Chapter 1

Bayesian stochastic model specification search for seasonal and calendar effects

Tommaso Proietti and Stefano Grassi

1.1 Introduction

Economic time series are typically available at the monthly frequency of observations. A key feature is the presence of seasonality and calendar effects, which account for much of the variation in the series. Modeling and extracting these component has thus constituted an important problem in the analysis of economic time series. See Zellner (1978), Zellner (1983) Nerlove et al. (1979), Hylleberg (1992), Peña et al. (2001), and Ghysels and Osborn (2001); Findley (2005) discusses some recent advances in seasonal adjustment.

Among the specification issues that have been debated by the literature on seasonality and its adjustment a prominent one deals with characterizing the nature of the seasonal and calendar effects as deterministic or stochastically evolving over time; see, among others, Canova and Hansen (1995), Hylleberg and Pagan (1997), Haywood and Tunnicliffe Wilson (2000), Koop and van Dijk (2000), Busetti and Harvey (2003), Dagum et al. (1993), Dagum and Quenneville (1993), Bell and Martin (2004).

This chapter deals with two research areas to which David Findley contributed significantly: model selection and stochastic models of seasonality. We apply a recently proposed Bayesian model selection technique, known as stochastic model specification search, (Frühwirth-Schnatter and Wagner, 2010) for characterising the nature of seasonality and calendar effects in macroeconomic time series. We illustrate that the methodology can be quite successfully applied to discriminate between stochastic and deterministic trends, seasonals and trading day effects. In particular, we formulate stochastic models for the components of an economic time series and decide on whether a specific feature of the series, i.e. the underlying level and/or a seasonal cycle are fixed or evolve.

The reference model is the unobserved component model known as the basic structural model (Harvey, 1989, BSM henceforth), which will be presented in Section 1.2. Section 1.3 discusses how stochastic model specification search (SMSS) can be applied for the selection of the components of the BSM. This hinges on the representation of the components in non-centered form and a convenient reparameterization of the standard deviation hyperparameters. Section 1.4 discusses the state space representation of the non-centered model and Markov Chain Monte Carlo (MCMC) inference via Gibbs sampling for model selection and Bayesian estimation of the hyperparameters and the components. We apply SMSS to a set of monthly U.S. and Italian macroeconomic time series; the results are presented in Section 1.5. We draw our conclusions in Section 1.6.

1.2 The Basic Structural Time Series Model

The basic structural model, proposed by Harvey and Todd (1983) for univariate time series and extended by Harvey (1989), postulates an additive decomposition of the series into a trend, a seasonal and an irregular component; calendar effects are modeled as regression effects. The name stems from the fact that it provides a satisfactory fit to a wide range of seasonal time series, thereby playing a role analogous to the Airline model in an unobserved components framework.

Let y_t denote a time series observed at $t = 1, 2, \ldots, n$; the BSM is formulated as follows:

$$y_t = \mu_t + S_t + C_t + \epsilon_t, \quad t = 1, \ldots, n, \tag{1.2.1}$$

where μ_t is the trend component, S_t is the seasonal component, C_t is the calendar component and $\epsilon_t \sim \text{NID}(0, \sigma_\epsilon^2)$ is the irregular component.

The trend component has a local linear representation:

$$\begin{aligned} \mu_t &= \mu_{t-1} + q_{t-1} + \eta_t, & \eta_t &\sim \text{NID}(0, \sigma_\eta^2), \quad t = 1, \ldots, n, \\ q_t &= q_{t-1} + \zeta_t, & \zeta_t &\sim \text{NID}(0, \sigma_\zeta^2) \end{aligned} \tag{1.2.2}$$

where q_t is the slope component and we assume that η_t and ζ_t are mutually uncorrelated and independent of ϵ_t and S_t (see Harvey, 1989 and West and Harrison, 1997).

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

$$S_t = \sum_{j=1}^{6} S_{jt}, \qquad \begin{bmatrix} S_{jt} \\ S_{jt}^* \end{bmatrix} = \begin{bmatrix} \cos\lambda_j & \sin\lambda_j \\ -\sin\lambda_j & \cos\lambda_j \end{bmatrix} \begin{bmatrix} S_{j,t-1} \\ S_{j,t-1}^* \end{bmatrix} + \begin{bmatrix} \varpi_{j,t} \\ \varpi_{j,t}^* \end{bmatrix}, j = 1, \ldots, 5, \tag{1.2.3}$$

and $S_{6,t} = -S_{6,t-1} + \varpi_{6t}$. The disturbances ϖ_{jt} and ϖ_{jt}^* are normally and independently distributed with common variance σ_ω^2 for $j = 1, \ldots, 5$, whereas $\text{Var}(\varpi_{6t}) = 0.5\sigma_\omega^2$. While S_{jt} is interpreted as the j-th seasonal cycle, the latent component S_{jt}^* is instrumental to casting the model in Markovian form.

Alternatively, the variance of the seasonal disturbances can be allowed to vary with the frequency, i.e. $\varpi_{jt} \sim \text{NID}(0, \sigma_j^2)$, $j = 1, \dots, 6$, $\varpi^*_{jt} \sim \text{NID}(0, \sigma_j^2)$, $j = 1, \dots, 5$.

In the sequel we will adopt an equivalent alternative representation for the seasonal component due to Hannan (1964), see also Hannan et al. (1970), and known as the evolving seasonal model:

$$\begin{aligned} S_t &= \textstyle\sum_{j=1}^{5}(a_{jt}\cos\lambda_j t + b_{jt}\sin\lambda_j t) + a_{6t}\cos\pi t, \\ a_{jt} &= a_{j,t-1} + \omega_{jt}, \quad \omega_{jt} \sim \text{NID}(0, \sigma_j^2) \\ b_{jt} &= b_{j,t-1} + \omega^*_{jt}, \quad \omega^*_{jt} \sim \text{NID}(0, \sigma_j^2) \end{aligned} \qquad (1.2.4)$$

and $\text{E}(\omega_{jt}\omega^*_{jt}) = 0$. This particular form can be easily represented in the non-centered form (see Section 1.3).

By trigonometric identities it is possible to prove that there is a one-to-one mapping between the two representations; in particular,

$$\begin{bmatrix} a_{jt} \\ b_{jt} \end{bmatrix} = \begin{bmatrix} \cos\lambda_j t & -\sin\lambda_j t \\ \sin\lambda_j t & \cos\lambda_j t \end{bmatrix} \begin{bmatrix} S_{jt} \\ S^*_{jt} \end{bmatrix}; \quad \begin{bmatrix} \omega_{jt} \\ \omega^*_{jt} \end{bmatrix} = \begin{bmatrix} \cos\lambda_j t & -\sin\lambda_j t \\ \sin\lambda_j t & \cos\lambda_j t \end{bmatrix} \begin{bmatrix} \varpi_{jt} \\ \varpi^*_{jt} \end{bmatrix}.$$

The random coefficients a_{jt} and b_{jt} are related to the amplitude of the j-th seasonal cycle as S_{jt} can be rewritten: $S_{jt} = \varphi_t \cos(\lambda_j t - \vartheta_t)$, where $\varphi_t = \sqrt{a_{jt}^2 + b_{jt}^2}$ is the time varying amplitude and $\vartheta_t = \tan^{-1}(b_{jt}/a_{jt})$ is the phase shift.

Calendar effects are due to the differential effects of trading days (TD) and to moving festivals; see Cleveland and Devlin (1982). The former are modeled as $TD_t = \sum_k \phi_k x_{kt}$, where x_{kt} are deterministic regressors defined as follows: letting D_{jt} denote the number of days of type j, $j = 1, \dots, 7$, occurring in month t, then $x_{kt} = D_{jt} - D_{7t}, k = 1, \dots, 6$. The regressors are the differential number of days of type j, $j = 1 \dots, 6$, compared to the number of Sundays, to which type 7 is conventionally assigned. See Bell and Hillmer (1983). If the effect of weekdays is the same, and Saturdays and Sundays are also the same, the trading day component is captured by a single explanatory variable, that is $x_t = D_{1t} - 5D_{2t}/2$, where D_{1t} is the number of weekdays in the month and D_{2t} is the number of Saturdays and Sundays.

As far as moving festivals are concerned, we consider Easter and Labor Day (U.S. time series); their effects are modeled in terms of the proportion of 7 days before Easter or Labor Day that fall in month t and subtracting their monthly long run average, computed over the first 400 years of the Gregorian calendar (1583-1982).

A time varying trading day component can be modeled by letting the coefficients ϕ_k evolve over time: $TD_t = \sum_{k=1}^{6} \phi_{kt} x_{kt}$ where x_{kt} were defined above and ϕ_{kt} are independent Gaussian random walks with common disturbance variance, $\phi_{kt} = \phi_{k,t-1} + \nu_{kt}, \nu_{kt} \sim \text{NID}(0, \sigma_\nu^2)$. Bell and Martin (2004) used this time-varying trading-day model with different disturbance variances.

1.3 Bayesian stochastic specification search for the BSM

This section illustrates how the stochastic model specification search recently proposed by Frühwirth-Schnatter and Wagner (Frühwirth-Schnatter and Wagner, 2010, FS-W henceforth) can be applied for the selection of the components of the BSM. The different specifications for the trend and the seasonal components are nested inside a more general state space model and are obtained by imposing exclusion restrictions, so that discriminating between deterministic and stochastic components amounts to performing variable selection within the regression framework considered by George and McCulloch (1993).

