ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 131, 1993

*Jan B. Gajda**

STRUCTURAL CHANGE IN REGRESSION PARAMETERS AND INTERACTIVE VARIABLES

1. INTRODUCTION

One of the strongest assumptions of econometric analysis is the assumption of the existence of structural relation with parameters that are *constant* (in time, for all groups of individuals, for all observations etc.). The opposite assumption would allow them to *vary freely* from observation to observation *without any specific pattern*.

The former, in spite of its restrictivity, has a deep pragmatic sense. It organizes the *way of thinking* and *research*. The *constant parameter* model explains the behavior of a variable (y) by the behavior of other variables (x-es). A parameter b_k, associated with explanatory variable x_k ($k = 1, \ldots, K$) characterizes the *strength* of the influence of x_k on y. Thus we have clear-cut differentiation between *causes* (explanatory variables) of the changes in y and *measures* (parameters) of the strength of their influence. The assumption of constant parameters, as a crude approximation of the reality, should rather be interpreted that the parameters change, but the changes are *tolerably* small, where the tolerance is determined by the purpose of the research (or by the outcome of tests for parameter constancy).

* Lecturer at the Institute of Econometrics and Statistics, University of Łódź.

The latter assumption may be closer to reality, however, it allows for everything. From pragmatic point of view its total flexibility seems to lead to nowhere. One cannot operatively investigate parameter changes without some idea about their character.

The intermediate versions of *flexible parameter* assumption allow for changes in structural parameters accordingly to *some specific pattern*. Changes may be deterministic or stochastic, continuous or discrete, with known or unknown breaking point, may cover some or all structural parameters. The pattern of changes is to be specified *a priori* by the researcher. This usually leads to some reparameterization of the original model in terms of new, constant hyperparameters.

In the *panta rei* world, where everything is in continuous movement, there exists *certain* risk associated with the *flexible parameter* assumption. Namely, too much flexibility in parameters may explain too much, leaving nothing to be explained by the model *variables*.

> *Modeling a changing economic structure by allowing response parameters to vary over observations may be a realistic approach, but the chances for misspecifications are many* (J u d g e et al. 1980, p. 398).

One may thus suggest a *pragmatic principle* that allows for the varying parameter assumption, provided that the *knowledge a priori* justifies such approach (i.e. we *know* or have good reasons to expect that the parameter changes are outside the tolerance limits of the constant parameter model).

The alternative approach, based much more on the sample information, calls for testing the parameter constancy and/or testing for the period(s) of structural change (if discrete) or the functional form (the "transition law" for the changing parameters, if continuous).

The changing pattern (instability) of parameters of the model under test need not be an inherent property of the modeled system. It may as well result from model misspecification - for example inadequate functional form or omitting an important variable (D z i e c h c i a r z 1989 discusses the problem in wider context). Testing for parameter constancy may thus lead to testing for the misspecification rather, than to a model with changing structural parameters. Testing stays outside the scope of this paper. The works H a c k l (ed.) (1989) and K r ä m e r, S o n n b e r g e r (1986) provide extensive material on this topic, H a c k l, W e s t l u n d (1985) present also an excellent bibliography.

2. MODEL WITH STRUCTURAL CHANGE IN INTERCEPT

We consider the simplest case of structural change in linear model - namely the estimation of discrete changes in intercept only (illustrated at Figure 3). Furthermore - we allow for multiple changes of intercept i.e. for more than two regimes.

The standard linear model

$$y_t = b_0 + \sum_{k=1}^{K} b_k x_{tk} + e_t, \qquad t = 1, \ldots, T \tag{2.1}$$

we rewrite for our purpose as follows:

$$y_t = \sum_{m=1}^{M} b_{0m} \cdot u_{tm} + \sum_{k=1}^{K} b_k x_{tk} + e_t, \qquad t = 1, \ldots, T \tag{2.2}$$

where: m - index of the regime, M - total number of regimes, u_{tm} - dummy variable associated with the m-th regime, assuming value 1 for this regime, zero otherwise[1].

