Noname manuscript No.
(will be inserted by the editor)

Stochastic trends and seasonality in economic time series: new evidence from Bayesian stochastic model specification search

Tommaso Proietti · Stefano Grassi

Received: date / Accepted: date

Abstract An important issue in modelling economic time series is whether key unobserved components representing trends, seasonality and calendar components, are deterministic or evolutive. We address it by applying a recently proposed Bayesian variable selection methodology to an encompassing linear mixed model that features, along with deterministic effects, additional random explanatory variables that account for the evolution of the underlying level, slope, seasonality and trading days. Variable selection is performed by estimating the posterior model probabilities using a suitable Gibbs sampling scheme.

The paper conducts an extensive empirical application on a large and representative set of monthly time series concerning industrial production and retail turnover. We find strong support for the presence of stochastic trends in the series, either in the form of a time-varying level, or, less frequently, of a stochastic slope, or both. Seasonality is a more stable component, although in at least 60% of the cases we were able to select one or more stochastic trigonometric cycles. Most frequently the time variation is found in correspondence with the fundamental and the first harmonic cycles.

An interesting and intuitively plausible finding is that the probability of estimating time-varying components increases with the sample size available. However, even for very large sample sizes we were unable to find stochastically varying calendar effects.

Keywords Nonstationarity · Variable selection · Linear Mixed Models · Seasonality.
J.E.L. Classification: E32, E37, C53.

Dipartimento di Economia e Finanza, Università di Roma, Tor Vergata and CREATES
Via Columbia 2, 00133 Rome, Italy,
E-mail: tommaso.proietti@uniroma2.it

School of Economics, Keynes College and CREATES
Canterbury, Kent, UK.
Tel: 00441227824715
E-mail: S.Grassi@kent.ac.uk.

1 Introduction

Economic time series, recorded at monthly time intervals, exhibit trends, seasonality and the effects due to the aliasing of the weekly cycle in economic activity. Modeling and extracting these components has represented an important problem in time series analysis. See Zellner (1978) Zellner (1983) Nerlove et al. (1979), Harvey (1989) Hylleberg (1992), Peña et al. (2001), and Ghysels and Osborn (2001), among others.

Figure 1 displays the index of industrial production for total manufacturing for Italy. The series shows trending behaviour and a strong seasonal pattern. It is definitely less straightforward to be able to spot the effect of trading days and moving festivals from the graph, but their contribution is also relevant. An interesting question is whether these components can be adequately represented by deterministic functions of time. For instance, the trend may be modelled by a time polynomial, and the seasonal component by a combination of sine and cosine functions with pre-specified frequencies. An alternative view is that these components are subject to random evolution, and thus we need more elaborate stochastic processes to model them. The time series literature offers methods for discriminating the deterministic generation hypothesis against the stochastic one. One approach is performing the class of seasonal unit root tests proposed by Hylleberg et al. (1990), which is based on the finite autoregressive representation of the series and tests for the presence of roots with unit modulus and zero or seasonal phase in the autoregressive polynomial. An alternative approach is to carry out the stationarity tests proposed by Canova and Hansen (1995) and extended by Busetti and Harvey (2003). In this paper we propose to investigate the issue as a model selection problem within a mixture model that encompasses both deterministic and stochastic generation hypotheses. For this purpose, we use the stochastic model specification search proposed by Frühwirth-Schnatter and Wagner (2010) (FS-W henceforth), and applied by Grassi and Proietti (2014) and Proietti and Grassi (2012). The mixture model nests the different specifications for the components, with the elements of the mixture representing the evolution of a particular unobserved component, such as a stochastic level, a stochastic slope, a stochastic trigonometric cycle defined at the fundamental frequency and at the harmonics. By setting up a suitable Gibbs sampling scheme we can sample the indicators of the mixture, as well as the model parameters and underlying state, and obtain a Monte Carlo estimate of the posterior probability for the various different specifications. Deterministic components are obtained by imposing exclusion restrictions. Hence, discriminating between deterministic and stochastic components amounts to performing variable selection within a regression framework that is similar to that considered by George and McCulloch (1993).

The central contribution of this paper lies with the empirical analysis, as we apply the methodology to a dataset consisting of 530 time series, with the aim of assessing the case for the presence of stochastic trends, seasonals and trading days effects in economic time series. For each of the series belonging to the dataset we perform model selection and evaluate the frequency by which time evolving components were selected. Since the available series are characterised by different lengths, we will be able to assess the role of the sample size in the probability of detecting time variation in the components. We find the evidence for the presence of stochastic trends overwhelming, most frequently as a result of a time-varying level, whereas the probability of detecting stochastic variation in the seasonal cycles depends crucially on the length of the available series and on the nature of the prior.

The paper is structured as follows. The reference model will be presented in Section 2, which also reviews the literature and contextualises the various specifications. Section 3 discusses how stochastic model specification search can be applied for the selection of the components of the linear mixed models. This hinges on a convenient reparameterization of the standard deviations of the disturbances that drive the components. Section 4 discusses the state space representation of the non-centered model and the prior specification. It also reports briefly the Gibbs sampling used for model selection and Bayesian estimation. After a brief description of the dataset, Section 5 presents the empirical results. In Section 6 we draw our conclusions. Appendix A provide a detailed description of the Gibbs sampling outlined in Section 4.

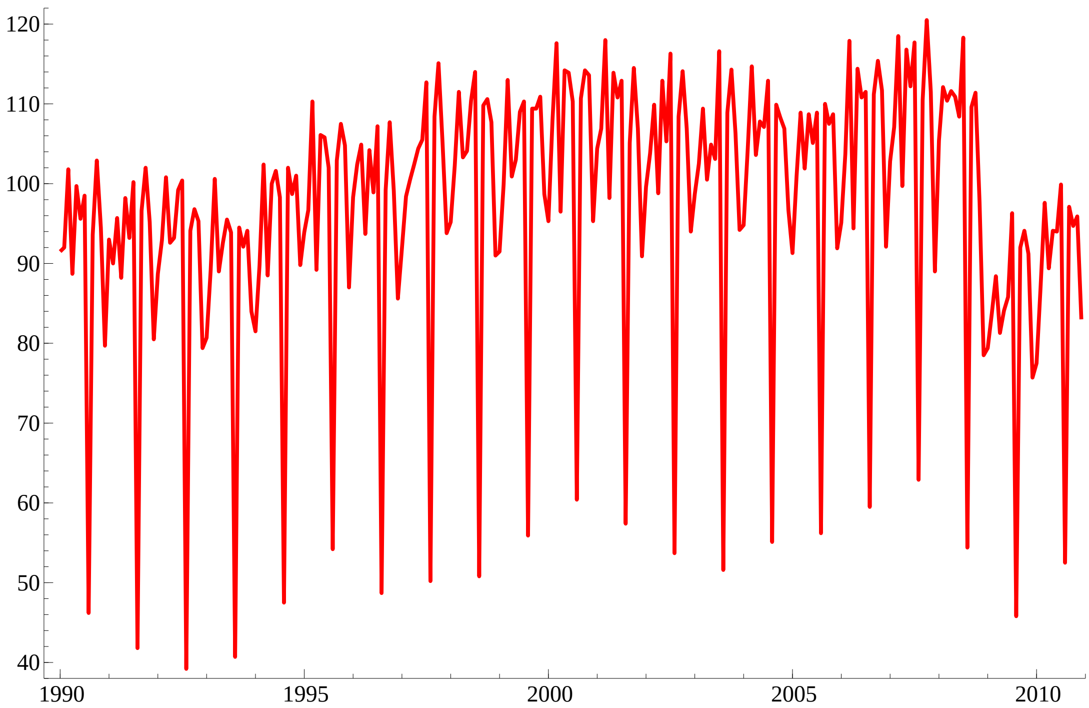

Fig. 1 Italy, Index of Industrial Production for Total Manufacturing. Source: Eurostat, Europa Database.

2 An encompassing linear mixed model with trend and seasonal effects

Let y_t denote a time series observed at $t = 1, 2, \ldots, n$. We focus on modelling y_t by a linear mixed model that accounts for a trend component, denoted μ_t, a seasonal component, S_t, a calendar component, C_t, and an irregular disturbance term, ϵ_t, specified as follows:

$$y_t = \mu_t + S_t + C_t + \epsilon_t, \quad t = 1, \ldots, n, \quad \epsilon_t \sim \text{NID}(0, \sigma_\epsilon^2). \tag{1}$$

The trend component has a deterministic linear part, and a random part, specified as follows:

$$\begin{aligned}
\mu_t &= \mu_0 + q_0 t + \sigma_\eta \tilde{\mu}_t + \sigma_\zeta \tilde{A}_t, \\
\tilde{\mu}_t &= \tilde{\mu}_{t-1} + \tilde{\eta}_t, & \tilde{\eta}_t \sim \text{NID}(0, 1), \\
\tilde{A}_t &= \tilde{A}_{t-1} + \tilde{q}_{t-1}, \\
\tilde{q}_t &= \tilde{q}_{t-1} + \tilde{\zeta}_t, & \tilde{\zeta}_t \sim \text{NID}(0, 1),
\end{aligned} \tag{2}$$

here $\tilde{\mu}_t$ is a random walk component with starting value $\tilde{\mu}_0 = 0$ and unit size; the parameter $\sigma_\eta \geq 0$ establishes the scale of this component. The process $\tilde{A}_t$ is an integrated random walk (such that $\tilde{q}_0 = \tilde{A}_0 = 0$), driven by standard normal disturbances, accounting for the random evolution of the slope; $\sigma_\zeta \geq 0$ is the scale parameter for the component. The trend component (2) can be rewritten equivalently using the following recursions:

$$\begin{aligned} \mu_t &= \mu_{t-1} + q_{t-1} + \eta_t, \; \eta_t \sim \text{NID}(0, \sigma_\eta^2), \\ q_t &= q_{t-1} + \zeta_t, \qquad\qquad \zeta_t \sim \text{NID}(0, \sigma_\zeta^2), \end{aligned} \tag{3}$$

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 trend model is related to cubic spline smoothing (see Wecker and Ansley, 1983). If $\sigma_\zeta = 0$ the trend is a random walk with constant drift; if $\sigma_\eta = 0$ and $\sigma_\zeta > 0$ the trend is an integrated random walk; finally, if both $\sigma_\eta = \sigma_\zeta = 0$ the trend is linear deterministic. The seasonal component results from the sum of six trigonometric cycles defined at the seasonal frequencies $\lambda_j = 2\pi j/12$, $j = 1, \ldots, 6$: the first is defined at the fundamental frequency, $\lambda_1 = \pi/6$ (corresponding to a period of 12 monthly observations) while the others are defined at the harmonic frequencies $\lambda_j = 2\pi j/12$, $j = 2, \ldots, 6$, (corresponding, respectively, 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).

