

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search http://ageconsearch.umn.edu aesearch@umn.edu

Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.

The Proportion of the Seasonal Period as a Season Index in Weekly Agricultural Data

Gloria Martín Rodríguez (gmartinr@ull.es) José Juan Cáceres Hernández (jcaceres@ull.es) University of La Laguna

Contributed Paper prepared for presentation at the International Association of Agricultural Economists Conference, Beijing, China, August 16-22, 2009

Copyright 2009 by Gloria Martín-Rodríguez and José Juan Cáceres-Hernández. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

The Proportion of the Seasonal Period as a Season Index in Weekly Agricultural Data

Gloria Martín Rodríguez (gmartinr@ull.es)

José Juan Cáceres Hernández (jcaceres@ull.es)

Department of Economía de las Instituciones, Estadística Económica y Econometría

University of La Laguna

España

Abstract

In this paper a seasonal model is proposed to deal with weekly agricultural seasonal patterns in which neither the length of the seasonal period nor the magnitude of the seasonal effects remain the same over time. To model this heterogeneous seasonal behaviour, the seasonal effect at a season is defined as a function of the proportion of the length of the seasonal period elapsed up to this season, and the seasonal pattern is modelled by means of evolving splines. The methodology is illustrated for weekly Canary tomato prices.

JEL classification: C22, Q17

Key words: weekly data, heterogeneous seasonal patterns, splines.

1. INTRODUCTION

In recent years weekly agricultural data are available over long periods of time. Data of this nature tend to show seasonal behaviours characterised by heterogeneity, as the length of the time period in which one of these periodic patterns is completed does not remain the same over time. For instance, weekly prices of an agricultural commodity are only registered when the product is on the market, and the marketing period may change from year to year depending on climate or managerial strategies.

Such data are not easy to handle using existing methods, and so there is a need for new models designed to cope with the complexities of these weekly time series. Modelling strategies should lead to parsimonious formulations as the number of seasons that define some periodic patterns is large, and they should be flexible enough to capture the heterogeneity of the seasonal variations. From this point of view, spline models (Poirier, 1976; Eubank, 1988), in which the seasonal effect at time t is defined as a function of the corresponding season, seem to be a suitable approach. In any case, the isolation of the trend signal needs to be made to make the specification of a seasonal model feasible. Harvey and Koopman (1993) and Harvey *et al.* (1997) decompose a time series into trend, seasonal and irregular components by using unobserved components time series models. To capture seasonal variations, they formulate stochastic periodic cubic splines which are required to match seasonal effects at specific seasons. Ferreira *et al.* (2000) also consider the seasonal component to be a smooth function of time and propose to specify trend and seasonal component by using two smoothing splines with different smoothing parameters.

Both of the two alternatives to model seasonal effects have been developed to provide enough flexibility to capture changing seasonal patterns. Recent works in time series literature seem to be inspired by the same idea (Orbe *et al.*, 2005; Pedregal and Young, 2006). The drawback of these procedures is that the restrictions over the coefficients introduced into the seasonal model are not derived from the changes observed in the seasonal pattern.

Unlike approaches that make the process variability more flexible to be feasible data are assumed to be generated from the selected model for the stochastic process, by means of spline functions the seasonal model can be easily adapted to changes observed in the magnitude and shape of the whole seasonal pattern over the seasonal period. This approach models explicitly changes in the mean of the stochastic process caused by changing seasonal variations. In this way, description and forecasting of this component is easier. The piecewise polynomial functional form and the periodicity and continuity restrictions which conventional splines are required to satisfy, provide such a model with a degree of regularity. On the other hand, splines make the definition of different seasonal models for different sample periods feasible and, therefore, enough flexibility is provided to capture heterogeneous seasonal patterns.