The model can be rewritten in the matrix form as follows:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{y}_M \end{bmatrix} = \left[\begin{array}{cccc|c} \mathbf{u} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{X}_1 \\ \mathbf{0} & \mathbf{u} & \ldots & \mathbf{0} & \mathbf{X}_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{u} & \mathbf{X}_M \end{array}\right] * \begin{bmatrix} \mathbf{b}_{01} \\ \mathbf{b}_{02} \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{b}_{0M} \\ \hline \mathbf{b} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{e}_M \end{bmatrix} \tag{2.3}$$

where:

$\mathbf{e}$ - vector with zero mean and a scalar covariance matrix $\sigma^2\mathbf{I}$;

---

[1] To simplify our formulas we assume equal number (Q) of observations associated with each regime. Hencefore $T = Q \times M$. The results stay valid for the nonequal case, as well, provided that the smallest subsample contains enough observations (roughly speaking not less than the number of parameters that supposedly undergo the structural change in given regime).

X_m - Q by K matrix of values of nonstochastic explanatory variables in m-th regime;

y_m, e_m - Q by 1 vectors of values of dependent variable and disturbances, respectively, in m-th regime;

u - Q by 1 vector of ones;

b - K by 1 vector of slope parameters, the same for all regimes;

$\mathbf{b}_0$ - M by 1 vector of varying intercepts, its m-th value b_{0m} represents the intercept in the regime m.

Using the Kronecker product, denoted by $\otimes$, we may rewrite (2.3) as follows:

$$\mathbf{y} = [\mathbf{I} \otimes \mathbf{u} \mid \mathbf{X}] * [\frac{\mathbf{b}_0}{\mathbf{b}}] + \mathbf{e} \tag{2.4}$$

where I denotes the unit matrix.

The least squares estimator of the parameters of (2.4) is given by the following formula:

$$\text{OLS est } \{[\frac{\mathbf{b}_0}{\mathbf{b}}]\} = \begin{bmatrix} (\mathbf{I} \otimes \mathbf{u})'(\mathbf{I} \otimes \mathbf{u}) * (\mathbf{I} \otimes \mathbf{u})'\mathbf{X} \\ \mathbf{X}'(\mathbf{I} \otimes \mathbf{u}) \qquad \mathbf{X}'\mathbf{X} \end{bmatrix}^{-1} \begin{bmatrix} (\mathbf{I} \otimes \mathbf{u})'\mathbf{y} \\ \mathbf{X}'\mathbf{y} \end{bmatrix} \tag{2.5}$$

Direct calculations show, that $(\mathbf{A} \otimes \mathbf{B}) * (\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$ (see T h e i l 1971, section 7.2), thus we have:

$$\text{OLS est } \{[\frac{\mathbf{b}_0}{\mathbf{b}}]\} = \begin{bmatrix} \mathbf{I} \otimes (\mathbf{u}'\mathbf{u}) & (\mathbf{I} \otimes \mathbf{u})'\mathbf{X} \\ \mathbf{X}'(\mathbf{I} \otimes \mathbf{u}) & \mathbf{X}'\mathbf{X} \end{bmatrix}^{-1} \begin{bmatrix} (\mathbf{I} \otimes \mathbf{u})'\mathbf{y} \\ \mathbf{X}'\mathbf{y} \end{bmatrix} \tag{2.6}$$

and finally:

$$\text{OLS est } \{[\frac{\mathbf{b}_0}{\mathbf{b}}]\} = \begin{bmatrix} \mathbf{I} \times \mathbf{Q} & (\mathbf{I} \otimes \mathbf{u})'\mathbf{X} \\ \mathbf{X}'(\mathbf{I} \otimes \mathbf{u}) & \mathbf{X}'\mathbf{X} \end{bmatrix}^{-1} \begin{bmatrix} (\mathbf{I} \otimes \mathbf{u})'\mathbf{y} \\ \mathbf{X}'\mathbf{y} \end{bmatrix} \tag{2.7}$$

since $\mathbf{u}'\mathbf{u} = Q$. Under the standard assumptions it is the best linear unbiased estimator (BLU) of $\mathbf{b}_0$ and **b** in (2.4).

Let us introduce a matrix $\mathbf{A}_{Q\times Q}$ defined as follows:

$$\mathbf{A} = \mathbf{I} - \frac{\mathbf{u}\,\mathbf{u}'}{Q} \tag{2.8}$$

For an observation matrix $\mathbf{Z}_{Q\times K}$ we have:

$$\mathbf{AZ} = \mathbf{Z} - \mathbf{u}(1/Q\ \mathbf{u}'\mathbf{Z}) = (z_{ij} - 1/Q \sum_{i=1}^{Q} z_{ij}), \quad i = 1, \dots, Q, \quad j = 1, \dots, K \tag{2.9}$$

i.e. A transforms observations in columns of **Z** into their deviations from the column mean.

