EUROPEAN UNIVERSITY INSTITUTE
DEPARTMENT OF ECONOMICS

EUI Working Paper **ECO** No. 2002/10

Seasonal Specific Structural Time Series Models

Tommaso Proietti

BADIA FIESOLANA, SAN DOMENICO (FI)

Published in Italy in May 2002
European University Institute
Badia Fiesolana
I-50016 San Domenico (FI)
Italy

Seasonal Specific Structural Time Series Models

Tommaso Proietti*

Dipartimento di Scienze Statistiche, Università di Udine and

Dept. of Economics, European University Institute

Abstract

This paper introduces the class of seasonal specific structural time series models, according to which each season follows specific dynamics, but is also tied to the others by a common random effects. This results in a dynamic variance components model that can account for some kind of periodic behaviour, such as periodic heteroscedasticity, and is tailored to deal with situations when one or a group of seasons behave differently. Trends and non periodic features can be still be extracted and their nature is discussed. Multivariate extensions entertain the case when cointegration pertains only to groups of seasons. We finally show that a circular correlation model for the idiosyncratic disturbances yields a periodic component that is isomorphic to a trigonometric seasonal component.

Keywords: Seasonal Heteroscedasticity; Circular Correlation; Unobserved Components; Seasonal Adjustment; Seasonal Cointegration; Periodic Processes.

*Address for correspondence: Department of Economics, European University Institute - Badia Fiesolana. Via dei Roccettini 9, San Domenico di Fiesole (FI) I-50016. Email: `tproiett@iue.it`

1 Introduction

The paper deals with a class of models for seasonal time series in an unobserved components framework, according to which each season follows specific dynamics but is also tied to the remaining seasons by a common disturbance.

The emphasis is on the time domain representation, for which the season are at the core of the modelling effort, rather than on the frequency domain, although the relationships between the two approaches will also be discussed, when we deal with the circular correlation model for the idiosyncratic or season specific disturbances.

In standard situations this class of models will produce nonperiodic difference stationary time series which admit the traditional decomposition into trends and a seasonal; in general, it is particularly well suited for situations in which one or a group of seasons behave differently. These are occurrences in which the constraint imposed by the trend-seasonal decomposition, namely that the latter component has a mean of zero over a number of consecutive observations equal to the seasonal period, is too binding. This is illustrated with respect to Italian industrial production for which the seasonal trough occurs in August, the traditional holiday period. If the August trough is particularly deep it will drag down the trend and the measurement of the underlying growth in the series will be affected. The rationale for introducing this class is that the information content of the seasons differs with respect to the long run behaviour of the series and if a subgroup is more variable (i.e. they behave more idiosyncratically), they should be appropriately discounted in extracting a non periodic signal that expresses the overall tendency of the series.

In a very extreme situation the value of the series in a particular season can be equal or around some fixed value (e.g. a structural zero), as in the production of some strongly seasonal items or in some historical demographic time series referring to periods of time when marriages were prohibited by religious prescriptions, so that, even if some events are observed, these hardly speak about the general dynamics of the series. In such cases the zeros can be interpreted as missing values and this is equivalent to setting the variance of the season to infinity.

The main results are presented with reference to the seasonal specific local level model, according to which each season evolves as a random walk with no drift term (section 2). We show that a decomposition is admissible into a non periodic component and a periodic component; in the definition of the latter the zero sum constraint is relaxed and the consequences are discussed (section 3). Hence, seasonal specific models introduce periodic features without affecting the possibility of extracting a non periodic signal, that provides an indication of the long run dynamics in the series.

Subsequently, this basic representation is extended to allow for the presence of slopes (section 4), yielding the so called seasonal specific local linear trend model. The latter is illustrated with reference to the Italian index of industrial production (section 5).

Multivariate extensions are provided that can deal with peculiar forms of seasonal or periodic cointegration that characterise only a subset of seasons. They move away from the usual notion of seasonal cointegration, that is defined in the frequency domain, and are illustrated with respect to a bivariate system of income and consumption in Sweden (section 7). The example shows that the lack of full seasonal cointegration can be explained mainly with the behaviour of the fourth quarter. Finally, section 8 presents some other extensions, dealing with circular correlation among the seasonal specific disturbances, and establishes the connection with the frequency domain representation of seasonality.

2 Seasonal Specific Local Level Model

Let us consider a time series, $y_t, t = 1, 2, \dots, T$, observed with periodicity s, and let $j = 1, \dots, s$, index the season to which the t-th observation refers. The *seasonal specific local level model* is formulated as follows:

$$\begin{aligned} y_t &= \mu_{jt} + \epsilon_{jt}, \quad j = 1, \cdots, s, \\ \mu_{j,t+1} &= \mu_{jt} + \eta_{jt}, \\ \eta_{jt} &= \eta_t + \eta^*_{jt}. \end{aligned} \tag{1}$$

According to (1) the seasons are characterised by a specific level, μ_{jt}, evolving as a random walk, driven by idiosyncratic disturbances, η^*_{jt}, and a common disturbance term, η_t, which bounds up their dynamics. The level is observed with superimposed noise, ϵ_{jt}, which may also have an error component structure, $\epsilon_{jt} = \epsilon_t + \epsilon^*_{jt}$, with ϵ^*_{jt} representing the idiosyncratic noise, and ϵ_t a source common to all the seasons. We assume throughout that all the disturbances are mutually independent, though the idiosyncratic ones, η^*_{jt} and ϵ^*_{jt}, may be correlated across the seasons (see section 8).

The seasonal specific local level model is closely related to the *form-free seasonal factor* dynamic linear model of West and Harrison (1997), with the relevant difference that it is further extended to allow for a common disturbance driving all the seasons and accounting for a uniform correlation among them.

Stacking the seasonal specific levels into the $s \times 1$ vector, $\boldsymbol{\mu}_t = [\mu_{1t}, \ldots, \mu_{st}]'$, the corresponding state space representation is:

$$\begin{array}{rcll} y_t & = & \boldsymbol{x}_t'\boldsymbol{\mu}_t + \epsilon_{jt}, & t = 1, \ldots, T, \\ \boldsymbol{\mu}_{t+1} & = & \boldsymbol{\mu}_t + \boldsymbol{i}\eta_t + \boldsymbol{\eta}^*_t & \mathsf{Var}(\boldsymbol{\eta}^*_t) = \boldsymbol{N} \end{array}$$

where the vector $\boldsymbol{x}_t' = [0, \ldots, 0, 1, 0, \ldots, 0]$ selects the relevant season and is characterised by the periodic property $\boldsymbol{x}_t = \boldsymbol{x}_{t-s}$; $\boldsymbol{i}$ is an $s \times 1$ vector of ones.

The measurement equation simply states that the observations arise from periodically sampling an $s \times 1$ random walk, plus a noise component. In the remainder we will assume for simplicity that $\epsilon_{jt} = \epsilon_t \sim \mathsf{NID}(0, \sigma^2_\epsilon)$, i.e. there is no seasonal idiosyncratic noise. Typically, $\boldsymbol{N}$ will be a diagonal matrix, although we can allow for correlated idiosyncratic disturbances, and a structured form for parameterising the correlation among the seasons, known as circular correlation, will be discussed in section 8; of course the model would not be identifiable if $\boldsymbol{N}$ spanned the space of $\boldsymbol{i}\boldsymbol{i}'$.

Model (1) is such that $\Delta_s y_t$, where $\Delta_s = 1 - L^s$ is the seasonal differencing operator, is periodically stationary (see Hipel and McLeod, 1994, for a review of periodic time series models), as $\Delta_s y_t = \eta_{t-1} + \eta^*_{j,t-1} + \Delta_s \epsilon_t$, where the right hand side is a zero mean periodic moving average process of order s. When the variance of the idiosyncratic disturbances is constant across the seasons, $\Delta_s y_t$ is stationary in the usual sense, i.e. there is no periodic effect.

Traditional models of seasonality (see Proietti, 2000) assume that $\boldsymbol{N}$ lies in the nullity of $\boldsymbol{i}$. For instance, in the Harrison and Stevens (1976, HS henceforth) seasonal model it is proportional to the matrix $\boldsymbol{I}_s - \boldsymbol{i}\boldsymbol{i}'/s$. In the next section we discuss the possibility of extracting periodic and a non periodic signals, highlighting the similarities and the differences with more traditional decompositions into trends and seasonals.

