Transformations and Seasonal Adjustment: Analytic Solutions and Case Studies

Tommaso Proietti and Marco Riani

3. December 2007

Transformations and Seasonal Adjustment: Analytic Solutions and Case Studies

Tommaso Proietti[*]
Università di Roma "Tor Vergata"

Marco Riani[†]
Università di Parma

This version: 3rd December 2007

[*]Dipartimento S.E.F. e ME.Q., Via Columbia 2, 00133 Roma, Italy (corresponding author). E-mail: tommaso.proietti@uniroma2.it.

[†]Dipartimento di Economia, Via Kennedy, 6 - 43100 Parma, Italy, Email: mriani@unipr.it

Abstract

We address the problem of seasonal adjustment of a nonlinear transformation of the original time series, such as the Box-Cox transformation of a time series measured on a ratio scale, or the Aranda-Ordaz transformation of proportions, which aims at enforcing two essential features: additivity and orthogonality of the components. The posterior mean and variance of the seasonally adjusted series admit an analytic finite representation only for particular values of the transformation parameter, e.g. for a fractional Box-Cox transformation parameter. Even if available, the analytical derivation can be tedious and difficult. As an alternative we propose to compute the two conditional moments of the seasonally adjusted series by means of numerical and Monte Carlo integration. The former is both fast and reliable in univariate applications. The latter uses the algorithm known as the simulation smoother and it is most useful in multivariate applications. We present several case studies dealing with robust seasonal adjustment under the square root and the fourth root transformation, the seasonal adjustment of the ratio of two series, and the adjustment of time series of proportions. Our overall conclusion is that robust seasonal adjustment under transformations can be carried out routinely and that the possibility of transforming the scale ought to be considered as a further option for improving the quality of seasonal adjustment.

Keywords: Structural Time Series Models; Box-Cox Transformation; Aranda–Ordaz Transformation; Simulation Smoother; Forward Search; Numerical Integration.

1 Introduction

The linear Gaussian model plays a central role in statistics; it is well understood and its features depend on the (conditional) first and second moments. Transformations aim at establishing a scale, different from the original measurements, for which the linear Gaussian model holds. For variables measured on a ratio scale with a strictly positive support, Tukey (1957) proposed the power transformation to achieve a model with simple structure, normal errors and constant error variance; this was later modified by Box and Cox (1964) and embodied into the model building process, so as to become what is commonly referred to as the Box-Cox transformation. Since then transformations have become a key element in regression analysis; see Atkinson (1985) and Cook and Weisberg (1999). Several modifications have been proposed to deal with negative observations and to extend the support of the transformed observation over the entire real interval; see among others John and Draper (1980), Bickel and Doksum (1981), and Yeo and Johnson (2000). Classes of parametric transformations for proportions were proposed by Guerrero and Johnson (1982), Aranda-Ordaz (1981) and Stukel (1988).

This paper deals with the seasonal adjustment of time series under a parametric nonlinear transformation of the original scale that depends on single parameter. In particular, we concentrate on the Box-Cox power transformation for positive time series, and on the Aranda-Ordaz transformation for proportions. Nevertheless, our approach is immediately generalizable to other parametric transformations that are continuous and invertible.

Seasonal adjustment rests upon two basic pillars: *additivity* and *orthogonality* of the seasonal and non seasonal components. This point is made strongly by Bell and Hillmer (1984, sec. 4.2), who state that "someone who does not want to make these assumptions is working on a different problem". This paper focusses on the situation when the two previous requirements are fulfilled on a scale other than the original scale of measurement and provides a model-based solution to the adjustment problem. Our linear Gaussian workhorse model is an unobserved components model known as the Basic Structural Model (Harvey, 1989). We use this specification, because the estimate of the seasonal component is robust to a misspecification in the nonseasonal part of the model. For example, results from Riani (1998) show that in most cases the estimate of the seasonal component inside the BSM is virtually unaffected by the omission of a stochastic cycle.

In this paper we shall assume that there is a scale at which the series admits an additive and orthogonal decomposition into seasonal and nonseasonal components. The seasonal effects are defined in terms of deviations from the underlying level component, so that their average over a yearly span has zero expectation.

The model is cast in the state space form and the Kalman filter and smoother provide the posterior mean and variance of the nonseasonal component on the transformed scale. The inverse transformation is a biased estimate and provides the conditional median, rather than the mean. We investigate the possibility of evaluating the posterior mean and variance of the nonseasonal component on the original scale in closed form. Extending a result due to Pankratz and Dudley (1987) we show that this occurs only for a fractional Box-Cox parameter; nevertheless, approximations proposed in other contexts, due to Taylor (1986) and Guerrero (1993) can be readily adapted to our problem. Two more general solutions are numerical integration and Monte Carlo integration. The paper shows that these two methods are both fast and reliable, which make seasonal adjustment under transformation a routine operation. Finally it is necessary to remark that, in general, prior outlier detection is performed on the original scale before estimating the transformation parameter. It is clear, however, that observations which seem atypical on the original scale may fit completely inside the bulk of the data once the observations have been transformed. In this paper we show how to evaluate the effect that each particular observations exerts on the fitted model under different transformation scales.

Current seasonal adjustment practice does not take into account the problem of seasonal adjustment under transformation; only SABL, a nonparametric seasonal adjustment procedure developed at Bell Laboratories and documented in Cleveland, Dunn and Terpenning (1978), performs the selection of a preliminary power transformation parameter that minimises the covariance between the level and the seasonal components. The issue of transforming the seasonally adjusted estimates on the original scale is not addressed explicitly. Shulman and McKenzie (1984) illustrate that the SABL estimates of the tranformation parameter differ significantly from the maximum likelihood estimates of an Arima model under the same power transformation. The X-12-Arima programme (Findley *et al.*, 1998) allows for the estimation of an Arima model with regression effects under the Box-Cox and the logistic transformation only for the purpose of obtaining forecast and backcast extensions with the naïve method, i.e. by simple inversion of the extrapolations made on the transformed scale. The Arima model-based seasonal adjustment procedure Tramo-Seats (Gomez and Maravall, 1997) performs a preliminary test for level versus logarithm specification, based on the estimate of the slope coefficient in the trimmed range-mean regression.

The question has to be raised as to why the transformation problem has not received sufficient recognition in the current seasonal adjustment practice. We can envisage three arguments: the first deals with the seasonal balance constraint, by which the expectation of the sum of the seasonal component over a calendar year is zero. According to a well established view the balance constraint should be enforced on the original measurement scale; to put it differently, the seasonally adjusted series should have the same expectation (average) as the original series over twelve consecutive monthly observations. This view is at the root of the treatment of the problem of seasonal adjustment under transformations by Thomson and Ozaki (2002), who propose ad hoc solutions with the specific intent of enforcing the seasonal balance constraint on the original scale. A second argument deals with contemporaneous aggregation: the seasonally adjusted aggregate should be equal to the aggregated sum of the seasonally adjusted sub-series. The consistency in aggregation requires that the series are not transformed as a necessary (though not sufficient) condition, and thus would not hold for the Box-Cox transformation. A third argument concerns the difficulties and the computational burden linked with the detection of influential observations and or of the outliers on the different transformation scales.

None of these arguments is compelling. Multiplicative adjustment, which is used frequently for economic time series already incorporates a different seasonal balance constraint, which refers to the geometric average, rather than the arithmetic. The view taken in this paper is that the possibly stochastic seasonal balance constraint needs to hold only on the transformed scale. The transformation parameter uniquely defines what type of seasonal balance constraint is enforced on the original scale; roughly speaking, if the power transformation parameter is 1, then the balance constraint is additive, if it is equal to 0 it is multiplicative, if it is -1 is harmonic, i.e. it is defined on the reciprocal of the series. Moreover, the conditions for consistency in cross-sectional aggregation are so stringent that the indirect seasonal adjustment of an aggregate is almost never used in practice. As concerns the third argument, in this paper we show how it is possible to robustly estimate the transformation parameter and at the same way to evaluate the effect that the different seasons exert on this estimate.

The plan of the paper is the following: section 2 presents the linear Gaussian model for the decomposition of a series into seasonal and nonseasonal components that will be used for seasonal adjustment. We first concentrate on the Box-Cox power transformation for positive observations and deal with the estimation of the transformation parameter in section 3. The evaluation of the posterior mean and variance of nonseasonal component is considered in section 4. The availability of closed form solutions is investigated and approximate solutions are reviewed. A more general approach is to evaluate the conditional moments by numerical and Monte Carlo integration using the simulation smoother (de Jong and Shephard, 1995). In section 5 we show how it is possible through the use of the forward search algorithm to obtain a robust estimate of the transformation parameter.

In section 6 The different estimation methods are applied to a well known case study concerning the Sales of an engineering company (Chatfield and Prothero, 1973), which calls for the fourth root transformation, and the Italian industrial production index for Leather and Shoes, for which the square root transformation is suggested. The assessment of the different methods leads to the conclusion that numerical integration is both fast and reliable in univariate applications.

We then turn to a bivariate application dealing with the seasonal adjustment of time series of tourist arrivals and overnight stays, along with the ratio of the latter to the former, which measures average stay (Section 7). The only option available for the seasonal adjustment of the ratio of two time series is Monte Carlo integration using the simulation smoother. Finally, in section 8 we consider the seasonal adjustment of time series of proportions. Among the available transformations, we concentrate on the class proposed by Aranda-Ordaz (1981), which is applied to the proportion of tourist arrivals in hotels. We draw our conclusions in section 9.

2 The basic structural model under transformations

The parametric linear and Gaussian model that we employ for the adjustment is the basic structural model (BSM henceforth, see Harvey, 1989). The BSM postulates an additive and orthogonal decomposition of a time series into unobserved components representing the trend, seasonality and the irregular component.

We assume that the BSM holds for a transformation $u_t(\lambda)$ of the original time series y_t, depending on a single transformation parameter λ. An important case is the Box-Cox (BC) transformation:

$$u_t(\lambda) = \begin{cases} \frac{y_t^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \ln y_t, & \lambda = 0 \end{cases} \tag{1}$$

see Box and Cox (1964). The above transformation is suitable for series measured on a ratio scale, which take only strictly positive values.

The BSM for the transformed series is formulated as follows:

$$u_t(\lambda) = \mu_t + \gamma_t + \sum_{k=1}^{K} \delta_k x_{kt} + \epsilon_t, \quad t = 1, \ldots, T, \tag{2}$$

where μ_t is the trend component, γ_t is the seasonal component, the x_{kt}'s are appropriate regressors that account for calendar effects, namely trading days, moving festivals (Easter) and the length of the month, and $\epsilon_t \sim \mathrm{NID}(0, \sigma_\epsilon^2)$ is the irregular component.

The trend component has a local linear representation:

$$\begin{array}{lclclcll} \mu_{t+1} & = & \mu_t & + & \beta_t & + & \eta_t, & \eta_t \sim \mathrm{NID}(0, \sigma_\eta^2), \\ \beta_{t+1} & = & & & \beta_t & + & \zeta_t, & \zeta_t \sim \mathrm{NID}(0, \sigma_\zeta^2), \end{array}$$

where β_t is the stochastic slope, that in turn evolves as a random walk; the disturbances η_t, ζ_t, are independent of each other and of any remaining disturbance in the model.

The seasonal component has a trigonometric representation, such that the seasonal effect at time t arises from the combination of six stochastic cycles: $\gamma_t = \sum_{j=1}^{6} \gamma_{jt}$, where, for $j = 1, \ldots, 5$,

$$\begin{array}{lcrll} \gamma_{j,t+1} & = & & \cos\lambda_j \gamma_{j,t} + \sin\lambda_j \gamma_{j,t}^* + \omega_{j,t} & \omega_{j,t} \sim \mathrm{NID}(0, \sigma_\omega^2) \\ \gamma_{j,t+1}^* & = & - & \sin\lambda_j \gamma_{j,t} + \cos\lambda_j \gamma_{j,t}^* + \omega_{j,t}^* & \omega_{j,t}^* \sim \mathrm{NID}(0, \sigma_\omega^2) \end{array}$$

and $\gamma_{6,t+1} = -\gamma_{6t} + \omega_{6t}$, $\omega_{6t}^* \sim \mathrm{NID}(0, \sigma_\omega^2/2)$; $\lambda_j = \frac{2\pi}{12} j$ is the seasonal frequency. The disturbances ω_{jt} and ω_{jt}^* are assumed to be normally and independently distributed with common variance σ_ω^2. All the disturbances are assumed to be mutually uncorrelated.

An alternative approach to model stochastic seasonality is derived by writing

$$\gamma_t = \mathbf{x}_t' \xi_t \qquad \xi_t = \xi_{t-1} + \boldsymbol{\omega}_t$$

where $\mathbf{x}_t' = [D_{1t}, \ldots, D_{st}]$, with $D_{jt} = 1$ in season j and 0 otherwise. The vector ξ_t contains the effects associated to each season and changes over time according to a multivariate random walk; $\boldsymbol{\omega}_t$ is a zero-mean multivariate white noise with covariance matrix

$$\text{Var}(\boldsymbol{\omega}_t) = \sigma_\omega^2[\mathbf{I}_s - \frac{1}{s}\mathbf{i}_s\mathbf{i}_s']$$

which enforces the constraint $\mathbf{i}_s'\text{Var}(\boldsymbol{\omega}_t) = \mathbf{0}$. This formulation is known in the literature as the Harrison and Stevens (HS) specification. The distinguishing feature of this approach is that it is formulated directly in terms of the effect of a particular season, thereby enhancing flexibility needed to model seasonal heteroscedasticity (that is when there are seasons which are 'more variables' than others, see Proietti, 1998). The appropriate action for this model to deal with heteroscedasticity is to define the covariance matrix of the seasonal innovations as follows:

$$\text{Var}(\boldsymbol{\omega}_t) = \mathbf{D} - \frac{1}{\mathbf{i}_s'\mathbf{D}\mathbf{i}_s}\mathbf{D}\mathbf{i}_s\mathbf{i}_s'\mathbf{D}$$

where $\mathbf{D}$ is a diagonal matrix, $\mathbf{D} = \text{diag}\{d_j, j = 1, \ldots, s\}$.