In particular, $S_t = \sum_{j=1}^{6} S_{jt}$, with each S_{jt} made up of a deterministic and a random component, for $j = 1, \ldots, 5$,

$$\begin{aligned} S_{jt} &= a_{j0} \cos \lambda_j t + b_{j0} \sin \lambda_j t + \sigma_j \left(\tilde{a}_{jt} \cos \lambda_j t + \tilde{b}_{jt} \sin \lambda_j t \right), \\ \tilde{a}_{jt} &= \tilde{a}_{j,t-1} + \tilde{\omega}_{jt}, \quad \tilde{\omega}_{jt} \sim \text{NID}(0,1), \;\; t = 1, \ldots, n, \\ \tilde{b}_{jt} &= \tilde{b}_{j,t-1} + \tilde{\omega}^*_{jt}, \quad \tilde{\omega}^*_{jt} \sim \text{NID}(0,1), \end{aligned} \tag{4}$$

with starting values $\tilde{a}_{j0} = \tilde{b}_{j0} = 0$, whereas, for $j = 6$,

$$\begin{aligned} S_{6t} &= a_{60}(-1)^t + \sigma_6 \tilde{a}_{6t}(-1)^t \\ \tilde{a}_{6t} &= \tilde{a}_{6,t-1} + \tilde{\omega}_{6t}, \quad \tilde{\omega}_{6t} \sim \text{NID}(0,1). \end{aligned} \tag{5}$$

The parameters a_{j0} and b_{j0}, $j = 1, \ldots, 6$, determine the amplitude of the fixed trigonometric cycles, whereas σ_j regulate the contribution of the random component. Using trigonometric identities, the j-th seasonal cycle in equation (4) can be rewritten as $S_{jt} = \varphi_t \cos(\lambda_j t - \vartheta_t)$, where

$$\varphi_t = \sqrt{\left(a_{j0} + \sum_{k=0}^{t-1} \tilde{\omega}_{j,t-k} \right)^2 + \left(b_{j0} + \sum_{k=0}^{t-1} \tilde{\omega}^*_{j,t-k} \right)^2}$$

is the time varying amplitude and

$$\vartheta_t = \tan^{-1} \left(\frac{a_{j0} + \sum_{k=0}^{t-1} \tilde{\omega}_{j,t-k}}{b_{j0} + \sum_{k=0}^{t-1} \tilde{\omega}^*_{j,t-k}} \right)$$

represents the phase shift. If $\sigma_1 = \cdots = \sigma_6 = 0$, the seasonal component is the sum of six perfectly deterministic cycles. The recursive representation of the j-th seasonal cycle is

$$\begin{bmatrix} S_{jt} \\ S^*_{jt} \end{bmatrix} = \begin{bmatrix} \cos \lambda_j & \sin \lambda_j \\ -\sin \lambda_j & \cos \lambda_j \end{bmatrix} \begin{bmatrix} S_{j,t-1} \\ S^*_{j,t-1} \end{bmatrix} + \begin{bmatrix} \varpi_{j,t} \\ \varpi^*_{j,t} \end{bmatrix}, j = 1, \ldots, 5, \tag{6}$$

and $S_{6,t} = -S_{6,t-1} + \varpi_{6t}$, where $\varpi_{jt} \sim \text{NID}(0, \sigma_j^2)$, $j = 1, \ldots, 6$, $\varpi^*_{jt} \sim \text{NID}(0, \sigma_j^2)$, $j = 1, \ldots, 5$. These recursions hold for $t = 1, \ldots, n$, with starting values $S_{j,0} = a_{j0}$ and $S^*_{j,0} = b_{j0}$. This representation is the one usually adopted in the time series literature (see Harvey, 1989 and West and Harrison, 1997).

The calendar component, C_t, plays an important role for the class of economic time series that we investigate in the paper; the component accounts for trading days (TD) effects and for moving festivals. The former are related to the fact that the number of weekdays and weekend days is not the same across the months. Let D_{jt} denote the number of days of type j, $j = 1, \dots, 7$, occurring in month t, and define $x_{kt} = D_{jt} - D_{7t}, k = 1, \dots, 6$, which is a contrast between the number of days of a particular type (Mondays, Tuesdays, ..., Saturdays), and the number of Sundays occurring in the same month. A time varying trading day component can be modelled as a regression component with time-varying coefficients:

$$\begin{array}{ll} TD_t = \sum_{k=1}^{6} \phi_{k0} x_{kt} + \sigma_\nu \left(\sum_{k=1}^{6} \tilde{\phi}_{kt} x_{kt} \right), & \\ \tilde{\phi}_{kt} \;\; = \tilde{\phi}_{k,t-1} + \tilde{\nu}_t, & \tilde{\nu}_t \sim \text{NID}(0,1). \end{array} \quad (7)$$

The coefficients associated with the regressors x_{kt} evolve as independent random walks with starting value $\tilde{\phi}_{k0} = 0$. Obviously, if $\sigma_\nu = 0$ the trading days effect are time invariant. As far as moving festivals are concerned, we focus on Easter and Labor Day (U.S. time series), and model their effects defining explanatory variable measuring the proportion of 7 days before Easter (x_{Et}) or Labor Day (x_{Lt}) that fall in month t and subtracting their monthly long run average, computed over the first 400 years of the Gregorian calendar (1583-1982). This treatment is quite standard in the literature; see Bell and Hillmer (1983), among others, and the references therein. Finally, the irregular component is a Gaussian white noise process, $\epsilon_t \sim \text{NID}(0, \sigma_\epsilon^2)$.

The linear mixed model proposed above sufficiently general to accommodate both deterministic and stochastic trends, seasonals and calendar effects. The specification with σ_j constant across j and $\sigma_{\tilde{\nu}} = 0$ is referred to as the basic structural model; see Harvey (1989). The representation used for the components is known as the non-centred (with respect to location and scale) representation; Frühwirth-Schnatter and Wagner (2010) (FS-W henceforth) and Strickland et al. (2007) discuss its advantages for Bayesian estimation of the model.

3 Bayesian stochastic specification search

A widely debated issue is whether trends, seasonals and trading day effects are deterministic or stochastically evolving over time; this translates into the following main specification issues with respect to the model set up in Section 2:

- when $\sigma_\zeta = 0$, *ceteris paribus*, the trend changes are white noise around a constant drift;
- when $\sigma_\eta = \sigma_\zeta = 0$, *ceteris paribus*, the trend is linear deterministic;
- when $\sigma_j = 0, \forall j$, *ceteris paribus*, seasonality is represented by a set of perfectly periodic deterministic components;
- when $\sigma_\nu = 0$, *ceteris paribus*, the TD coefficients are time invariant.

In the econometric literature formal statistical tests are available for discriminating deterministic trends from stochastic ones. When seasonality is absent, unit root tests, see Dickey and Fuller (1979) and Phillips and Perron (1988), test the null of integration versus a stationary alternative; for the Bayesian approach to unit root testing see De Jong and Whiteman (1991), Koop (1992), Sims (1988), Sims and Uhlig (1991), Phillips (1991), Schotman and van Dijk (1991), Phillips and Perron (1994), among others. On the contrary, the tests proposed by Nyblom and Makelainen (1983) and Kwiatkowski et al. (1992) take trend stationarity as the null hypothesis against the alternative of integration. Unit root tests were extended to the seasonal case by Hylleberg et al. (1990), whereas the extension for stationarity tests was proposed by Canova and Hansen (1995), and Busetti and Harvey (2003). Other important references on whether seasonality is stochastically evolving over time include Hylleberg and Pagan (1997) and Koop and van Dijk (2000). The issue as to whether trading days affects are time varying has been addressed by Dagum et al. (1993), Dagum and Quenneville (1993), Bell and Martin (2004).

We can decide on the above main specification issues using the specification search methodology proposed by FS-W. The approach starts with the linear mixed model representation presented in Section 2 and proceeds to the reparameterization of the hyperparameters representing standard deviations as regression parameters with unrestricted support, as it will be illustrated shortly.

It should be noticed that the linear mixed model is identified up to sign switches that operate on both the standard deviations and on the underlying stochastic components. Consider, for instance the trend component in equation (3): if we replace $\sigma_\eta \tilde{\mu}_t$ by the product $(-\sigma_\eta)(-\tilde{\mu}_t)$, i.e. we switch the sign to both the elements, we obtain an observationally equivalent representation, characterised by exactly the same likelihood. FS-W came up with the clever idea of replacing $\sigma_\eta \tilde{\mu}_t$ with $\beta_\mu \mu_t^*$, where, for $t = 1, \dots, n$,

$$\beta_\mu \mu_t^* = \begin{cases} \sigma_\eta \tilde{\mu}_t, & \text{with probability } 0.5 \\ (-\sigma_\eta)(-\tilde{\mu}_t), & \text{with probability } 0.5. \end{cases}$$

Hence, the sign switch is the outcome of a Bernoulli random experiment, with 50% success probability. According to this setting, the parameter β_μ can take any real value and it would be suitable to set up a normal prior for it centred in zero. The same reasoning can be applied to 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 $(-\sigma_\nu)\left(-\sum_{k=1}^{6}\tilde{\phi}_{kt}x_{kt}\right)$ and $\sigma_\nu\left(\sum_{k=1}^{6}\tilde{\phi}_{kt}x_{kt}\right)$. The likelihood function is symmetric around zero along the parameter space and multimodal, if the true standard deviations are larger than zero, as resulting from the identifiability issue. This feature will be later exploited to judge whether the posteriors of σ_η, $\sigma_\zeta, \sigma_j, j = 1, \dots, 6$, and σ_ν, are far away from zero or sufficiently close to it. The random switch process can be formalised by defining independent Bernoulli random variates with success probability 0.5, $\mathsf{B}_\mu, \mathsf{B}_A, \mathsf{B}_{sj}, j = 1, \dots, 6, \mathsf{B}_{TD}$, so that we can use the reparameterisation $\sigma_\eta \tilde{\mu}_t = \beta_\mu \mu_t^*$, 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 $\beta_A = (-1)^{\mathsf{B}_A}\sigma_\zeta, A_t^* = (-1)^{\mathsf{B}_A}\tilde{A}_t$, $\sigma_j\left(\tilde{a}_{jt}\cos\lambda_j t + \tilde{b}_{jt}\sin\lambda_j t\right) = \beta_{sj}U_{jt}^*$, $\beta_{sj} = (-1)^{\mathsf{B}_{sj}}\sigma_j, U_{jt}^* = (-1)^{\mathsf{B}_{sj}}\left(\tilde{a}_{jt}\cos\lambda_j t + \tilde{b}_{jt}\sin\lambda_j t\right)$, for $j = 1, \dots, 6$, and

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

As stated above, the reparameterisation aims at transforming a standard deviation into a regression parameter in a linear mixed model, so that the selection of an evolutive component is related to the inclusion of a particular explanatory variable. The different specifications for the trend and the seasonal components 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). Although in principle we could conduct variable selection for any of the explanatory variables in the model, for our purposes, it will suffice to carry it out on the slope term $q_0 t$, 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 thus introduce nine binary indicator variables $\gamma_\mu, \gamma_A, \gamma_{sj}, j = 1, \dots, 6, \gamma_{TD}$, taking value 1 if the random effects $\mu_t^*, A_t^*, U_{jt}, j = 1, \dots, 6, \left(\sum_{k=1}^{6}\Phi_{kt}^* x_{kt}\right)$ 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 $q_0 t$

is included in the model. The ten indicators can be further collected in the multinomial vector $\Upsilon = (\gamma_\mu, \gamma_A, \gamma_{sj}, j = 1, \ldots, 6, \gamma_{TD}, \delta)$.

As there are $U = 10$ possible variables to select, considering all the possible values of Υ, there are $K = 2^U = 1024$ possible models in competition, which are nested in the specification:

$$\begin{array}{lll}
y_t & = \mu_t + S_t + C_t + \epsilon_t, & \epsilon_t \sim \text{NID}(0, \sigma^2_\epsilon), \\
\\
\mu_t & = \mu_0 + \delta q_0 t + \gamma_\mu \beta_\mu \mu^*_t + \gamma_A \beta_A A^*_t, & \\
\mu^*_t & = \mu^*_{t-1} + \tilde{\eta}_t, & \tilde{\eta}_t \sim \text{NID}(0, 1), \\
A^*_t & = A^*_{t-1} + \tilde{q}_{t-1}, & \\
\tilde{q}_t & = \tilde{q}_{t-1} + \tilde{\zeta}_t, & \tilde{\zeta}_t \sim \text{NID}(0, 1), \\
\\
S_t & = \sum_{j=1}^5 (a_{j0} \cos \lambda_j t + b_{j0} \sin \lambda_j t) + a_{60}(-1)^t + \sum_{j=1}^6 \gamma_{sj} \beta_{sj} U^*_{jt}, & \\
U^*_{jt} & = A^*_{jt} \cos \lambda_j t + B^*_{jt} \sin \lambda_j t, \;\; j = 1, \ldots, 5, & U^*_{6t} = A^*_{6t} \cos \pi t, \\
A^*_{jt} & = A^*_{j,t-1} + \tilde{\omega}_{jt}, & \tilde{\omega}_{jt} \sim \text{NID}(0, 1), \\
B^*_{jt} & = B^*_{j,t-1} + \tilde{\omega}^*_{jt}, & \tilde{\omega}^*_{jt} \sim \text{NID}(0, 1), \\
\\
C_t & = \sum_{k=1}^6 \phi_{k0} x_{kt} + \gamma_{TD} \beta_{TD} \left(\sum_{k=1}^6 \Phi^*_{kt} x_{kt} \right) + \phi_E x_{Et} + \phi_L x_{Lt}, & \\
\Phi^*_{kt} & = \Phi^*_{k,t-1} + \tilde{\nu}_t, & \tilde{\nu}_t \sim \text{NID}(0, 1),
\end{array} \tag{8}$$

where we have defined $A^*_{jt} = (-1)^{\mathsf{B}_{sj}} \tilde{a}_{jt}, B^*_{jt} = (-1)^{\mathsf{B}_{sj}} \tilde{B}_{jt}, \Phi^*_{kt} = (-1)^{\mathsf{B}_{TD}} \phi^*_{kt}$.

