International Journal of Forecasting 32 (2016) 180–202

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Outlier detection in structural time series models: The indicator saturation approach

Martyna Marczak [a,*], Tommaso Proietti [b,c]

[a] University of Hohenheim, Germany
[b] Università di Roma "Tor Vergata", Italy
[c] CREATES, Denmark

ARTICLE INFO

Keywords:
Indicator saturation
Seasonal adjustment
Structural time series model
Outliers
Structural change
General-to-specific approach
State space model

ABSTRACT

Structural change affects the estimation of economic signals, such as the growth rate or the seasonally adjusted series. One important issue that has attracted a great deal of attention in the seasonal adjustment literature is its detection by an expert procedure. The general-to-specific approach to the detection of structural change, which is currently implemented in Autometrics via indicator saturation, has proven to be both practical and effective in the context of stationary dynamic regression models and unit-root autoregressions. By focusing on impulse- and step-indicator saturation, we use Monte Carlo simulations to investigate the performance of this approach for detecting additive outliers and level shifts in the analysis of nonstationary seasonal time series. The reference model is the basic structural model, featuring a local linear trend, possibly integrated of order two, stochastic seasonality and a stationary component. Further, we apply both kinds of indicator saturation to the detection of additive outliers and level shifts in the industrial production series of five European countries.

1. Introduction

Structural change affects the estimation of economic signals, such as the growth rate or the seasonally adjusted series. An important issue is its detection by an expert procedure. Automatic outlier detection is already implemented in official seasonal adjustment procedures, like TRAMO-SEATS (Gómez & Maravall, 1996) and X-12 ARIMA (and its enhanced version X-13 ARIMA-SEATS). Both procedures consist of two main stages. First, the observed time series is modeled by means of a seasonal ARIMA (SARIMA) model with possible regression effects, which may include outlier effects. Second, based on the model identified, the series is decomposed into different components, e.g., trend or seasonal components, according to the so-called canonical decomposition (TRAMO-SEATS), or by using a cascade filter (X-12 ARIMA). Outlier detection is carried out in the first stage, and follows a specific-to-general approach based on sequential addition (potential outliers are identified one at a time), followed by backward deletion.

In this paper, we take a new look at the detection of structural change in seasonal economic time series. In particular, we consider the structural time series approach proposed by Harvey (1989) and West and Harrison (1997), according to which a parametric model for the series is formulated in terms of unobserved components directly. The reference model for adjustment purposes is the basic structural model (BSM) that was proposed by Harvey and Todd (1983) for univariate time series, and extended to the multivariate case by Harvey (1989). The BSM postulates an additive decomposition of the series into trend,

* Correspondence to: University of Hohenheim, Department of Economics, Schloss, Museumsfluegel, D-70593 Stuttgart, Germany.
E-mail address: marczak@uni-hohenheim.de (M. Marczak).

http://dx.doi.org/10.1016/j.ijforecast.2015.04.005

seasonal and irregular components. Although this model is relatively simple, it is flexible and provides a satisfactory fit for a wide range of seasonal time series. The model can be represented in state space form, which enables efficient algorithms, such as the Kalman filter and smoother, to be used for likelihood evaluation, prediction, and the estimation of the unobserved components. We refer to Durbin and Koopman (2012) for a comprehensive and up-to-date treatment of state space methods.

Seasonal adjustment using structural time series models is well established, and can be performed by the specialized software STAMP 8 (Koopman, Harvey, Doornik, & Shephard, 2009). However, in contrast to the officially used software packages for seasonal adjustment, STAMP 8 offers only a basic facility for the automatic treatment of outliers. This explains the necessity for the investigation of different approaches to outlier detection in this particular framework.

