data and demographic information in the form of vital statistics; secondly, we employ a multiple restricted forecasting technique, with its compatibility testing companion, to analyze the official goals for population growth proposed by the Government. Thus, in order to estimate unavailable annual population data of the Metropolitan Zone of Mexico City (MZMC), we combine decennial census data with annual vital statistics using temporal disaggregation. The combination involves multiple time-series data, since we consider that the MZMC is composed by the Central City and three concentric rings, as shown in Figure 1. The geographic units (delegations and municipalities) that compose these rings are available from the authors on request. On the other hand, to evaluate the feasibility of the official goals for the PGR of each ring, we combine the targets with the annual disaggregated series. Thus, we generate multiple restricted forecasts with a vector auto-regressive (VAR) model and carry out compatibility testing.

In Mexico, demographic data can be obtained from several sources of information, two of the most important are: (1) censuses carried out every 10 years (the most recent in year 2000) by the National Institute of Statistics and Geography (INEGI) [15,16], and (2) annual data on vital

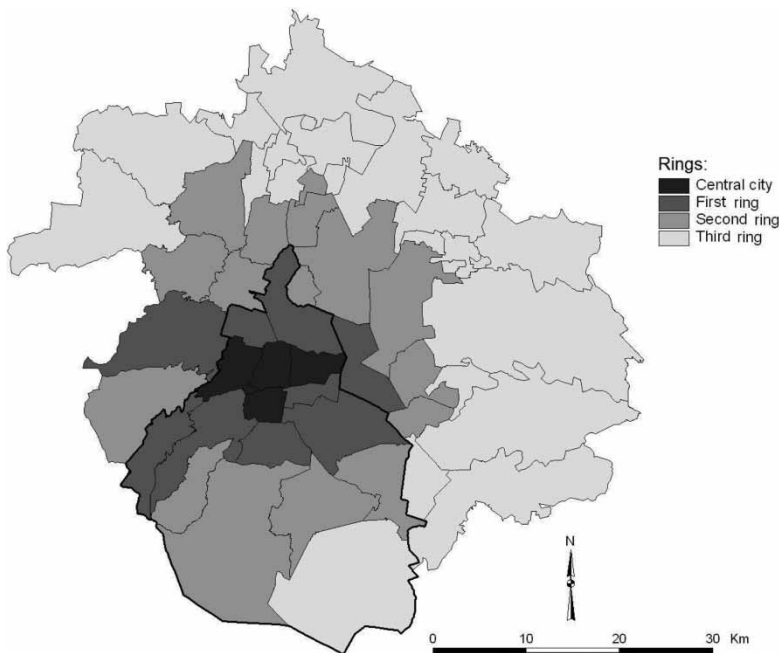


Figure 1. MZMC and its composition in concentric rings.

statistics given by births and deaths from 1940 up to 2000, for the Federal District (DF) and the State of Mexico (SM). These data can be obtained from the Secretariat of Health [27,28] since 1940 up to 1993 and from INEGI [17,18] for years 1994–2000.

We propose to disaggregate low-frequency demographic time-series data on cumulative PGRs, available every decade, with the aid of auxiliary data observed with high frequency (annual vital statistics). Then, the resulting annual estimates will follow the annual pattern provided by the auxiliary data and satisfy the restrictions imposed by the census data. We apply a temporal disaggregation procedure to the census population series for each and every ring, including the Central City, and the resulting estimates will be reasonable in demographic terms, since the population of the rings will add up to the total population for the MZMC. The disaggregation technique that we will use is that proposed by Guerrero and Nieto [10]. Then, we shall employ multiple restricted forecasting, with its corresponding compatibility testing procedure, to evaluate the demographic targets established for the PGRs of each ring. These targets appear in the population program for the DF [7]. The previous time-series techniques have already appeared in the literature, but their typical applications are in economics. We now apply them in demography with success. That is our main contribution, besides providing a new result that serves to take into account that the multiple restricted forecasting technique is applied to a vector of estimated time series, which is then considered to have a measurement error.

Some substantial results obtained in this work are the following. When using the temporal disaggregation technique, we obtained annual series estimates of cumulative PGR that behave as expected, according to the demographic logic. Besides, adequacy of the estimates for all rings was validated empirically by comparing them against data coming from an interdecade population counting. Then, when applying multiple restricted forecasting with the official targets as restrictions, we observed some incompatibility with the demographic dynamics and concluded that the proposed targets are not feasible. As a result, we proposed some other targets that became statistically compatible with the historical behavior (to reach this conclusion we performed a test at the 5% significance level). In particular for the case of Mexico, we did not find any trace of a previous work that focus on the demographic problem we deal with, neither with an approach similar to ours, nor with other approaches, to disaggregate series or to evaluate targets.

The rest of this paper is organized as follows. In Section 2, we present the temporal disaggregation technique to be used and describe the procedure for multiple restricted forecasting, with its companion compatibility testing (for estimated processes). In addition, we show how to incorporate measurement error variability for variables measured with error (in our case, obtained by temporally disaggregating the census data). Section 3 illustrates the application of the aforementioned techniques to the four rings included in the MZMC. First, with temporal disaggregation we obtain estimated annual series of cumulative PGR for each ring and years 1940–2000. The second application provides us with multiple restricted forecasts for the concentric rings and allows us to analyze their respective compatibilities with the official targets. We make some comments about these targets and deduce feasible goals for the future PGR. In Section 4, we conclude with some final comments. The Appendixes show how we: (i) corrected the vital statistics series for outlying observations, (ii) generated the preliminary series required by the disaggregation procedure, and (iii) incorporated the measurement error in the restricted forecasting procedure, for a proper combination of the goals with the annual estimated series.

2. Methodology

2.1 Temporal disaggregation of multiple time series

Several proposals aimed at solving the temporal disaggregation problem of multiple time series are generalizations of univariate disaggregation procedures. The main limit of those methods is

that they assume specific structures for the random error involved: white noise [3,26], random walk [4], or multivariate AR(1) [25]. Therefore, they can be considered as general devices that are usually applied without taking into account the particular features of the data under study. As such, they are general rather than data-specific procedures and their appropriateness cannot be judged empirically. Rather than assuming a specific structure *a priori*, we follow Guerrero and Nieto's suggestion [10] of deducing the structure from the observed data and assume only that a VAR model of order p is appropriate to capture the dynamics of the random error. So, the main advantage of this approach is its objectivity, since it is fully supported and suggested by the data. Moreover, it is derived from theoretical results and produces a statistically optimal estimate of the disaggregated multiple time series. We consider these to be key elements and they should be underlined because the proposed approach will be employed for the first time (as far as we know) to demographic information. Besides, it is important to note that the vital statistics available correspond to the DF and the SM, not to the rings, and this fact precluded the use of the alternative disaggregation procedures.

It should be noticed that instead of disaggregating the multiple population time series, we could have used a different approach such as the Mixed Data Sampling (MIDAS) regressions as in [2,6]. With such an approach, we could merge data with different frequencies of observation (say decennial and annual) into a single regression equation to produce efficient forecasts of the higher frequency series. With this approach we would not require disaggregating the time series and could proceed directly to generate forecasts, but then we would require to extend that idea to the multiple equation case and derive the corresponding restricted forecasting formulas for such a situation. Thus, it would be interesting to employ the MIDAS approach in a future work and analyze its eventual improvement of the multiple unrestricted and restricted forecasts.

