WORKING PAPERS AND STUDIES

ISSN 1725-4825

2005 EDITION

Temporal disaggregation and seasonal adjustment

COPYRIGHT

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

ISSN 1725-4825
ISBN 92-79-01308-4
Cat. No. KS-DT-05-011-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 92-79-01308-4

Temporal Disaggregation and Seasonal Adjustment

Tommaso Proietti*
Università di Udine

Filippo Moauro†
ISTAT

31st March 2005

Abstract

The paper discusses the main issues arising in the construction of quarterly national accounts estimates, adjusted for seasonality and calendar effects, obtained by disaggregating the original annual actual measurements using related monthly indicators.

It proposes and implements an approach that hinges upon the estimation of a bivariate basic structural time series model at the monthly frequency, accounting for the presence of seasonality and calendar components. The monthly frequency enables more efficient estimation of calendar component.

The main virtue of this approach is to enable adjustment and temporal disaggregation to be carried out simultaneously. The proposed methodology also complies with the recommendations made by the Eurostat - European Central Bank task force on the seasonal adjustment of Quarterly National accounts.

Keywords: Structural Time Series Models; Calendar effects; Kalman filter and smoother; Temporal Aggregation.

*Dipartimento di Scienze Statistiche, Università di Udine Italy. E-mail:proietti@dss.uniud.it.

†ISTAT, Direzione Centrale Contabilità Nazionale, Via Agostino Depretis 74/B, 00184 Roma, Italy. E-mail:moauro@istat.it. The views expressed in this paper are those of the authors and do not necessarily reflect the positions or policies of Istat.

1 Introduction

This paper is concerned with the temporal disaggregation of economic flow series that are available only at the annual frequency of observations; the resulting quarterly or monthly estimates incorporate the information available from related indicators at the higher frequency, but as the indicators are affected by seasonal and calendar variation, there arise the problem of adjusting the estimates for those effects.

Seasonality and calendar components explain a relevant part of the fluctuations of economic aggregates. While the former refers to the intra-year movements in economic activity caused by various factors, among which climatic and institutional ones are prominent, calendar effects result from essentially three sources (see Cleveland and Devlin, 1980, Bell and Hillmer, 1983): i) weekly periodicity: the level of economic activity depends on the day of the week. The aggregation of weekly seasonal effects into a monthly series is referred to as trading day (TD) or working day (WD) effects, according as to whether the series refers to sales or production. ii) moving festivals, such as Easter, which change their position in the calendar from year to year. iii) the different length of the month or quarter: once TD/WD and seasonal effects are accounted for, what residues is the leap year effect.

Providing quarterly national accounts estimates corrected for seasonality and calendar components satisfies a well established information need for both business cycle and structural analyses; this is officially recognised in Eurostat's Handbook of Quarterly National Accounts (Eurostat, 1999). A task force established by Eurostat and the European Central Bank (Eurostat, 2002) has also set forth some guidelines for calendar adjustment, some of which motivate this contribution: in particular, the use of regression methods is recommended in the place of proportional adjustment, with the regressors constructed so as to take into account the country specific holidays; when available, adjustment should be performed on monthly series, as calendar effects are more easily identified at that frequency.

The Italian Statistical Institute, Istat, has started trading day adjustment of quarterly national accounts in June 2003 and publishes seasonally adjusted and trading day corrected series since then. See Istat (2003) for a description of the methodology. The French methodology is documented in Insee (2004). Essentially, the current practice involves at least three operations: a separate seasonal and calendar adjustment of the indicator series, and two temporal disaggregation of the annual aggregate using the two versions of the indicator. The disaggregation method adopted is based on the technique proposed by Chow and Lin (1971).

We argue that this is unnecessarily complicated; indeed, the main aim of the paper is to show that all these operations can easily brought under the same umbrella. Within the unifying framework represented by the estimation of a multivariate structural time series model formulated at the higher time interval, seasonal adjustment of the indicators and the correction for calendar variation are carried out in one step. The multivariate setup also provides a more consistent framework for using the information on related series.

The plan of the paper is the following: the next section introduces the disaggregated basic structural model with regression effects which lay at the basis of our approach. Section 3 discusses the effects of temporal aggregation on the seasonal component and considers the consequences on modelling and data dissemination policies. The modelling of the calendar component is considered in section 4. Section 5 illustrates the statistical treatment of the model, whereas section 6 presents a real life example.

2 The Bivariate Basic Structural Model

The basic structural model (BSM henceforth), proposed by Harvey and Todd (1983) for univariate time series and extended by Harvey (1989) to the multivariate case, postulates an additive decomposition of the series into a trend, a seasonal and an irregular component. Its name stems from the fact that it provides a satisfactory fit to a wide range of seasonal time series, thereby playing a role analogous to the Airline model in an unobserved components framework; see also Maravall (1985).

Without loss of generality we focus on a bivariate series $y_t = [y_{1t}, y_{2t}]'$, where t is time in months; in the sequel y_{1t} will represent the indicator series, whereas y_{2t} is subject to temporal aggregation, being observed only at the annual frequency.

The BSM is such that each of the component series has the following representation:

$$y_{it} = \mu_{it} + \gamma_{it} + x_{it}'\delta_i + \epsilon_{it}, \quad i = 1, 2; \quad t = 1, \ldots, n, \quad \epsilon_{it} \sim \mathrm{NID}(0, \sigma^2_{i\epsilon})$$

where the series specific trend, μ_{it}, is a local linear component:

$$\begin{array}{rcccccll} \mu_{i,t+1} & = & \mu_{it} & + & \beta_{it} & + & \eta_{it}, & \eta_{it} \sim \mathrm{N}(0, \sigma^2_{i\eta}) \\ \beta_{i,t+1} & = & & & \beta_{it} & + & \zeta_{it}, & \zeta_{it} \sim \mathrm{N}(0, \sigma^2_{i\zeta}) \end{array} \tag{1}$$

The disturbances η_{it} and ζ_{it} are mutually and serially uncorrelated, but are contemporaneously correlated with the disturbances η_{jt} and ζ_{jt}, respectively, affecting the same equation of the trend for the other series.

