Modelling structural change using broken sticks

Paul White[1], Don J. Webber[2] and Angela Helvin[1]

[1] *Department of Mathematics and Statistics, University of the West of England, Bristol, UK*
[2] *Department of Economics, University of the West of England, Bristol, UK*

Abstract

This paper presents a "broken stick" method to test for structural breaks in a regression model. The method is illustrated using output data across the EU and the results are bootstrapped to identify statistical significance.

Keywords: Chow test; Broken stick regression

JEL Classification: C12; C22; E32

Corresponding Author: Dr Don J Webber, School of Economics, Bristol Business School, University of the West of England, Bristol, BS16 1QY. E-mail: Don.Webber@uwe.ac.uk. Tel.: (+44/0) 117 328 2741. Fax: (+44/0) 117 328 2289

1. Introduction

In regression analyses the concept of lack-of-fit may be paraphrased as the "need to have different regression equations in different parts of the regression space to adequately model a response variable, *Y*". The Chow test (Chow, 1960) is often used to determine whether a regression model can be improved upon by incorporating a structural break into a regression model. If such a break is warranted then the analyst may choose to proceed with separate regression equations or may incorporate a break in the model by a judicious use of dummy variables so as to model a differential effect of an explanatory variable *X* with the response *Y*. Others (see Andrews (1993), Hansen (1992) and Greene (1999)) have suggested alternative methods to identify structural changes.

This paper presents an alternative approach to modelling structural breaks through the use of what may be termed a "broken stick" model and is applied to EU-wide GDP per capita data. Broken stick regression is the modelling of two or more intersecting straight lines with the knots (break points) forming the piecewise linear regression either identified prior to data collection based on theoretical considerations or break points identified in exploratory analyses. Section 2 presents the broken stick models, Section 3 reviews the data, the application of the models and assessments of their statistical significance are provided in Sections 4 – 7, and Section 8 concludes.

2. Broken Stick Models

Let X denote a potential explanatory variable of a response *Y*. Define

$$X_1 = \begin{cases} X & X \le k \\ k & X > k \end{cases}$$

$$X_2 = \begin{cases} 0 & X \le k \\ (X - k) & X > k \end{cases}$$

and consider a model defined by

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

In the above model formation ε denotes a random error component and $\beta_0 + \beta_1 X_1 + \beta_2 X_2$ denotes the structural part of the model. For $X \le k$ the structural part of the above model reduces to $Y = \beta_0 + \beta_1 X$. For $X > k$ the structural part of the model reduces to $Y = \beta_0 + (\beta_1 - \beta_2)k + \beta_2 X$. If $X = k + \delta x$, $\delta x > 0$ then in the limit as $\delta x \to 0$, $Y \to \beta_0 + \beta_1 X$. Accordingly the above model specification is piecewise continuous and β_1 indicates the rate of change of X with *Y* for $X \le k$ and β_2 indicates the rate of change of X with Y for $X > k$. In the special case of $\beta_1 = \beta_2$ the model logically reduces to a simple linear relationship between *X* and *Y*.

A model formulation allowing a double "break" with a break point at X = k_1 and another break point at X = k_2 (with $k_2 > k_1$) is

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

where

$$X_1 = \begin{cases} X & if \ X \le k_1 \\ k_1 & if \ X > k_1 \end{cases}$$

$$X_2 = \begin{cases} 0 & if \ X \le k_1 \\ X - k_1 & if \ k_1 \le X \le k_2 \\ k_2 - k_1 & if \ X > k_2 \end{cases}$$

$$X_3 = \begin{cases} 0 & if \ X \le k_2 \\ X & if \ X > k_2 \end{cases}$$

The above specification provides a piecewise continuous linear model. In this model β_1 denotes the rate of change of X with Y for $X \le k_1$; β_2 indicates the rate of change of X with Y for $k_1 < X \le k_2$ and β_3 indicates the rate of change X with Y for X > k_2.