Let us first define $\mathbf{Z}_{Dt} = (Z_{it}, \dots, Z_{kt})'$ as the k -dimensional column vector of non-observable variables at time t , for $t = 1, \dots, mn$, where n is the number of complete periods and m is the intraperiod frequency ($m = 10$ years in a decade), while $\mathbf{Z}_D = (\mathbf{Z}'_{D1}, \dots, \mathbf{Z}'_{Dmn})'$ is a stacked vector that contains the vectors \mathbf{Z}_{Dt} . Besides, \mathbf{W}_{Dt} and \mathbf{W}_D are defined as vectors of preliminary data corresponding to \mathbf{Z}_{Dt} and \mathbf{Z}_D , respectively. We want to estimate \mathbf{Z}_D on the basis of \mathbf{W}_D and the identity

$$\mathbf{Y}_D = C_D \mathbf{Z}_D, \tag{1}$$

where \mathbf{Y}_D is a kn -dimensional vector that contains the aggregated data of \mathbf{Z}_D , and C_D is a known $kn \times kmn$ constant matrix. The following result was established in [10].

PROPOSITION 1 *The best linear unbiased estimator of \mathbf{Z}_D , given \mathbf{W}_D and \mathbf{Y}_D is*

$$\hat{\mathbf{Z}}_D = \mathbf{W}_D + A_D (\mathbf{Y}_D - C_D \mathbf{W}_D) \tag{2}$$

with

$$\text{Cov}(\hat{\mathbf{Z}}_D - \mathbf{Z}_D | \mathbf{W}_D) = (I_{knm} - A_D C_D) \Pi^{-1} (P \otimes \Sigma_a) \Pi^{-1'} \tag{3}$$

in which

$$A_D = \Pi^{-1} (P \otimes \Sigma_a) \Pi^{-1'} C'_D [C_D \Pi^{-1} (P \otimes \Sigma_a) \Pi^{-1'} C'_D]^+, \tag{4}$$

where the superscript + denotes Moore–Penrose inverse. The $k m n \times k m n$ matrix Π is built from the matrix coefficients π_1, \dots, π_p of the polynomial involved by the VAR model, as follows

$$\Pi = \begin{bmatrix} I_k & 0 & 0 & 0 & 0 \\ -\pi_1 & I_k & 0 & 0 & 0 \\ & \vdots & & \ddots & \vdots \\ -\pi_p & -\pi_{p-1} & -\pi_{p-2} & 0 & 0 \\ 0 & -\pi_p & -\pi_{p-1} & 0 & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & -\pi_1 & I_k \end{bmatrix}.$$

Moreover, P is an $m n \times m n$ positive definite matrix derived from the data and Σ_a is the error variance–covariance matrix of the VAR model. We refer the reader to the original paper [10] for details on these definitions and the method itself. The operational procedure derived from this proposition consists of two stages. In the first stage, we obtain a preliminary disaggregated series $\{W_D\}$ on the basis of the theory underlying the phenomenon under study and fit a VAR model with deterministic terms to that series. From such a model and the expressions in Proposition 1 (with $P = I$) we obtain a tentatively estimated series $\{\bar{Z}_D\}$ and test for whiteness of the series produced by $\hat{\Pi}(\bar{Z}_D - W_D)$. If this series behaves as white noise we conclude that the tentative series is statistically supported and call it the final disaggregated series. Otherwise, we go to the second stage where we look for a VAR representation of the differences in order to obtain an estimate of the matrix P and derive the final estimate \hat{Z}_D , using again Proposition 1.

2.2 Multiple unrestricted forecasts

Let $Z_t = (Z_{it}, \dots, Z_{kt})'$ be a vector of k variables observed at time t , for $t = 1, \dots, N$. In our case, the multiple time series $\{Z_t\}$ comes from an application of the disaggregation procedure and admits the VAR(p) representation

$$\Pi(B)Z_t = \Lambda D_t + a_t, \tag{5}$$

where $\Pi(B)$ is a matrix polynomial of order $p < \infty$ in the backshift operator B such that $BX_t = X_{t-1}$ for every variable X and subindex t . D_t is a vector containing the deterministic variables (usually a constant and a linear trend), Λ is a matrix of coefficients that capture the deterministic effects, and $\{a_t\}$ is a k -dimensional Gaussian zero-mean white noise process with positive definite covariance matrix $E(a_t a_t') = \Sigma_a$.

Further, let $Z = (Z'_1, \dots, Z'_N)'$ be the vector of known data and let $Z_F = (Z'_{N+1}, \dots, Z'_{N+H})'$ be the vector of future values, with $H \geq 1$ the forecast horizon. The optimal linear forecast of Z_{N+h} , in minimum mean square (MSE) error sense, is given for $h = 1, \dots, H$, by

$$E(Z_{N+h}|Z) = \Lambda D_t + \Pi_1 E(Z_{N+h-1}|Z) + \dots + \Pi_p E(Z_{N+h-p}|Z) \tag{6}$$

with $E(Z_{N+h}|Z) = Z_{N+h}$ for $h \leq 0$. The corresponding forecast errors are given by

$$Z_F - E(Z_F|Z) = \Psi a_F, \tag{7}$$

where $a_F = (a'_{N+1}, \dots, a'_{N+H})' \sim N(0, I_H \otimes \Sigma_a)$ and Ψ is the $kH \times kH$ lower triangular matrix with the identity I_k in its main diagonal, Ψ_1 in its first subdiagonal, Ψ_2 in its second subdiagonal,

and so on. Where the Ψ matrices are obtained recursively from the following expressions:

$$\Psi_0 = I_k, \quad \Psi_j = \Pi_j + \Pi_{j-1}\Psi_1 + \Pi_{j-2}\Psi_2 + \dots + \Pi_1\Psi_{j-1} \text{ for } j = 1, \dots, H - 1, \quad (8)$$

with $\Pi_j = 0$ if $j > p$ or $j < 0$ [29]. Thus, the multiple unrestricted forecasts are conditionally unbiased and their MSE matrix is given by

$$\text{MSE} = \Psi(I_H \otimes \Sigma_a)\Psi'. \quad (9)$$

2.3 Multiple restricted forecasts

We now consider that some additional information is available in the form of a vector $\mathbf{Y} = (Y_1, \dots, Y_M)'$ that imposes $M \geq 0$ independent linear restrictions on the future values of the vector \mathbf{Z} . These restrictions come from an external source to the time-series model and are related to \mathbf{Z}_F by means of

$$\mathbf{Y} = \mathbf{C}\mathbf{Z}_F + \mathbf{u}, \quad (10)$$

where $\mathbf{u} \sim N(\mathbf{0}_M, \Sigma_u)$. In our case, the restrictions are targets on the population rate of growth and in order to test for their compatibility with the unrestricted forecasts, we assume they are certain, so that $\Sigma_u = 0$. Besides, \mathbf{C} is an $M \times kH$ matrix of rank M given by $\mathbf{C} = [\mathbf{c}_1 \dots \mathbf{c}_M]'$ where $\mathbf{c}_m = (c_{m,1}, \dots, c_{m,kH})$ for $m = 1, \dots, M$.

Using Equations (7) and (10) Pankratz [24], showed that the optimal restricted forecast of \mathbf{Z}_F is

$$\mathbf{Z}_{F,H}^R = E(\mathbf{Z}_F|\mathbf{Z}) + \mathbf{A}[\mathbf{Y} - \mathbf{C}E(\mathbf{Z}_F|\mathbf{Z})] \quad (11)$$

with

$$\text{MSE}(\mathbf{Z}_{F,H}^R) = (I_H - \mathbf{A}\mathbf{C})\Sigma_{E(\mathbf{Z}_F|\mathbf{Z})} \quad \text{and} \quad \mathbf{A} = \Sigma_{E(\mathbf{Z}_F|\mathbf{Z})}\mathbf{C}'\Omega^{-1}, \quad (12)$$

where

$$\Sigma_{E(\mathbf{Z}_F|\mathbf{Z})} = \Psi(I_H \otimes \Sigma_a)\Psi' \quad \text{and} \quad \Omega = \mathbf{C}\Psi(I_H \otimes \Sigma_a)\Psi'\mathbf{C}'. \quad (13)$$

Expressions (11)–(13) can be obtained also by applying Theorem 1 of Nieto and Guerrero [23] without the normality assumption required by Pankratz's result.