The point at issue is that all of these seasonal models may not be suitable enough when, from time to time, there are no observations corresponding to some of the seasons that belong to the seasonal period or, above all, in time series in which the length of the seasonal period does not remain the same over time. Besides a model able to capture seasonal variations, there is a need to reflect on the concept of seasonal pattern. For instance, the points in time at which the marketing period of an agricultural commodity in the international market begins and ends are likely to change from year to year. Furthermore, the product could be assumed to reach the highest quality level around the middle of the harvesting period and this circumstance implies that the best prices are obtained during this period of the year. In this hypothetical situation, the seasonal effect at a point in time in the weekly price series depends on the proportion of the whole seasonal period which corresponds to this point in time. The formulation of seasonal effects as functions of such a proportion means that the periodicity of the seasonal pattern is measured in a different scale than the usual one. Obviously, this new approach is only useful when the length of the seasonal period does not remain the same over time. This being the case, the new formulation does not bring about an increase in mathematical or statistical complexity, but provides the seasonal model with a noticeable gain in flexibility and parsimony. In this sense, a formulation is developed in the following section. In Section 3 the proposal is applied to a weekly series of tomato prices in which the number of weeks during the marketing period is not fixed. Finally, in Section 4 the concluding remarks are stated.

2. SEASONAL SPLINE

In a time series model formulated as

$$y_t = \mu_t + \gamma_t + \varepsilon_t, \ t = 1, ..., T , \tag{1}$$

where μ_t and γ_t are the trend or level component and the seasonal component, and ε_t is the irregular component, the seasonal pattern could be assumed to be completed in a period whose length does not remain the same over time. Suppose that the series is divided in *m* sub-periods of time in which the length of the seasonal period is s_c , c = 1,...,m. Let γ_t be defined as $\gamma_t = \gamma_w$ if the observation at time *t* corresponds to season j_c belonging to subperiod *c* in such a way that $w = \frac{j_c}{s_c}$. Note that *w* is the proportion of the seasonal period elapsed up to season j_c , $j_c = 1,...,s_c$, and, therefore, $0 < w \le 1$. If seasonal variation is also assumed to change in a smooth manner from one season to the following one, a periodic cubic spline is a suitable model for these variations. That is,

$$\gamma_w = g(w) + \xi_w, \qquad (2)$$

where ξ_w is a residual term and g(w) is a third degree piecewise polynomial function,

$$g(w) = g_i(w) = g_{i,0} + g_{i,1}w + g_{i,2}w^2 + g_{i,3}w^3, \ w_{i-1} \le w \le w_i, \ i = 1, \dots, k,$$
(3)

where $w_0 = 0$ and $w_k = 1$. The continuity of the spline function and its first and second derivatives are enforced by the following conditions, respectively,

$$g_{i,0} + g_{i,1}w_i + g_{i,2}w_i^2 + g_{i,3}w_i^3 = g_{i+1,0} + g_{i+1,1}w_i + g_{i+1,2}w_i^2 + g_{i+1,3}w_i^3,$$
(4a)

$$g_{i,1} + 2g_{i,2}w_i + 3g_{i,3}w_i^2 = g_{i+1,1} + 2g_{i+1,2}w_i + 3g_{i+1,3}w_i^2$$
(4b)

and

$$2g_{i,2} + 6g_{i,3}w_i = 2g_{i+1,2} + 6g_{i+1,3}w_i,$$
(4c)

for i = 1, ..., k - 1. The continuity between two consecutive complete seasonal periods, which are possibly completed in a different number of seasons, is assured by means of the three following conditions

$$g_{k,0} + g_{k,1} + g_{k,2} + g_{k,3} = g_{1,0},$$
(5a)

$$g_{k,1} + 2g_{k,2} + 3g_{k,3} = g_{1,1} \tag{5b}$$

and

$$2g_{k,2} + 6g_{k,3} = 2g_{1,2}.$$
 (5c)

Note that the spline is periodic as a function of the proportion of the seasonal period, but it may not be periodic in terms of the season.