3 Orthogonal decomposition into periodic and non-periodic components

When $\boldsymbol{N}$ lies in the nullity of $\boldsymbol{i}$, y_t the seasonal specific levels can be decomposed into a common level component, and a purely seasonal component, which arise respectively from the orthogonal projection of $\boldsymbol{\mu}_t$ on the subspaces spanned by $\boldsymbol{i}$ and $\boldsymbol{I}_s - s^{-1}\boldsymbol{i}\boldsymbol{i}'$.

We show that a non periodic component, describing the long run evolution of the series devoid of periodic features can be extracted in the more general framework provided by (1) when $\boldsymbol{N}$ is non singular. In particular, we can decompose $\boldsymbol{\mu}_t$ into an overall non periodic (NP) component and a periodic (P) component, which allows to rewrite (1) as follows:

$$\begin{aligned} y_t &= \bar{\mu}_t + \bar{\gamma}_t + \epsilon_t, \\ \bar{\mu}_t &= \boldsymbol{w}'\boldsymbol{\mu}_t, \\ \bar{\gamma}_t &= \boldsymbol{x}_t'\boldsymbol{\gamma}_t, \quad \boldsymbol{\gamma}_t = (\boldsymbol{I}_s - \boldsymbol{i}\boldsymbol{w}')\boldsymbol{\mu}_t. \end{aligned} \tag{2}$$

Here $\bar{\mu}_t$ denotes the NP component, whereas $\bar{\gamma}_t$ is the periodic one; both are defined in terms of weighted linear combinations, respectively of rank 1 and $s-1$ of the vector containing the seasonal specific levels $\boldsymbol{\mu}_t$.

The NP component results from the contemporaneous aggregation of $\boldsymbol{\mu}_t$ with weights provided by:

$$\boldsymbol{w} = \frac{\boldsymbol{N}^{-1}\boldsymbol{i}}{\boldsymbol{i}'\boldsymbol{N}^{-1}\boldsymbol{i}}, \qquad \boldsymbol{w}'\boldsymbol{i} = 1; \tag{3}$$

by definition, the elements of $\boldsymbol{w}$ sum up to one. The transition equation for $\bar{\mu}_t$ is established by multiplying both sides of that for $\boldsymbol{\mu}_t$ by $\boldsymbol{w}'$ and noticing $\boldsymbol{w}'\boldsymbol{i} = 1$:

$$\bar{\mu}_{t+1} = \bar{\mu}_t + \eta_t + \boldsymbol{w}'\boldsymbol{\eta}_t^*. \tag{4}$$

Hence the NP component is a univariate random walk driven by two sources of variation: the common disturbance η_t and a weighted average of the disturbances specific to each season; its size is thus

$$\mathsf{Var}(\Delta\bar{\mu}_{t+1}) = \sigma^2_\eta + \boldsymbol{w}'\boldsymbol{N}\boldsymbol{w} = \sigma^2_\eta + (\boldsymbol{i}'\boldsymbol{N}^{-1}\boldsymbol{i})^{-1}.$$

Writing $\boldsymbol{\mu}_t = \boldsymbol{i}\boldsymbol{w}'\boldsymbol{\mu}_t + (\boldsymbol{I} - \boldsymbol{i}\boldsymbol{w}')\,\boldsymbol{\mu}_t$, and defining $\boldsymbol{\gamma}_t = (\boldsymbol{I}_s - \boldsymbol{i}\boldsymbol{w}')\,\boldsymbol{\mu}_t$, the periodic component, $\bar{\gamma}_t = \boldsymbol{x}_t'\boldsymbol{\gamma}_t$, is generated by systematically sampling the singular multivariate random walk:

$$\boldsymbol{\gamma}_{t+1} = \boldsymbol{\gamma}_t + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t = (\boldsymbol{I} - \boldsymbol{i}\boldsymbol{w}')\,\boldsymbol{\eta}^*_t,$$

with disturbance covariance matrix

$$\boldsymbol{\Omega} = \mathsf{Var}(\boldsymbol{\omega}_t) = \boldsymbol{N} - \frac{\boldsymbol{i}\boldsymbol{i}'}{\boldsymbol{i}'\boldsymbol{N}^{-1}\boldsymbol{i}},$$

that is singular, since

$$\boldsymbol{\Omega}\boldsymbol{w} = \boldsymbol{0},$$

as it is easily checked. Finally,

$$\mathsf{Cov}(\boldsymbol{\omega}_t, \eta_t + \boldsymbol{w}'\boldsymbol{\eta}^*_t) = \boldsymbol{0},$$

so that NP and P define orthogonal components.

Hence, result (2) is based on an oblique projection of $\boldsymbol{\mu}_t$ using the projection matrices $\boldsymbol{i}\boldsymbol{w}'$ and $\boldsymbol{I}_s - \boldsymbol{i}\boldsymbol{w}'$.

As we have seen, the NP component is defined as a weighted average of the season specific trends, so that if $\boldsymbol{N}$ is diagonal, more variable seasons will be downweighted; as a matter of fact, the weights will be inversely related to the variance of the idiosyncratic disturbances.

We have used the term periodic component rather than seasonal component since, in situations like those illustrated in section (5), $S(L)\bar{\gamma}_t$, where $S(L) = 1 + L + \cdots + L^{s-1}$ is the seasonal summation operator, is not a zero mean MA(q) process, with $q \leq s - 2$, as it holds for traditional seasonal models (see Proietti, 2000). This point is further discussed in section 3.1.

The traditional trend-seasonal decomposition where the trend is defined as a simple average of the seasonal specific random walks, $\bar{\mu}_t = \boldsymbol{i}'\boldsymbol{\mu}_t/s$, is also an option, but this will lead to correlated components. The P component would be a genuine seasonal component, with the property $S(L)\bar{\gamma}_t \sim MA(q)$, which arises as a consequence of $\boldsymbol{\Omega i} = \mathbf{0}$, but the resulting trend would be more variable than that implied by the decomposition (2); this is so because the contribution of the seasonal specific disturbances to the variance of the changes in the NP component would be $s^{-2}\boldsymbol{i}'\boldsymbol{N}\boldsymbol{i}$, rather than $(\boldsymbol{i}'\boldsymbol{N}^{-1}\boldsymbol{i})^{-1}$, and it is easy to demonstrate that the former is greater when $\boldsymbol{N}$ is not a scalar matrix. Consider for simplicity the case in which $\boldsymbol{N}$ is diagonal: $(\boldsymbol{i}'\boldsymbol{N}^{-1}\boldsymbol{i})^{-1}$ is the harmonic mean of the elements of $\boldsymbol{N}$ scaled by $1/s$, whereas $s^{-2}\boldsymbol{i}'\boldsymbol{N}\boldsymbol{i}$ is the arithmetic mean scaled by the same factor, $1/s$. It is well know that the former is always dominated by the latter so $s^{-2}\boldsymbol{i}'\boldsymbol{N}\boldsymbol{i} \geq (\boldsymbol{i}'\boldsymbol{N}^{-1}\boldsymbol{i})^{-1}$. According to the orthogonal decomposition (2), a disturbance that is idiosyncratic to a specific season is projected along the subspace $\boldsymbol{i}$, spanning the space of the trend common to all the seasons, in the direction $\boldsymbol{w}$, rather than orthogonally.

If $\boldsymbol{N} = \sigma^2_{\eta*}\boldsymbol{I}_s$, the model is in fact a variant of the basic structural model (BSM, Harvey, 1989), according to which the series is decomposed into a random walk trend and a seasonal component with a Harrison and Stevens (1976) representation. As a matter of fact, $\boldsymbol{w} = \frac{1}{s}\boldsymbol{i}$ and $\boldsymbol{\Omega} = \boldsymbol{I}_s - s^{-1}\boldsymbol{i}\boldsymbol{i}'$. Note, however, that the unweighted average the idiosyncratic disturbances $\boldsymbol{\eta}^*_t$ enters the trend equation: $\bar{\mu}_{t+1} = \bar{\mu}_t + \eta_t + s^{-1}\boldsymbol{i}'\boldsymbol{\eta}^*_t$, with disturbance variance $\sigma^2_\eta + \sigma^2_{\eta^*}/s$. Therefore, the seasonal specific model is strictly slightly different from the BSM because of the presence of this feature, although the periodic component is a pure seasonal component and is orthogonal from the non periodic one which can be called a trend by all means. In the BSM with HS seasonality the seasonal disturbances are produced in a space orthogonal to the trend (see also the discussion at the end of section 5).