A comparison of the various representations of a seasonal component and a discussion of the implications for forecasting are given in Proietti (2000). One of the purposes of this paper is to check how the presence of seasonal heteroscedasticity may affect the estimate of the transformation parameter.

Trading day (working day) effects are modelled as fixed effects through the inclusion of appropriate regressors (see Cleveland and Devlin, 1982, Bell and Hillmer, 1984). Letting x_{jt} denote the number of days of type j, $j = 1, \ldots, 7$, occurring in month t and assuming that the effect of a particular day is constant, the trading day effect is given by:

$$TD_t = \sum_{k=1}^{6} \delta_k \left(x_{kt} - x_{7t}\right)$$

The regressors are the differential number of days of type k, $k = 1 \ldots, 6$, compared to the number of Sundays, to which type 7 is conventionally assigned. The Sunday effect on the series is then obtained as $\left(-\sum_{k=1}^{6} \delta_k\right)$. This ensures that the TD effect is zero over a period corresponding to multiples of the weekly cycle. The regressors are sometimes corrected to take into account the national calendars.The only moving festival that we consider is Easter; its effects are modelled as

$E_t = \delta_E h_t$ where h_t is the proportion of 7 days before Easter that fall in month t. Subtracting the long run average, computed over the first 400 years of the Gregorian calendar (1583-1982), from h_t yields the regressor $h_t^* = h_t - \bar{h}_t$, where $\bar{h}_t$ takes the values 0.354 and 0.646 respectively in March and April, and zero otherwise. Finally, the length of month (LOM) regressor results from subtracting from the number of days in each month, $\sum_{k=1}^{7} D_{kt}$, its long run average, which is 365.25/12.

3 Estimation and inference for the BSM under transformation

For a given value of the transformation parameter the BSM can be cast in state space form. The Kalman filter enables the evaluation of the likelihood via the prediction error decomposition. See Durbin and Koopman (2001) and Harvey and Proietti (2005) for a review. The maximum likelihood estimates can be obtained by a quasi-Newton algorithm, such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm (see Press *et alii*, 1992, sec. 10.7). Diagnostic checking can be carried out on the standardized one-step-ahead prediction errors, also known as the innovations, so as to detect any departure from the stated assumptions and possibly take the corrective actions against them.

As far as the estimation of the transformation parameter is concerned, we maximise the profile likelihood corrected so as to take into account the change of scale of the observations. For this purpose the logarithm of the Jacobian, $\ln J(\lambda) = (\lambda - 1) \sum_t \ln(y_t)$, must be added to the log-likelihood of the transformed observations $\mathcal{L}\{u_1(\lambda), \dots, u_T(\lambda); \tilde{\boldsymbol{\theta}}(\lambda)\}$, maximized over the (transformed) hyperparameters $\tilde{\boldsymbol{\theta}}(\lambda)$, where, e.g. $\sigma_\eta^2 = \exp(2\theta_1(\lambda))$, $\sigma_\zeta^2 = \exp(2\theta_2(\lambda))$, $\sigma_\omega^2 = \exp(2\theta_3(\lambda))$, and $\sigma_\epsilon^2 = \exp(2\theta_4(\lambda))$. For the treatment of nonstationary and regression effects see de Jong (1991) and Koopman (1997).

Hence, the maximum likelihood estimate of the transformation parameter emerges as the solution of the problem:

$$\max_\lambda \left\{ \mathcal{L}\{u_1(\lambda), \dots, u_T(\lambda); \tilde{\boldsymbol{\theta}}(\lambda)\} + \ln J(\lambda) \right\}.$$

This can be done in practice by performing a grid search over the range of values of λ. Equivalently, we can maximize the uncorrected log-likelihood of the normalized observations $u_t(\lambda) / \prod_t y_t^{\lambda-1}$

(Atkinson, 1985).

Alternative approximate methods for estimating λ are spread-level regression and added variable regression. In the former case the series $u_t(\lambda)$ is divided into yearly non overlapping subsets consisting of 12 observations; then either the standard deviation of the yearly subsets, or the interdecile range, is regressed on a constant and either the corresponding yearly averages of $u_t(\lambda)$ or the median. The estimated regression slope provides an approximate estimate of $\lambda - 1$; if it results not significantly different from zero, no transformation ($\lambda = 1$) is suggested.

The idea behind added variable estimation of λ (Atkinson, 1985, ch. 6) is to consider the first order Taylor series expansion of $u_t(\lambda)$ about a maintained value λ_0 (e.g. 0 or 1): $u_t(\lambda) = u_t(\lambda_0) + (\lambda - \lambda_0)w_t(\lambda_0)$, with $w_t(\lambda_0) = (\partial u_t(\lambda)/\partial\lambda)|_{\lambda=\lambda_0}$. If for some λ, $u_t(\lambda) = \mu_t + \gamma_t + \sum_k \delta_k x_{kt} + \epsilon_t$, then the approximate linear model is

$$u_t(\lambda_0) = \mu_t + \gamma_t + \sum_k \delta_k x_{kt} + \delta^* w_t(\lambda_0) + \epsilon_t,$$

with $\delta^* = \lambda_0 - \lambda$. The augmented model is estimated including among the regressors the additional variable $w_t(\lambda_0)$. Significant regression denotes the need for a transformation and provides a preliminary estimate of the correct λ as $\hat{\lambda} = \lambda_0 - \hat{\delta}$. The t test on the additional constructed variable $w_t(\lambda_0)$ is known in the statistic literature as "score test statistic for transformation".

Conditional on the λ and $\boldsymbol{\theta}(\lambda)$ parameter estimates the Kalman smoother provides the conditional expectations of the latent components given all the available observations, along with their conditional variance. These inference are employed in the next section to produce estimates of the seasonally adjusted series on the untransformed scale.

4 Seasonal Adjustment and the Box-Cox Transformation

Let us write (2) as $u_t(\lambda) = u_t^* + \gamma_t + \sum_k \delta_k x_{kt}$, where $u_t^* = \mu_t + \epsilon_t$ is the seasonally adjusted series on the transformed scale, and denote by $\tilde{u}_t^* = \mathrm{E}(u_t^*|\mathcal{F}_T)$ and $V_t = \mathrm{Var}(u_t^*|\mathcal{F}_T)$, respectively the posterior mean and variance of u_t^*, $\mathcal{F}_t$ being the information set at time t. These inferences are delivered by the Kalman filter and smoother applied to the relevant linear state space model. See e.g. Durbin and Koopman (2001) for details.

We define the seasonally adjusted series on the original scale as the inverse transformation of the nonseasonal component u_t^*, $y_t^* = u^{-1}(u_t^*)$, where $u^{-1}(\cdot)$ is the inverse transformation. For the Box-Cox transformation:

$$y_t^* = \begin{cases} (1+\lambda u_t^*)^{1/\lambda}, & \lambda \neq 0, \\ \exp(u_t^*), & \lambda = 0. \end{cases}$$

The estimator of the seasonally adjusted series is thus

$$\tilde{y}_t^* = \mathrm{E}(y_t^*|\mathcal{F}_T) = \int u^{-1}(u_t^*) f(u_t^*|\mathcal{F}_T) du_t^*. \tag{3}$$

whereas the conditional variance of the estimation error for the seasonally adjusted series is defined as:

$$\mathrm{Var}(y_t^*|\mathcal{F}_T) = \int \left[u^{-1}(u_t^*) - \tilde{y}_t^*\right]^2 f(u_t^*|\mathcal{F}_T) du_t^* = \mathrm{E}(y_t^{*2}|\mathcal{F}_T) - \tilde{y}_t^{*2}. \tag{4}$$

As is well-known, the conditional expectation is the optimal estimator under quadratic loss. The above integrals do have a closed form solution only in particular cases, namely $\lambda = 0$, and $\lambda = 1/p, p = 1, 2, 3, \ldots$, as it will be seen shortly.

Notice that the naïve estimator of the SA series,

$$\hat{y}_t^* = \begin{cases} (1+\lambda \tilde{u}_t^*)^{1/\lambda}, & \lambda \neq 0, \\ \exp(\tilde{u}_t^*), & \lambda = 0, \end{cases} \tag{5}$$

provides the *median* of the conditional distribution of y_t^*, given the observations.

For $\lambda = 0$ using the properties of the lognormal distribution we have that:

$$\mathrm{E}(y_t^*|\mathcal{F}_T) = \exp\left(\tilde{u}_t^* + \frac{V_t}{2}\right) = \hat{y}_t^* \exp\left(\frac{V_t}{2}\right). \tag{6}$$

$$\mathrm{Var}(y_t^*|\mathcal{F}_T) = \exp\left(2\tilde{u}_t^* + V_t\right) \cdot \left(\exp(V_t) - 1\right). \tag{7}$$

4.1 Analytical solutions

For general λ, in the appendix we prove the following theorem.

Theorem 1: the mean and the variance of the seasonally adjusted series in the original scale are given by the two following expressions:

$$E(y_t^*|\mathcal{F}_T) = \tilde{y}_t^* = \hat{y}_t^* \left[1 + \sum_{j=1}^{\infty} k_{2j}(t) a_j(t)\right] \tag{8}$$

$$\text{Var}(y_t^*|\mathcal{F}_T) = \hat{y}_t^{*^2} \left[\sum_{j=1}^{\infty} k_j^2(t) a_j(t) + 2\sum_{j=1}^{\infty}\sum_{r=1}^{\infty} k_j(t) k_{j+2r}(t) a_{j+r}(t) - \left(\sum_{j=1}^{\infty} k_{2j}(t) a_j(t)\right)^2\right] \tag{9}$$

where[1]

$$a_j(t) = \text{E}\left[(u_t^* - \tilde{u}_t^*)^{2j}|\mathcal{F}_T\right] = \frac{(2j)!}{j!2^j} V_t^j \qquad \text{and} \qquad k_j(t) = \frac{1}{j!}\left(\prod_{k=1}^{j-1}(1-\lambda k)\right)\hat{y}_t^{*-\lambda j}$$

The results follow from the Taylor series expansion of the reverse transformation. Notice that for $\lambda = 0$ $k_j(t) = (j!)^{-1}$ and the term of (8) is simply the expansion of $\exp(V_t/2)$. This method was proposed originally by Neyman and Scott (1960), who however did not consider explicitly time series applications and did not give the analytical exact solution for $\lambda = 1/p$, with p integer. An alternative approach for expressing the time series forecasts on the original scale, based on Hermite polynomial expansion, was suggested by Granger and Newbold (1976). The expression (8) was derived by Pankratz and Dudley (1987) for the simple power transformation y_t^λ using a different argument. For integer $p = 1/\lambda$ they write the inverse transformation as $u_t^{*p} = (\tilde{u}_t^* + \sqrt{V_t} w_t)^p = \tilde{u}_t^{*p}(1 + \frac{\sqrt{V_t}}{\tilde{u}_t^*} w_t)^p$, where $w_t \sim N(0,1)$. They then consider the expansion of the binomial and take the expectation.

The expressions in square brackets in equations (8) and (9) are the multiplicative correction terms that have to be applied to the naïve estimator of the SA series or to its square in order to produce the conditional mean and the conditional variance in the original scale.

An alternative expression for the variance is derived as follows. Defining $\hat{V}_t^*$ the naïve estimate of the variance resulting from the application of the Delta method,

$$\hat{V}_t^* = V_t\left[\frac{du^{-1}(u_t^*)}{du_t^*}\Big|_{u_t^*=\tilde{u}_t^*}\right]^2 = V_t \hat{y}_t^{*2(1-\lambda)},$$

[1] We adopt the convention that when $j = 1$ the product in brackets in k_j equals 1, $\prod_{i=1}^{0} x_i = 1$.

then we can rewrite (9) as:

$$\mathrm{Var}(y_t^*|\mathcal{F}_T) = \hat{V}_t^* \left[1 + \sum_{j=2}^{\infty} \bar{k}_j^2(t)\bar{a}_j(t) + 2\sum_{j=1}^{\infty}\sum_{r=1}^{\infty} \bar{k}_j(t)\bar{k}_{j+2r}(t)\bar{a}_{j+r}(t) - V_t\left(\sum_{j=1}^{\infty} \bar{k}_{2j}(t)\bar{a}_j(t)\right)^2\right], \tag{10}$$

where $\bar{k}_j(t) = k_j(t)\hat{y}_t^{*\lambda}$ and $\bar{a}_j(t) = a_j(t)/V_t$. According to expression (10), the exact variance can be seen as the product of the naïve variance resulting from the Delta method and a correction factor.