All the specifications will include 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 seasonal and fixed calendar effects, corresponding to $\Upsilon = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$. When $\Upsilon = (0, 0, 0, 0, 0, 0, 0, 0, 0, 1)$, the model features a deterministic linear trend and a perfectly deterministic seasonal component (assuming there is no moving festival):

$$y_t = \mu_0 + q_0 t + \sum_{j=1}^5 (a_{j0} \cos \lambda_j t + b_{j0} \sin \lambda_j t) + a_{60}(-1)^t + \sum_{k=1}^6 \phi_{k0} x_{kt} + \epsilon_t. \tag{9}$$

In turn, $\Upsilon = (1, 0, 1, 0, 0, 0, 0, 0, 0, 1)$ corresponds to a specification featuring a stochastic trend, represented by a random walk with constant nonzero drift, fixed seasonals and calendar effects, and a stochastic seasonal cycle defined at the fundamental frequency:

$$\begin{array}{c}
y_t = \mu_0 + q_0 t + \sigma_\eta \tilde{\mu}_t + \sum_{j=1}^5 (a_{j0} \cos \lambda_j t + b_{j0} \sin \lambda_j t) + a_{60}(-1)^t + \\
\sigma_1 \left(\tilde{a}_{1t} \cos \lambda_1 t + \tilde{b}_{1t} \sin \lambda_1 t \right) + \sum_{k=1}^6 \phi_{k0} x_{kt} + \epsilon_t.
\end{array} \tag{10}$$

The different models will be labelled by

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

where Υ_u is the u-th element of the vector Υ, $u = 1, \ldots, U$. For instance, $\Upsilon = (1, 0, 1, 0, 0, 0, 0, 0, 0, 1)$ is model M_{642}.

4 Statistical Treatment

Depending on the value of Υ, the models nested in (8) admit the following state space representation:

$$\begin{aligned} y_t &= x'_{\delta,t}\rho_\delta + z'_{\gamma,t}\alpha_{\gamma,t} + \epsilon_t, \quad \epsilon_t \sim \text{NID}(0,\sigma^2_\epsilon), \quad t = 1,\dots,n, \\ \alpha_{\gamma,t} &= T_\gamma \alpha_{\gamma,t-1} + R_\gamma u_{\gamma,t}, \qquad u_{\gamma,t} \sim \text{NID}(0, I), \end{aligned} \tag{11}$$

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

$$\begin{aligned} x_{\delta,t} &= (1, \delta t, \cos\lambda_1 t, \sin\lambda_1 t, \dots, \cos\pi t, x_{1t}, \dots, x_{6t}, x_{Et}, x_{Lt})', \\ \rho_\delta &= (\mu_0, q_0, a_{10}, b_{10}, \dots, a_{60}, \phi_1, \dots, \phi_6, \phi_E, \phi_L)', \\ z_{\gamma,t} &= (\gamma_\mu\beta_\mu, \gamma_A\beta_A, 0, \gamma_{s1}\beta_{s1}\cos\lambda_1 t, \gamma_{s1}\beta_{s1}\sin\lambda_1 t, \dots, \gamma_{s6}\beta_{s6}\cos\pi t, \\ &\qquad \gamma_{TD}\beta_{TD}x_{1t}, \dots, \gamma_{TD}\beta_{TD}x_{6t})', \\ \alpha_{\gamma,t} &= (\mu^*_t, A^*_t, \tilde{q}_t, A^*_{1t}, B^*_{1t}, \dots, A^*_{6t}, \Phi^*_{1t}, \dots, \Phi^*_{6t}), \end{aligned}$$

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

In the sequel we will denote by α the collection of the latent states $\{\alpha_{\gamma,t}, t = 0,1,\dots,n\}$, and by ψ_Υ the appropriate subset of the parameters $(\mu_0, q_0, a_{10}, b_{10}, \dots, a_{60}, \ \phi_{10}, \dots, \phi_{60}, \beta_\mu, \beta_A, \beta_{s1}, \dots, \beta_{s6}, \beta_{TD})$ that enter the model for a particular value of Υ.

In matrix notation, we write the 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.

Model selection entails the computation of the posterior model probabilities $\pi(M_k|y) \propto \pi(M_k)$ $\pi(y|M_k)$, where y denotes the collection of time series values $\{y_t, t = 1,\dots,n\}$. The evaluation of the marginal likelihood $\pi(y|M_k)$ for each model is computationally intensive; in fact, it would be unfeasible to compute the posterior model probabilities for each of the 1024 specifications and select the specification characterised by the largest. It is feasible instead to draw samples from the posterior distribution of Υ given the data by MCMC methods, as by a suitable design of the priors the full conditional posterior distribution of the multinomial vector Υ is available in closed form. A suitable Gibbs sampling (GS) scheme can in fact be devised which enables Υ to be sampled along with the model parameters and states. After the GS scheme has converged, we estimate $\pi(\Upsilon|y)$, by the proportion of times a particular specification was drawn.

The prior has the following conditional independence structure:

$$\pi(\Upsilon, \psi_\Upsilon, \sigma^2_\epsilon, \alpha) = \pi(\Upsilon)\pi(\sigma^2_\epsilon)\pi(\psi_\Upsilon|\Upsilon, \sigma^2_\epsilon)\pi(\alpha|\Upsilon),$$

where individual factors are given as follows.

- As far as $\pi(\Upsilon)$ is concerned, we write $\pi(\Upsilon) = \pi(\gamma_\mu)\pi(\gamma_A)\cdots\pi(\gamma_{TD})\pi(\delta)$ and, following Ley and Steel (2009), we adopt a hierarchical prior for the probability of including a specific component. For instance, we write $\pi(\gamma_\mu = 1) = \varsigma$, where ς is a beta random variable with hyperparameters $h_a, h_b > 0$, $\varsigma \sim \mathcal{B}(h_a, h_b)$, and the same prior is specified for the other indicator variables. We set $h_a = 1$, as in Ley and Steel (2009), and $h_b = 4$, which corresponds to an expected model size, $\bar{M} = 10\frac{1}{1+h_b}$ equal to 2 (there are 10 variables to select from).
  Hence, the prior puts most mass on the null model. This is a mildly conservative prior that requires some data evidence to favour the inclusion of regressors. We compare the results with the alternative prior assumption that the models $M_k, k = 1,...,K$, are equally likely a priori, that is $\pi(\Upsilon) = 2^{-10}$, which amounts to setting $\varsigma = 0.5$, so that the expected model size is $\bar{M} = 5$.

– For σ_ϵ^2 we adopt a hierarchical inverse Gamma (IG) prior: $\sigma_\epsilon^2 \sim \mathrm{IG}(c_0, C_0)$, where $C_0 \sim \mathrm{G}(g_0, G_0)$, $\mathrm{G}(\cdot)$ denoting the Gamma distribution, $c_0 = 2.5$, $g_0 = 5$, and $G_0 = g_0/[0.75\mathrm{Var}(y_t)(c_0 - 1)]$. The hierarchical prior is intended to make the posteriors less sensitive to the choice of the hyperparameters of the IG distribution.
– Denoting the i-th element of the parameter vector ψ_Υ by ψ_{Υ_i}, $i = 1\dots,p$, we set $\pi(\psi_\Upsilon|\sigma_\epsilon^2) = \prod_{i=1}^p \pi(\psi_{\Upsilon_i}|\sigma_\epsilon^2)$, where all the priors are conjugate. For instance, $q_0|\sigma_\epsilon^2 \sim \mathrm{N}(0, d_0\sigma_\epsilon^2)$, etc.
For the constant term and the coefficients $a_{j0}, j = 1,\dots,6, b_{j0}, j = 1,\dots,5, \phi_{k0}, k = 1,\dots,6$ we adopt the uninformative priors $\pi(\mu_0|\sigma_\epsilon^2) \propto 1$.
For the remaining parameters, $\beta_\mu,\dots,\beta_{TD}$, two prior specifications will be considered:
  1. **Fixed scale priors**. A distinctive feature of the FS-W methodology is the adoption of Gaussian priors, centered at zero, for the parameters $\beta_\mu, \beta_A, \beta_{sj}, \beta_{TD}$:
$$\beta_\mu|\sigma_\epsilon^2 \sim \mathrm{N}(0, \kappa_\mu\sigma_\epsilon^2), \quad \beta_A|\sigma_\epsilon^2 \sim \mathrm{N}(0, \kappa_A\sigma_\epsilon^2),$$
$$\beta_{sj}|\sigma_\epsilon^2 \sim \mathrm{N}(0, \kappa_j\sigma_\epsilon^2), j = 1,\dots,6, \quad \beta_{TD}|\sigma_\epsilon^2 \sim \mathrm{N}(0, \kappa_{TD}\sigma_\epsilon^2).$$
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. $\beta_\mu = \pm\sigma_\eta$, $\beta_\mu|\sigma_\epsilon^2 \sim \mathrm{N}(0, \kappa_\mu\sigma_\epsilon^2)$, instead of the usual inverse Gamma prior for the variance parameter σ_η^2.
  2. **Fractional priors**. We also consider the fractional prior of O'Hagan (1995), which, in the context of model (11), see Frühwirth-Schnatter and Wagner (2010), is defined as: $p(\psi_\Upsilon \mid \sigma_\epsilon^2) \propto p(y \mid \psi_\Upsilon, \sigma_\epsilon^2)^b$, where $b > 0$ is a fraction between 0 and 1, close to zero. The fractional prior can be interpreted as the posterior of a non-informative prior and a fraction b of the data y and amount in our case to setting: $\psi_\Upsilon \mid \sigma_\epsilon^2 \sim N\left((Z_\Upsilon' Z_\Upsilon)^{-1}(Z_\Upsilon)y, \sigma_\epsilon^2(Z_\Upsilon' Z_\Upsilon)^{-1}/b\right)$.
– The prior distribution for α is given directly by the Gaussian dynamic model (11):
$$\pi(\alpha|\Upsilon) = \pi(\alpha_{\gamma 0})\prod_{t=1}^n \pi(\alpha_{\gamma t}|\alpha_{\gamma,t-1}),$$
with $\alpha_{\gamma t}|\alpha_{\gamma,t-1} \sim \mathrm{N}(T_\gamma\alpha_{\gamma,t-1}, R_\gamma R_\gamma')$ and $\alpha_{\gamma,0} = 0$.

The estimation of the posterior density $\pi(\Upsilon, \psi_\Upsilon, \sigma_\epsilon^2, \alpha|y)$ is performed by a Gibbs sampling scheme that can be sketched as follows. After specifying a set of initial values $\Upsilon^{(0)}, \sigma_\epsilon^{2(0)}, \alpha^{(0)}, \psi_\Upsilon^{(0)}$, the following steps are iterated for $i = 1, 2, \dots, M$:

a. Draw $\Upsilon^{(i)} \sim \pi(\Upsilon|\alpha^{(i-1)}, y)$;
b. Draw $\sigma_\epsilon^{2(i)} \sim \pi(\sigma_\epsilon^2|\Upsilon^{(i)}, \alpha^{(i-1)}, y)$;
c. Draw $\psi_\Upsilon^{(i)} \sim \pi(\psi_\Upsilon|\Upsilon^{(i)}, \sigma_\epsilon^{2(i)}, \alpha^{(i-1)}, y)$;
d. Draw $\alpha^{(i)} \sim \pi(\alpha|\Upsilon^{(i)}, \sigma_\epsilon^{2(i)}, \psi_\Upsilon^{(i)}, y)$.

The above complete conditional densities are available, up to a normalizing constant, from the form of the likelihood and the prior. Details are provided in appendix 6.

5 Empirical Results

Our application deals with data set consisting of 530 monthly time series for 10 Euro Area countries, the UK, and the US, referring to the index of industrial production and retail turnover. This is a large and representative sample, with 379 series referring to the index of industrial production (IIP) and 151 to the index of retail turnover (RT). The breakdown of the series by country and their sample period is available in Table 1.