When $\boldsymbol{N}$ is a rank zero matrix the model collapses to a local level model with deterministic seasonality; in effect, we can write: $\boldsymbol{\mu}_t = \boldsymbol{i}\bar{\mu}_t + \boldsymbol{\mu}_\theta$, where $\boldsymbol{\mu}_\theta$ is a $s \times 1$ vector orthogonal to $\boldsymbol{i}$ and $\bar{\mu}_{t+1} = \bar{\mu}_t + \eta_t$ is the trend common to all the seasons.

The orthogonal decomposition was defined assuming a non singular matrix $\boldsymbol{N}$. However, some degeneracies can be easily accommodated. Assume, for instance, that $\boldsymbol{N}$ is a diagonal matrix with some zero elements, e.g. $\sigma^2_{\eta^* j} = 0$ for some j. In this case the non zero weights correspond to the zero elements in the diagonal of $\boldsymbol{N}$; genuine information about the NP component is provided by those seasons in which only the common disturbance enter the transition equation. For instance, if $\boldsymbol{N} = \text{diag}(\boldsymbol{0}_{s_1}, \sigma^2 \boldsymbol{I}_{s_2})$ then, $\boldsymbol{w} = [\boldsymbol{i}_{s_1}/s_1, \boldsymbol{0}]'$, where $\boldsymbol{i}_{s_1}$ is an $s_1 \times 1$ vector of ones. In the extreme case when only one season belongs to the first group, the NP component is defined only in terms of this season, since it is revealed in that season.

3.1 Mean Correction and Seasonality

As we saw in the previous section, when $\boldsymbol{w} = s^{-1}\boldsymbol{i}$, $\bar{\gamma}_t$ is a seasonal component in the usual sense since $\boldsymbol{\Omega i} = \boldsymbol{0}$, and the seasonal sums $S(L)\bar{\gamma}_t$ are a zero mean MA(q) process with $q \leq s-2$. This is in general no longer true for the orthogonal NP-P decomposition, for which $\boldsymbol{w} \neq s^{-1}\boldsymbol{i}$, since $\boldsymbol{N}$ is not necessarily a scalar matrix; in the general case, the matrix $\boldsymbol{\Omega}$ lies in the nullity of $\boldsymbol{w}$, rather than $\boldsymbol{i}$. Assuming that w_j (modulo s) denotes the element of $\boldsymbol{w}$ associated with the j-th season, that the P process has started at time $t = 1$, and defining the weighted moving average of s consecutive terms

$$w_j(L)\bar{\gamma}_t = w_j\bar{\gamma}_t + w_{j-1}\bar{\gamma}_{t-1} + \cdots + w_{j-s+1}\bar{\gamma}_{t-s+1},$$

the periodic component is such that $w_j(L)\bar{\gamma}_t$ is a zero mean periodic MA(q) process, $q \leq s-2$. On the contrary, the process $S(L)\bar{\gamma}_t$ has a mean of $(s^{-1}\boldsymbol{i} - \boldsymbol{w})'\boldsymbol{\mu}_1$, which is not necessarily equal to zero. Correspondingly, the NP component will be generated by (4) with starting value $\boldsymbol{w}'\boldsymbol{\mu}_1$ and in a situation in which the weights are different for each season, this will produce a component which, loosely speaking, passes through the seasons having larger weights.

To correct for this situation we may add $(s^{-1}\boldsymbol{i} - \boldsymbol{w})'\boldsymbol{\mu}_1$ to the non periodic component and subtract it from the periodic one to yield components that comply with the common definition of trends and seasonals, although it must be stressed that $S(L)\bar{\gamma}_t$ does not have a finite MA(q) representation (although it is stationary around a zero mean).

4 Seasonal specific local linear trend model

The seasonal specific local level model discussed above is not suitable for a range of macroeconomic time series displaying upward or downward trends. A stochastic slope component needs to be brought into the equations describing the seasons' dynamics, giving the *seasonal specific local linear trend model*:

$$\begin{aligned} y_t &= \boldsymbol{x}_t'\boldsymbol{\mu}_t + \epsilon_t, & t &= 1, \dots, T, \\ \boldsymbol{\mu}_{t+1} &= \boldsymbol{\mu}_t + \boldsymbol{\beta}_t + \boldsymbol{i}\eta_t + \boldsymbol{\eta}_t^*, & \boldsymbol{\eta}_t^* &\sim \mathsf{NID}(\boldsymbol{0}, \boldsymbol{N}_\eta) \\ \boldsymbol{\beta}_{t+1} &= \boldsymbol{\beta}_t + \boldsymbol{i}\zeta_t + \boldsymbol{\zeta}_t^*, & \boldsymbol{\zeta}_t^* &\sim \mathsf{NID}(\boldsymbol{0}, \boldsymbol{N}_\zeta) \end{aligned} \tag{5}$$

according to which the seasonal trends are represented by random walks with stochastic drifts. Again the seasonal specific slopes are driven by a common disturbance, ζ_t, and a disturbance specific to the season, ζ_{jt}^*, $j = 1, \dots, s$. The observation at time t arises from sistematically sampling (via the selection vector $\boldsymbol{x}_t$) a vector IMA(2,1) process with common and idiosyncratic variance components. The reduced form of the model is such that $\Delta\Delta_s y_t$ is a periodic stationary moving average process of order $s+1$; if the specific variances are all equal, then a non periodic model arises.

Model (5) has presumably too many sources of variation, but it may sensibly restricted to provide a suitable representation for seasonal economic time series: if for a scalar and positive q we can express $\boldsymbol{N}_\zeta = q\boldsymbol{N}_\eta$ (homogeneity), then there exists a unique orthogonal decomposition into a non periodic component (with local linear representation) and a periodic one (with seasonal slopes):

$$\begin{aligned} y_t &= \bar{\mu}_t + \bar{\gamma}_t + \epsilon_t, \\ \bar{\mu}_t &= \boldsymbol{w}'\boldsymbol{\mu}_t, \\ \bar{\mu}_{t+1} &= \bar{\mu}_t + \bar{\beta}_t + \eta_t + \boldsymbol{w}'\boldsymbol{\eta}_t^*, \\ \bar{\beta}_{t+1} &= \bar{\beta}_t + \zeta_t + \boldsymbol{w}'\boldsymbol{\zeta}_t^*, \\ \bar{\gamma}_t &= \boldsymbol{x}_t'\boldsymbol{\gamma}_t, \quad \boldsymbol{\gamma}_t = (\boldsymbol{I}_s - \boldsymbol{i}\boldsymbol{w}')\boldsymbol{\mu}_t, \\ \boldsymbol{\gamma}_{t+1} &= \boldsymbol{\gamma}_t + \boldsymbol{\beta}_t^* + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t = (\boldsymbol{I} - \boldsymbol{i}\boldsymbol{w}')\,\boldsymbol{\eta}_t^*, \\ \boldsymbol{\beta}_{t+1}^* &= \boldsymbol{\beta}_t^* + \boldsymbol{\omega}_t^*, \quad \boldsymbol{\omega}_t^* = (\boldsymbol{I} - \boldsymbol{i}\boldsymbol{w}')\,\boldsymbol{\zeta}_t^*, \end{aligned} \tag{6}$$

where $\boldsymbol{w}$ is given as in (3), with $\boldsymbol{N}$ replaced by $\boldsymbol{N}_\eta$.

The slope will only be featured by the NP component when $\boldsymbol{N}_\zeta = \mathbf{0}$, which under homogeneity arises for $q = 0$. The model with common slopes will be plausible for most macroeconomic time series (see the illustration in the next section). In such case the decomposition is like (6) but with $\boldsymbol{\beta}_t^* = \mathbf{0}$ (no seasonal slopes) and

$$\bar{\mu}_{t+1} = \bar{\mu}_t + \beta_t + \eta_t + \boldsymbol{w}'\boldsymbol{\eta}_t^*, \qquad \beta_{t+1} = \beta_t + \zeta_t.$$

If further $\boldsymbol{N}_\eta = \sigma^2_{\eta*}\boldsymbol{I}_s$, we recover the variant of the *basic structural model* (BSM, see Harvey, 1989) with $\boldsymbol{w} = s^{-1}\boldsymbol{i}$, so that the non periodic component is a local linear trend, which is shocked also by an unweighted average of the seasonal specific disturbances, η^*_{jt}, and the periodic one is the HS seasonal component, such that $\boldsymbol{\Omega}\boldsymbol{i} = \mathbf{0}$.