We shall transform the nondummy variables of our model into deviations:

$$\mathbf{W} = \begin{bmatrix} \mathbf{A}'\mathbf{X}_1 \\ \mathbf{A}'\mathbf{X}_2 \\ \vdots \\ \mathbf{A}'\mathbf{X}_M \end{bmatrix} = \begin{bmatrix} (x_{ij,1} - \bar{x}_{j,1}) \\ (x_{ij,2} - \bar{x}_{j,2}) \\ \vdots \\ (x_{ij,M} - \bar{x}_{j,M}) \end{bmatrix} \begin{matrix} \leftarrow & i = 1, \ldots, Q \\ \leftarrow & i = Q + 1, \ldots, 2Q \\ & \vdots \\ \leftarrow & i = (M-1)Q + 1, \ldots, T \\ & j = 1, \ldots, K. \end{matrix} \quad (2.10)$$

(similarly $\mathbf{v} = \begin{bmatrix} \mathbf{A}'\mathbf{y}_1 \\ \mathbf{A}'\mathbf{y}_2 \\ \vdots \\ \mathbf{A}'\mathbf{y}_M \end{bmatrix}$),

from their *regime means* $\bar{x}_{j,m}$. Using the formula for the inverse of a block-diagonal matrix[2]:

$$\begin{bmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{E}^{-1}(\mathbf{I} + \mathbf{F}\mathbf{D}^{-1}\mathbf{G}\mathbf{E}^{-1}) & -\mathbf{E}^{-1}\mathbf{F}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{G}\mathbf{E}^{-1} & \mathbf{D}^{-1} \end{bmatrix}$$

where $\mathbf{D} = \mathbf{H} - \mathbf{G}\mathbf{E}^{-1}\mathbf{F}$ (provided **E** and **D** are nonsingular), after necessary rearrangement, we may write (2.7) as follows:

$$\hat{\mathbf{b}} = [\mathbf{X}'(\mathbf{I} \otimes \mathbf{A})\mathbf{X}]^{-1}\mathbf{X}'(\mathbf{I} \otimes \mathbf{A})\mathbf{y} = [\mathbf{X}'(\mathbf{I} \otimes \mathbf{A})'(\mathbf{I} \otimes \mathbf{A})\mathbf{X}]^{-1} \mathbf{X}'(\mathbf{I} \otimes \mathbf{A})'(\mathbf{I} \otimes \mathbf{A})\mathbf{y} \quad (2.11)$$

or

$$\hat{\mathbf{b}} = (\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{v} \quad (2.12)$$

for the slopes[3] and

$$\hat{\mathbf{b}}_{0m} = \bar{\mathbf{y}}_m - \bar{\mathbf{x}}'_{,m}\hat{\mathbf{b}}, \quad m = 1, \ldots, M \quad (2.13)$$

for the intercepts,

---

[2] Compare J u d g e et al. (1982), p. 480, G o l d b e r g e r (1963), p. 27.

[3] Computer estimates of residual variance based on (2.11) will be $\hat{\mathbf{e}}'\hat{\mathbf{e}}/(MT-K)$ rather, than appropriate, resulting from the direct application of (2.7) $\hat{\mathbf{e}}'\hat{\mathbf{e}}/(MT - (M + K))$. It is thus advisable to recalculate it, along with the proper adjustment of such expressions as standard errors of estimation, interval estimates, t-statistics and so on.

where:

$\bar{y}_m$ - the mean of the y in the m-th regime;

$\bar{x}_{,m}$ - K by 1 vector of the means of x-es in the m-th regime, with elements $\bar{x}_{j,m}$, $j = 1, \ldots, K$.

This is a generalization of the classical result that the least squares estimate of slopes $\hat{b}$ depend *solely* on the variability (deviations from mean) of the model variables, while the intercept(s) assume value(s), that makes the mean residual equal zero[4]. Consistent estimation of, changing from regime to regime, intercepts of (2.3) requires consistent estimation of slopes. As we have shown - BLU estimators of the latter (contained in $\hat{\mathbf{b}}$) depend *solely* on the variability of y and x-es calculated for *each regime separately* (variability around respective *regime means),* while the changing intercept assume value making the mean residual equal *zero within each regime*[5].

