Volume Title: Annals of Economic and Social Measurement, Volume 2, number 4

Volume Author/Editor: Sanford V. Berg, editor

Volume Publisher:

Volume URL: http://www.nber.org/books/aesm73-4

Publication Date: October 1973

Chapter Title: The Estimation Of Structural Shifts By Switching Regressions

Chapter Author: Stephen Godfeld, Richard Quandt

Chapter URL: http://www.nber.org/chapters/c9938

Chapter pages in book: (p. 473 - 483)

Annals of Economic and Social Measurement, 2/4, 1973.

THE ESTIMATION OF STRUCTURAL SHIFTS BY SWITCHING REGRESSIONS

BY STEPHEN M. GOLDFELD AND RICHARD E. QUANDT

This paper surveys several econometric techniques for dealing with switching regressions. More general formulations, designed to produce maximum likelihood estimates, are introduced, and the problems of numerical optimization discussed. Also examined are extensions to Markov models, simultaneous equations, and switching of causal directions.

I. INTRODUCTION

In recent years, increasing attention has been devoted to problems of parameter variation in regression models. This variation has been modeled in two principal ways. The first of the approaches typically allows for an infinite number of possible parameter values and for random parameter variations. In this case the appropriate econometric technique is the random coefficients regression model or one of its particular varieties, such as the error-components model or the linear dynamic recursive model ([8], [12], [22], [32], [37]).

Alternatively, the number of possible parameter changes may be finite (usually small) where we may call each possible state of the parameter vector a regime. In time series applications these regimes may be associated with such things as the state of the business cycle or other more fundamental structural changes. In cross section work different regimes may be posited to hold for behavioral units with different characteristics (e.g., asset size, income, and whether or not rationing is imposed on the unit in a particular market). In either event the appropriate econometric technique is the switching regression model.

The switching regressions model can be formulated as follows. Assume that n observations are available on a dependent variable y and on p independent variables $x_1, \ldots, x_p$. Denote the ith observation on the x's by the vector x_i. There may be reason to believe that the observations on y were generated by two distinct regression equations or regimes: i.e.,

$$y_i = x_i'\beta_1 + u_{1i} \qquad i \in I_1 \tag{1.1}$$

and

$$y_i = x_i'\beta_2 + u_{2i} \qquad i \in I_2 \tag{1.2}$$

where i indexes observations, I_1 and I_2 are the sets of indices for which the two different regression equations hold, u_{1i} and u_{2i} are error terms (customarily but not necessarily assumed to be distributed as $N(0, \sigma_1^2)$ and $N(0, \sigma_2^2)$) and finally where the β_1, β_2 are the vectors of regression coefficients. In the most general case one would assume that $(\beta_1, \sigma_1^2) \neq (\beta_2, \sigma_2^2)$, although in particular instances one or more of the parameters may be thought to have identical values in the two regimes.[1]

The circumstance which makes the estimation of (1.1) and (1.2) nontrivial and which makes testing the null hypothesis that no switch occurred (i.e., that there is only one regime) also nontrivial is that the investigator is assumed to have no exact prior knowledge about how to classify data points with respect to the two regimes (1) and (2).[2] In the absence of such knowledge, clearly one must impose some further structure on the problem if it is to be tractable. As we shall see below this may be accomplished in a variety of ways both deterministically and in the spirit of the random coefficients model. However, before describing these methods we shall indicate some substantive applications of the switching model.

Some applications. Several recent econometric models have posited the existence of a switch in a regression equation. The manner in which the sample of observations was separated into subsamples corresponding to the two regimes varies from case to case. We describe three such models.[3]

Hamermesh [21] is concerned with estimating a wage equation according to which the negotiated annual wage change for the ith firm in the tth period, W_{it}, is a linear function of the inverse of the unemployment rate U_t and the annual percentage change in the consumer price index c_t. He assumes that a threshold effect is present and manifests itself at $c = 2.0$. Hence he posits two regimes given by

$$W_{it} = \beta_1 + \beta_2 U_t^{-1} + \beta_3 c_t + u_t \quad \text{if } c_t \le 2$$

$$W_{it} = \gamma_1 + \gamma_2 U_t^{-1} + \gamma_3 c_t + v_t \quad \text{if } c_t > 2.$$

The mechanism by which the two regimes are separated is given here a priori: in principle it would be desirable to estimate an unknown c^* such that the first regime holds when $c_t \le c^*$ and the second in the converse case.

