Computational Statistics and Data Analysis 58 (2013) 4–14

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

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

Removing seasonality under a changing regime: Filtering new car sales

Michael A. Thornton *

School of Economics, University of Reading, Whiteknights, PO Box 217, RG6 6AH, United Kingdom

ARTICLE INFO

Article history:
Available online 24 June 2011

Keywords:
Seasonal adjustment
Signal extraction
Frequency response
Butterworth lowpass filter

ABSTRACT

The use of filters for the seasonal adjustment of data generated by the UK new car market is considered. UK new car registrations display very strong seasonality brought about by the system of identifiers in the UK registration plate, which has mutated in response to an increase in the frequency with which the identifier changes, while it also displays low frequency volatility that reflects UK macroeconomic conditions. Given the periodogram of the data, it is argued that an effective seasonal adjustment can be performed using a Butterworth lowpass filter. The results of this are compared with those based on adjustment using X-12 ARIMA and model-based methods.

1. Introduction

The registration system for new cars in the UK has long involved a component signifying the period in which the vehicle was first registered. This feature, which means that cars bought and registered at certain points of the year appear to be newer for longer, has made the pattern of new car purchase highly seasonal. To illustrate, August is usually a relatively quiet month in all industries in the northern hemisphere but, because an annual identifier incorporated into registrations changed on August 1, it was the month with the highest number of new car registrations every year in the UK for 28 years.

Such strong seasonality is, in itself, a challenge for modellers and analysts. A range of techniques have been developed to remove seasonality from data, including empirical procedures such as X-12 ARIMA, see Findley et al. (1998), and model-based procedures such as the TRAMO-SEATS procedure of Gomez and Maravall (1996), applied for example in Maravall (2006), or the structural time series approach, see for example Harvey and Todd (1983) or Hindrayanto et al. (2010). In this instance, the task is complicated by a seasonal break.

Over time, it became clear that this August effect was dominating the year. In response, the registration system was modified in 1999 so that the identifier changed every six months, in September and March. It was hoped that this would smooth sales to something more like the patterns seen in continental Europe. In reality the seasonality mutated, over an identifiable interval and in an identifiable way, but did not disappear: the standard deviation in monthly sales over a calender year has fallen a little as a proportion of average monthly sales but remains at around 60%. That mutation transformed both the amplitude and frequency of the seasonal cycles while also inducing a modest phase shift: after a period of annual identifiers, there is one seven month identifier before the new regime of six-monthly identifiers. This is not the type of problem for which traditional seasonal adjustment procedures were designed.

As noted in Ghysels and Osborn (2001) there is a long tradition from model-based seasonal adjustment procedures of treating the removal of seasonality as a signal extraction problem. Given the nature of the mutation in seasonal pattern and the relatively low power displayed by the series at non-trend, non-seasonal frequencies, we also make use of a filter that

* Tel.: +44 0 118 378 5344; fax: +44 0 118 378 4029.
E-mail address: m.a.thornton@reading.ac.uk.
URL: http://www.personal.reading.ac.uk/~qb901444/.

0167-9473/$ – see front matter
doi:10.1016/j.csda.2011.06.021

can be defined in terms of its response to frequencies contained within certain bands. This is the Butterworth lowpass filter as discussed by Pollock (2000), Gomez (2001), Harvey and Trimbur (2003), Pollock (2006) and Proietti (2007). In econometrics it has largely been used for identifying business cycle activity, but has been used by Pollock (2000) for removing seasonal components in Swiss unemployment data. The Butterworth filter has two interpretations. The first, which is fully parametric, as the optimal signal extraction filter for an underlying unobserved components model. The second as a semi-parametric technique to split a series into two parts based on two non-overlapping frequency ranges. With a semi-parametric interpretation, it is set up to cope well with mutations in amplitude and phase, enabling the identification of trend-cycle elements across the full span of the series. Other techniques, which are designed around a fixed or slowly evolving seasonal pattern, would be forced into treating the data generated by each of the regimes somewhat differently.

The underlying trends in UK new car data are of interest beyond presenting a methodological challenge. The motor industry is widely seen both as a leading indicator, and as a wider linchpin, of the manufacturing sector, particularly during the recent economic downturn. Following the lead in continental Europe, from May 2009 to March 2010, the UK ran a car scrappage incentive scheme, whereby purchasers of new cars received £2000 toward the cost of a new car if they scrap a vehicle that is more than 10 years old. In total 400,000 purchases benefited from this additional funding. Such schemes and their potential to create further cycles in activity have been analysed from a theoretical viewpoint by, among others, Adda and Cooper (2000).

The paper is split into 5 sections. Section 2 discusses the nature of seasonality and popular methods of seasonal adjustment, making the case for the use of a frequency-specific filter as a part of that tool-kit. Section 3 presents the Butterworth lowpass filter. Section 4 explains the changes that took place in the series and applies the filter to the full sample and to two sub-samples representing the different regimes, comparing the results with those using X-12 ARIMA and TRAMO-SEATS, and Section 5 concludes.

2. Seasonal adjustment

Here we follow the definition in Nerlove (1964, pp. 264):

> In the more general case, then, we may define seasonality as that characteristic which gives rise to spectral peaks at seasonal frequencies.

Seasonal factors tend to fall into three broad categories, as noted by Hylleberg (1992): climate; convention (including the timing of religious festivals); and repetitive practices (such as tax years, accountancy periods and the UK registration plate system). Some of these will remain fixed, some may vary but are always predictable, and some may offer departures around a roughly regular pattern, causing spectral power to concentrate around rather than appearing directly on the seasonal frequencies.

2.1. Seasonal adjustment procedures

Procedures for removing seasonal factors from data fall into two broad camps: empirical and model-based techniques. Empirical techniques use statistical smoothing methods without presupposing that the data are generated by an underlying model. The procedure developed at the US Census Bureau, Variant X-11, see Shishkin et al. (1967), is still the major part of adjustment procedures used in most statistical agencies. An excellent treatment of the way the program is assembled can be found in Hylleberg (1992), Ghysels and Perron (1993) and Wallis (1982). The more recent X-12 ARIMA, see Findley et al. (1998), builds on X-11, improving diagnostics, the treatment of outliers and enabling the use of ARIMA-generated out of sample values in the smoothing. Such techniques, however, have a number of drawbacks for econometric work. Firstly, the use of moving averages with long lags and leads means that a definitive figure for the adjusted series will not be available for a number of years. Secondly, the procedure provides no insight into what seasonality might be and no framework in which we can examine its relationship with the trend component. Thirdly, these procedures smooth data in connection with the rest of the sample leaving a series of data-points which are no longer independent realisations. The number of degrees of freedom lost will depend on a number of issues including the choice of procedure for smoothing outliers and trading day effects.

Model-based procedures, on the other hand, suppose that the series to be treated can be decomposed into unobserved components, perhaps after taking logarithms. A typical form splits a series (X_t) into trend-cycle (TC_t), seasonal (S_t) and irregular (I_t) components,

$$
\begin{aligned}
X_t &= TC_t + S_t + I_t,\\
TC_t &= \frac{\gamma(L)}{\delta(L)}\epsilon_t,\\
S_t &= \frac{\psi(L)}{s(L)}\upsilon_t\\
I_t &= \frac{\beta(L)}{\alpha(L)}\nu_t,
\end{aligned}
\qquad (1)
$$