Table 1 Breakdown of the series by country, sample period, number of time series.

	Index of industrial production		*Index of Retail sales*	
Country	Sample period	Number of series	Sample period	Number of series
Austria	1996.1-2010.12	28	1999.1-2010.12	6
Belgium	1995.1-2010.12	27	1998.1-2010.12	15
Finland	1990.1-2010.12	20	1995.1-2010.12	14
France	1990.1-2010.12	28	1994.1-2010.12	14
Germany	1991.1-2010.12	28	1994.1-2010.12	15
Greece	2000.1-2010.12	29	1995.1-2010.12	13
Italy	1990.1-2010.12	27	2000.1-2010.12	14
Netherlands	1990.1-2010.12	22	1996.1-2010.12	9
Portugal	1995.1-2010.12	20	1995.1-2010.12	9
Spain	1980.1-2010.12	28	2000.1-2010.12	14
UK	1990.1-2010.12	28	2000.1-2010.12	14
US	1947.1-2010.12	94	1992.1-2010.12	14

For the IIP we consider series from Sectors B (Mining and quarrying), C (Manufacturing), D (Energy). The series for the manufacturing sectors are from those identified by two digits of the NACE statistical classifications of economic activities (sectors C1-C31). For the US we consider the 63 series for Market and Industry Group and the 32 series for Special Aggregates and Selected Detail (see `http://www.federalreserve.gov/releases/g17/table1_2.htm` for more details).

For retail turnover we focus on the series available with code starting with G47 (Retail trade, except of motor vehicles and motorcycles). The sources of the series are Eurostat (`http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/`), the Federal Reserve and the US Census Bureau. All the series are analysed in logarithms.

5.1 Stochastic model specification search results

Stochastic model specification search was conducted using the hierarchical prior on the model space, discussed in Section 3, implying an expected model size consisting of 2 random effects (in addition to the time-invariant explanatory variables). As far as the priors on the regression parameters are concerned, we considered two values for the scale factors $d_0 = \kappa_\mu = \kappa_A = \kappa_j = \kappa_{TD} = \kappa$, namely $\kappa = 10$ and $\kappa = 100$, and two values for the fraction b of the fractional prior described in the previous section, namely 10^{-4} and 10^{-5}.

The Gibbs Sampling scheme was iterated 100,000 times, including a burn–in sample of 20,000 draws. Hence, all the results presented refer to 80,000 MCMC draws. For each series y_{it}, $i = 1, \dots, N$, we recorded the 10 modal models, denoted $\Upsilon_{ik(r)}$, $r = 1, \dots, 10$, visited by the GS scheme, as well as the number of times they were visited. The estimated posterior probability of the model are

$$\hat{\pi}_{ik(r)} = \frac{c_{ik(r)}}{c_i},$$

where $c_{ik(r)}$ denotes the number of times model M_k was selected for series i, whereas c_i is the total number of draws (which is actually invariant with i, $c_i = 80,000$).

Limiting ourselves to the first 10 modal models is motivated by storage constraints, and is not at all restrictive. For instance, in the fixed-scale prior case with $\kappa = 100$ the median percentage of draws absorbed by the top 10 modal models across the 530 series amounts to 99.41%, and the distribution of $\sum_{r=1}^{10} \hat{\pi}_{ik(r)}$, i.e. of the total probability attached to the 10 modal specification, turns out to be highly concentrated around 100%, as shown by figure 2.

Fig. 2 Distribution of $\sum_{r=1}^{10} \hat{\pi}_{ik(r)}$ for the 530 series in our dataset.

The output of the stochastic model specification search are the 5,300 indicator vectors

$$\Upsilon_{ik(r)} = \left(\gamma_{\mu,k(r)}, \gamma_{A,k(r)}, \ldots, \delta_{k(r)}\right), \quad i = 1, \ldots, 530, \quad r = 1, \ldots, 10,$$

referring to the 10 modal models for the 530 time series and the associated estimated posterior probabilities $\hat{\pi}_{ik(r)}$. Their distribution is displayed in figure 3, separately for the two prior assumptions, in the model space: the horizontal axis refers to $M_k, k = 1, \ldots, 1024$, and each bar is proportional to the average number of times the corresponding specification was visited per series.

The empirical evidence can be summarised as follows:

- The posterior distribution of Υ_i is highly concentrated around some modal models; in particular, model M_{513}, corresponding to $\Upsilon = (1, 0, 0, 0, 0, 0, 0, 0, 0, 0)$, plays a prominent role. This specification features a stochastic level, no slope and a fixed seasonal pattern. Another important specification in the case of fixed-scale priors is M_{514}, which is as the same as M_{513}, but with a time-invariant non zero slope.
- The heterogeneity of the distribution $\{M_{ik(r)}, \pi_{ik(r)}, i = 1, \ldots, 530, r = 1, \ldots, 10\}$ varies according to the prior assumptions: typically, it is smaller for less informative priors and if a fractional prior is taken.
- The empirical evidence for a stochastically evolving seasonal component is larger for the fractional priors. In particular, when $b = 10^{-4}$, the modal model is M_{1022}, corresponding to

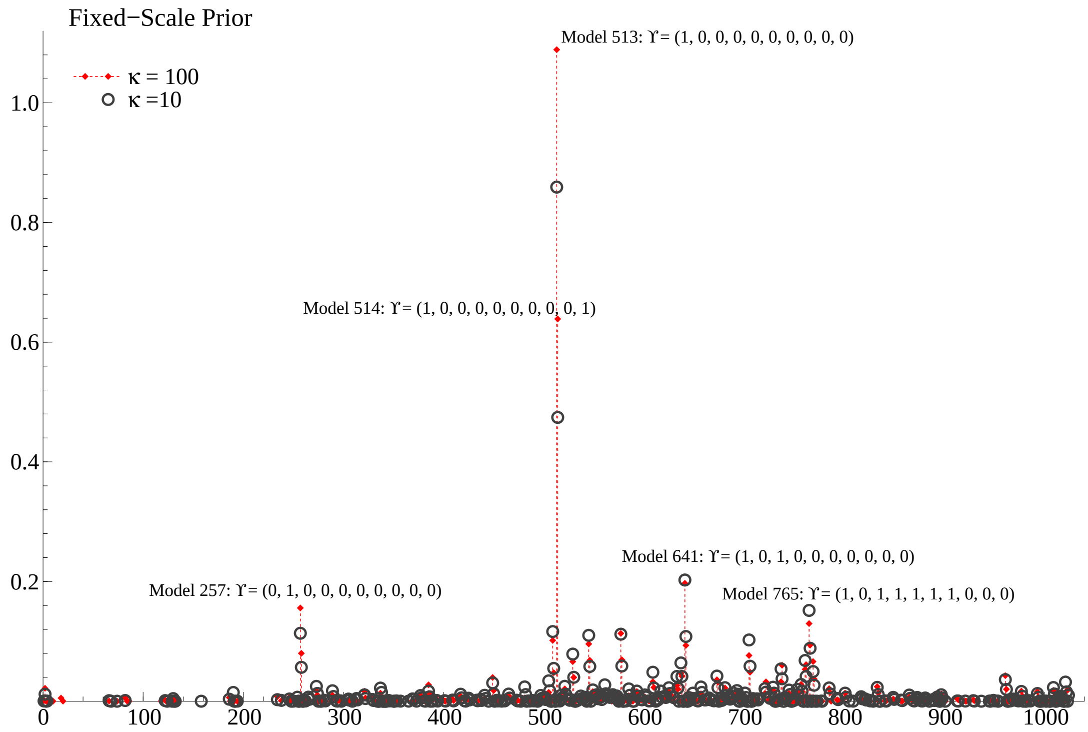

Fig. 3 Distribution of selected models in the model space. The horizontal axis refers to $M_k, k = 1, \ldots, 1024$. Each symbol is proportional to the average number of times the corresponding specification was visited per series. The corresponding Υ vector is also reported.

$\Upsilon = (1, 1, 1, 1, 1, 1, 1, 1, 0, 1)$. This model features stochastic level, slope and seasonals, and a fixed calendar component.

– Specifications featuring some but not all stochastically evolving seasonal cycles are non ignorable: for instance, models M_{641} and M_{765} play an important role and are such that only the fundamental (M_{641}) cycle and all the cycle except that defined at the Nyquist frequency (M_{765}) are time-varying.

To obtain an estimate of the marginal probabilities that a particular component is present in series i we compute

$$\hat{\pi}_i = \frac{\sum_{r=1}^{10} c_{ik(r)} \Upsilon_{ik}}{\sum_{r=1}^{10} c_{ik(r)}},$$

where, with obvious notation, $c_{ik(r)}$ is the number of times model M_k was selected as the r-th modal model for series i. This is a 1×10 vector containing the estimated probability that each component is present in the first ten modal specifications that were recorded, as it amounts to summing up the values $c_{ik(r)}/\sum_{r=1}^{10} c_{ik(r)}$ over those specifications which contain the component, i.e. for which the elements of Υ are equal to one. However, as we have illustrated, the modal specifications absorb the quasi totality of the GS draws, so that the vector $\hat{\pi}_i$ can be thought of as an approximation to the true marginal probabilities given the data.

The distribution of the $\hat{\pi}_i, i = 1, \ldots, 530$, can be effectively represented graphically using a biplot; see Gower and Hand (1996) and Greenacre (2010). The latter is a two-dimensional display that is based on the singular value decomposition of the 530×10 matrix Π, obtained by stacking the row vectors $\hat{\pi}_i$.

In the graph, obtained using the BiplotGUI package, described in La Grange et al. (2009), the individual series are represented as points and the columns are represented as calibrated axes, as advocated by Gower and Hand (1996). For simplicity we only present the case referring to the fixed-scale prior with $\kappa = 100$. Points are marked by a circle if they refer to the IIP series, whereas the RT series are marked by a square.

The interpretation of the biplot is such that (the best rank 2 approximation to) the individual probabilities $\hat{\pi}_{ik(r)}$ are obtained from the orthogonal projection of the point representing the series on the calibrated axis representing the r-th component. Moreover, the Euclidean distances among the points are an approximation to the Mahalanobis distances between the vectors $\hat{\pi}_i$, so that series that follow similar models are represented close in the plane. The calibrated axes are defined by the eigenvectors corresponding to the two largest eigenvalues of the covariance matrix of Π. The orientation of the axes can be gauged from the position of the labels. For instance slope and level span essentially the same subspace, which we can label the trend subspace, but move along opposite directions, which is a consequence of the negative correlation between the corresponding columns of the matrix Π. Also, the trend subspace is almost orthogonal to the space spanned by the seasonal components. This is interesting, and conforms our expectation that the probability of detecting a time-varying seasonality should not depend on the trend; on the contrary, the probability of detecting a time-varying slope is at odds with that of detecting a stochastic level.

The cluster of points in the lower left part of the graph refers to series for which the level is stochastic, the slope is fixed (the probability of selecting this component is well below the average) and seasonality is stable, i.e. the projection of those points along the Seas1-6 axes is low. This is the most numerous cluster. The points to the left will display stochastic seasonality as well. On the contrary, the set of points on the top left corner are characterised by low probabilities for stochastic level and seasonal cycles, but the probability that the slope is stochastic is high. The biplot also illustrates that most industrial production series feature a stochastic level and a fixed slope.

The vectors $\hat{\pi}_i$ can be further aggregated across the series to yield:

$$\hat{\pi} = \sum_{i=1}^{N} \hat{\pi}_i w_i, \quad w_i = \sum_{r=1}^{10} c_{ik(r)} / \sum_{i=1}^{N} \sum_{r=1}^{10} c_{ik(r)}.$$

Here the weight is needed to adjust for the fact that the number of draws pertaining to the first 10 modal models varies with the series.

The elements of this vector are presented in Table 2 for the 379 industrial production series (IIP), the 151 retail turnover series (RT) and for all the 530 series. The main result is that the marginal probability of having a stochastic level component is very high, and it is much higher for the industrial production series. The probability of having a stochastic slope is higher for the retail series, and it is very low for the IIP series, which confirms the patterns displayed in the biplot. Thus, we find that the trend component in the IIP series has a different characterisation than the RT series. The marginal probability of a stochastically time-varying seasonal trigonometric cycle is always less than 50% and tends to decrease with j, as order of the harmonic cycles increases. Time-varying trading day effects are never detected.