Given that the number of unknown parameters, 4k, is larger than the number of demanded conditions, 3k, the parameters of the vector $G_{3kx1} = (g_{1,1}, g_{1,2}, g_{1,3}, ..., g_{k,1}, g_{k,2}, g_{k,3})'$ are expressed as functions of the parameter vector $G_{kx1}^* = (g_{1,0}, g_{2,0}, ..., g_{k,0})'$. In a matrix form, the continuity restrictions can be written as TG = R, where

| | w_1 | w_1^2 | w_1^3 | $-w_1$ | $-w_1^2$ | $-w_1^3$ | 0 | 0 | 0 | 0 | 0 | 0 |
|---------------|-------|---------|----------|--------|-----------|-----------|---------------|---------------|---------------|------------|--------------|----------------|
| $T_{3kx3k} =$ | ÷ | ÷ | ÷ | ÷ | ÷ | ÷ | ÷ | ÷ | ÷ | ÷ | ÷ | : |
| | 0 | 0 | 0 | 0 | 0 | 0 | w_{k-1} | w_{k-1}^{2} | w_{k-1}^{3} | $-w_{k-1}$ | $-w_{k-1}^2$ | $-j_{k-1}^{3}$ |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |
| | 1 | $2w_1$ | $3w_1^2$ | -1 | $-2w_{1}$ | $-3w_1^2$ | 0 | 0 | 0 | 0 | 0 | 0 |
| | ÷ | ÷ | ÷ | ÷ | ÷ | ÷ | ÷ | ÷ | ÷ | ÷ | ÷ | : |
| | 0 | 0 | 0 | 0 | 0 | 0 | 1 | $2w_{k-1}$ | $3w_{k-1}^2$ | -1 | $-2w_{k-1}$ | $-3w_{k-1}^2$ |
| | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 3 |
| | 0 | 2 | $6w_1$ | 0 | -2 | $-6w_{1}$ | 0 | 0 | 0 | 0 | 0 | 0 |
| | ÷ | ÷ | ÷ | : | ÷ | ÷ | : | ÷ | ÷ | : | : | : |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | $6w_{k-1}$ | 0 | -2 | $-6w_{k-1}$ |
| | 0 | -2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 6 |

and

$$R_{3kx1} = \begin{bmatrix} -g_{1,0} + g_{2,0} \\ \vdots \\ -g_{k-1,0} + g_{k,0} \\ -g_{k,0} + g_{1,0} \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

Therefore, $G = T^{-1}R$. If the inverse of the matrix T is expressed by

$$T^{-1}{}_{3kx3k} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,3k} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,3k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{3k,1} & a_{3k,2} & \cdots & a_{3k,3k} \end{bmatrix}$$

and g_r , r = 1, ..., 3k, are defined in such a way that $g_1 = g_{1,1}, g_2 = g_{1,2}, g_3 = g_{1,3},$ $g_4 = g_{2,1}, g_5 = g_{2,2}, g_6 = g_{2,3}, ..., g_{3k-2} = g_{k,1}, g_{3k-1} = g_{k,2}$ and $g_{3k} = g_{k,3}$, then

$$g_{r} = \alpha_{r,1}g_{1,0} + \alpha_{r,2}g_{2,0} + \dots + \alpha_{r,k-1}g_{k-1,0} + \alpha_{r,k}g_{k,0}, \qquad (6)$$

where $\alpha_{r,1} = -a_{r,1} + a_{r,k}$, $\alpha_{r,i} = a_{r,i-1} - a_{r,i}$, i = 2, ..., k, for r = 1, ..., 3k. Now, the spline can be formulated as a function of the parameters $g_{i,0}$, i = 1, ..., k, as follows,

$$g(w) = g_{1,0}X_{1,w} + g_{2,0}X_{2,w} + \dots + g_{k-1,0}X_{k-1,w} + g_{k,0}X_{k,w},$$
(7)

where

$$X_{i,w} = D_{i,w} + (\alpha_{1,i}w + \alpha_{2,i}w^2 + \alpha_{3,i}w^3)D_{1,w} + (\alpha_{4,i}w + \alpha_{5,i}w^2 + \alpha_{6,i}w^3)D_{2,w} + \dots + (\alpha_{3k-2,i}w + \alpha_{3k-1,i}w^2 + \alpha_{3k,i}w^3)D_{k,w}$$
(8)

for
$$i = 1, ..., k$$
, $D_{i,w} = \begin{cases} 1, & \text{if } w_{i-1} \le w < w_i \\ 0, & \text{in other case} \end{cases}$, $i = 1, ..., k-1$, and $D_{k,w} = \begin{cases} 1, & \text{if } w_{k-1} \le w \le 1 \\ 0, & \text{in other case} \end{cases}$