For $\lambda = 1/p, p = 1, 2, \ldots$, it is immediate to see that the series $k_1(t), k_2(t), \ldots$ contains only p terms different from zero. For example, for $\lambda = 1$, $k_1(t) = 1/\hat{y}_t^*$ and $k_2(t) = k_3(t) = \cdots = 0$ so that $\tilde{y}_t^* = \hat{y}_t^*$ and $\mathrm{var}(y_t^*|\mathcal{F}_T) = V_t$ as obvious. In the case of the square root transformation $(\lambda = 0.5)$ $k_1(t) = 1/\sqrt{\hat{y}_t^*}$, $k_2(t) = 1/(4\hat{y}_t^*)$, $a_1(t) = V_t$, $a_2(t) = 3V_t^2$ and $k_3(t) = k_4(t) = \cdots = 0$ so that

$$\tilde{y}_t^* = \hat{y}_t^* \left[1 + \frac{1}{4}\frac{V_t}{\hat{y}_t^*}\right] \tag{11}$$

and

$$\mathrm{Var}(y_t^*|\mathcal{F}_T) = \hat{y}_t^* V_t + \frac{1}{8}V_t^2.$$

Using similar arguments we give in Table 1, for the most common values of λ, the exact correction factors for the mean and the variance which must be applied to the naïve estimator of the seasonally adjusted series $\hat{y}_t^*$ in order to find the true conditional mean and variance in the original scale.

This table clearly shows that the correction term depends on the ratio between the variance (raised to some power) of the SA series on the transformed scale and the value of the naïve estimator (raised to some power of λ). If this is small, the correction is negligible. More precisely, we have that:

Lemma 1 The correction factor for the mean which we call $\psi_\mu(\lambda, \hat{y}_t^*, V_t)$ satisfies the following properties:

(i) $\psi_\mu(\lambda, \hat{y}_t^*, V_t) \leq 1$ for $\lambda \geq 1$ and $\psi_\mu(\lambda, \hat{y}_t^*, V_t) \geq 1$ for $\lambda \leq 1$.

(ii) $\psi_\mu(\lambda, \hat{y}_t^*, V_t) = 1$ when $\lambda = 1$

Table 1: Exact correction factors which have to be applied to the naïve estimator of the seasonally adjusted series and of the variance, in order to obtain the conditional mean and the conditional variance in the original scale for the most important fractional values of λ.

	Mean	Variance
λ	Correction factor for $\hat{y}_t^*$	Correction factor for $\hat{V}_t^*$
1/2	$1 + \frac{1}{4}\frac{V_t}{\hat{y}_t^*}$	$1 + \frac{1}{8}V_t\hat{y}_t^{*-1}$
1/3	$1 + \frac{1}{3}\frac{V_t}{\hat{y}_t^{*2/3}}$	$1 + \frac{4}{9}V_t\hat{y}_t^{*-2/3} + \frac{5}{243}V_t^2\hat{y}_t^{*-4/3}$
1/4	$1 + \frac{3}{8}\frac{V_t}{\hat{y}_t^{*1/2}} + \frac{3}{256}\frac{V_t^2}{\hat{y}_t^*}$	$1 + \frac{21}{32}V_t\hat{y}_t^{*-1/2} + \frac{3}{32}V_t^2\hat{y}_t^{*-1} + \frac{3}{32}V_t^3\hat{y}_t^{*-3/2}$

(iii) $\psi_\mu(\lambda, \hat{y}_t^*, V_t) \to 1^-$ when $\lambda \to +\infty$ and $\psi_\mu(\lambda, \hat{y}_t^*, V_t) \to +\infty$ when $\lambda \to -\infty$ if $\hat{y}_t^* > 1$. $\psi_\mu(\lambda, \hat{y}_t^*, V_t) \to -\infty$ when $\lambda \to +\infty$ and $\psi_\mu(\lambda, \hat{y}_t^*, V_t) \to 1^+$ when $\lambda \to -\infty$ if $\hat{y}_t^* < 1$.

(iv) $\psi_\mu(\lambda, \hat{y}_t^*, V_t)$ is non increasing for $\lambda \leq \frac{1}{2\ln \hat{y}_t^*} + 1$ and non decreasing for $\lambda \geq \frac{1}{2\ln \hat{y}_t^*} + 1$ if $\hat{y}_t^* > 1$. $\psi_\mu(\lambda, \hat{y}_t^*, V_t)$ is non decreasing for $\lambda \leq \frac{1}{2\ln \hat{y}_t^*} + 1$ and non increasing for $\lambda \geq \frac{1}{2\ln \hat{y}_t^*} + 1$ if $\hat{y}_t^* < 1$.

The proofs are straightforward but tedious. Details are given in a technical report by the authors.

Figure 1 shows the correction factor as a function of λ for 6 different combinations of values of $\hat{y}_t^*$ and V_t. The symbol of the square which is drawn in correspondence of $\lambda = 0$ denotes the value obtained applying directly the formula given in equation (6). It is clear that if the correction factor is neglected there is negative (positive) bias for $\lambda < 1$ ($\lambda > 1$) which can be more or less severe depending on the problem under study. The first 2 left panels show that if the ratio between the naïve estimator and the value of the variance of the SA series in the transformed scale is greater than a certain threshold and if the estimated λ is greater (smaller) than 1 and $\hat{y}_t^*$ is greater (smaller) than 1, the correction which must be applied to the naïve estimator can be overlooked. Finally, notice that in this case the value of the minimum is for the panels in the top row is $\lambda \approx -2.40$ while that for the bottom panels is $\lambda \approx 2.06$.

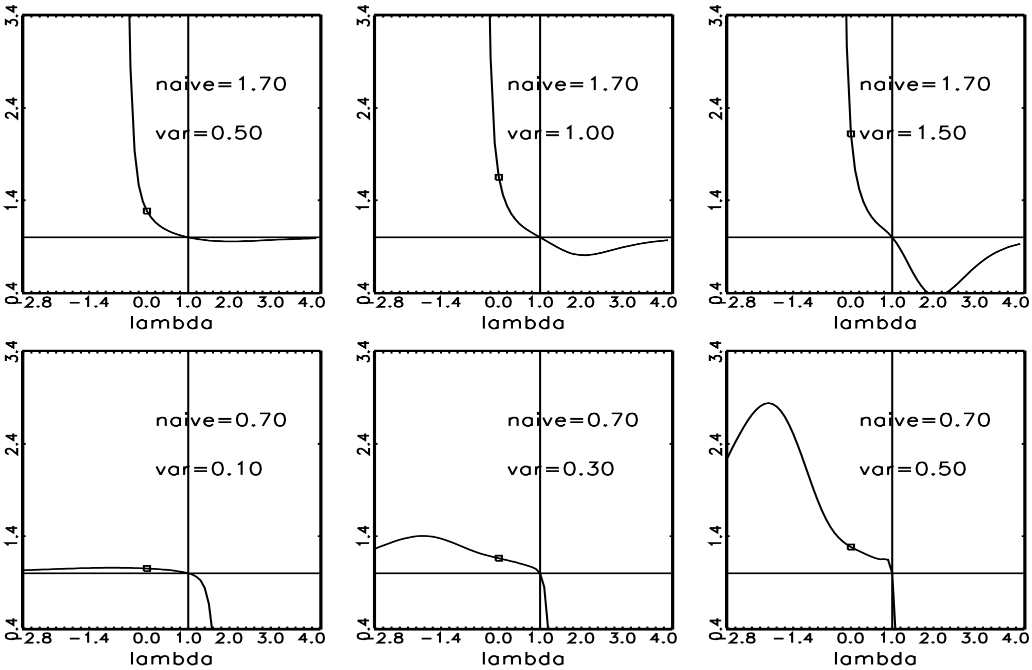


Figure 1: Correction factor which must be applied to the naïve estimator of the seasonally adjusted series to obtain the conditional mean in the original scale as a function of λ for six different combinations of $\hat{y}_t^*$ (naive) and V_t (var).

Lemma 2 The correction factor for the naïve variance which we call $\psi_\sigma(\lambda, \hat{y}_t^*, V_t)$ satisfies the following properties:

(i) $\psi_\sigma(\lambda, \hat{y}_t^*, V_t) = 1$ when $\lambda = 1$

(ii) $\psi_\sigma(\lambda, \hat{y}_t^*, V_t) \to 1^+$ when $\lambda \to +\infty$ and $\psi_\sigma(\lambda, \hat{y}_t^*, V_t) \to +\infty$ when $\lambda \to -\infty$ if $\hat{y}_t^* > 1$. $\psi_\sigma(\lambda, \hat{y}_t^*, V_t) \to 1^+$ when $\lambda \to -\infty$ and $\psi_\sigma(\lambda, \hat{y}_t^*, V_t) \to +\infty$ when $\lambda \to +\infty$ if $\hat{y}_t^* < 1$.

(iii) $\psi_\sigma(\lambda, \hat{y}_t^*, V_t)$ is non decreasing (non increasing) for $\frac{4}{5} + \frac{1}{2\ln \hat{y}_t^*} - k_{\hat{y}_t^*} \leq \lambda \leq \frac{4}{5} + \frac{1}{2\ln \hat{y}_t^*} + k_{\hat{y}_t^*}$ if $\hat{y}_t^* > 1$ ($\hat{y}_t^* < 1$). $k_{\hat{y}_t^*} = \sqrt{\frac{1}{25} + \frac{1}{4\ln^2 \hat{y}_t^*}}$.

These three properties are illustrated graphically in Figure 2. The symbol of the square which is drawn in correspondence of $\lambda = 0$ denotes the value obtained applying directly the formula given in equation (7). It is interesting to notice that when the correction factor for the mean tends to one ($\psi_\mu(\lambda, \hat{y}_t^*, V_t) \to 1$) also the correction factor for the naïve variance tends to 1 ($\psi_\sigma(\lambda, \hat{y}_t^*, V_t) \to 1^+$). On the other hand, when the correction factor for the mean becomes very large in absolute value ($\psi_\mu(\lambda, \hat{y}_t^*, V_t) \to \infty$), the correction for the variance goes to plus infinity, ($\psi_\sigma(\lambda, \hat{y}_t^*, V_t) \to +\infty$).

4.2 Approximate and computational solutions

Taylor (1986) proposed an approximate correction for the case $\lambda \neq 0$, which amounts to neglecting higher order terms in the expansion (13):

$$\tilde{y}_t^{*T} = \hat{y}_t^* \left[1 + \frac{1}{2}(1-\lambda)\frac{V_t}{\hat{y}_t^{*2\lambda}}\right].$$

This estimate is exactly equal to equation (11) so it is exact only for $\lambda = 0.5$.

Guerrero (1993) proposed a solution which is coincident with the exact solution in the logarithmic case ($\lambda = 0$) and is approximate for $\lambda \neq 0$. In our notation, it can be written as follows:

$$\tilde{y}_t^{*G} = \tilde{y}_t^* \left\{\frac{1}{2} + \frac{1}{2}\left[1 + 2\lambda(1-\lambda)\frac{V_t}{\hat{y}_t^{*2\lambda}}\right]^{1/2}\right\}^{1/\lambda}.$$

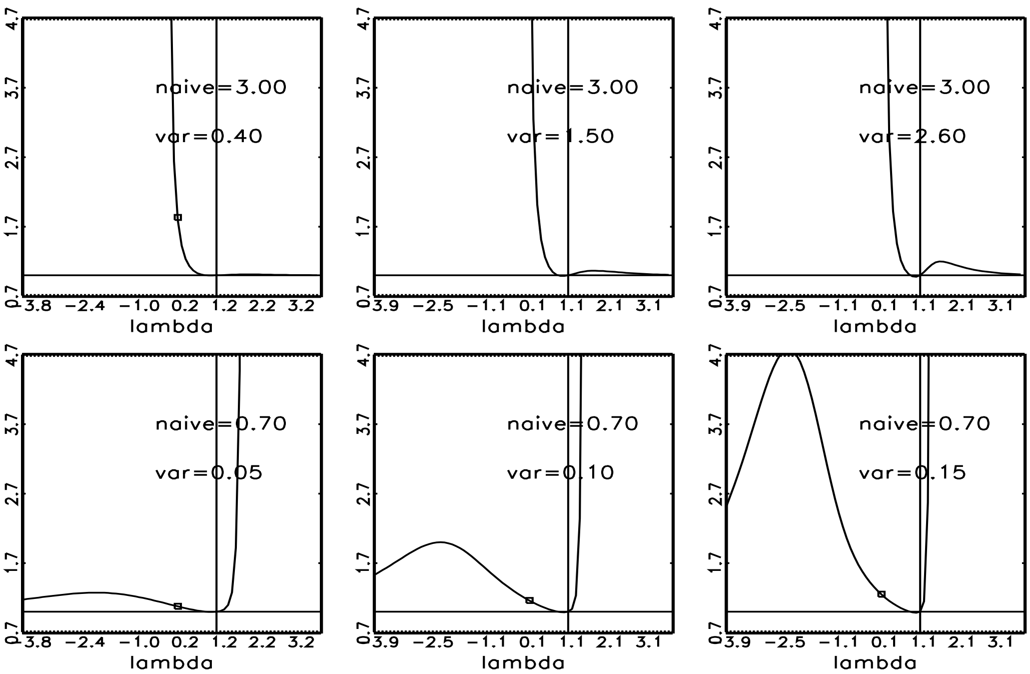


Figure 2: Correction factor which must be applied to the naïve estimator of the variance of the seasonally adjusted series to obtain the conditional variance in the original scale as a function of λ for six different combinations of $\hat{y}_t^*$ (naive) and V_t (var).