5 Illustration: Italian Industrial Production

Our first illustration concerns the Italian monthly Industrial production series available from the period 1960.1-1999.7 (Source: OECD Statistical Compendium). The series, plotted in figure 1, displays a strong seasonal feature with two relevant seasonal troughs occurring in August and December, related to institutional factors, namely holidays.

We can think of systematically sampling the series so as to build 12 yearly time series, one for each month; each individual time series could be modelled as a local linear trend (plus noise), but it would be linked to the others due to a common disturbance source, with the effect of making them vary together, so that a weighted combination of them is devoid of long run dynamics. We also would expect that the idiosyncratic component is stronger in August and December, allowing these seasons to drift away somewhat from the other months.

The seasonal specific local linear trend model (5), with no idiosyncratic slope disturbances ($\boldsymbol{N}_\zeta = \mathbf{0}$) and idiosyncratic homoscedastic disturbances ($\boldsymbol{N}_\eta = \sigma^2_{\eta*}\boldsymbol{I}_s$) was fitted using Ox 3.0 (Doornik, 2001) and the library of state space algorithms Ssfpack by Koopman *et al.* (1999). The estimated parameters are: $\hat{\sigma}^2_\epsilon = 1399 \times 10^{-7}$, $\hat{\sigma}^2_\eta = 2476 \times 10^{-7}$; $\hat{\sigma}^2_{\eta*} = 329 \times 10^{-7}$; $\hat{\sigma}^2_\zeta = 0$; the value of the maximised log-likelihood is 859.86. The standardised Kalman filter innovations suffer from excess kurtosis, resulting in a highly significant Bowman and Shenton normality test statistic (133.40), whereas the portmanteau test statistic computed on the first 12 residual autocorrelations, Q(12) = 14.89, is not significant.

If we allow the variance of the idiosyncratic disturbance to be greater in August and December we get the following results: $\hat{\sigma}^2_\epsilon = 1515 \times 10^{-7}$, $\hat{\sigma}^2_\eta = 2411 \times 10^{-7}$; $\hat{\sigma}^2_{\eta *} = 11 \times 10^{-7}$, for all seasons excluding August and December; $\hat{\sigma}^2_{\eta_A *} = 3110 \times 10^{-7}$, for August; $\hat{\sigma}^2_{\eta_D *} = 681 \times 10^{-7}$ for December; $\hat{\sigma}^2_\zeta = 0$. Notice that the slope is constant and common to all the seasons. This extension provides a substantial improvement in the fit: the log-likelihood is 966.34 (hence the test of the hypothesis that the idiosyncratic variances are constant is strongly rejected according to the likelihood ratio statistic) and the normality test statistic is substantially reduced (20.62), though it is still significant; moreover, Q(12)=13.30.

We now define the NP component as $\boldsymbol{w}'\tilde{\boldsymbol{\mu}}_{t|T}$, where $\tilde{\boldsymbol{\mu}}_{t|T}$ denotes the smoothed estimates of the vector of seasonal specific trends and $100\boldsymbol{w}' = [9.98\boldsymbol{i}'_7, 0.04, 9.98\boldsymbol{i}'_3, 0.16]$; the P component is correspondingly defined as $\boldsymbol{x}'_t(\boldsymbol{I}_s - \boldsymbol{i}\boldsymbol{w}')\tilde{\boldsymbol{\mu}}_{t|T}$. The rationale of the decomposition is that, given the weighting pattern, the NP component is almost completely unaffected by the values of the series of August and December.

The different variability of the seasons can also be accommodated by the basic structural model with seasonal heteroscedasticity as in Proietti (1998), which is specified as follows:

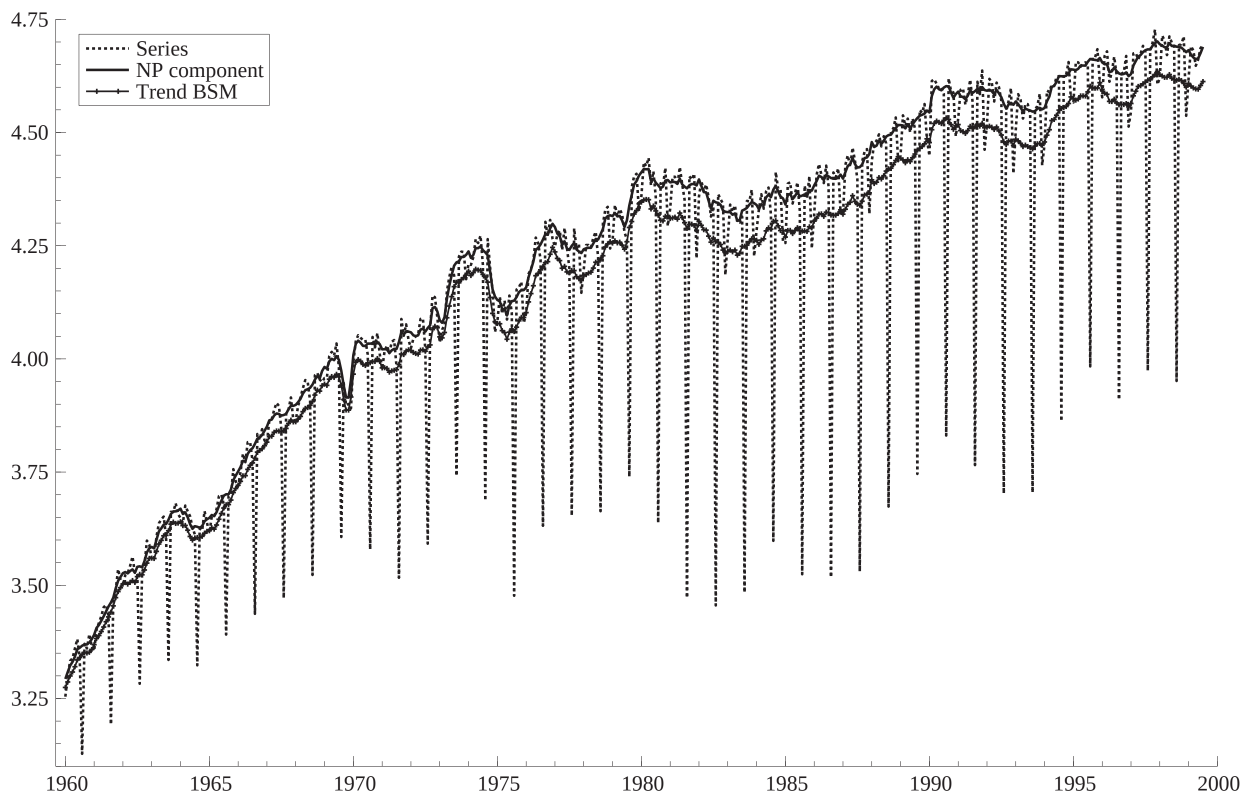
$$\begin{aligned}
y_t &= \mu_t + \gamma_t + \epsilon_t, \quad \epsilon_t \sim \mathsf{NID}(0, \sigma^2_\epsilon) \\
\\
\mu_{t+1} &= \mu_t + \beta_t + \eta_t, \;\; \eta_t \sim \mathsf{NID}(0, \sigma^2_\eta) \\
\beta_{t+1} &= \beta_t + \zeta_t, \;\; \zeta_t \sim \mathsf{NID}(0, \sigma^2_\zeta) \\
\gamma_t &= \boldsymbol{x}'_t \boldsymbol{\gamma}_t, \\
\boldsymbol{\gamma}_{t+1} &= \boldsymbol{\gamma}_t + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim \mathsf{NID}(\mathbf{0}, \boldsymbol{\Omega}),
\end{aligned}$$

where $\boldsymbol{\Omega}$ lies in the null space of $\boldsymbol{i}$,

$$\boldsymbol{\Omega} = \boldsymbol{D} - \frac{\boldsymbol{D}\boldsymbol{i}\boldsymbol{i}'\boldsymbol{D}}{\boldsymbol{i}'\boldsymbol{D}\boldsymbol{i}}$$

and $\boldsymbol{D}$ is a diagonal matrix which has constant values except for August and December.

Figure 1: Italian Industrial Production, 1960.1-1999.7. Original series in logarithms, non periodic component and trend component extracted by the BSM with seasonal heteroscedasticity.