2.4 Compatibility testing

Combining information should be judged from an empirical point of view, because the restrictions imposed to the series by the population goals may contradict the observed behavior of the series. To this end, we use the following statistic proposed by Guerrero and Peña [11,12]

$$\mathbf{K} = \mathbf{d}'\Omega^{-1}\mathbf{d} \sim \chi_M^2, \quad (14)$$

where $\mathbf{d} = \mathbf{Y} - \mathbf{C}E(\mathbf{Z}_F|\mathbf{Z})$. Then, $\mathbf{Y} - \mathbf{C}E(\mathbf{Z}_F|\mathbf{Z})$ lies in the compatibility region if the calculated statistic K_{calc} is not greater than $\chi_M^2(\alpha)$, the $(1 - \alpha)$ th quantile of a Chi-square distribution with M degrees of freedom, and we declare \mathbf{Y} incompatible with $\mathbf{C}E(\mathbf{Z}_F|\mathbf{Z})$ at the $100\alpha\%$ significance level if $K_{\text{calc}} > \chi_M^2(\alpha)$. We can also use partial compatibility test statistics, denoted as K_{par} , to evaluate the compatibility of specific restrictions with unrestricted forecasts.

2.5 VAR forecasting and compatibility testing with estimated processes

In a VAR model with estimated parameters the forecasts are conditionally unbiased and asymptotically valid [5]. Also, it can be shown that the vector of optimum restricted forecasts with an estimated process is given by

$$\hat{\mathbf{Z}}_{F,H}^R = E(\hat{\mathbf{Z}}_F|\mathbf{Z}) + \hat{A}[\mathbf{Y} - CE(\hat{\mathbf{Z}}_F|\mathbf{Z})], \tag{15}$$

where

$$\hat{A} = \hat{\Sigma}_{E(\hat{\mathbf{Z}}_F|\mathbf{Z})} C' \hat{\Omega}^{-1}, \quad \hat{\Omega} = C \hat{\Sigma}_{E(\hat{\mathbf{Z}}_F|\mathbf{Z})} C' \tag{16}$$

and

$$\hat{\Sigma}_{E(\hat{\mathbf{Z}}_F|\mathbf{Z})} \approx \hat{\Sigma}_{E(\mathbf{Z}_F|\mathbf{Z})} + N^{-1} \hat{\Sigma}_a. \tag{17}$$

Moreover, its estimated MSE matrix is given by

$$\widehat{\text{MSE}}(\hat{\mathbf{Z}}_{F,H}^R) = (I - \hat{A}C) \hat{\Sigma}_{E(\hat{\mathbf{Z}}_F|\mathbf{Z})}. \tag{18}$$

Compatibility testing should also be modified for estimated processes. Gomez and Guerrero [8], showed that the appropriate test statistic is given by

$$\bar{K} = \frac{\hat{\mathbf{d}}' \hat{\Omega}^{-1} \hat{\mathbf{d}}}{M} \sim F_{M,T-Mp-1}, \tag{19}$$

where $\hat{\mathbf{d}} = \mathbf{Y} - CE\hat{\mathbf{Z}}_F|\mathbf{Z}$. So that, \mathbf{Y} is not in the compatibility region at the α significance level if $\bar{K}_{\text{calc}} \geq F_{M,T-Mp-1}(\alpha)$ with $F_{M,T-Mp-1}(\alpha)$ being the $(1 - \alpha)$ th quantile of the $F_{M,T-Mp-1}$ distribution. This statistic will be used below for examining compatibility between official targets and unrestricted forecasts. Similarly, we will apply partial compatibility test statistics, \bar{K}_{par} , to evaluate the compatibility between specific restrictions and unrestricted forecasts.

2.6 Incorporating measurement error variability

From now on, we denote Central City population by ccp_t , first ring population by frp_t , second ring population by srp_t , third ring population by trp_t , MZMC population by mzmc_t , DF population by dfp_t , and SM population by smp_t . In this application, it is very important to note that the multiple VAR forecasts are not obtained from actual observations of the variables of interest, but from estimated data (hence, measured with error) derived as an application of the disaggregation technique. This is an important point that must be emphasized because VAR forecasts are generally produced from observed time series, which is not the case here. In fact, the VAR model is used to forecast an unobserved disaggregated multiple time series which came out from an unbiased estimation procedure. Hence, the estimated series will be assumed to be equal to the true, but unobserved, time series plus an error term, that we call a measurement error. In Appendix 3 we show how to incorporate the measurement error variability into the restricted forecasting formula.

Thus, to take into account these measurement errors into the forecasts we define the 80×80 matrix of estimated measurement error variances

$$\hat{\Sigma}_\varepsilon = I_{20} \otimes \text{diag}(\bar{\sigma}^2(\text{ccp}), \bar{\sigma}^2(\text{frp}), \bar{\sigma}^2(\text{srp}), \bar{\sigma}^2(\text{trp})), \tag{20}$$

where \otimes denotes the Kronecker product and every element in the diagonal matrix is an average of the respective elements $\hat{\sigma}_t^2(\text{ccp})$, $\hat{\sigma}_t^2(\text{frp})$, $\hat{\sigma}_t^2(\text{srp})$, and $\hat{\sigma}_t^2(\text{trp})$ for $t = 1981, 1982, \dots, 2000$. These estimated variances are taken from the diagonal of the covariance matrices produced by the disaggregation procedure. We considered the errors of the last two decades because the forecasts are required for a 20-year horizon, from 2001 to 2020. The matrix $\hat{\Sigma}_e$ was added to Equations (17) and (18) to include the effect of measurement errors. Without it we could get a false idea of the variability associated with the multiple restricted forecasts, the compatibility test would not be strictly valid, and the evaluation of official targets would lead to erroneous conclusions.

It is convenient to mention that Nieto [22] has provided another approach to solve essentially the same problem considered here. His solution produces optimal forecast in the context of the so-called ex ante prediction of unobservable multivariate time series. Thus, it would be interesting to apply his results in a future work that postulates a multivariate structural model.

3. Applications

3.1 Application 1: temporal disaggregation

In this application, temporal disaggregation of the census data is equivalent to interpolate them by annual figures. We require first preliminary series for each concentric ring and to get them it was necessary to correct the annual births series for outlying observations (in Appendix 1 we apply the technique proposed by Gómez *et al.* [9]). Then, we employed the algorithm in Appendix 2 to focus the problem from a demographic, rather than a statistical point of view, see for example [1]. We did that to obtain a better subject matter interpretation of the resulting annual population series.

The computations were performed with the packages E-Views 5.1 (Quantitative Micro Software) and Matlab 7 (MathWorks, Inc.). The data are available from the authors on request. Let zccp_t , zfrp_t , zsrp_t , and ztrp_t be the non-observable variables at time representing the cumulative PGR of the Central City and the rings. The number of complete periods is $n = 6$ (decades) and $m = 10$ is the number of annual observations in a decade. Let $\mathbf{zccp} = (\text{zccp}'_1, \dots, \text{zccp}'_{mn})'$ be a stacked vector of the mn values of zccp . The vectors \mathbf{zfrp} , \mathbf{zsrp} , and \mathbf{ztrp} are defined similarly. Then, we define the vectors of preliminary series \mathbf{wccp} , \mathbf{wfrp} , \mathbf{wsrp} , and \mathbf{wtrp} corresponding to \mathbf{zccp} , \mathbf{zfrp} , \mathbf{zsrp} , and \mathbf{ztrp} .