The critical point is the selection of the number and position of knots. A previous approximation of seasonal variations shows the main changes in the shape of the seasonal pattern and, consequently, the number of break points can be decided. In this sense, the deviations $\hat{\gamma}_t^{l}$ from the original series to a moving average series with period s_c can be used. Then, the decision has been adopted to select the combination of locations that minimises the residual sum of squares when the following regression model

$$\hat{\gamma}_{t}^{1} = g_{1,0}X_{1,t} + g_{2,0}X_{2,t} + \dots + g_{k-1,0}X_{k-1,t} + g_{k,0}X_{k,t} + \xi_{t}, \qquad (9)$$

is estimated, where $X_{i,t} = X_{i,w}$, i = 1,...,k, if the observation at time *t* corresponds to season j_c in such a way that the proportion of the seasonal period elapsed up to this season is *w*.

The previous specification is flexible enough to capture a seasonal pattern in which γ_w evolves over time. The seasonal pattern in the *m* sub-periods in which the series is divided, can be jointly modelled as $\gamma_t = g(t) + \xi_t$, where g(t) is the changing spline

$$g(t) = \sum_{c=1}^{m} \left[g_{1,0}^{c} X_{1,t} + \dots + g_{k,0}^{c} X_{k,t} \right] D_{c,t}^{sp}, \qquad (10)$$

where
$$D_{c,t}^{sp} = \begin{cases} 1, & t \in sub - period \ c \\ 0, & in \ other \ case \end{cases}$$
, $c = 1, \dots, m$, and $X_{i,t} = X_{i,w}$, $i = 1, \dots, k$, if the

observation at time t corresponds to season j_c in sub-period c in such a way that $w = \frac{J_c}{s_c}$.

If $g_{i,0}^c = g_{i,0}$, i = 1,...,k, c = 1,...,m, occurs, the previous model provides a valuable gain in parsimony with regards to traditional models in which the seasonal variation is defined as a function of the season. However, these parameters could also be assumed to evolve over time, and the locations of break points are not necessarily the same for each one of the sub-periods.

When the shape of the seasonal pattern remains the same, the higher the frequency in which changes in the length of the seasonal period are observed, the higher the gain in parsimony. Changes in the specific seasons in which the seasonal variation is completed, even though the number of seasons does not change, can be another source of parsimony. Given that trend and seasonal components are unobservable, the specification of one of them has consequences in the specification of the other one. Therefore, the procedure to isolate an approximation of the seasonal component should be able to capture its most relevant features in order to formulate the model. In particular, the chosen approach should show the changes in the shape of the seasonal pattern. In this sense, suitable procedures are based on calculating moving averages with a changing period defined in terms of the number of seasons in which the seasonal variation is completed.

3. AN APPLICATION TO A WEEKLY SERIES OF TOMATO PRICES

The usefulness of the methodological proposal is illustrated using a weekly series for Canary tomato prices in German markets from 1987/1988 to 2006/2007 harvests. These prices may be taken as an approximation to mainland European tomato prices. Price statistics have been obtained from the daily series at the wholesale markets in Germany (Berlin, Colonia, Frankfurt, Hamburg and Munich) provided by the International Trade General Secretary of the Spanish Ministry for Industry, Tourism and Trade (see the COMEX webpage http://www.mcx.es/IPRECOM/)¹. The weekly prices for a box of six kg have been calculated as an average of the modal daily prices corresponding to the fruit diameters of 47-57mm and 57-67 mm.

Actually, there is no Canary tomato export for some weeks, especially during the summer period, and therefore, neither there are prices for Canary fruit at these dates. This pattern is a rational response guided by the search for profitability; there are no exports in summer because Northern European countries and Canary supplies converge in this season and due to this Canary tomato prices would be low. However, the extent of the harvest has

¹ Missing observations at some weeks have been completed from the weekly data published by ZMP (Zentrale Markt und Preisberichtstelle) or by the provincial exporter association of *Santa Cruz de Tenerife* (*ACETO*) in its export harvest reports.

evolved over time. During the 1980s, the harvest often finished in early May, when the overlap of different supplies caused lower prices, and eventually forced Canary producers to pay compensatory taxes. The new trade situation of the Canary Islands with regard to the EU since July 1991, when reference prices were substituted by supply prices, and the full integration into the CAP since January 1st 1993, that meant the abolition of reference/supply prices, brought about the export period continue until June from the campaign 1991/1992.