More generally a model comprising B broken sticks with break points k_b (b=1, ⋯⋯, $B-1$), $k_{b-1} < k_b$, is defined by

$$Y = \beta_0 + \sum_{b=1}^{B-1} \beta_b X_b + \varepsilon$$

where

$$X_b = \begin{cases} 0 & if \ k_{b-1} > X \\ X - k_{b-1} & if \ k_{b-1} \le X < k_b \\ k_b & if \ X > k_b \end{cases}$$

3. Data

We applied the above models to EUROSTAT-sourced GDP per capita (in 1995 prices) data for the 1980 – 2008 period. The observations are raw averages for all countries included in the sample for each year. Note the data also includes Cyprus, Czech Republic, Estonia, Hungary, Lithuania, Latvia, Malta, Poland, Slovenia and Slovak Republic from 1990 onwards only.[1] There appears to be (at least) one important break point in this dataset, as captured in Figure 1. In addition to the accession of countries, the early 1990s was plagued by severe structural changes and a decline in output and high inflation rates, 1993 saw the advent of The Maastricht Treaty, 1999 saw the official beginning of European Monetary Union and EU membership widened in several years.[2] All of these events may be associated with a change in the slope of the path on the graph.

4. Applying the Single Break Model

Consider a broken stick model with a single break at $X = k$ and structural specification $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$. For the sample EU data we consider GDP per capita to be the dependent variable (Y) and code the years (1980, 1981, …, 2008) with the values $X = 1, 2, \ldots, 29$. Table 1 summarises the coefficient of determination, R^2, for each ordinary least squares fit for each possible single break point for the sample data. Under this approach the value of k that minimises the within sample error sum of squares is $k = 16$, corresponding to the year 1996. Estimated parameters ($\hat{\beta}_0 = 14{,}026$; $\hat{\beta}_1 = 79.4$; $\hat{\beta}_2 = 312.0$) and a summary of standard tests of significance ($H_0 : \beta_j = 0$) is given in Table 2.

5. Assessing Significance of the Single Break Model

The standard tests of significance, such as those reported in Table 2, do not readily answer the question of interest; namely whether the inclusion of a single breakpoint provides a real improvement in overall fit compared with a simple linear specification. Such an improvement would only be apparent if in fact $\beta_1 \neq \beta_2$. Bootstrap procedures (Davidson and Hinkley, 1997) provide a means to validly answer this question.

A first stage in the bootstrap assessment of significance of a single break model is to fit a simple linear model and to obtain the predicted values ($\hat{y}_i$) and sample residuals (e_i). A new bootstrap data set adhering to a simple linear specification is obtained by sampling the residuals with replacement to create bootstrap residuals r_i and to form a new data set (x_i, $\hat{y}_i + r_i$), $i = 1, \ldots, n$. This newly created sample is used to fit the single break model and a measure of overall model fit (e.g. R^2) is recorded. Repeating this process with B bootstrap samples produces an empiric distribution for which the corresponding observed sample statistic (e.g. R^2) may be judged against.

[1] Data for Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the UK are included throughout the period.

[2] During this period, accessions to EU membership took place in 1986 (Spain and Portugal), 1995 (Austria, Finland and Sweden), 2004 (Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovak Republic and Slovenia), and 2007 (Bulgaria and Romania).

To assess the statistical significance of the one-break model, two versions of the bootstrap algorithm have been implemented. In the first case (Case A1) we consider the breakpoint to be fixed at k = 16 and in the second case we consider the value of k to be determined by the data (Case A2). Algorithmically, for Case A1,

(1a) generate a bootstrap sample using the simple linear (no break) model
(2a) for the given bootstrap sample fit the one break model with k = 16
(3a) record the R^2 value for the fitted model.

Steps (1a), (2a), (3a) are repeated B times. The proportion of times that the R^2 values from bootstrap samples exceed the corresponding value of R^2 for the sample data is recorded. This proportion gives a bootstrap estimate of the p-value for testing $H_0 : \beta_1 = \beta_2$ against $H_1 : \beta_1 \neq \beta_2$ assuming k = 16.

In the second case, Case A2, the algorithm implemented has the form

(1b) generate a bootstrap sample using the simple linear (no break) model
(2b) fit the single break model using the given bootstrap sample; let k be bootstrap sample dependent and choose the value of k to be that value which maximises the within sample goodness-of-fit, R^2, for the given sample
(3b) record the R^2 value for the fitted model.

Steps (1b), (2b), (3b) are repeated B times. The proportion of times that the R^2 values from bootstrap samples exceed the corresponding value of R^2 for the sample data is recorded. This proportion gives a bootstrap estimate for testing $H_0 : \beta_1 = \beta_2$ against $H_1 : \beta_1 \neq \beta_2$ without a pre-specification of k.

