WORKING PAPERS AND STUDIES

ISSN 1725-4825

2005 EDITION

Theory and applications of univariate and multivariate models for temporal disaggregation

COPYRIGHT

Luxembourg: Office for Official Publications of the European Communities, 2005

ISSN 1725-4825
ISBN 92-79-01328-9
Cat. No. KS-DT-05-031-EN-N

EUROPEAN COMMISSION

eurostat

EURO INDICATORS

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server (http://europa.eu.int).

Luxembourg: Office for Official Publications of the European Communities, 2005

ISSN 1725-4825
ISBN 92-79-01328-9

Theory and applications of univariate and multivariate models for temporal disaggregation

Gian Luigi Mazzi[†], Filippo Moauro[‡] and Giovanni Savio[†]

[†]*Statistical Office of the European Communities, Eurostat*
5, rue Alphonse Weicker, Luxembourg, L-2920
E-mail: Gianluigi.Mazzi@cec.eu.int and Giovanni.Savio@cec.eu.int
[‡]*Istituto Nazionale di Statistica, Istat*
Via A. Depretis 74/B, 00184 Rome, Italy
E-mail: Moauro@istat.it

1. INTRODUCTION

Infra-annual macro-economic statistics represent nowadays a key tool for economic policy-making and business cycle analysis. These need, of course, reliable data to be carried out. However, reliability is only one of the quality characteristics data should have. Users need long, homogeneous and high-frequency time series, possibly covering a wide spectrum of key economic phenomena. Furthermore, geographical aggregates - such as those for the EU/EMU zones - are increasingly requested and analysed for a number of reasons, notably for monetary purposes.

Despite these pressures, official statistics have still shortcomings – such as lack of stability over time and thus insufficient length, definition at lower frequencies than those traditionally needed for business cycle analysis, lack of timeliness etc. – that could be in part remedied with the use of analytical tools.

It is an increasing belief that such tools can play an important role in the short-medium term in gathering momentum for our statistics, and in filling the existing gaps and shortcomings. However, the use of such techniques should be viewed as a temporary solution, while waiting for more structural improvements to take effect, i.e. in the data collection and elaboration processes at the national level.

In particular, what has emerged thanks to the efforts of the 'benchmarking community' and its interaction with official statisticians, is that these techniques can give a clear contribution to improve specific aspects of data quality, such as coherence, timeliness and completeness. Optimal benchmarking methods can contribute to an increase in accuracy of final estimates and relevance of our statistics. In the short-medium term, when resources are fixed and the capacity of national and international statistical systems to react to abrupt or planned changes are limited, these techniques often succeed in temporarily satisfy users' needs.

In a very broad sense, benchmarking techniques are those processes optimally combining two or more sources of measurement to obtain reliable estimates of the series under investigation. Though traditional fields of application of benchmarking techniques are national accounts, censuses and demographic data, employment, administrative records and cross-section data, it is essentially in the first field that these techniques have known a wider application, both in Eurostat and in many EU and non-EU NSIs.

The interest of Eurostat for benchmarking issues dates back to the beginning of the '90s, when it become clear that these could become a quite natural estimation framework for Quarterly National Accounts. Since then, it has become more and more evident that backward estimation, extrapolation, temporal disaggregation, and benchmarking techniques could play a central role in meeting important challenges for official statistics.

The discipline of benchmarking has made considerable advancements over the last decade promoting new theoretical developments, improvements in traditional methods, new ideas in applications, and the development of freely available software and tools (Di Fonzo 2003c; Abad and Quilis 2005; Quilis 2005; Palate 2005). Nowadays, benchmarking is widely used - sometimes more than thought - in national and international statistical agencies, and it is a continuing source for research and empirical studies in statistics, econometrics and, more generally, time series analysis.

Eurostat has played a leading role in all these developments. The official reference to the so-called 'indirect method' of estimation dates back to the European System of Accounts, ESA 1995 (Eurostat 1996, par. 2.04), and it has been further developed in the Eurostat Handbook on Quarterly National Accounts (Eurostat 1999), with two Chapters of about 50 pages dedicated to the theory and use of 'Mathematical and statistical methods' in the estimation of national accounts figures.

Benchmarking techniques have known a wider dissemination at the EU level and worldwide through the release by Eurostat of a statistical/econometric tool for temporal disaggregation called Ecotrim. This software, already developed during the first half of the last decade, has undergone further refinements up to the current releases, which can run in different environments, and are routinely used for a number of purposes in many EU and non-EU NSIs and by many researcher.

Eurostat, as well as other international organisations such as the OECD, have regarded the issue of benchmarking and temporal disaggregation as a high priority in its work with NSIs, trying to keep abreast of the numerous developments and to stimulate further research in the field.

Indeed, theoretical and empirical advances in benchmarking and temporal disaggregation have been so impressive in the last 5-10 years, that Eurostat and the OECD have decided to jointly organise a Workshop on 'Frontiers in Benchmarking Techniques and Their Application to Official Statistics', held in Luxembourg on April 2005.

The Workshop has been a great success, attracting more then one hundred and fifty participants from all over the world. The quality of both invited and contributed papers has been very high; the same holds for the quality of the related talks[1]

[1] All reference material from the Workshop, including access to freely available software or descriptions of how to obtain it, is attached to the workshop program available at the Euroindicators dedicated section of the Eurostat site. The Workshop featured keynote lectures by eminent researchers in benchmarking, such as S. Brown, T. Di Fonzo, V. Guerrero, B. Quenneville, A. Ranbaldi, T. Proietti, A. Trabelsi, M. Weale.

The papers presented during the Workshop encompassed recent advances in several important areas for benchmarking, such as the use of dynamic and state-space models, model-based approaches, balancing techniques, and back-recalculation of time series. Specific sections have been dedicated to the comparison of alternative methods for benchmarking, the use of such techniques in official statistics, and recent developments of tools and software.

Eurostat is strongly committed to further promoting and enhancing the use of such benchmarking techniques, where appropriate, with NSIs and its partners and to encourage the use and experimentation of new benchmarking methods and tools. The Conference organised by Istat has to be considered a leading initiative in this area, as it gives the opportunity to have an integrated view of the discipline, to promote exchanges amongst research and official statistics, and to concretely apply recent developments to national data.

The primary target of Eurostat is to extend the Ecotrim tool in order to include such recent developments and to reflect the results of the numerous studies carried out with the support of external experts in the last five years. In this respect a specific project has been arranged with the Commission DG-JRC of Ispra, Italy, and the results are expected by the first half of 2007. The project foresees both the development of a new Ecotrim tool and the preparation of a new software for back-recalculation and the reconstruction of long time series, another important area of research with clear links with benchmarking issues (Di Fonzo 2003b; Bournay and Ladiray 2005; Caporin and Sartore 2005).

This paper aims at giving an overview of such recent theoretical developments and their links with traditional methods (Sections 2). We concentrate here on temporal disaggregation techniques because this is the focus of the event organised by Istat. After that, we move to a short description of a multivariate method for temporal disaggregation, which uses seemingly unrelated time series equations to solve the problem of missing observations (Section 3, an Appendix contains further details on the approach). Section 4 is dedicated to a comparison between univariate and multivariate methods for time disaggregation. Some case studies trying to shed light on the difficult task of the choice among different disaggregation methods are discussed in Section 5. Section 6 contains our conclusions.

2. CLASSICAL METHODS AND RECENT DEVELOPMENTS IN TEMPORAL DISAGGREGATION

Temporal disaggregation has been extensively considered in the econometric and statistical literature and numerous solutions have been proposed. Broadly speaking, two alternative methods have been considered so far:

- methods which do not use related series but rely upon purely mathematical criteria or time series models to derive a smooth path for the unobserved series;
- methods which make use of the information obtained from related indicators observed at the desired higher frequency.

The first approach comprises the model-based methods (Stram and Wei 1986; Wei and Stram 1990) relying on the $ARIMA$ representation of the series to be disaggregated (e.g., see Eurostat 1999, Ch. 6, for a survey and a taxonomy of temporal disaggregation methods). The latter approach includes, among others, the adjustment procedure developed by Denton (1971) and the methods proposed by Chow and Lin (1971), Fernández (1981) and Litterman (1983).

The last 5-10 years refinements of these traditional techniques have of course benefited from the advances of the econometric, statistical and economic literature in the last two decades or more. As such, they are less prone to critics that instead could be addressed to the traditional approaches indicated above. For example, Di Fonzo (2003b, pg. 2) states that these approaches 'sometimes demonstrates obsolete and not responding to the increasing demand for more sophisticated and/or theoretically well founded statistical and mathematical methods in estimating national accounts figures.'.