Davis, Dempster and Wildavsky [9], [10] attempt to explain the budgetary process of U.S. government agencies. Letting x_t represent the appropriation requested by the Bureau of the Budget and y_t the appropriation passed by Congress, the simplest of their models takes the form

$$x_t = \beta y_{t-1} + u_t$$

$$y_t = \gamma x_t + v_t$$

where u_t and v_t are normally distributed errors. Because of the change in administrations over time and other possible causes of changes in decision structures they posit the possibility of two regimes, i.e.,

$$(\beta, \gamma) = (\beta_1, \gamma_1) \quad \text{if } t \le t^*$$

$$(\beta, \gamma) = (\beta_2, \gamma_2) \quad \text{if } t > t^*.$$

[1] Special constraints are imposed on the problem if it is assumed that the equations representing the two regimes intersect at some particular point. See Ando [1], Hudson [25], Hinkley [23], [24], and Gallant and Fuller [16].

[2] With such knowledge, hypothesis testing can be accomplished, at least under certain circumstances, by the Chow test [6]. The corresponding estimation problem is solved by obtaining the least squares regression from the pooled data if the Chow test produces insignificant results and by obtaining separate least squares regressions in the opposite case.

[3] Also see models by Sengupta and Tintner [33], [34], Gordon [20], and Fair and Jaffee [13].

They identify the unknown t^* by an examination of the residuals from the equations and the Chow F-statistic for varying t^*.

Silber [35] is concerned with explaining the spread between the interest rate on federal agency securities and comparable maturity Treasury securities as a function of the size of the agency issue. He posits a model of the form

$$y_t = a_1 + b_1 s_t + c x_t + u_t \qquad s_t \le s^*$$

$$y_t = a_2 + b_2 s_t + c x_t + v_t \qquad s_t > s^*$$

where y is the spread, x is a set of other variables (whose coefficients remain constant), s is the size of the issue and s^* is the critical size. Silber estimated this model by use of a variant of a technique to be described below and found strong support for the switching hypothesis.[4]

2. Theoretical Results

Several econometric approaches have been introduced to deal with switching regressions under a variety of conditions. The principal difference among conditions is whether nature's choice between the two regimes is assumed to be stochastic, i.e., depend on unknown probabilities λ and $1 - \lambda$ respectively or deterministic in the sense that it depends on the comparison of an observable variable z with an unknown threshhold or cutoff value z_0, where z may either be one of the regressors or an entirely extraneous variable. A special case of this latter mechanism is one in which the variable z is the time index of the observations.

Deterministic switching based on time index. Assume that (1.1) holds for $i \le i^*$ and (1.2) holds $i > i^*$. Quandt has proposed ([28], [31]) that the two regimes be estimated by first maximizing the likelihood conditional on i^*

$$L(y|i^*) = \left(\frac{1}{2\pi}\right)^{n/2} \sigma_1^{-i^*}\sigma_2^{-(n-i^*)} \exp\left\{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{i^*}(y_i - x_i'\beta_1)^2 - \frac{1}{2\sigma_2^2}\sum_{i=i^*+1}^{n}(y_i - x_i'\beta_2)^2\right\} \tag{2.1}$$

and then choosing as the estimate for i^* that value which maximizes the maximal likelihoods $L(y|i^*)$. For testing the null hypothesis that no switch took place a likelihood ratio test is suggested with the likelihood ratio being given by $\mu = \hat\sigma_1^{i^*}\hat\sigma_2^{(n-i^*)}/\hat\sigma^n$ where $\hat\sigma$ is the estimated standard deviation of the residuals from a single regression over the entire sample.[5]

The previous technique provides a method for both estimation and testing. There are several other techniques which just address the testing problem. Brown and Durbin [6] have introduced a test based on recursive residuals defined in the

[4] This example is one in which it seems desirable to impose a "meeting condition," i.e., that at $s_t = s^*$ the two regimes should give the same y_t. Silber did not do this but his unconstrained estimates nearly satisfied the condition. See also footnote 1.