The derivation of Guerrero's result proceeds from the following argument: using the second order Taylor approximation of u_t^* around $\tilde{y}_t$, taking the conditional expectation, and using the approximation $\text{Var}(y_t^*|\mathcal{F}_T) \approx \tilde{y}_t^{*2(1-\lambda)} V_t$, yields

$$1 + \lambda \tilde{u}_t^* = \tilde{y}_t^{*\lambda} + \frac{1}{2}\lambda(\lambda - 1)V_t \tilde{y}_t^{*-\lambda}$$

which is a quadratic equation in $\tilde{y}_t^{*\lambda}$. Solving for $\tilde{y}_t^*$ provides the Guerrero approximation.

For general λ there are three possible ways of evaluating $\text{E}(y_t^*|\mathcal{F}_T)$ and $\text{Var}(y_t^*|\mathcal{F}_T)$:

- Monte Carlo evaluation using the simulation smoother: the latter is used to draw repeated samples from the conditional distribution of $u^* = \{u_1^*, \ldots, u_T^*\}$, given the available observations.
- Numerical integration with respect to the normal density, $f(u_t^*|\mathcal{F}_T)$, whose moments $\tilde{u}_t^*$ and V_t^* are provided by the Kalman filter and smoother.
- Direct application of equations (8) and (9) truncating the summations to a particular order.

As concerns the first method (Monte Carlo evaluation), sampling from the posterior distribution of the latent components or disturbances has been considered in detail: Carlin, Polson and Stoffer (1992) proposed a single move state sampler, which however usually is very inefficient due to the high correlation between the unobserved components, especially when they are weakly evolutive. Gamerman (1998) proposed a single move disturbance sampler, which is more efficient since the disturbances driving the components are much less persistent and autocorrelated over time. Along with reparameterization, an effective strategy is blocking, through the adoption of a multimove sampler as in Carter and Kohn (1994) and Früwirth Schnatter (1994), who focus on sampling the unobserved components. Again, a more efficient multimove sampler can be constructed by focusing on the disturbances, rather than the states. This is the idea underlying the simulation smoother proposed by de Jong and Shephard (1995).

Letting $\varsigma_t = [\eta_t, \zeta_t, \epsilon_t]'$ denote the vector of disturbances that drive the nonseasonal component of the series, $u_t^* = \mu_t + \epsilon_t$, the simulation smoother hinges on the following factorisation of the joint posterior density:

$$f(\varsigma_0, \ldots, \varsigma_T|\mathcal{F}_T) = f(\varsigma_T|\mathcal{F}_T) \prod_{t=0}^{T-1} f(\varsigma_t|\varsigma_{t+1}, \ldots, \varsigma_T; \mathcal{F}_T).$$

Conditional random vectors are generated recursively. In the forward step we run the KF and the innovations, their covariance matrix and the Kalman gain are stored. In the backwards sampling step conditional random vectors are generated recursively from $\varsigma_t|\varsigma_{t+1}, \ldots, \varsigma_T; \mathcal{F}_T$; the algorithm keeps track of all the changes in the mean and the covariance matrix of these conditional densities. The simulated disturbances are then integrated into the trend using the transition equation and a draw from $u_1^*, \ldots, u_T^*|\mathcal{F}_T$ is obtained. A computationally faster simulation smoother has been recently developed by Durbin and Koopman (2002).

5 Robust estimation of the transformation parameter and seasonal adjustment

As is well known the estimated transformation and related test statistic may be sensitive to the presence of one, or several, atypical observations. In addition, it is important to remark that outliers in one transformed scale may not be atypical in another scale. Therefore, it is important to choose a transformation which does not depend on the presence of particular observations. In this paper in order to provide a robust estimate of the transformation parameter we use the forward search approach in the way suggested by Atkinson and Riani (2000) and extended to time series by Riani (2004). Finally, in order to quantify the effect of each observation on the choice of the transformation parameter, we use the fan plot (Atkinson and Riani, 2002).

The algorithm is both efficient and robust. It is efficient because it makes use of the Gaussian likelihood machinery underlying the Kalman filter. It is robust because the outliers enter in the last steps of the procedure and their effect on the statistics of interest is clearly depicted. More generally, this approach allows evaluation of the inferential effect that each time period, either outlying or not, exerts on the fitted model.

One major advantage of the forward search over other high-breakdown techniques is that a number of diagnostic measures can be computed and monitored as the algorithm progresses. The focus of this paper is to produce forward plots of the approximate score statistic described in section 3 for testing the significance of the set of constructed variables for different values λ_0, using a separate search for each λ_0. The trajectories of the score tests can be combined in a single picture named the "fan plot". If the number of observations is not large (i.e. less than 200), generally the five most common values of λ_0 $(-1, -0.5, 0, 0.5, 1)$ are sufficient for selecting the appropriate transformation. On the other hand, when the sample size is large we have to consider a finer grid of values of λ_0. The monitoring of the fan plot for the different specifications of the seasonal component (trigonometric, HS or heteroscedastic HS) inside the basic structural model enables to appraise how robust is our estimate of the transformation parameter to the various parameterisations of γ_t.

An additional novelty of this paper is that we implement for the first time in time series the so called proportional forward search in order to have in the subset the same proportion of months which are present in the overall sample. More precisely, let R_l be the ratio between the number of seasons of kind l in the subset (m_l) and in the full sample (n_l): $R_l = m_l/n_l$, $l = 1, \ldots, s$. Given a subset of size $m = m_1, \ldots, m_s$ in every step of the search in order to determine how to progress from subset size m to subset size $m + 1$ we preliminary consider the season(s) with the smallest R_l. Among these we increase by one unit the group which has the smallest $(m_l + 1)$th ordered one step ahead standardized prediction residual. The new subset of size $m + 1$ will be formed by the $m_l + 1$ units with the smallest one step ahead standardized squared residual for the season which was least represented in the subset and by the m_j $j \neq l = 1, \ldots, s$, units with the smallest one step ahead squared standardized prediction residuals for the other seasons. In this way in each step the subset has a composition of months which reflects as much as possible the structure of the overall sample.

6 Illustrations

In this section we propose two illustrations dealing with seasonal adjustment under the square root transformation and $\lambda = \frac{1}{4}$, for which the mean and the variance of the posterior distribution of the seasonally adjusted series admit an analytic representation. These case study are used to evaluate the differences that emerge from standard additive and multiplicative seasonal adjustment, which use $\lambda = 1$ and $\lambda = 0$, respectively, and to assess the reliability of the numerical and Monte Carlo methods for evaluating (3) and (4). All the computations were performed using Ox 3.x by Doornik (2001) and the library of state space function SsfPack 2.3 by Koopman et al. (1999). The numerical integration for (8) is implemented using the QuadPack function QAGS, see Piessens et al. (1983); QuadPack is a Fortran library for univariate numerical integration ('quadrature') using adaptive rules.

6.1 Sales X data

Our first illustration deals with a well known case study, concerning the monthly sales of a engineering company (company X), from January 1965 to May 1971, that was presented and studied by Chatfield and Prothero (1973) as a case study in Box-Jenkins forecasting methods.

The plot of the series (see the first panel of Figure 3) reveals that the amplitude of the seasonal pattern is increasing over time as the trend increases, but the evidence is that the logarithmic transformation is overtransforming the series, i.e. the amplitude decreases as the trend increases on the transformed scale.

In his discussion of the Chatfield and Prothero paper Tunnicliffe-Wilson suggested the adoption of the Box-Cox transformation with parameter $\lambda = 0.34$, which he estimated by maximum likelihood using only the first 60 observations; for the full sample consisting of 77 observations Box and Jenkins (1973) using range-mean plots suggested the value 0.25, which is also the value estimated by Chatfield and Prothero in their reply. That value is confirmed by Guerrero (1993), by a different method, which looks at the variance stabilising properties of the transformation.

When the BSM is estimated under the Box-Cox transformation, the profile likelihood for the parameter λ is reported in the second panel of Figure (3). The horizontal line is drawn at $\mathcal{L}_{\max} - 0.5\chi^2(0.95)$, where $\mathcal{L}_{\max}$ is the value of the corrected profile likelihood evaluated at the maximum and $\chi^2(0.95)$ is the 95-th percentile of the χ^2 distribution with 1 degree of freedom (3.84).

The logarithmic transformation and the value $\lambda = 1$ are clearly rejected and the maximum likelihood estimate is $\tilde{\lambda} = 0.27$. It is worth noticing that the differences in results with respect to other authors can be attributed to the fact that we use a different model and that we include a calendar component in our model, which turns out to be significant. As the value 0.27 is not significantly different from 0.25, our subsequent analyzes will use the value $\lambda = 0.25$, for which, as we have seen in the previous section, the conditional mean and the variance of the seasonally adjusted series admit a closed form solution.

The maximum likelihood estimates of the variance parameters are $\tilde{\sigma}^2_\eta = 0.1108$; $\tilde{\sigma}^2_\zeta = \tilde{\sigma}^2_\omega = 0.0000$; and $\tilde{\sigma}^2_\epsilon = 0.1728$. As a result the slope is fixed and seasonality is deterministic. The Bowman and Shenton normality test takes the value 0.5; some residual correlation is left, as by the Ljung-Box portmanteau test statistic with 12 autocorrelation, which takes the value 23.34.

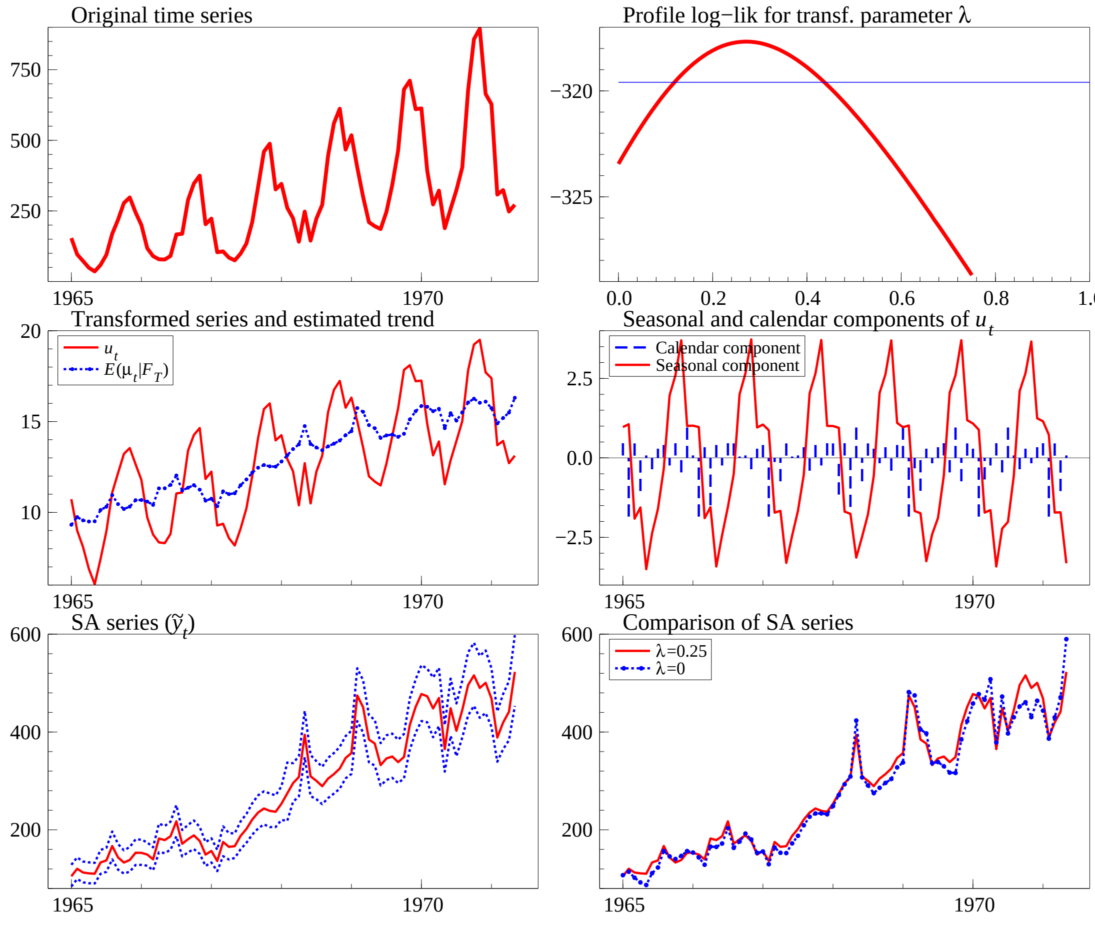
The central panels present the transformed series along with the seasonally adjusted series, $\tilde{u}^*_t$, the seasonal and calendar component on the transformed scale. The bottom panel displays the estimates of the seasonally adjusted series on the original scale, that is $\tilde{y}^*_t$ along with their 95% highest density region. The computations were made by numerical integration, but as we argue below these are undistinguishable from the exact estimates and from the Monte Carlo estimates using a suitably large number of replications. It is interesting to notice, as we have seen theoretically in the previous section, that the width of the confidence interval of the seasonally adjusted series in the original scale increases as the trend increases.

The last panel compares the estimates of the SA series arising for the estimated transformation parameter with that arising in the case of the logarithmic transformation. The graph highlights that the differences can be relevant and the Box-Cox transformation is indeed an issue in seasonal adjustment.

