Vidyodaya J. Arts, Sci., Lett., Vol. 6 No. 1, 1978, pp. 1 - 8

Estimation of Parameters of Switching Regression Model When the Switching Point is Unknown

by

W. A. JAYATISSA

Department of Economics, Vidyodaya Campus

INTRODUCTION:

In econometric literature there are several testing procedures to detect structural changes in various econometric models when the parameters changing points (switching points) are a priori known. However, there are some situations we have strong reasons for suspecting that the parameters of a model are not constant over the given sample period but we do not have a priori knowledge about the parameters changing points. In the literature this variation has been modelled in two principal ways. The first approach typically allows for an infinite number of possible parameters values and for random parameter variations. In this case the appropriate econometric techniques is based on the random coefficient regression model. In the second approach we assume that the number of possible parameters changes may be finite (usually small). In this paper we shall consider the second approach by assuming that the number of parameters changes within a given sample is one. In otherwords there are only two regimes associated with a given sample.

This type of problem does arise in econometric studies. As an example consider the consumption function $C = a+bY$. Aggregate consumption depends upon the level of aggregate (disposable) income. In addition, it may be hypothesised that consumption depends non-linearly on other factors such as the state of expectations concerning the future of the economy, the volume of instalment buying, the level of the interest rate, etc. These other variables may have the effect of altering the parameters of the consumption function in the following fashion: when the critical outside variable (say, the rate of interest) i satisfies

$$i < i^*$$

then

$$C = a_1 + b_1 Y$$

and when

$$i \geqslant i^*$$

$$C = a_2 + b_2 Y$$

where i^* is the level of the outside variable in question.

ESTIMATION OF PARAMETERS OF SWITCHING REGRESSION MODEL WHEN THE SWITCHING POINT IS UNKNOWN

In general, one may not be able to identify the critical outside variable and, as a result of this, one may not be able to state at what time the system $C = a + bY$ changes from one regime to the other.

There are two fundamental problems associated with this type of regression model.

(i) Given a sample of observations how can we test whether there is any shift of coefficients or not?

(ii) If there is a shift of coefficients how can we estimate the switching point and the parameters of the model?

Several studies have been done with respect to the first problem. Therefore we limit our study to consider only the second problem.

2. Bayesian Analysis of a Switching Regression Model

2.1 Basic Assumptions

Let us consider the multiple regression model

$$y_i = x_{i.} \beta_1 + e_i \qquad i = 1, \ldots\ldots, t \qquad (1)$$

$$y_i = x_{i.} \beta_2 + e_i \qquad i = t+1, \ldots\ldots, T \qquad - (2)$$

where y_i is the i th observation on the dependent variable, $x_{i.}$ is the i th observation on the k independent variables, β_1, β_2 are the parameter vectors for two regimes, e_i is the i th value of the disturbance term and t is the switching time point. To analyse this model we make the following assumptions:

(1) The e_i, $i=1, \ldots\ldots, T$ are normally and independently distributed, each with mean zero and common variance σ^2.

(ii) The $x_{i.}$ $i = 1, \ldots\ldots, T$, are fixed nonstochastic variables or they are random variables distributed independently of e_i, with a probability density function not involving the parameters β_1, β_2, and σ^2.

(iii) All unknown parameters have independent prior distributions and β_1, β_2 and $\log \sigma^2$ are uniform over the whole real line. The assumption of the uniform distributions means that our prior knowledge about β_1, β_2 and $\log \sigma^2$ is vague compared with that obtained from the observations.

(iv) Prior distribution of t is p(t). The only assumption about it is that $p(t \mid t=1) = \ldots\ldots\ldots = p(t \mid t=k-1) = o$, $p(t \mid t=T-(k-1)) = \ldots\ldots = p(t \mid t=T) = o$. This assumption guarantees that both regimes consist at least k observations.

2.2 Posterior Distributions of the Parameters.

Let $\theta = (\beta_1, \beta_2, \sigma^2, t)$ be the parameter vector of the model. To form the likelihood function under assumptions (i) and (iii) we write the joint probability density function for y and x, namely

$$p(y,x \mid \theta, \mu) = p(y \mid x, \theta)\, g(x \mid \mu) \quad -(3)$$

where μ denotes the parameters of the marginal probability density function of x. Since by assumption μ does not involve any parameter in θ, the likelihood function for θ can be formed from the first factor on the right hand side of the equation (3).