If one calculates the estimate (2.12) on the basis of matrices X and y expressed as deviations from the *global* column means $\bar{x}_j = 1/T \, \Sigma_{i=1}^{T} x_{ij}$ (as opposed to the formula (2.10) where deviations are calculated from the *regime means* $\bar{x}_{j,m}$), one gets[6]:

$$\tilde{\mathbf{W}} = \mathbf{A}_T' \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{X}_M \end{bmatrix} = \begin{bmatrix} (x_{ij,1} - \bar{x}_j) \\ (x_{ij,2} - \bar{x}_j) \\ \cdot \\ \cdot \\ \cdot \\ (x_{ij,M} - \bar{x}_j) \end{bmatrix} =$$

[4] See for example G o l d b e r g e r (1963), p. 182-184. If T is sufficiently large - one can estimate separate regressions for each regime (the approach similar to that of Chow in his test for parameter stability, see C h o w 1960). Observe that in this approach the mean residual equals zero in *each regime separately*. The trouble with the separate regressions is the lack of guarantee, that the estimates of coefficients that *do not* undergo the structural change will assume the same values in different regimes.

[5] The result is closely related to the question of intra-class and inter-class variability in the analysis of variance, see for example K l e i n (1972), p. 109-116, G o l d b e r g e r (1963), p. 227-231.

[6] The symbol $\mathbf{A}_T$ reminds that this $\mathbf{A}$ matrix is of order T rather than Q as in (2.10).

$$= \begin{bmatrix} (x_{ij,1} - \bar{x}_{j,1}) \\ (x_{ij,2} - \bar{x}_{j,2}) \\ \vdots \\ (x_{ij,M} - \bar{x}_{j,M}) \end{bmatrix} + \begin{bmatrix} (\bar{x}_{j,1} - \bar{x}_{j}) \\ (\bar{x}_{j,2} - \bar{x}_{j}) \\ \vdots \\ (\bar{x}_{j,M} - \bar{x}_{j}) \end{bmatrix} =$$

$$= \begin{bmatrix} (x_{ij,1} - \bar{x}_{j,1}) \\ (x_{ij,2} - \bar{x}_{j,2}) \\ \vdots \\ (x_{ij,M} - \bar{x}_{j,M}) \end{bmatrix} + \begin{bmatrix} u\,(\bar{x}_{1} - \bar{x})' \\ u\,(\bar{x}_{2} - \bar{x})' \\ \vdots \\ u\,(\bar{x}_{M} - \bar{x})' \end{bmatrix} = \mathbf{W} + \mathbf{R} \qquad (2.14)$$

where: $\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_K \end{bmatrix}$ - vector of global means, and

$\bar{\mathbf{x}}_{,m} = \begin{bmatrix} \bar{x}_{1,m} \\ \bar{x}_{2,m} \\ \vdots \\ \bar{x}_{K,m} \end{bmatrix}$ - vector of means calculated for the m-th regime;

R is a matrix conformable with **W**, in which the observations of **W** are replaced by the respective difference between the global mean and the regime means.

Now the OLS estimator (2.15) uses $\tilde{\mathbf{W}}$ and $\tilde{\mathbf{v}}$ (defined analogously to $\tilde{\mathbf{W}}$):

$$\tilde{\mathbf{b}} = (\tilde{\mathbf{W}}'\tilde{\mathbf{W}})^{-1}\tilde{\mathbf{W}}'\tilde{\mathbf{v}} = (\mathbf{W}'\mathbf{W} + \mathbf{R}'\mathbf{R})^{-1}(\mathbf{W}'\mathbf{y} + \mathbf{R}'\mathbf{y}) =$$
$$= (\mathbf{W}'\mathbf{W} + \mathbf{R}'\mathbf{R})^{-1}\mathbf{W}'\mathbf{y} + (\mathbf{W}'\mathbf{W} + \mathbf{R}'\mathbf{R})^{-1}\mathbf{R}'\mathbf{y} \qquad (2.15)$$

For a nonsingular matrix **A** and a matrix $\mathbf{B} = \mathbf{R}'\mathbf{R}$ we may write: $(\mathbf{A} + \mathbf{B})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{R}'(\mathbf{I} + \mathbf{R}\mathbf{A}^{-1}\mathbf{R}')^{-1}\mathbf{R}\mathbf{A}^{-1}$ (see R a o (1973)).