Given that an exact solution exists for $\tilde{y}^*_t$ and Var$(y^*_t|\mathcal{F}_T)$, we can evaluate the accuracy of the various estimates that have been proposed. The estimation methods that are compared are

- The naïve estimate (the conditional median) $\tilde{y}^*_t = (1 + 0.25\tilde{u}^*_t)^4$.
- Monte Carlo integration using the simulation smoother: M independent samples, $u^{(i)*}_t, i = 1, \ldots, M$, are drawn from the conditional distribution $u^*_t|\mathcal{F}_T \sim \mathrm{N}(\tilde{u}^*_t, V^*_t)$, which is done recursively by the simulation smoother. The seasonally adjusted series is estimated by average $\tilde{y}^{*MC}_t = \frac{1}{M}\sum_{i=1}^M \left[1 + 0.25u^{(i)*}_t\right]^4$. The variance of the SA series is estimated by $\tilde{\mathrm{Var}}(y^*_t|\mathcal{F}_T) = \frac{1}{M}\sum_{i=1}^M \left[1 + 0.25u^{(i)*}_t\right]^8 - \left(\tilde{y}^{*MC}_t\right)^2$. Results are presented for the number of replications $M = 500, 1000, 2500, 5000, 10000$. An antithetic variable was used: for each draw $\varsigma^{(i)}_t$ the antithetic variable is given by $\varsigma^{(i)\dagger}_t = 2\mathrm{E}(\varsigma_t|\mathcal{F}_T) - \varsigma^{(i)}_t$. The conditional expectation $\mathrm{E}(\varsigma_t|\mathcal{F}_T)$ is provided by the disturbance smoother (Koopman, 1993).
- Numerical integration using the QuadPack function QAGS, available in Ox 3.4; the finite

Figure 3: Sales of Company X.



integration interval is defined as $[\tilde{u}_t^* - 8\sqrt{V}_t, \tilde{u}_t^* + 8\sqrt{V}_t]$, where $\tilde{u}_t$ and V_t are evaluated by the Kalman filter and smoother applied to the transformed observations.

- The Taylor estimation method based on a Taylor's approximation:

$$\tilde{y}_t^{*T} = \hat{y}_t^* \left[1 + \frac{3}{8}\frac{V_t}{\sqrt{\hat{y}_t^*}}\right].$$

- The method proposed by Guerrero:

$$\tilde{y}_t^{*G} = \hat{y}_t^* \left\{\frac{1}{2} + \frac{1}{2}\left[1 + \frac{3}{8}\frac{V_t}{\sqrt{\hat{y}_t^*}}\right]^{1/2}\right\}^4$$

Table 2 reports the mean error of method j, $\text{ME}_j = T^{-1}\sum_{t=1}^T(\tilde{y}_t^{*j} - \tilde{y}_t^*)$, where the subtrahend is given by the exact expression given in Table 1, the mean square error, $\text{MSE}_j = T^{-1}\sum_{t=1}^T(\tilde{y}_t^{*j} - \tilde{y}_t^*)^2$, the mean percent error, $\text{MPE}_j = 100T^{-1}\sum_{t=1}^T\left[(\tilde{y}_t^{*j} - \tilde{y}_t^*)/\tilde{y}_t^*\right]$ and the mean absolute percent error, $\text{MAPE}_j = 100T^{-1}\sum_{t=1}^T\left[|\tilde{y}_t^{*j} - \tilde{y}_t^*|/\tilde{y}_t^*\right]$.

In this application the ratio $V_t/\sqrt{\hat{y}_t^*}$ is very small (6×10^{-5} on average) and thus the naïve estimate has a good performance. It should be recalled that the last two columns present percent values. It is also evident from the table that the Taylor and Guerrero approximations are very accurate for this application. Numerical integration is the most accurate; the performance of Monte Carlo integration depends on the number of replications that are used. The convergence to the true conditional mean is not very fast. This is due to the correlation between the random draws that results from the persistence of the nonseasonal component of the series. The use of an antithetic variable greatly improves the performance.

The second part of the table displays the same statistics with reference to the problem of estimating the conditional variance $\text{Var}(y_t^*|\mathcal{F}_T)$. It must be remarked that the Taylor and Guerrero methods do not provide an estimate of this feature. Again, numerical integration provides the fastest and most reliable method of estimating $\text{Var}(y_t^*|\mathcal{F}_T)$.

6.2 Italian industrial production of LS sector

Our second illustration deals with the estimation of the seasonally adjusted series and of its posterior variance according to (3) and (4) with reference to the industrial production index for the

Table 2: Sales X data: Accuracy of different estimation methods

Estimation of $\tilde{y}_t^* = \mathrm{E}(y_t^*\|\mathcal{F}_T)$				
Method	Mean Error	Mean Square Error	Mean Percent Error	MAPE
Naive	-0.48452941	0.26217669	-0.18820985	0.18820985
MC Int 500	-0.00046981	0.00169539	-0.00037667	0.01294316
MC Int 1000	-0.00587377	0.00044615	-0.00227596	0.00646891
MC Int 2500	0.00131137	0.00016612	0.00029774	0.00396751
MC Int 5000	0.00014997	0.00007092	-0.00016578	0.00243353
MC Int 10000	0.00173609	0.00003516	0.00060559	0.00175783
Num Int	0.00000000	0.00000000	0.00000000	0.00000000
Taylor	-0.00007648	0.00000001	-0.00003273	0.00003273
Guerrero	0.00003824	0.00000000	0.00001636	0.00001636
Estimation of $\mathrm{Var}(y_t^*\|\mathcal{F}_T)$				
Method	Mean Error	Mean Square Error	Mean Percent Error	MAPE
MC Int 500	20.3325	1276.3999	4.5745	6.1875
MC Int 1000	10.1550	440.8200	1.9854	3.5628
MC Int 2500	4.5773	178.4940	0.7700	1.8158
MC Int 5000	1.0162	103.1112	0.0189	1.3430
MC Int 10000	-1.1800	34.7828	-0.4323	0.9942
Num Int	0.0000	0.0000	0.0000	0.0000

Leather and Shoes (LS) sector, available for the period 1981.1-2005.2 (source Istat, base 2000 = 100, 290 observations), under the Box-Cox transformation. We notice in passing that the official seasonal adjustment performed by the Italian National Statistical Institute (Istat) is carried out on the untransformed series (i.e. $\lambda = 1$) using the software Tramo-Seats.

The plot of the original series (see the top left hand panel of Figure 4) reveals that the amplitude of the seasonal component decreases with the trend. The dominant feature is the seasonal trough occurring in August. The likelihood ratio test of $H_0 : \lambda = 1$ is significant and the maximum likelihood estimate of the transformation parameter is $\hat{\lambda} = 0.501$, corresponding to the square root transformation. The profile likelihood for the transformation parameter is plotted in the top right hand panel of Figure 4.

Clearly we have to establish whether the square root transformation is due to the presence of particular observations or it is diffused throughout the data. Finally, we need to know what is the effect on the estimated λ of the months of August or whether there are other months whose variance of the seasonal movements is much greater than the others, but are obscured by the high fluctuations of the month of August.

To start answering all these questions in Figure 5 we produce a series of fan plots for $\lambda = (0, 0.25, 0.5, 0.75)'$. The top left panel of Figure 5, which uses a trigonometric specification for the seasonal component and a non proportional forward search, shows that the log transformation is always rejected throughout.

In Figure 4 we had seen that the values of $\lambda = 0.25$ and $\lambda = 0.75$ were both at the boundary of the acceptance region. The FS enables to state that the value 0.75 is always strongly rejected throughout the search and that at the end there is a set of observations which brings the value of the score close to the acceptance region. The same upward trend is visible in the curves associated with $\lambda = 0.25$ and $\lambda = 0.5$. In the case this set of units brings the values of the score from value -3 to a value around 3.

The monitoring of the seasons inside subset (bottom right panel) clearly shows that the units entering the subset in steps 270-290 all belong to the month of August. The effect of the month of August is even more pronounced if we consider the HS specification for stochastic seasonality (see top right panel in Figure 5).

In order to understand whether this is due to seasonal heteroscedasticity we redo the fan plot allowing the variance of the month of August to be different from that of the other months. The resulting fan plot, which is given in the bottom left panel, shows that the presence of heteroscedastic seasonality for the month of August does seem to alter our conclusions about the transformation parameter.

A major benefit of the fan plot is that it clearly enables to appreciate the effect that the different months and/or different subperiods exert on the estimate of the transformation parameter. As is well known, the FS provides an ordering of the data from those most in agreement with a suggested model (which enter the first steps) to those least in agreement with it (which are included in the final steps). For example, the bottom right panel shows that the seasons which are most difficult to model are those associated with the months of November and August. However, while the effect of the introduction of the months of November (steps 230-260) does not change appreciably the value of the score test, it is clear the effect that the months of August exert on the estimated λ.

Figure 6 shows the new fan plot respectively for trigonometric (top panel) and HS specification (bottom panel) for a proportional forward search. Both plots show that if we consider subsets which contain the same proportion of months as that of the original sample the curves for the different values of λ are more stable and the values associated with the square root transformation in the central and final part of the search always lie inside the confidence bands.

The monitoring of the estimates of the hyperparameters on the square root scale (not given here for lack of space) show that in this scale the values of the variances of the underlying components are stable together with the t-statistics for the trading days and there are not sudden jumps due to the presence of atypical observations.

As a result of this analysis, the BSM was used on the transformed observations $u_t = 2(y_t^{1/2} - 1)$. The use of the trigonometric seasonal specification gave the following maximum likelihood estimates of the variance parameters $\tilde{\sigma}^2_\eta = 0.01556$; $\tilde{\sigma}^2_\zeta = 0.00003$ $\tilde{\sigma}^2_\omega = 0.00079$; and $\tilde{\sigma}^2_\epsilon = 0.05640$. There is a significant calendar component in the series, the coefficients associated to the working days being positive and those associated to the week-end being negative. The diagnostics are satisfactory, and normality is accepted (the Bowman and Shenton normality test results 3.76, with p-value=0.15).

The transformed series and its estimated components are displayed in the central panels of Figure 4. The seasonal pattern reduces its amplitude over time. The point estimates of the seasonally adjusted series on the original scale of measurement are reproduced in the bottom left panel, along with the 95% highest density region of the posterior distribution of y_t^*. The method is numerical integration as in the previous section. The last graph compares these estimates with those emerging from the BSM adapted to the untransformed series ($\lambda = 1$); on average the latter display a positive bias, but there are relevant differences that go well beyond a level change. In particular, the differences are substantial with respect to August.

Given that analytical solutions are available, $\tilde{y}_t^* = \hat{y}_t^* \left[1 + \frac{1}{4}\frac{V_t}{\hat{y}_t^*}\right]$ and $\text{Var}(y_t^*|\mathcal{F}_T) = \hat{y}_t^* V_t + \frac{1}{8}V_t^2$, we can assess the accuracy of the various estimation methods considered in the previous subsection. It must be stressed that the Taylor method gives only in this case ($\lambda = 0.5$) an exact solution. Table 3 reports the ME, the MSE, the MPE and the MAPE for the seasonally adjusted series and its variance. The most accurate method is numerical integration, which has an excellent performance also for the estimation of the conditional variance; Monte Carlo integration is more accurate than Guerrero's method and both outperform the naïve estimate.

7 The Seasonal Adjustment of the Ratio of two Time Series

The purpose of this section is twofold: to illustrate the estimation of the Box-Cox tranformation parameter for a bivariate time series and the estimation of the seasonally adjusted ratio of the two constituent series. The motivation arises in the context of the analysis of tourism trends. Statistical information on tourism in Italy concerns the monthly number of arrivals and the number of overnight stays of residents and non-residents at hotel and other establishments (camping sites, tourist farms, private accommodations, mountain huts, company vacation facilities, vacation facilities for youth, etc.). An arrival is defined as a person who arrives at a collective accommodation establishment or at private tourism accommodation and checks in. A night spent (or overnight stay) is each night that a guest actually spends or is registered in a collective accommodation establishment or in private tourism accommodation. Average stay (AS) represents the ratio of overnight stays to the number of arrivals. This variable is important for the assessment of trends in tourism.

Table 3: Index of Industrial Production, sector DC: Accuracy of different estimation methods

Estimation of $\tilde{y}_t^* = \mathrm{E}(y_t^*\vert\mathcal{F}_T)$				
Method	Mean Error	Mean Square Error	Mean Percent Error	MAPE
Naive	-0.00780061	0.00006267	-0.00779884	0.00779884
MC Int 500	0.00006245	0.00000026	0.00006111	0.00041684
MC Int 1000	-0.00002007	0.00000011	-0.00001500	0.00025849
MC Int 2500	0.00002431	0.00000005	0.00002980	0.00017916
MC Int 5000	0.00000782	0.00000003	0.00001050	0.00012541
MC Int 10000	-0.00000872	0.00000001	-0.00000713	0.00008970
Num Int	0.00000000	0.00000000	0.00000000	0.00000000
Taylor	0	0	0	0
Guerrero	-0.00390034	0.00001567	-0.00389946	0.00389946
Estimation of $\mathrm{Var}(y_t^*\vert\mathcal{F}_T)$				
Method	Mean Error	Mean Square Error	Mean Percent Error	MAPE
MC Int 500	0.0256	0.0423	0.8170	5.4020
MC Int 1000	-0.0102	0.0175	-0.1816	3.3153
MC Int 2500	0.0082	0.0081	0.2965	2.2887
MC Int 5000	0.0025	0.0043	0.1005	1.5791
MC Int 10000	-0.0039	0.0023	-0.0995	1.1266
Num Int	0.0000	0.0000	0.0000	0.0000

Let y_{1t} denote the number of arrivals and let y_{2t} that of overnight stays. These series are highly seasonal, as it can be seen from the plot of the monthly time series for hotel establishments in Italy, plotted in the first panel of Figure 7. The interest of the analysis lies in the seasonal adjustment of the two series (say y^*_{1t}, y^*_{2t}) and in the estimation of the ratio y^*_{2t}/y^*_{1t}, which is the seasonally adjusted average stay. The raw series, y_{2t}/y_{1t} is plotted in the bottom left hand panel. Also for this indicator seasonality is the most prominent source of variation. There are several ways of going about the above inferential problem; the traditional one is to adjust the three series separately, with the consequence that the estimates of the SA ratio are not related to those of y^*_{1t} and y^*_{2t}. Another suboptimal strategy is to estimate y^*_{1t} and y^*_{2t} separately and then compute the ratio of the estimates; clearly this yields biased estimates of the ratio.