The seasonal component, γ_{it}, arises from the combination of six stochastic cycles defined at the seasonal frequencies $\lambda_j = 2\pi j/s$, $j = 1, \ldots, 6$, λ_1 representing the fundamental frequency (corresponding to a period of 12 monthly observations) and the remaining being the five harmonics (corresponding to periods of 6 months, i.e. two cycles in a year, 4 months, i.e. three cycles in a year, 3 months, i.e. four cycles in a year, 2.4, i.e. five cycles in a year, and 2 months):

$$\gamma_{it} = \sum_{j=1}^{6} \gamma_{ijt}, \qquad \begin{bmatrix} \gamma_{ij,t+1} \\ \gamma^*_{ij,t+1} \end{bmatrix} = \begin{bmatrix} \cos\lambda_j & \sin\lambda_j \\ -\sin\lambda_j & \cos\lambda_j \end{bmatrix} \begin{bmatrix} \gamma_{ij,t} \\ \gamma^*_{ij,t} \end{bmatrix} + \begin{bmatrix} \omega_{ij,t} \\ \omega^*_{ij,t} \end{bmatrix}, j = 1, \ldots, 5, \tag{2}$$

and $\gamma_{i6,t+1} = -\gamma_{i6t} + \omega_{i6t}$. For the i-th series, the disturbances ω_{ijt} and ω^*_{ijt} are normally and independently distributed with common variance $\sigma^2_{i\omega}$ for $j = 1, \ldots, 5$, whereas $\text{Var}(\omega_{ijt}) = \text{Var}(\omega^*_{ijt}) = 0.5\sigma^2_{i\omega}$ (see Proietti, 2000, for further details).

The symbol ϵ_{it} denotes the irregular component, which is taken to be series specific, in that it is also uncorrelated with ϵ_{jt}. This restriction, which is not critical and can be removed at will, attributes this source of variation to series specific measurement error.

The vector x_t is a $K \times 1$ vector of regressors accounting for calendar effects, which will be specified in section 4 and δ_i is a vector of unknown regression coefficients for the $i-th$ series.

According to the model specification, the indicator variable y_{1t} and the national account flow y_{2t} form a *Seemingly Unrelated Time Series Equations* system (Harvey, 1989). There is no cause and effect relationship among them, but they are subject to the same underlying economic environment. In particular, the first series can be viewed as a partial, possibly noisier, measurement of the same underlying phenomenon.

3 The effects of temporal aggregation on the seasonal component

The flow series y_{2t} is not observed; the actual observations pertain to the yearly series

$$Y_{2\tau} = \sum_{k=0}^{11} y_{2,12\tau-k}, \quad \tau = 1, 2, \ldots, [n/12],$$

where $[a/b]$ denotes the integer part of a/b.

As the sum of 12 consecutive values of γ_{2t} is a zero mean invertible moving average process of order equal to 10 months, it immediately follows that the aggregation of γ_{2t}, $\sum_{k=0}^{11} \gamma_{2,12\tau-k}$, yields a pure white noise, which, without the aid of external information on the indicator series, would be indistinguishable from the aggregation of the series specific measurement error, that is $\sum_{k=0}^{11} \epsilon_{2,12\tau-k}$.

As the seasonal disturbances in y_{2t} are contemporaneously correlated with those driving the seasonal component in the indicator, in principle the bivariate model could identify the component resulting from aggregation of γ_{2t}, as the white noise source of variation that is independent of $\sum_{k=0}^{11} \epsilon_{2,12\tau-k}$ and which is due to the interaction with the disturbances ω_{1jt}'s.

However, in the situations typically occurring in practice, where seasonality has a slow and weak evolution and sample sizes are small, this source of variation is negligible to an extent that trying to disentangle it from the measurement error would be asking too much of the available data.

One possibility is to assume it away, as will soon be argued. An alternative feasible strategy is to borrow the seasonal pattern from the indicator as suggested by Moauro and Savio (forthcoming). This is also what is prevailing in current practice adopted by statistical national offices, which produce disaggregate estimates according to the scheme: $\hat{y}_{2t} = b_0 + b_1 y_{1t} + e_t$, where b_0 and b_1 are the generalised least squares estimates of the regression coefficients based on the Chow and Lin (1971) model, and e_t is the distribution residual.

The estimates $\hat{y}_{2t}$ are referred to as "raw"; trading day adjusted series are produced by the same scheme, in which y_{1t} is replaced by a corrected series. The assumption underlying these operations is that the seasonal component in the national accounts aggregate is proportional to that in the indicator, the factor of proportionality being the same b_1 that relates the annual series.

The conditions under which the seasonal behavior of the aggregate series can be borrowed from y_{1t} via standard generalised regression are indeed rather stringent. Not only common seasonal features are required but also a restricted covariance structure in the nonseasonal component.

Denoting by $z_t = [z_{1t}, z_{2t}]'$ the nonseasonal component, we rewrite the disaggregate bivariate model as $y_{it} = z_{it} + \gamma_{it}, \ i = 1, 2$. Assume now that $\gamma_{2t} = \lambda\gamma_{1t}$ (proportional seasonal component) and that the nonseasonal component follows a seemingly unrelated system of equations:

$$z_t = \frac{\theta(L)}{\phi(L)}\kappa_t, \quad \kappa_t \sim \text{NID}(0, \Sigma_\kappa), \quad \Sigma_\kappa = \begin{pmatrix} \sigma^2_{1\kappa} & \sigma_{12,\kappa} \\ \sigma_{12,\kappa} & \sigma^2_{2\kappa} \end{pmatrix}, \tag{3}$$

where $\theta(L)$ and $\phi(L)$ are suitable scalar lag polynomials.