Thus we have:

$$\begin{aligned}\tilde{\mathbf{b}} &= [(\mathbf{W}'\mathbf{W})^{-1} - (\mathbf{W}'\mathbf{W})^{-1}\mathbf{R}'[\mathbf{I} + \mathbf{R}\mathbf{A}^{-1}\mathbf{R}']^{-1}\,\mathbf{R}(\mathbf{W}'\mathbf{W})^{-1}]\,\mathbf{W}'\mathbf{y} + \\ &+ [(\mathbf{W}'\mathbf{W})^{-1} - (\mathbf{W}'\mathbf{W})^{-1}\mathbf{R}'[\mathbf{I} + \mathbf{R}\mathbf{A}^{-1}\mathbf{R}']^{-1}\,\mathbf{R}(\mathbf{W}'\mathbf{W})^{-1}]\,\mathbf{R}'\mathbf{y} = \\ &= (\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{y} - (\mathbf{W}'\mathbf{W})^{-1}\mathbf{R}'[\mathbf{I} + \mathbf{R}\mathbf{A}^{-1}\mathbf{R}']^{-1}\,\mathbf{R}(\mathbf{W}'\mathbf{W})^{-1}\,\mathbf{W}'\mathbf{y} + \\ &+ [(\mathbf{W}'\mathbf{W})^{-1} - (\mathbf{W}'\mathbf{W})^{-1}\mathbf{R}'[\mathbf{I} + \mathbf{R}\mathbf{A}^{-1}\mathbf{R}']^{-1}\,\mathbf{R}(\mathbf{W}'\mathbf{W})^{-1}]\,\mathbf{R}'\mathbf{y} = \\ &= \hat{\mathbf{b}} - \{(\mathbf{W}'\mathbf{W})^{-1}\mathbf{R}'[\mathbf{I} + \mathbf{R}\mathbf{A}^{-1}\mathbf{R}']^{-1}\mathbf{R}(\mathbf{W}'\mathbf{W})^{-1}\,\mathbf{W}'\mathbf{y} + \\ &+ [(\mathbf{W}'\mathbf{W})^{-1} - (\mathbf{W}'\mathbf{W})^{-1}\mathbf{R}'[\mathbf{I} + \mathbf{R}\mathbf{A}^{-1}\mathbf{R}']^{-1}\,\mathbf{R}(\mathbf{W}'\mathbf{W})^{-1}]\,\mathbf{R}'\mathbf{y}\} \qquad (2.16)\end{aligned}$$

We have proven, that the OLS estimator $\tilde{\mathbf{b}}$ based on *global* means differs from the OLS estimator $\hat{\mathbf{b}}$ based on the *regime* means (let us remind - the BLU one) by quite complicated expression. Thus, in general, the former is biased, even asymptotically, i.e. inconsistent; it depends also on the *inter-regime* variability of the regime means of variables. On the other hand, $\hat{\mathbf{b}}$ does depend solely on the *intra-regime* variability while the *inter-regime* differences in means are reflected in changing intercepts. Only in very special cases $\hat{\mathbf{b}}$ may be consistent.

We may summarize the results obtained insofar:

(i) the OLS estimator $\hat{\mathbf{b}}$ (2.5) is best linear unbiased (BLU);

(ii) the OLS estimator (2.12) is identical with the OLS estimator (2.5), and thus is the best linear unbiased (BLU);

(iii) one gets the same BLU estimator $\hat{\mathbf{b}}$, if one introduces relevant dummy variable for each regime (compare 2.7 and 2.12) into model, without transforming the other variables into deviations;

(iv) if the nondummy variables of the model are expressed as deviations from *global* (to be precise - *other than regime* means), the OLS estimator (2.15) of slopes is, in general, biased and inconsistent;

(v) thus if one *does not* introduce the dummy variable for *each* regime, one obtains biased and inconsistent OLS estimator of the type (2.16), dependent, inter alia, on the *intra-regime* variability of the *regime* means around the *global* means.

3. INTERACTIVE VARIABLES AND MODEL WITH OVERLAPPING REGIMES OF STRUCTURAL CHANGE

In the following we shall assume, that the structural change manifests itself in changes of slope (i.e. there exists a change in the response of y with respect to the unit change of some x) in known time period.

If the change takes place with respect to slopes of all (non-dummy) explanatory variables in the same periods - one may accordingly divide the sample into as many subsamples as there are regimes and write the model as a direct continuation of the original model with changing intercept (2.3):

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_M \end{bmatrix} = \left[\begin{array}{cccc|cccc} u & 0 & \dots & 0 & X_1 & 0 & \dots & 0 \\ 0 & u & \dots & 0 & 0 & X_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & & & \cdot \\ 0 & 0 & \dots & u & 0 & 0 & \dots & X_M \end{array}\right] * \begin{bmatrix} b_{01} \\ b_{02} \\ \cdot \\ \cdot \\ \cdot \\ b_{0M} \\ \hline b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_M \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_M \end{bmatrix} \quad (3.1)$$

However, the size of the vector b increases considerably, accompanied by a rapid decrease of degrees of freedom.