Our strategy is to provide a simultaneous solution to the three problems, which entails the specification and the estimation of a bivariate BSM for the two series. The latter is such that each of the component series, after a Box-Cox transformation with common parameter λ, $u_{it} = (y^\lambda_{it} - 1)/\lambda$, $\lambda \neq 0$, and $u_{it} = \ln y_{it}$, $\lambda = 0$, $i = 1, 2$, admits an additive and orthogonal decomposition into components whose disturbances may be contemporaneously correlated. The assumption that the transformation parameter is common to the series can be tested, and removed if necessary. Denoting by bold symbols the 2×1 vector containing the elements of the two series, e.g. $\mathbf{u}_t = [u_{1t}, u_{2t}]'$, $\boldsymbol{\mu}_t = [\mu_{1t}, \mu_{2t}]'$, $\boldsymbol{\eta}_t = [\eta_{1t}, \eta_{2t}]'$, and so forth, the BSM has the following representation:

$$\mathbf{u}_t = \boldsymbol{\mu}_t + \boldsymbol{\gamma}_t + \mathbf{X}'_t\boldsymbol{\delta} + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \text{NID}(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon), \quad i = 1, 2; \quad t = 1, \ldots, T,$$

$$\begin{array}{rcccccl} \boldsymbol{\mu}_{t+1} & = & \boldsymbol{\mu}_t & + & \boldsymbol{\beta}_t & + & \boldsymbol{\eta}_t, \qquad \boldsymbol{\eta}_t \sim \text{NID}(\mathbf{0}, \boldsymbol{\Sigma}_\eta), \\ \boldsymbol{\beta}_{t+1} & = & & & \boldsymbol{\beta}_t & + & \boldsymbol{\zeta}_t, \qquad \boldsymbol{\zeta}_t \sim \text{NID}(\mathbf{0}, \boldsymbol{\Sigma}_\zeta). \end{array}$$

$$\boldsymbol{\gamma}_t = \sum_{j=1}^{6} \boldsymbol{\gamma}_{jt}, \qquad \begin{bmatrix} \boldsymbol{\gamma}_{j,t+1} \\ \boldsymbol{\gamma}^*_{j,t+1} \end{bmatrix} = \left(\begin{bmatrix} \cos\lambda_j & \sin\lambda_j \\ -\sin\lambda_j & \cos\lambda_j \end{bmatrix} \otimes \mathbf{I}_2 \right) \begin{bmatrix} \boldsymbol{\gamma}_{jt} \\ \boldsymbol{\gamma}^*_{jt} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\omega}_{jt} \\ \boldsymbol{\omega}^*_{jt} \end{bmatrix},$$

$j = 1, \ldots, 5$, and $\boldsymbol{\gamma}_{6,t+1} = -\boldsymbol{\gamma}_{6t} + \boldsymbol{\omega}_{6t}$, with $\boldsymbol{\omega}_{jt} \sim \text{NID}(\mathbf{0}, \boldsymbol{\Sigma}_\omega), \boldsymbol{\omega}^*_{jt} \sim \text{NID}(\mathbf{0}, \boldsymbol{\Sigma}_\omega), j = 1, 2, 3, 4, 5$, and $\boldsymbol{\omega}_{6t} \sim \text{NID}(\mathbf{0}, 0.5\boldsymbol{\Sigma}_\omega)$. Finally, $\mathbf{X}_t = \mathbf{x}'_t \otimes \mathbf{I}_2$, where $\mathbf{x}_t$ is a vector containing the regressors accounting for calendar effects, and $\boldsymbol{\delta} = [\boldsymbol{\delta}'_1, \boldsymbol{\delta}'_2]'$.

Analogously to the univariate case, we define the nonseasonal component in the transformed series as $\mathbf{u}^*_t = \boldsymbol{\mu}_t + \boldsymbol{\epsilon}_t$ and $\mathbf{y}^*_t = u^{-1}(\mathbf{u}^*_t)$ is the nonseasonal component in the original scale. The interest lies in

$$\mathrm{E}\left(\frac{y_{2t}^*}{y_{1t}^*}\middle|\,\mathcal{F}_T\right) = \int \frac{u^{-1}(u_{2t}^*)}{u^{-1}(u_{1t}^*)} f(u_{2t}^*, u_{1t}^*|\mathcal{F}_T) du_{1t}^* du_{2t}^* = \int \frac{y_{2t}^*}{y_{1t}^*} f(u_{2t}^*, u_{1t}^*|\mathcal{F}_T) du_{1t}^* du_{2t}^*,$$

which is the optimal estimator of the seasonally adjusted series under quadratic loss, and Var $\left(\frac{y_{2t}^*}{y_{1t}^*}\middle|\,\mathcal{F}_T\right)$.

An explicit solution exists only when both variables are transformed in logarithms: in such case, $\ln(y_{2t}^*) - \ln(y_{1t}^*)$ has a Gaussian distribution and thus the ratio y_{2t}^*/y_{1t}^* is log-normal; hence $\mathrm{E}(y_{2t}^*/y_{1t}^*|\mathcal{F}_T) = \exp\left(\tilde{u}_{2t}^* - \tilde{u}_{1t}^* + 0.5(V_{1t} + V_{2t} - 2C_{12t})\right)$; where $V_{it}, i = 1, 2$, is the conditional variance of the i-th SA series and C_{12t} is the conditional covariance between the SA series. For $\lambda \neq 0$ no closed form solution is available and we resorted to Monte Carlo integration using the simulation smoother.

When the univariate BSM is applied to each of the series separately, the maximum likelihood estimate of the transformation parameter is 0.36 for arrivals (the 95% confidence interval is [0.06, 0.68]), 0.60 for overnight stays (the 95% confidence interval is [0.36,0.84]), and 0.68 for average stays. The likelihood is however so flat in the last case that the 95% confidence interval includes both 0 and 1.

Estimation of the bivariate BSM for arrivals and stays yields a point estimate $\tilde{\lambda} = 0.62$ and a 95% interval estimate $\lambda \in [0.39, 0.84]$, which rules out both the logarithmic transformation and the analysis of the original scale. The top right panel is a plot of the profile likelihood for the λ parameter, which is assumed common to the two series. If we augment the model with two added variable, namely $\partial u_{1t}/\partial\lambda$ in the first equation, and $\partial u_{2t}/\partial\lambda$ in the second, both evaluated at $\lambda = \tilde{\lambda}$, to test that either the transformation parameter applied to the first series or that pertaining to the second differ from the maximum likelihood estimate (see section 3), the t-statistics associated to the added variables are not significant.

The parameter estimates resulted:

$$\mathbf{\Sigma}_\eta = \begin{bmatrix} 20658.9 & 37733.9 \\ 37733.9 & 96876.8 \end{bmatrix}, \quad \mathbf{\Sigma}_\zeta = \mathbf{0},$$

$$\mathbf{\Sigma}_\omega = \begin{bmatrix} 2526.9 & 5065.0 \\ 5065.0 & 13919.1 \end{bmatrix}, \quad \mathbf{\Sigma}_\epsilon = \begin{bmatrix} 0.00000331 & -0.00000000 \\ -0.00000000 & 0.01212355 \end{bmatrix}.$$

The diagnostics are satisfactory, with the normality statistics computed on the standardized innovations taking the values 6.99 and 3.30 for the two series. The plot of the seasonally adjusted series, along with their 95% highest density region, evaluated by Monte Carlo integration using the simulation smoother with $M = 10000$ replications and one antithetic variable, are presented in the central panels of Figure 7, respectively for arrivals and stays. Obviously, since the integral is computed with respect to a univariate conditional density, numerical integration may as well be used here. The evaluation of the seasonally adjusted average stays requires a two-dimensional integration over the joint posterior density $f(u_{1t}^*, u_{2t}^*|\mathcal{F}_T)$, for which the Monte Carlo method is feasible. The estimated SA average stays series is displayed in the last panel of the Figure along with its 95% highest density region and the corresponding series that would be obtained when both series are transformed into logarithms. We recall that only in this case an exact solution exists. After a period of stability or even slight increase in average stay, there has been a significant reduction over the more recent years. These tendencies are also broadly present in the series estimated imposing $\lambda = 0$, but the level and the dynamics of the estimates differ.

8 Seasonal adjustment of proportions under transformations

The Box-Cox transformation is suitable for series that have a lower bound, typically 0. For time series of proportions, that are such that the observation are constrained to lie between 0 and 1, it may not be appropriate, unless the actual values have limited spread, e.g. they range from 0.03 to 0.15, as it is usually the case for time series of unemployment rates. Several parametric transformations map the (0,1) range to the real interval; Atkinson (1985) discusses the folded power transformation (Mosteller and Tukey (1977), page 92) $u_t(\lambda) = y_t^\lambda + (1-y)^\lambda, \quad 0 < y_t < 1$ which yields the untransformed observations for $\lambda = 1$ and the logit transformation for λ approaching 0, $u_t(0) = \ln(y_t/(1-y_t))$. One serious drawback is that the transformation is not invertible, that is y_t is an implicit function of u_t. Guerrero and Johnson (1982) proposed to apply the Box-Cox tranformation to the odds ratio $p_t = y_t/(1-y_t)$, i.e. $u_t(\lambda) = (p_t^\lambda - 1)/\lambda$, which yields the logit transformation for $\lambda = 0$ and $1/(1-y_t)$ for $\lambda = 1$. The inverse transformation can be calculated explicitly, but the fact that $u_t(\lambda)$ fails to give the untransformed observations for any value of λ is often regarded as a limitation.

Aranda-Ordaz (1981, AO henceforth) proposed a class of transformation that does not suffer from the above drawbacks:

$$u_t(\lambda) = \frac{2}{\lambda}\frac{y_t^\lambda - (1-y_t)^\lambda}{y_t^\lambda + (1-y_t)^\lambda} = \frac{2}{\lambda}\frac{p_t^\lambda - 1}{p_t^\lambda + 1}, \tag{12}$$

For $\lambda \to 0$ it yields the logit transformation, $u_t(0) = \ln(y_t/(1-y_t))$, and the untransformed series for $\lambda = 1$, $u_t(1) = 2(2y_t - 1)$. The reverse transformation is:

$$y_t = \frac{1}{1+\exp(-v_t)}, v_t = \begin{cases} \frac{1}{\lambda}\ln\left(\frac{2+\lambda u_t}{2-\lambda u_t}\right), & \lambda \neq 0, \\ u_t, & \lambda = 0 \end{cases}$$

In this section we shall adopt the AO transformation; other types and generalisations are considered in Stukel (1988).

Analogously to the Box-Cox case, we define the seasonally adjusted series as the inverse transformation of the nonseasonal component u_t^*, $y_t^* = [1+\exp(-v_t^*)]^{-1}$, where $v_t^* = u_t^*$ if $\lambda = 0$ and $v_t^* = \lambda^{-1}\ln\left[(2+\lambda u_t^*)/(2-\lambda u_t^*)\right]$, otherwise. The optimal estimator of the seasonally adjusted series is given by

$$\int [1+\exp(-v_t^*)]^{-1} f(u_t^*|\mathcal{F}_T) du_t^*$$

where $f(u_t^*|\mathcal{F}_T)$ is the density function of a Gaussian distribution with mean $\tilde{u}_t^*$ and variance V_t, both evaluated by the Kalman filter and smoother for the linear model on the transformed scale. No analytical solution is available unless $\lambda = 1$. We thus estimate the seasonally adjusted series via numerical and Monte Carlo integration.

Our illustrative example is the monthly proportion of tourist arriving at hotels; this can be considered as the market share of hotels with respect to the totality of private and public accommodation establishments. The plot of the original time series, see the first panel of Figure 8, shows that this proportion is highly seasonal, the seasonal trough occurring in August, and slightly declining over the most recent years. Seasonal adjustment is crucial for assessing the underlying tendencies.

Estimation of the BSM under the AO transformation yields a point estimate $\tilde{\lambda} = 0.65$, which is significantly different from 0 (logit) and from 1 (untransformed series); see the second panel of Figure 8. The maximum likelihood estimates of the variance parameters of the BSM are $\tilde{\sigma}_\eta^2 = 0.00010209$; $\tilde{\sigma}_\zeta^2 = 0, \tilde{\sigma}_\omega^2 = 0.00001312$; and $\tilde{\sigma}_\epsilon^2 = 0.00044716$. There is no significant effect connected to the number of days of the week, but it is interesting to report that the Easter effect is significant and it is adverse to the share of tourist arrivals absorbed by hotels. This is in line with our expectations, since Easter is a period when preferences shift towards tourist farms and private accommodations. The diagnostics are satisfactory and normality is not rejected.