The stochastic model specification search methodology proposed by FS-W hinges on two basic ingredients: the first is the reparameterization of the unobserved components μ_t, S_t and C_t, in non-centered form, with respect to location and scale (see also Gelfand et al., 1995, Frühwirth-Schnatter, 2004 and Strickland et al., 2007). The second is the reparameterization of the hyperparameters representing standard deviations as regression parameters with unrestricted support. The choice of the prior and the conditional independence structure of the reparameterized model enable the definition of a very efficient MCMC estimation strategy based on Gibbs sampling.

1.3.1 Non-centered representation of the random components

The non-centered representation of the trend component is obtained as follows. Denoting by μ_0 and q_0 the initial values of the level and slope components, the trend (1.2.2) can be reparameterized as follows:

$$\begin{aligned} &\mu_t = \mu_0 + q_0 t + \sigma_\eta \tilde{\mu}_t + \sigma_\zeta \tilde{A}_t, \quad && t = 1, \ldots, n, \\ &\tilde{\mu}_t = \tilde{\mu}_{t-1} + \tilde{\eta}_t, && \tilde{\eta}_t \sim \text{NID}(0,1), \\ &\tilde{A}_t = \tilde{A}_{t-1} + \tilde{q}_{t-1}, && \tilde{q}_t = \tilde{q}_{t-1} + \tilde{\zeta}_t, \quad \tilde{\zeta}_t \sim \text{NID}(0,1), \end{aligned} \tag{1.3.1}$$

so that $\tilde{\mu}_0 = \tilde{A}_0 = \tilde{q}_0 = 0$, and $\tilde{\zeta}_t = \zeta_{t-1}/\sigma_\zeta$. Thus, in the non-centred representation the mean function is explicitly written as a linear function of time and the stochastic part is the combination of a random walk and an integrated random walk, both starting off at the origin and driven by standardized independent disturbances.

The non-centered representation of the j-th seasonal cycle is obtained as follows. Denoting by a_{j0} and b_{j0} the initial values of the coefficients,

$$\begin{aligned} S_{jt} &= a_{j0}\cos\lambda_j t + b_{j0}\sin\lambda_j t + \sigma_j\left(\tilde{a}_{jt}\cos\lambda_j t + \tilde{b}_{jt}\sin\lambda_j t\right), \quad && j = 1, \ldots, 5 \\ S_{6t} &= a_{j0}(-1)^t + \sigma_6 \tilde{a}_{6t}(-1)^t \\ \tilde{a}_{jt} &= \tilde{a}_{j,t-1} + \tilde{\omega}_{jt}, && \tilde{\omega}_{jt} \sim \text{NID}(0,1), \\ \tilde{b}_{jt} &= \tilde{b}_{j,t-1} + \tilde{\omega}^*_{jt}, && \tilde{\omega}^*_{jt} \sim \text{NID}(0,1). \end{aligned} \tag{1.3.2}$$

Hence, the non-centered representation of the seasonal component is obtained as $S_t = \sum_{j=1}^{6} S_{jt}$, with S_{jt} given as in (1.3.2).

Alternatively, the non-centered representation of the j-th seasonal cycle can be defined as:

$$\begin{aligned} S_{jt} &= a_{j0}\cos\lambda_j t + b_{j0}\sin\lambda_j t + \sigma_j \tilde{S}_{jt}, && j = 1, \ldots, 5 \\ \tilde{S}_{jt} &= \cos\lambda_j \tilde{S}_{j,t-1} + \sin\lambda_j \tilde{S}^*_{j,t-1} + \tilde{\varpi}_{jt}, && \tilde{\varpi}_{jt} \sim \text{NID}(0,1), \\ \tilde{S}^*_{jt} &= -\sin\lambda_j \tilde{S}_{j,t-1} + \cos\lambda_j \tilde{S}^*_{j,t-1} + \tilde{\varpi}^*_t, && \tilde{\varpi}^*_{jt} \sim \text{NID}(0,1). \\ S_{6t} &= a_{j0}(-1)^t + \sigma_6 \tilde{S}_{6t}, && \tilde{S}_{6t} = -\tilde{S}_{6,t-1} + \tilde{\varpi}_{6t}, \tilde{\varpi}_{6t} \sim \text{NID}(0,1). \end{aligned} \tag{1.3.3}$$

The non-centered representation of the trading days component is:

$$\begin{aligned} TD_t &= \textstyle\sum_{k=1}^{6} \phi_{k0} x_{kt} + \sigma_\nu \left(\sum_{k=1}^{6} \tilde{\phi}_{kt} x_{kt} \right) \\ \tilde{\phi}_{kt} &= \tilde{\phi}_{k,t-1} + \tilde{\nu}_t, \qquad \tilde{\nu}_t \sim \text{NID}(0,1). \end{aligned} \tag{1.3.4}$$

1.3.2 Reparameterization of the BSM

The non-centered representation is useful not only for the efficiency of Bayesian estimation by Markov chain Monte Carlo (MCMC) methods (in particular, when e.g. σ_η^2 is small in comparison to σ_ϵ^2), but also since it paves the way to performing model selection in a regression framework via the stochastic search variable selection (SSVS) approach proposed by George and McCulloch (1993).

The non-centered representation for the components is identified up to sign switches that operate on both the standard deviations and on the underlying stochastic components. For instance the trend component with $(-\sigma_\eta)(-\tilde{\mu}_t)$ replacing $\sigma_\eta \tilde{\mu}_t$ in (1.3.1) is observationally equivalent, i.e. it has the same likelihood. The same can be said of the pairs $(-\sigma_\zeta)(-\tilde{A}_t)$ and $(\sigma_\zeta)(\tilde{A}_t)$, $(-\sigma_j)\left\{-\left(\tilde{a}_{jt}\cos\lambda_j t + \tilde{b}_{jt}\sin\lambda_j t\right)\right\}$ and $\sigma_j\left(\tilde{a}_{jt}\cos\lambda_j t + \tilde{b}_{jt}\sin\lambda_j t\right)$, and so forth. As a consequence, the likelihood function is symmetric around zero along the σ_η, $\sigma_\zeta, \sigma_j, \sigma_\nu$, dimensions and multimodal, if the true standard deviations are larger than zero. This fact can be exploited to quantify how far the posterior of σ_η, $\sigma_\zeta, \sigma_j, j = 1, \ldots, 6$, and σ_ν, is removed from zero.

As a matter of fact, defining independent Bernoulli random variables with success probability 0.5, $\mathsf{B}_\mu, \mathsf{B}_A, \mathsf{B}_{sj}, j = 1, \ldots, 6, \mathsf{B}_{TD}$, we can equivalently write $\sigma_\eta \tilde{\mu}_t = \beta_\mu \mu_t^*$, where $\beta_\mu = (-1)^{\mathsf{B}_\mu}\sigma_\eta$, and $\mu_t^* = (-1)^{\mathsf{B}_\mu}\tilde{\mu}_t$,; similarly, $\sigma_\zeta \tilde{A}_t = \beta_A A_t^*$, where $\beta_A = (-1)^{\mathsf{B}_A}\sigma_\zeta$, $A_t^* = (-1)^{\mathsf{B}_A}\tilde{A}_t$,

$$\sigma_j\left(\tilde{a}_{jt}\cos\lambda_j t + \tilde{b}_{jt}\sin\lambda_j t\right) = \beta_{sj} U_{jt}^*, \quad \beta_{sj} = (-1)^{\mathsf{B}_{sj}}\sigma_j, U_{jt}^* = (-1)^{\mathsf{B}_{sj}}\left(\tilde{a}_{jt}\cos\lambda_j t + \tilde{b}_{jt}\sin\lambda_j t\right),$$

for $j = 1, \dots, 6$, and

$$\sigma_\nu \left(\sum_k \phi_{kt} x_{kt} \right) = \beta_{TD} \Phi_t^*, \quad \beta_{TD} = (-1)^{\mathsf{B}_{TD}} \sigma_\nu, \Phi_t^* = (-1)^{\mathsf{B}_{TD}} \left(\sum_k \phi_{kt} x_{kt} \right).$$

Replacing into the expressions for the components yields:

$$\begin{array}{lcll}
y_t & = & \mu_t + S_t + C_t + \epsilon_t, & \epsilon_t \sim \text{NID}(0, \sigma_\epsilon^2), \\
\\
\mu_t & = & \mu_0 + q_0 t + \beta_\mu \mu_t^* + \beta_A A_t^*, & \\
\mu_t^* & = & \mu_{t-1}^* + \tilde{\eta}_t, & \tilde{\eta}_t \sim \text{NID}(0,1), \\
A_t^* & = & A_{t-1}^* + \tilde{q}_{t-1}, & \\
\tilde{q}_t & = & \tilde{q}_{t-1} + \tilde{\zeta}_t, & \tilde{\zeta}_t \sim \text{NID}(0,1), \\
\\
S_t & = & \sum_{j=1}^5 (a_{j0} \cos \lambda_j t + b_{j0} \sin \lambda_j t) + a_{60}(-1)^t + \sum_{j=1}^6 \beta_{sj} U_{jt}^*, & \\
U_{jt}^* & = & A_{jt}^* \cos \lambda_j t + B_{jt}^* \sin \lambda_j t, \quad j = 1, \dots, 5, & U_{6t}^* = A_{6t}^* \cos \pi t, \\
A_{jt}^* & = & A_{j,t-1}^* + \tilde{\omega}_{jt}, & \tilde{\omega}_{jt} \sim \text{NID}(0,1), \\
B_{jt}^* & = & B_{j,t-1}^* + \tilde{\omega}_{jt}^*, & \tilde{\omega}_{jt}^* \sim \text{NID}(0,1), \\
\\
C_t & = & \sum_{k=1}^6 \phi_{k0} x_{kt} + \beta_{TD} \left(\sum_{k=1}^6 \Phi_{kt}^* x_{kt} \right) + \phi_E x_{Et}, & \\
\Phi_{kt}^* & = & \Phi_{k,t-1}^* + \tilde{\nu}_t, & \tilde{\nu}_t \sim \text{NID}(0,1).
\end{array} \quad (1.3.5)$$

where we have posited $A_{jt}^* = (-1)^{\mathsf{B}_{sj}} \tilde{a}_{jt}, B_{jt}^* = (-1)^{\mathsf{B}_{sj}} \tilde{B}_{jt}, \Phi_{kt}^* = (-1)^{\mathsf{B}_{TD}} \phi_{kt}^*$.