Recent developments have essentially followed two distinct lines of research:

1. univariate dynamic regression models - usually represented as autoregressive distributed lag (ADL) models -, possibly using non-linearly transformed data;

2. univariate or multivariate approaches that use formulations in terms of unobserved components and structural time series models (and possibly non-linearly transformed data) and Kalman filtering techniques to get optimal estimates of missing observations by a smoothing algorithm.

The first line of research comprises the work by Gregoir (1995), Santos Silva and Cardoso (2001), Santos Silva (2005), Mitchell et al. (2005), Mitchell and Weale (2005) and Guerrero (2005). The rationale under these papers have been further investigated and extended by Di Fonzo (2003a, 2003b)[2].

The second line of research has been exploited by Gudmundsson (1999), Hotta and Vasconcellos (1999), Proietti (1999) and Gómez (2000). The original idea dates back to Harvey and Pierce (1984), and has been further developed in the framework of structural time series models by Durbin and Quenneville (1997), Harvey (1989) and, in a recent application, by Harvey and Chung (2000), where use is made of seemingly unrelated structural time series (SUTSE) models to obtain timely estimates of the underlying change in unemployment. In the spirit of the work by Harvey and Chung (2000), Moauro and Savio (2005) have further developed the use of SUTSE models for temporal disaggregation.

Other fields of research at the border of the two lines discussed above are the use of nonlinear transformations of the data (Proietti 2004a; Di Fonzo 2003a; Mitchell et al. 2005) and the contemporaneous disaggregation and seasonal adjustment of time series in multivariate structural models (Moauro and Proietti 2005).

[2] A representation of the dynamic models in state-space form has been carried out in Proietti 2004b.

Both the rationale and the numerous dynamic formulations underlying dynamic models for temporal disaggregation are quite well known in the literature. Contrarily, the use of multivariate SUTSE models for time disaggregation is relatively recent. Therefore, the next Section briefly introduces this class of models.

3. MULTIVARIATE STRUCTURAL TIME SERIES MODELS AND THEIR USE FOR TEMPORAL DISAGGREGATION

These models jointly represent in a multivariate framework the series to be disaggregated and the set of related time series. The framework used in SUTSE models represents a multivariate generalization of standard structural time series models (see, e.g., Harvey 1989; Fernández and Harvey 1990; Harvey and Koopman 1997).

Given a cross-section of time series $\mathbf{y}_t = (y_{1t}, \ldots, y_{Nt})'$, it is assumed that each y_{it}, $i = 1, 2, \ldots, N$ and $t = 1, 2, ..., n$, is not directly related with the others, although the series are subject to similar influences. $\mathbf{y}_t$ is expressed in terms of additive N-dimensional unobserved components, e.g. level $\boldsymbol{\mu}_t$, slope $\boldsymbol{\beta}_t$, cycle $\boldsymbol{\psi}_t$, seasonality $\boldsymbol{\gamma}_t$ and irregular $\boldsymbol{\xi}_t$, which can be contemporaneously correlated.

SUTSE models can be formulated in a number of ways: the natural starting point is the multivariate local linear trend (LLT) model, where $\mathbf{y}_t$ consists of a stochastic trend plus a white noise:

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\mu}_t + \boldsymbol{\xi}_t, & \boldsymbol{\xi}_t &\sim NID(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\xi}}), & (1)\\ \boldsymbol{\mu}_t &= \boldsymbol{\mu}_{t-1} + \boldsymbol{\beta}_t + \boldsymbol{\eta}_t, & \boldsymbol{\eta}_t &\sim NID(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\eta}}), & (2)\\ \boldsymbol{\beta}_t &= \boldsymbol{\beta}_{t-1} + \boldsymbol{\zeta}_t, & \boldsymbol{\zeta}_t &\sim NID(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\zeta}}), & (3) \end{aligned}$$

where the $\boldsymbol{\Sigma}_h$'s, $h = \boldsymbol{\xi}, \boldsymbol{\eta}$ and $\boldsymbol{\zeta}$, are the covariance matrices of system disturbances, $\boldsymbol{\xi}_t$, $\boldsymbol{\eta}_t$ and $\boldsymbol{\zeta}_t$, assumed to be mutually uncorrelated in all time periods.

The LLT model may take a variety of forms: if $\boldsymbol{\Sigma}_{\boldsymbol{\zeta}} = \mathbf{0}$ the stochastic slope reduces to a fixed slope and the trend reduces to a multivariate random walk with drift (RWD); when $\boldsymbol{\Sigma}_{\boldsymbol{\eta}} = \mathbf{0}$, while $\boldsymbol{\Sigma}_{\boldsymbol{\zeta}}$ is positive semi-definite, a smooth trend or integrated random walk (IRW) is obtained; $\boldsymbol{\Sigma}_{\boldsymbol{\eta}} = \boldsymbol{\Sigma}_{\boldsymbol{\zeta}} = \mathbf{0}$ implies a deterministic linear trend. Different forms also arise when restrictions on the covariance matrices $\boldsymbol{\Sigma}_h$'s are introduced. The restrictions can concern the rank of any of the $\boldsymbol{\Sigma}_h$'s, implying a common component restriction, and/or proportionality of the $\boldsymbol{\Sigma}_h$'s to each other, that is homogeneity.

The LLT collapses to the local level (LL) model when there is no slope component. Then, the system is defined by equation (1) and by a random walk (RW): $\boldsymbol{\mu}_{t+1} = \boldsymbol{\mu}_t + \boldsymbol{\eta}_t$. The restriction $\boldsymbol{\Sigma}_{\boldsymbol{\eta}} = \mathbf{0}$ leads the RW to become a constant level. In both cases the LLT and the LL models can allow for more complicated expressions by introducing a cyclical and/or a seasonal component.

Equations (1)-(3) can be written in a more compact form in the following SSF:

$$\begin{aligned}\boldsymbol{\alpha}_{t+1} &= \mathbf{T}_t\boldsymbol{\alpha}_t + \mathbf{H}_t\boldsymbol{\varepsilon}_t,\ \boldsymbol{\alpha}_1 \sim N(\mathbf{0}, \mathbf{P}), & (4)\\ \mathbf{y}_t &= \mathbf{Z}_t\boldsymbol{\alpha}_t + \mathbf{G}_t\boldsymbol{\varepsilon}_t, & (5)\end{aligned}$$

where the system matrices $\mathbf{T}_t$, $\mathbf{H}_t$, $\mathbf{Z}_t$ and $\mathbf{G}_t$ can be time-varying to allow for missing observations. The state vector $\boldsymbol{\alpha}_t$ is such that $\boldsymbol{\alpha}_t = (\boldsymbol{\mu}_t', \boldsymbol{\beta}_t')'$, $\boldsymbol{\varepsilon}_t \sim NID(\mathbf{0}, \mathbf{I})$ and, dropping the subscripts, the system matrices are $\mathbf{T} = \begin{pmatrix} \mathbf{I}_N & \mathbf{I}_N \\ \mathbf{0} & \mathbf{I}_N \end{pmatrix}$, $\mathbf{H} = \begin{pmatrix} \boldsymbol{\Gamma}_\eta & \boldsymbol{\Gamma}_\zeta \\ \mathbf{0} & \boldsymbol{\Gamma}_\zeta \end{pmatrix}$, $\mathbf{Z} = [\mathbf{I}_N, \mathbf{0}]$ and $\mathbf{G} = [\mathbf{0}, \mathbf{0}, \boldsymbol{\Gamma}_{\boldsymbol{\varepsilon}}]$, with the $\boldsymbol{\Gamma}_{\mathbf{h}}$'s lower triangular matrices such that $\boldsymbol{\Sigma}_{\mathbf{h}} = \boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Gamma}_{\mathbf{h}}'$ and $\mathbf{h} = \boldsymbol{\eta}$, $\boldsymbol{\zeta}$, $\boldsymbol{\xi}$.

SUTSE models are estimated in the time domain by using the Kalman Filter (KF). Once their State-Space Form (SSF) have been set up, the KF yields the one-step ahead prediction errors and the Gaussian log-likelihood function via the prediction error decomposition. The system matrices $\mathbf{T}_t$, $\mathbf{H}_t$, $\mathbf{Z}_t$, and $\mathbf{G}_t$ of the SSF (4)-(5) depend on a set of unknown parameters, denoted by $\boldsymbol{\varphi}$. Then, numerical optimization routines can be used to maximize the log-likelihood function with respect to $\boldsymbol{\varphi}$. Once $\boldsymbol{\varphi}$ has been estimated, the output of the KF may be used for different purposes, such as forecasting, diagnostic checking and smoothing. In particular, the backward recursions given by the smoothing algorithm yield optimal estimates of the unobserved components. The treatment of the diffuse initial conditions is efficiently approached by using the methods proposed by Koopman (1997) and Koopman and Durbin (2003).