Furthermore, the beginning of some harvests also moved forward by some weeks with regard to the usual one, which is located around the week 40. These anomalous observations have been deleted. The removed observations correspond to 1995 (week 27), 1996 (week 27), 1999 (weeks 27 to 29), and 2001 (weeks 36 to 38). However, the missing observations corresponding to 2003 (weeks 1 and 42), 2004 (week 1) and 2006 (week 52) were not deleted because they are located in the harvesting period. This decision involves the relevant seasonal variation is assumed to be defined during the export period. In fact, missing prices in summer period do not provide relevant economic information beyond the one which is derived from the length of the non-export period, which is provided implicitly by the modelled seasonal variation. So, a new series is obtained and it will be referred to as $\{y_t\}_{t=1,...,717}$ (Figure 1). As shown in Table 1, the length of the export period oscillates between 29 and 39 weeks. Therefore, the methodology proposed in the previous section is one option that seems very well suited for the purpose of modelling the seasonal pattern.

Some stylized facts about such a pattern are clear from the data observed. The lowest prices are located at the beginning of the harvesting period until the campaign 1990/1991. However, minimum prices are usually observed at the end of the harvesting period from the 1991/1992 harvest. As mentioned, Canary exporters were bound to observe a reference prices system until the beginning of the 1990s. Such a system prevented prices laying down reference levels from the end of April. Once the reference/supply prices were abolished, the increasing overlap of different supplies caused lower prices at the end of the harvest. On the other hand, the highest prices are usually located between the end of December and the beginning of February or between March and April.

As regards to the long term movement, an approximation can be extracted by calculating moving averages with period s_c , corresponding to observations that belong to a seasonal pattern whose period is s_c , $\{ma_t\}_{t=1,\dots,717}^2$. From the data registered and the calculated moving averages in Figure 1, several periods differing by the long-term movement can be distinguished. The average prices by campaign in Table 1 show that the stability of export volumes lets exporters to obtain prices around 8 euros per box until the beginning of the 1990s. Surely as a response to the new trade regulation, the significant export boost until 1996, when a trade agreement between the EU and Morocco was implemented (Reg (CE) 747/2001; Reg (CE) 37/2004; Reg (CE) 503/2005), brought about the average prices lie down until a level around 6 euros/box. The minimum average price is registered in 1996/1997 campaign. During the following harvests, the decrease in export levels explains the moderate increase in average prices, in spite of some quality problems. Prices near 8 euros/box are registered in 2001/2002 and 2004/2005 campaigns, but, in general, the average prices registered in the last harvests in the sample period oscillate between 6 and 7 euros/box. Unobserved components time series models (Harvey, 1989) provide a suitable statistical tool to capture these instabilities in the trend signal.

To estimate the structural model, seasonal variations need to be formulated by means of spline functions. To this end, the first step consists of estimating an approximation of seasonal effects. In this sense, the deviations from the original series to the moving averages series are previous estimates of the seasonal variations, $\{\hat{\gamma}_t^1\}_{t=1,...,717}^{-3}$. The number and locations of break points are selected from the series $\{\hat{\gamma}_t^1\}_{t=1,...,717}$. To obtain a more parsimonious formulation, the break points are assumed to be the same for all harvests. These points are also assumed to belong to the set $\{\frac{l}{1000}\}_{l=1,...,1000}$ that divide the continuous interval (0,1) into short enough subintervals. From the results of estimating the regression

 $^{^{2}}$ When missing observations are located into the period which is used to calculate the moving average, such observations are deleted, and the period is extended in such a way that the number of observations which is used to calculate the moving average is the corresponding to the length of the harvesting period.