By this reparameterization a standard deviation is transformed into a regression coefficient and SSVS can be applied. Hence the selection of a randomly evolving component is related to the inclusion of a particular regressor.

In principle, we could conduct variable selection for any of the explanatory variables; however, for the computational feasibility of the stochastic search we consider specifications that always include as explanatory variables the constant term, the set of 11 sine and cosine terms at the seasonal frequencies, the six trading days regressors and the moving festivals regressors, so that the most elementary model is a model with a constant level, deterministic seasonals and fixed calendar effects. Variable selection is carried out on the slope term $a_0 t$, on the random walk and integrated random walk components μ_t^*, A_t^*, on the six stochastic terms U_{jt}^* and on $\left(\sum_{k=1}^6 \Phi_{kt}^* x_{kt}\right)$.

We now introduce nine binary indicator variables $\gamma_\mu, \gamma_A, \gamma_{sj}, j = 1, \ldots, 6, \gamma_{TD}$, taking value 1 if the random effects $\mu_t^*, A_t^*, U_{jt}, j = 1, \ldots, 6, \left(\sum_{k=1}^6 \Phi_{kt}^* x_{kt}\right)$ are present and 0 otherwise, along with a binary indicator for the linear trend component, δ, taking values (0,1) according to whether the term $a_0 t$ is included in the model. The ten indicators can be further collected in the multinomial vector $\Upsilon = (\gamma_\mu, \gamma_A, \gamma_{sj}, j = 1, \ldots, 6, \gamma_{TD}, \delta)$.

Hence, there are $K = 2^{10} = 1024$ possible models in competition. These are nested in the specification:

$$\begin{aligned} y_t \;=\; & \mu_0 + \delta q_0 t + \gamma_\mu \beta_\mu \mu_t^* + \gamma_A \beta_A A_t^* + \textstyle\sum_{j=1}^5 (a_{j0} \cos \lambda_j t + b_{j0} \sin \lambda_j t) + a_{60}(-1)^t + \\ & \textstyle\sum_{j=1}^6 \gamma_{sj} \beta_{sj} U_{jt}^* + \sum_{k=1}^6 \phi_{k0} x_{kt} + \gamma_{TD} \beta_{TD} \left(\sum_{k=1}^6 \Phi_{kt}^* x_{kt}\right) + \phi_E x_{Et} + \epsilon_t, \end{aligned} \tag{1.3.6}$$

The different models will be labelled by

$$M_k, \quad k = 1 + \sum_{u=1}^{U} 2^{U-u} \Upsilon_u,$$

where Υ_u is the u-th element of the vector Υ, $u = 1, \ldots, U$.

1.3.3 The restricted BSM with a single variance parameter and model reparameterization

Deciding whether a single variance parameter should be used instead of six for the seasonal component is one of the most important specification issues in formulating a seasonal model.

The trigonometric seasonal model with a single variance parameter is nested within the model for S_t in (1.3.5), as it arises when $\sigma_j = \sigma_\omega, j = 1, \ldots, 5$, and $\sigma_6 = 2^{-1/2}\sigma_\omega$. Under these restrictions, the non-centered representation for the seasonal component becomes

$$\begin{array}{rcll} S_t & = & \sum_{j=1}^{5}(a_{j0}\cos\lambda_j t + b_{j0}\sin\lambda_j t) + a_{60}(-1)^t + \sigma_\omega U_t & \\ U_t & = & \sum_{j=1}^{5}\left(\tilde{a}_{jt}\cos\lambda_j t + \tilde{b}_{jt}\sin\lambda_j t\right) + 2^{-1/2}\tilde{a}_{6t}(-1)^t & \\ \tilde{a}_{jt} & = & \tilde{a}_{j,t-1} + \tilde{\omega}_{jt}, & \tilde{\omega}_{jt} \sim \mathrm{NID}(0,1), \\ \tilde{b}_{jt} & = & \tilde{b}_{j,t-1} + \tilde{\omega}^*_{jt}, & \tilde{\omega}^*_{jt} \sim \mathrm{NID}(0,1). \end{array} \quad (1.3.7)$$

where U_t is a single explanatory variable resulting from combining six non-centered orthogonal stochastic cycles.

More generally, this model can be nested within the more general specification that we have considered in the previous section, by decomposing the frequency specific variance parameters as follows:

$$\sigma_j = \sigma_\omega + (\sigma_j - \sigma_\omega), j = 1, \ldots, 5, \quad \sigma_6 = 2^{-1/2}\sigma_\omega + (\sigma_6 - 2^{-1/2}\sigma_\omega),$$

where

$$\sigma_\omega = \frac{\sum_{j=1}^{5}\sigma_j + \sqrt{2}\sigma_6}{5 + \sqrt{2}}$$

is a weighted average of the individual parameters. Further, we denote by $\sigma^*_j = \sigma_j - \sigma_\omega, j = 1, \ldots, 5, \sigma^*_6 = (\sigma_6 - 2^{-1/2}\sigma_\omega)$, the deviations from the mean. These coefficients are such that $\sum_{j=1}^{5}\sigma^*_j + \sqrt{2}\sigma^*_6 = 0$, and thus we can express the last coefficient σ^*_6 as a linear combination of the others, namely

$$\sigma^*_6 = -2^{-1/2}\sum_{j=1}^{5}\sigma^*_j.$$

Replacing in (1.3.3), we can reparameterize the seasonal component as follows:

$$\begin{array}{rcl} S_t & = & \sum_{j=1}^{5}(a_{j0}\cos\lambda_j t + b_{j0}\sin\lambda_j t) + a_{60}(-1)^t + \sigma_\omega U_t + \sum_{j=1}^{5}\sigma^*_j U^\dagger_{jt} \\ U_t & = & \sum_{j=1}^{5}\left(\tilde{a}_{jt}\cos\lambda_j t + \tilde{b}_{jt}\sin\lambda_j t\right) + 2^{-1/2}\tilde{a}_{6t}(-1)^t \\ U^\dagger_{jt} & = & \tilde{a}_{jt}\cos\lambda_j t + \tilde{b}_{jt}\sin\lambda_j t - 2^{-1/2}\tilde{a}_{6t}(-1)^t. \end{array} \quad (1.3.8)$$

Hence, the reparameterized non-centered form of the model features six random regressors, the first being a weighted average and the remaining five representing weighted contrasts between the $j-th$ and the last non-centered stochastic cycles. If the restrictions were to hold, the coefficients σ_j^* would equal zero and the model reduces to (1.3.7)

This treatment shows that deciding this specification issue can be considered as a model selection issue. Again, defining a suitable set of coefficient $\beta_s = \sigma_\omega(-1)^{\mathsf{B}_s}$ $\beta_{sj}^* = \sigma_j^*(-1)^{\mathsf{B}_{sj}}, j = 1,\dots,5$, where $\mathsf{B}_s, \mathsf{B}_{sj}$ are IID Bernoulli random variables, and indicators $\gamma_s, \gamma_{sj}^\dagger, j = 1,\dots,5,$, taking value 1 if the random effects $\tilde{U}_t = (-1)^{\mathsf{B}_s} U_t, \tilde{U}_{jt} = (-1)^{\mathsf{B}_{sj}} U_{jt}^\dagger, j = 1,\dots,5$, are present and 0 otherwise, we can write the 2^{10} possible models arising from the reparameterisation as follows:

$$\begin{aligned} y_t &= \mu_0 + \delta q_0 t + \gamma_\mu \beta_\mu \mu_t^* + \gamma_A \beta_A A_t^* + \textstyle\sum_{j=1}^5 (a_{j0}\cos\lambda_j t + b_{j0}\sin\lambda_j t) + a_{60}(-1)^t + \\ & \quad \gamma_s \beta_s \tilde{U}_t + \textstyle\sum_{j=1}^5 \gamma_{sj}^\dagger \beta_{sj}^* \tilde{U}_{jt}^* + \sum_{k=1}^6 \phi_{k0} x_{kt} + \gamma_{TD}\beta_{TD}\left(\sum_{k=1}^6 \Phi_{kt}^* x_{kt}\right) + \phi_E x_{Et} + \epsilon_t, \end{aligned} \tag{1.3.9}$$

When the single variance parameter restriction is enforced ($\beta_{sj}^* = 0, j = 1,\dots 5,$), the number of possible models reduces to $2^5 = 32$. For instance, model M_{32} has $\gamma_\mu = \gamma_A = \gamma_s = \gamma_{TD} = \delta = 1$, which corresponds to the unrestricted local linear trend model with stochastic levels and slopes, stochastic seasonality with a single variance parameter and time-varying trading days effects.