In a more general and realistic approach changes in different slopes happen in different time periods, producing the model with *overlapping regimes* (for different x-es regimes change in different periods). One can still apply the model (3.1). However the assumption, that the regime for the *whole* model changes whenever *at least one* parameter undergoes structural change, decreases degrees of freedom at a dramatic rate. There is also no guarantee, that the estimates of the parameters, that *stay unchanged in several regimes* will assume tne same values in all these regimes (some estimation with constraints on parameters may be necessary).

The model shown below seem to be more appropriate for this case

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_M \end{bmatrix} = \left[\begin{array}{cccc|cccc|c} u & u & \dots & u & z_1 & z_1 & \dots & z_1 & X_1 \\ 0 & u & \dots & u & 0 & z_2 & \dots & z_2 & X_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & & & z_{M-1} & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & & & & \cdot \\ 0 & 0 & \dots & u & 0 & 0 & \dots & 0 & X_M \end{array}\right] \times \begin{bmatrix} b_{01} \\ b_{02} \\ \cdot \\ \cdot \\ \cdot \\ b_{0M} \\ \hline \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_{M-1} \\ \hline b \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_M \end{bmatrix} \quad (3.2)$$

or:

$$\tilde{y} = [\tilde{U} : \tilde{Z} : \tilde{X}][\tilde{b}] + [\tilde{e}] \quad (3.3)$$

where $\begin{bmatrix} z_1 \\ z_2 \\ \cdot \\ \cdot \\ \cdot \\ z_m \\ 0 \end{bmatrix}$ denotes a submatrix of the matrix $\begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ \cdot \\ X_M \end{bmatrix}$

containing *only* those variables (columns) of the latter which have parameters, that shift in the regime m. The dummies, indicating the observations covered by the m-th regime[7], are collected in the first submatrix of the right hand side of (3.2)[8]. We shall write it down analogously to the formula (2.2)

$$y_t = b_{0M} + \sum_{m=1}^{M} \beta_{0,m} \cdot u_{t,m} + \sum_{m=1}^{M} \sum_{k \in N_m} \beta_{k,m} \cdot (u_{t,m} \cdot x_{tk,m}) +$$

$$+ \sum_{k=1}^{K} b_k \cdot x_{tk} + e_t, \qquad t = 1, \dots, T \quad (3.4)$$

where: N_m - set of indices of those variables, whose coefficients undergo structural change in the regime m, the product

[7] For sake of simplicity we assume in (3.2) that all regimes start in the first period. This assumption can easily be relaxed.

[8] Thus one can first select the columns (variables) of X associated with parameters that are expected to change in the regime m, and then multiply them observation after observation, by the respective dummy of the regime m to get the submatrix Z_m as a result of interaction between dummy and the chosen columns of X.

$(u_{tm} \cdot x_{tk,m})$ defines the *interactive variables* that allow for the change in the slopes.

The interpretation is similar to that of alternative reparameterization of the dummy variables in the model with changing intercepts. The element $\beta_{k,m}$ of the vector β_m measures the difference between the slope, associated with the variable x_k in the basic regime m_0 (in our case $m_0 = M$) and in the regime m, while β_{0m} accounts for the induced intercept change (to be interpreted in the same way as the difference of two intercepts). If for some variable x_k several regimes overlap

$$u_{t,m_1} = u_{t,m_2} = u_{t,m_3} = 1, \quad \text{for} \quad t \in (t_1, t_2),$$

then in the overlapping subperiod (t_1, t_2) coefficient associated with x_k differs from the base one by $\beta_{k,m_1} + \beta_{k,m_2} + \beta_{k,m_3}$.

Comparison of (3.1) and (3.2) reveals the main gain of the latter specification. In the former one has to use M dummies and $(M - 1) \times K$ interactive variables, while in the latter we use the same dummies, but only $\sum_{m=1}^{M} k_m$ interactive variables (where k_m is the number of indices in N_m, i.e. the number of structural changes undergoing in the given regime. This approach implicitly imposes the mentioned earlier equality constraints on the parameters that stay stable in several regimes.

The OLS estimator of (3.4) - as a special case of the BLUE (2.3) it is BLUE itself.

