QUANTITATIVE METHODS IN ECONOMICS
Vol. XIII, No 2, 2012, pp. 31 – 39

COMPARISON OF THE BEEF PRICES IN SELECTED COUNTRIES OF THE EUROPEAN UNION

Stanisław Jaworski
Department of Econometrics and Statistics
Warsaw University of Life Sciences – SGGW
e-mail: stanislaw_jaworski@sggw.pl

Abstract: Functional data analysis is used to examine beef price differences in selected countries of the European Union from 2006 to 2011. The prices are modeled as functional observations. The analysis is conducted in three steps relating to three kinds of functional data analysis. First the observations are smoothed with roughness penalty. Then functional principal analysis is applied. Finally functional analysis of variance is used to reveal significant difference between two given groups of countries.

Keywords: B-splines basis system, functional principal component analysis, functional analysis of variance, permutation tests

INTRODUCTION

The goal of the paper is to compare beef prices in European countries since 2006 to 2011. The price data are collected monthly and come from the website of Ministry of Agricultural and Rural Development (http://www.minrol.gov.pl). The main characteristic of the prices is that they don't change rapidly. Two consecutive prices are unlikely to be too different from each other, so it seems reasonable to turn the raw price data into smooth functions and think of the observed data as single entities, rather than as a sequence of individual observations. A linear combinations of basis functions is used as a method for representing smooth functions. The basis function approach is designed to reveal the most important type of variation from the smoothed prices. A key technique in the approach is a functional principal component analysis.

The particular aim of the paper is to find out if there is a significant difference between beef prices considering old and new members of the European

Union. In that case dependent variable is modeled as a functional observation so the methodology needed is a functional analysis of variance.

It is assumed that the first group, referring to the old members, consists of Belgium, Denmark, Germany, Greece, France, Spain, Ireland, Italy, Luxemburg, Nederland, Austria, Portugal, Finland, Sweden and United Kingdom. The second group of the new members consists of Czech Republic, Estonia, Latvia, Lithuania, Poland, Slovenia and Slovakia.

METHODS

It is assumed that the beef price y_{ij} in time t_j related do the i-th country has the form

$$y_{ij} = x_i(t_j) + \varepsilon_{ij}, \quad j = 1, 2, \ldots, N \tag{1}$$

where ε_{ij} is an unspecified random error and $x_i(t) = \sum_{i=1}^{K} c_{ik}\phi_k(t)$ is a smoothed price expressed as a linear combination of B - splines basis system $\{\phi_k\}$ (see Ramsay, Hooker, Graves (2009), p. 35). The coefficients $\{c_{ik}\}$ of the expansion are determined by minimizing, for each i , the least squares criterion

$$\sum (y_{ij} - x_i(t_j))^2 + \lambda \int \left[D^2 x_i(s)\right]^2 ds \tag{2}$$

Details of this approach can be found in Ramsey and Silverman (2005). The parameter λ is fixed. It can be selected arbitrarily or by minimizing Generalized Cross-Validation (GCV) measure (see Ramsay and Silverman (2005), p.97).

The smoothed prices are used in a functional principal component analysis (see Besse and Ramsey (1986), Ramsey and Dalzell (1991) and Besse, Cardot and Ferraty (1997)). In the analysis the weight functions $\xi_1, \xi_2, \ldots, \xi_K$ are chosen consecutively. Each consecutive weight function ξ_m maximize

$$\frac{1}{n}\sum_{i=1}^{n}\left(\int \xi_m(s)x_i(s)ds\right)^2 \tag{3}$$

subject to

$$\int \xi_k(s)\xi_m(s)ds = 0 \text{ and (for } k < m) \tag{4}$$

The vector $f_m = (f_{1m}, f_{2m}, \ldots, f_{nm})$ where $f_{im} = \int \xi_m(s)x_i(s)ds$, $i = 1, 2, \ldots, n$, is called the m-th principal component. The percentage of variability of the first m components is expressed as

$$\frac{\sum_{j=1}^{m}\sum_{i=1}^{n} f_{ij}^{\,2}}{\sum_{j=1}^{K}\sum_{i=1}^{n} f_{ij}^{\,2}}100\% \tag{5}$$