If z_t results from the sum of several orthogonal components, $z_t = \sum_j \frac{\theta_j(L)}{\phi_j(L)}\kappa_{jt}, \ \kappa_{jt} \sim \text{NID}(0, \Sigma_{j\kappa})$, such as $z_t = \mu_t + \epsilon_t$, then (3) requires *homogeneity* (see Harvey, 1989, sec. 8.3), which amounts to $\Sigma_{j\kappa} = q_j\Sigma_\kappa$, where Σ_κ is a constant matrix and q_j is a proportionality factor which equals 1 for a selected component.

If further $\sigma_{12,\kappa} = \lambda\sigma^2_{2\kappa}$, then it is possible to write

$$y_{2t} = \lambda y_{1t} + z^*_{1t}, z^*_{1t} = \frac{\theta(L)}{\phi(L)}\kappa^*_{1t}, \kappa^*_{1t} \sim \text{NID}(0, \sigma^2_{1\kappa} - \lambda^2\sigma^2_{2\kappa})$$

and thus we can safely attribute the portion λ of the seasonality in the indicator to the aggregate series.

The restrictions under considerations are testable, say by the LR principle, although the properties of such a test are yet to be investigated. The test might become integral part of the modelling strategy and an example is shown in a next section.

We believe that the strategy of giving up the idea of estimating the seasonality in y_{2t} altogether is more neutral. Thus, in the sequel we shall assume that

$$\sum_{k=0}^{11} \gamma_{2,12\tau-k} = 0, \tag{4}$$

en lieu of $\mathrm{E}\left(\sum_{k=0}^{11}\gamma_{2,12\tau-k}\right)=0$. Notice that (4) strictly holds when seasonality is deterministic (that is $\sigma^2_{2\omega}=0$).

In the light of the previous discussion, the "raw" series are more a statistical artifact, than a useful addition to the supply of official economic statistics. If the primary interest of the investigation were the seasonal fluctuations on their own, it is more sensible and informative to study the monthly indicators from the outset.

A final important point arises as a consequence of (4). The simplification preserves the accounting relationship that the sum of the disaggregated series over 12 months adds up exactly to the annual total, which would not hold otherwise. As for the series corrected for the calendar component, this would sum up to the annual estimate with the calendar effects removed.

In conclusion the proposed solution has the additional merit of complying with the recommendation of the Eurostat/ECB task force concerning time consistency with annual data (recommendation 3.c):

> *Time consistency of adjusted data should be maintained for practical reasons. The reference aggregates should be the annual total of quarterly raw data for seasonally adjusted data and annual total of quarterly data corrected for trading day effects for seasonally and trading day adjusted data. Exceptions from the time consistency may be acceptable if the seasonality is rapidly changing.*

In situations were seasonality is not rapidly changing, our assumption seems plausible.

4 Calendar components

Calendar effects have been introduced as regression effects in the model equation for y_{it}. Three sets of regressors are defined to account for each of the three sources of variation mentioned in the introduction.

Trading day (working day) effects occur when the level of activity varies with the day of the week, eg. it is lower on Saturdays and Sundays.

Letting D_{jt} denote the number of days of type j, $j=1,\dots,7$, occurring in month t and assuming that the effect of a particular day is constant, the differential trading day effect for series i is given by:

$$TD_{it}=\sum_{j=1}^{6}\delta_{ij}\left(D_{jt}-D_{7t}\right)$$

The regressors are the differential number of days of type j, $j=1\dots,6$, compared to the number of Sundays, to which type 7 is conventionally assigned. The Sunday effect on the i-th series is then obtained as $-\sum_{j=1}^{6}\delta_{ij}$. This expedient ensures that TD effect is zero over a period corresponding to multiples of the weekly cycle.

The regressors are then corrected to take into account the national calendars: for instance, if the Christmas falls on a Monday, one unit should be deducted from D_{1t} and reassigned to D_{7t} if for that particular application a holiday can be assimilated to a Sunday. This type of correction is recommended by Eurostat and is adopted in this paper, giving:

$$TD_{it} = \sum_{j=1}^{6} \delta^*_{ij} \left(D^*_{jt} - D^*_{7t} \right).$$

It is often found that the effect of the working day from Mondays to Friday is not significantly different and that it helps to avoid collinearity among the regressors to assume that $\delta^*_{ij} = \delta^*_i$ for $j = 1, \dots 5$; in such case a single regressor can validly be employed, writing

$$TD_{it} = \delta^*_i D^*_t, \quad D^*_t = \sum_{j=1}^{5} D^*_{jt} - \frac{5}{2} D^*_{6t}.$$

The only moving festival in the Italian case concerns Easter; its effect is modelled as $E_t = \delta h_t$ where h_t is the proportion of 7 days before Easter that fall in month t. Subtracting 1/12 from h_t yields a regressor $h^*_t = h_t - 1/12$ which has a zero mean over the calendar year.

Finally, the length of month (LOM) regressor results from subtracting from the number of days in each month, $\sum_j D_{jt}$, its long run average, which is 365.25/12.

What are the consequences of temporal aggregation from the monthly frequency to the annual one? The holiday effect becomes constant ($h_\tau = 1$, $h^*_t = 0$), whereas the LOM regressor takes the value 3/4 in leap years and -1/4 in normal years, describing a four year cycle, which is an identifiable though not necessarily significant effect.

As shown by Cleveland and Devlin, the presence of trading day effects in a monthly time series induces a peak in the spectrum at the frequency $0.348 \times 2\pi$ in radians, and a secondary peak at $0.432 \times 2\pi$. For yearly data the relevant frequencies are $0.179 \times 2\pi$ and $0.357 \times 2\pi$, corresponding to a period of 5.6 years and 2.80 years, respectively. In conclusion, the presence of a calendar component in yearly data produces peaks at the frequencies 0.358π (TD), 0.5π (leap year), 0.714π (TD) and π (leap year).