The temporal restrictions are specified by means of $\mathbf{Y}_D = (I_6 \otimes C_0)\mathbf{Z}_D$, where $C_0 = [\mathbf{0}_9 \ 1]$ with $\mathbf{0}_9$ a 9-dimensional zero vector. The six elements of the vector \mathbf{Y}_D are the cumulative PGR for the rings, coming from the census data, i.e. for the years ending in 0 from 1950 up to 2000. No contemporaneous restrictions are used in this case, since they are considered implicitly by the temporal restrictions. Therefore, the multivariate application of this technique became a univariate application, and we applied the disaggregation procedure to each univariate time series (for each ring) separately.

In the first stage, we built an autoregressive model to represent the behavior of the preliminary series for each ring. In the second stage, we used another autoregressive model for the differences between the tentatively estimated series and the preliminary series. We present the estimation results in Table 1.

Standard errors for the disaggregated series were obtained as square roots of the elements in the diagonal of the estimated covariance matrix (3). Then, we obtained prediction intervals (PI) from these estimates. These PI look as 'bubbles' in Figure 2, because there is no uncertainty associated with the observed values for the census years.

The polynomials of order 4 in the models for the second and third rings look strange, but they were required to get a stationary behavior of their stochastic structures, since the method assumes that all kind of non-stationarities in the data can be taken into account by way of deterministic

Table 1. Estimated autoregressive models used for univariate temporal disaggregation (*t*-statistics in parentheses).

Rings	First stage	Second stage
ccp_t	$\widehat{wccp}_t = 1.881 \widehat{wccp}_{t-1} - 0.883 \widehat{wccp}_{t-2},$ $\hat{\Sigma} = 7.42 \times 10^{-5}$	$\hat{\phi}_1 = 3.572, \hat{\phi}_2 = -4.820, \hat{\phi}_3 = 2.913,$ $\hat{\phi}_4 = -0.666,$ $\hat{\Sigma} = 7.41 \times 10^{-8}$
frp_t	$\widehat{wfrp}_t = 0.014t - 0.001t^2 + 0.986 \widehat{wfrp}_{t-1}$ $- 0.110 \widehat{wfrp}_{t-10}, \hat{\Sigma} = 1.16 \times 10^{-4}$	$\hat{\phi}_1 = 1.903, \hat{\phi}_2 = -0.916,$ $\hat{\phi}_{10} = 0.686, \hat{\phi}_{11} = -1.339,$ $\hat{\phi}_{12} = 0.664, \hat{\Sigma} = 1.33 \times 10^{-5}$
srp_t	$\widehat{wsrp}_t = 0.001 t^2 - 2.17 \times 10^{-5} t^3$ $+ 1.43 \times 10^{-7} t^4 + 1.236 \widehat{wsrp}_{t-1}$ $- 0.411 \widehat{wsrp}_{t-2}, \hat{\Sigma} = 3.26 \times 10^{-4}$	$\hat{\phi}_1 = 2.375, \hat{\phi}_2 = -2.047,$ $\hat{\phi}_3 = 0.650, \hat{\phi}_{10} = 0.851,$ $\hat{\phi}_{11} = -2.042, \hat{\phi}_{12} = 1.777,$ $\hat{\phi}_{13} = -0.572, \hat{\Sigma} = 2.52 \times 10^{-5}$
trp_t	$\widehat{wtrp}_t = 2.10 \times 10^{-4} t^2 - 2.39 \times 10^{-8} t^4$ $+ 1.30 \widehat{wtrp}_{t-1} - 0.466 \widehat{wtrp}_{t-2},$ $\hat{\Sigma} = 7.49 \times 10^{-5}$	$\hat{\phi}_1 = 2.514, \hat{\phi}_2 = -2.176,$ $\hat{\phi}_{13} = 0.659, \hat{\Sigma} = 2.35 \times 10^{-5}$

elements. Thus, all models used in the first and second stages have characteristic polynomials with roots outside the unit circle. Moreover, we could not reject the white noise hypotheses for the univariate residuals at the 5% significance level. As it was expected, the annual disaggregated series satisfy the temporal restrictions imposed by the observed census data. Some selected results appear in Table 2.

To validate the previous results empirically, we made use of the data obtained in an interdecade population counting carried out in Mexico in 1995. Table 3 shows the observed population figures obtained in that counting [19,20]. All the corresponding values for the rings fall within the 95% PI for the disaggregated values.

3.2 Application 2: evaluating population goals

To obtain multiple unrestricted forecasts, we first estimated a VAR model for the population series selecting its order by the likelihood ratio testing scheme with upper bound $p = 5$. The deterministic element in each equation of the VAR model was only a constant. The results are

$$H_0^1 : \pi_5 = 0 \text{ vs. } H_1^1 : \pi_5 \neq 0, \quad \chi^2(16) = 6.69,$$

$$H_0^2 : \pi_4 = 0 \text{ vs. } H_1^2 : \pi_4 \neq 0 | \pi_5 = 0, \quad \chi^2(16) = 11.85,$$

$$H_0^3 : \pi_3 = 0 \text{ vs. } H_1^3 : \pi_3 \neq 0 | \pi_5 = \pi_4 = 0, \quad \chi^2(16) = 13.96,$$

$$H_0^4 : \pi_4 = 0 \text{ vs. } H_1^4 : \pi_2 \neq 0 | \pi_5 = \pi_4 = \pi_3 = 0, \quad \chi^2(16) = 64.17,$$

Downloaded By: [Silva, Eliud] At: 13:54 6 February 2011

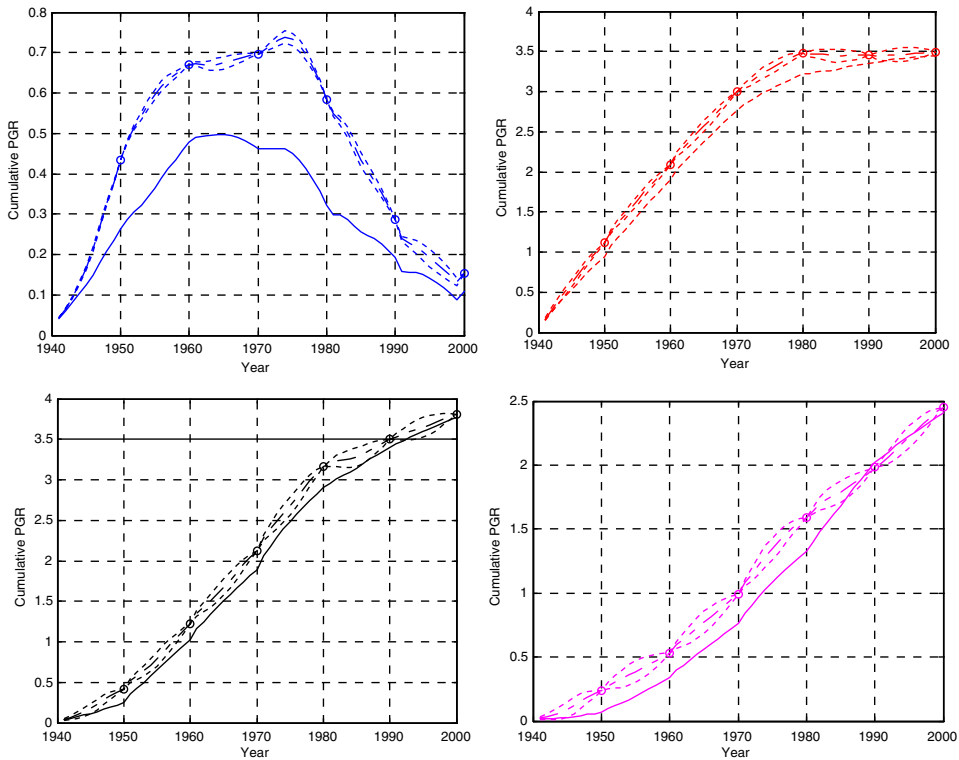


Figure 2. Temporal disaggregation of $ccpt$ (top left), $frpt$ (top right), $srpt$ (bottom left), $trpt$ (bottom right). Solid lines denote preliminary series, dashed lines are disaggregate series with their 95% prediction intervals and dots are census data.