The use of SUTSE models for temporal disaggregation is straightforward in the KF framework, the disaggregation problem being treated as a missing observation problem. What is required in this case is an adjustment of the SSF of the general multivariate model through the use of a cumulator variable for each series contained in the vector $\mathbf{y}_t$ subject to temporal aggregation (see the Appendix for details).

In synthesis, the steps one concretely could follow when a SUTSE model is used for temporal disaggregation closely resemble a general-to-specific approach to time series modelling. These steps are:

1. start from the general multivariate LLT model, augmented in order to include seasonal and cyclical components, if appropriate;

2. once a first estimation is run, some of the variances of the system in the LLT model can be fixed by the system estimation itself because they are approximately equal to zero;

3. test for the form of the trend component and the existence of common factors can be carried out using parametric and non-parametric tests (see the Appendix for further details and references);

4. if restrictions are accepted by the data, the parameters of the models can be estimated again by imposing the associated (rank) restrictions in order to obtain a more parsimonious model.

4. COMPARISON OF UNIVARIATE AND MULTIVARIATE METHODS

The main traditional methods for temporal disaggregation proposed by the literature hypothesize a simple linear univariate relationship between the unknown low-frequency variable y_{1t} and the high-frequency related time series y_{2t}. The most important difference among the various methods lies in the structure imposed on the disturbances of the hypothesized econometric relationship. Starting from the high frequency variables regression, $y_{1t} = \alpha + \beta y_{2t} + u_t$, the key issue - after temporal aggregation and model estimation using the observed series - is the identification of the covariance matrix of the disturbances of the high frequency model from the estimated covariance matrix of the low-frequency model. This model is derived from the available low-frequency data, possibly by imposing an $ARIMA$ structure on the data generating process of $u_t \mid y_{2t}$. This structure is assumed to be an $AR(1)$ process by Chow and Lin (1971), an $I(1)$ process by Fernández (1981), and an $ARIMA(1,1,0)$ process by Litterman (1983). Though the proposal of Stram and Wei (1986) encompasses the other models - the authors considers a general $ARIMA(p,d,q)$ structure of the data generation process of the aggregated series or, if an indicator is available, an $ARIMA(p,d,q)$ model for the disturbances - NSI's often use the Chow and Lin's family procedures because of their computational simplicity (see Bloem, Dippelsman and Mæhke 2001). The quadratic minimization approach suggested by Denton (1971), strongly favoured by Bloem, Dippelsman and Mæhke (2001, Ch. 6), can be easily seen as a special case of the least squares approach of Chow and Lin (1971) (see also Fernández 1981).

Implicit assumptions of the univariate approaches are the weak exogeneity of y_{2t} and the existence of a behavioural relation between y_{1t} and y_{2t}. As stated by Harvey (1989, pp. 463-465), none of these assumptions are necessarily fulfilled in current practices.

Then, the question naturally arises as to when is a multivariate approach *equivalent* to the univariate approaches utilised for temporal disaggregation by most NSIs so far. In order to theoretically compare the various methods, it is important to note that the reduced form of the LLT model is multivariate $IMA(2,2)$. Therefore, comparable models can be obtained only if restrictions are imposed on its form, and in particular on the MA component.

Let us assume that $N = 2$ and consider the following factorization of the covariance matrix $\mathbf{\Sigma}_h$'s for the h-th component, $h = \boldsymbol{\xi}$, $\boldsymbol{\eta}$ and $\boldsymbol{\zeta}$:

$$\mathbf{\Sigma}_h = \begin{pmatrix} \boldsymbol{\sigma}_{1h}^2 & \rho_h \boldsymbol{\sigma}_{1h} \boldsymbol{\sigma}_{2h} \\ \rho_h \boldsymbol{\sigma}_{1h} \boldsymbol{\sigma}_{2h} & \boldsymbol{\sigma}_{2h}^2 \end{pmatrix}. \tag{6}$$

Starting from the multivariate LL model, whose reduced form is multivariate $IMA(1,1)$, Harvey (1989, 1996) has shown the conditions to obtain fully efficient estimates of the parameters of interest from the univariate model, namely the hypotheses for y_{2t} to be weakly exogenous. However, these conditions lead to disturbances which are either $IMA(1,1)$ or white noise. The first case arises if the following results hold: (a) $\mathbf{\Sigma}_{\boldsymbol{\xi}}$ and $\mathbf{\Sigma}_{\boldsymbol{\eta}}$ are positive definite and the system is homogeneous, $\mathbf{\Sigma}_{\boldsymbol{\eta}} = q_\eta \mathbf{\Sigma}_{\boldsymbol{\xi}}$, with known homogeneity coefficient; (b) $\sigma_{2\xi} = 0$; (c) $\sigma_{2\eta} = 0$; (d) $\theta_\eta = \theta_\xi$ in the factorization of the covariance matrix discussed in the Appendix for the case of common components. The second case occurs if $\boldsymbol{\Sigma_\eta} = \mathbf{0},$ or if $\sigma_{2\xi} = 0$ with the further restriction that $\rho_h = 1$, namely common levels.

The condition that must be satisfied to obtain a single equation form equal to the model proposed by Fernández (1981) with $IMA(1,0)$ disturbances is, in some respects, more stringent, as we need $\boldsymbol{\Sigma_\xi} = \mathbf{0}$ (the two series are equal to their trend components). The same result is obtained by making the same assumption, but the starting point here is the LLT model with fixed slope. In this case, the final univariate model contains a time trend as an additional regressor. Further, weak exogeneity is only obtained when the slope is known and identical for both series.

The model proposed by Chow and Lin (1971) is obtained from a modified LL model where the level component is $AR(1)$ with autoregressive coefficient equal for both series, namely $\boldsymbol{\mu}_t = \phi\boldsymbol{\mu}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim NID(\mathbf{0}, \boldsymbol{\Sigma_\eta})$, with ϕ scalar and known. The conclusions here are the same as those discussed above for the model proposed by Fernández (1981). In this situation weak exogeneity can also be obtained by assuming the homogeneity restriction with q_ξ known. However, in the majority of cases q_ξ is unknown, then y_{2t} is no longer weakly exogenous and there is a loss of information in neglecting the equation for y_{2t}. In these cases, the single equation estimator is only asymptotically efficient. Similar conclusions can be obtained from a modified LL model with common levels and common $AR(1)$ disturbances. Because y_{1t} and y_{2t} are co-integrated, it follows from the results in Stock (1987) that the obtained estimates are (super-)consistent and efficient in large samples even if y_{2t} and the disturbances themselves are correlated by construction. Note that the hypothesis of common levels actually reduces the integration order of the error component, so that it is no longer an $I(1)$ process.

If the starting point is the modified LLT model with autoregressive slopes and $\boldsymbol{\Sigma_\xi} = \mathbf{0}$, weak exogeneity is obtained if the model is trend homogeneous, $\boldsymbol{\Sigma_\zeta} = q_\zeta\boldsymbol{\Sigma_\eta}$, with q_ζ known. Furthermore, as stated before, unless the autoregressive parameters are equal in the two equations and known, there is a loss of efficiency in parameters estimates if these are based on single equation estimation: in other words, we do not have weak exogeneity. However, the resulting disturbances of the single equation for y_{1t} are $ARIMA(1,1,1)$: then, a moving average component is added to the disturbances of the model proposed by Litterman (1983) and the final result can be considered as closer to the wider assumptions of the approach by Stram and Wei (1986) and Wei and Stram (1990).

It should be noted that the assumptions of common levels and slopes, even when $\boldsymbol{\Sigma_\xi} = \mathbf{0}$, do not lead us to obtain weak exogeneity for y_{2t} in the equation of interest. Therefore, as noticed for the model by Fernández (1981), the model favoured by Litterman (1983) is by itself in conflict with the likely property of co-integration between y_{1t} and y_{2t}.

In a univariate context, most of the classical models for temporal disaggregation are particular cases of the class of the $ADL(1,1)$ models of the form $y_{1t} = \phi y_{1t-1} + \alpha + \beta_0 y_{2t} + \beta_1 y_{2t-1} + u_t$, either on levels or first difference of the series (Di Fonzo 2003a; Proietti 2004b). The $ADL(1,1)$ model encompasses both the model of Chow and Lin (1971), when $\beta_1 = -\phi\beta_0$ (common factors restriction), and the model by Fernández (1981), when the condition $\phi = 1$ is further imposed. The scheme suggested by Santos Silva and Cardoso (2001) is again a particular case of the $ADL(1,1)$ model, when $\beta_1 = 0$ (partial adjustment). Finally, the model proposed by Litterman is a special case of the same $ADL(1,1)$ model in first differences, when the common factor restriction is imposed.

The $ADL(1,1)$ models have, of course, the great merit of generalizing the models traditionally used in the context of time disaggregation. As such, their strenght consists in designing a bridge toward more genuine dynamic extensions that do not impose *a priori* restrictions on the parameters of the model before these being validated by the data.