Table 2 summarises the fitted simple linear regression model used to generate the bootstrap samples and also summarises the single break model using k = 16 (corresponding to 1996). All effects are statistically significant; the positive sign for the coefficients indicates that under this formulation the rate of change up to 1996 is positive (79.39 euros per capita per year) and this increases to 311.83 euros per capita per year. Application of the bootstrap algorithm, Case A1, using B = 5,000 bootstrap samples provides an estimated p-value of 0.019 i.e. 97 of the 5,000 bootstrap samples had an R^2 values that exceeded the observed sample R^2 (91.1%) statistic. Application of the bootstrap sample algorithm, Case A2, using B = 5,000 bootstrap samples provides an estimated p-value of 0.146 i.e. approximately fifteen percent of the bootstrap samples produced an R^2 value that exceeded the observed sample value of R^2 = 91.1%, when k was determined by maximizing the bootstrap sample of R^2.

6. Applying the Two Break Model

Consider a broken stick model with a break at $X = k_1$ and another break at $X = k_2$ and with structural specification Y = $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$. In application there are two major situations to consider; namely when $(k_1,\ k_2)$ are specified by external theoretical considerations, and otherwise when $(k_1,\ k_2)$ are to be estimated using sample data.

For the sample EU data all possible pairs of values for k_1 and k_2 have been considered. For each possible pair the R^2 coefficient of determination has been obtained. Figure 2 provides a summary contour plot of the values of R^2 for each possible double break. The "top" five models, in terms of within sample goodness-of-fit are (k_1, k_2, R^2) = (10, 11, 98.1), (k_1, k_2, R^2) = (10, 12, 97.9), (k_1, k_2, R^2) = (10, 13, 97.4), (k_1, k_2, R^2) = (9, 13, 97.0) and (k_1, k_2, R^2) = (9, 14, 96.7). Taken at face value the differences in the R^2 values between these models seem quite minimal. From the possibilities listed a potential criticism of the first four models is the relatively small gap between breakpoints which may reflect local minima arising as chance random patterns in the data and on this basis the fifth model will tentatively be taken as a good two-break model for the sample data.

7. Assessing Significance of the Double break model

The double break model Y = $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$ reduces to the simple linear model when $\beta_1 = \beta_2 = \beta_3$. Bootstrapping may be used to test the null hypothesis H_0 : $\beta_1 = \beta_2 = \beta_3$ against H_1: "At least two values of β_j differ".

If break points can be considered to be pre-specified then the bootstrap algorithm for these hypotheses for the case under consideration would be along the following lines:

(1c) generate a bootstrap sample using the simple linear (no break) model
(2c) for the given bootstrap sample fit the double break model with $k_1 = 9$ and k_2 = 14
(3c) record the R^2 value for the fitted model.

Steps (1c), (2c), (3c) are repeated *B* times. The proportion of times that the R^2 values from bootstrap samples exceed the corresponding value of R^2 for the sample data is recorded. This proportion gives a bootstrap estimate for testing H_0 : $\beta_1 = \beta_2 = \beta_3$ against H_1: "At least two values of β_j differ" assuming a pre-specification of k_1 and k_2. If however k_1 and k_2 are not pre-specified and are estimate from the data from a position of ignorance then the values of k_1 and k_2 may be determined for each bootstrap sample to be those values that minimise within bootstrap sample variation.

The above bootstrap algorithms have been applied using the sample data (*B* = 5000 in both instances). In these bootstrap samples the derived values of R-squared for the bootstrapped two break models never exceed the sample R-squared value (96.7%) for the given two break model. Accordingly, the two-break model provides a statistically significant improvement over the simple linear (no break) specification and this is irrespective of whether k1 and k2 are considered fixed with value 9 and 14 respectively or whether they are considered as values uncovered from the data under a position of initial ignorance.

We may also entertain whether the double break model may be considered as an improvement on the single break model. The double break model Y = $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$ reduces to the single break model when $\beta_1 = \beta_2$ or when β_2 =

β_3. Bootstrapping may be used to test the null hypothesis H_0 : $\beta_1 = \beta_2$ or $\beta_2 = \beta_3$ against H_1: $\beta_1 \neq \beta_2$ and $\beta_2 \neq \beta_3$. In this set of bootstrap model the best fitting single break model will be used to generate the predicted values and residuals. A comparison of the given two break model (k1 = 9; k2 = 14) against similar two break bootstrap models with the null hypothesis specified by the best one break model (k1 = 16) indicates that the given two break model is a statistically significant improvement over the one break model ($p < 0.001$). Similarly a comparison of the given two break model (k1 = 9; k2 = 14) against the best possible two break model for each bootstrap sample with the null hypothesis specified by the best one break model (k1 = 16) has been undertaken. In these cases 30 bootstrap samples out of 5000 bootstrap samples give a higher R-squared value than the given sample R-squared value (96.7%). Accordingly, even after allowing for chance effects from "over analysing the data" the given two break model may be deemed to be a significant improvement over the one break model ($p = 0.0006$).