The difference between beef prices for considered groups is investigated by functional analysis of variance. In formal terms, we have a number of countries in each group $g = 2$, and the model for the m th price function in the g th group, indicated by Price_{mg}, is

$$\mathrm{Price}_{mg}(t) = \mu(t) + \alpha_g(t) + \varepsilon_{mg}(t). \tag{6}$$

The function μ is the grand mean function, and therefore indicates the average mean price profile across all of countries. The terms α_g are the specific effects on price of being in group g. It is required that they satisfy the following constraint

$$\sum_g \alpha_g(t) = 0 \text{ for all } t \tag{7}$$

The residual function ε_{mg} is the unexplained variation specific to the m th price within group g.

As in ordinary analysis of variance F-ratio and t-test statistics can be calculated. Let denote them as $F(t)$ and $T(t)$ accordingly. These statistics can be used point-wise but it is desired to account for significant difference at different times. So the following statistics can be considered:

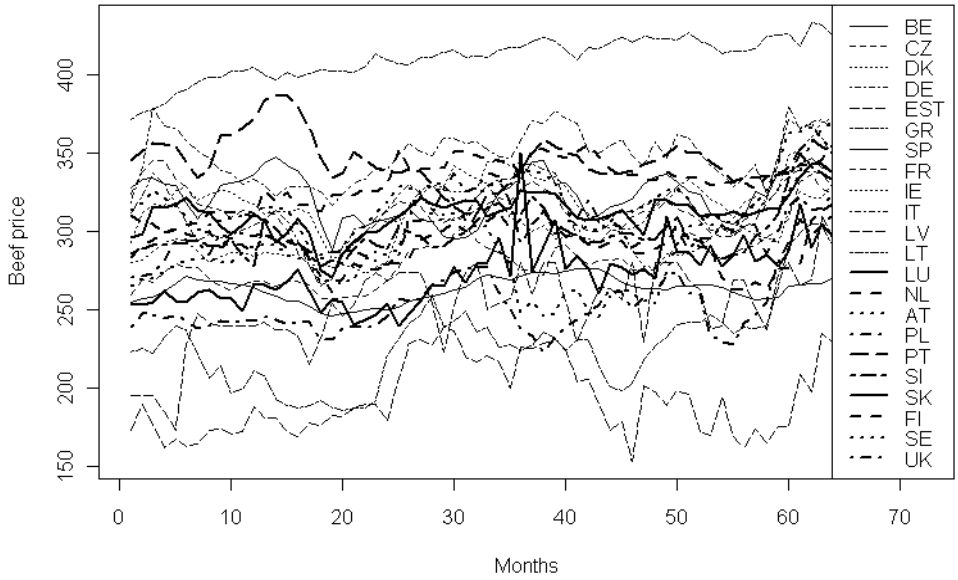
$$F = \sup_t F(t) \text{ and } T = \sup_t |T(t)| \tag{8}$$

A permutation-based significance value can obtained for these statistics.

DATA ANALYSIS AND RESULTS

In this chapter the European mean beef prices are considered. The following countries are taken into account: Belgium (BE), Czech Republic (CZ), Denmark (DK), Germany (DE), Estonia (EST), Greece (GR), Spain (SP), France (FR), Ireland (IE), Italy (IT), Latvia (LV), Lithuania (LT), Luxembourg (LU), Netherlands (NL), Austria (AT), Poland (PL), Portugal (PT), Slovenia (SI), Slovakia (SK), Finland (FI), Sweden (SE), United Kingdom (UK). The beef prices are presented in Figure 1.