1.4 Statistical Treatment

Depending on the value of Υ, the models nested in (1.3.6) admit the following state space representation:

$$\begin{aligned} y_t &= x_{\delta,t}'\rho_\delta + z_{\gamma,t}'\alpha_{\gamma,t} + \epsilon_t, \quad \epsilon_t \sim \mathrm{NID}(0,\sigma_\epsilon^2), \quad t = 1,\dots,n, \\ \alpha_{\gamma,t} &= T_\gamma \alpha_{\gamma,t-1} + R_\gamma u_{\gamma,t}, \quad u_{\gamma,t} \sim \mathrm{NID}(0, I), \end{aligned} \tag{1.4.1}$$

where $\alpha_{\gamma,0} = 0$, and

$$
\begin{aligned}
x_{\delta,t} &= (1, \delta t, \cos\lambda_1 t, \sin\lambda_1 t, \ldots, \cos\pi t, x_{1t}, \ldots, x_{6t}, x_{Et})' \\
\rho_\delta &= (\mu_0, q_0, a_{10}, b_{10}, \ldots, a_{60}, \phi_1, \ldots, \phi_6, \phi_E)', \\
z_{\gamma,t} &= (\gamma_\mu \beta_\mu, \gamma_A \beta_A, 0, \gamma_{s1}\beta_{s1}\cos\lambda_1 t, \gamma_{s1}\beta_{s1}\sin\lambda_1 t, \ldots, \gamma_{s6}\beta_{s6}\cos\pi t, \gamma_{TD}\beta_{TD}x_{1t}, \ldots, \gamma_{TD}\beta_{TD}x_{6t})', \\
\alpha_{\gamma,t} &= (\mu_t^*, A_t^*, \tilde{q}_t, A_{1t}^*, B_{1t}^*, \ldots, A_{6t}^*, \Phi_{1t}^*, \ldots, \Phi_{6t}^*),
\end{aligned}
$$

$$
T_\gamma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & I_{12} \end{pmatrix} \qquad R_\gamma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & I_{12} \end{pmatrix}.
$$

We will assume that the models $M_k, k = 1, \ldots, K$, are equally likely a priori, that is $p(M_k) \propto 1$, or equivalently $p(\Upsilon) = 2^{-U}$, where $p(\cdot)$ denotes the density or the probability function of the argument.

As far as model selection is concerned, it would be prohibitively expensive to compute the posterior model probabilities for each of the 2^U models and select that specification which has the largest. The evaluation of the marginal likelihood for each model is computationally intensive and the accuracy may be poor (see the discussion in FS-W and the references therein). Rather than computing the posterior probabilities of all the possible models, it is computationally more attractive to simulate samples from their posterior distribution by MCMC methods. In particular, exploiting the conditional independence structure of the model, and given the availability of the full conditional posterior distribution of Υ in closed form, the multinomial vector Υ is sampled along with the model parameters by using a Gibbs sampling scheme and a stochastic search of the most likely explanation of the observed time series is sought. After the Gibbs sampling scheme has converged, model selection (and averaging, if one wishes) can be based on $p(\Upsilon|y)$, as estimated by the proportion of times a particular specification was drawn.

1.4.1 Prior specification

Let y denote the collection of time series values $\{y_t, t = 1, \ldots, n\}$ and α denote that of the latent states $\{\alpha_{\gamma,t}, t = 0, 1, \ldots, n\}$; also let ψ_Υ collect the appropriate subset of the parameters $(\mu_0, q_0, a_{10}, b_{10}, \ldots, a_{60}, \phi_{10}, \ldots, \phi_{60}, \beta_\mu, \beta_A, \beta_{s1}, \ldots, \beta_{s6}, \beta_{TD})$ that enter the model for a particular value of Υ.

The prior assumes a conditional independence structure between each block of variables, such that:

$$p(\Upsilon, \psi, \sigma^2_\epsilon, \alpha) = p(\Upsilon)p(\sigma^2_\epsilon)p(\psi|\Upsilon, \sigma^2_\epsilon)p(\alpha|\Upsilon).$$

As stated before, the prior distribution over the model space is uniform, that is $p(\Upsilon) = 2^{-U}$.

For the irregular variance a hierarchical inverse Gamma prior (IG) is adopted, $\sigma^2_\epsilon \sim IG(c_0, C_0)$, where $C_0 \sim Ga(g_0, G_0)$, $Ga(\cdot)$ denoting the Gamma distribution, $c_0 = 2.5$, $g_0 = 5$, and $G_0 = g_0/[0.75\text{Var}(y_t)(c_0 - 1)]$, as in FS-W. The hierarchical prior makes the posterior distributions less sensitive to the choice of the hyperparameters of the IG distribution; it obviously requires an additional sampling step where C_0 is sampled conditional on σ^2_ϵ from the conditional Gamma posterior $C_0|\sigma^2_\epsilon \sim Ga(g_0+c_0, G_0+1/\sigma^2_\epsilon)$ at each sweep of the sample.

For the parameter vector ψ_Υ, if we denote its generic element by $\psi_{\Upsilon i}$, $p(\psi_\Upsilon|\Upsilon, \sigma^2_\epsilon) = \prod_i p(\psi_i|\sigma^2_\epsilon)$, where all the priors are conjugate; for instance, $\beta_\mu|\sigma^2_\epsilon \sim N(0, \kappa_\mu\sigma^2_\epsilon)$, $\beta_A|\sigma^2_\epsilon \sim N(0, \kappa_A\sigma^2_\epsilon)$, $q_0|\sigma^2_\epsilon \sim N(0, d_0\sigma^2_\epsilon)$, etc. For the constant term and the coefficients $a_{j0}, j = 1, \ldots, 6, b_{j0}, j = 1, \ldots, 5, \phi_{k0}, k = 1, \ldots, 6$ we adopt the uninformative priors, e.g. $p(\mu_0|\sigma^2_\epsilon) \propto 1$.

A distinctive feature of the stochastic specification search methodology proposed by Frühwirth-Schnatter and Wagner (2010) is the adoption of Gaussian priors, centered at zero, for the parameters $\beta_\mu, \beta_A, \beta_{sj}, \beta_{TD}$. Not only this allows conjugate analysis, but FS-W show that inference will benefit substantially from the use of a normal prior for e.g. $\beta_\mu = \pm\sigma_\eta$, $\beta_\mu|\sigma^2_\epsilon \sim N(0, \kappa_\mu\sigma^2_\epsilon)$, in lieu of the usual inverse Gamma prior for the variance parameter σ^2_η. In fact, a major problem arising when the IG prior is used is the high sensitivity of the posterior distribution of the variance parameters to the hyperparameters of the IG distribution, when the true variance is close to zero; as a result the MCMC draws will mix very slowly or even lack convergence. On the contrary, the posterior distribution of the β coefficients is not too sensitive to the choice of the prior variance and Monte Carlo inference is much more efficient.

Notice that $\beta_\mu|\sigma_\eta, \gamma_\mu = 1$, is a random variable which takes the values $-\sigma_\eta$ and σ_η with probabilities both equal to 1/2 so that a Gaussian prior centered at zero is reasonable; furthermore, this choice amounts to specifying a hierarchical mixture prior to the parameter β_μ, of the form $p(\beta_\mu) = (1-\gamma_1)I_0 + \gamma_1 N(0, \kappa\sigma_\epsilon^2)$ where I_0 is a degenerate density with point mass at zero, see Smith and Kohn (1996). As pointed out in George and McCulloch (1997), this prior entails that a stochastic trend will be included if β_μ can be distinguished from zero irrespective of its absolute size.

Finally, the prior for α is provided by the Gaussian dynamic model (1.4.1), so that,

$$p(\alpha) = p(\alpha_{\gamma 0}) \prod_{t=1}^{n} p(\alpha_{\gamma t}|\alpha_{\gamma,t-1}),$$

with $\alpha_{\gamma t}|\alpha_{\gamma,t-1} \sim N(T_\gamma \alpha_{\gamma,t-1}, R_\gamma R'_\gamma)$ and $\alpha_{\gamma,0} = 0$.

1.4.2 MCMC Estimation

Model selection requires the evaluation of the posterior probability function of the multinomial vector Υ, denoted $p(\Upsilon|y)$. Also, for the selected model we are interested in the marginal posterior distributions of the parameters $p(\psi|y)$ and the states $p(\alpha|y)$. The required posteriors are not available in closed form, but we are capable of drawing samples from them by Markov chain Monte Carlo methods and, in particular, by a Gibbs sampling (GS) scheme that we now are going to discuss in some detail. The GS scheme produces correlated random draws from the posteriors by repeatedly sampling an ergodic Markov chain whose invariant distribution is the target density; see e.g. Robert and Casella (2004) and Gamerman and Lopes (2007). In essence, it defines a homogeneous Markov Chain such that the transition kernel is formed by the full conditional distributions and the invariant distribution is the unavailable target density.