However, the $ADL(1,1)$ model is again nested in the multivariate SUTSE framework, when the series are characterised by a multivariate stationary $AR(1)$ process with drift and different autoregression coefficients. The conclusions here are the same as those reported above: the major consequence in neglecting the multivariate nature of the model is a loss of information and inefficiency in the estimation of the unknown parameters.

5. RESULTS OF THE ANALYSES

In this Section we analyse the results obtained by applying a number of disaggregation methods proposed so far by the literature. The results of the structural approach have been generated using the Ox program, version 3.0 (see Doornik 2001), and the SsfPack package (see Koopman, Shephard and Doornik 1999, 2002), while for the other methods we have mainly benefited from the program Ecotrim developed by Eurostat (see Eurostat 1999) and on Gauss routines developed by T. Di Fonzo on behalf of Eurostat (see Di Fonzo 2003c).

The first set of comparison is taken from Moauro and Savio (2005), whilst other examples are either part of research conducted at Eurostat, or specific runs carried out for presentation at the Istat Conference.

5.1 The OECD data set

The first set of analyses has been carried out by considering a wide data set drawn from OECD (2002). The time series utilised refer to the twelve biggest OECD countries in terms of GDP at current prices in 2001. Eight sets of bivariate data are used: 1) Industrial production index and deliveries in manufacturing; 2) GDP and industrial production index; 3) Consumer and producer price indices; 4) Private consumption and GDP; 5) GDP deflator and consumer price index; 6) Broad and narrow money supply; 7) Short-term and long-term interest rates; 8) Imports evaluated on a f.o.b. (free on board) and c.i.f. (costs, insurances and freights) basis. In total, 96 cases (8 data sets times 12 countries) are considered. Whenever available in the OECD data set, unadjusted data have been preferred to the corresponding seasonally adjusted series. The eight sets of data have been chosen in order to cover almost all the various situations which typically occur in practice.

The results of the disaggregations have been compared with the actual data using standard statistics, such as root mean squared or mean absolute errors. In what follows, we report the results obtained in terms of root mean squared percentage errors (RMSPE) only to save space the results obtained with other criteria do not change the ordering of results and the conclusions. The results of the comparisons among the various methods for disaggregation and some synthetic measures of their relative performances are displayed in Tables 1-3 (see Moauro and Savio 2005 for more details on the models chosen and the tests carried out for the form of the final model).

The results indicate that the multivariate structural approach is likely to be more accurate than the other methods in virtually all the distributions of time series. Table 1 indicates for each experiment and country the best and worst performing method in terms of RMSPE. Table 2 shows that the SUTSE approach has an overall probability to have the lowest RMSPE close to 55%. Over each competitor, the percentage varies from about 78% (against the univariate structural model) to about 92% (against Chow-Lin's approach), see Table 2, first column. For all the experiments, the average gain in terms of RMSPE varies from about 14% in the case of Litterman's procedure, to some 60% for Denton's model (Table 2, first row). Among the methods using related time series, the approach by Litterman seems to be, on average, the second-best solution for time disaggregation issues. This is probably due to the fact that this approach is the closest to the SUTSE approach because, on the one hand, it requires fewer restrictions on the form of the underlying multivariate system, and on the other hand it is more able to resemble the time behaviour of highly non-stationary time series. Almost at the same level, the approaches by Denton (1971) and Chow and Lin (1971) have quite unsatisfactory performances with respect to the other methods: they have lower RMSPEs, even compared with methods which do not use related time series, seldom above 40% of cases (Table 3). Furthermore, these methods have the highest RMSPEs in the whole sample in 34.4% and 27.1% of cases respectively (Table 2).

When no indicator is available, the best solution seems again to be offered by the use of a structural approach, with a gain in terms of accuracy over the approaches by Denton and Stram-Wei of about 15%. Another interesting result emerges from our experiments. Other things being equal, the use of a related series can have an impact on the accuracy of final estimates which greatly depends on the method used for time disaggregation. In fact, while the gains obtained passing from the univariate structural model to the SUTSE approach are substantial (about 35%, with an increase of the probability success of 42.8% and a zeroing of the probability failure), for Denton's approach the use of a related time series seems even to worsen final outcomes (a reduction of 3% in accuracy and a consistent increase of the probability failure). Therefore, what this limited experiment seems to indicate is that the choice of the method for time disaggregation can be even more relevant than the use of a good reference series, even if the use of this series can substantially add in terms of accuracy when an appropriate framework for time disaggregation is chosen. Furthermore, as noted before, the gains from using the SUTSE over the univariate structural approach can be low when the series show similar behaviour. This is what emerges from cases C and E in our data set where, on average, the target and the related series are characterised by similar patterns. Here the gain in terms of RMSPE from using an indicator series is 2.4% (against the average of 35.1% in the whole data set) and the probability success decreases to 62.5% from an average of 78.1.

.

5.2 Extensions to cases where $N > 2$

Extensions of the applications presented above to cases where $N > 2$ are straightforward even with SUTSE models, but in such a context an issue to be addressed is the choice of the related series. In a time disaggregation framework it is quite natural to select the related series by looking at the performance of the competing models in terms of forecast (or extrapolation, following the terminology used by Chow and Lin 1971). This could be viewed as an extension of what envisaged by the Eurostat (1996) (par. 12.04) for the choice of the method:

> *The choice between the different indirect procedures must above all take into account the minimisation of the forecast error for the current year, in order that the provisional annual estimates correspond as closely as possible to the final figures.*

Therefore, the alternative models can be evaluated on the basis of the out-of-sample RMSPEs obtained by comparing the true low frequency data with the sum of the high frequency extrapolated data.

Here we report the results obtained with an extended version of the Private consumption-GDP dataset that includes investment, inflation, the money supply and a short-term interest rate. A similar dataset has been extensively used in the framework of Real Business Cycle literature in order to study the 'stylised facts' of business fluctuations and the relative importance of nominal/real shocks in explaining the bulk of economic fluctuations.

Table 4 shows the results of the forecasting comparisons for Canada and the USA. In order to make the comparisons easy, we have included under Model 1, Case A, the in-sample RMSPEs for the bivariate models discussed in the previous paragraph (Table 1); the column Model 1, Case B, reports the corresponding results for the out-of-sample forecasts.

Following King, Plosser, Stock and Watson (1991), two extensions of the original Private consumption-GDP dataset are considered. The first consists in a trivariate system including Private consumption, GDP and Gross fixed capital formation (Model 2), the second is based on a six-variable dataset obtained by adding the deflated money supply (broad money), the GDP deflator growth rate and the short-term interest rate (Model 3).

In the comparisons the univariate approaches have not been considered for obvious reasons, and Denton's model because it is an adjustment method that can be used if $N \leq 2$. The sample for the forecasting evaluation has been kept equal to a fourth of the number of annual data (5 years for Canada, 10 years for the USA).

It is not easy to draw conclusions from this limited experiment. However, three points need to be noticed: (a) the SUTSE approach seems once again to be successful over competitors; (b) according to the out-of-sample RMSPEs, the final rank does not dramatically change from those reported under Model 1, Case A, for both countries; (c) in the forecasting examples, the differences in terms of RMSPEs among the various methods are bigger than those obtained for the in-sample exercises.

In the Canada example, the six-dimension model is the best solution for all approaches except for Litterman's, whilst for the USA a bivariate system should generally be preferred. In almost all cases there is a gain in extending the dataset from $N = 3$ to $N = 6$, but for the USA the best choice seems to be - in three out of four cases - a bivariate Consumption-GDP model.

As noticed above, once the approach for disaggregation has been chosen, one could get substantial gains from choosing the more appropriate set of related series. This set and its dimension could depend upon a number of factors, notably the nature of the disaggregation at hand, the country, and the preferred approach. For example, looking at the SUTSE model, the RMSPE obtained for the USA is about 20% higher for the full system than for the bivariate case, whilst for Canada it is about half. Hence, the suggestion for NSIs would be, of course, to choose the related series on a case-by-case basis, perhaps periodically checking their forecasting behaviour out-of-sample once a new set of low-frequency benchmark data set becomes available.

5.2 An example using Italian QNA series

In this illustration we show the inadequacy of some classical approaches to face a disaggregation problem of the ones routinely faced in current practices. The example takes the series of the annual Value Added and the quarterly Industrial Producton Index of the Metal Sector, defined over the sample 1977-2003. These series have been analysed e.g. in Proietti (2004), who compares the performances of some classical (Chow-Lin, Fernández and Litterman) approaches to quarterly disaggregation of value added in terms of in-sample forecasting outcomes.