When this model is fitted to the series we get the following results: $\hat{\sigma}_\epsilon^2 = 1444 \times 10^{-7}$, $\hat{\sigma}_\eta^2 = 2341 \times 10^{-7}$; the diagonal elements of $\boldsymbol{D}$ are $\hat{d} = 96 \times 10^{-7}$ for all the months excluding August and December, $\hat{d}_A = 57.189$ for August; $\hat{d}_D = 248 \times 10^{-7}$ for December; $\hat{\sigma}_\zeta^2 = 0$. Moreover, Q(12)=16.72, the Bowman and Shenton test statistic is 25.08, and the log-likelihood is 930.93, so, although the models are not nested, the seasonal specific one yields a higher likelihood.

Figure 1 compares the trend extracted by the two competing models; the difference lies in the fact that the BSM with seasonal heteroscedasticity retains the zero sum constraint for the seasonal component so that if August is more variable all the remaining months adjust to this and the greater variability is smeared on the other seasons; correspondingly the trend is dragged down in the middle of the sample due to the behaviour of August. In the seasonal specific model the zero sum constraints, loosely speaking, applies only with respect to the 10 seasons excluding August and December. If a mean correction is applied, along the lines of section 3.1, the same considerations hold since this would produce only a constant downward shift along the vertical axis. Notice that the differences are not fully accounted by this vertical shift: in 1973-74 the series tend to get closer due to a less deep August trough.

6 Multivariate Seasonal Specific Models

Multivariate generalisations are relatively straightforward. We can devise a system of seemingly unrelated time series equations according to which each of the univariate time series is represented as a seasonal specific structural time series model and the disturbances are contemporaneously correlated.

For simplicity, we focus on a bivariate time series, (y_{1t}, y_{2t}) and on the seasonal specific local linear trend (SSLLT) model with no idiosyncratic slopes, for which y_{kt}, $k = 1, 2$, is represented as:

$$\begin{array}{rcll} y_{kt} & = & \boldsymbol{x}_t'\boldsymbol{\mu}_{kt} + \epsilon_{kt}, & \epsilon_{kt} \sim \mathsf{NID}(0, \sigma^2_{k\epsilon}), \\ \boldsymbol{\mu}_{k,t+1} & = & \boldsymbol{\mu}_{kt} + \boldsymbol{i}\beta_{kt} + \boldsymbol{i}\eta_{kt} + \boldsymbol{\eta}^*_{kt}, & \boldsymbol{\eta}^*_{kt} \sim \mathsf{NID}(\boldsymbol{0}, \boldsymbol{N}_k) \\ \beta_{k,t+1} & = & \beta_{kt} + \zeta_{kt}, & \zeta_{kt} \sim \mathsf{NID}(0, \sigma^2_{k\zeta}) \end{array} \tag{7}$$

where $\epsilon_{kt}, \eta_{kt}, \boldsymbol{\eta}^*_{kt}, \zeta_{kt}$ are mutually uncorrelated at all leads and lags, $\boldsymbol{N}_k$ is a diagonal matrix and each of the disturbances are contemporaneoulsly correlated with the corresponding disturbance in the other series; we shall denote the correlation coefficients respectively by $\rho_\epsilon = \mathsf{Corr}(\epsilon_{1t}, \epsilon_{2t})$, $\rho_\eta = \mathsf{Corr}(\eta_{1t}, \eta_{2t})$, $\rho_\zeta = \mathsf{Corr}(\zeta_{1t}, \zeta_{2t})$, and finally $\rho_{\eta^* j} = \mathsf{Corr}(\eta^*_{1,jt}, \eta^*_{2,jt}), j = 1, \ldots, s$.

Any linear combination of the two series will also have a SSLLT representation; however, if for some j the idiosyncratic disturbances are perfectly correlated, that is $\rho_{\eta^* j} = \pm 1$, there is a common single source of seasonal specific disturbances in season j and there exists a linear combination that has no idiosyncratic feature corresponding to those seasons. As a consequence, the non periodic component that can be extracted from that linear combination will depend solely on that season, where it is fully revealed. In the homoscedastic case, if $\rho_{\eta^*, j} = 1$ for all js then the series are seasonally cointegrated in the usual sense (a linear combination displays only deterministic seasonality).

7 Illustration: Income and Consumption in Sweden

This illustrations refers to a bivariate data set consisting of quarterly real per capita income and non-durables consumption (logarithms) in Sweden analysed by Franses (1996, pp. 202-207) as a case study in periodic cointegration, and available for the period 1963.1-1988.4.

The results from fitting the univariate SSLLT model with no slope idiosyncratic disurbances ($\boldsymbol{N}_\zeta = 0$) are reported in the first two columns of table 1. As far as income is concerned, the LR test of the hypothesis that the seasonal specific level disturbances

η_{jt*}, $j = 1, 2, 3, 4$, are homoscedastic is not significant, and the second column of the table reports the parameter estimates for the restricted model. The model provides a good fit, apart from a significant negative residual autocorrelation at lag 8.

As for consumption, a plot of the seasonal factors extracted by the basic structural model fitted in STAMP 6 (Koopman *et al.*, 2000) showed that the second and the third quarter are less variable for the consumption series. The likelihood for the homoscedastic model ($\sigma^2_{\eta_j *} = \sigma^2_{\eta *}$) is 269.98 whereas that for the model with $\sigma^2_{\eta 1*} = \sigma^2_{\eta 4*}$ and $\sigma^2_{\eta 2*} = \sigma^2_{\eta 3*}$ is 272.77. The LR test of the hypothesis that all idiosyncratic variances are equal is 5.77 which leads to reject the null of homoscedastic disturbances (p-value=0.02); although the evidence for heteroscedasticity is not overwhelming, the fact that the idiosyncratic variance is relatively small in the second and third quarters underlies the fact that the non periodic component is almost perfectly revealed in this seasons, which is reflected in the optimal weighting pattern for the extraction of the NP component, which is approximately $\boldsymbol{w} = [0, 1/2, 1/2, 0]$.

Each component series has 7 sources of disturbances which may be correlated or even common across the two series. We now fit the bivariate seasonal specific model (7), which allows for contemporaneous correlation so that each series has a seasonal specific local linear trend representation, but each disturbance in the model for income is correlated only with the corresponding disturbance in consumption.

The first model we entertain is the homoscedastic model with the idiosyncratic variances $\hat{\sigma}^2_{\eta_j *}$ that are constant across j both for income and consumption and with common correlation coefficient, ρ_{η^*}; if this parameter was equal to ± 1 the series would be seasonally cointegrated, i.e. a linear combination would display only deterministic seasonality. The maximum likelihood estimates of the correlation parameters resulted $\hat{\rho}_{\eta *} = 0.27$, $\hat{\rho}_\eta = 0.46, \hat{\rho}_\zeta = 1.00, \hat{\rho}_\epsilon = 0.18$, so that only the slopes are driven by a common disturbance. The value of the maximised likelihood is 483.48. Notice that the common slope disturbances are perfectly correlated, but there is no seasonal cointegration.

In order to explore the possibility that lack of cointegration is due to a subset of seasons, we consider the same homoscedastic model, but allowing the correlation of the

Table 1: Income and Consumption in Sweden, parameter estimates and diagnostics for univariate and bivariate seasonal specific local linear trend models (variance parameters are multiplied by 10^7). Q(p) is the univariate or multivariate portmanteau test statistic for residual autocorrelation, based on the first p autocorrelations, and Norm. is the Bowman and Shenton normality test.