[5] The evidence in [31] suggested some problems with this test but more recently it has been found to be of use for certain ranges of the true value of i^* [15]. We have found, and it is also reported in [5] that a Chow-test, used with caution and as if i^* were known a priori, is also satisfactory.

following way. Let $\hat{\beta}_i$ be the least squares estimate of β based on the first i observations and let X_i be the matrix having as its rows the vectors $x'_1, x'_2, \ldots, x'_i$. Then, defining

$$w_i = \frac{y_i - x'_i\hat{\beta}_{i-1}}{[1 + x'_i(X'_{i-1}X_{i-1})^{-1}x_i]^{1/2}} \qquad i = p+1, \ldots, n$$

it can be shown that under the null hypothesis of no switch, $w_i \sim N(0, \sigma^2)$. The test for shifting β is based on departures from zero of the cumulative sums

$$C_i = \frac{1}{s}\sum_{j=p+1}^{i} w_j \qquad i = p+1, \ldots, n$$

where $s^2 = \Sigma_{j=p+1}^{n} w_j^2/(n-p)$. At the 0.05 level of significance the null hypothesis is rejected if the sequence of C_i's crosses either the line connecting $(p, 0.948\sqrt{n-p})$ and $(n, 2.844\sqrt{n-p})$ or the line connecting $(p, -0.0948\sqrt{n-p})$ and $(n, -2.844\sqrt{n-p})$.

Farley and Hinich [14] and Farley, Hinich and McGuire [15] devise an alternative specification based on the assumption that the unknown switching point is equally likely to have occurred at each value of the index i. If i^* were known, the null hypothesis that the regression coefficients before and after i^* are the same could be tested by estimating the regression

$$y_i = x'_i\beta + z'_i\delta + u_i$$

where $z_i = x_i$ if $i > i^*$ and $z_i = 0$ otherwise and testing the hypothesis $\delta = 0$. Since i^* is unknown, they propose replacing z_i by the sum of all possible z_i's; hence z_i becomes ix_i. The null hypothesis then is that $\delta = 0$ in the regression

$$y_i = x'_i(\beta + i\delta) + u_i.$$

Some finite sample comparisons of this test with the likelihood ratio test proposed by Quandt and with the Chow test based on the assumption that $i^* = i/2$ are reported in [15].

Deterministic switching based on other variables. Each of the previous three procedures can be adapted to the situation in which the switching mechanism is controlled by a single variable with observations z_i, provided that there is no serial correlation of the disturbances and there are no lags present. One simply rearranges the observations in increasing (or decreasing) order of z_i and applies the previous techniques with no essential change.

A recent and more general formulation, due to Goldfeld and Quandt [19] assumes that there exist variables with observations $z_{i1}, \ldots, z_{is}$ $(i = 1, \ldots, n)$ and that nature selects between regimes 1 and 2 according to whether $\Sigma_{l=1}^{s} \pi_l z_{il} \leq 0$ or >0 where the π_l are unknown coefficients. (The simplest possible case of this type is when $s = 2$, $z_{i2} = -1$ and $\pi_1 = 1$ *a priori*. In that case the classification depends on the comparison of a single z-variable, z_{i1}, with an (unknown) cutoff point π_2 and is formally the same as the problems of Hamermesh or Silber). Letting $D_i = 0$ if $\Sigma_{l=1}^{s} \pi_l z_{il} \leq 0$ and $D_i = 1$ otherwise, the two regimes may be combined by multiplying (1.1) by $(1 - D_i)$, (1.2) by D_i and adding, which yields

$$y_i = x'_i[(1 - D_i)\beta_1 + D_i\beta_2] + (1 - D_i)u_{1i} + D_iu_{2i} \tag{2.2}$$

in which the β's, σ's and D's must be estimated. In order to render this problem tractable, D_i may be approximated by a continuous function. One approximation that has been successful is given by[6]

$$D_i = \int_{-\infty}^{\sum_{l=1}^{s} \pi_l z_{il}} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{\xi^2}{2\sigma^2}\right\} d\xi. \tag{2.3}$$

The loglikelihood function is

$$\log L = -\frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{i=1}^{n} \log\left[\sigma_1^2(1 - D_i)^2 + \sigma_2^2 D_i^2\right] - \frac{1}{2}\sum_{i=1}^{n} \frac{(y_i - x_i'[\beta_1(1 - D_i) + \beta_2 D_i])^2}{\sigma_1^2(1 - D_i)^2 + \sigma_2^2 D_i^2}. \tag{2.4}$$