Thus the likelihood of y, given x, for both models (1) and (2) is

$$p(y \mid X, \theta)\, \alpha\, \sigma^{-T} \exp\left\{ -\frac{1}{2\sigma^2}\left[(Y_1 - X_1\beta_1)^1 (Y_1 - X_1\beta_1) + (Y_2 - X_2\beta_2)^1 (Y_2 - X_2\beta_2) \right]\right\}$$

where $Y_1{}^1 = (y_1, \ldots\ldots, y_t)$, $X_1{}^1 = (x^1{}_1, \ldots\ldots, x^1{}_t)$ — (4)

and $Y^1{}_2 = (y_{t+1}, \ldots\ldots, y_T)$, $X^1{}_2 = (x^1{}_{t+1}, \ldots\ldots, x^1{}_T)$

According to these definitions we can see that Y_1, Y_2, X_1, and X_2 all are function of t.

If we define

$$\hat{\beta}_i = (X^1_i X_i)^{-1} X_i{}^1 Y_i \quad \text{and} \quad S_i^2 = (Y_i - X_i\hat{\beta}_i)^1 (Y_i - X_i\hat{\beta}_i)$$

for $i = 1,2$, we can write the likelihood function (4) in the following way :

$$\text{1. e } p(Y \mid X, \theta)\, \alpha\, \sigma^{-T} \exp\left\{ -\frac{1}{2\sigma^2}\left[S_1^2 + (\beta_1 - \hat{\beta}_1)^1 X^1{}_1 X_1 (\beta_1 - \hat{\beta}_1) + S_2^2 + (\beta_2 - \hat{\beta}_2)^1 X^1{}_2 X_2 (\beta_2 - \hat{\beta}_2) \right]\right\} \quad -(5)$$

Since the joint posterior distribution of the parameters

$$p(\theta \mid X, Y) = \frac{p(Y \mid X, \theta) p(t)}{\sigma^2}$$

where p (t) is the prior probability density function of t, we can write

$$p(\theta \mid X, Y) \propto \sigma^{-(T+2)} \exp \left\{ -\frac{1}{2\sigma^2} \left[S^2_1 + S_2^2 + (\beta_1 - \hat{\beta}_1)^1 X_1^1 X_1 (\beta_1 - \hat{\beta}_1) + (\beta_2 - \hat{\beta}_2)^1 X^1_2 X_2 (\beta_2 - \hat{\beta}_2) \right] \right\} p(t) \qquad — (6)$$

Let us define $S^2 = S_1^2 + S_2^2$, $t = k, \ldots\ldots\ldots, T-k$.

Then the posterior probability density function of t is given by

$$p(t \mid X, Y) \propto \int_0^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(\theta \mid X, Y)\, d\beta_1\, d\beta_2\, d\sigma^2$$

$$\propto \left| X^1_1 X_1 \right|^{-\frac{1}{2}} \left| X^1_2 X_2 \right|^{-\frac{1}{2}} (S^2)^{-\frac{(T-2k)}{2}} p(t) \qquad — (7)$$

Posterior probability density function of β_1

$$p(\beta_1 \mid X, Y) \propto \sum_{t=k}^{T-K} \int_0^{\infty} \int_{-\infty}^{\infty} p(\theta \mid X, Y)\, d\beta_2\, d\sigma^2$$

$$\propto \sum_{t=k}^{T-k} \frac{\left| X^1_1 X_1 \right|^{\frac{1}{2}}}{(S^2)^{K/2}} \left[1 + \frac{(\beta_1 - \hat{\beta}_1)^1 X_1^1 X_1 (\beta_1 - \hat{\beta}_1)}{S^2} \right]^{-\left(\frac{T-k}{2}\right)} p(t \mid X, Y) \qquad — (8)$$

Posterior probability density function of β_2

$$p(\beta_2 \mid X, Y) \propto \sum_{t=k}^{T-k} \int_0^{\infty} \int_{-\infty}^{\infty} p(\theta \mid X, Y)\, d\beta_1\, d\sigma^2$$

$$\alpha \sum_{t=k}^{T-k} \frac{|X_2^1 X_2|^{\frac{1}{2}}}{(S^2)^{K/2}} \left[1 + \frac{(\beta_2 - \hat{\beta}_2)^1 X_2^1 X_2 (\beta_2 - \hat{\beta}_2)}{S^2} \right]^{-\frac{(T-k)}{2}} p(t \mid X, Y) \qquad -(9)$$

and posterior probability density function of σ^2

$$p(\sigma^2 \mid X, Y) \ \alpha \sum_{t=k}^{T-k} \int_{-oo}^{oo} \int_{-oo}^{oo} p(\theta \mid X, Y)\, d\beta_1\, d\beta_2$$

$$\alpha \sum_{t=k}^{T-k} \left(\frac{S^2}{\sigma^2}\right)^{\frac{(T-2k)}{2}} \frac{e^{-\frac{S^2}{2\sigma^2}}}{\sigma^2} p(t \mid X, Y) \qquad -(10)$$

In the Baysian approach for positive definite quadratic loss function the mean of the posterior density functions, if they exist are the point estimates of the parameters. Using this result we can obtain the point estimates of the unknown parameters of the regression models (1) and (2).