	Univariate			Bivariate I		Bivariate II	
	Income	Cons.		Income	Cons.	Income	Cons.
$\hat{\sigma}^2_\eta$	759	251	$\hat{\sigma}^2_\eta$	644	179	754	180
			$\hat{\rho}_\eta$	0.41		0.47	
$\hat{\sigma}^2_{\eta 1*}$	1154	125	$\hat{\sigma}^2_{\eta 1*}$	1095	97	832	179
			$\hat{\rho}_{\eta 1*}$	0.82		0.95	
$\hat{\sigma}^2_{\eta 2*}$	1154	2	$\hat{\sigma}^2_{\eta 2*}$	1095	97	832	28
			$\hat{\rho}_{\eta 2*}$	1.00		1.00	
$\hat{\sigma}^2_{\eta 3*}$	1154	2	$\hat{\sigma}^2_{\eta 3*}$	1095	97	832	28
			$\hat{\rho}_{\eta 3*}$	1.00		1.00	
$\hat{\sigma}^2_{\eta 4*}$	1154	125	$\hat{\sigma}^2_{\eta 4*}$	1095	97	832	179
			$\hat{\rho}_{\eta 4*}$	-0.29		-0.38	
$\hat{\sigma}^2_\zeta$	2	2	$\hat{\sigma}^2_\zeta$	7	3	6	3
			$\hat{\rho}_\zeta$	1.00		1.00	
$\hat{\sigma}^2_\epsilon$	273	731	$\hat{\sigma}^2_\epsilon$	540	720	935	692
			$\hat{\rho}_\epsilon$	-0.27		-0.06	
log Lik	207.41	272.77	log Lik	487.34		489.81	
Q(4)	0.56	0.67	Q(4)	6.17		6.33	
Q(8)	12.86*	2.63	Q(8)	26.46		24.45	
Norm.	1.57	7.12*	Norm.	8.53		7.17	

idiosyncratic level disturbances, $\rho_{\eta^* j}$, to vary with the season. The results are reported in table 1 under the header "Bivariate I". It is noticeable that the specific disturbances of the second and third quarter are common, those of the first quarter are strongly and positively correlated, whereas for the fourth quarter they are negatively correlated. This is perhaps not surprising since the seasonal effect associated with the fourth quarter has been declining over time for the consumption series and the same did not occur for income.

The fact that the fourth quarter is mainly responsible for the lack of seasonal cointegration is confirmed by the second bivariate model that was estimated; this differs from the previous only for the idiosyncratic variances of the consumption equation, which are allowed to be different for the second and third quarter in accordance to the univariate findings.

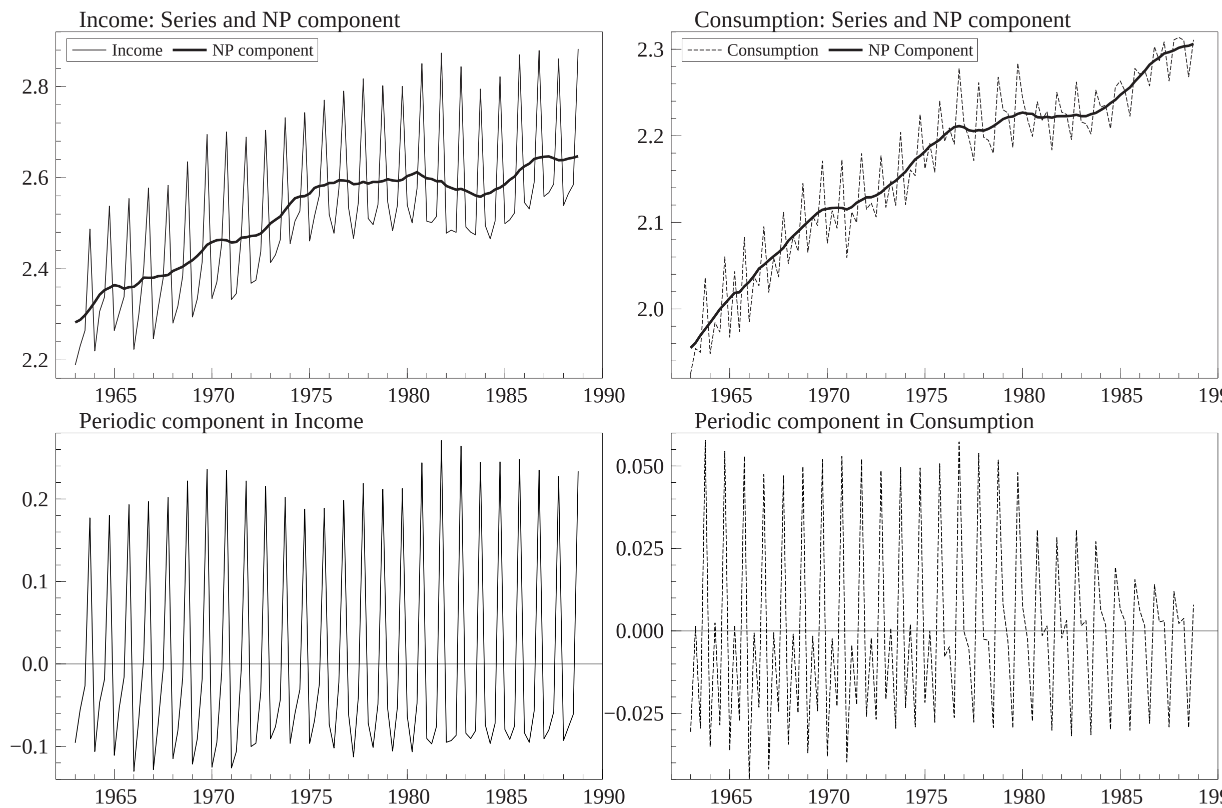
The irregular component is not a negligible source of variation for both series, as the likelihood ratio test of the hypothesis $\sigma_\epsilon^2 = 0, k = 1, 2$ confirms (see Harvey, 1989, ch. 5), but the estimated correlation coefficient is not significantly different from zero. The common level disturbance is positively and significantly correlated across the series, whereas the slope variations are perfectly and positively correlated. The statistics $Q(P)$ for the bivariate models refer to the bivariate portmanteau test based on the first P crosscovariance matrices and are not significant.

Figure 2 displays the smoothed estimates of trend and seasonal component, with a mean correction for the consumption series conducted according to section 3.1. The estimates of the seasonal component point out quite clearly that the two series behave quite differently in the fourth and the first quarter. In particular, the seasonal effect of the fourth quarter shows a marked decline in the last part of the sample period.

8 Other extensions and relation with trigonometric seasonality

Up to now we have entertained the case when $\boldsymbol{N}$ (or $\boldsymbol{N}_\eta$) is a diagonal matrix. Can we allow for some correlation among the seasons and what is a plausible model for this

Figure 2: Income and consumption in Sweden: series with trends and seasonal components extracted by bivariate model II.



correlation? Since the seasons can be represented on a circle, a circular correlation model would be a relevant option; this implies, e.g. that the correlation among any two adjacent seasons is the same and can be made operational specifying $\boldsymbol{N}$ as a circulant matrix. A circulant is a Toeplitz matrix in which n_{jk} is a function of $j-k$ modulo s; each column of $\boldsymbol{N}$ is equal to the previous column rotated downwards by one element.

Let $\boldsymbol{P}$ be the permutation matrix:

$$\boldsymbol{P} = \begin{bmatrix} \mathbf{0}_{s-1} & \boldsymbol{I}_{s-1} \\ 1 & \mathbf{0}'_{s-1} \end{bmatrix},$$

characterised by the orthogonality property: $\boldsymbol{P}^{-1} = \boldsymbol{P}'$, and assume that the number of seasons is even.

The matrix $\boldsymbol{N}$ can be written as a linear combination of the powers of $\boldsymbol{P}$ and their transpose ($\boldsymbol{P}^0 = \boldsymbol{I}_s$):

$$\boldsymbol{N} = \sigma^2_{\eta*}\left[\boldsymbol{I}_s + \sum_{k=1}^{s/2-1}\rho_k(\boldsymbol{P}^k + \boldsymbol{P}^{k'}) + \frac{1}{2}\rho_{s/2}(\boldsymbol{P}^{s/2} + \boldsymbol{P}^{s/2'})\right], \tag{8}$$

where ρ_k denotes the correlation between any two seasons that are k time units apart. As shown in Anderson (1971, Theorem 6.5.3),

$$\frac{1}{2}(\boldsymbol{P}^k + \boldsymbol{P}^{k'}) = \boldsymbol{H}\boldsymbol{\Sigma}_k\boldsymbol{H}',$$

where $\boldsymbol{H}$ is the $s\times s$ orthonormal matrix ($\boldsymbol{H}'\boldsymbol{H} = \boldsymbol{H}\boldsymbol{H}' = \boldsymbol{I}_s$)

$$\boldsymbol{H} = \frac{1}{\sqrt{s}}\begin{bmatrix} 1 & \sqrt{2}\cos\frac{2\pi}{s} & \sqrt{2}\sin\frac{2\pi}{s} & \sqrt{2}\cos\frac{4\pi}{s} & \cdots & -1 \\ 1 & \sqrt{2}\cos\frac{2\pi}{s}2 & \sqrt{2}\sin\frac{2\pi}{s}2 & \sqrt{2}\cos\frac{4\pi}{s}2 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & \sqrt{2}\cos\frac{2\pi}{s}(s-1) & \sqrt{2}\sin\frac{2\pi}{s}(s-1) & \sqrt{2}\cos\frac{4\pi}{s}(s-1) & \cdots & 1 \\ 1 & \sqrt{2}\cos\frac{2\pi}{s}s & \sqrt{2}\sin\frac{2\pi}{s}s & \sqrt{2}\cos\frac{4\pi}{s}s & \cdots & -1 \end{bmatrix}$$