The seasonally adjusted proportion of arrivals in Hotels is displayed, along with its 95% highest density region, in the third panel of Figure 8. The estimation method that was used is numerical integration. For the sake of comparison the plot also reports the estimate arising when $\lambda = 1$ (no transformation). Again, the two estimates differ non only for a level shift, but also the dynamics are affected; namely, the decline in the seasonally adjusted proportion is slower when $\lambda = 1$.

9 Conclusive remarks

This paper has investigated the issue of seasonal adjustment under the Box-Cox power transformation of time series measured on a ratio scale, which are bounded from below by 0, and the Aranda-Ordaz transformation for time series of proportions.

The rationale behind the transformation is to enhance several desirable features of the maintained measurement model: linearity, additivity and orthogonality of components, normality of the disturbances driving the components.

In this paper we have concentrated on the Box Cox transformation applied to the basic structural model. However the idea of imposing the seasonal constraint on the transformed scale, perform seasonal adjustment and then transforming back into the original series can be applied to more complicated models like the so called transformation/weighting (see for example Carroll and Ruppert, 1991) models, where not only the response is transformed, but also the part used to fit the mean model and the disturbance term to take into account heteroscedasticity. The extension to other parametric classes that are continuous in the transformation parameters and invertible is straightforward. Continuity is required for likelihood based inferences on the transformation parameter; invertibility is necessary to re-express the nonseasonal component on the original scale.

This paper has documented that adjustment is both feasible and relevant. It is feasible, since there are computationally efficient and accurate methods of estimating the conditional mean and variance of the seasonally adjusted series that are applicable in the absence of a closed form solution. It is relevant, since the estimates may differ relevantly from those obtained using either the untransformed observations or the logarithms. Our case studies concerned cases when seasonality is the most prominent source of variation of the data as it occurs for industrial production, tourism and sales.

For univariate analysis numerical integration is both fast and reliable and it is recommended in the place of approximate methods and Monte Carlo integration using the simulation smoother. The latter may require a large number of replications when the nonseasonal component is highly persistent and weakly evolutive. However, it can be made as accurate as needed by increasing the number of replications and by using variance reduction techniques.

Finally it must be remarked that if in univariate models it is possible to choose between different solutions, in multivariate applications the use of the simulation smoother is the unique option available. A multivariate application that we have in mind is indirect seasonal adjustment when a cross sectional seasonally adjusted aggregate is obtained as $Y_t^* = \sum_{i=1}^{N}(1 + \lambda_i u_{it}^*)^{1/\lambda_i}$. Then $\mathrm{E}(Y_t^*|\mathcal{F}_T)$ can be evaluated by Monte Carlo integration using the simulation smoother.

The focus of this paper was seasonal adjustment. However, our method can be easily extended to find the estimate of all the other components on the original scale (e.g. the detrended series). In other words, once the two conditional moments of the detrended series in the transformed scaled are found using the KFS, the detrended series on the original scale can be computed using numerical or Monte Carlo integration or the exact analytic solution described in the paper.

A relevant topic that we did not address concerns the assessment of the reliability of the seasonally adjusted series, taking into account the additional source of uncertainty determined by the selection of the transformation parameter from the data. In fact, all the proposed inferences were conditional on the scale selected. In the context of regression analysis Bickel and Doksum (1981) showed that the spread of the marginal distribution of the estimators of the regression parameters is much larger than that of the conditional distribution, given the estimated transformation parameter. In the wake of this results one might want to investigate and quantify the increase in the variance due to the transformation parameter uncertainty. Actually, this point is highly controversial. However Carroll and Ruppert (1981) give a general result which indicates that the cost of estimating extra nuisance parameters such as λ for prediction is not large. Furthermore, Box and Cox (1982) and Hinkley and Runger (1984) argue that the variance inflation is irrelevant and illusory, as the linear model parameters have meaning only with reference to a particular scale and thus all relevant inferences can only be conditional on the selected transformation parameter.

Appendix: Proof of theorem 1

We start considering the Taylor series expansion of the inverse transformation $(1+\lambda u_t^*)^{1/\lambda}$ around $\tilde{u}_t^*$:

$$\begin{aligned}(1+\lambda u_t^*)^{1/\lambda} &= (1+\lambda\tilde{u}_t^*)^{1/\lambda} + (1+\lambda\tilde{u}_t^*)^{1/\lambda-1}(u_t^*-\tilde{u}_t^*) + \\ &\quad \tfrac{1}{2}(1-\lambda)(1+\lambda\tilde{u}_t^*)^{1/\lambda-2}(u_t^*-\tilde{u}_t^*)^2 + \cdots \\ &= \hat{y}_t^*\left[1+\textstyle\sum_{j=1}^{\infty}\frac{1}{j!}\left(\prod_{k=1}^{j-1}(1-\lambda k)\right)(\hat{y}_t^*)^{-\lambda j}(u_t^*-\tilde{u}_t^*)^j\right]\end{aligned}$$

Now, taking the expectation of both sides with respect to the Gaussian density $f(u_t^*|\mathcal{F}_T)$ and remembering that the central j-order moment is zero if j is odd, after some manipulation we obtain that:

$$E(y_t^*|\mathcal{F}_T) = \tilde{y}_t^* = \hat{y}_t^*\left[1+\sum_{j=1}^{\infty}\frac{1}{j!2^j}\left(\prod_{k=1}^{2j-1}(1-\lambda k)\right)\frac{V_t^j}{\hat{y}_t^{*2\lambda j}}\right] \tag{13}$$

If we denote with

$$a_j(t) = \frac{(2j)!}{j!2^j}V_t^j \qquad \text{and} \qquad k_j(t) = \frac{1}{j!}\left(\prod_{k=1}^{j-1}(1-\lambda k)\right)\hat{y}_t^{*-\lambda j},$$

equation (13) can be rewritten as:

$$\mathrm{E}(y_t^*|\mathcal{F}_T) = \tilde{y}_t^* = \hat{y}_t^*\left[1+\sum_{j=1}^{\infty}k_{2j}(t)a_j(t)\right] \tag{14}$$

The second noncentral moment is given by

$$\begin{aligned}\mathrm{E}(y_t^{*^2}|\mathcal{F}_T) &= \hat{y}_t^{*^2}\mathrm{E}\left[1+\sum_{j=1}^{\infty}k_j(t)(u_t^*-\tilde{u}_t^*)^j\right]^2 \qquad (15)\\ &= \hat{y}_t^{*^2}\mathrm{E}\left[1+\sum_{j=1}^{\infty}k_j^2(t)(u_t^*-\tilde{u}_t^*)^{2j}+2\sum_{j=1}^{\infty}\sum_{r>j}k_j(t)k_r(t)(u_t^*-\tilde{u}_t^*)^{j+r}+2\sum_{j=1}^{\infty}k_j(t)(u_t^*-\tilde{u}_t^*)^j\right]\end{aligned}$$

Taking the expectation of both sides with respect to the Gaussian density $f(u_t^*|\mathcal{F}_T)$ we obtain:

$$\mathrm{E}(y_t^{*^2}|\mathcal{F}_T) = \hat{y}_t^{*^2}\left[1+\sum_{j=1}^{\infty}k_j^2(t)a_j(t)+2\sum_{j=1}^{\infty}\sum_{r=1}^{\infty}k_j(t)k_{j+2r}(t)a_{j+r}(t)+2\sum_{j=1}^{\infty}k_{2j}(t)a_j(t)\right] \tag{16}$$

On the other hand, the square of the first moment can be written as:

$$[E(y_t^*|\mathcal{F}_T)]^2 = \hat{y}_t^{*^2}\left[1+(\sum_{j=1}^{\infty}k_{2j}(t)a_j(t))^2+2\sum_{j=1}^{\infty}k_{2j}(t)a_j(t)\right].$$

After some manipulations we obtain that:

$$\text{Var}(y_t^*|\mathcal{F}_T) = \hat{y}_t^{*^2} \left[\sum_{j=1}^{\infty} k_j^2(t) a_j(t) + 2 \sum_{j=1}^{\infty} \sum_{r=1}^{\infty} k_j(t) k_{j+2r}(t) a_{j+r}(t) - \left(\sum_{j=1}^{\infty} k_{2j}(t) a_j(t) \right)^2 \right]. \tag{17}$$

Acknowledgements

The authors are grateful A.C. Atkinson, W.R. Bell, V.M. Guerrero and two anonymous referees for critically reading the manuscript and providing valuable suggestions which have led to an improvement in the presentation.

References

Aranda-Ordaz F. J. (1981). On two families of transformations to additivity for binary response data, *Biometrika*, 68, 357–363.

Atkinson, A. C. (1985). *Plots, Transformations and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis*, New York: Oxford University Press.

Atkinson, A. C. and Riani M. (2000). *Robust Diagnostic Regression Analysis*, New York: Springer Verlag.

Atkinson, A. C. and Riani M. (2002). Tests in the fan plot for robust, diagnostic transformations in regression, *Chemometrics and Intelligent Laboratory Systems*, 60, 87–100.

Bell W. R. and Hillmer S.C. (1984). Issues involved with the seasonal adjustment of economic time series, *Journal of Business and Economic Statistics*, 2, 291–320.

Bickel P. J. and Doksum K. A. (1981). An analysis of transformations revisited. *Journal of the American Statistical Association*. 76, 296-311.

Box G. E. P. and Cox D.R. (1964). An analysis of transformations (with discussion), *Journal of the Royal Statistical Society, B*, 26, 211–246.

Box, G.E.P. and Cox D.R. (1982). An analysis of transformation revisited, rebutted. *Journal of the American Statistical Association*, 77, 209–210.

Box G. E. P. and Jenkins, G.M. (1973). Some Comments on a Paper by Chatfield and Prothero and on A Review by Kendall. *Journal of the Royal Statistical Society, A*, 136, 337-352.

Carlin, B.P., Polson, N.G. and Stoffer, D.S. (1992). A Monte Carlo Approach to Nonnormal and Nonlinear State Space Modeling. *Journal of the American Statistical Association*, 87, 493–500.

Carroll, R.J. and Ruppert, D. (1981). On Prediction and the Power Transformation Family. *Biometrika*, 68, 609-615.

Carroll, R.J. and Ruppert, D. (1991). Prediction and Tolerance Intervals With Transformation and/or Weighting. *Technometrics*, 33, 197-210.

Carter, C.K. and Kohn, R. (1994). On Gibbs Sampling for State Space Models. *Biometrika*, 81, 541-553.

Chatfield C., and Prothero, D. L. (1973). Box-Jenkins seasonal forecasting: problems in a case-study. *Journal of the Royal Statistical Society, A*, 136, 295–315.

Cleveland, W.S., and Devlin, (1982). Calendar Effects in Monthly Time Series: Modeling and Adjustment. *Journal of the American Statistical Association,* 77, 520–528.

Cleveland, W.S., Dunn, D.M. and Terpenning, I.J. (1978). SABL – A Resistant Seasonal Adjustment Procedure with Graphical Methods for Interpretation and Diagnosis. In A. Zellner, editor, *Seasonal Analysis of Economic Time Series*, 201-231. U.S. Department of Commerce, Bureau of the Census, Washington, D.C.

Cook, R. D. and Weisberg, S. (1999). *Applied Regression Including Computing and Graphics*, New York: Wiley.

de Jong P. (1991). The Diffuse Kalman Filter, Annals of Statistics, 19, 1073-1083.

de Jong P. and Shephard N.G. (1995). The simulation smoother for time series models, *Biometrika*, 82, 339–350.

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

Durbin, J., and Koopman S.J. (2001). *TimeSeries Analysis by State Space Methods*. Oxford University Press, Oxford.

Durbin, J., and Koopman S.J. (2002). A simple and efficient simulation smoother for state space time series analysis *Biometrika*, 89, 603-615.

Findley D.F., Monsell B.C., Bell W.R., Otto M.C. and Chen B.C. (1998). New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program (with Discussion), *Journal of Business and Economic Statistics*, 16, 127-177.

Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models, *Journal of Time Series Analysis*, 15, 183-202.

Gamerman, D. (1998). Markov chain Monte Carlo for dynamic generalised linear models. *Biometrika*, 85, 215-227.

Gómez V. and Maravall A. (1997). *Programs TRAMO and SEATS: Instructions for the User (Beta Version: June 1997)*, Working Paper 97001, Ministerio de Economia y Hacienda, Direccion General de Analisis y Programacion Presupuestaria, Madrid, Spain.

Granger, C. W. J. and Newbold, P. (1976). Forecasting transformed time series. *Journal of the Royal Statistical Society, B*, 38, 189-203.

Guerrero, V. M. (1993). Time-series analysis supported by power transformations, *Journal of Forecasting*, 12, 37–48.

Guerrero V.M. and Johnson R.A. (1982). Use of the Box Cox transformation with binary response models, *Biometrika*, 69, 309–314.

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

Harvey A. C. and Proietti T. (2005). *Readings in Unobserved Components Models*, Advanced Texts in Econometrics, Oxford University Press, Cambridge UK.

Hinkley D. V. and Runger G. (1984). The analysis of transformed data. *Journal of the American Statistical Association*, 79, 302-320.

John J. A. and Draper N.R. (1980). An alternative family of transformations, *Applied Statistics*, 29, 190–197.

Koopman, S. J. (1993). Disturbance smoother for state space models. *Biometrika*, 80, 117-26.

Koopman, S. J. (1997). Exact initial Kalman filter and smoother for non-stationary time series models. *Journal of the American Statistical Association*, 92, 1630-1638.

Koopman, S.J., Shephard, N. and Doornik, J.A. (1999). Statistical algorithms for models in state space using SsfPack 2.2. *Econometrics Journal*, 2, 113-166.