Table 2. Disaggregated series: preliminary and final, with standard errors ($100 \times SE$).

Year	Prelim.	Final	SE	Prelim.	Final	SE	Prelim.	Final	SE	Prelim.	Final	SE
		ccp_t			frp_t			srp_t			trp_t	
1941	0.042	0.043	0.001	0.156	0.163	0.031	0.040	0.042	0.018	0.019	0.022	0.017
1942	0.061	0.065	0.003	0.257	0.276	0.057	0.057	0.065	0.047	0.020	0.031	0.045
1943	0.080	0.092	0.005	0.352	0.388	0.078	0.075	0.094	0.079	0.021	0.045	0.076
1944	0.101	0.126	0.008	0.443	0.499	0.092	0.094	0.130	0.107	0.024	0.065	0.104
1945	0.126	0.168	0.011	0.532	0.610	0.098	0.117	0.176	0.125	0.029	0.092	0.123
1946	0.149	0.213	0.012	0.617	0.718	0.096	0.139	0.225	0.129	0.033	0.120	0.128
1947	0.179	0.269	0.012	0.705	0.828	0.085	0.167	0.282	0.117	0.044	0.155	0.117
1948	0.207	0.324	0.010	0.788	0.931	0.066	0.194	0.335	0.089	0.052	0.186	0.090
1949	0.235	0.379	0.006	0.867	1.027	0.038	0.219	0.381	0.049	0.058	0.212	0.050
1950	0.264	0.434	0.000	0.945	1.115	0.000	0.247	0.417	0.000	0.067	0.236	0.000
...
1991	0.157	0.239	0.012	3.365	3.444	0.070	3.440	3.528	0.080	2.059	2.016	0.055
1992	0.156	0.228	0.023	3.380	3.446	0.115	3.482	3.560	0.148	2.104	2.062	0.104
1993	0.156	0.219	0.033	3.395	3.451	0.145	3.524	3.591	0.200	2.149	2.114	0.142
1994	0.150	0.206	0.040	3.406	3.456	0.162	3.562	3.619	0.231	2.190	2.167	0.167
1995	0.141	0.192	0.043	3.415	3.462	0.168	3.598	3.648	0.242	2.228	2.219	0.175
1996	0.130	0.177	0.043	3.422	3.469	0.164	3.632	3.679	0.233	2.265	2.270	0.168
1997	0.117	0.163	0.039	3.429	3.476	0.147	3.666	3.712	0.203	2.302	2.320	0.145
1998	0.103	0.148	0.030	3.435	3.484	0.117	3.700	3.747	0.151	2.338	2.368	0.107
1999	0.088	0.132	0.017	3.442	3.489	0.072	3.733	3.781	0.082	2.374	2.413	0.058
2000	0.109	0.153	0.000	3.447	3.492	0.000	3.766	3.811	0.000	2.410	2.454	0.000

Downloaded By: [Silva, Eliud] At: 13:54 6 February 2011

Table 3. Observed and disaggregated cumulative PGR values for 1995.

Rings	Observed (interdecade counting)	Lower 95% limit	Estimated by disaggregation	Upper 95% limit
ccp _t	0.195	0.171	0.192	0.213
frp _t	3.485	3.279	3.462	3.545
srp _t	3.706	3.529	3.648	3.767
trp _t	2.255	2.132	2.219	2.305

so that $p = 2$ was deemed reasonably adequate. The estimated arrays $\hat{\pi}_1, \hat{\pi}_2, \hat{D}_t,$ and $\hat{\Sigma}_z$ are as follows (t -values in parentheses and $-$ denotes a non-significant coefficient, at the 5% level)

$$\hat{\pi}_1 = \begin{bmatrix} 1.516 & - & - & - \\ (10.02) & & & \\ - & 1.505 & - & - \\ & (8.06) & & \\ - & - & 1.387 & - \\ & & (6.18) & \\ - & - & - & 1.307 \\ & & & (6.07) \end{bmatrix}, \quad \hat{\pi}_2 = \begin{bmatrix} -0.619 & - & - & - \\ (-4.22) & & & \\ - & -0.451 & - & - \\ & (-2.27) & & \\ - & - & - & - \\ - & - & - & - \end{bmatrix},$$

$$\hat{D}_t = \begin{bmatrix} - \\ 0.039 \\ (1.98) \\ - \\ - \end{bmatrix}, \quad \hat{\Sigma}_z = 10^{-5} \begin{bmatrix} 6.64 & -0.13 & 1.28 & 2.63 \\ -0.13 & 16.2 & 16.9 & 5.67 \\ 1.28 & 16.9 & 34.2 & 13.0 \\ 2.63 & 5.67 & 13.0 & 8.59 \end{bmatrix}.$$

The residuals produced individual Ljung–Box statistics that do not led us to reject the white noise hypotheses at the 5% significance level. That is,

$$\begin{aligned} \{zccp_t\} : Q(10) &= 16.047, & Q(20) &= 30.005, & Q(30) &= 35.244, \\ \{zfrp_t\} : Q(10) &= 9.122, & Q(20) &= 22.267, & Q(30) &= 29.001, \\ \{zsrp_t\} : Q(10) &= 12.157, & Q(20) &= 16.757, & Q(30) &= 21.084, \\ \{ztrp_t\} : Q(10) &= 20.386^*, & Q(20) &= 30.079, & Q(30) &= 33.705 \end{aligned}$$

(* in this case, the individual Ljung–Box statistic does not reject the white noise hypothesis at the 1% significance level).

We also computed the multivariate portmanteau statistic $\hat{Q}_h = T^2 \sum_{j=1}^h (T - j)^{-1} \text{tr} (\hat{C}'_j \hat{C}_0^{-1} \hat{C}_j \hat{C}_0^{-1})$ where $\hat{C}_j = \sum_{t=j+1}^T \hat{a}_t \hat{a}'_{t-j} / T$ and \hat{a}_t are the k -dimensional residuals of the estimated VAR(p) model, with $\hat{Q}_h \sim \chi^2(k^2(h - p))$, where $k = 4$ is the number of variables, $p = 2$ is the lag order of the fitted model, and h was chosen as 6 (for details on the use of this test, see [21]). We obtained $\hat{Q}_6 = 43.466$ and compared this value with $\chi^2(64)_{0.95} = 83.68$ so that we could not reject the white noise hypothesis for the errors at the 5% level.

3.3 Unrestricted forecasts

The 2005 Population interdecade census reported population figures for every geographic unit considered in the MZCM [13,14]. However, in 2009 INEGI made some adjustments to those figures and produced new estimated population figures. The official cumulative PGR, its forecast for each and every ring, and its corresponding 95% PI are shown in Table 4. There we see that all the official cumulative PGR figures fall within its probability interval.

Table 4. Officially estimated and forecasted cumulative PGR values for 2005.

Rings	Estimated figure (interdecade counting)	Lower 95% limit	Unrestricted forecast	Upper 95% limit
ccp_t	0.147	0.034	0.104	0.175
frp_t	3.477	3.350	3.458	3.567
srp_t	3.866	3.856	4.008	4.160
trp_t	2.694	2.534	2.627	2.720

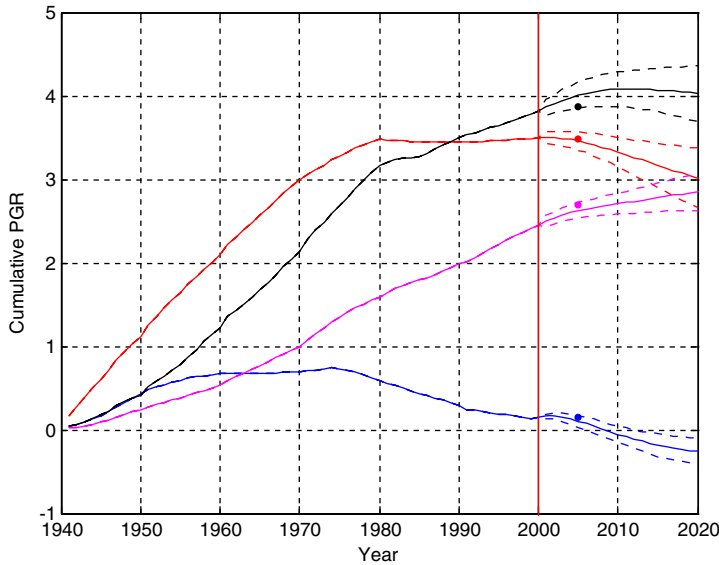


Figure 3. Unrestricted forecasts for 2001–2020 with 95% prediction intervals. Dots are official figures for ccp_t , trp_t , frp_t and srp_t (from bottom to top).