³ Moving averages corresponding to the first and the last sixteen observations have been obtained by using observations corresponding to 1986/1987 and 2007/2008 campaigns.

model in Equation (9), a six-segment spline has been specified, and the selected break points were located at the proportions of the seasonal period $w_1 = 0.36$, $w_2 = 0.58$, $w_3 = 0.61$, $w_4 = 0.65$ and $w_5 = 0.99$. The number of break points is selected in order to the spline captures the main changes in the shape of the observed seasonal pattern, whereas the set of locations of these points is the one that provides the best adjustment.

According to the results of estimating the parametric model

$$\hat{\gamma}_{t}^{1} = \sum_{c=1}^{20} \left[g_{1,0}^{c} X_{1,t} + \dots + g_{6,0}^{c} X_{6,t} \right] D_{c,t}^{sp} + \xi_{t} , \qquad (11)$$

a new approximation of seasonal variations is obtained as

$$\sum_{c=1}^{20} \left[\hat{g}_{1,0}^{c} X_{1,t} + \ldots + \hat{g}_{6,0}^{c} X_{6,t} \right] D_{c,t}^{sp} .$$
(12)

Once these estimates are corrected in such a way that the area under the spline function is equal to zero over each harvest, that is to say,

$$\sum_{i=1}^{6} \int_{w_{i-1}}^{w_i} \left(\hat{g}_{i,0}^c + \hat{g}_{i,1}^c w + \hat{g}_{i,2}^c w^2 + \hat{g}_{i,3}^c w^3 \right) dw = 0, \qquad (13)$$

new estimates $\{\hat{\gamma}_{t}^{2}\}_{t=1,\dots,717}$ are obtained (Figure 2). Parameter estimates $\hat{g}_{i,1}^{c}, \hat{g}_{i,2}^{c}, \hat{g}_{i,3}^{c}$ are obtained according to Equation (6). Note that the parametric spline formulation provides estimates of the seasonal variation at missing values.

The previous estimates of the seasonal variations do not show a common pattern over the sample. The magnitudes of the seasonal effects at the beginning of the harvesting period are similar to the ones at the end of such a period until the harvest 1996/1997. From the 1997/1998 harvest, the value of the seasonal effect at the beginning is higher than the one observed at the end. The differences are outstanding in harvests 1999/2000 and 2001/2002, whereas they are not significant during the harvests 2000/2001, 2002/2003, 2004/2005 and 2005/2006. In more general terms, noticeable seasonal effects are observed from the harvests 1987/1988 to 1990/1991, also in harvest 1999/2000, and above all in harvest 2001/2002; whereas most of harvests are characterised by more moderate effects (from 1991/1992 to 1998/1999, 2000/2001, and from 2002/2003 to 2006/2007).

On the other hand, some harvests show irregular seasonal variations (1991/1992, 1992/1993, 2002/2003). However, a period with higher and positive seasonal effects is usually observed (harvests 1993/1994 to 1998/1999, 2001/2002, 2003/2004, 2005/2006). Two or even three of these periods are also frequent (see harvests 1988/1989, 1989/1990, 1990/1991, 2000/2001, 2006/2007, and harvests 1987/1988, 1999/2000, 2004/2005, respectively). Obviously, these seasonal effects correspond to the highest prices, and, as mentioned, such prices are located around the end of the first third of the campaign (between the end of December and the beginning of February) or around the last third of the campaign (during March or April).

In spite of the differences observed in the seasonal patterns, once the parametric model

$$\hat{\gamma}_{t}^{2} = \sum_{c=1}^{20} \left[g_{1,0}^{c} X_{1,t} + \dots + g_{6,0}^{c} X_{6,t} \right] D_{c,t}^{sp} + \xi_{t}$$
(14)

is estimated, F tests lead to the conclusion that the seasonal patterns in 93/94 and 94/95 are homogeneous. The same conclusion is obtained for 95/96, 96/97 and 97/98 harvests.