In the limited sample here considered, the series show a clear tendency to move together, at least as far as concerns their levels and slopes. A univariate estimation of the $ADL(1,1)$ model carried out on the annualised series gives the following results (t-stats in parentheses under the coefficients):

$$y_{1t} = \underset{(2.82)}{-9092.83} + \underset{(2.67)}{0.477393} y_{1t-1} + \underset{(5.27)}{652.348} y_{2t} - \underset{(0.41)}{69.1383} y_{2t-1} + \underset{(0.07)}{37.1578} t.$$

The common factors restriction imposed by the Chow-Lin (1971) model is rejected by the data, as we obtain a $\chi^2(1)$ statistic equal to 5.3450 with a p-value equal to 0.0208. The same holds for the restrictions imposed by Fernandez (1981), with a test $\chi^2(2) = 16.108$, $p = 0.0003$. The estimation of the $ADL(1,1)$ model in first differences and the hypothesis of common factors restriction leading to Litterman's model is again rejected, with a $\chi^2(1) = 27.614$.

The annualised series could be described by a multivariate LLT model, with results given by:

$$\begin{array}{lll} \widetilde{\sigma}_{1\eta} = 36632.00 & \widetilde{\sigma}_{2\eta} = 80.95 & \widetilde{\rho}_{\eta} = 1.00 \\ \widetilde{\sigma}_{1\zeta} = 0.00 & \widetilde{\sigma}_{2\zeta} = 0.00 & \widetilde{\rho}_{\zeta} = 0.00 \\ \widetilde{\sigma}_{1\xi} = 27462.00, & \widetilde{\sigma}_{1\xi} = 67.40, & \widetilde{\rho}_{\xi} = 1.00 \end{array}$$

The starting estimation of the multivariate model should lead us to further investigate a LL model with common levels. Final estimates of such a system are as follows:

$$\begin{aligned} y_{1t} &= {\mu^{\dagger}}_t + \xi_{1t}, \\ y_{2t} &= 0.00127{\mu^{\dagger}}_t - 5.5174 + \xi_{2t}, \end{aligned}$$

where ${\mu^{\dagger}}_t$ is a univariate random walk. Thus we have the following relationship between the level components as they appear in the two series: $\mu_{2t} = 0.00127\mu_{1t} - 5.5174$.

Here again the alternative models can be evaluated on the basis of the out-of-sample RMSPEs obtained by comparing the true low frequency data with the sum of the high frequency extrapolated data. As in the preceding example, the sample for the forecasting evaluation has been kept equal 6 years, a fourth of the number of annual data. The RMSPE for the multivariate LL model is equal to 3.0544, slightly lower than the RMSPE obtained with the model of Fernández (3.0587), but with a good gain over the models by Chow-Lin (3.1362) and Litterman (3.2718)[3].

5.2 An estimation of monthly value added in Industry for the Euro-zone

Here we take the quarterly value added in Industry for the Euro-zone, defined over the sample 1991.1-2005.1 and the corresponding monthly industrial production index defined over the same sample period. The situation is quite close to the case discussed above, the exception being that the estimated relation gives a clear indication in favor of the $ADL(1,1)$ or the Chow-Lin model, as the common factor restriction is accepted by our data ($\chi^2(1) = 0.12392$, $p = 0.7248$):

$$y_{1t} = -\underset{(3.38)}{26889.5} + \underset{(5.49)}{0.581376} y_{1t-1} + \underset{(13.5)}{2632.24} y_{2t} - \underset{(3.96)}{1497.60} y_{2t-1} + \underset{(1.07)}{32.0415} t.$$

The results obtained in terms of forecasting RMSPEs over the last five quarters indicate for the $ADL(1,1)$ model a value equal to 4.0012, followed by the Chow-Lin (4.03309), and the partial adjustment scheme (4.13655). The assumptions underlying the model by Fernandez are clearly rejected by our quarterly data ($\chi^2(2) = 16.764$, $p = 0.0002$), and its RMSPE in the forecasting exercise is equal to 4.63721. The same applies for Litterman's model, with a RMSPE equal to 4.83834. In this case, the estimation of the SUTSE $(2, 0, 2)$ places in a intermediate position, with an error equal to 4.36451.

[3] The estimation of the annual relationship leads us to suspect that the partial adjustment representation could be a good parsimonious candidate in the $ADL(1,1)$ family. In effect, in the annual estimated relation reported above the restriction implied by a partial adjustment model is decidedly accepted by the data, $\chi^2(1) = 0.20414$, $p = 0.6514$.

6. CONCLUSIONS

The last few years have been characterised by substantial advances in temporal disaggregation techniques which have exploited different but interrelated aspects of the discipline, among these: a) the use of general dynamic models (and the related problem of initial conditions); b) the introduction of non-linear transformation of the data; c) the extensions to multivariate models for time disaggregations casted in the family of unobserved component models.

Once taken into account the availability of data and the limited samples statisticians have often to cope with, the use of genuine dynamic models, such as the $ADL(1,1)$, seems to be a valid starting point.

The use of standard models for time disaggregation often implies a reduction process that in effect has never been tested before the disaggregations start. Another implicit assumption of some schemes is the lack of co-integration between the series to be disaggregated and the set of related time series, a circumstance not likely in a time disaggregation context. Imposing these models on hundreds of disaggregations routinely carried out without checking their likelihood in the observed low-frequency data seems to be at least questionable.

There are models in the $ADL(1,1)$ family, such as the partial adjustment model, with a respectable pedigree in economic analysis but, again, imposing a priori such models in a world where the restrictions are not supported by the data will transpire to generate results which are an artifact of selecting, for example, a partial adjustment principle. Note that we do not make a critique of the principle, but of the way the principle is actually implemented.

A further step toward a model-based (and data-based) approach consists in the use of a multivariate model to temporal disaggregation, such as the one embedded in the class of SUTSE models. Since there is usually no behavioural relationship between the series to be disaggregated and the set of related variables, the SUTSE model could be a more appropriate framework than the traditional univariate regression approaches to represent and solve temporal disaggregation issues. The SUTSE approach is based on the use of related time series, but in this context the term 'related' has a different and more appealing meaning, as the implicit assumption is that the series to be disaggregated and the set of related series are simply affected by a similar environment. Consequently, the series should move together and measure similar things, though none of them necessarily causes the other in any statistical or economic sense.

In this framework, common component restrictions - such as common trends, cycles and seasonalities - can be tested and imposed quite naturally. This represents a further departure from traditional literature on time disaggregation. The SUTSE approach is flexible enough to allow for almost any kind of disaggregation problem (i.e. annual to quarterly, annual to monthly, quarterly to monthly ...) and to handle interpolation, distribution and extrapolation of both raw and seasonally adjusted time series. In this respect, another advantage of this methodology is that it can allow for simultaneous disaggregation and seasonal adjustment of the series, whereas NSIs usually undertake these two procedures separately (first disaggregation, then seasonal adjustment of the estimated raw series).

APPENDIX

Both interpolation and distribution find an optimal and general solution in the KF framework where they are treated as missing observation problems. The Kalman filtering and smoothing (KFS) allows for an adjustment in the dimension of the data and, in particular, in the system matrices $\mathbf{Z}_t$ and $\mathbf{G}_t$. Moreover, if for certain values of t no observations are available, the KFS can be simply run by skipping the updating equations, without affecting the validity of the prediction error decomposition.

While in the interpolation case the SSF introduced in Section 2 remains valid, in the distribution case the model and the observed timing intervals are different. By extending the discussion in Harvey (1989, p. 309) and Harvey and Chung (2000), we indicate with δ the model frequency and with $\delta_1^+, \ldots, \delta_N^+$ the frequencies at which the unobserved disaggregated flows $y_{1,t}, \ldots, y_{N,t}$ are observed. Model and observed frequencies are such that their ratios, denoted $\delta_i = \delta/\delta_i^+$, are integers for each i. Then, the aggregates are such that:

$$y_{i,t}^{\dagger} = \sum_{r=0}^{\delta_i - 1} y_{i,t-r}, \quad t = \delta_i, 2\delta_i, \ldots. \tag{7}$$