In Figure 3, we show the multiple unrestricted forecasts together with their PIs and the official figures reported in 2009.

3.4 Restricted forecasts and compatibility testing

In 1997, the DF Government presented a population program [7]. Part B of that program included intended growth rates for the Central City and the rings. The specific demographic goals for population growth of the rings are: (a) to reach a growth rate of 0.4% between 2006 and 2010 and 0.9% between 2010 and 2020 for the Central City; (b) to reduce the growth rate to 0.3% between 2000 and 2003, increase it to 0.5% in 2006–2010, and reduce it to 0.3% between 2010 and 2020 for the first ring; (c) to reduce the growth rate to 1.2% between 2000 and 2003, to 1.1% between 2003 and 2006, to 0.7% of 2006–2010, and to 0.5% in the following decade for the second ring; (d) to reduce the growth rate to 2.4% between 2000 and 2003, to 2.2% between 2003 and 2006, to 0.8% between 2006 and 2010, and to 0.7 in the following decade for the third ring. We understood these values as goals to be reached at the end of every period and translated them into binding restrictions to be imposed on the forecasts of the cumulative PGR, as shown in Table 5.

Downloaded By: [Silva, Eliud] At: 13:54 6 February 2011

Table 5. Restricted forecasts for the concentric rings.

Restricted forecast	Year			
	2003	2006	2010	2020
\widehat{zccp}_t	Without target	Without target	$\widehat{Fzccp}_{2005} + 0.004^5$	$\widehat{zccp}_{2010} + 0.009^{10}$
\widehat{zfrp}_t	$zfrp_{2000} + 0.003^3$	Without target	$\widehat{Fzfrp}_{2005} + 0.005^5$	$\widehat{zfrp}_{2010} + 0.003^{10}$
\widehat{zsrp}_t	$zsrp_{2000} + 0.012^3$	$\widehat{zsrp}_{2003} + 0.011^3$	$\widehat{zsrp}_{2006} + 0.007^4$	$\widehat{zsrp}_{2010} + 0.005^{10}$
\widehat{ztrp}_t	$ztrp_{2000} + 0.024^3$	$\widehat{ztrp}_{2003} + 0.022^3$	$\widehat{ztrp}_{2006} + 0.008^4$	$\widehat{ztrp}_{2010} + 0.007^{10}$

Note: $zfrp_{2000}$, $zsrp_{2000}$, $ztrp_{2000}$ are observed 2000 census data. \widehat{Fzccp}_{2005} and \widehat{Fzfrp}_{2005} are unrestricted forecasts of ccp_t and frp_t produced by the VAR(2) model.

To take the previous restrictions into account, we define the Y vector as in Equation (10) and the C matrix with the following structure

$$C = \begin{bmatrix} 0_{3 \times 9} & I_3 & 0_{3 \times 68} \\ 0_{2 \times 22} & I_2 & 0_{2 \times 56} \\ 0_{4 \times 36} & I_4 & 0_{4 \times 40} \\ & 0_{4 \times 76} & I_4 \end{bmatrix},$$

where $0_{i \times j}$ are $i \times j$ zero matrices and I_i are i -dimensional identity matrices.

We carried out compatibility testing of these goals and obtained the value $K_{calc} = 3.45$ which is significant at the 5% level, as compared with an $F_{M,T-Mp-1}$ distribution with $M = 13$ and $T - Mp - 1 = 33$ degrees of freedom. Therefore, the goals are jointly incompatible with the expected behavior of the multiple population series. However, the partial compatibility tests indicate that the goals \widehat{zccp}_{2010} and \widehat{zccp}_{2020} may be considered compatible at the 0.07% and 0.05% significance level, respectively.

Although the set of goals established in the population program are not jointly compatible, we shall elaborate on them and make a proposal on the PGRs for the rings. The idea is to find a set of population targets that are compatible with the empirical evidence provided by the annual disaggregated series. Our proposal looks for population targets that produce a smooth population pattern, in agreement with the demographic logic, if no catastrophic or anomalous situation occurs. By so doing, we obtained the multiple compatibility test statistic $K_{calc} = 0.74$ with p -value 7.12%.

Since the unrestricted forecast for the Central City population has a clear decreasing trend, we suggest reaching a cumulative PGR of 0 at the end of 2010 and fix a negative cumulative PGR of 0.18% at the end of 2020. For the first ring, all the targets of population growth were compatibles, but to get a smooth pattern of population we propose a cumulative PGR of 3.5% at the end of 2003, 3.38% in 2010, and 3% in 2020.

Our proposal for the cumulative PGR, based on the demographic dynamics presented by the series for the second ring, is 3.9% at the end of 2003 and 3.98% in 2006, then it should go up to 4.06% in 2010 and 4.1% at the end of 2020. For the third ring, our proposal is to modify only the first target at the end of 2003, that is, to reach a cumulative PGR of 2.55%, then reach 2.64% at the end of 2006, 2.68% in 2010, and 2.77% in 2020. In Table 6 we see that all the individual restrictions of our proposal are compatible with the disaggregated series at the 5% significance level, so that they are empirically supported.

Finally, in Figure 4 we can see the expected behavior of the population series for the Central City and the rings. The observed patterns are reasonably smooth for all the series except for the Central City population. We think this is a consequence of imposing constraints on that series that essentially tend to lower its cumulative PGR, so that the restricted forecasts have to bend the smooth curve in order to fulfill the constraints. In summary, we conclude that the goals proposed

Table 6. Compatibility testing for growth rates with our proposal.

Restriction	K_{parc}	$M, T - Mp - 1$	Significance
frp ₂₀₀₃	0.073	1, 57	0.788
srp ₂₀₀₃	0.550	1, 57	0.461
trp ₂₀₀₃	0.400	1, 57	0.530
srp ₂₀₀₆	0.368	1, 57	0.547
trp ₂₀₀₆	0.013	1, 57	0.910
ccp ₂₀₁₀	1.196	1, 57	0.279
frp ₂₀₁₀	0.374	1, 57	0.544
srp ₂₀₁₀	0.031	1, 57	0.862
trp ₂₀₁₀	0.197	1, 57	0.659
ccp ₂₀₂₀	1.052	1, 57	0.309
frp ₂₀₂₀	0.004	1, 57	0.951
srp ₂₀₂₀	0.190	1, 57	0.665
trp ₂₀₂₀	0.394	1, 57	0.533

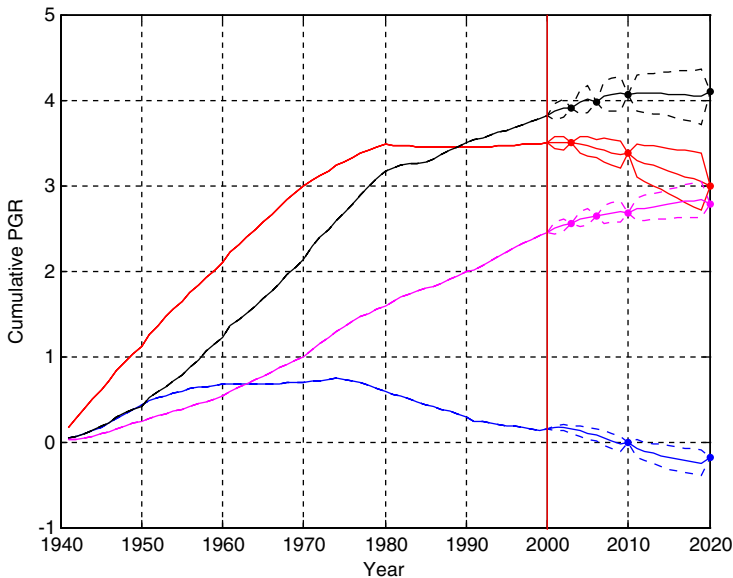


Figure 4. Restricted forecasts and 95% prediction intervals for the rings with our proposed goal for ccp_t , trp_t , frp_t and srp_t (from bottom to top).

for the Central City should be different than those presented in the official population program, while those for the rings must be in general only slightly different.