Mosteller F. and Tukey J.W. (1977), *Data analysis and regression*, Addison-Wesley, Reading Mass..

Neyman J.N. and Scott E.L. (1960). Correction for bias introduced by a transformation of variables. The Annals of Mathematical Statistics. 31, 3, pp. 643-655.

Pankratz, A. and Dudley, U. (1987). Forecasts of power-transformed series. *Journal of Forecasting*, 6, 239-248.

Piessens, R., de Doncker-Kapenga, E., Uberhuber, C. and Kahaner, D. (1983) *QUADPACK: A Subroutine Package for Automatic Integration*. Berlin: Springer.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T. (1992). *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge, UK.

Proietti T. (1998) Seasonal heteroscedasticity and trends. *Journal of Forecasting*, 17, 1-17.

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

Riani, M. (1998) Weights and Robustness of Model Based Seasonal Adjustment. *Journal of Forecasting*, 17, 1, 19-34.

Riani, M. (2004) Extensions of the Forward Search to Time Series. *Studies in Non Linear Dynamics and Econometrics*, 8, 2, 1-23.

Shulman H.B. and McKenzie S.K. (1984), *A study of pre-adjustment transformations*, Statistical Research Division Report Series number CENSUS/SRD/RR-84/34.

Stukel T.A. (1988) Generalised logistic models. *Journal of the American Statistical Association*, 83, 426-431.

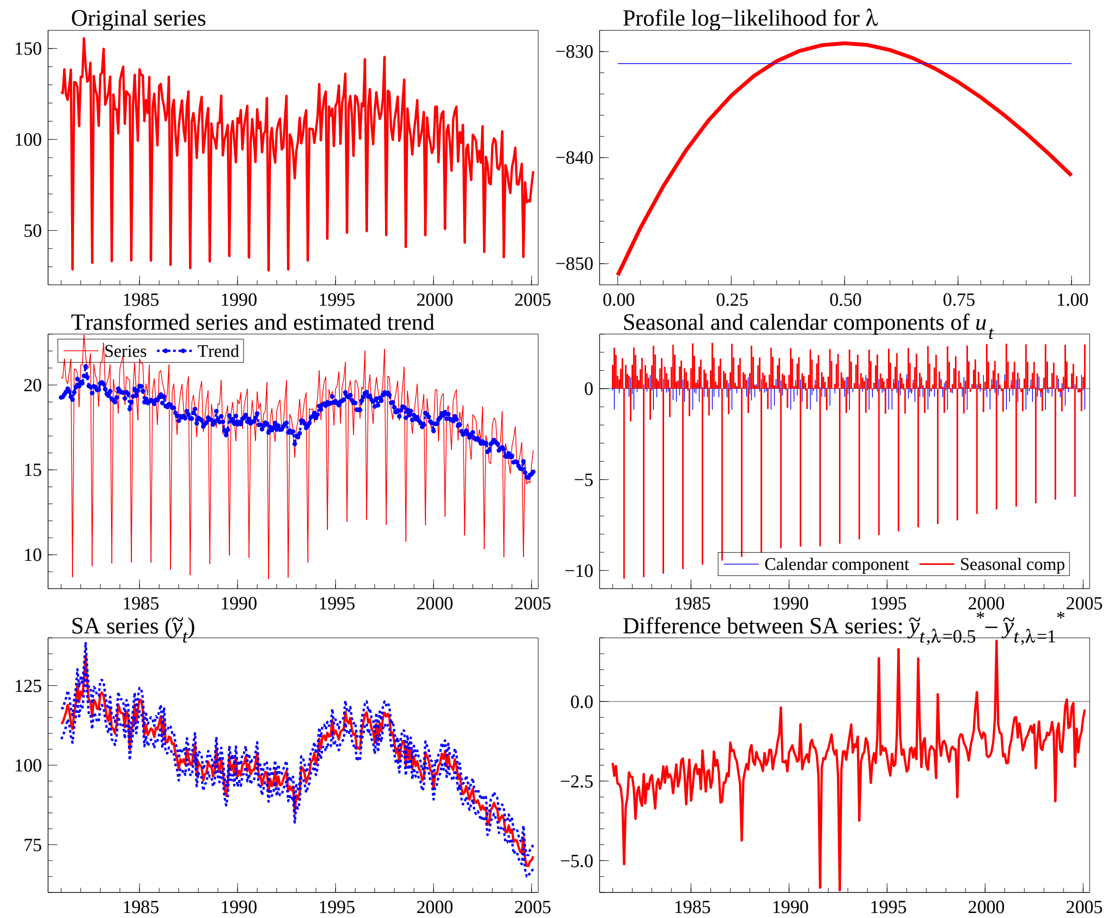
Taylor, J. M. G. (1986) The retransformed mean after a fitted power transformation. *Journal of the American Statistical Association*, 81, 114-118.

Thomson P. and Ozaki, T. (2002) Transformation and trend-seasonal decomposition, Invited paper. *Proceedings of the 3rd International Symposium on Frontiers of Time Series Modeling, Tokyo, Japan,*, 197–212.

Tukey, J. W. (1957) On the comparative anatomy of transformations. *Annals of Mathematical Statistics*, 28, 602-632.

Yeo, I. K. and Johnson, R. A. (2000) A new family of power transformation to improve normality or symmetry, *Biometrika*, 87, 954-959.

Figure 4: Index of Industrial Production, Sector DC: Leather and Shoes.



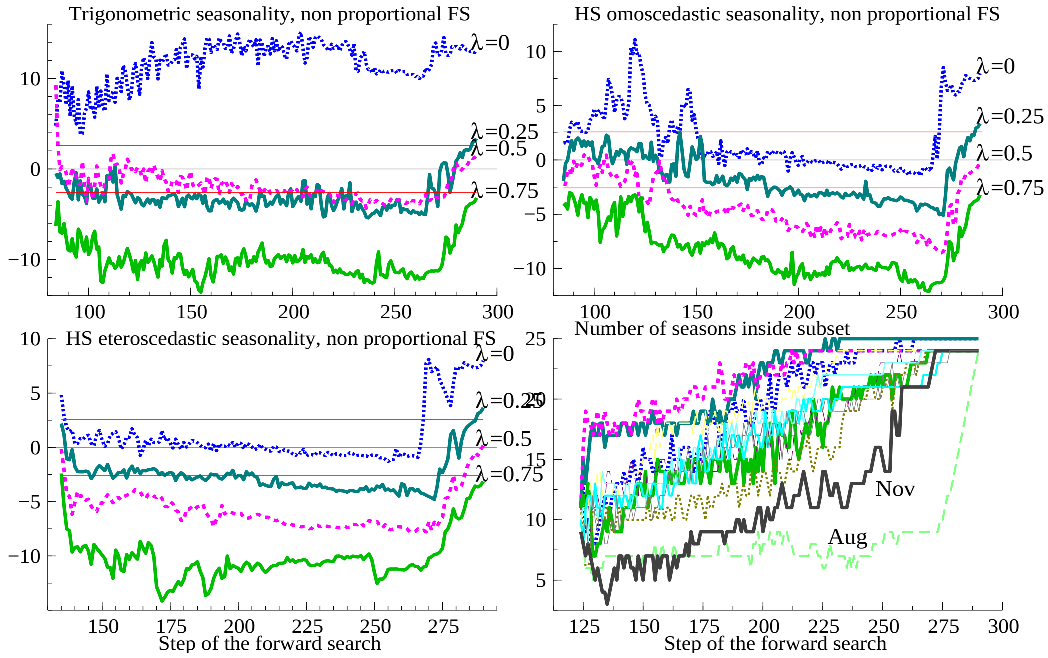


Figure 5: Robust estimate of the transformation parameter for different specifications of the seasonal component using a non proportional FS.

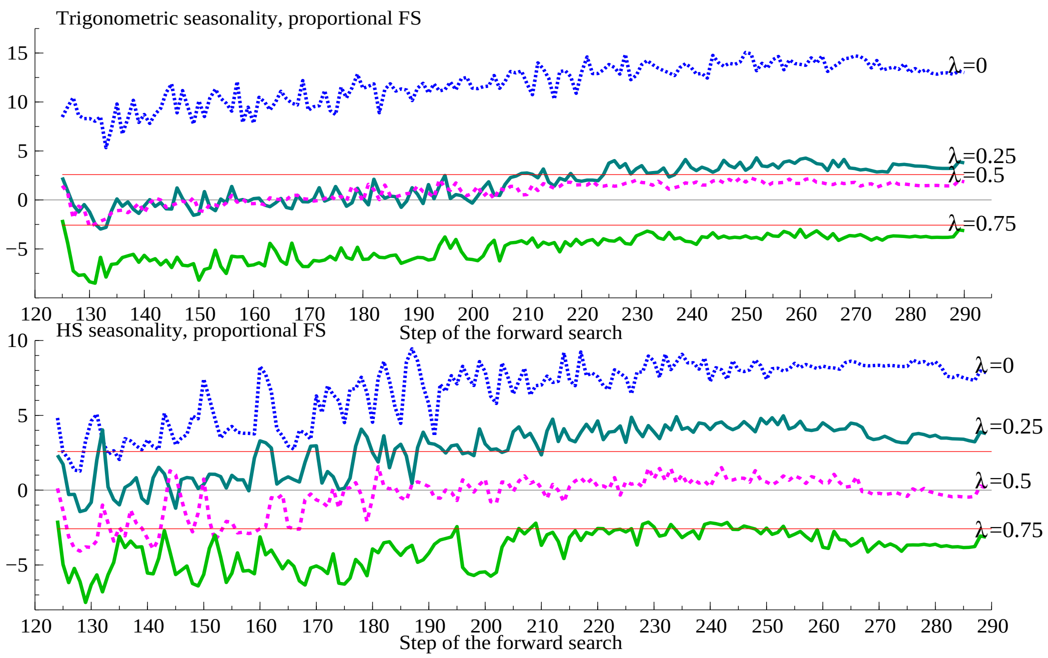


Figure 6: Robust estimate of the transformation parameter for different specifications of the seasonal component using a proportional FS.

Figure 7: Seasonal adjustment of the ratio of two time series: average stay at hotel establishments, Italy, 1997.1-2005.10.

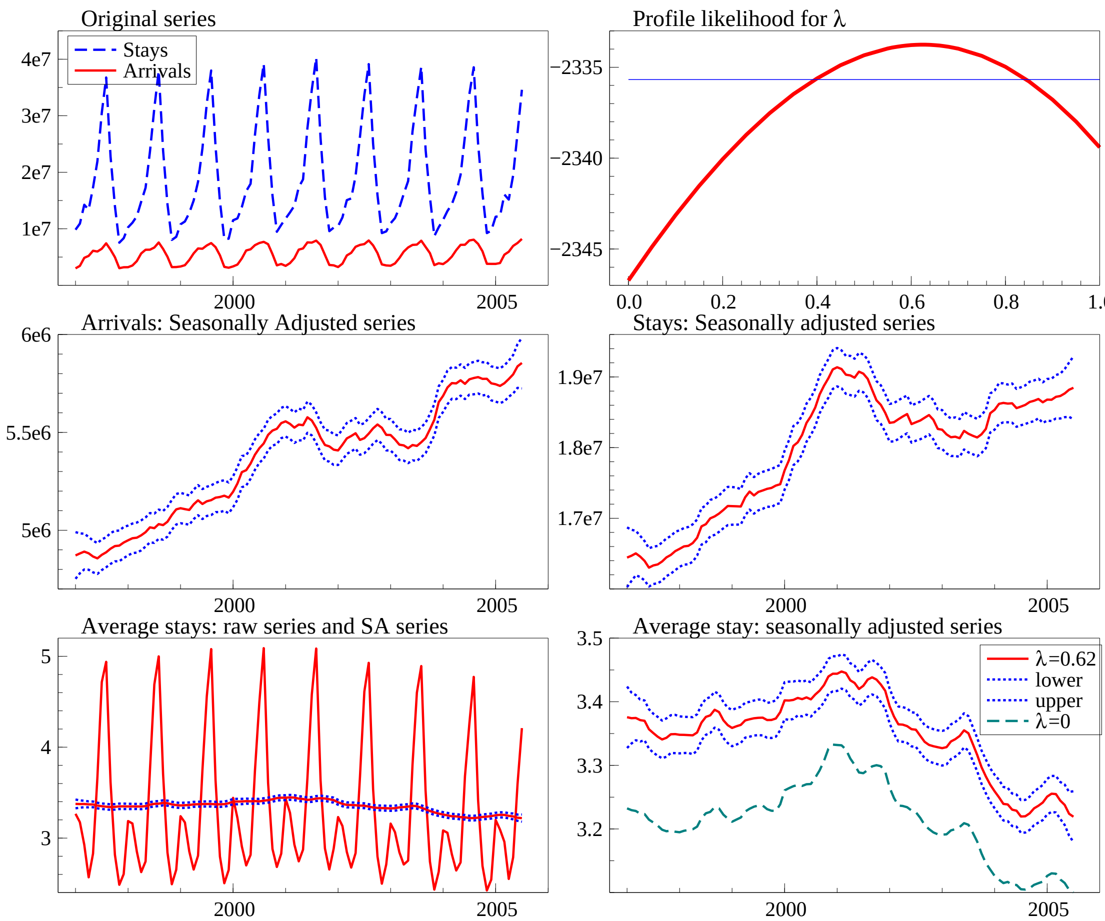


Figure 8: Proportion of total tourist arrivals in Hotels, Italy, 1997.1-2005.10.