Replacing D_i by (2.3) in (2.4), the likelihood function may be maximized with respect to the β's, π's and the σ introduced in (2.3) which has been interpreted to measure the goodness of the discrimination between the regimes. Unless discrimination is perfect, some of the estimated $\hat{D}_i$ will not be exactly 0 or 1. One variant of the above D-method which handles this problem is to estimate in a second stage separate regressions as in (1.1) and (1.2) where the sets I_1 and I_2 are defined by[7]

$$I_1 = \{i \mid \Sigma \hat{\pi}_l z_{il} \leq 0\}$$

$$I_2 = \{i \mid \Sigma \hat{\pi}_l z_{il} > 0\}.$$

Let the maximum of the likelihood function (the logarithm of which is (2.4)) be denoted by $L(\hat{\beta}_1, \hat{\beta}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\pi})$, and the maximum under the null hypothesis by $L(\hat{\beta}, \hat{\sigma})$. The natural likelihood ratio test statistic is

$$\mu = \frac{L(\hat{\beta}, \hat{\sigma})}{L(\hat{\beta}_1, \hat{\beta}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\pi})}$$

and $-2\log\mu$ appears in finite samples to be well approximated by the χ^2 distribution with $p + s + 2$ degrees of freedom.

Stochastic choice of regimes. On the assumption of normality of error terms the dependent variable y has the following probability density function (pdf) in the two regimes:[8]

$$f_{1i} = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2}(y_i - x_i'\beta_1)^2\right\} \tag{2.5}$$

$$f_{2i} = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2}(y_i - x_i'\beta_2)^2\right\}. \tag{2.6}$$

[6] Alternatives are the Cauchy integral $D_i = 1/2 + 1/\pi \tan^{-1}(\sum_{l=1}^{s} \pi_l z_{il})$ or the logistic $D_i = (1 + \exp(-\sum_{l=1}^{s} \pi_l z_{il}))^{-1}$ where the scale parameter corresponding to σ in (2.3) has been suppressed. See Bacon and Watts [2]. In some cases it is possible to dispense with the approximation of D_i. See, for example, Gallant and Fuller [16].

[7] An alternative in either case is not to estimate σ in the approximation (2.3) but to fix it as some small value.

[8] The reader will observe that we have replaced the variances σ_1^2, σ_2^2 by a common variance σ^2. As was pointed out to us by P. A. V. B. Swamy, this will insure the consistency of the maximum likelihood estimator. Actually, all that is necessary is that $\sigma_1^2 = k\sigma_2^2$ where k is known. For simplicity, here and in what follows, we have assumed $k = 1$.

It has been suggested by Quandt ([29], [30]) that one may think of nature choosing regimes 1 and 2 with unknown probabilities λ and $1 - \lambda$. The pdf of y_i then is[9]

$$h(y_i) = \lambda f_{1i} + (1 - \lambda) f_{2i} \tag{2.7}$$

and the appropriate loglikelihood function is

$$\log L = \sum_{i=1}^{n} \log [\lambda f_{1i} + (1 - \lambda) f_{2i}] \tag{2.8}$$

which is to be maximized with respect to the parameters of (2.5), (2.6) and λ. Tests of the null hypothesis again may employ the natural likelihood ratio.

3. Extensions of the Analysis

Simple extensions. Both the D-method and the λ-method may be extended to the case of more than two regimes. If r regimes are postulated, the pdf corresponding to (2.7) in the λ-method becomes

$$h(y_i) = \sum_{j=1}^{r} \lambda_j f_{ji}$$

with $\Sigma_{j=1}^{r} \lambda_j = 1$. For the D-method we define $r - 1$ sets of variables D_i^j ($j = 1, \ldots, r - 1$) similarly to (2.3). For convenience also define $D_i^0 = 1$ and $D_i^r = 0$. The equation representing the kth regime is then multiplied by $\prod_{j=0}^{k-1} D_i^j \prod_{j=k}^{r} (1 - D_i^j)$ and the resulting equations are added together to form a composite equation.