4. Conclusions

We presented first an application of a temporal disaggregation technique to a demographic time series. The most time-consuming part of such a technique involved the generation of an appropriate preliminary series from demographic considerations. We think it was worth doing it this way because the quality of the final results depends heavily on the quality of such a series. This task is much simpler to perform in economic contexts, because usually there are economic indexes that play the role of a basic auxiliary variable when obtaining a preliminary estimate of the unobserved series. In the case considered here, we were forced to perform a meticulous search for demographic data and events by geographic unit and year.

Downloaded By: [Silva, Eliud] At: 13:54 6 February 2011

The application of restricted forecasting and compatibility testing to demographic data was carried out in order to evaluate the feasibility of the targets proposed in an official population program for the MZMC. This analysis indicates that before suggesting demographic goals, it is necessary to evaluate their empirical feasibility in an objective way.

In general, it can be noticed that temporal disaggregation produced convincing results because we took great care to generate an adequate preliminary series. Also, it was possible to determine new goals consistent with the population dynamics. Of course, the methodological strategies presented here can also be used to solve similar problems with demographic information in other developing countries or in any other geographic zone where the need of combining demographic information arises. In fact, according with the results obtained in this paper, we could say that both temporal disaggregation and restricted forecasting are efficient statistical tools that serve to consolidate this type of data. Moreover, we are convinced that growth population programs could be made feasible and monitored afterwards with this kind of analysis.

This article provides evidence on the appropriateness and accessibility of specialized statistical techniques, that have been developed and traditionally employed for economic analysis, in the demographic field. We hope this work motivates specialists in these fields to identify new potential applications or possibilities of methodological developments that will ultimately help practitioners to get more information from their data and support better decision-making. It is worth stressing that the procedures applied here can be used with other kind of demographic data, such as those related with fertility, marriage, divorce, and migration. By so doing, we could evaluate the feasibility of official population programs for the population determinants jointly, in different contexts and around the world.

Acknowledgements

The authors gratefully acknowledge the comments provided by Fátima Juárez and Ismael Sánchez to a previous version of this paper. They also thank two anonymous referees and the editor in charge for their valuable comments that helped to improve the presentation of this work. V.M. Guerrero thanks Asociación Mexicana de Cultura, A.C. for providing support to carry out this work through a Professorship on Time Series Analysis and Forecasting in Econometrics. We also acknowledge support from the MICINN Grant SEJ2007-64500.

References

- [1] G. Chow and A. Lin, *Best linear unbiased interpolation, distribution and extrapolation of time series by related series*, Rev. Econ. Stat. 53 (1971), pp. 372–375.
- [2] M. Clements and A. Galvão, *Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the United States*, J. Bus. Econ. Stat. 26 (2008), pp. 546–554.
- [3] T. Di Fonzo, *The estimation of m disaggregate time series when contemporaneous and temporal aggregates are known*, Rev. Econ. Stat. 72 (1990), pp. 178–182.
- [4] T. Di Fonzo, *Temporal disaggregation of a system of time series when the aggregate is known: Optimal vs. adjustment methods*, Workshop on Quarterly National Accounts, Paris, 1994, pp. 1–22.
- [5] J.M. Dufour, *Unbiasedness of predictions from estimated vector autoregressions*, Econometric Theory 1 (1985), pp. 387–402.
- [6] E. Ghysels, P. Santa-Clara, and R. Valkanov, *Predicting volatility: Getting the most out of return data sampled at different frequencies*, J. Econometrics 131 (2006), pp. 59–95.
- [7] Gobierno del DF y Consejo de Población del Distrito Federal, *Programas de Población del Distrito Federal 2001–2006: Nivel normativo (Metas)*. Dirección de Política Poblacional, 1997. Available at <http://www.df.gob.mx/secretarias/social/copodf/index.html/> [Accessed 30 June 2006].
- [8] N. Gómez and V.M. Guerrero, *Restricted forecasting with VAR models: An analysis of a test for joint compatibility between restrictions and forecasts*, Int. J. Forecast. 22 (2006), pp. 751–770.
- [9] V. Gómez, A. Maravall, and D. Peña, *Missing observations in ARIMA models: Skipping approach versus additive outlier approach*, J. Econometrics 88 (1999), pp. 341–363.
- [10] V.M. Guerrero and F.H. Nieto, *Temporal and contemporaneous disaggregation of multiple economic time series*, Test 8(2) (1999), pp. 459–469.

- [11] V.M. Guerrero and D. Peña, *Linear combination of restrictions and forecasts in time series analysis*, J. Forecast. 19 (2000), pp. 103–122.
- [12] V.M. Guerrero and D. Peña, *Combining multiple time series predictors: A useful inferential procedure*, J. Statist. Plann. Inference 116 (2003), pp. 249–276.
- [13] INEGI, *II Censo de población y vivienda 2005: Distrito Federal*. Dirección General de Estadística, México, 2005. Available at <http://www.inegi.org.mx/est/contenidos/espanol/proyectos/conteos/conteo2005/bd/consulta2005/pt.asp?s=est&c=10401> [Accessed 19 February 2009].
- [14] INEGI, *II Censo de población y vivienda 2005: Estado de México*, Dirección General de Estadística, México, 2005. Available at <http://www.inegi.org.mx/est/contenidos/espanol/proyectos/conteos/conteo2005/bd/consulta2005/pt.asp?s=est&c=10401> [Accessed 19 February 2009].
- [15] INEGI, *Censos Generales de Población y Vivienda: Distrito Federal*, Dirección General de Estadística, México, 1940–2000.
- [16] INEGI, *Censos Generales de Población y Vivienda: Estado de México*, Dirección General de Estadística, México, 1940–2000.
- [17] INEGI, *Estadísticas vitales: Distrito Federal*, Sistema Municipal de Base de Datos (SIMBAD), México 1994–2000. Available at <http://www.inegi.gob.mx/inegi/default.asp/> [Accessed 30 June 2006].
- [18] INEGI, *Estadísticas vitales: Estado de México*, SIMBAD, México, 1994–2000. Available at <http://www.inegi.gob.mx/inegi/default.asp/> [Accessed 30 June 2006].
- [19] INEGI, *I Censo de población y vivienda 1995: Distrito Federal*, Dirección General de Estadística, SIMBAD, México, 1995. Available at <http://www.inegi.gob.mx/inegi/default.asp/> [Accessed 8 December 2006].
- [20] INEGI, *I Censo de población y vivienda 1995: Estado de México*, Dirección General de Estadística, SIMBAD, México, 1995. Available at <http://www.inegi.gob.mx/inegi/default.asp/> [Accessed 8 December 2006].
- [21] H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*, Springer, New York, 2005.
- [22] F.H. Nieto, *Ex post and ex ante prediction of unobservable multivariate time series: A structural-model based approach*, J. Forecast. 26 (2007), pp. 53–76.
- [23] F.H. Nieto and V.M. Guerrero, *Kalman filter for singular and conditional state-space models when the system state and the observational error are correlated*, Stat. Probab. Lett. 22 (1995), pp. 303–310.
- [24] A. Pankratz, *Time series forecasts and extra-model information*, J. Forecast. 8 (1989), pp. 75–83.
- [25] J. Pavía, *Desagregación conjunta de series anuales: Perturbaciones AR(1) multivariante*, Investigaciones económicas 24(3) (2000), pp. 727–737.
- [26] N. Rossi, *A note on the estimation of disaggregate time series when the aggregate is known*, Rev. Econ. Stat. 64 (1982), pp. 695–696.
- [27] Secretariat of Health, *Compendio histórico: Estadísticas vitales 1893–1993, Distrito Federal*, Dirección General de Estadística Informática y Evaluación, México, 1993.
- [28] Secretariat of Health, *Compendio histórico: Estadísticas vitales 1893–1993, Estado de México*. Dirección General de Estadística Informática y Evaluación, México, 1993.
- [29] W.W.S. Wei, *Time series analysis*. Addison-Wesley, California, 1990.