The probability of a stochastically evolving seasonal component is larger if the prior is less diffuse (i.e. if κ and the fraction b are smaller) and more generally if a fractional prior is considered.

Additional important stylized facts are obtained by presenting the joint frequencies by which given stochastic components are detected. In particular, we focus on the joint frequency distribution for the indicators γ_μ, γ_A and $I(\sum_j \gamma_j > 0)$; the latter is the indicator for the presence of at least one stochastic cycle at anyone of the seasonal frequencies. We refer for simplicity to the fixed-scale prior case with $\kappa = 100$. This analysis is presented in table 3, separately for the complete dataset

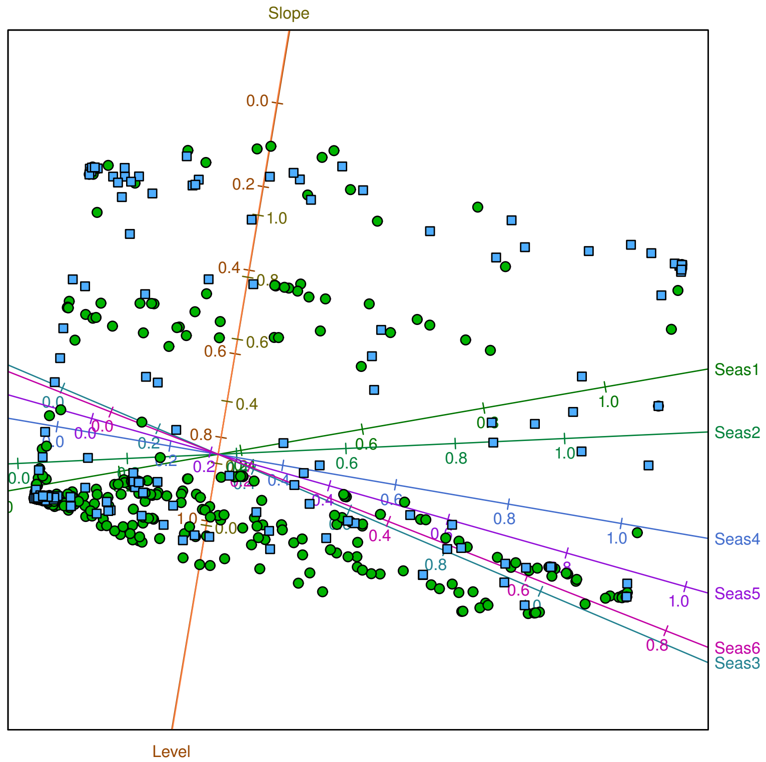

Fig. 4 Principal components biplot of $\hat{\pi}_i, i = 1, \ldots, 530$. Circles represent industrial production (IIP) series, and squares represent retail sales (RT) series. The orientation of the calibrated axes is provided by the position of the labels.

(Total, 530 series), the industrial production series (IIP, 379 series), the retail turnover series (RT, 151 series), the subset of very long time (VL) series, consisting of the US industrial production series (for which more than sixty years of data are available), the subset of long time series (L), featuring 195 series having more less than 30 but more that 18 years of data; (all the series except the US RT series belong to the IIP group), the subset of medium sized series (M), featuring more than 12 and no more than 18 years of data (149 series are in this group), and finally the subset of short time series, with at most 12 years of data (this subsets comprises 96 time series).

The table reports the proportion of the MCMC draws referring to the 10 modal specifications that featured a particular combination of the indicators. A stochastic trend (either $\gamma_\mu = 1$ or $\gamma_A = 1$, or both) is detected in most occurrences; only in the case of short time series a completely deterministic trend was found in 2.54% of the draws. The modal representation for short time series is (1,0,0), i.e. it features a stochastic level, but deterministic slope and seasonals. If we consider the entire dataset, the modal representation (42.62%) features a stochastic level and at least one stochastic seasonal cycle.

An interesting finding is that the frequency by which stochastic slope and seasonality are detected depends inversely on the sample size. The percentage of specifications featuring all three

Table 2 Probability of identifying a particular component for different prior specification.

	Fixed-Scale Prior, $\kappa = 100$			Fixed-Scale Prior, $\kappa = 10$		
Component of Υ	IIP	RT	Total	IIP	RT	Total
Stochastic Level γ_μ	0.93	0.62	0.84	0.92	0.57	0.82
Stochastic Slope γ_A	0.17	0.39	0.23	0.17	0.42	0.24
Stochastic Seas1 γ_{s1}	0.34	0.42	0.36	0.38	0.46	0.40
Stochastic Seas2 γ_{s2}	0.35	0.41	0.37	0.40	0.45	0.41
Stochastic Seas3 γ_{s3}	0.31	0.36	0.32	0.38	0.41	0.39
Stochastic Seas4 γ_{s4}	0.25	0.36	0.28	0.31	0.43	0.35
Stochastic Seas5 γ_{s5}	0.19	0.26	0.21	0.24	0.30	0.26
Stochastic Seas6 γ_{s6}	0.14	0.18	0.15	0.17	0.21	0.18
Time-Varying Calendar γ_{TD}	0.00	0.00	0.00	0.00	0.00	0.00
Drift δ	0.35	0.49	0.39	0.34	0.38	0.35
	Fractional Prior $b = 10^{-5}$			Fractional Prior $b = 10^{-4}$		
Component of Υ	IIP	RT	Total	IIP	RT	Total
Stochastic Level γ_μ	0.95	0.75	0.89	0.93	0.71	0.87
Stochastic Slope γ_A	0.32	0.54	0.38	0.38	0.64	0.46
Stochastic Seas1 γ_{s1}	0.42	0.66	0.49	0.50	0.74	0.57
Stochastic Seas2 γ_{s2}	0.52	0.72	0.58	0.60	0.78	0.65
Stochastic Seas3 γ_{s3}	0.53	0.71	0.58	0.60	0.78	0.65
Stochastic Seas4 γ_{s4}	0.48	0.74	0.55	0.55	0.79	0.62
Stochastic Seas5 γ_{s5}	0.43	0.67	0.50	0.52	0.73	0.58
Stochastic Seas6 γ_{s6}	0.33	0.57	0.40	0.42	0.66	0.49
Time-Varying Calendar γ_{TD}	0.00	0.00	0.00	0.00	0.00	0.00
Drift δ	0.35	0.49	0.39	0.43	0.56	0.47

Table 3 Joint frequency distribution of the three indicators γ_μ, γ_A and $I(\sum_j \gamma_j > 0)$ for the complete dataset (Total), the industrial production series subset (IIP), the retail turnover series subset (IIP), the subsets consisting of very long time series (VL), long time series (L), medium sized (M), and short time series (S).

γ_μ	γ_A	$I(\sum_j \gamma_j > 0)$	Total	IIP	RT	VL	L	M	S
0	0	0	0.44	0.13	1.22	0.00	0.00	0.00	2.54
0	0	1	0.93	0.14	2.90	0.00	0.39	0.01	4.85
0	1	0	4.49	2.20	10.27	0.00	1.18	7.66	10.70
0	1	1	9.35	3.82	23.26	1.26	8.39	18.19	4.64
1	0	0	32.89	36.01	25.03	25.11	22.31	39.57	49.30
1	0	1	42.62	46.92	31.78	32.69	64.21	33.39	26.49
1	1	0	1.97	2.44	0.78	9.62	0.42	0.13	0.34
1	1	1	7.31	8.32	4.77	31.32	3.11	1.05	1.15

random components (corresponding to the triple (1,1,1) of the indicators) is 31.32 for the U.S. industrial production series, which are all classified as VL (very long). This percentage decreases quite rapidly as the sample sizes decreases. The different results for the IIP and RT subsets may be also the consequence of the different sample sizes of the series making up the two groups, the RT series being much shorter.

A further analysis that we conducted was to assess the sensitivity of the results to different prior model probabilities assumptions. In particular, we considered the fixed-scale prior with $\kappa = 100$, but assuming that the models are equally likely a priori (yielding a larger expected model size a priori). The heterogeneity of the posterior model distribution decreased, as can be seen also by comparing[1] the estimated selection probabilities, displayed in table 4 with those reported in the upper left panel of table 2.

Table 4 Probability of identifying a particular component estimated using a uniform prior on the model space.

Component of Υ	IIP	RT	Total
Stochastic Level γ_μ	0.93	0.60	0.84
Stochastic Slope γ_A	0.18	0.43	0.25
Stochastic Seas1 γ_{s1}	0.39	0.46	0.41
Stochastic Seas2 γ_{s2}	0.42	0.47	0.43
Stochastic Seas3 γ_{s3}	0.41	0.44	0.42
Stochastic Seas4 γ_{s4}	0.34	0.47	0.37
Stochastic Seas5 γ_{s5}	0.24	0.34	0.27
Stochastic Seas6 γ_{s6}	0.18	0.23	0.19
Time-Varying Calendar γ_{TD}	0.00	0.00	0.00
Drift δ	0.34	0.38	0.35

Finally to check MCMC convergence we can run the first 1000 draws, for each time series, both from a full ($\Upsilon = (1,1,1,1,1,1,1,1,1,1)$) and a sparse ($\Upsilon = (1,1,0,0,0,0,0,1,1,0)$) model and we can check, if the selected model and the corresponding posterior probabilities are more or less the same. Table 5 reports an experiment that refers to a subset of the series available, namely the general index for each country. The table reports the first (M_1) and second (M_2) selected model in the unrestricted (S_1) and restricted (S_2) case. The table shows that the procedure is quite reliable, in fact the selected models are the same and the corresponding posterior probabilities are almost the same in both cases.

5.2 Illustrative example

After getting to the broad picture emerging from our analysis, we discuss the results for the Italian industrial production series plotted in figure 1. The series is fairly representative and serves to illustrate the sensitivity to the prior assumptions, amongst other things. In table 6 we report the first three models that were visited more frequently by the Gibbs sampler; the last column reports the posterior probability of the model, estimated by MCMC as $\hat{\pi}_{ik}$.

The evidence is overwhelmingly in favour of a stochastic level component, as, regardless of our prior assumption, the random walk component μ_t^* is featured by all the modal specifications. There is little or no evidence for the presence of a slope component: a stochastic slope is never selected, whereas a constant slope is featured e.g. by specification M_{514}. As for stochastic seasonality, the selection results depend on the nature of the prior. The evidence for an evolutive seasonal component is strong for the fractional prior and rather weak for the fixed scale prior. The choice of the prior also affects the distribution of the estimated posterior probabilities, which is more heterogeneous for the fractional prior, a circumstance that we have already observed.

Do these differences matter for characterising the time series properties of the data? To address this issue, figure 5 compares the posterior mean of the trend, seasonal and calendar components estimated by Gibbs sampling for models M_{513} (selected under a fixed-scale prior with $\kappa = 100$) and

[1] The results are based on 80,000 draws, with burn-in sample of 20,000.

Table 5 Sensitivity to starting values. S_1 represent the full model, corresponding to $\Upsilon = (1,1,1,1,1,1,1,1,1,1)$. S_2 represent the null restricted model (e.g model 791) corresponding to $\Upsilon = (1,1,0,0,0,0,0,1,1,0)$. Finally M_1 and M_2 represent the first and the second most selected model. The percentage of the selected model is reported in parenthesis. Prior specification $\kappa = 100$.

	Austria		Belgium		Finland		France	
	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2
S_1	514 (56.20)	513 (25.50)	553 (49.39)	569 (19.98)	641 (53.54)	130 (40.08)	513 (71.03)	514 (28.70)
S_2	514 (57.60)	513 (34.50)	553 (54.93)	569 (22.08)	641 (51.12)	130 (42.41)	513 (72.40)	514 (27.21)
	Germany		Greece		Italy		Netherlands	
	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2
S_1	513 (71.87)	514 (27.16)	257 (70.10)	258 (26.30)	513 (51.00)	529 (23.60)	513 (70.68)	514 (28.61)
S_2	513 (67.78)	514 (31.37)	257 (70.68)	258 (25.74)	513 (46.10)	529 (22.70)	513 (70.80)	514 (28.50)
	Portugal		Spain		UK		US	
	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2
S_1	513 (69.26)	514 (27.20)	421 (36.05)	417 (21.30)	641 (72.46)	642 (22.10)	889 (23.60)	865 (18.17)
S_2	513 (58.10)	514 (40.20)	421 (42.30)	417 (21.64)	641 (69.90)	642 (23.30)	889 (26.23)	865 (18.53)

M_{569} (selected under a fractional prior with $b = 10^{-5}$). The main differences arise with respect to the seasonal component, which towards the end of the sample is subject to an amplitude reduction in the fractional case. The estimated trend and the calendar components do not vary relevantly, implying that the two specifications will differ also for the estimated contribution of the irregular component. As a result, the different specifications may impact on the seasonally adjusted series.