Another straightforward generalization is to assume that the probability λ in the λ-method is itself a function of some variable z. The resulting procedure is a hybrid between the D-method and the λ-method. The likelihood function is as before.[10]

A Markov model. It is an essential feature of the λ-method that the probability that nature selects regime 1 or 2 at the ith trial is independent of what state the system was in on the previous trial. Goldfeld and Quandt [19] recently relaxed this assumption by positing that the transitions of the system between the two states is governed by the constant transition matrix T. If $\lambda_i' = (\lambda_{1i}\ 1 - \lambda_{1i})$ denotes the vector of probabilities that regimes 1 or 2 will be chosen at the ith trial, we have

$$\lambda_i' = \lambda_{i-1}' T$$

and

$$\lambda_i' = \lambda_0' T^i.$$

It is straightforward to express the elements of λ_i' in terms of the elements of T. The loglikelihood function (2.8) is then written as

$$\log L = \sum_{i=1}^{n} \log [\lambda_{1i} f_{1i} + (1 - \lambda_{1i}) f_{2i}] \tag{3.1}$$

which needs to be maximized with respect to β_1, β_2, σ^2 and the elements of T.[11]

[9] The formulation is closely related to the question of mixture distributions. See [4], [11], [38]. It is obviously also spiritually close to the random coefficients model with only two possible values for the coefficient vector.

[10] For more detail see [17].

[11] See footnote 8.

A further extension is possible if one assumes that the elements of T are themselves functions of some extraneous variable z.[12]

Serial correlation of disturbances. None of the methods discussed so far has treated the case of error structures involving autocorrelation. In ordinary regression models it is customary to introduce autocorrelation by assuming a first-order (more rarely, a second-order) Markov process for the error term as in

$$u_t = \rho u_{t-1} + \varepsilon_t. \tag{3.2}$$

In the present case more alternatives arise, partly because of the regime-switching mechanism and partly because one may wish to approach the problem either with the D-method or the λ-method.

The first possibility is to assume that

$$\begin{aligned} u_{1t} &= \rho_1[(1 - D_{t-1})u_{1t-1} + D_{t-1}u_{2t-1}] + \varepsilon_{1t} \\ u_{2t} &= \rho_2[(1 - D_{t-1})u_{1t-1} + D_{t-1}u_{2t-1}] + \varepsilon_{2t} \end{aligned} \tag{3.3}$$

if the D-method is employed, where $\varepsilon_{1t} \sim N(0, \sigma_1^2)$ and $\varepsilon_{2t} \sim N(0, \sigma_2^2)$ and independent of each other.

The equivalent assumption for the λ-method (with $\sigma_1^2 = \sigma_2^2$) is

$$\begin{aligned} u_{1t} &= \rho_1 u_{1t-1} + \varepsilon_{1t} && \text{with probability } \lambda^2 \\ u_{1t} &= \rho_1 u_{2t-1} + \varepsilon_{1t} && \text{with probability } \lambda(1 - \lambda) \\ u_{2t} &= \rho_2 u_{1t-1} + \varepsilon_{2t} && \text{with probability } \lambda(1 - \lambda) \\ u_{2t} &= \rho_2 u_{2t-1} + \varepsilon_{2t} && \text{with probability } (1 - \lambda)^2. \end{aligned} \tag{3.4}$$

The essence of the assumption is that there are two autocorrelation coefficients each associated with one of the regimes, which are applied to the error term of the previous period, irrespective of which regime *that* error term came from. The appropriate likelihood functions can be derived but are not presented here because of their relative complexity.

An alternative specification, originally suggested to the authors by J. D. Sargan, posits that if in period t regime 1 operates and in $t - 1$ regime 1 operated as well, the error term follows the usual Markov process: if in period $t - 1$ regime 2 operated (i.e., a switch took place) then a nonautocorrelated error term is generated. Accordingly, for the D-method

$$\begin{aligned} u_{1t} &= (1 - D_{t-1})(\rho_1 u_{1t-1} + \varepsilon_{1t}) + D_{t-1}\varepsilon_{1t} = (1 - D_{t-1})\rho_1 u_{1t-1} + \varepsilon_{1t} \\ u_{2t} &= D_{t-1}(\rho_2 u_{2t-1} + \varepsilon_{2t}) + (1 - D_{t-1})\varepsilon_{2t} = D_{t-1}\rho_2 u_{2t-1} + \varepsilon_{2t} \end{aligned} \tag{3.5}$$

and for the λ-method

$$\begin{aligned} u_{1t} &= \rho_1 u_{1t-1} + \varepsilon_{1t} && \text{with probability } \lambda^2 \\ u_{1t} &= \varepsilon_{1t} && \text{with probability } \lambda(1 - \lambda) \\ u_{2t} &= \varepsilon_{2t} && \text{with probability } \lambda(1 - \lambda) \\ u_{2t} &= \rho_2 u_{2t-1} + \varepsilon_{2t} && \text{with probability } (1 - \lambda)^2. \end{aligned} \tag{3.6}$$