The GS scheme can be sketched as follows. Specify a set of initial values $\Upsilon^{(0)}, \sigma_\epsilon^{2(0)}, \alpha^{(0)}, \psi^{(0)}$. For $i = 1, 2, \ldots, M$, iterate the following operations:

a. Draw $\Upsilon^{(i)} \sim p(\Upsilon|\alpha^{(i-1)}, y)$

b. Draw $\sigma_\epsilon^{2(i)} \sim p(\sigma_\epsilon^2|\Upsilon^{(i)}, \psi^{(i-1)}, \alpha^{(i-1)}, y)$

c. Draw $\psi^{(i)} \sim p(\psi|\Upsilon^{(i)}, \sigma_\epsilon^{2(i)}, \alpha^{(i-1)}, y)$

d. Draw $\alpha^{(i)} \sim p(\alpha|\Upsilon^{(i)}, \sigma_\epsilon^{2(i)}, \psi^{(i)}, y)$

The above complete conditional densities are available, up to a normalizing constant, from the form of the likelihood and the prior.

For the sake of notation, let us write the regression model as $y = Z_\Upsilon \psi_\Upsilon + \epsilon$, where y and ϵ are vectors stacking the values $\{y_t\}$ and $\{\epsilon_t\}$, respectively, and the generic row of matrix Z_Υ contains the relevant subset of the explanatory variables.

Step a. is carried out by sampling the indicators with probabilities proportional to the conditional likelihood of the regression model, as

$$\begin{aligned} p(\Upsilon|\alpha, y) &\propto p(\Upsilon)p(y|\Upsilon, \alpha) \\ &\propto p(y|\Upsilon, \alpha), \end{aligned}$$

which is available in closed form (see below).

Under the normal-inverse Gamma conjugate prior for $(\psi_\Upsilon, \sigma_\epsilon^2)$

$$\sigma_\epsilon^2 \sim IG(c_0, C_0), \quad \psi_\Upsilon|\sigma_\epsilon^2 \sim N(0, \sigma_\epsilon^2 D_\Upsilon),$$

where D_Υ is a diagonal matrix with elements κ_μ, κ_A, etc., steps b. and c. are carried out by sampling from the posteriors

$$\begin{aligned} \sigma_\epsilon^2|\Upsilon, \alpha, y &\sim IG(c_{T*}, C_{T*}) \\ \psi_\Upsilon|\Upsilon, \sigma_\epsilon^2, \alpha, y &\sim N(m, \sigma_\epsilon^2 S) \end{aligned}$$

where

$$S = \left(Z'_\Upsilon Z_\Upsilon + D_\Upsilon^{-1}\right)^{-1}, \qquad m = SZ'_\Upsilon y$$
$$c_{T*} = c_0 + T^*/2, \qquad C_{T*} = C_0 + \tfrac{1}{2}\left(y'y - m'S^{-1}m\right).$$

Finally,

$$p(y|\Upsilon, \alpha) \propto \frac{|S|^{0.5}}{|D_\Upsilon|^{0.5}} \frac{\Gamma(c_{T^*})}{\Gamma(c_0)} \frac{C_0^{c_0}}{C_{T^*}^{c_{T^*}}},$$

see e.g. Geweke (2005), where $\Gamma(\cdot)$ denotes the Gamma function.

The sample from the posterior distribution of the latent states, conditional on the model and its parameters, in step d., is obtained by the conditional simulation smoother proposed by Durbin and Koopman (2002) for linear and Gaussian state space models.

Finally, the draw of the parameters β_μ, β_A, $\beta_{sj}, j = 1, \ldots, 6$, β_{TD} are obtained by performing a final random sign permutation. This is achieved by drawing independently Bernoulli random variables B_μ, B_A, $\mathsf{B}_{sj}, j = 1, \ldots, 6, \mathsf{B}_{TD}$ with probability 0.5, and recording $(-1)^{\mathsf{B}_\mu}(\sigma_\eta, \tilde{\mu}_t)$, $(-1)^{\mathsf{B}_A}(\sigma_\zeta, \tilde{A}_t, a_t)$, etc.

1.5 Empirical Results

We apply Bayesian stochastic specification search to a set of U.S. and Italian macroeconomic time series, listed in Table 1.1, which were selected for their relevance in the measurement of the macroeconomy. All the series are transformed into logarithms, except for the U.S. monthly inflation rate, which is computed as the logarithmic change of the consumer price index with respect to the previous month.

Table 1.1 Approximately Here

We start by discussing the results for the specifications with a single seasonal variance parameter, based on 60,000 MCMC draws, 20,000 of which constituted the burn-in sample. For this case there are $K = 32$ models, as Υ is a vector of five indicator variables with elements $(\gamma_\mu, \gamma_A, \gamma_s, \gamma_{TD}, \delta)$.

Table 1.2 reports the percentage of MCMC replicates by which model $M_k, k = 1+16\gamma_\mu + 8\gamma_A + 4\gamma_s + 2\gamma_{TD} + \delta$, was selected. It also reports the value of the Deviance Information Criterion (DIC). The latter is a measure of model fit (see e.g. Gelman and Rubin, 2004) computed as follows:

$$DIC_\Upsilon = \hat{D}_\Upsilon(y;\psi,\sigma^2_\epsilon) + \{\hat{D}_\Upsilon(y;\psi,\sigma^2_\epsilon) - D_\Upsilon(y;\hat{\psi},\hat{\sigma}^2_\epsilon)\}, \tag{1.5.1}$$

where

$$\hat{D}_\Upsilon(y;\psi,\sigma^2_\epsilon) = \frac{1}{R}\sum_{i=1}^{R} D_\Upsilon(y|\psi^{(i)},\sigma^{2(i)}_\epsilon)$$

is the average value of the deviance

$$D_\Upsilon(y|\psi,\sigma^2_\epsilon) = -2\ln p(y|\psi,\sigma^2_\epsilon)$$

computed over the R posterior simulations of (ψ,σ^2_ϵ) for the specification Υ obtained by the Gibbs sampling scheme; the conditional likelihood $p(y|\psi,\sigma^2_\epsilon)$ is evaluated by the Kalman filter for the relevant state space model with parameter values (ψ,σ^2_ϵ). The term in parenthesis in (1.5.1), where $D_\Upsilon(y;\hat{\psi},\hat{\sigma}^2_\epsilon)$ represents the deviance evaluated at the posterior means $\hat{\psi} = R^{-1}\sum_i \psi^{(i)}, \hat{\sigma}^2_\epsilon = R^{-1}\sum_i \sigma^{2(i)}_\epsilon$, measures the number of effective parameters in the model.

The main evidence can be summarized as follows.

1. The specification with time-varying trading days is never selected.

2. The modal specification has $\Upsilon = (1,0,1,0,0)$ in four cases (US.HS, US.IR, US.Imp and IT.IP): the trend is a driftless random walk and stochastic seasonals.

3. The specifications selected for the US.UR and IT.TA, and US.IP, M_{18} and M_{25}, respectively, do not feature stochastic seasonality. Model M_{18} features a random walk trend with constant drift and fixed seasonal and calendar effects; model M_{25} differs only for the trend model, which is local linear.

4. For US.CC and US.CPI the models two most frequently selected specifications are M_{29} and M_{30}; they both feature a local linear trend and stochastic seasonal, the only

difference relating to the fact that the slope component is nonzero at the beginning of the sample period only for the latter.

5. The models selected for US.CPI and its first differences, US.IR, can be easily reconciled as M_{21}, M_{22} are the same as M_{29}, M_{30}, but with a nonstochastic slope. However, notice that if $\sigma^2_\eta > 0$ then the model for the irregular should be replaced by a moving average component of order one.

6. The two U.S. retail sales series feature models M_{13} and M_{14} as modal specifications; they entail a fixed level, a stochastic slope, stochastic seasonality and the initial slope is zero (M_{13}) or nonzero (M_{14}).

7. Models selected more frequently have lower DIC values.

Table 1.2 Approximately Here

Turning to the selection of seasonal models with variance parameters varying with the trigonometric components, we present in Table 1.3 the first three modal specification that were selected, along with the posterior model probabilities $100 \times \hat{p}(\Upsilon|y)$ estimated by the Gibbs sampling scheme and the DIC.

The results confirm that for the series considered in the application trading days effects can be safely considered as fixed, rather than evolving over time, the marginal probability $P(\gamma_{TD} = 1)$ being virtually zero in all the cases. The main evidence arising from Table 1.3 can be summarized as follows.

- Trends and seasonals are better characterized as stochastic, rather than deterministic. The results are in broad agreement with the analysis of the restricted model, except for US.UR and IT.TA, and US.IP, for which some of the trigonometric cycles are not fixed when the variance parameters are allowed to vary with the frequency of the cycle.
- For US.IP and US.UR the three modal models are such that the trigonometric component defined at the fundamental frequency $\lambda_1 = \pi/6$ is not stochastic. On the contrary, the only components that are stochastically evolving for IT.TA are the fundamental and the first harmonic.

- There is a lot of variation across the series as to which trigonometric cycles are time-varying or fixed. The broad evidence arising from Table 1.3 is that the number of occurrences in which the cycle at λ_j is selected as stochastically evolving decreases with j; quite often the cycle defined at the $\lambda_6 = \pi$ frequency (six cycles per year) is fixed.