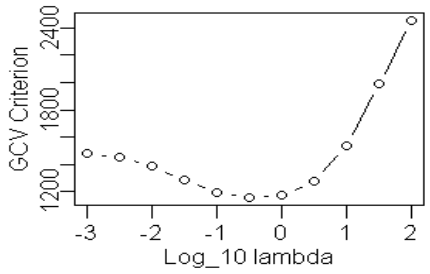
Figure 1. Beef prices in selected countries



Source: own preparation

The data were smoothed with B-splines system with roughness penalty parameter $\lambda = 1$. The choice of the parameter's value was based on generalized cross-validation plot (Figure 2).

Figure 2. GCV plot



.Source: own preparation

The generalized cross-validation measure (GCV) is popular in the spline smoothing literature and was developed by Craven and Wahba (1979). The criterion is usually expressed as

$$GCV(\lambda) = \left(\frac{n}{n - df(\lambda)}\right)\left(\frac{SSE}{n - df(\lambda)}\right) \tag{9}$$

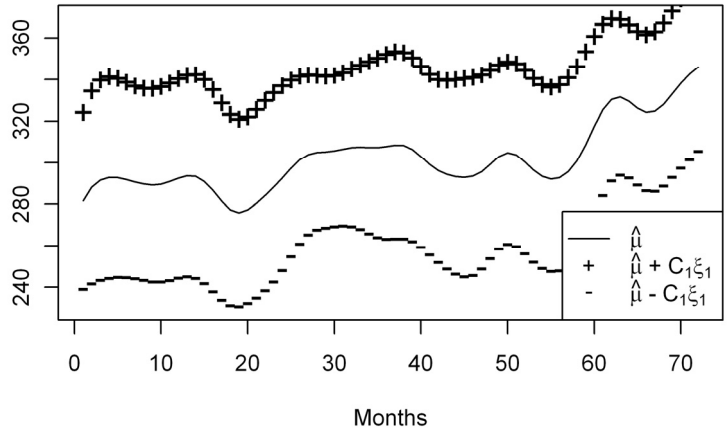
where SSE is a sum of squared errors and $df(\lambda)$ is a degrees of freedom for a spline smooth. Figure 2 shows the variation of the generalized cross-validation statistic GCV (that is $GCV(\lambda) = \sum_i GCV_i(\lambda)$, where index i relates to x_i) over a range of $\log_{10}(\lambda)$ values

The generalized cross validation measure is minimized at $\lambda = 0.1$ and it is slightly higher for $\lambda = 1$. The bigger value was chosen because GCV criterion yields under-smoothing (see C.Gu (2002))

Next the smoothed data were explored by functional principal components analysis. The first two principal components account for 95% of the total variation.

A method found to be helpful in interpreting the components is to examine plots of the overall mean function $\hat{\mu}(t) = \frac{1}{n}\sum_{i=1}^{n} x_i(t)$ and the functions obtained by adding and subtracting a suitable component functions: $\hat{\mu} \pm C_m \xi_m$, $m = 1,2$, where $C_m = \frac{1}{n}\sum_i f_{im}^2$. Figures 3 and 4 show such plots of the price data. In each case the solid curve is the overall mean price and the other two curves clarify the effect of a given principal component. The effect of the first principal component of variation (covering 92% of the total variation) means that the greatest variability between countries corresponds to variability of its average beef price levels. As can be seen from the plot in Figure 3 this source of variability represents countries where, between 2006 and 2011, prices are at either low or high level. The i-th country for which the score f_{i1} is high have much higher than average beef prices.