8. Conclusion

The Chow test has been employed in econometrics for over 40 years. In this paper we propose an alternative “broken stick” method to identify structural breaks in a regression model. The method is illustrated using output data across the EU and the results are bootstrapped to identify statistical significance.

References

Andrews, D. W. K. 1993. Tests for parameter instability and structural change with unknown change point. Econometrica 61(4), 821-856

Chow, G. C. 1960. Tests of equality between sets of coefficients in two linear regressions. Econometrica 28(3), 591-605.

Davidson, A. C. and D. V. Hinckley, 1997. Bootstrap methods and their applications. Cambridge University Press, Cambridge

Hansen, B. E. 1992. Tests for parameter instability in regressions with I(1) processes. Journal of Business and Economic Statistics 10(3), 321-336

Greene, C. A. 1999. Testing for a break at an unknown change-point: A test with known size in small samples. Economics Letters 64, 13-16

Table 1. R^2s

Break	R^2	$R^2(adj)$		Break	R^2	$R^2(adj)$		Break	R^2	$R^2(adj)$
1	0.829	0.823		10	0.850	0.838		19	0.900	0.892
2	0.831	0.818		11	0.865	0.855		20	0.893	0.884
3	0.832	0.819		12	0.880	0.871		21	0.885	0.876
4	0.832	0.819		13	0.893	0.884		22	0.878	0.869
5	0.832	0.819		14	0.903	0.896		23	0.868	0.858
6	0.833	0.820		15	0.909	0.902		24	0.863	0.852
7	0.834	0.821		16	0.911	0.905		25	0.856	0.845
8	0.837	0.824		17	0.910	0.904		26	0.849	0.837
9	0.841	0.829		18	0.906	0.899		27	0.840	0.828

Table 2: Summary of standard tests of significance

	No Break			One Break (k = 16)			Two Break (k1 = 9, k2 = 14)		
Effect	Coef	t	p	Coef	t	p	Coef	t	p
Intercept (b0)	13192	47.91	<0.001	14026	53.16	<0.001	13050	57.16	<0.001
Gradient b1	183.58	11.45	<0.001	79.4	3.27	0.003	296.9	7.91	<0.001
Gradient b2				312.0	10.89	<0.001	−250.1	−5.27	<0.001
Gradient b3							330.8	20.14	<0.001
R^2	82.9%			91.1%			96.8%		