- Model uncertainty often concerns marginal aspects, such as the presence of a non-zero slope term at the initial time, or a specific trigonometric component.

Table 1.3 Approximately Here

Hereby we provide a more detailed analysis of Italian IP series. Figure 1.1 displays the estimated posterior densities of some of the parameters of the saturated BSM model, which is (1.3.6) with $\Upsilon = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$. The estimates are based on MCMC draws obtained by running the Gibbs sampler for 40,000 iterations after a burn–in of 20,000.

When the posterior of the parameters $\beta_\Lambda, \Lambda = \{\mu, A, s_1, \ldots, s_6, TD\}$ is bimodal and sufficiently removed from zero, the corresponding true variance parameter is different from zero and the associated random component contributes significantly to the evolution of the series. This is the case of β_μ (stochastic level) and the seasonal parameters $\beta_{sj}, j = 1, 2, 3, 4$, whereas β_{s5} and β_{s6} have some density around zero. On the contrary, the posterior of β_A is concentrated around zero, which points to a fixed slope; moreover the distribution of q_0 is such that the initial slope is not significantly different from zero, so that the specification of the trend component reduces to a driftless random walk. Also, trading days effects are fixed.

Figure 1.1 Approximately Here

When SMSS is applied by running a MCMC sampling scheme that draws samples from the posterior distribution of the indicators, the specification with maximal estimated posterior probability is $\Upsilon = (1, 0, 1, 1, 1, 1, 1, 0, 0, 0)$, corresponding to M_{761}, which is a restricted BSM with no slope, a fixed trigonometric cycle at the Nyquist frequency, and fixed calendar effects. The estimated posterior model probability is 0.3. Figure 1.2 shows the estimated posterior means of the unobserved components (along with the 95% credible interval for the trend), whereas Figure 1.3 displays the estimated posteriors of the trading days effects, $\phi_{k0}, k = 1, \dots, 6$, along with the posterior of the Sunday effect, obtained as $\phi_{70} = -\sum_{j=1}^{6} \phi_{k0}$. Model uncertainty deals essentially with the time variation of the seasonal trigonometric cycles defined at the frequencies λ_5 and λ_6 (see Table 1.3).

Figure 1.2 Approximately Here

Figure 1.3 Approximately Here

The saturated BSM model can be estimated also using the reparameterization discussed in section 1.3.2. In particular, pronounced bi-modality of the posterior distribution of the parameter β_s (presented in the top left panel of figure 1.4) points out that seasonality is stochastic rather than deterministic. Also, the posterior distributions of the differential parameters $\beta^*_{sj}, j = 1, \dots, 5$ point out that the variance parameters σ_j are likely to differ significantly from $\sqrt{2}\sigma_6$, except for σ_5. Hence, the seasonal model could be simplified by expressing the frequency specific variances in terms of five, rather than six, unrestricted parameters. Overall, the evidence is against the specification with a single variance parameter.

Figure 1.4 Approximately Here

The model with frequency specific variance parameters is usually a substantial improvement over the specification with a single variance parameter σ^2_ω. To illustrate this point, Figure 1.5 compares the posterior distribution of the Easter regression coefficient ϕ_E for the unrestricted model (1.3.6) and the specification enforcing the restriction $\sigma^2_j = \sigma^2_\omega, j = 1, \dots, 5, \sigma^2_6 = 0.5\sigma^2_\omega$. Similar considerations can be made for the precision by which the unobserved components are estimated: the bottom panel compares the 95% credible intervals of the trend component for the two specifications.

Figure 1.5 Approximately Here

A final point deals with the comparison of the saturated model (M_{1024}) with the selected model (see Table 1.3). For the series investigated in this paper model selection has little effect on the estimation of the seasonally adjusted series, although it may affect the trend and the irregular, or the seasonal and the calendar components, individually. However, once model selection has been carried out once, conditioning on the selected model may improve the efficiency and timeliness of the Gibbs sampling scheme (the convergence statistics, see e.g. Geweke (2005), not reported for brevity, are always satisfactory for the restricted model, whereas they may fail for the unrestricted model).

1.6 Conclusions

We have applied a recent methodology, Bayesian stochastic model specification search (Frühwirth-Schnatter and Wagner, 2010), for the selection of the unobserved components (level, slope, seasonal cycles, trading days effects) that are stochastically evolving over time.

SMSS hinges on two basic ingredients: the non-centered representation of the unobserved components and the reparameterization of the hyperparameters representing standard deviations as regression parameters with unrestricted support. The choice of the prior and the conditional independence structure of the model enable the definition of a very efficient MCMC estimation strategy based on Gibbs sampling. Indeed, our first general conclusion is that, transcending the model selection problem, Bayesian estimation of the BSM should be carried out by using the approach suggested by Frühwirth-Schnatter and Wagner (2010).

Our empirical illustrations have dealt with a limited data set consisting of 11 time series, so that we can envisage an extension of this research that gathers further empirical evidence by processing a much larger set of data. However, there are some regularities that we have drawn from our case studies. The first is that, somewhat disappointingly, trading day effects are time-invariant. A possible explanation is that the series available are possibly too short to enable us to detect small variations induced by the calendar; moreover, some of the TD variation may be absorbed by seasonal cycles defined at higher frequencies.

A second conclusion is that the specification with six frequency specific variance parameters proves superior to that using a single parameter, yielding more precise estimates of the unobserved components and the regression effects. We also suspect that the latter can induce a bias towards selecting deterministic models of seasonality. We leave to future research discriminating between the two representations as a model selection problem, by comparing their posterior probabilities.

The selection of a BSM specification among the 2^{10} possible ones has led in all the cases to models with one or more seasonal cycles being characterized as deterministic. The overall result is that the set of time series analyzed display stochastically evolving trends and seasonality.

Finally, our stochastic model specification search was carried out for a version of the BSM with trigonometric seasonality. In the future we would like to apply the methodology to alternative models for seasonal time series, featuring a stochastic dummy seasonal model (see e.g. West and Harrison (1997)), where the individual monthly effects may be evolving over time.

Acknowledgments

The Authors wish to thank the Editors and the Referees for their comments that helped improve the paper. Financial support from CREATES, funded by the Danish National Research Foundation, is gratefully acknowledged by Stefano Grassi.

Bibliography

Bell, W. R. and Hillmer, S. C. (1983). Modeling time series with calendar variation. *Journal of the American Statistical Association*, 78:526–534.

Bell, W. R. and Martin, D. E. K. (2004). Modeling time-varying trading day effects in monthly time series. In *ASA Proceedings of the Joint Statistical Meetings: Alexandria*. American Statistical Association.

Busetti, F. and Harvey, A. (2003). Seasonality tests. *Journal of Business and Economic Statistics*, 21:420–436.

Canova, F. and Hansen, B. (1995). Are seasonal patterns constant over time? a test for seasonal stability. *Journal of Business and Economic Statistics*, 13:237–252.

Cleveland, W. and Devlin, J. (1982). Calendar effects in monthly time series: modeling and adjustment. *Journal of of the American Statistical Association*, 77:520–528.

Dagum, E. and Quenneville, B. (1993). Dynamic linear models for time series components. *Journal of Econometrics*, 55:333–351.

Dagum, E., Quenneville, B., and Sutradhar, B. (1993). Trading-day variations multiple regression models with random parameters. *International Statistical Review*, 60:57–73.

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

Findley, D. (2005). Some recent developments and directions in seasonal adjustment. *Journal of Official Statistics*, 21(2):343–365.

Frühwirth-Schnatter, S. (2004). Efficient bayesian parameter estimation for state space models based on reparameterizations. In Harvey, A. C., Koopman, S. J., and Shephard, N., editors, *State Space and Unobserved Component Models: Theory and Applications*, pages 123–151. Cambridge University Press.

Frühwirth-Schnatter, S. and Wagner, H. (2010). Stochastic model specification search for gaussian and partial non-gaussian state space models. *Journal of Econometrics*, 154(1):85–100.

Gamerman, D. and Lopes, F. H. (2007). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman and Hall/CRC Texts in Statistical Science.

Gelfand, A., Sahu, S., and Carlin, B. (1995). Efficient parameterizations for normal linear mixed models. *Biometrika*, 82:479–488.

Gelman, A., C. J. S. H. and Rubin, D. (2004). *Bayesian Data Analysis (2nd ed.).* Chapman and Hall/CRC Texts in Statistical Science.

George, E. I. and McCulloch, R. (1993). Variable selection via gibbs sampling. *Journal of the American Statistical Association*, 88:881– 889.

George, E. I. and McCulloch, R. (1997). Approaches for bayesian variable selection. *Statistica Sinica*, 7:339–373.

Geweke, J. (2005). *Contemporary Bayesian Econometrics and Statistics*. Wiley Series in Probability and Statistics.

Ghysels, E. and Osborn, D. (2001). *The econometric analysis of seasonal time series*. Cambridge: Cambridge University Press.

Hannan, E. J. (1964). The estimation of a changing seasonal pattern. *Journal of the American Statistical Association*, 59:1063–1077.

Hannan, E. J., Terrell, R. D., and Tuckwell, N. (1970). The seasonal adjustment of economic time series. *International Economic Review*, 11:24–52.