In conclusion, the calendar component has detectable effects on an annually aggregated time series; thus, one possibility is to let the vector δ^*_2 measuring the corresponding effects unrestricted. An alternative parsimonious strategy is to assume that $\delta^*_2 = \kappa\delta^*_1$ for a scalar κ, which amounts to assume that the calendar effects on the second series are proportional to those affecting the first. This would require the estimation of a single coefficient. The difference with the unrestricted approach is that the disaggregated time series including the calendar component would feature the Easter effect, which would otherwise be absent.

5 Statistical treatment

The state space methodology provides the necessary inferences, starting from the estimation of unknown parameters, such as the variances of the disturbances driving the components, the regression coefficients, the estimation of the disaggregated values y_{2t} and the assessment of their reliability. Moreover, diagnostic checking can be carried out on the model's innovations, so as to detect and possibly take the corrective actions against any departure from the stated assumptions.

As a first step, the monthly bivariate model, with temporal aggregation concerning solely the second variable, is cast in the state space form using an approach due to Harvey (1989, sec. 6.3, 2001), which translates the aggregation problem into a missing value problem. According to this approach, the following cumulator variable is defined for the second variable:

$$y_{2t}^c = \psi_t y_{2,t-1}^c + y_{2t}, \quad \psi_t = \begin{cases} 0, & t = 12(\tau-1)+1, \tau = 1, \ldots, [n/12] \\ 1, & \text{otherwise} \end{cases} \tag{5}$$

In the case of monthly flows whose annual total is observed,

$$\begin{array}{cccc} y_{21}^c = y_{21}, & y_{22}^c = y_{21} + y_{22}, & \cdots & y_{2,12}^c = y_{21} + \ldots + y_{2,12}, \\ y_{2,13}^c = y_{2,13}, & y_{2,14}^c = y_{2,13} + y_{2,14}, & \cdots & y_{2,24}^c = y_{2,13} + \ldots + y_{2,24}, \\ \vdots & \vdots & \cdots & \vdots \end{array}$$

Only a systematic sample of every 12-th value of y_{2t}^c process is observed, $y_{2,12\tau}^c, \tau = 1, \ldots [n/12]$, so that all the remaining values are missing.

The cumulator is included in the state vector and the state space representation if formed. The associated algorithms, and in particular the Kalman filter and smoother are used for likelihood evaluation, and estimation of the missing observations and thus of the disaggregated values of the series. The smoothed estimates of the monthly series are then aggregated to the quarterly frequency. All the computations concerning the illustrations presented in the next section were carried out in Ox[1]. The statistical treatment of the model was performed using the augmented Kalman filter and smoother due to de Jong (1991, see also de Jong and Chu-Chun-Lin, 1994), suitably modified to take into account the presence of missing values, which is accomplished by skipping certain updating operations. More technical details, which we purposively omit for brevity, and computer programmes are available from the authors.

6 Illustrations

This section presents two illustrations, both based on Italian time series released by Istat, dealing with the problem of disaggregating the annual production resulting from the

[1]Ox is a matrix programming language developed by J.A. Doornik (2001).

National Accounts (NA), given the availability of the monthly industrial production (IP) index. The first concerns the estimation of a bivariate basic structural model under the hypothesis of null seasonality for the annual flow; the latter implements the estimation of a trend homogeneous system with proportional restriction on the seasonal component and test for the attribution of seasonality on the annual flow.

6.1 BSM with seasonal adjustment and temporal disaggregation

The application is referred to the electrical and optical equipment industry (subsection DL of the Nace Rev.1 economic activity classification), with the NA annual aggregate measured at constant price 1995 = 100, covering the years from 1977 to 2003; on the other hand, the monthly index has base year 2000 = 100, is seasonal unadjusted, with sample period from January 1977 to December 2003. The original series are plotted in figure 1.

The first step of the analysis was to estimate the bivariate basic structural model under temporal aggregation. Maximum likelihood estimation produced the following parameter estimates:

$$\begin{array}{lll} \hat{\sigma}_{1\eta} = 1.334, & \hat{\sigma}_{2\eta} = 6.137, & \hat{\rho}_{\eta} = 0.737, \\ \hat{\sigma}_{1\epsilon} = 1.916, & \hat{\sigma}_{2\epsilon} = 0.035, & \\ \hat{\sigma}_{1\omega} = 0.224, & & \end{array}$$

with $\sigma_{1\zeta} = \sigma_{2\zeta} = 0$, where the suffix 1 denotes the industrial production index and 2 the output, whereas ρ_η represents the correlation between η_{1t} and η_{2t}; the maximised log-likelihood is equal to -1056.891.

These results show that for both the series the trend features a constant slope, since its disturbance variance is zero; as a result the trend is a bivariate random walk with a constant drift, with positively, but not perfectly, correlated disturbances (ρ_η is estimated equal to 0.737). This suggests that the series are not cointegrated.

The non-zero value for the seasonal variance parameter $\sigma^2_{1\omega}$ indicates that the seasonal pattern changes in the sample period. The seasonal pattern extracted for the monthly industrial production index is plotted in figure 2 along with the resulting seasonally adjusted series.