[12] [18] also contains several other switching models. One model allows the choice of a regime to depend on the temporal pattern of regime choices. Another allows for a hybrid transition regime between the two pure regimes. Wilton [39] has also considered a special case of this last problem.

The corresponding likelihood functions can again be derived but are also omitted here. In either formulation estimates of all parameters can be obtained by maximizing the likelihood function.[13]

Switching in simultaneous equations. A two-regime problem may be said to exist in a system of simultaneous equations if

$$B_1 y_i + \Gamma_1 z_i = u_{1i}, \qquad u_{1i} \sim N(0, \Sigma_1), \qquad i \in I_1 \tag{3.7}$$

and

$$B_2 y_i + \Gamma_2 z_i = u_{2i}, \qquad u_{2i} \sim N(0, \Sigma_2), \qquad i \in I_2 \tag{3.8}$$

where $B_1, B_2, \Gamma_1, \Gamma_2$ are the usual coefficient matrices. y_i and z_i the ith observation on the vectors of G endogenous and K exogenous variables respectively and I_1 and I_2 the index sets defined in (1.1) and (1.2). The formulation of (3.7) and (3.8) allows for various special cases such as the case in which only one equation in the system is subject to switching: in that event B_1 and B_2 are the same except for the row corresponding to the switching equation and similarly for Γ_1 and Γ_2.

Either the D-method or the λ-method (with $\Sigma_1 = \Sigma_2$) may be applied to the problem, depending on the specification of the switching mechanism as described in Section 2. In the case of the D-method we define

$$B_i = (1 - D_i)B_1 + D_i B_2$$

$$\Gamma_i = (1 - D_i)\Gamma_1 + D_i \Gamma_2$$

$$\Sigma_i = (1 - D_i)^2 \Sigma_1 + D_i^2 \Sigma_2 .$$

The joint pdf for the vector y_i then is

$$h(y_i) = (2\pi)^{-G/2} (\det \Sigma_i)^{-1/2} |\det B_i| \exp\{(-\tfrac{1}{2}(B_i y_i + \Gamma_i z_i)' \Sigma_i^{-1} (B_i y_i + \Gamma_i z_i)\} \tag{3.9}$$

from which the loglikelihood function is obtained as $\log L = \Sigma \log h(y_i)$. In the case of the λ-method we have

$$h(y_i) = \lambda h_1(y_i) + (1 - \lambda) h_2(y_i) \tag{3.10}$$

where $h_1(y_i)$ and $h_2(y_i)$ are the joint pdf's for y_i under (3.7) and (3.8) respectively. The loglikelihood is again straightforward.[14]

In the case of simultaneous equations, it is necessary to verify that in an econometric model incorporating switching between regimes the parameters are identified. It is plausible to assume that all parameters are identified separately in (3.7) and (3.8). It can then be shown that the λ-combination leaves the composite system identified. It can also be shown that the composite system is identified under the D-method if (a) all D_i equal 0 or 1 exactly, or (b) if each equation in (3.7) satisfies the same *a priori* restrictions as the corresponding equation in (3.8).

Switching of causal directions. It is interesting to consider the possibility that the difference between two regimes may consist only in which variable is dependent

[13] For a related contribution to the autocorrelation problem see Maddala and Nelson [26].

[14] Barten and Bronsard [3] have considered the application of two stage least squares when the shift points are known a priori. It is possible to combine a multivariate generalization of the technique described at the beginning of Section 2 with the Barten–Bronsard method to yield a two stage procedure with unknown shift points. This will be the subject of a forthcoming paper.

(endogenous) and which is independent (exogenous). For simplicity we shall consider the single equation case.