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

Haywood, J. and Tunnicliffe Wilson, G. (2000). Selection and estimation of component models for seasonal time series. *Journal of Forecasting*, 19:393–417.

Hylleberg, S. (1992). *Modelling Seasonality.* Oxford: Oxford University Press.

Hylleberg, S. and Pagan, A. (1997). Seasonal integration and the evolving seasonals model. *International Journal of Forecasting*, 13:329–340.

Koop, G. and van Dijk, H. (2000). Testing for integration using evolving trend and seasonals models: A bayesian approach. *Journal of Ecoometrics*, 97:261–291.

Nerlove, M., Grether, D. M., and Carvalho, J. L. (1979). *Analysis of economic time series: a synthesis.* New York: Academic Press.

Peña, D., Tiao, G., and Tsay, R. (2001). *A Course in Time Series Analysis.* New York: J. Wiley and Sons.

Robert, P. and Casella, G. (2004). *Monte Carlo Statistical Methods.* New York: Springer.

Strickland, C. M., Martin, G., and Forbes, C. (2007). Parameterisation and efficient mcmc estimation of non-gaussian state space models. *Computational Statistical and Data Analysis*, 97(52):2911–2930.

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

Zellner, A. (1978). *Seasonal Analysis of Economic Time Series.* US Dept. of Commerce-Bureau of the Census.

Zellner, A. (1983). *Applied Time Series Analysis of Economic Data.* US Dept. of Commerce-Bureau of the Census.

Table 1.1: Dataset used in the study.

Series description	Sample period	Name
U.S. Housing Starts Total	1960.1 - 2010.2	US.HS
U.S. Industrial Product index	1986.1 - 2010.1	US.IP
U.S. Retail Sales Total	1960.1 - 2008.3	US.RSt
U.S. Retail with Food less Auto	1960.1 - 2008.3	US.RSla
U.S. Unemployment Rate	1960.1 - 2009.8	US.UR
U.S. Consumer Price Index	1960.1 - 2009.8	US.CPI
U.S. Monthly Inflation Rates	1960.2 - 2009.8	US.IR
U.S. Consumer Credit Total	1992.1 - 2009.12	US.CC
U.S. Imports of Crude Oil (Quantity)	1973.1 - 2009.7	US.Imp
Italian Industrial Production	1990.1 - 2010.1	IT.IP
Italian Tourist Arrivals	1990.1 - 2009.10	IT.TA

Table 1.2: BSM with single seasonal variance parameter. Percentage by which model $M_k, k = 1+16\gamma_\mu+8\gamma_A+4\gamma_s+2\gamma_{TD}+\delta$, is selected in 40,000 MCMC draws. Deviance Information Criterion (DIC) values are in parenthesis.

	Model											
Series	M_9	M_{10}	M_{13}	M_{14}	M_{17}	M_{18}	M_{21}	M_{22}	M_{25}	M_{26}	M_{29}	M_{30}
US.HS	0	0	0	0	4 (−1044.0)	5 (−1058.1)	82 (−1193.6)	9 (−1072.0)	0	0	0	0
US.IP	0	0	0	0	0	0	0	0	67 (−3579.2)	33 (−3557.3)	0	0
US.RSt	0	0	53 (−1028.6)	41 (−1009.3)	0	0	0	0	1 (−975.4)	1 (−974.2)	1 (−965.1)	2 (−1001.9)
US.RSla	0	0	30 (−1026.4)	68 (−1057.3)	0	0	0	0	0	0	0	2 (−1014.6)
US.UR	0	0	0	0	0	70 (−1944.1)	20 (−1851.9)	10 (−1850.9)	0	0	0	0
US.CPI	0	0	0	0	0	0	0	0	0	0	30 (−1818.0)	70 (−4739.7)
US.IR	0	0	0	0	0	0	65 (−5217.9)	35 (−5214.7)	0	0	0	0
US.CC	0	0	0	6 (−1017.1)	0	0	0	0	0	0	57 (−5101.0)	43 (−4595.9)
US.Imp	0	0	0	0	1 (−599.11)	0	70 (−689.34)	29 (−601.97)	0	0	0	0
IT.IP	0	0	0	10 (−914.89)	5 (−904.09)	5 (−905.40)	41 (−918.12)	34 (−914.65)	0	1 (−903.22)	4 (−905.70)	0
IT.TA	5 (−867.91)	3 (−868.33)	0	0	29 (−873.91)	59 (−877.05)	0	0	2 (−873.69)	1 (−872.35)	0	0

Table 1.3: First three modal specifications selected by the Gibbs sampling scheme, estimated posterior probabilities $100 \times \hat{\pi}(\Upsilon|y)$ (in parentheses). The vector Υ has elements $(\gamma_\mu, \gamma_A, \gamma_{sj}, j = 1, \ldots, 6, \gamma_{TD}, \delta)$.

Series	Fist Selected Model	%	DIC	Second Selected Model	%	DIC	Third Selected Model	%	DIC
US.HS	$\Upsilon = (1,0,1,1,0,0,0,0,0,0)$	35	-1208.7	$\Upsilon = (1,0,1,1,1,0,0,0,0,0)$	22	-1205.4	$\Upsilon = (1,0,1,0,0,0,0,0,0,0)$	15	-1185.3
US.IP	$\Upsilon = (1,1,0,1,1,1,1,0,0,0)$	30	-3822.5	$\Upsilon = (1,1,0,1,1,1,1,0,0,1)$	25	-3810.7	$\Upsilon = (1,1,0,1,1,1,0,0,0,0)$	10	-3769.9
US.RSt	$\Upsilon = (0,1,1,1,0,1,0,0,0,0)$	37	-1074.9	$\Upsilon = (0,1,1,1,0,0,0,0,0,0)$	31	-1020.1	$\Upsilon = (0,1,1,1,0,0,0,0,0,1)$	17	-1057.5
US.RSla	$\Upsilon = (0,1,1,1,1,1,1,1,0,1)$	30	-1153.5	$\Upsilon = (1,1,1,1,1,1,1,1,0,0)$	24	-1136.1	$\Upsilon = (0,1,1,1,1,1,1,1,0,1)$	15	-1146.1
US.UR	$\Upsilon = (1,0,0,1,1,1,1,1,0,0)$	40	-2115.3	$\Upsilon = (1,0,0,1,1,1,1,1,0,1)$	20	-2070.8	$\Upsilon = (1,1,0,1,1,1,1,1,0,0)$	16	-2068.7
US.CPI	$\Upsilon = (1,1,1,0,0,0,0,0,0,1)$	65	-5250.6	$\Upsilon = (1,1,1,1,0,0,0,0,0,1)$	34	-5249.2	$\Upsilon = (1,1,1,0,1,0,0,0,0,1)$	1	-5242.3
US.IR	$\Upsilon = (1,0,1,1,0,0,0,0,0,0)$	70	-7123.5	$\Upsilon = (1,0,1,1,1,0,0,0,0,0)$	24	-7113.2	$\Upsilon = (1,0,1,1,1,1,0,0,0,1)$	6	-7109.5
US.CC	$\Upsilon = (1,1,0,1,1,1,0,0,0,0)$	23	-4694.2	$\Upsilon = (1,1,0,1,1,1,0,1,0,0)$	17	-4685.4	$\Upsilon = (1,1,0,1,1,1,0,0,0,1)$	16	-4642.5
US.Imp	$\Upsilon = (1,0,1,0,0,0,0,0,0,0)$	38	-677.03	$\Upsilon = (1,0,0,0,0,0,0,0,0,0)$	31	-673.26	$\Upsilon = (1,0,1,1,0,0,0,0,0,0)$	17	-661.66
IT.IP	$\Upsilon = (1,0,1,1,1,1,1,0,0,0)$	30	-979.12	$\Upsilon = (1,0,1,1,1,1,1,1,0,0)$	24	-974.51	$\Upsilon = (1,0,1,1,1,1,0,0,0,0)$	10	-951.67
IT.TA	$\Upsilon = (1,0,1,1,0,0,0,0,0,1)$	45	-913.10	$\Upsilon = (1,0,1,1,1,0,0,0,0,1)$	28	-912.73	$\Upsilon = (1,1,1,1,0,0,0,0,0,0)$	10	-905.16

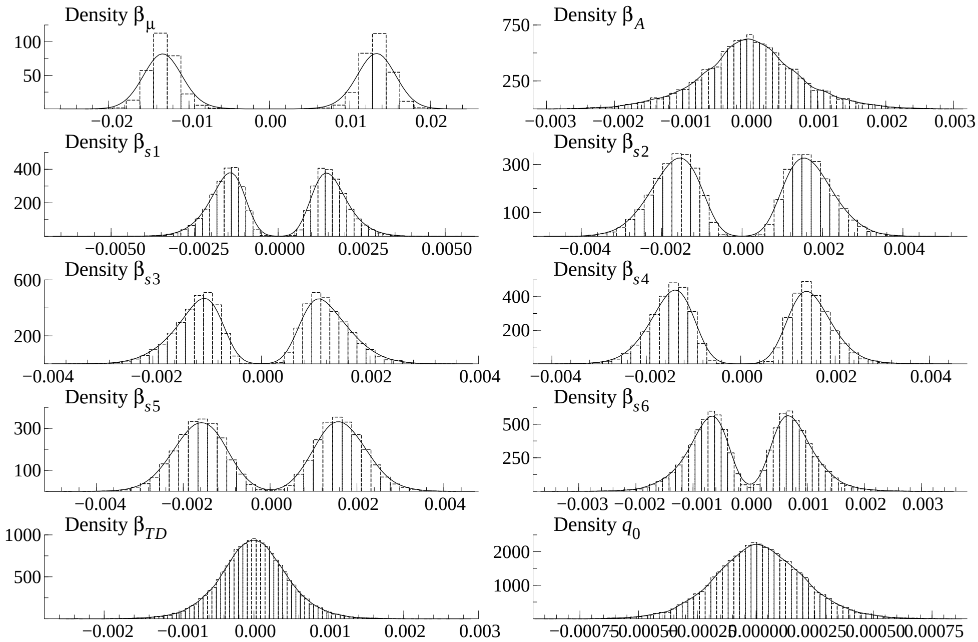

Figure 1.1: IT.IP series: estimated posterior densities of the parameters β_μ, β_A, $\beta_{sj}, j = 1, \ldots, 6,$, β_{TD} and q_0.