5.3 Convergence and Diagnostics

The specification selected according to its posterior model probability is estimated by MCMC and the convergence of the chain can be assessed by means of the Geweke statistic. In this subsection we will consider the case of a fixed scale prior with $\kappa = 100$. Let $\psi^{(i)}$ denote the i-th draw of a generic parameter ψ produced by the GS scheme, after the initial burn-in period. Let also $\bar{\psi}_a$ denote the average of the first $\mathcal{M}_a$ draws, $\bar{\psi}_b$ that of the last $\mathcal{M}_b$ draws, which are taken sufficiently remote to prevent any overlap, the Geweke's convergence statistic (Geweke, 1992, 2005) is defined as

$$C_G = \frac{\bar{\psi}_a - \bar{\psi}_b}{\sqrt{V_{L,a}/\mathcal{M}_a + V_{L,b}/\mathcal{M}_b}},$$

where

$$V_{L,k} = c_{0,k} + 2\sum_{j=1}^{\mathcal{M}_k - 1} w_j c_{j,k}, \quad k = a, b,$$

is the long run variance of the parameter sample path for the $\mathcal{M}_k$ draws, based on a weighted combination of the autocovariances of the draws at lag j, $c_{j,k}$, with weights w_j that are decreasing in j and ensure that $V_{L,k} \geq 0$. A customary choice is the set of linearly declining weights $w_j = \frac{l-j}{l+1}, j = 1, \ldots, l$, where l is the truncation parameter.

Table 6 Estimation results for the Italian Industrial Production Total Series.

M_k	γ_μ	γ_A	γ_{s1}	γ_{s2}	γ_{s3}	γ_{s4}	γ_{s5}	γ_{s6}	γ_{TD}	δ	$100 \times \hat{\pi}_{ij}$
Fixed-Scale Prior, $\kappa = 100$											
513	1	0	0	0	0	0	0	0	0	0	50.93
529	1	0	0	0	0	1	0	0	0	0	23.59
514	1	0	0	0	0	0	0	0	0	1	13.45
Fixed-Scale Prior, $\kappa = 10$											
M_k	γ_μ	γ_A	γ_{s1}	γ_{s2}	γ_{s3}	γ_{s4}	γ_{s5}	γ_{s6}	γ_{TD}	δ	$100 \times \hat{\pi}_{ij}$
529	1	0	0	0	0	1	0	0	0	0	41.12
513	1	0	0	0	0	0	0	0	0	0	22.05
530	1	0	0	0	0	1	0	0	0	1	18.10
Fractional Prior, $b = 10^{-5}$											
M_k	γ_μ	γ_A	γ_{s1}	γ_{s2}	γ_{s3}	γ_{s4}	γ_{s5}	γ_{s6}	γ_{TD}	δ	$100 \times \hat{\pi}_{ij}$
569	1	0	0	0	1	1	1	0	0	0	24.88
570	1	0	0	0	1	1	1	0	0	1	20.50
762	1	0	1	1	1	1	1	0	0	1	11.43
Fractional Prior, $b = 10^{-4}$											
M_k	γ_μ	γ_A	γ_{s1}	γ_{s2}	γ_{s3}	γ_{s4}	γ_{s5}	γ_{s6}	γ_{TD}	δ	$100 \times \hat{\pi}_{ij}$
765	1	0	1	1	1	1	1	1	0	0	28.22
761	1	0	1	1	1	1	1	0	0	0	20.61
633	1	0	0	1	1	1	1	0	0	0	7.94

In our particular case, we set $\mathcal{M}_a = 20,000$, $\mathcal{M}_b = 40,000$ and $l = 100$ for all the parameters. Recalling that the total number of MCMC draws is 80,000, the first block represents the first 25% of the draws, whereas the second is formed by the last 50% of the draws; the two blocks are thus well separated.

Table 7 refers to a subset of the series available, namely the general index for each country, and reports, for each of the parameters of the modal specification, the Geweke convergence diagnostic, C_G, and the simulation inefficiency factors (SIF), calculated according to Kim et al. (1998) as

$$\hat{R}_{\mathcal{B}_m} = 1 + 2\frac{\mathcal{B}_m}{\mathcal{B}_m - 1}\sum_{j=1}^{\mathcal{B}_m} K(j/\mathcal{B}_m)\hat{\rho}(j), \tag{12}$$

where $K(\cdot)$ the Parzen kernel, $\mathcal{B}_m$ is the bandwidth, chosen equal to 1500, and $\hat{\rho}(j)$ the j-th order autocorrelation of the draws $\psi^{(i)}$. Large values imply that the draws are strongly and positively autocorrelated, so that the mixing of the chain is very slow and the information content of an additional draw is small.

The convergence statistics are highly satisfactory; in particular, with only a few exceptions (referring e.g. to a_{40} and a_{50} for IIP Germany, γ_δ for IIP Austria, etc.), C_G is almost never significant. Also, SIF is small in most cases; it is typically larger for the initial level μ_0.

As a misspecification check we compute the Ljung-Box test statistic of the null of no residual autocorrelation. The statistic is based on the pseudo-innovations $v_t = y_t - \tilde{y}_{t|t-1}$, where $\tilde{y}_{t|t-1} = \mathrm{E}(y_t|y_1, \ldots, y_{t-1}, \tilde{\psi}_\Upsilon)$, and $\tilde{\psi}_\Upsilon$ is the posterior mean of the parameters. The pseudo-innovations are computed by the Kalman filter, see Durbin and Koopman (2012). Obviously, v_t

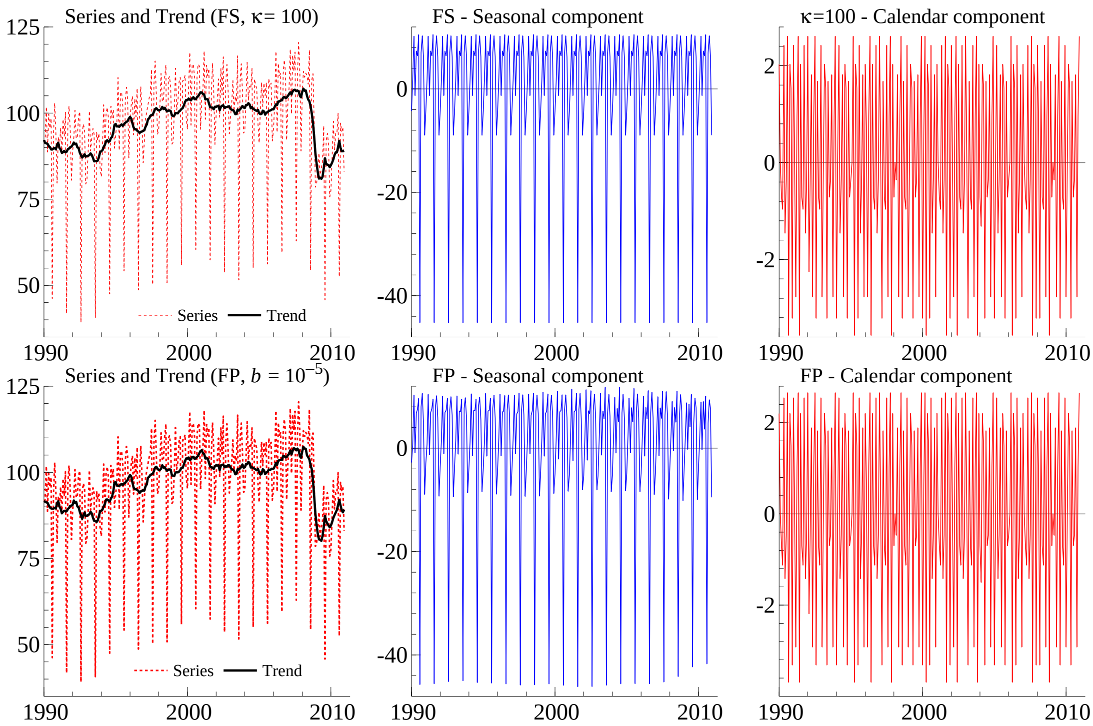

Fig. 5 Italian Index of Industrial Production, Total. Posterior mean of the trend, seasonal and calendar components estimated by Gibbs sampling for models M_{513} (selected under a fixed-scale prior with $\kappa = 100$) and M_{569} (selected under a fractional prior with $b = 10^{-5}$.

is not a true innovation sequence, which is $y_t - \mathrm{E}(y_t|y_1, \ldots, y_{t-1})$. The latter could be computed by a particle filter (see e.g. Durbin and Koopman (2012), chapter 12), which however entails being able to simulate from $\pi(\psi_\Upsilon|y_1, \ldots, y_{t-1})$. The test is based on the first 24 autocorrelations of the standardized innovations, where again the variance of the pseudo-innovations is given by the Kalman filter. Table 7 reports the p-value based on the chi-squared null distribution. Only for three series the null is rejected at the 5% level. We can tentatively conclude that the model is sufficiently general to capture the evolution of most of the series without any major misspecification.

6 Conclusions

An important issue in modelling economic time series is whether the most relevant unobserved components, representing trends, seasonality and calendar components, are deterministic or evolutive. Adopting the approach by Frühwirth-Schnatter and Wagner (2010), we argue that deciding upon this issue amounts to performing Bayesian variable selection in a linear mixed models that features, along with deterministic linear trends, trigonometric seasonals and trading days regressors, additional random explanatory variables that account for the evolution of the underlying level, slope, seasonality and trading days. Variable selection is performed by estimating the posterior model probabilities by MCMC, via a Gibbs sampling scheme that samples the indicator variables for the different specifications, along with the parameters and the unobserved states.

The paper has conducted an extensive empirical application on a large and representative set of monthly time series concerning industrial production and retail turnover. We find strong support for the presence of a stochastic trend in the series, either in the form of a time-varying level, or, more rarely, of a stochastic slope, or both. We estimate the probability of detecting a stochastic trend close to 1. There is however a difference in the trend model for the industrial production series and the retail turnover, as for the latter a stochastic slope is more likely to be found, whereas for the former the slope is either fixed or zero in most of the cases.

The evidence for a stochastically evolving seasonal pattern is stronger under a fractional prior and our results depend also on the prior assumptions. As it is well known, the fractional prior is tailored for situations when prior information is weak, and thus it seems appropriate for our case. Nevertheless, the paper does not aim at reaching a conclusion on what prior assumptions should be taken; on the contrary, it has a more exploratory intent and aims at presenting the empirical evidence, as well as at evaluating the sensitivity to different prior assumptions. We conclude that although seasonality is a more stable component, even for a fixed-scale prior, in about 60% of the series we were able to select at least one stochastic trigonometric cycle out of the six possible cycles. Most frequently the time variation is found in correspondence with the fundamental and the first harmonic frequencies.

An interesting intuitive finding is that the probability of estimating time-varying components increases with the sample size available. As pointed out by one of the referees, this outcome partly results from having a fixed scale prior for different sample sizes. The fixed-scale prior specification can lead to overfitting (i.e. a more complex model with stochastically evolving components is selected) for longer time series, as shown by (Casella et al., 2009). On the contrary, the fractional prior automatically corrects for this problem.

The fact that a more flexible model is chosen for larger sample sizes is particularly true of seasonality. However, even for very large sample sizes (such as those available for the US industrial production) we were unable to find stochastically varying calendar effects.

Acknowledgements

The authors wish to thank the Reviewers and the Editor for their careful review and constructive criticism that led to several improvements in the paper. Tommaso Proietti gratefully acknowledges financial support by the Italian Ministry of Education, University and Research (MIUR), PRIN Research Project 2010-2011 - prot. 2010J3LZEN, *Forecasting economic and financial time series: understanding the complexity and modelling structural change*. Both authors acknowledge support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation.