Let us suppose that $\mathbf{y}_t$ is generated by the LLT model (1)-(3), but that some elements of $\mathbf{y}_t$ are observed in temporally aggregated form. Following Harvey (1989, p. 313) and Harvey and Chung (2000), let us define the cumulator as:

$$y_{i,t-\delta_i+r}^{c} = \sum_{j=1}^{r} y_{i,t-\delta_i+j}, \quad r = 1, \ldots, \delta_i, \quad i = 1, \ldots, N \tag{8}$$

so that $y_{i,t}^c = y_{i,t}^{\dagger}$ for $t = \delta_i, 2\delta_i, \ldots$. The cumulator can also be written as:

$$\begin{aligned} \mathbf{y}_t^c &= \mathbf{C}_t \mathbf{y}_{t-1}^c + \boldsymbol{\mu}_t + \boldsymbol{\xi}_t = \\ &= \mathbf{C}_t \mathbf{y}_{t-1}^c + \boldsymbol{\mu}_{t-1} + \boldsymbol{\beta}_{t-1} + \boldsymbol{\eta}_t + \boldsymbol{\zeta}_t + \boldsymbol{\xi}_t, \end{aligned} \tag{9}$$

where $\mathbf{C}_t = diag\left(c_{1t}, c_{2t}, \ldots, c_{Nt}\right)$ and:

$$c_{it} = \begin{cases} 0 & t = 1, \delta_i + 1, 2\delta_i + 1, \ldots \\ 1 & \text{otherwise.} \end{cases} \tag{10}$$

Then, the state vector in equations (4)-(5) becomes $\boldsymbol{\alpha}_t = \left(\boldsymbol{\mu}_t', \boldsymbol{\beta}_t', \mathbf{y}_t^{c'}\right)'$ and the system matrices are given by

$$\mathbf{T}_t = \begin{pmatrix} \mathbf{I}_N & \mathbf{I}_N & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N & \mathbf{0} \\ \mathbf{I}_N & \mathbf{I}_N & \mathbf{C}_t \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} \boldsymbol{\Gamma}_\eta & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Gamma}_\zeta & \mathbf{0} \\ \boldsymbol{\Gamma}_\eta & \boldsymbol{\Gamma}_\zeta & \boldsymbol{\Gamma}_\xi \end{pmatrix}, \quad \mathbf{Z} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{I}_N \end{pmatrix}, \tag{11}$$

with $\mathbf{y}_0^c \equiv \mathbf{0}$, and $\mathbf{G} = \mathbf{0}$. Note that, though the form above is not the most parsimonious, it allows for the use of fast Kalman filtering and smoothing as suggested by Koopman and Durbin (2000). Further, the system matrix $\mathbf{T}$ is now time-varying as denoted by the subscript t.

The model above can account for cyclical components. In structural time series models the cyclical component is a linear function of sines and cosines. When a distribution problem occurs, the treatment of this component is similar to the treatment of the level. In a cycle plus noise model, the state vector is defined as $\boldsymbol{\alpha}_t = \left(\boldsymbol{\psi}_t', \boldsymbol{\psi}_t^{*\prime}, \mathbf{y}_t^{c\prime}\right)'$, and the system matrices become $\mathbf{T}_t = \begin{pmatrix} \mathbf{C}_{\boldsymbol{\Psi}} & \mathbf{S}_{\boldsymbol{\Psi}} & \mathbf{0} \\ -\mathbf{S}_{\boldsymbol{\Psi}} & \mathbf{C}_{\boldsymbol{\Psi}} & \mathbf{0} \\ \mathbf{C}_{\boldsymbol{\Psi}} & \mathbf{S}_{\boldsymbol{\Psi}} & \mathbf{C}_t \end{pmatrix}$ and $\mathbf{H}=\begin{pmatrix} \boldsymbol{\Gamma}_{\boldsymbol{\kappa}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Gamma}_{\boldsymbol{\kappa}} & \mathbf{0} \\ \boldsymbol{\Gamma}_{\boldsymbol{\kappa}} & \mathbf{0} & \boldsymbol{\Gamma}_{\xi} \end{pmatrix}$, where $\mathbf{C}_{\boldsymbol{\Psi}} = \rho\cos\lambda \cdot \mathbf{I}_N$, $\mathbf{S}_{\boldsymbol{\Psi}} = \rho\sin\lambda \cdot \mathbf{I}_N$, the frequency of the cyclical component is $0 \leq \lambda \leq \pi$ and the damping factor is $0 < \rho \leq 1$. The matrix $\boldsymbol{\Gamma}_{\boldsymbol{\kappa}}$ is obtained by the Cholesky decomposition of the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\kappa}}$ for the cyclical disturbances $\boldsymbol{\kappa}_t$.

Though SUTSE models do not necessarily require common components, common factor restrictions on level, slope, cycle and irregular can be introduced by imposing rank restrictions on the covariance matrices of the disturbances driving the components of interest. In this respect, it is possible to use a factorization of the covariance matrix $\boldsymbol{\Sigma}_{\mathbf{h}}$ such as $\boldsymbol{\Sigma}_{\mathbf{h}} = \boldsymbol{\Theta}_{\mathbf{h}}\mathbf{D}_{\mathbf{h}}^2\boldsymbol{\Theta}_{\mathbf{h}}'$, $\mathbf{h} = \boldsymbol{\eta}$, $\boldsymbol{\zeta}$, $\boldsymbol{\xi}$, $\boldsymbol{\kappa}$, where $\boldsymbol{\Theta}_{\mathbf{h}}$ is a lower triangular matrix with 1's along the principal diagonal and with $\mathbf{D}_{\mathbf{h}}$ diagonal (note that $\boldsymbol{\Gamma}_{\mathbf{h}} = \boldsymbol{\Theta}_{\mathbf{h}}\mathbf{D}_{\mathbf{h}}$). A model with r common restrictions for the $\mathbf{h}$-th component, $1 < r \leq N$, reduces the dimension of $\mathbf{D}_{\mathbf{h}}$ to $N - r$ and $\boldsymbol{\Theta}_{\mathbf{h}}$ to a full rank $(N \times (N - r))$ matrix. The limit case given by $r = N$ lets the $\mathbf{h}$-th component be deterministic since $\boldsymbol{\Sigma}_{\mathbf{h}} = \mathbf{0}$.

The LLT model in (1) can also be extended to deal with seasonal time series. In the dummy seasonal (DS) model the multivariate seasonal component $\boldsymbol{\gamma}_t$ is such that:

$$S(L)\,\boldsymbol{\gamma}_t = \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim NID(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\omega}})\,, \tag{12}$$

where $S(L) = 1 + L + ... + L^{\delta-1}$, with L the lag operator and $\boldsymbol{\omega}_t$ the vector of disturbances driving the seasonal pattern.

Hotta and Vasconcellos (1999) discuss the aggregation problem of the DS model for univariate time series. When a flow variable is aggregated across time, the form (12) does not change if δ is not a multiple of the aggregation period. If that is the case, the seasonality is unobserved and it is confused with the irregular component.

In a multivariate context, we consider the case in which seasonality is observed only for some elements of $\mathbf{y}_t^{\dagger}$. Then, the limited seasonal information has to be distributed among all the elements of $\mathbf{y}_t^{\dagger}$ by restricting in some way model (12). A first solution is the common seasonal model: if r is the number of elements of $\mathbf{y}_t^{\dagger}$ for which seasonality is not observed, both $\boldsymbol{\gamma}_t$ and $\boldsymbol{\omega}_t$ will become $((N - r) \times 1)$ vectors. Thus, for a simple seasonal plus irregular model, $\mathbf{y}_t^{\dagger}$ is represented by:

$$\mathbf{y}_t^{\dagger} = \boldsymbol{\Theta}_{\boldsymbol{\omega}}\boldsymbol{\gamma}_t + \overline{\boldsymbol{\gamma}}_t + \boldsymbol{\xi}_t, \tag{13}$$

where $\boldsymbol{\Theta}_{\boldsymbol{\omega}}$ follows the definition of the previous Section and $\overline{\boldsymbol{\gamma}}_t$ is a $(N \times 1)$ vector of fixed seasonal effects.

When $\overline{\gamma}_t$ is set to zero, model (13) leads to the stronger restriction of *similar seasonals*. In other words, the seasonal pattern is proportional among the elements of $\mathbf{y}_t^\dagger$. Finally, seasonality is *identical* when $r = N - 1$ and $\boldsymbol{\Theta}_{\boldsymbol{\omega}}$ is further restricted to a N vector of ones.

The strategy of model building adopted in Moauro and Savio (2005) consists in starting from the general multivariate LLT model, augmented in order to include seasonal and cyclical components if appropriate. Once a first estimation is run, some of the variances of the system in the general LLT model can be fixed by the system estimation itself because they are approximately equal to zero.

In this reduction strategy, a number of parametric tests for the form of the trend component and for the existence of common factors (see, e.g. Harvey 2001; Nyblom and Harvey 2000, 2001) are of great help. These tests are based on the innovations obtained from the model estimated at the lower frequency at which the series are observed. These tests start from estimating the nuisance parameters of the unrestricted model, then the Kalman filter and smoother with the appropriate variances set to zero in order to extract the innovations is run.