Let the two regimes be given by

$$y_i = a_1 + b_1 x_i + u_{1i} \qquad i \in I_1 \tag{3.11}$$

$$x_i = a_2 + b_2 y_i + u_{2i} \qquad i \in I_2 \tag{3.12}$$

where, in the first regime x_i and in the second regime y_i is treated as nonstochastic and identical in repeated samples, and where $u_{1i} \sim N(0, \sigma_1^2)$, $u_{2i} \sim N(0, \sigma_2^2)$. A case in point might be where either x or y but not both could be chosen as an exogenous policy instrument and the policy maker shifts between instruments at unknown points of time. More realistically such a problem is likely to be found in the context of a macroeconometric model of the simultaneous equations variety.

It is obvious that if either (3.11) or (3.12) were estimated on the assumption that all observations were generated by it, the estimates would not be consistent. We have explored the possibility of estimating such a model by both the D and λ techniques but we have encountered conceptual problems in each instance. Therefore, the proper method for estimating this rather interesting model remains an open question.

4. Concluding Remarks

Numerous approaches exist to the several specifications of switching regression equations. Some of these such as Quandt ([28], [31]), Brown and Durbin [6], and Farley and Hinich [14] can easily be incorporated in standard regression packages for computation. Others, namely the D- and λ-methods and their variants, are designed to produce maximum likelihood estimates and invariably involve problems of numerical optimization. These problems have been found soluble both in sampling experiments and in realistic contexts. On the basis of fairly extensive Monte Carlo experiments in single-equation models and somewhat more restricted experiments in simultaneous equation models both the D- and λ-method appear to have acceptable small sample properties. The Fair and Jaffee model of the housing market [12] was reestimated using both methods as well as the Markov generalization of the λ-method and yielded reasonable conclusions in each instance.

Princeton University

References

[1] Ando, A., "On a Problem of Aggregation," *International Economic Review*, 12 (1971), 306–311.

[2] Bacon, D. W. and D. G. Watts, "Estimating the Transition Between Two Intersecting Straight Lines," *Biometrika*, 58 (1971), 525–534.

[3] Barten, A. P. and L. S. Bronsard, "Two-Stage Least-Squares Estimation with Shifts in the Structural Form," *Econometrica*, 38 (1970), 938–941.

[4] Behboodian, J., "On the Distribution of a Symmetric Statistic From a Mixed Population," *Technometrics*, 14 (1972), 919–923.

[5] Bhattracharyya, G. K. and R. A. Johnson, "Non-parametric Tests for Shift at an Unknown Time Point," *Annals of Math. Stat.*, 39 (1968), 1731–43.

[6] Brown, R. L. and J. Durbin, "Methods of Investigating Whether a Regression Relationship is Constant Over Time," Paper presented at the European Statistical Meeting, Amsterdam, 1968.