Table 7 Convergence diagnostics for the selected models for the industrial production selected series. Geweke diagnostics (C_G) and simulation inefficiency factors (SIF). The table also reports the P-values of Ljung-Box statistics (*: significant at 5% level).

	Austria		Belgium		Finland		France		Germany		Greece	
Parameter	C_G	SIF	C_G	SIF	C_G	SIF	C_G	SIF	C_G	SIF	C_G	SIF
σ_ϵ^2	-1.49	30.30	1.27	75.32	2.37	4.63	-1.53	26.18	1.01	15.78	-1.42	4.79
γ_μ	0.24	0.80	-0.66	0.89	0.95	0.87	0.43	0.66	0.76	1.12	-	-
γ_A	-	-	-	-	-	-	-	-	-	-	-0.04	1.06
γ_{s1}	-	-	-	-	1.39	0.57	-	-	-	-	-	-
γ_{s2}	-	-	-	-	-	-	-	-	-	-	-	-
γ_{s3}	-	-	1.41	0.48	-	-	-	-	-	-	-	-
γ_{s4}	-	-	-	-	-	-	-	-	-	-	-	-
γ_{s5}	-	-	0.49	1.15	-	-	-	-	-	-	-	-
γ_{s6}	-	-	-	-	-	-	-	-	-	-	-	-
γ_{TD}	-	-	-	-	-	-	-	-	-	-	-	-
γ_δ	-	-	-	-	-	-	-	-	-	-	-	-
a_{10}	-0.64	3.00	0.80	14.23	-0.14	15.80	-0.23	18.37	1.09	8.67	-0.28	0.62
b_{10}	-0.32	6.91	1.12	14.37	1.88	29.87	1.06	13.75	1.70	5.10	1.97	0.74
a_{20}	-0.42	0.63	0.68	0.89	1.87	1.21	0.02	0.70	0.41	1.72	0.65	1.73
b_{20}	-1.81	2.61	-0.12	6.76	1.25	1.81	0.86	3.39	0.87	5.02	-0.56	1.33
a_{30}	-0.34	1.50	1.90	14.87	-0.26	0.72	-0.08	0.81	1.60	2.34	0.73	0.73
b_{30}	-0.07	2.94	1.25	46.45	0.41	0.96	0.60	2.28	0.14	2.06	0.05	0.91
a_{40}	0.76	1.05	2.06	3.95	0.14	1.47	-0.75	1.19	-2.51	2.00	0.57	1.20
b_{40}	0.83	0.88	0.64	3.48	-0.51	1.05	-0.06	2.43	0.69	2.68	-0.14	1.05
a_{50}	-1.77	1.82	0.01	24.84	1.14	0.30	-1.34	2.19	-2.41	0.84	-2.37	1.70
b_{50}	0.38	1.19	-0.92	9.41	1.73	0.83	1.89	1.65	-1.09	1.46	0.52	0.91
a_{60}	1.10	1.35	0.56	1.12	-0.83	0.16	1.59	2.59	-0.14	2.11	-0.80	0.93
ϕ_1	0.08	3.33	-0.04	1.13	1.95	1.81	-1.02	2.09	-0.29	4.17	0.62	1.07
ϕ_2	-0.28	1.20	0.47	2.40	-0.20	1.44	0.47	1.85	-0.16	2.58	0.34	1.02
ϕ_3	-0.40	0.98	-0.70	4.81	-0.09	0.90	0.89	3.37	0.13	1.51	-0.66	0.67
ϕ_4	1.56	0.91	-0.10	3.37	-0.38	0.74	-1.48	2.13	1.72	1.64	-0.98	0.98
ϕ_5	-1.79	1.50	1.17	1.92	0.81	0.17	0.25	0.97	-1.99	0.54	0.32	0.33
ϕ_6	1.23	1.24	0.00	3.05	-0.79	0.59	-0.09	2.52	-0.18	0.68	-0.46	0.54
ϕ_E	0.02	0.93	-1.03	1.88	1.53	0.87	1.11	1.80	-0.77	0.77	-0.61	1.10
μ_0	0.52	181.18	-1.71	543.54	-1.29	33.53	-1.40	262.93	-2.62	229.84	-0.70	8.06
	Ljung-Box		Ljung-Box		Ljung-Box		Ljung-Box		Ljung-Box		Ljung-Box	
	0.1677		0.0269*		0.2072		0.1002		0.1056		0.5833	

Table 7 (continued) Convergence diagnostics for the selected models for the industrial production selected series. Geweke diagnostics (C_G) and simulation inefficiency factors (SIF). The table also reports the P-values of Ljung-Box statistics (*: significant at 5% level).

	Italy		Netherlands		Portugal		Spain		UK		US	
Parameter	C_G	SIF	C_G	SIF	C_G	SIF	C_G	SIF	C_G	SIF	C_G	SIF
σ^2_ϵ	-1.28	24.62	-1.04	11.24	0.22	32.36	1.29	13.37	-0.37	81.49	-1.88	18.74
γ_μ	0.44	0.67	-1.16	1.08	1.00	0.70	-	-	1.32	0.88	-0.96	0.67
γ_A	-	-	-	-	-	-	0.78	0.90	-	-	0.70	1.09
γ_{s1}	-	-	-	-	-	-	-0.30	0.70	0.55	0.70	-	-
γ_{s2}	-	-	-	-	-	-	-	-	-	-	1.17	1.20
γ_{s3}	-	-	-	-	-	-	-1.56	1.06	-	-	0.09	0.88
γ_{s4}	-	-	-	-	-	-	-	-	-	-	-0.57	1.40
γ_{s5}	-	-	-	-	-	-	-	-	-	-	0.29	0.65
γ_{s6}	-	-	-	-	-	-	1.11	0.90	-	-	-	-
γ_{TD}	-	-	-	-	-	-	-	-	-	-	-	-
γ_δ	-	-	-	-	-	-	-	-	-	-	-	-
a_{10}	-0.17	16.03	-1.00	4.03	0.36	1.87	1.94	38.66	-0.11	210.35	-0.01	75.78
b_{10}	0.98	13.37	-0.99	6.40	0.74	4.64	0.45	49.38	-0.91	186.40	-0.02	109.17
a_{20}	-0.04	0.73	-1.57	0.76	0.48	0.83	0.67	0.30	-0.09	1.00	-0.58	135.89
b_{20}	0.80	3.54	0.90	1.65	0.71	2.37	0.10	1.34	-0.29	5.93	-0.34	156.97
a_{30}	0.07	0.76	-0.33	1.30	-0.63	1.56	1.50	19.99	0.44	1.06	-1.28	36.21
b_{30}	0.54	2.34	0.87	2.05	-0.81	1.18	0.81	9.46	0.26	4.57	-0.50	24.19
a_{40}	-0.56	1.15	0.65	1.39	-1.93	0.71	0.44	1.08	-0.56	0.22	-0.33	81.01
b_{40}	-0.09	2.42	-0.68	1.06	0.60	1.77	0.14	1.30	-0.71	1.93	0.78	68.06
a_{50}	-1.42	2.25	0.15	1.11	-1.44	1.06	1.17	1.01	1.56	0.53	0.32	26.54
b_{50}	1.94	1.64	1.46	0.53	1.22	0.85	-1.96	1.91	2.27	0.67	0.93	47.79
a_{60}	1.70	2.48	0.69	0.88	-0.82	0.70	-1.05	20.26	0.49	1.94	-0.08	6.00
ϕ_1	-1.02	1.90	-0.98	1.71	1.20	1.56	0.44	0.74	0.36	2.76	0.75	6.78
ϕ_2	0.46	1.78	2.13	1.03	-1.29	2.28	-0.06	0.77	-0.22	1.02	0.93	5.10
ϕ_3	0.82	3.31	-0.71	1.14	0.95	1.59	-0.06	0.92	0.28	3.27	-0.35	9.36
ϕ_4	-1.44	1.96	-1.06	0.84	-0.16	1.59	0.02	0.85	-0.39	4.35	-1.08	4.69
ϕ_5	0.25	1.06	1.97	-0.03	0.57	1.63	-0.10	0.87	-0.07	3.34	1.03	9.78
ϕ_6	-0.11	2.38	-0.94	0.39	-0.27	1.61	-0.94	2.95	0.41	2.31	1.18	6.13
ϕ_E	1.12	1.87	1.27	1.27	-1.53	0.45	-1.14	1.90	0.92	2.19	1.25	3.87
μ_0	-1.14	279.07	0.64	303.47	-2.13	99.65	-0.89	20.10	1.30	663.08	1.00	213.17
	Ljung-Box		Ljung-Box		Ljung-Box		Ljung-Box		Ljung-Box		Ljung-Box	
	0.1130		0.5742		0.0859		0.0513		0.0747		0.9139	

Table 7 (continued) Convergence diagnostics for the selected models for the selected retail series. Geweke diagnostics (C_G) and simulation inefficiency factors (SIF). The table also reports the P-values of Ljung-Box statistics (*: significant at 5% level).

	Austria		Belgium		Finland		France		Germany		Greece	
Parameter	C_G	SIF	C_G	SIF	C_G	SIF	C_G	SIF	C_G	SIF	C_G	SIF
σ^2_ϵ	1.05	4.51	2.80	3.90	-1.78	89.35	1.07	35.79	-0.53	6.08	1.12	43.89
γ_μ	-1.16	1.47	-1.07	1.39	-	-	-	-	-1.36	1.13	0.11	1.62
γ_A	-	-	-	-	-0.60	1.91	-0.65	0.28	-	-	-	-
γ_{s1}	-	-	-	-	-2.43	0.37	-0.90	0.95	0.36	0.64	-	-
γ_{s2}	-	-	-	-	-0.62	0.47	0.51	1.62	-	-	-	-
γ_{s3}	-	-	-	-	0.35	0.46	0.01	0.71	-	-	-	-
γ_{s4}	-	-	-	-	0.72	1.08	-0.65	0.45	-	-	-	-
γ_{s5}	-	-	-	-	-0.17	0.83	0.22	1.56	-	-	-	-
γ_{s6}	-	-	-	-	-0.93	1.47	-0.09	1.27	-	-	-	-
γ_{TD}	-	-	-	-	-	-	-	-	-	-	-	-
γ_δ	-2.63	21.34	-1.02	17.54	-	-	-	-	-	-	0.16	28.95
a_{10}	-0.13	5.93	4.07	3.00	-0.83	33.05	1.55	48.81	0.05	22.75	-0.82	8.30
b_{10}	1.26	3.11	1.30	1.80	-2.35	44.00	-0.46	70.95	0.19	33.16	-0.64	5.53
a_{20}	1.74	1.55	0.03	1.03	-1.92	38.93	1.11	14.75	0.12	0.26	0.35	1.83
b_{20}	1.45	2.00	-0.42	0.85	0.13	37.03	-2.05	38.79	-0.67	2.64	-0.18	5.80
a_{30}	1.12	1.58	1.39	1.13	-1.58	48.82	1.00	37.67	0.02	1.10	-0.92	1.55
b_{30}	0.49	1.93	-0.80	1.03	-1.95	14.53	-1.13	38.87	0.62	0.92	-1.28	0.71
a_{40}	0.45	1.09	-0.07	0.39	-2.66	45.69	-1.44	47.42	-1.33	1.36	0.13	1.16
b_{40}	1.22	0.90	0.98	0.80	2.58	76.17	-1.17	23.82	-0.74	1.45	-0.14	2.90
a_{50}	0.68	1.28	-0.53	1.34	1.71	13.24	-0.15	33.27	1.28	0.13	0.79	1.48
b_{50}	0.30	1.23	-0.93	2.11	-0.46	38.96	-1.50	15.16	-1.15	2.18	-0.10	1.25
a_{60}	0.27	0.90	1.19	0.73	0.48	37.09	0.02	30.02	0.03	0.63	-1.41	2.37
ϕ_1	0.41	1.11	-0.03	1.08	0.29	11.76	1.22	3.63	-0.46	1.20	-0.05	0.52
ϕ_2	-0.22	0.43	-0.61	1.36	-0.19	6.26	-0.18	3.27	-1.75	1.53	-0.98	2.04
ϕ_3	0.29	0.14	0.26	1.33	0.80	4.51	-0.93	1.02	1.54	0.70	2.61	2.22
ϕ_4	-1.17	0.67	-0.73	1.34	-0.52	3.61	1.34	1.50	0.27	0.90	-2.22	1.53
ϕ_5	1.90	1.10	0.97	0.83	0.98	9.97	1.49	3.17	-0.93	1.45	-0.68	1.52
ϕ_6	-1.23	2.06	0.20	2.05	-0.25	8.91	-0.50	1.49	-0.57	1.38	1.62	3.14
ϕ_E	-1.01	0.92	2.11	1.35	-1.58	4.16	1.06	2.40	-0.60	0.69	0.12	1.03
μ_0	-0.14	17.30	-0.84	57.83	0.72	80.76	1.07	81.05	-0.76	121.06	0.46	186.23
	Ljung-Box		Ljung-Box		Ljung-Box		Ljung-Box		Ljung-Box		Ljung-Box	
	0.6731		0.3548		0.0000*		0.0847		0.4153		0.0001*	

Table 7 (continued) Convergence diagnostics for the selected models for the selected retail series. Geweke diagnostics (C_G) and simulation inefficiency factors (SIF). The table also reports the P-values of Ljung-Box statistics (*: significant at 5% level).