and $\boldsymbol{\Sigma}_k = \mathrm{diag}(1, \cos\frac{2\pi}{s}k, \cos\frac{2\pi}{s}k, \cos\frac{4\pi}{s}k, \ldots, \cos\pi k)$. Hence, it follows that the spectral decomposition of the $\boldsymbol{N}$ matrix is $\boldsymbol{N} = \boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}'$, with

$$\begin{aligned} \boldsymbol{\Sigma} &= \boldsymbol{I}_s + 2\textstyle\sum_{k=1}^{s/2-1}\rho_k\boldsymbol{\Sigma}_k + \rho_{s/2}\boldsymbol{\Sigma}_{s/2} \\ &= \mathrm{diag}(\sigma_0, \sigma_1, \sigma_1, \sigma_2, \sigma_2, \ldots, \sigma_{s/2-1}, \sigma_{s/2-1}, \sigma_{s/2}). \end{aligned}$$

It should be noted that positive definiteness of $\mathbf{\Sigma}$ imposes constrains on the coefficients ρ_k; for instance, for $\rho_k = 0, k > 1$, ρ_1 can take values in [-0.5,0.5].

Now, since $s^{-1/2}\boldsymbol{i}$ is the eigenvector of $\boldsymbol{N}$ corresponding to the characteristic root $\sigma_1 = 1 + 2\sum_k \rho_k + \rho_{s/2}$, the vector of weights of the orthogonal decomposition, defined by (3), is $\boldsymbol{w} = s^{-1}\boldsymbol{i}$, as $\boldsymbol{N}^{-1}\boldsymbol{i} = \sigma_0^{-1}\boldsymbol{i}$ and the denominator is $s\sigma_0^{-1}$; the periodic component is the purely seasonal component $\bar{\gamma}_t = \boldsymbol{x}_t'\boldsymbol{\gamma}_t$, $\boldsymbol{\gamma}_{t+1} = \boldsymbol{\gamma}_t + \boldsymbol{\omega}_t$, with $\mathbf{\Omega} = \tilde{\boldsymbol{H}}\tilde{\mathbf{\Sigma}}\tilde{\boldsymbol{H}}'$, where $\tilde{\boldsymbol{H}}$ is the $s \times (s-1)$ matrix obtained deleting the first column of $\boldsymbol{H}$ and $\tilde{\mathbf{\Sigma}}$ is the diagonal $(s-1) \times (s-1)$ obtained deleting the first row and column from $\mathbf{\Sigma}$.

The trigonometric representation of the seasonal model is obtained as follows:

$$\bar{\gamma}_t = \boldsymbol{z}_t'\boldsymbol{\tau}_t, \quad \boldsymbol{\tau}_{t+1} = \boldsymbol{\tau}_t + \boldsymbol{\kappa}_t,$$

with $\boldsymbol{z}_t' = \boldsymbol{x}_t'\tilde{\boldsymbol{H}}$ and $\mathsf{Var}(\boldsymbol{\kappa}_t) = \tilde{\mathbf{\Sigma}}$. The vector $\boldsymbol{\tau}_t$ measures the time variation in the coefficients associated to $s/2$ trigonometric cycles defined at the seasonal frequencies $\lambda_j = 2\pi j/s, j = 1, \ldots, s/2$, where all the cycles are scaled by $\sqrt{2}$ except the last (defined at π).

The time series properties of seasonal specific models with circular correlation are such that the autocorrelation function does not vary with the season. To introduce periodic effects need to break the circular structure of the $\boldsymbol{N}$ matrix, e.g. allowing the ρ_k's to vary with the season or introducing heteroscedastic idiosyncratic disturbances. When the $\boldsymbol{N}$ is a scalar matrix, the periodic component is a HS seasonal component and this is equivalent to a trigonometric model of seasonality with $\tilde{\mathbf{\Sigma}} = \boldsymbol{I}_{s-1}$. This follows immediately elaborating results in Proietti (2000).

The just established equivalence between circular correlation and trigonometric seasonals highlight the connection between frequency and time domain methods for dealing with seasonal time series. In the latter the ultimate object of interest are the seasons, whereas in the former attention focusses on how fundamental and harmonic cycles combine to yield an overall seasonal pattern. For instance, a trigonometric seasonal model for which attributes the same disturbance variance at all frequencies corresponds to a model in which the seasonal specific disturbances are homoscedastic and uncorrelated across the seasons.

9 Illustration: Consumer Price Index for Women Apparel

This section illustrates how a suitable specification of the $\boldsymbol{N}$ matrix can handle seasonal features that are dealt with in the frequency domain. The series under investigation is the U.S. consumer price index for women apparel goods, for the period 1970.1-2000.8, made available electronically by the Bureau of Labor Statistic.

The main stylised fact is the presence of 2 cycles per year, the first concerning apparel goods for the spring-summer season (February-July), the second those for the autumn-winter season (August-January). A situation like this is smartly accommodated by a trigonometric model of seasonality such that most of the variation is attributable to the cycle defined at the frequency $\pi/3$.

In fact the BSM with trigonometric seasonality with different variances for the fundamental frequency and harmonics produces an excellent fit; basically only the fundamental and the first two harmonics are needed. It is less easy to interpret the trigonometric cycles at $\pi/6$ and $\pi/2$, but there may be two explanations for their significance: first and foremost they are a consequence of aggregation of several apparel item which have a similar but not exactly synchronous seasonal behaviour (2 cycles per year); at a more disaggregated level there are series (for instance women footwear) where only the cycle at $\pi/3$ is needed. Secondly, the autumn-winter season in some years can start in September rather than in August or end in February rather than in January, and so forth. In general, we must admit that each seasonal cycle need not have a particular interpretation in terms of an identifiable economic mechanism; as a matter of fact, harmonics also serve to allow for asymmetries in the overall seasonal cycle.

It is also possible to verify that in this situation X-12-Arima underadjusts the cycle at $\pi/3$ so the SA series displays some seasonal feature that is pronounced at the end of the 80's.

Provided that we know what a sensible model for the series is, we know ask what pattern of correlation among the seasons is capable of interpreting this strong feature of the series. Of course we expect that $\boldsymbol{N} = \sigma^2_{\eta*}\boldsymbol{I}_s$ is not suitable since this would yield an equivalent trigonometric model with all the cycles receiving the same weight. We also could estimate the 7 parameter ($\rho_k, k = 1, \dots, 6$, and $\sigma^2_{\eta*}$) of the circular model but we want to achieve parsimony imposing a pattern on the ρ_k coefficients.

In order to achieve this we can explore the following parameterisations:

1. First order circular correlation model: $\rho_k = \phi\rho_{k-1}, \rho_0 = 1, k \leq 6$
2. Second order circular correlation model: $\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}$, $\rho_0 = 1, \rho_1 = \phi_1/(1 - \phi_2), k \leq 6$

by which the pattern of the ρ_k's is parameterised in terms of one and two coefficients, respectively. The first order model is inadequate for our series since a positive value for ϕ will emphasise the fundamental frequency and a negative one the π frequency; in fact it would give $\sigma_1 > \sigma_2 > \cdots > \sigma_6$ for ϕ positive and the inequality is reversed for ϕ negative, whereas we know that σ_2, the variance associated to the cycle at $\pi/3$, should be the dominating parameter.

The second order model can produce this as we show by fitting the following seasonal specific model:

$$\begin{aligned} y_t &= \boldsymbol{x}'_t\boldsymbol{\mu}_t + \epsilon_t, & t &= 1, \dots, T, \\ \boldsymbol{\mu}_{t+1} &= \boldsymbol{\mu}_t + \boldsymbol{\beta}_t + \boldsymbol{i}\eta_t + \boldsymbol{\eta}^*_t, & \mathsf{Var}(\boldsymbol{\eta}^*_t) &= \boldsymbol{N}, \\ \boldsymbol{\beta}_{t+1} &= \boldsymbol{\beta}_t + \boldsymbol{i}\zeta_t, & \mathsf{Var}(\zeta_t) &= \sigma^2_\zeta. \end{aligned}$$

with $\mathsf{Var}(\eta_t) = \sigma^2_\eta$, and $\boldsymbol{N}$ is specified as in (8) with $\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}$.