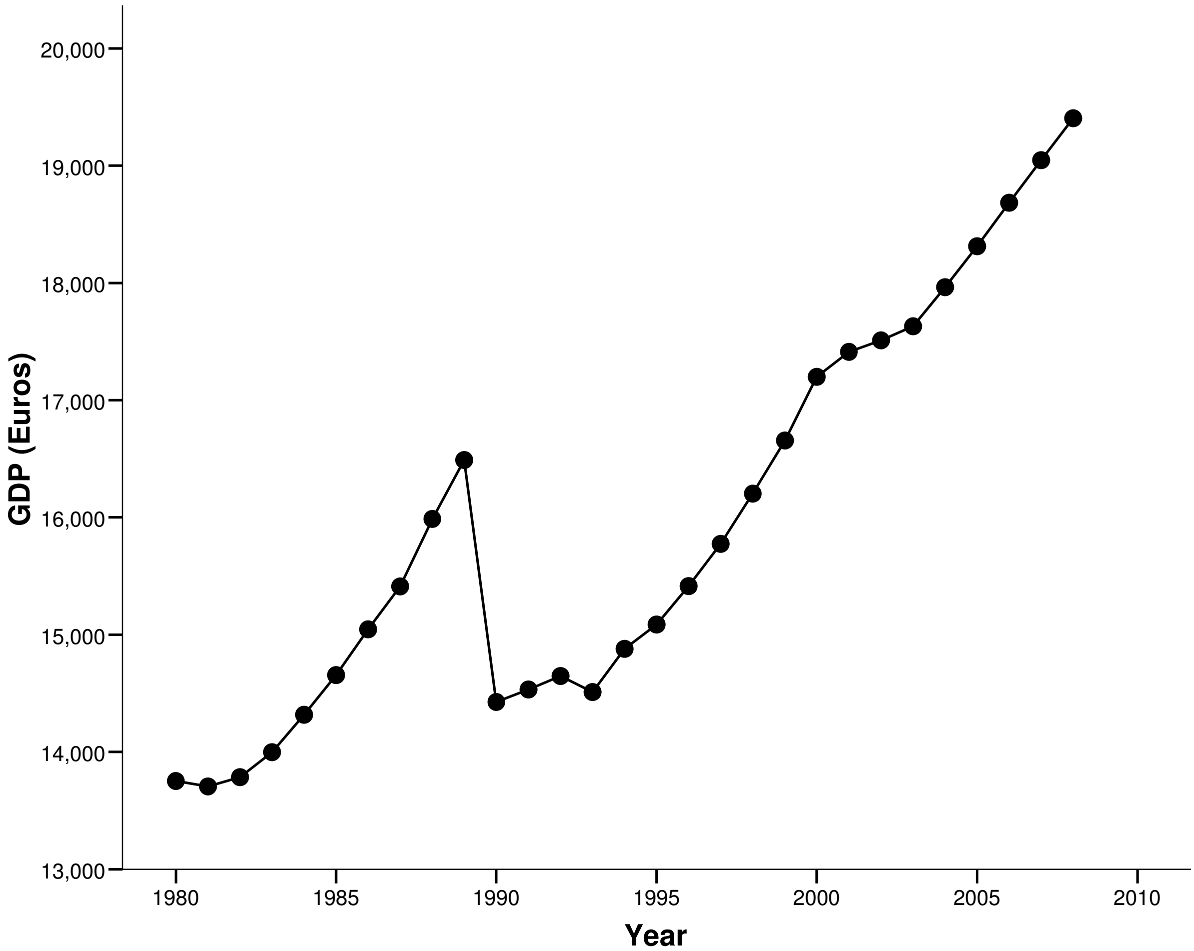

Figure 1 GDP per capita in Euros for EU

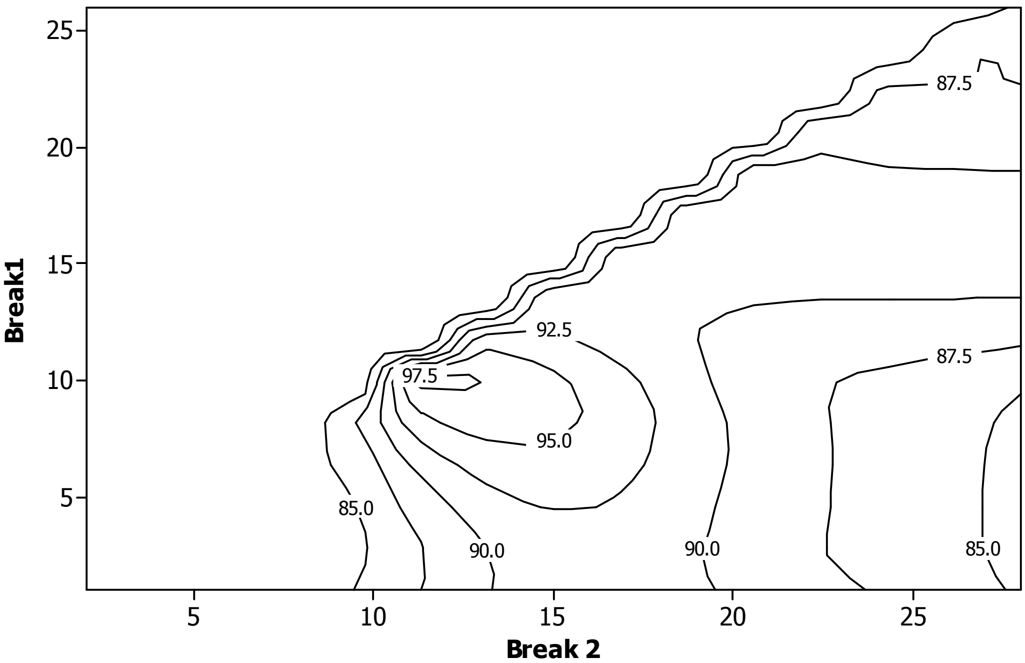

Figure 2: Contour plot