	Italy		Netherlands		Portugal		Spain		UK		US	
Parameter	C_G	SIF	C_G	SIF	C_G	SIF	C_G	SIF	C_G	SIF	C_G	SIF
σ_ϵ^2	1.83	11.18	1.88	20.32	-1.12	32.97	2.57	11.86	-1.05	134.17	1.22	22.66
γ_μ	0.06	0.75	-	1753,25	1.36	0.56	-	-	-0.27	1.09	-	-
γ_A	-	-	-0.85	0.88	-	-	0.55	0.59	-	-	1.30	0.82
γ_{s1}	1.43	0.77	-	-	0.94	1.17	-	-	-	-	1.11	0.90
γ_{s1}	0.56	1.32	-	-	-	-	-	-	-	-	-0.39	1.42
γ_{s2}	-	-	-	-	-0.39	0.37	-	-	-	-	-	-
γ_{s3}	-	-	-	-	-	-	-	-	-	-	-	-
γ_{s4}	-	-	-	-	-	-	-	-	-	-	-	-
γ_{s5}	-	-	-	-	-	-	-	-	-	-	-	-
γ_{TD}	-	-	-	-	-	-	-	-	-	-	-	-
γ_δ	-0.39	19.21	-	-	0.91	29.25	-	-	0.80	20.92	0.29	34.59
a_{10}	0.79	22.09	0.57	1.48	-1.32	24.88	-0.84	1.03	-0.06	9.86	0.03	22.51
b_{10}	-0.40	56.06	0.17	2.39	-1.31	49.47	1.35	1.14	-0.24	30.98	0.72	5.10
a_{20}	0.89	32.67	-1.34	1.50	-1.23	1.76	-0.52	1.12	0.35	1.50	1.21	13.80
b_{20}	0.12	27.58	0.58	0.74	1.30	4.31	1.49	0.96	-0.28	9.20	0.03	1.22
a_{30}	-0.94	0.81	-1.18	0.09	0.49	9.74	-0.41	0.58	-0.47	0.84	-0.80	0.71
b_{30}	-1.82	1.83	0.87	1.42	0.11	23.47	0.49	0.75	1.13	3.78	0.14	0.47
a_{40}	-1.68	1.91	-1.52	0.66	-0.29	1.11	1.05	0.67	-0.56	1.75	-1.96	1.77
b_{40}	-0.25	1.01	1.26	0.84	-0.09	0.07	-0.28	1.62	0.85	3.08	-0.86	0.76
a_{50}	-0.28	1.12	0.50	0.80	-0.11	1.21	-1.54	1.30	-0.01	1.70	1.31	0.90
b_{50}	-0.33	0.94	-0.22	0.89	0.12	1.07	-0.81	0.44	0.79	1.91	0.47	0.82
a_{60}	1.13	1.74	2.29	0.87	0.99	3.04	-1.12	0.87	0.52	2.32	-1.12	0.51
ϕ_1	-0.71	1.91	0.12	0.37	-0.65	-0.11	-0.48	0.58	-0.30	3.07	1.49	0.35
ϕ_2	1.76	0.97	-0.09	0.19	-0.22	2.08	0.09	1.24	1.49	1.35	0.69	0.96
ϕ_3	-0.84	0.31	0.39	2.45	0.54	2.95	0.05	1.08	-2.50	1.12	-1.84	1.24
ϕ_4	-0.74	0.84	-0.58	2.17	0.90	1.57	-1.44	1.14	0.12	1.11	1.11	1.58
ϕ_5	0.42	0.77	-0.29	1.09	-1.58	1.41	1.44	1.41	0.38	1.39	-0.28	0.93
ϕ_6	-0.03	0.56	-0.52	0.56	1.60	1.76	-1.05	1.36	-0.10	1.12	0.52	1.55
ϕ_E	-1.05	1.09	0.00	0.65	-0.88	2.21	-0.98	1.37	0.02	2.51	1.58	2.07
μ_0	-0.23	122.28	-1.11	35.14	1.46	115.67	-0.81	16.92	-0.83	291.19	-0.38	24.58
	Ljung-Box		Ljung-Box		Ljung-Box		Ljung-Box		Ljung-Box		Ljung-Box	
	0.1864		0.0550		0.1935		0.3090		0.0832		0.0582	

Appendix A: Markov Chain Monte Carlo Estimation

This appendix provides details of the Gibbs sampling scheme outlined at the end of section 3. Recall that the linear mixed model was defined 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

$$\pi(\Upsilon|\alpha, y) \propto \pi(\Upsilon)\pi(y|\Upsilon, \alpha)$$

which is available in closed form (see below). All the 2^{10} combinations of indicators are sampled jointly in a multimove way as in FS-W, conditional on the latent process α.

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

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

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

$$\begin{array}{c}\sigma_\epsilon^2|\Upsilon, \psi_\Upsilon, \alpha, y \sim \mathrm{IG}(c_T, C_T)\\ \psi_\Upsilon|\Upsilon, \sigma_\epsilon^2, \alpha, y \sim \mathrm{N}(m, \sigma_\epsilon^2 S)\end{array}$$

where

$$\begin{array}{ll} S \ = \left(Z_\Upsilon' Z_\Upsilon + D_\Upsilon^{-1}\right)^{-1}, & m \ = S Z_\Upsilon' y,\\ c_T = c_0 + T/2, & C_T = C_0 + \frac{1}{2}\left(y'y - m'S^{-1}m\right).\end{array}$$

Finally,

$$\pi(y|\Upsilon, \alpha) \propto \frac{|S|^{0.5}}{|D_\Upsilon|^{0.5}} \frac{\Gamma(c_T)}{\Gamma(c_0)} \frac{C_0^{c_0}}{C_T^{c_T}},$$

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).

The draws of the parameters β_μ, β_A, $\beta_{sj}, j = 1, \dots, 6$, β_{TD} are obtained by performing a final random sign permutation. This is achieved by drawing independently Bernoulli random variables B_μ, B_A, $\mathsf{B}_{sj}, j = 1, \dots, 6, \mathsf{B}_{TD}$ 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.

The starting values are obtained by iterating for the full model (with all the indicators being equal to 1) the above GS scheme 1000 times, with initial values of the hyperparameters in ψ_Υ set equal to zero.

Under a fractional prior the marginal likelihood is evaluated as follows:

$$\begin{array}{l} c_T = c_0 + \frac{1-b}{2}T, \quad C_T = C_0 + \frac{(1-b)}{2}\left(y'y - m'\left(Z_\Upsilon' Z_\Upsilon\right)^{-1} m\right),\\ \pi \quad (y \mid \Upsilon, \alpha) \propto \frac{b^{q/2}}{2\pi^{T(1-b)/2}} \frac{\Gamma(c_T)}{\Gamma(c_0)} \frac{C_0^{c_0}}{C_T^{c_T}},\end{array} \tag{13}$$

where q is the dimension of ψ_Υ, and m is as given above. The free parameter in the fractional prior specification is b that in our study is set to be 10^{-4} and 10^{-5}.

A key assumption is that σ_ϵ^2 is strictly greater than zero, i.e. the irregular component is always present.

References

Bell, W. R. and Hillmer, S. C. (1983). Modeling time series with calendar variation. *Journal of Business and Economic Statistics*, 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. C. (2003). Seasonality tests. *Journal of Business and Economic Statistics*, 21:420–436.

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

Casella, G., Girón, F. J., Martínez, M. L., and Moreno, E. (2009). Consistency of bayesian procedures for variable selection. *The Annals of Statistics*, 37:1207–1228.

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

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

De Jong, D. N. and Whiteman, C. H. (1991). The case for trend-stationarity is stronger than we thought. *Journal of Applied Econometrics*, 6:413–421.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74:427–431.

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

Durbin, J. and Koopman, S. J. (2012). *Time Series Analysis by State Space Methods. Second Edition.* Oxford University Press, Oxford, UK.

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

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

Geweke, J. (1992). *Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments.* Bayesian Statistics 4. In J. J. Heckman and E. Leamer In J. M. Bernado and J. O. Berger and A. P. Dawid and A. F. M. Smith (eds.). University Press, Oxford. Pages 169–193.

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.

Gower, J. C. and Hand, D. J. (1996). *Biplots. Monographs on Statistics and Applied Probability.* Chapman & Hall, London, UK.

Grassi, S. and Proietti, T. (2014). Characterising economic trends by bayesian stochastic model specification search. *Computational Statistics & Data Analysis*, 71:359–374.

Greenacre, M. (2010). *Biplots in Practice.* Fundación BBVA.

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

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

Hylleberg, S., Engle, R. F., Granger, C. W. J., and Yoo, B. S. (1990). Seasonal integration and cointegration. *Journal of Econometrics*, 44:215–238.

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

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies*, 65:361–393.

Koop, G. (1992). Objective bayesian unit root tests. *Journal of Applied Econometrics*, 7:65–82.

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

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54:159–178.

La Grange, A., Roux, N., and Gardner-Lubbe, S. (2009). Biplotgui: Interactive biplots in r. *Journal of Statistical Software*, 30:1–37.

Ley, E. and Steel, M. F. J. (2009). On the effect of prior assumptions in bayesian model averaging with applications to growth reression. *Journal of Applied Econometrics*, 4:651–674.

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

Nyblom, J. and Makelainen, T. (1983). Comparisons of tests for the presence of random walk coefficients in a simple linear model. *Journal of the American Statistical Association*, 78:865–864.

O'Hagan, A. (1995). Fractional bayesian factors for model comparison. *Journal of the Royal Statistical Society, Series B.*, 57:99–118.

Peña, D., Tiao, G. C., and Tsay, R. S. (2001). *A Course in Time Series Analysis*. New York: J. Wiley and Sons.
Phillips, P. C. B. (1991). To criticize the critics: An objective bayesian analysis of stochastic trends. *Journal of Applied Econometrics*, 6:333–364.
Phillips, P. C. B. and Perron, P. (1988). Posterior odds testing for a unit root with data-based model selection. *Biometrika*, 75:335–346.
Phillips, P. C. B. and Perron, P. (1994). Testing for a unit root in time series regression. *Econometric Theory*, 10:774–808.
Proietti, T. and Grassi, S. (2012). *Bayesian stochastic model specification search for seasonal and calendar effects.* Economic Time Series: Modeling and Seasonality. In William R. Bell, Scott H. Holan and Tucker S. McElroy (eds.). CRC Press, Pages 431–455.
Schotman, P. C. and van Dijk, H. K. (1991). On bayesian routes to unit roots. *Journal of Applied Econometrics*, 6:387–401.
Sims, C. A. (1988). Bayesian skepticism on unit root econometrics. *Journal of Economic Dynamic and Control*, 12:463–474.
Sims, C. A. and Uhlig, H. (1991). Understanding unit rooters: A helicopter tour. *Econometrica*, 59:1591–1599.
Strickland, C. M., Martin, G. M., and Forbes, C. S. (2007). Parameterisation and efficient MCMC estimation of non-Gaussian state space models. *Computational Statistical and Data Analysis*, 97:2911–2930.
Wecker, W. E. and Ansley, C. F. (1983). The signal extraction approach to nonlinear regression and spline smoothing. *Journal of the American Statistical Association*, 78:81–89.
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.