These hypotheses about the seasonal pattern can be introduced into a structural model formulated as

$$y_{t} = \mu_{t} + \sum_{c=1,...,6,12,...,19} \left[g_{1,0}^{c} X_{1,t} + ... + g_{6,0}^{c} X_{6,t} \right] D_{c,t}^{sp} \\ + \left[g_{1,0}^{7,8} X_{1,t} + ... + g_{6,0}^{7,8} X_{6,t} \right] \left[D_{7,t}^{sp} + D_{8,t}^{sp} \right] \\ + \left[g_{1,0}^{9,10,11} X_{1,t} + ... + g_{6,0}^{9,10,11} X_{6,t} \right] \left[D_{9,t}^{sp} + D_{10,t}^{sp} + D_{11,t}^{sp} \right] \\ + \left[g_{1,0}^{20} X_{1,t} + ... + g_{5,0}^{20} X_{5,t} \right] D_{20,t}^{sp} \\ + \lambda_{51,1995} I_{51,1995} + \lambda_{26,2000} I_{26,2000} + \varepsilon_{t}$$
(15)

where the level component is modelled as

$$\mu_t = \mu_{t-1} + \beta_t + \eta_t \,, \tag{16}$$

and the slope component is

$$\beta_t = \beta_{t-1} + \lambda_{41,1991} I_{41,1991} + \zeta_t \,. \tag{17}$$

Note that impulse intervention variables, $I_{\ell,p}$, have been included to capture anomalous observations located at week ℓ of year p. In order to avoid multicollinearity problems, the

regressor $X_{6,t}D_{20,t}^{sp}$ is dropped. The model is estimated once each missing observations in the original series is substituted by an estimate obtained as a sum of an average of the contiguous values of the moving averages, plus the estimate of $\hat{\gamma}_t^2$ in Equation (14), plus the magnitude of the seasonal correction applied to $\hat{\gamma}_t^2$ at the corresponding time point. The estimating results lead to assume that the variances of the disturbance terms in level and slope equations are null. The results of the final model's estimation are shown in Table 2, Figure 1 and Figure 2. The estimates of the seasonal variations from the structural time series model have been corrected in such a way that the area under the spline function over each campaign is equal to zero. The correction applied to the estimates of the long term component.

The estimates of the level at the second half of the sample show jumps between the end of a campaign and the beginning of the following one. This behaviour is noticeable at the first week of the harvests 2001/2002 and 2004/2005, when an increase of the average price is registered. However, given the specification of the structural model, there is no need to add an outlying intervention to the level component. On the other hand, note that the intervention variable in the slope term captures the end of the growth period in the long term movement.

4. CONCLUDING REMARKS

The methodology designed has been shown to be a suitable statistical tool to capture heterogeneous seasonal patterns in weekly agricultural data, particularly in a time series whose length of the seasonal period does not remain the same over the sample. The formulation of the seasonal effect at a season as a function of the proportion of the length of the seasonal period elapsed up to this season provides a gain in parsimony; and, above all, such a formulation makes feasible the modelling of seasonal variations in which the length of the seasonal period is not fixed. So, there is no need to remove relevant observations or to divide the sample in an artificial way. Note that, in these settings, the use of conventional formulations usually implies that the series should be considered to be a set of different series corresponding to different sub-samples, which may give rise to distortions in the estimates of the long term movement.

On the other hand, the methodology proposed is flexible enough to capture instabilities in the magnitude of seasonal effects, and the model can be easily adapted to the changes observed in the magnitude and shape of the whole seasonal pattern over the seasonal period. That is to say, the flexibility of the model does not depend on the variance of the disturbance terms that drive the dynamic of the time varying parameters in conventional stochastic formulations. From the changes observed in the shape of the seasonal pattern, the model is formulated in order to identify the changes in the mean of the stochastic process which captures the seasonal variation. Therefore, the elements of the spline specification are not always derived from an optimization procedure. Furthermore, splines adjusted to the data provide smooth estimates of seasonal effects, which make the description of the seasonal pattern easier and, therefore, help farmers who have to take the seasonal nature of agricultural economics into account.

REFERENCES

Eubank RL. Spline smoothing and nonparametric regression. M. Dekker: New York; 1988.

Ferreira E, Núñez-Antón V, Rodríguez-Póo J. Semiparametric approaches to signal extraction problem in economic time series. Computational Statistics & Data Analysis 2000; 33; 315-333.