In the LL model, a test for the null hypothesis that $\boldsymbol{\Sigma}_{\boldsymbol{\eta}} = \mathbf{0}$ against the homogeneous alternative $\boldsymbol{\Sigma}_{\boldsymbol{\eta}} = q\boldsymbol{\Sigma}_{\boldsymbol{\xi}}$, $q > 0$, is given by $\eta(r, N) = tr\left(\mathbf{S}^{-1}\mathbf{C}\right)$, with $\mathbf{C} = n^{-2}\sum_{i=1}^{n}\left(\sum_{t=1}^{i}\boldsymbol{\nu}_t\right)\left(\sum_{t=1}^{i}\boldsymbol{\nu}_t\right)'$, $\mathbf{S} = n^{-1}\sum_{t=1}^{n}\boldsymbol{\nu}_t\boldsymbol{\nu}_t'$ and with $\boldsymbol{\nu}_t$ indicating the innovations. In the LLT with constant slope, the test above is indicated with $\eta'(r, N)$, with different rejection regions from $\eta(r, N)$. In the IRW model, the test for $\boldsymbol{\Sigma}_{\boldsymbol{\zeta}} = \mathbf{0}$ against the homogeneous alternative $\boldsymbol{\Sigma}_{\boldsymbol{\zeta}} = q\boldsymbol{\Sigma}_{\boldsymbol{\xi}}$ is given by $\zeta = tr\left(\mathbf{S}^{-1}\mathbf{T}\right)$, where $\mathbf{T} = n^{-2}\sum_{i=1}^{n}\left[\sum_{s=1}^{i}\sum_{r=1}^{s}\boldsymbol{\nu}_t\right]\left[\sum_{s=1}^{i}\sum_{r=1}^{s}\boldsymbol{\nu}_t'\right]$ (see Nyblom and Harvey 2001). The test against a stochastic slope with $\boldsymbol{\Sigma}_{\boldsymbol{\eta}} > \mathbf{0}$ is again given by $\eta(r, N)$. Furthermore, one might test for a specified number r of common levels/slopes, that is one might test for the rank of the relevant covariance disturbance. The test is constructed on the sum of the $N - r$ smallest eigenvalues of the matrices constructed on innovations.

REFERENCES

Abad, A. and E. Quilis (2005). Software to perform temporal disaggregation of economic time series: a Matlab libray and its Excel interface. Paper presented at the joint Eurostat-OECD *Workshop on Frontiers in Benchmarking Techniques and Their Application to Official Statistics*, 7-8 April, Luxembourg.

Bloem, A. M., R. J. Dippelsman and N. Ø. Mæhke (2001). *Quarterly National Accounts Manual - Concepts, Data Sources, and Compilation.* Washington DC: International Monetary Fund.

Bournay, J. and D. Ladiray (2005). Using benchmarking techniques to backcast French annual national accounts. Paper presented at the joint Eurostat-OECD *Workshop on Frontiers in Benchmarking Techniques and Their Application to Official Statistics*, 7-8 April, Luxembourg.

Caporin, M. and D. Sartore (2005). Methodological aspects of time series back-recalculation. Paper presented at the joint Eurostat-OECD *Workshop on Frontiers in Benchmarking Techniques and Their Application to Official Statistics*, 7-8 April, Luxembourg.

Chow, G. C. and A. Lin (1971). Best linear unbiased interpolation, distribution and extrapolation of time series by related series. *The Review of Economics and Statistics 53*, 372-5.

Denton, F. T. (1971). Adjustment of monthly or quarterly series to annual totals: an approach based on quadratic minimization. *Journal of the American Statistical Association 66*, 99-102.

Di Fonzo, T. (2003a). Temporal disaggregation using related series: log-transformation and dynamic extension. *Rivista Internazionale di Scienze Economiche e Commerciali 50*, 371-400.

———— (2003b). Constrained retropolation of high-frequency data using related series. A simple dynamic model approach. *Statistical Methods & Applications 12*, 109-19.

———— (2003c). Temporal disaggregation of economic time series: towards a dynamic extension. *Working Papers and Studies.* Luxembourg: European Communities.

Doornik, J. A. (2001). *Ox 3.0 - An Object-Oriented Matrix Programming Language.* London: Timberlake Consultants Ltd.

Durbin J. and B. Quenneville (1997). Benchmarking by state space models. *International Statistical Review 65*, 23-48.

Eurostat (1996). *European System of Accounts, ESA 1995.* Luxembourg: European Communities.

———— (1999). *Handbook on Quarterly National Accounts.* Luxembourg: European Communities.

Fernández, R. B. (1981). A methodological note on the estimation of time series. *The Review of Economics and Statistics 63*, 471-6.

Fernández, F. J. and A. C. Harvey (1990). Seemingly unrelated time series equations and a test for homogeneity. *Journal of Business & Economic Statistics 8*, 71-81.

Gómez, V. (2000). Estimating missing observations in ARIMA models with the Kalman filter. Forthcoming in *Annali di Statistica.* Rome: Istat.

Gregoir, S. (1995). Propositions pour une désagrégation temporelle basée sur des modèles dynamiques simples. *Mimeograph.* Paris: Insee (mimeo).

Gudmundsson, G. (1999). Disaggregation of annual flow data with multiplicative trends. *Journal of Forecasting 18*, 33-37.

———(2005). Some time series considerations in disaggregation. Paper presented at the joint Eurostat-OECD *Workshop on Frontiers in Benchmarking Techniques and Their Application to Official Statistics*, 7-8 April, Luxembourg.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

——— (1996). Intervention analysis with control groups. *International Statistical Review 64*, 313-28.

——— (2001). Testing in unobserved components models. *Journal of Forecasting 20*, 1-19.

Harvey, A. C. and C. Chung (2000). Estimating the underlying change in unemployment in the UK. *Journal of the Royal Statistical Society A 163*, 303-39.

Harvey, A. C. and A. Jaeger (1993). Detrending, stylized facts and the business cycle. *Journal of Applied Econometrics 8*, 231-47.

Harvey, A. C. and S. J. Koopman (1997). Multivariate structural time series models (with comments). In *System Dynamics in Economic and Financial Models*, C. Heij, J. M. Shumacher, B. Hanzon and C. Praagman (eds.). Chichester: John Wiley & Sons Ltd, 269-98.

Harvey, A. C. and R. G. Pierce (1984), Estimating missing observations in economic time series. *Journal of the American Statistical Association 79*, 125-31.

Hendry, D. F. (1995). *Dynamic Econometrics*. Oxford: Oxford University Press.

Hotta, L. K. and K. L. Vasconcellos (1999). Aggregation and disaggregation of structural time series models. *Journal of Time Series Analysis 20*, 155-71.

King, R. G., C. I. Plosser, J. H. Stock and M. W. Watson (1991). Stochastic trends and economic fluctuations. *The American Economic Review 81*, 819-40.

Koopman, S. J. (1997). Exact initial Kalman filtering and smoothing for non-stationary time series models. *Journal of the American Statistical Association 92*, 1630-8.

Koopman, S. J., and J. Durbin (2000). Fast filtering and smoothing for multivariate state space models. *Journal of Time Series Analysis 21*, 281-96.

——— (2003). Filtering and smoothing of state vector for diffuse state-space models. *Journal of Time Series Analysis 24*, 85-98.

Koopman, S. J., N. Shephard and J. A. Doornik (1999). Statistical algorithms for models in state space using SsfPack 2.2. *The Econometrics Journal 2*, 107-60.

———(2002). SsfPack 3.0beta02: statistical algorithms for models in state space. February (see www.ssfpack.com).

Leila, H. and A. Trabelsi (2005). A polynomial method for temporal disaggregation. Paper presented at the joint Eurostat-OECD *Workshop on Frontiers in Benchmarking Techniques and Their Application to Official Statistics*, 7-8 April, Luxembourg.

Litterman, R. B. (1983). A random walk, markov model for distribution of time series. *Journal of Business & Economic Statistics 1*, 169-73.

Mitchell, J., R. J. Smith, M. R. Weale, S. Wright and E.L Salazar (2005). An indicator of monthly GDP and an early estimate of quarterly GDP growth. *The Economic Journal F 115*, F108-29.

Mitchell, J. and M. Weale (2005b). A monthly indicator of GDP for the Euro area: an application of the regression approach to intepolation. Paper presented at the joint Eurostat-OECD *Workshop on Frontiers in Benchmarking Techniques and Their Application to Official Statistics*, 7-8 April, Luxembourg.

Moauro, F. and T. Proietti (2005). Temporal disaggregation and seasonal adjustment. Paper presented at the joint Eurostat-OECD *Workshop on Frontiers in Benchmarking Techniques and Their Application to Official Statistics*, 7-8 April, Luxembourg.

Moauro, F. and G. Savio (2005). Temporal disaggregation using multivariate structural time series models. *The Econometrics Journal 8*, 214-34.

Nyblom, J. and A. C. Harvey (2000). Tests of common stochastic trends. *Econometric Theory 16*, 176-99.