The model specification also included 3 regressors representing the calendar effects: the single trading days regressor D_t^*, the Easter variable h_t^* using seven days before Easter, and the length of the month (LOM) variable; the trading day variable accounts for Italian specific holidays (e.g. New Year's Day, Easter Monday, First of May, 8th of December, Christmas, etc.). The estimated coefficients for the industrial production index, denoted respectively by $\hat{\delta}_1^*$, $\hat{\delta}_1^{Easter}$ and $\hat{\delta}_1^{LOM}$ have been

$$\hat{\delta}_1^* = \underset{(0.060)}{0.947} \qquad \hat{\delta}_1^{Easter} = \underset{(0.800)}{-2.774} \qquad \hat{\delta}_1^{LOM} = \underset{(1.283)}{2.024}$$

where in parenthesis are reported the standard errors. The overall effect is shown in figure 2. All the parameters are significant, with the exception of LOM, and have the expected sign. For the second series the calendar effects have been restricted to be proportional to the coefficients of the indicators: the estimated scale factor resulted 2.955, with standard error 0.238. We may thus conclude that the calendar effect on the NA aggregate is significant.

The restrictions on calendar effects associated to the annual series plays a particular role in model specification: by extending this exercise to the other 15 industries covering manufacturing, it has been verified that it is rarely possible to obtain reliable results setting unconstrained coefficients on calendar regressors; further, imposing proportionality, 10 cases converged towards a 0 value for the scale factor, 3 cases gave positive but unreliable values, with only two other cases (cars and machinery and equipment n.e.c.) resulting acceptable. In other words it is clear that moving from a finer towards a larger timing interval reduces accuracy in the estimates of calendar components, at point that these are not more detectable under annual temporal aggregation.

Figure 3 plots the Kalman filter innovations for the IP series and the NA aggregate, along with the standard error confidence interval and the density of standardised innovations. Diagnostics based on these values suggest a satisfactory specification for both the equations: in particular the Box-Ljiung statistic based on 15 and 6 autocorrelations for the two series respectively is equal to 10.078 and 0.183; the Bowman-Shenton normality test is 8.054 for the IP index and 0.915 for the annual NA aggregate.

The smoothed estimates of the disaggregate NA production series are available at both the monthly and quarterly observation frequency. They are presented in unadjusted form in figure 4, along with their 95% upper and lower confidence limits. The size of the confidence interval informs on the reliability of the estimates and embodies the uncertainty surrounding the estimation of the calendar effects (but not that ascribed to the estimation of the hyperparameters - namely the variance parameters).

The quarterly estimates, adjusted for calendar effects, are presented and compared to the raw ones in the last two panels of figure 4. The last plot refers to the estimated growth rates on an annual basis and highlights that not only the adjusted series is smoother, but that the adjustment influences the location and sharpness of turning points.

6.2 The homogeneous system

This example implements an homogeneous system for the Italian series referred to manufacture of cars. Similarly to the first illustration the NA aggregate refers to production at constant prices 1995 = 100 for the years from 1997 to 2003 and the IP series to a monthly index with base 2000 = 100 and sample January 1977-December 2003. The plot of both the series is shown in figure 5.

The estimation of a homogeneous BSM under proportional seasonality, i.e. $\gamma_{2t} = \lambda\gamma_{1t}$ with the suffixes 1 and 2 denoting the IP index and the output series respectively, gave the following maximum likelihood parameter estimates:

$$\hat{\sigma}_{1\epsilon} = 3.746, \quad \hat{\sigma}_{2\epsilon} = 9.441, \quad \hat{\rho}_{\epsilon} = 0.928,$$
$$\hat{\sigma}_{1\omega} = 0.135, \quad \widehat{\lambda} = 2.521, \quad \widehat{q} = 0.669,$$

where $\widehat{q}$ is the homogeneity proportionality factor among the irregular and the level components ($\Sigma_\eta = q\Sigma_\epsilon$), the slope is supposed to be deterministic and $\hat{\rho}_\epsilon$ represents the correlation between the irregular components ϵ_{1t} and ϵ_{2t}; the maximised log-likelihood is equal to -1193.723.

The specification of homogeneity chosen in this case reproduce the model of trend of the first illustration, i.e. a bivariate random walk with constant drift; the series are not cointegrated since correlation among disturbances is estimated equal to 0.928.

The seasonal pattern on the IP series is time varying given that $\hat{\sigma}_{1\omega}$ is a non-zero value. Figure 6 shows the estimated seasonal component along with calendar effects. In this regard only the coefficient related to the single trading days regressor resulted significant, with value and standard error estimated respectively equal to 0.972 and 0.090; the Easter and the LOM regressors were dropped from model specification. The estimate of trading days for the annual series is 3.028 with standard error 0.911.

The diagnostics based on Kalman filter innovations, figure 7, show a satisfactory specification for both the series: in particular the Box-Ljiung statistic based on 15 autocorrelations of standardised innovations is equal to 14.772 for the IP series; it is 0.668 for the NA aggregate based on 6 autocorrelations. The Bowman-Shenton normality test is 15.042 and 0.668 for the IP index and the annual NA aggregate.

Figure 8 presents the smoothed disaggregated estimates. In the first two panels the seasonal unadjusted estimates with calendar effects removed are shown under both the monthly and quarterly frequency along with their 95% upper and lower confidence interval. In the third panel a monthly seasonal adjusted comparison is performed between the disaggregated levels of the NA series and the indicator. Finally, in the forth panel the comparison concerns the annual growth rates of the series.

The most important difference with the first illustration is that the particular form of the model considered here gives the advantage to estimate disaggregated values including the seasonal component. Moreover, the attribution of seasonality to the aggregate series seems to be justified in this example, since there is not significant difference in the log-likelihood values whether the further restriction $\sigma_{12,\epsilon} = \lambda\sigma^2_{2\epsilon}$ is specified. The consequence is that the LR test proposed in section 3 is zero and the null hypothesis of identical regression coefficients among the seasonal and non seasonal components might be not rejected.

7 Conclusions

This article has proposed a disaggregation strategy for the estimation of quarterly national account series that has several advantages over current practice. The strategy is a novel application of the ideas contained in Harvey (1989) and Harvey and Chung (2000).

The estimates arise from fitting a multivariate structural time series model formulated at the monthly interval, which relates the annual national account series to the corresponding monthly indicator. The monthly frequency allows more accurate estimation of the calendar effects.