[7] Chow, G., "Tests of the Equality Between Two Sets of Coefficients in Two Linear Regressions," *Econometrica*, 28 (1960), 561–605.
[8] Cooley, T. F. and E. C. Prescott, "An Adaptive Regression Model," *International Economic Review*, 14 (1973), 364–371.
[9] Davis, Otto A., M. A. H. Dempster, and A. Wildavsky, "A Theory of the Budgetary Process," *American Political Science Review*, LX (1966), 529–547.
[10] Davis, Otto A., M. A. H. Dempster, and A. Wildavsky, "On the Process of Budgeting II: An Empirical Study of Congressional Appropriations," in *Studies in Budgeting*, ed. by R. F. Byrne, *et al.*, North Holland Publishing, 1971.
[11] Day, N. E., "Estimating the Components of a Mixture of Normal Distributions," *Biometrika*, 56 (1969), 463–474.
[12] Duncan, D. B. and S. D. Horn, "Linear Dynamic Recursive Estimation from Viewpoint of Regression Analysis," *Journal of the American Statistical Association*, 67 (1972), 815–821.
[13] Fair, R. C., and D. M. Jaffee, "Methods of Estimation for Markets in Disequilibrium," *Econometrica*, 40 (1972), 497–514.
[14] Farley, J. U. and M. J. Hinich, "A Test for A Shifting Slope Coefficient in a Linear Model," *Journal of the American Statistical Association*, 65 (1970), 1320–1329.
[15] Farley, J. U., M. J. Hinich, and T. W. McGuire, "Testing for a Shift in The Slopes of a Multivariate Linear Time Series Model," Carnegie Mellon University, Working Paper 77-70-1, Jan. 1973.
[16] Gallant, A. R. and W. A. Fuller, "Fitting Segmented Polynomial Regression Models Whose Join Points Have to be Estimated," *Journal of the American Statistical Association*, 68 (1973), 144–147.
[17] Goldfeld, S. M., H. H. Kelejian, and R. E. Quandt, "Least Squares and Maximum Likelihood Estimation of Switching Regressions," Econometric Research Program, Research Memorandum No. 130, Princeton University, November 1971.
[18] Goldfeld, S. M. and R. E. Quandt, "A Markov Model for Switching Regressions," *Journal of Econometrics*, 1 (1973), 3–16.
[19] Goldfeld, S. M. and R. E. Quandt, *Nonlinear Methods in Econometrics*, North Holland Press, 1972, ch. 9.
[20] Gordon, R. J., "Wage Price Controls and the Shifting Phillips Curve," *Brookings Papers on Economic Activity* (1972, #2), pp. 385–430.
[21] Hamermesh, D. S., "Wage Bargains, Threshold Effects and the Phillips Curve," *Quarterly Journal of Economics*, LXXXIV (1970), 501–517.
[22] Hildreth, C. and J. P. Houck, "Some Estimators for a Linear Model with Random Coefficients," *Journal of the American Statistical Association*, 63 (1968), 584–595.
[23] Hinkley, D. V., "Inference About The Intersection in Two-Phase Regression," *Biometrika*, 56 (1969), 495–504.
[24] Hinkley, D. V., "Inference in Two-Phase Regression," *Journal of the American Statistical Association*, 66 (1971), 736–743.
[25] Hudson, D. J., "Fitting Segmented Curves Whose Join Points Have to be Estimated," *Journal of the American Statistical Association*, 61 (1966), 1097–1129.
[26] Maddala, G. S. and F. D. Nelson, "Limited Dependent Variable Methods for the Estimation of Markets in Disequilibrium," Department of Economics, University of Rochester, Discussion Paper 73-7, 1973.
[27] McGee, V. E. and W. T. Carleton, "Piecewise Regression," *Journal of the American Statistical Association*, 65 (1970), 1109–1124.
[28] Quandt, R. E., "The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes," *Journal of the American Statistical Association*, 53 (1958), 873–880.
[29] Quandt, R. E., "A Further Approach to the Estimation of Switching Regressions," Econometric Research Program Research Memorandum No. 122, Princeton University, March 1971.
[30] Quandt, R. E., "A New Approach to Estimating Switching Regressions," *Journal of the American Statistical Assocation*, 67 (1972), 306–310.
[31] Quandt, R. E., "Tests of the Hypothesis That a Linear Regression System Obeys Two Separate Regimes," *Journal of the American Statistical Association*, 55 (1960), 324–330.
[32] Rosenberg, B., "Estimation of Stationary Stochastic Regression Parameters Reexamined," *Journal of the American Statistical Association*, 67 (1972), 650–654.
[33] Sengupta, J. K. and G. Tintner, "An Approach to a Stochastic Theory of Economic Development with Applications," *Problems of Economic Dynamics and Planning*, Essays in Honour of Michael Kalecki, PWN Polish Scientific Publishers, 1966.
[34] Sengupta, J. K. and G. Tintner, "On Some Aspects of Trend in the Aggregative Models of Economic Growth," *Kyklos*, 16 (1963), 47–61.

[35] Silber, W. L., "The Market for Federal Agency Securities: Is There an Optimum Size of Issue?", mimeo.
[36] Sprent, P., "Some Hypotheses Concerning Two-Phase Regression Lines," *Biometrics*, 17 (1961), 634–45.
[37] Swamy, P. A. V. B., "Efficient Inference in a Random Coefficients Regression Model," *Econometrica*, 38 (1970), 311–323.
[38] Thomas, E. A. C., "Distribution Free Tests for Mixed Probability Distributions," *Biometrika*, 56 (1969), 475–484.
[39] Wilton, D. A., "Structural Shift with an Inter-structural Transition Function," Institute for Economic Research, Queens University, Discussion Paper #92, Oct. 1972.