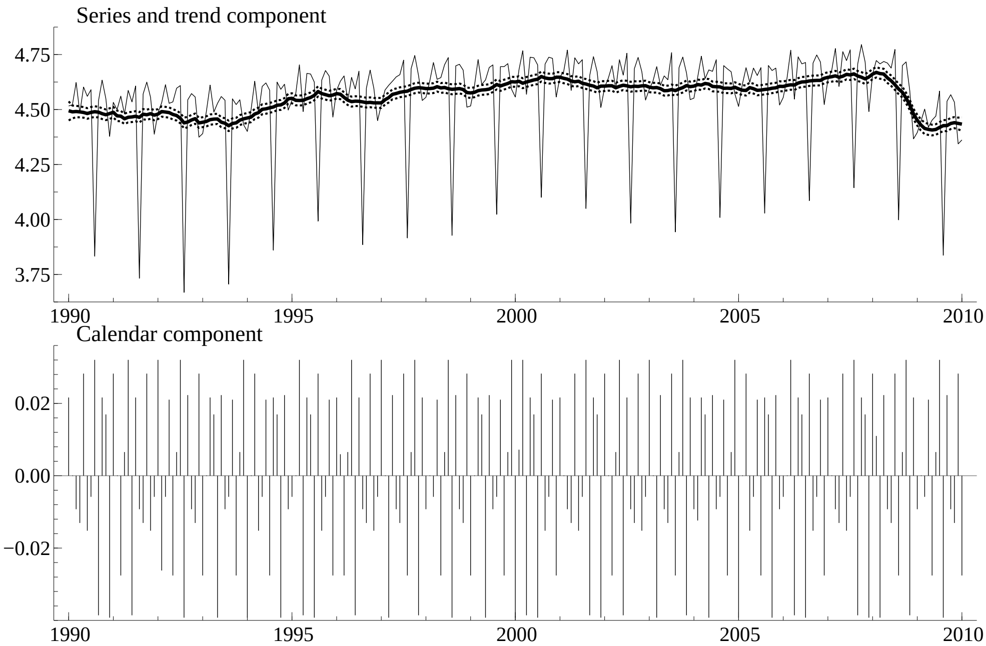

Figure 1.2: IT.IP series: posterior means of the unobserved components.

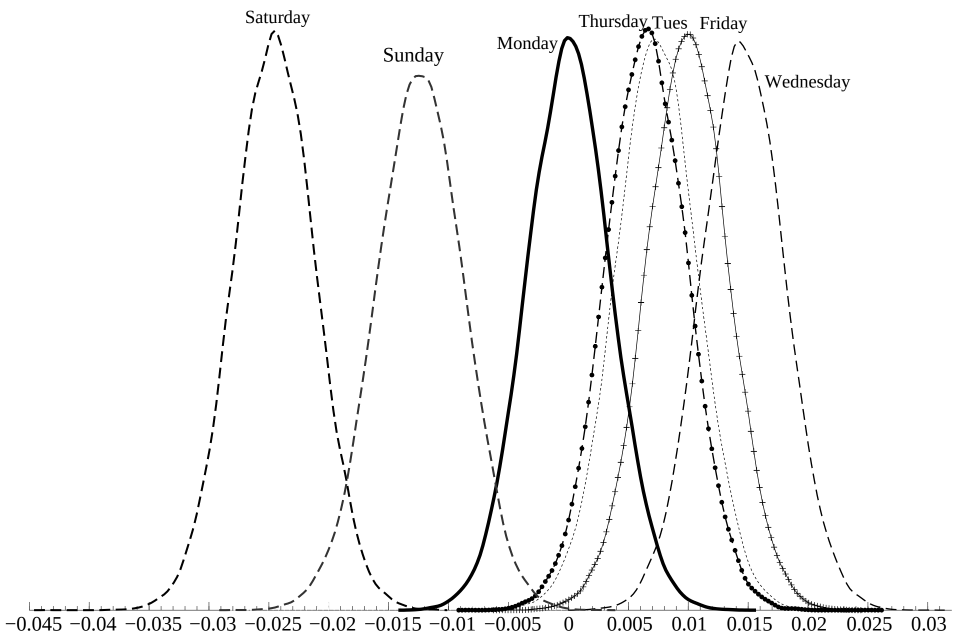

Figure 1.3: IT.IP series: estimated posterior densities of the trading days coefficients ϕ_{k0}, $k = 1, \ldots, 6$. The Sunday effect has been obtained as $\phi_{70} = -\sum_1^6 \phi_{k0}$.

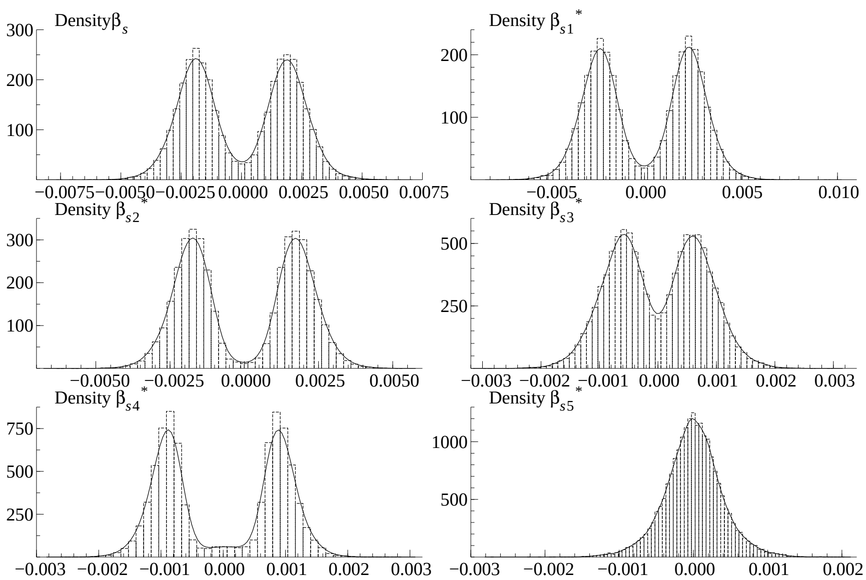

Figure 1.4: IT.IP series. Posterior densities of the seasonal parameters β_s and $\beta^*_{sj}, j = 1, \ldots, 5$, of the reparameterized model 1.3.9.

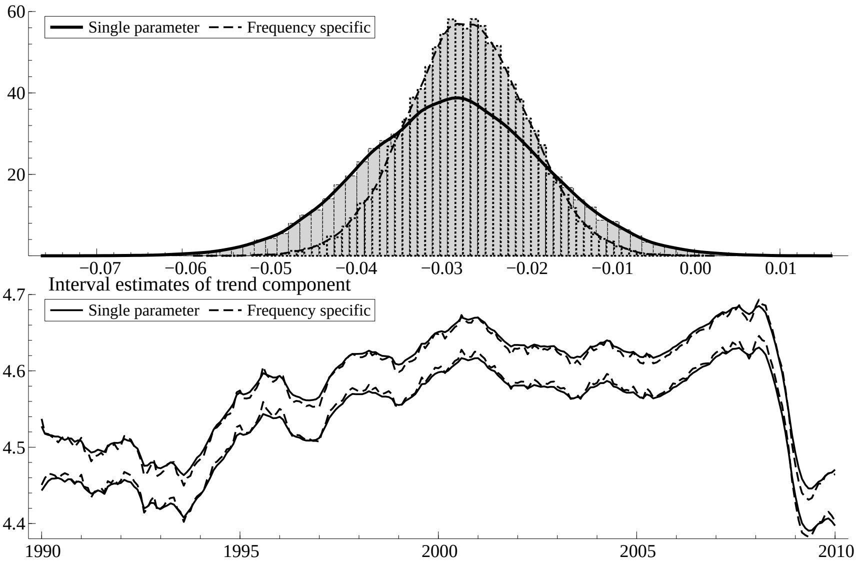

Figure 1.5: IT.IP series. Upper panel: posterior densities of the Easter coefficient for model (1.3.6) with frequency specific coefficients and the restricted specification with $\sigma_j^2 = \sigma_\omega^2, j = 1, \ldots, 5, \sigma_6^2 = 0.5\sigma_\omega^2$ (single variance parameter). Lower panel: interval estimates of trend component.