Maximum likelihood estimation of the unknown parameters, the estimation of the disaggregated observations and their reliability, diagnostic checking and the assessment of goodness of fit are achieved through the state space methodology.

The approach yields automatically "raw" and adjusted estimates without the need to iterate the disaggregation procedure.

Simultaneity and statistical modelling render the proposed strategy more transparent.

From a more philosophical standpoint the approach has the merit of moving away from the exogeneity assumption underlying the disaggregation methods based on a regression framework, such as Chow-Lin (1971), according to which the indicator is considered as an explanatory variable.

Although we have illustrated the bivariate case, which is nevertheless the leading case of interest for statistical agencies, the approach is immediately extended to higher dimensional systems and other frequencies of observations.

Acknowledgments

This work was carried out for the Istat Commission on temporal disaggregation methods for the estimation of quarterly national accounts. The authors wish to thank Tommy di Fonzo, Cosimo Vitale, and Marco Marini for discussion on the issues covered in the paper. The usual disclaimers apply.

References

Chow, G., and Lin, A. L. (1971). Best Linear Unbiased Interpolation, Distribution and Extrapolation of Time Series by Related Series, *The Review of Economics and Statistics*, 53, 4, 372-375.

de Jong, P. (1991). The diffuse Kalman filter, *Annals of Statistics*, 19, 1073-1083.

de Jong, P., and Chu-Chun-Lin, S. (1994). Fast Likelihood Evaluation and Prediction for Nonstationary State Space Models, *Biometrika*, 81, 133-142.

Eurostat (1999). *Handbook on quarterly accounts*, Luxembourg.

Eurostat (2002). *Main results of the ECB/Eurostat Task Force on Seasonal Adjustment of Quarterly National Accounts in the European Union.* Informal working group on Seasonal Adjustment Fifth meeting Luxembourg, 25 Ű- 26 April 2002.

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

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

Harvey, A.C. and Chung, C.H. (2000) Estimating the underlying change in unemployment in the UK, *Journal of the Royal Statistics Society, Series A*, 163, 303-339.

Moauro F. and Savio G. (forthcoming). Temporal Disaggregation Using Multivariate Structural Time Series Models. Econometrics Journal.

Insee (2004). Methodology of French quarterly national accounts. Available at http://www.insee.fr/en/indicateur/cnat_trim/methodologie.htm

Istat (2003). *Principali caratteristiche della correzione per i giorni lavorativi dei conti economici trimestrali.* Direzione centrale della Contabilità Nazionale. Available at www.istat.it, Istat, Rome.

Proietti T. (2000). Comparing seasonal components for structural time series models. *International Journal of Forecasting*, 16, 2, p. 247-260.

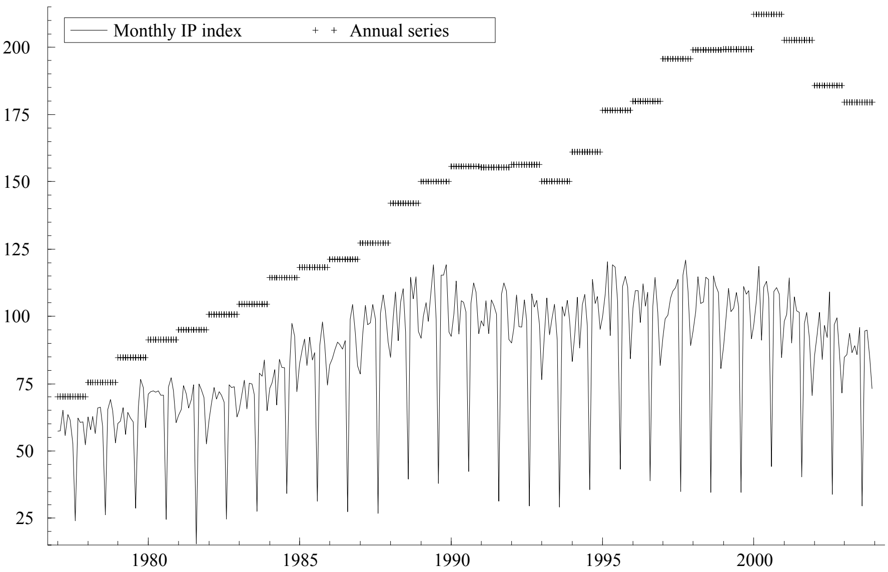

Figure 1: Annual production and monthly industrial production index for electrical and optical equipment.

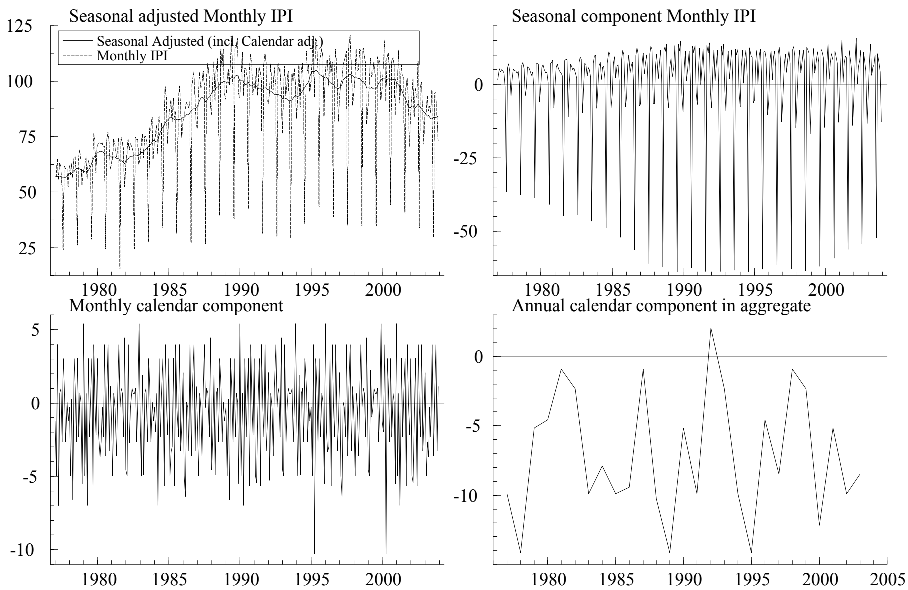

Figure 2: Monthly and quarterly calendar effects estimated for the electrical and optical equipment production series.