First we consider the posterior probability density functions of β_1 and β_2 By comparing with multivariate student t distributions we can see that the posterior distributions of β_1 and β_2 are sum of multivariate student t distributions with mean β_1 and β_2 respectively.

Hence we obtain

$$E(\beta_1 \mid X,Y) = \sum_{t=k}^{T-k} \hat{\beta}_1\, p(t \mid X, Y) \qquad -(11)$$

and

$$E(\beta_2 \mid X,Y) = \sum_{t=k}^{T-k} \hat{\beta}_2\, p(t \mid X, Y) \qquad -(12)$$

By evaluating the integral $\int_{o}^{oo} \sigma^2 p(\sigma^2 \mid X, Y)\, d\sigma^2$ it can easily be seen that

$$E(\sigma^2 \mid X,Y) = \frac{1}{T-k-2} \sum_{t=k}^{T-k} S^2 \, p\,(t \mid X, Y) \qquad (13)$$

The mean of the posterior distribution of t can be calculated by using the following expression

$$E(t \mid X,Y) = \sum_{t=k}^{T-k} t \; p\,(t \mid X, Y) \qquad - (14)$$

Here E stands for mathematical expectation of the corresponding estimators.

Once we obtained the estimates of unknown parameters we can construct confidence intervals for these parameters.

Let g^1 be 1xk vector of constants. Then from (8) the posterior probability density function of $\alpha = g^1 \beta_1$ can easily be written as

$$p\,(\alpha \mid X, Y) \propto \sum_{t=k}^{T-k} \frac{1}{(c\,S)^{\frac{1}{2}}} \left[1 + \frac{(\alpha - \hat{\alpha})^2}{c\,S^2} \right]^{-\frac{(v+1)}{2}} p\,(t \quad X, Y) \qquad - (15)$$

where $v = T\text{-}2\,k$, $c = g^1 \left(X_1^{\,1} X_1 \right)^{-1} g$ and $\hat{\alpha} = g^1 \hat{\beta}_1$

By comparing with univariate t distribution we can see that

$$\frac{(\hat{\alpha} - \alpha)}{(c\,s^2/v)^{\frac{1}{2}}} \sim t_v$$

This fact can be utilised to make inference about α .

$$\therefore \Pr \left\{ \alpha \in \left[(e, f) \mid X, Y \right] \right\} = \sum_{t=k}^{T-k} \Pr$$

$$\left\{ t_v \in \left[t^1 (e, t), t^1 (f, t) \right] \right\} \; p\,(t \mid X, Y)$$

$$\text{Where } t^1 \; (u, t) = \frac{(u - \hat{\alpha})}{(c\, s^2/v)^{\frac{1}{2}}} \qquad \text{— (16)}$$

The confidence intervals for β_2 can also be obtained using similar procedure. Now consider the posterior probability density function of σ^2.

$$p\,(\sigma^2 \mid X, Y)\ \alpha \sum_{t=k}^{T-k} \left\{ \frac{S^2}{\sigma^2} \right\}^{v/2} \frac{e^{-S^2/2\sigma^2}}{\sigma^2}\ p\,(t \mid X, Y)$$

This expression implies that $\left(\frac{S^2}{\sigma^2}\right)$ has the "Chi-square" (X^2) distribution with v degrees of freedom.

$$p\,(X^2) = \frac{1}{\Gamma(v/_2)\, 2^{v/2}} \, (X^2)^{v/_2 - 1}\, e^{-X^2/_2} \; ; X^2 > o$$

Using this fact we can construct confidence intervals for σ^2 as follows.

$$Pr \left\{ \sigma^2 \in \left[(e, f) \mid X, Y \right] \right\} = \sum_{t=k}^{T-k} Pr \left\{ \chi^2_v \in \left(\frac{S^2}{f}, \frac{S^2}{e} \right) \right\} p\,(t \mid X, Y) \qquad \text{— (17)}$$

The expressions (11), (12), (13) and (14) can be used to get the estimates for β_1, β_2, σ^2 and t respectively and the expressions (16) and (17) can help to construct the confidence intervals for β's and σ^2.

To see the validity of the above result we need great deal of computations and hence we delete it from this paper.

REFERENCES:

1. Box G. E. P. and G. C. Tiao (1973), **Bayesian Inference in Statistical Analysis,** Addision-Wesley Publishing Company.

2. Ferreira, P. E. (1975) "A Bayesian Analysis of a Switching Regression Model: Known number of Regimes", *Journal of the American Statistical Association* 70,370–374.

3. Quandt, E. R. (1958), "The estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes". *Journal of the American Statistical Association* 53, 873–880.

4. Zelliner, A. (1971). *An Introduction to Bayesian Inference in Econometrics,* John Wiley & Sons. New York.