Along with the case:

(i) $\tilde{U}$ contains some dummy variables of binary type, associated with different regimes; 1 indicates that the related observation belongs to some regime, 0 otherwise (one gets multiregime model with parameters changing in a discrete manner),

- the one discussed above, we may distinguish further important cases:

(ii) $\tilde{U}$ contains a special dummy variable - namely the time trend (or some function of it, say, quadratic one $\gamma_0 + \gamma_1 \times t + \gamma_2 \times t^2$), one gets time-varying parameter model with structural changes *continuously* shifting in time.

(iii) $\tilde{U}$ contains some variables from X (or - generally speaking - some other variable or function - in the simplest case powers of some x-es) one gets the model with parameters varying along with these variables or functions, the structural changes have *continuous* character implying that y depends on x-es in a continuous, nonlinear fashion.

Combination of the above cases leads us to very flexible model, in which the interactive variables allow for discrete, continuous and/or time-dependent structural change.

4. SHIFTS IN INTERCEPTS INDUCED BY SHIFTS IN SLOPES - A GRAPHICAL PRESENTATION

We shall devote a section to a graphical illustration of a phenomenon associated with the case when one of the factors forming interactive variables is a binary (zero-one) variable[9].

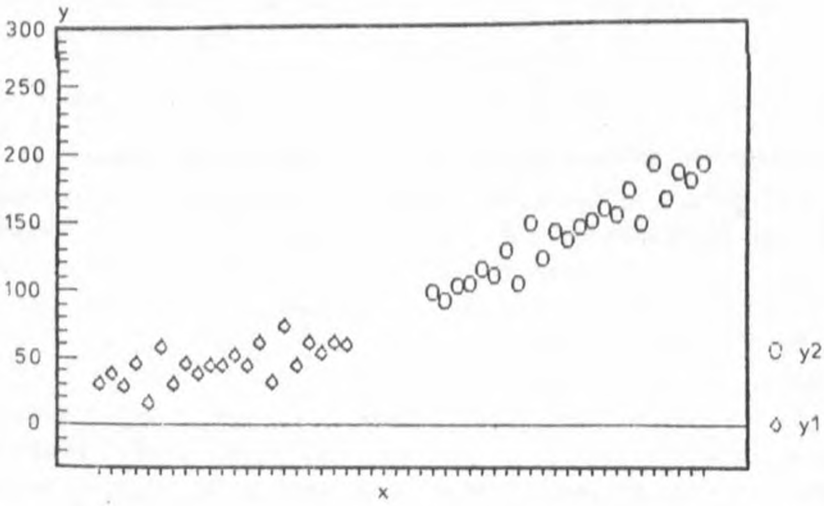

Fig. 1. Observations generated by a relationship working in two regimes

---

[9] In this paragraph we follow some earlier work of G a j d a (1985). The general formulas for the analysis of misspecification bias apply here: for the model $y = X_1\beta_1 + X_2\beta_2 + e$, OLS est $(\beta_1) = b_1 = (X_1'X_1)^{-1}X_1'y$ with bias $\beta_1 - Eb_1 = (X_1'X_1)^{-1}X_1'X_2\beta_2$. However the *implicit* and *explicit* intercept changes are harder to demonstrate within such framework.

Figure 1 illustrates one of the most frequently encountered cases of discrete structural change (with known regime changing point) - the case when the slope b_k associated with some explanatory variable x_k (i.e. the response parameter) changes. To capture the phenomenon the following function is frequently chosen:

$$y = b_{oo} + b_{11} \times (x\, u_1) + b_{12} \times (x\, u_2) + e \qquad (4.1)$$

where u_1 is a binary dummy, having 1 in the first regime, u_2 - analogously defined dummy for the second regime.

> *A shift in the value of the coefficient of a predetermined variable does not cause special problems. If, say, only one shift occurs, one defines two predetermined variables to replace the original one. The vector of observations for the first of these consists of the observations on the original variable, with the exception of those observations for which the second value is supposed to hold. These latter observations are replaced by zero's. The vector of observations on the second variable is simply the difference between the vector of observations for the original variable and the one for the first variable* state Barten and Bronsard 1970.

One may use an alternative parameterization of (4.1):

$$y = b_{oo} + b_1 \times x + \Delta b_1 \times (x\, u_2) + e \qquad (4.2)$$

where b_1 represents the slope for the first regime, and Δb_1 represents the *change* in the slope after the first regime ends. One finds b_{12} of (4.1) as $b_1 + \Delta b_1$.