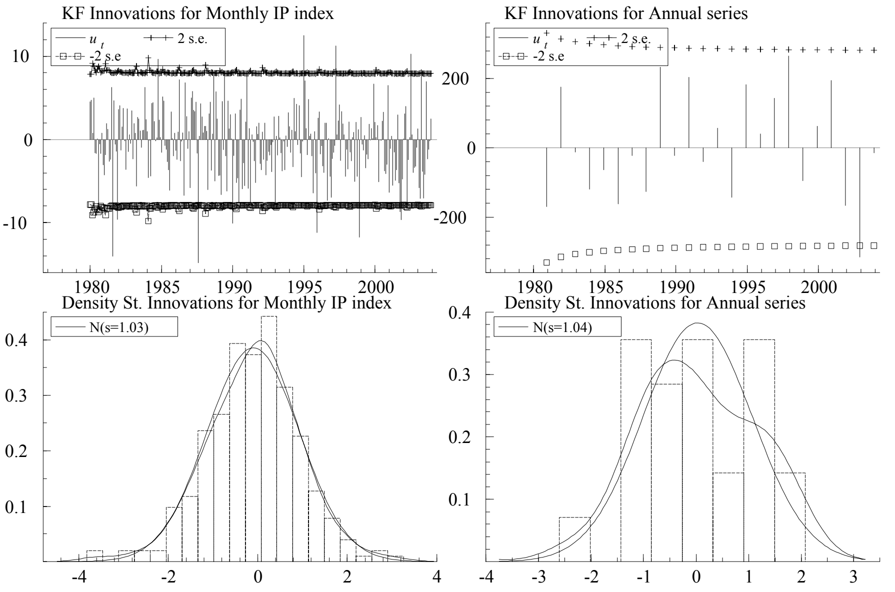

Figure 3: Kalman filter innovations of the model for production of eletrical and optical equipment

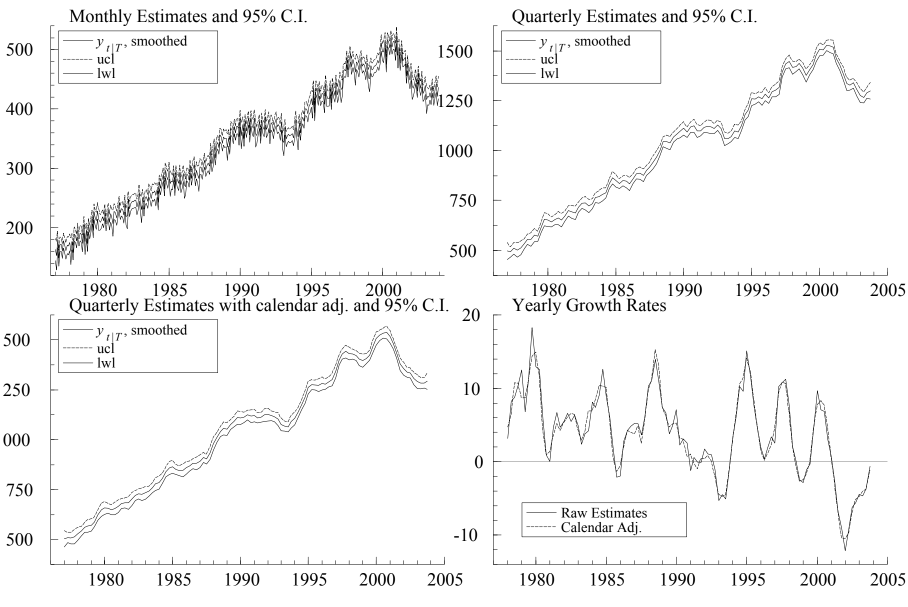

Figure 4: Monthly and quarterly disaggregated production series.