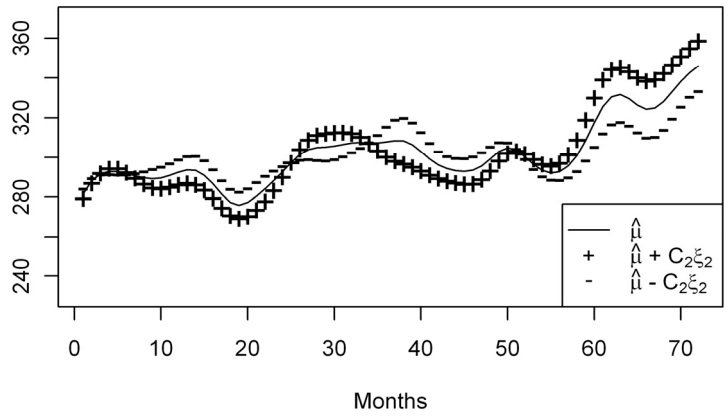
Figure 3. The first principal component curves



Source: own preparation

The second source of variation (Figure 4) is more complicated then the first one. It corresponds to countries where beef prices changed its relative level three times since 2006. This source of variation covers merely 3% of total variation.

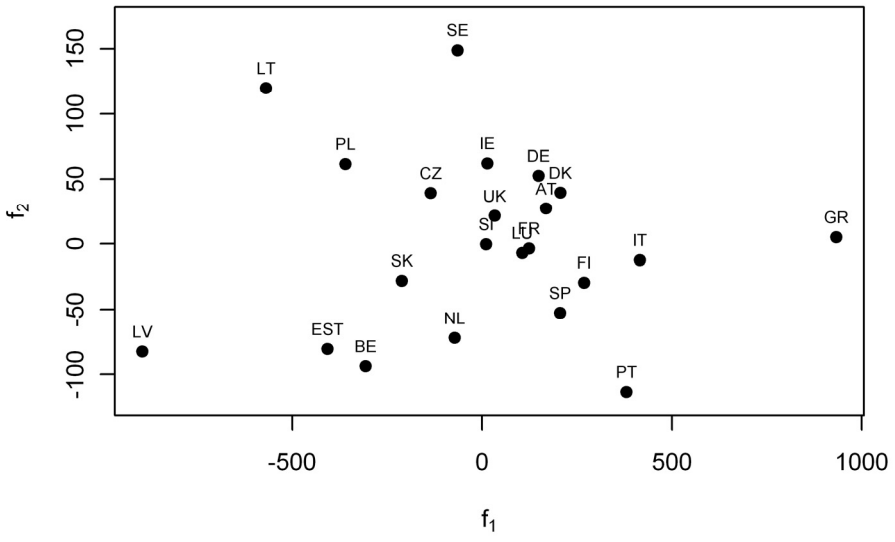
Figure 4. The second principal component curves



Source: own preparation

A good insight into the differences between countries can be made by plotting principal scores (Figure 5).

Figure 5. Principal components of beef prices: the set of scores $\{(f_{i1}, f_{i2}) : i = 1,2,\ldots,n\}$.



Source: own preparation

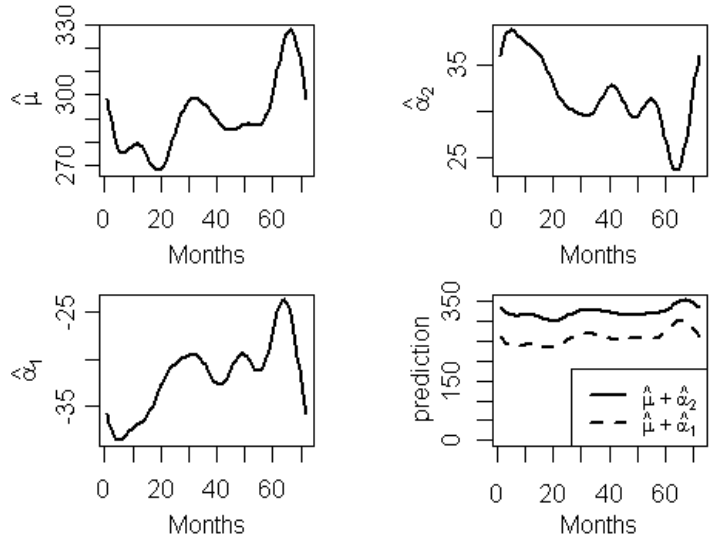
Majority of former members of Eastern Bloc are placed to the left side of the plot. It means that the countries have low beef prices since 2006. Opposite side of the plot relates to the countries with high beef prices in the period, for example to Greece, Italy and Portugal.