——— (2001). Testing against smooth stochastic trends. *Journal of Applied Econometrics 16*, 415-29.

OECD (2002). *Main Economic Indicators.* Paris: Organisation for Cooperation and Economic Development, July.

Palate, J. (2005). Reusable components for benchmarking using Kalman filters. Paper presented at the joint Eurostat-OECD *Workshop on Frontiers in Benchmarking Techniques and Their Application to Official Statistics*, 7-8 April, Luxembourg.

Proietti, T. (1999). Distribution and interpolation revisited: a structural approach. *Statistica 58*, 411-32.

——— (2004a). On the estimation of nonlinearly aggregated mixed models. *Working Papers and Studies.* Luxembourg: European Communities (revised version forthcoming in *Journal of Computational and Graphical Statistics*).

——— (2004b). Temporal disaggregation by state space methods: dynamic regression methods revisited. *Working Papers and Studies.* Luxembourg: European Communities.

Quilis, E. (2005). Benchmarking techniques in the Spanish quarterly national accounts. Paper presented at the joint Eurostat-OECD *Workshop on Frontiers in Benchmarking Techniques and Their Application to Official Statistics*, 7-8 April, Luxembourg.

Santos Silva, J. M. C. (2005). Monthly estimates for the Portuguese unemployment rate. Paper presented at the joint Eurostat-OECD *Workshop on Frontiers in Benchmarking Techniques and Their Application to Official Statistics*, 7-8 April, Luxembourg.

Santos Silva, J. M. C. and F. N. Cardoso (2001). The Chow-Lin method using dynamic models. *Economic Modelling 18*, 269-80.

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

Stram, D. O. and W. W. S. Wei (1986). A methodological note on the disaggregation of time series totals. *Journal of Time Series Analysis 7*, 293-302.

Wei, W. W. S. and D. O. Stram (1990). Disaggregation of time series models. *Journal of the Royal Statistical Society*, *Series B 52*, 453-67.

Table 1. Results of time disaggregations; methods with lower and higher RMSPE.

COUNTRY	Can	Mex	Usa	Aus*	Jap	Kor	Fra	Deu	Ita	Nld	Esp	Gbr
					A. Quarterly industrial production - Monthly deliveries							
Lower	**STS**	**STS**	C-L	FNZ	**STS**	STR	DTU	**STS**	**STS**	STR	**STS**	STR
Higher	DTU	DTM	DTM	LTM	DTM	DTM	DTM	LTM	S-W	DTM	S-W	DTM
					B. Annual GDP - Quarterly industrial production							
Lower	**STS**	**STS**	**STS**	**STS**	STR	**STS**	**STS**	**STS**	**STS**	**STS**	LTM	LTM
Higher	S-W	S-W	C-L	C-L	DTM	DTU	C-L	DTM	DTM	C-L	C-L	C-L
					C. Quarterly consumer prices - Monthly producer prices							
Lower	S-W	S-W	LTM	**STS**	**STS**	**STS**	S-W	**STS**	**STS**	**STS**	DTU	STR
Higher	C-L	C-L	C-L	C-L	DTM	C-L	C-L	C-L	C-L	DTM	C-L	C-L
					D. Annual private consumption expenditures- Quarterly GDP							
Lower	**STS**	**STS**	**STS**	**STS**	DTM	STR	DTM	FNZ	STR	**STS**	STR	**STS**
Higher	DTM	LTM	C-L	C-L	S-W	LTM	DTU	LTM	FNZ	DTM	LTM	DTU
					E. Annual GDP deflator - Quarterly consumer prices							
Lower	LTM	STR	**STS**	**STS**	S-W	LTM	STR	STR	**STS**	**STS**	S-W	**STS**
Higher	DTU	C-L	DTM	DTM	DTM	STR	DTM	DTM	DTU	DTM	C-L	C-L
					F. Quarterly broad money supply - Monthly narrow money supply							
Lower	**STS**	**STS**	**STS**	**STS**	**STS**	**STS**	**STS**	**STS**	**STS**	**STS**	DTU	**STS**
Higher	DTM	C-L	DTM	DTM	DTM	C-L	DTM	DTM	DTM	DTM	C-L	DTM
					G. Annual short-term interest rates - Monthly long-term interest rates							
Lower	LTM	**STS**	LTM	LTM	**STS**	C-L	LTM	LTM	DTU	LTM	LTM	**STS**
Higher	DTM	DTU	DTM	DTM	LTM	S-W	DTM	C-L	S-W	S-W	S-W	LTM
					H. Annual imports c.i.f. - Quarterly imports f.o.b.							
Lower	**STS**	**STS**	DTM	**STS**	**STS**	DTM	DTM	**STS**	DTM	**STS**	DTM	DTM
Higher	S-W	S-W	STR	S-W	S-W	DTU	S-W	STR	S-W	DTU	DTU	S-W

Note: Legenda for countries - Can=Canada, Mex=Mexico, USA=The United States of America, Aus=Australia, Jap=Japan, Kor=North Korea, Fra=France, Deu=Germany, Ita=Italy, Nld=Netherlands, Esp=Spain, Gbr=Great Britain.
Legenda for methods - STS= SUTSE (in bold), C-L= Chow-Lin, FNZ= Fernández, LTM= Litterman, DTM= Denton with related series, STR= Structural, DTU= Denton without related series, S-W= Stram-Wei.
* In cases A and C annual-quarterly exercise instead of quarterly-monthly.

Table 2. Synthesis of results: percentage of cases where the method has the lowest/highest RMSPE

METHODS	SUTSE	Chow-Lin	Fernández	Litterman	Denton	Structural	Denton univ.	Stram-Wei
% lowest RMSPE	54.2	2.1	2.1	12.5	8.3	11.4	4.2	5.2
% higher RMSPE	0.0	27.1	0.0	8.3	34.4	5.2	8.3	16.7

Note: The Table shows the proportion of times that each method has the lowest/highest RMSPE. If the RMSPE were all equal, the number of times one particular model has the lowest/highest RMSPE should be distributed as a Binomial $(96, 0.125)$. Using the Normal approximation to the Binomial, one can derive a 95% confidence interval $p \pm 0.066$ for each proportion p reported in the Table.

Table 3. Synthesis of results and gains/losses in terms of RMSPEs.

METHODS	SUTSE	Chow-Lin	Fernández	Litterman	Denton	Structural	Denton univ.	Stram-Wei
SUTSE	-	147.2	123.1	114.3	158.8	135.5	154.1	155.4
Chow-Lin	91.7	-	83.6	77.7	107.9	92.1	104.7	105.6
Fernández	86.5	26.0	-	92.9	129.0	110.1	125.2	126.2
Litterman	81.3	31.3	41.7	-	138.8	118.5	134.7	135.9
Denton	88.5	52.1	71.9	71.9	-	85.4	97.1	97.9
Structural	78.1	36.5	45.8	54.2	33.3	-	113.7	114.7
Denton univ.	82.3	41.7	57.3	62.5	36.5	76.0	-	100.8
Stram-Wei	83.3	43.8	59.4	65.6	36.5	79.2	59.4	-

Note: The Table shows, in the lower part, the percentage of successes of the method in column over that in row, in the upper part the geometric average of the ratios of RMSPE of the method in column over the method in row. A geometric mean is used for its reciprocity properties.

Table 4. Results of time disaggregations, $N \geq 2$: RMSPEs.

	Canada				USA			
	Model 1		Model 2	Model 3	Model 1		Model 2	Model 3
	Case A	Case B			Case A	Case B		
SUTSE	0.345 (1)	0.694 (1)	0.582 (1)	**0.329 (1)**	0.314 (1)	**0.533 (1)**	0.731 (1)	0.639 (1)
Chow-Lin	0.393 (4)	0.993 (4)	1.012 (4)	**0.341 (2)**	0.465 (4)	**0.796 (3)**	0.970 (4)	0.842 (4)
Fernández	0.372 (2)	0.809 (2)	0.758 (3)	**0.433 (3)**	0.434 (3)	0.844 (4)	0.945 (3)	**0.761 (3)**
Litterman	0.384 (3)	0.869 (3)	**0.656 (2)**	0.785 (4)	0.333 (2)	**0.649 (2)**	0.812 (2)	0.737 (2)

Note: In Model 1, Case A, $N = 2$ (Consumption and GDP), and RMSPEs are computed in-sample; in Model 1, Case B, $N = 2$ (Consumption and GDP), and RMSPEs are computed out-of-sample; in Model 2, $N = 3$ (Consumption, GDP and Investments), and RMSPEs are computed out-of-sample; in Model 3, $N = 6$ (Consumption, GDP, Investments, Money supply, Interest rate and Inflation rate), and RMSPEs are computed out of sample. In parentheses is reported the ranking of the method in the experiment, in bold the best performing method.