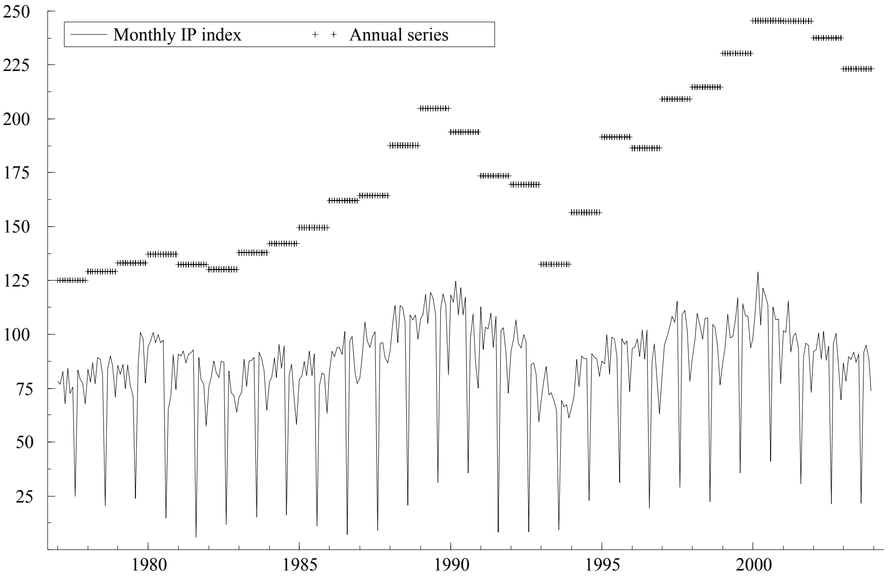

Figure 5: Annual production and monthly industrial production of cars

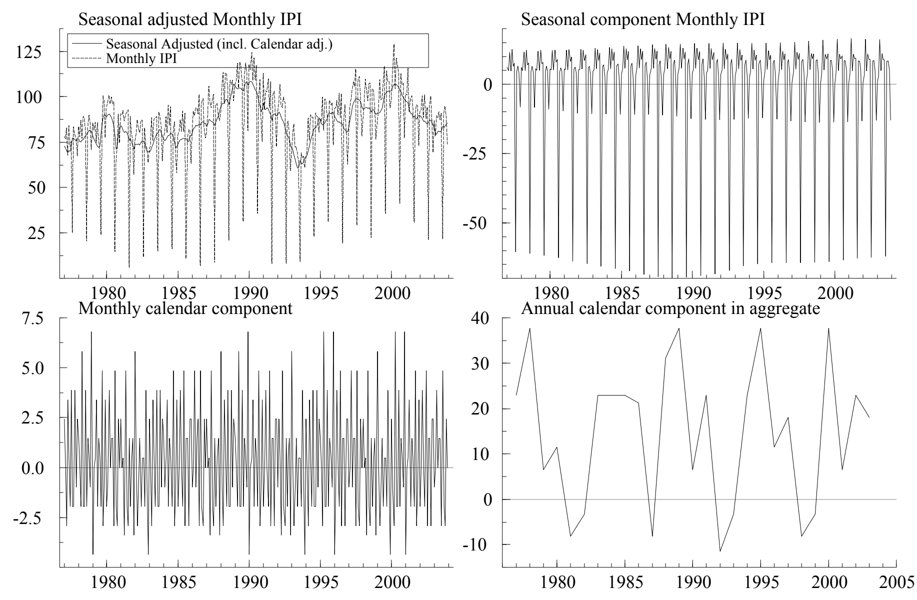

Figure 6: Monthly and quarterly calender effects estimated for the cars series

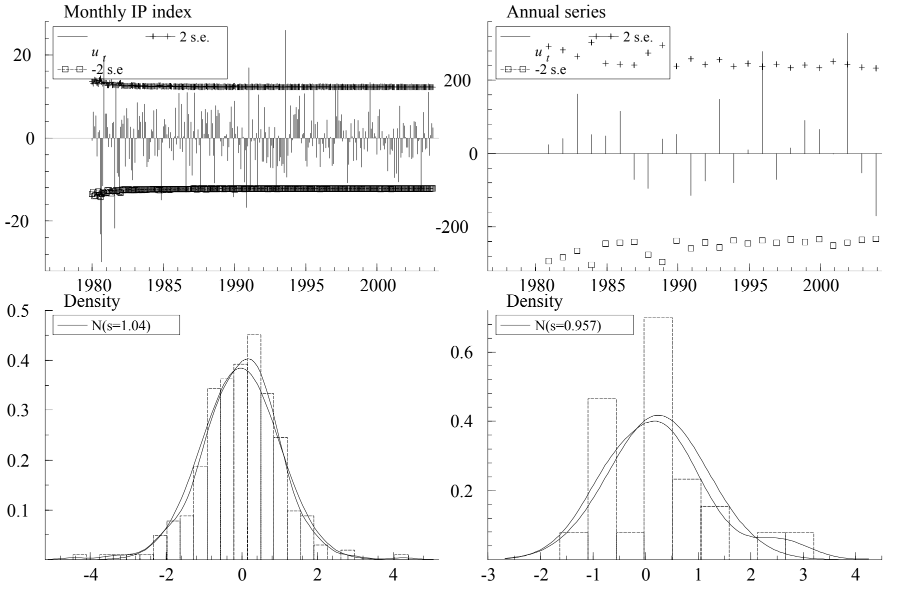

Figure 7: Kalman filter innovations of the model for production of cars

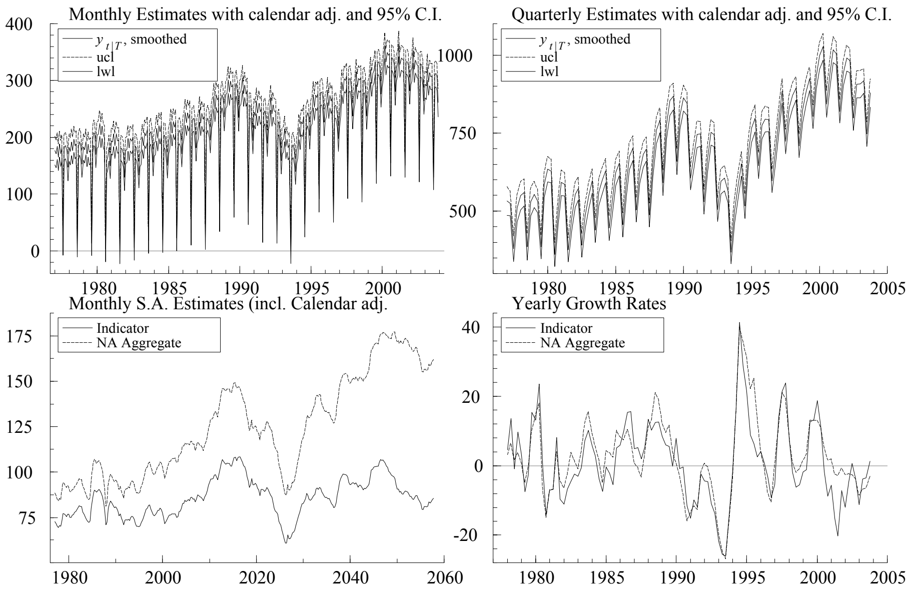

Figure 8: Mnthly and quarterly disaggregated series for production of cars