Appendix 1. Correcting the annual birth series for outliers

We used annual vital statistics of births and deaths for years $t = 1941, 1942, \dots, 2000$ for the DF and the SM. The vital statistics in year t are: ddf_t deaths in the DF, bdf_t births in the DF, $ngdf_t = bdf_t - ddf_t$ natural growth in the DF, dsm_t deaths in the SM, bsm_t births in the SM, and $ngsm_t = bsm_t - dsm_t$ natural growth in the SM. We corrected the annual $\{bsm_t\}$ series for outliers in years 1983, 1984, and 1985. This series was regressed on marriages nsm_t of the SM, a dummy variable $i77_t$ that accounts for a structural change in 1977 due to a birth control policy (valued 0 from 1940 up to 1976 and 1 from 1977 onwards), three dummies for the aforementioned outliers i_{it} with $i = 1, 2, 3$, and lagged values of births bsm_{t-1} to account for its own inertia. Optimal corrected values were obtained as indicated in [9], from the model expressed in logarithms and $t = 1941, 1942, \dots, 2000$,

$$\log(bsm_t) = \beta_1 i77_t + \beta_2 \log(nsm_{t-1}) + \sum_{i=3}^5 \beta_i i_{it} + \beta_6 \log(bsm_{t-1}) + a_t,$$

where a_t is a random error from a 0-mean Gaussian white noise process. Since the dummy variable i_{3t} was not significant at the 5% level, it was omitted. The results of the final estimated model are

(*t*-statistics in parentheses),

$$\hat{\beta}_1 = -0.11, \quad \hat{\beta}_2 = 0.21, \quad \hat{\beta}_3 = -0.68, \quad \hat{\beta}_4 = -0.24, \quad \hat{\beta}_6 = 0.83$$

$$\begin{matrix} (-2.13) & (-3.18) & (-6.75) & (-2.30) & (-15.42) \end{matrix}$$

with $R^2 = 0.977$ and $\hat{\sigma}_a = 0.098$. The Ljung–Box Q statistic for serial autocorrelation took on the followings values $Q(3) = 1.23$, $Q(5) = 2.98$, and $Q(10) = 13.17$, which did not provide evidence of inadequacy of the model.

Appendix 2. Obtaining the preliminary series

The algorithm to get preliminary series for the rings from data on the DF and the SM goes as follows.

- I. Calculate $ngdf_t$ and $ngsm_t$, for $t = 1941, 1942, \dots, 2000$.
- II. Calculate dfp_t and smp_t as partial populations of the DF and the SM $dfp_t = dfp_{2000-n} + \sum_{j=2000-n+1}^{2000} ngdf_j$ and $smp_t = smp_{2000-n} + \sum_{j=2000-n+1}^{2000} ngsm_j$ with $n = 60$.
- III. Consider the population proportions of the rings with respect to $dfp_{t^*} + smp_{t^*}$ for the census years $t^* = 1940, 1950, \dots, 2000$, represented by $rccp_{t^*}, rfrp_{t^*}, rsrp_{t^*}$, and $rtrp_{t^*}$. Calculate the proportions for the intercensus years t , given by $rccp_t, rfrp_t, rsrp_t$, and $rtrp_t$, assuming a linear behavior. The series of estimated proportions satisfy the relationship $rmzmc_p_t = rccp_t + rfrp_t + rsrp_t + rtrp_t$ for $t = 1941, 1942, \dots, 2000$.
- IV. Calculate the series of population proportions for the rings with $ccp_t^+ = rccp_t(dfp_t + smp_t)$, $srp_t^+ = rsrp_t(dfp_t + smp_t)$, and $trp_t^+ = rtrp_t(dfp_t + smp_t)$ for $t = 1941, 1942, \dots, 2000$.
- V. Calculate the differences attributable to migration for the rings: $migccp_{t^*} = ccpc_{t^*} - ccp_{t^*}^+$, $migrfp_{t^*} = frpc_{t^*} - frp_{t^*}^+$, $migrsrp_{t^*} = srpc_{t^*} - srp_{t^*}^+$, and $migrtrp_{t^*} = trpc_{t^*} - trp_{t^*}^+$ for $t^* = 1940, 1950, \dots, 2000$, where $ccpc_{t^*}, frpc_{t^*}, srpc_{t^*}$, and $trpc_{t^*}$ are census populations at years t^* . Suppose migration behaves uniformly in time and add one-tenth of these differences to the annual estimates series obtained in the previous step.
- VI. Calculate $ccp_t = ccpc_{t^*} + 0.1 * migccp_{t^*}$, $frp_t = frpc_{t^*} + 0.1 * migrfp_{t^*}$, $srp_t = srpc_{t^*} + 0.1 * migrsrp_{t^*}$, and $trp_t = trpc_{t^*} + 0.1 * migrtrp_{t^*}$ for $t = 1941, 1942, \dots, 2000$.
- VII. Finally, estimate the cumulative PGR between 1940 and each year from 1941 to 2000, for every ring, as $r = \log_e(p_t/p_{1940})$.

Appendix 3. Incorporating measurement error variability

Let $\{Z_t^*\}$ and $\{Z_t\}$ be series of observable and unobservable values, respectively, that admit stationary VAR representations of order $1 \leq p < \infty$ with all non-stationarities taken into account by the deterministic effects, so that their Wold’s representations are $Z_t^* = D_t + \Psi(B)a_t$ and $Z_t = D_t + \Psi(B)a_t + \epsilon_t$ with D_t a vector of deterministic components that includes the constant term, $a_t \sim N(0, \Sigma_a)$, and ϵ_t is the measurement error which we assume is uncorrelated with the whole sequence $\{a_t\}$. Then, we recall the notation in Section 2.2 to write

$$Z^* = D + \Psi a_F \quad \text{and} \quad Z = D + \Psi a_F + \epsilon_F \tag{A1}$$

with $a_F \sim N(0, I_H \otimes \Sigma_a)$ and $\epsilon_F \sim N(0, \Sigma_\epsilon)$ and $E(a_F \epsilon_F') = 0$, where Σ_ϵ is given in Equation (20) and D is a vector of deterministic elements. We then write $Z_F = D + \Psi \delta_F$, with

$$\delta_F \sim N(0, (I_H \otimes \Sigma_a) + \Psi^{-1} \Sigma_\epsilon \Psi^{-1'}) \tag{A2}$$

to obtain again expressions (11)–(13), with

$$A = [\Psi(I_H \otimes \Sigma_a)\Psi' + \Sigma_\epsilon]C'\Omega^{-1} \tag{A3}$$

and

$$\Omega = C[\Psi(I_H \otimes \Sigma_a)\Psi' + \Sigma_\epsilon]C'. \tag{A4}$$