Here, we follow the indicator saturation (IS) approach, which is a new but very promising strand of research on outlier detection. It was proposed by Hendry (1999), and constitutes a general-to-specific approach. In his seminal work, Hendry (1999) introduced the impulse-indicator saturation (IIS) as a test for an unknown number of breaks, occurring at unknown times, with unknown durations and magnitudes. The procedure relies on the addition of a pulse dummy as an intervention at every observation in the sample. Significant dummies at individual points in time indicate additive outliers. The properties of this method have been studied by Castle, Doornik, and Hendry (2012), Hendry, Johansen, and Santos (2008) and Johansen and Nielsen (2009). Economic applications of IIS have been provided by, e.g., Ericsson and Reisman (2012), Hendry and Mizon (2011), and Hendry and Pretis (2013).

Recently, various other types of indicator saturation have been discussed in the literature as well, related to different types of intervention functions, representing level shifts, slope changes, etc. Considering different indicator functions should assist in the search for the most appropriate types of structural change. For example, Doornik, Hendry, and Pretis (2013) consider the so-called step-indicator saturation (SIS), which models level shifts based on step interventions. From a computational point of view, IIS and its extensions pose the problem of having more regressors than observations, which can be solved by dividing all dummies into blocks and selecting over blocks; see, e.g., Hendry and Krolzig (2004). A more elaborate search algorithm, which also accounts for collinearity between indicators, is provided by Autometrics (Doornik, 2009c), which is an integral part of PcGive (Doornik & Hendry, 2013). A recent alternative to the indicator saturation performed in PcGive is offered by the package "gets" in the R programming language (Pretis, Reade, & Surracat, 2015). Even though indicator saturation has proven to be both practical and effective in the context of the stationary dynamic regression model, its performance in the structural time series models framework has not been examined to date.

This paper's contribution to the literature lies in the fact that it is the first to combine seasonal adjustment using BSM with the general-to-specific approach to outlier detection. The method presented here differs substantially from the procedures in TRAMO-SEATS and X-12 ARIMA, in both the modeling and the outlier detection strategy. In the first step, we assess the performance of indicator saturation via Monte Carlo simulations. Next, we provide an empirical application of the method considered to raw industrial production series in France, Germany, Italy, Spain and the UK over the time period 1991.M1–2014.M1. In our analysis, we apply impulse-indicator saturation (IIS) and step-indicator saturation (SIS). The reason for this specific choice is twofold. Pulse and step dummies are both the simplest and the most flexible ways of modeling structural changes. Moreover, our greatest interest in the empirical exercise lies in the question of whether the procedure is capable of identifying a potential level shift corresponding to the economic and financial crisis that began in Europe at around the end of 2008.

The remainder of the article is organized as follows. Section 2 describes the framework for modeling seasonal time series with outlying observations and location shifts. In particular, we set out the basic structural model with calendar effects in Section 2.1, then present the concept of indicator saturation and explain how it is integrated into the current framework in Section 2.2. Section 3 summarizes our findings on the performances of IIS and SIS, obtained by Monte Carlo simulations. First, we discuss the findings on the detection power of IIS and SIS in relation to differing settings for the data generating process and outlier detection. Then, we examine two aspects that relate to the situation without any outliers: the null rejection frequency, and the impacts of IIS and SIS on the estimated model parameters. In Section 4, IIS and SIS are applied to real data for detecting outliers and level shifts. Section 5 concludes.

2. Modeling framework

2.1. The basic structural time series model

The BSM postulates an additive and orthogonal decomposition of a time series into unobserved components that represent the trend, the seasonality and the irregular component. If y_t denotes a time series observed at $t = 1, 2, \ldots, T$, the decomposition can be written as follows:

$$y_t = \mu_t + \gamma_t + \sum_{k=1}^{K} \delta_{xk} x_{kt} + \epsilon_t, \quad t = 1, \ldots, T, \tag{1}$$

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

The trend component has a local linear representation:

$$\begin{aligned} \mu_{t+1} &= \mu_t + \beta_t + \eta_t \\ \beta_{t+1} &= \beta_t + \zeta_t, \end{aligned} \tag{2}$$

where η_t and ζ_t are mutually and serially uncorrelated normally distributed random shocks, with zero means and variances σ_η^2 and σ_ζ^2, respectively.

The seasonal component can be modeled as a combination of six stochastic cycles with the common variance σ_ω^2. The individual stochastic cycles have trigonometric