The conclusions drawn from principal component analysis mean that there can be a significant difference between beef prices with respect to if a country is an old or new member of European Union. Functional analysis of variance model is used in the paper to confirm this supposition. Estimated parameters: μ, α_1, α_2 and prediction are presented in Figure 6.

The estimate of α_1 is negative. It suggests that the new members have lower beef prices then the old ones. Although the parameter is changing over time the prediction plot in Figure 6 suggests that the difference between the two groups is relatively constant. The hypothesis that α_1 is equal to zero was verified by permutation test. It is presented in Figure 7. The dashed line gives the permutation 0.05 critical value for the T - statistic and the dotted curve the permutation critical value for the point-wise T(t) - statistic

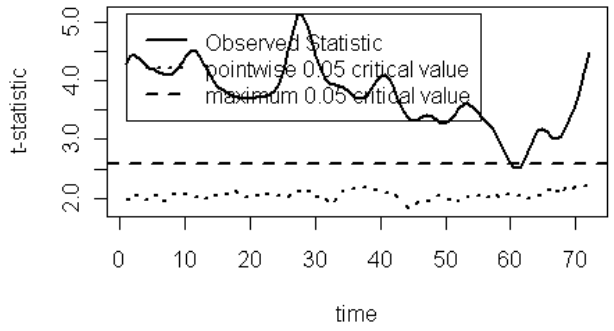
The test confirmed that there was a statistically significant difference between beef prices for the two considered groups.

Figure 6. Parameters and prediction of functional analysis of variance model



Source: own preparation

Figure 7. Permutation -based significance values for $T(t)$ and T statistics



Source: own preparation

SUMMARY

Some features of mean beef prices across European countries were uncovered with functional data analysis. The most important type of beef price variations was revealed with help of smoothing techniques and functional component analysis. The difference between two separated groups of countries was investigated in terms of functional analysis of variance and permutation tests.

Some conclusions can be drawn. Apart from the overhead beef price in Europe increased since 2006 the source of the greatest beef price variability didn't change over the investigated time (see Figure 3). The new members of European Union have still lower beef prices than the old ones. The difference is relatively constant as can be seen from prediction plot in Figure 6. It is interesting to note that such countries as Latvia, Lithuania, Estonia, where the beef prices are at low level (see Figure 5), are not the eurozone member states and the countries with high level of beef prices, for example Greece, Italy and Portugal, are highly indebted now. It means that it would be interesting to involve more explanatory variables for the price data analysis and provide more sophisticated model of functional regression then the presented functional variance model.

REFERENCES

Besse P.C., Cardot H. and Ferraty F. (1997) Simultaneous nonparametric regressions of unbalanced longitudinal data. Computational Statistics and Data Analysis, Vol. 24, 255-270.

Besse P.; Ramsay J.O. (1986) Principal components analysis of sampled functions. Psychometrika, 51, 285-311.

Craven, P. and Wahba, G. (1979) Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation, Numerische Mathematik, 31, 377-403.

Gu, C. (2002) Smoothing Spline ANOVA Models", Springer.

Ramsay J.O, Dalzell C.J (1991) Some tools for functional data analysis (with discussion). Journal of the Royal Statistical Society, Series B, 53:539-572.

Ramsay J.O, Silverman B. W. (2005) Functional Data Analysis. Second Edition, Springer, NY.

Ramsay J.O, Hooker G., Graves S. (2009) Functional Data Analysis with R and MATLAB. Springer.