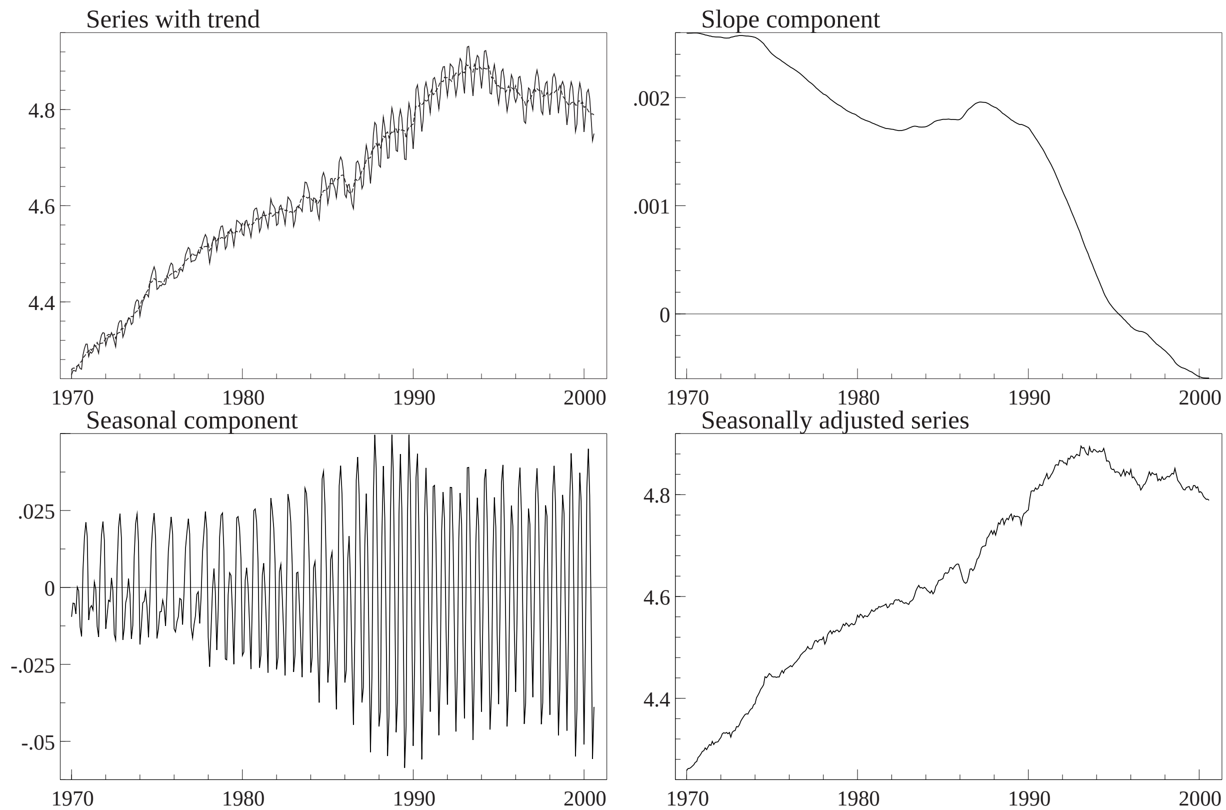
The estimated parameters are (log-likelihood = 1105): $\hat{\sigma}^2_\epsilon = 0$; $\hat{\sigma}^2_\eta = 541 \times 10^{-7}$; $\hat{\sigma}^2_{\eta*} = 30 \times 10^{-7}$; $\hat{\sigma}^2_\zeta = 2 \times 10^{-8}$; moreover, $\hat{\phi}_1 = 0.91$ and $\hat{\phi}_2 = -.72$, which imply the following pattern for the ρ_k's:

$$\rho_1 = 0.53, \rho_2 = -0.24, \rho_3 = -0.60, \rho_4 = -0.38, \rho_5 = 0.09, \rho_6 = 0.35;$$

this is perfectly sensible since the cycle at $\pi/3$ implies a positive correlation at leads and lags of six months and a maximum negative correlation three months apart.

The implied trigonometric disturbance variances are the diagonal elements of the matrix $\tilde{\boldsymbol{K}}$, which result:

Figure 3: U.S. CPI for Women Apparel: smoothed estimates of components arising from the seasonal specific local linear trend model with second order circular correlation.



$$
\begin{array}{lll}
\text{1 cycle per year } (\lambda_1 = \pi/6): & \hat{\sigma}_1 = & 4 \times 10^{-7}; \\
\text{2 cycles per year } (\lambda_2 = \pi/3): & \hat{\sigma}_2 = & 10 \times 10^{-7}; \\
\text{3 cycles per year } (\lambda_3 = \pi/2): & \hat{\sigma}_3 = & 1 \times 10^{-7}; \\
\text{4 cycles per year } (\lambda_4 = 2\pi/3): & \hat{\sigma}_4 = & 4 \times 10^{-8}; \\
\text{5 cycles per year } (\lambda_5 = 5\pi/6): & \hat{\sigma}_5 = & 1 \times 10^{-8}; \\
\text{6 cycles per year } (\lambda_6 = \pi): & \hat{\sigma}_6^2 = & 2 \times 10^{-8}.
\end{array}
$$

Figure 3 presents the smoothed estimates of the components of CPI women apparel, where the trend is defined as $\boldsymbol{w}'\tilde{\boldsymbol{\mu}}_{t|T} = \frac{1}{s}\boldsymbol{i}'\tilde{\boldsymbol{\mu}}_{t|T}$; and the seasonal as $\boldsymbol{x}_t'(\boldsymbol{I}_s - s^{-1}\boldsymbol{i}\boldsymbol{i}')\tilde{\boldsymbol{\mu}}_{t|T}$. The estimates of the seasonal component show the fact that the spring-summer season becomes more and more prominent during the 80ies and the 90ies.

10 Conclusions

This paper has introduced a class of unobserved components models for seasonal time series that hinges on the breakdown of variation sources into sources specific to a particular season and common sources affecting all the seasons. The relative size of the idiosyncratic disturbances can make one season drift away persistently from the other seasons.

Seasonal specific models are periodic, that is they imply first and second order moments that vary with the seasons, but we showed that we can meaningfully extract a non periodic process that is close to the notion of a trend in a seasonal time series; this process is driven by the common disturbance and a suitable weighted average of the idiosyncratic disturbances; the weights discount the more variable seasons, as they are less informative on the overall dynamics. An orthogonal periodic component can also be extracted and the paper has pointed out how it departs from traditional models of seasonality.

When seasonal specific disturbances are homoscedastic, the traditional decomposition into orthogonal trend and seasonality is obtained, the only difference being that the trend is driven by the common disturbance and a simple arithmetic average of the idiosyncratic disturbances. The same holds if we assume a circular correlation structure for them, in which case we have shown that the seasonal component has an isomorphic trigonometric representation, such that the relative importance of the fundamental frequency and the harmonics vary as a function of the circular correlation parameters.

Multivariate extensions are possible and we showed how they can entertain the idea that subsets of season are subject to the same influences, so that partial seasonal cointegration takes place with respect to particular seasons: the bivariate example concerning income and consumption in Sweden served to illustrate the fact that income and consumption cointegrate in the second and third quarters, but not in the first and the fourth quarters.

Overall the class of models proposed introduces periodicity without affecting the possibility of extracting signals that are an expression the long run behaviour. Therefore, they furnish a reasonable comprimise between increasing model complexity in the presence of strong seasonal effects, and preserving the decomposability of the time series.

References

Anderson, (1971) *The Statistical Analysis of Time Series*, John Wiley & Sons, Inc., New York.

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

Franses, P.H. (1996). *Periodicity and Stochastic Trends in Economic Time Series.* Oxford University Press, Oxford U.K.

Hipel, R.W., and McLeod, A.I. (1994). *Time Series Modelling of Water Resources and Environmental Systems.* Developments in Water Science, 45, Elsevier, Amsterdam.

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

Koopman S.J., Harvey, A.C., Doornik, J.A., and Shephard, N. (2000), *STAMP 6.0. Structural Time Series Analyser, Modeller and Predictor.* London, Timberlake Consultants Press, London.

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

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

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

West, M., and Harrison, J. (1997), *Bayesian Forecasting and Dynamic Models*, 2nd ed., Springer-Verlag, New York.