Harvey AC. Forecasting, Structural Time Series and the Kalman Filter. Cambridge University Press: Cambridge; 1989.

Harvey AC, Koopman SJ. Forecasting hourly electricity demand using time varying splines. Journal of the American Statistical Association 1993; 88; 1228-1236.

Harvey AC, Koopman SJ, Riani M. The modelling and seasonal adjustment of weekly observations. Journal of Business and Economic Statistics 1997; 15; 354-368.

Orbe S, Ferreira E, Rodríguez-Póo J. Nonparametric estimation of time varying parameters under shape restrictions. Journal of Econometrics 2005; 126; 53-77.

Pedregal DJ, Young PC. Modulated cycles, and approach to modelling periodic components from rapidly sampled data. International Journal of Forecasting 2006; 22; 181-194.

Poirier DJ. The Econometric of Structural Change with Special Emphasis on Spline Functions. North Holland Publishing Company: Amsterdam; 1976.

| Harvest | Week of | Length of the | Average | Minimur | n price | Maximum price | |
|-----------|----------|-----------------|---------|----------|---------|---------------|-------|
| | the year | seasonal period | price | Week | Price | Week | Price |
| 1987/1988 | 41-20 | 32 | 7,91 | 44 | 3,58 | 11 | 12,60 |
| 1988/1989 | 43-19 | 29 | 7,53 | 43 | 4,60 | 14 | 12,36 |
| 1989/1990 | 43-22 | 32 | 8,41 | 43,44,45 | 4,60 | 5 | 12,73 |
| 1990/1991 | 41-22 | 34 | 7,96 | 44 | 4,81 | 50 | 11,53 |
| 1991/1992 | 41-26 | 38 | 7,41 | 26 | 5,34 | 11 | 10,95 |
| 1992/1993 | 40-25 | 38 | 6,97 | 23 | 4,39 | 4 | 8,84 |
| 1993/1994 | 40-26 | 39 | 6,39 | 22 | 4,13 | 13 | 9,93 |
| 1994/1995 | 40-25 | 38 | 5,89 | 24 | 3,57 | 11 | 9,18 |
| 1995/1996 | 41-26 | 38 | 6,19 | 23 | 3,73 | 51 | 11,04 |
| 1996/1997 | 42-26 | 37 | 5,61 | 21 | 3,91 | 51 | 9,00 |
| 1997/1998 | 41-23 | 35 | 6,00 | 23 | 3,71 | 52 | 9,08 |
| 1998/1999 | 42-26 | 37 | 6,21 | 26 | 3,36 | 51 | 9,92 |
| 1999/2000 | 40-25 | 38 | 6,44 | 23 | 3,78 | 13 | 10,23 |
| 2000/2001 | 39-25 | 39 | 6,67 | 23,25 | 4,77 | 48 | 9,30 |
| 2001/2002 | 40-22 | 35 | 7,99 | 21 | 3,13 | 13 | 14,70 |
| 2002/2003 | 41-25 | 37 | 6,06 | 45 | 4,14 | 11 | 8,39 |
| 2003/2004 | 41-24 | 36 | 6,00 | 23 | 4,21 | 12 | 8,97 |
| 2004/2005 | 42-26 | 37 | 7,49 | 26 | 4,00 | 3 | 11,72 |
| 2005/2006 | 42-24 | 35 | 5,84 | 23 | 4,09 | 19 | 8,33 |
| 2006/2007 | 41-21 | 33 | 6,71 | 48 | 4,15 | 14 | 10,66 |

Table 1. Prices level characteristics

Note: Due to the presence of leap years in the sample, there are some observations corresponding to the week 53. The starting point of the harvest has been moved forward by one week in these cases.

Table 2. Estimates of intervention parameters

| Parameter | $\lambda_{41,1991}$ | $\lambda_{51,1995}$ | $\lambda_{26,2000}$ |
|-----------|---------------------|---------------------|---------------------|
| Estimate | -0.048676 | 4.8850 | 4.9761 |





Figure 2. Seasonal effects (previous approximation and final estimates)