Figure 2 shows the same data with the two true regression lines - y1 and y2 in respective regimes. Comparing lines y1 and y2 one clearly sees, that the change in the slope induces *implicit* change Δ▯ of intercept, the researcher may be unaware of. If one overlooks that fact (as Barten and Bronsard did) one estimates the model represented on the Figure 2 by the (solid) line *fitted*, obtaining (sometimes severely) biased estimators, since the above formulas incorrectly assume, that lines of both regimes have the same intercept. The case illustrated on Figures 1 and 2 stresses, that whenever there exist the structural break in slope *and the regression line is assumed to be continuous*[10] in the sample, a change in

---

[10] Compare the problem of *joint points* in Judge et al. (1980), p. 387-388.

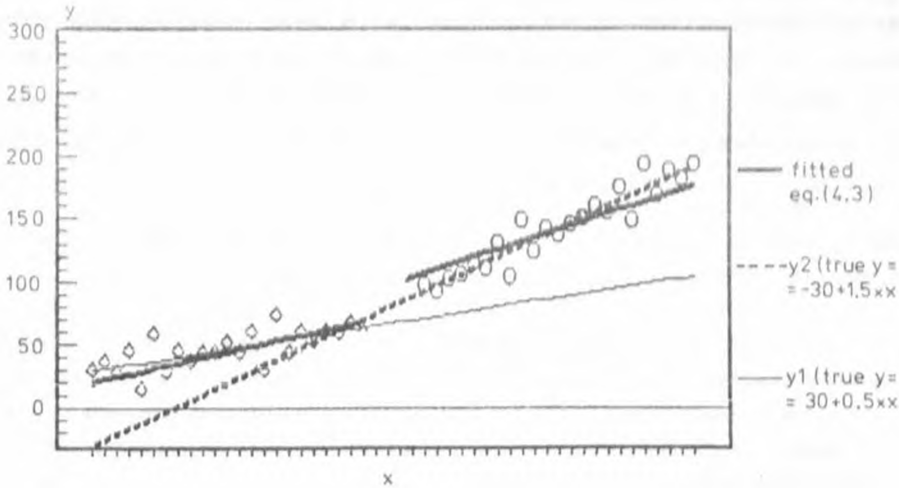

Fig. 2. Observations of Figure 1 with true regimes and fitted values of model with single interactive variable (no dummy)

the intercept is to be expected, as well. We call it the *implicit* intercept change, as it is *implied* by the change in the slope.

The data generating mechanism for our charts is as follows:

$$y_t = \begin{cases} 30 + 0.5x_t + e_t & \text{for } t = 1, \ldots, 24 \\ -30 + 1.5x_t + e_t & \text{for } t = 25, \ldots, 50 \end{cases}$$

$$e_t \sim N(0,15), \quad x_t = (t - 1) \times 3.$$

The estimation of variant (4.1) i.e. with intercept of both segments forced to be the same, gives the following result:

$$\hat{y} = \underset{(4.1)}{20.82} + \underset{(5.5)}{0.71} \times x + \underset{(3.5)}{0.34} \times (x\, u_2) \qquad (4.3)$$

(t)

$R^2 = 0.9314$, MAPE = 15.94, SE = 13.4[11],

while the results for the model with dummy included are:

---

[11] R^2 - coefficient of determination, MAPE - mean absolute percentage error, SE standard error of regression. The use of Clopper Almons's multipurpose econometric software package "G" for IBM-PC (see A l m o n 1988), is gratefully acknowledged.

$$\hat{y} = 29.67 + 0.52 \times x - 62.5 \times u_2 + 1.002 \times (x\, u_2) \qquad (4.4)$$

(t) (6.9) (4.8) (5.4) (7.0)

$R^2 = 0.9583$, MAPE = 13.49, SE = 10.48.

In terms of explained variability of y the two equations do not differ much. However, the estimates differ considerably. When compared with the true parameters - the differences are well pronounced in the estimates of the first equation, and nonsignificant in the case of the second equation.

Table 1

Comparison of estimates of models with dummy and without dummy

Explained variable: y

Parameter	Estimate with dummy present	Estimate with dummy absent
$b_0 = 30$	29.670	20.817
$\Delta b_0 = -60$	-62.500	none
$b_1 = 0.5$	0.520	0.709
$\Delta b_1 = 1.0$	1.002	0.343

Source: The author's calculation.

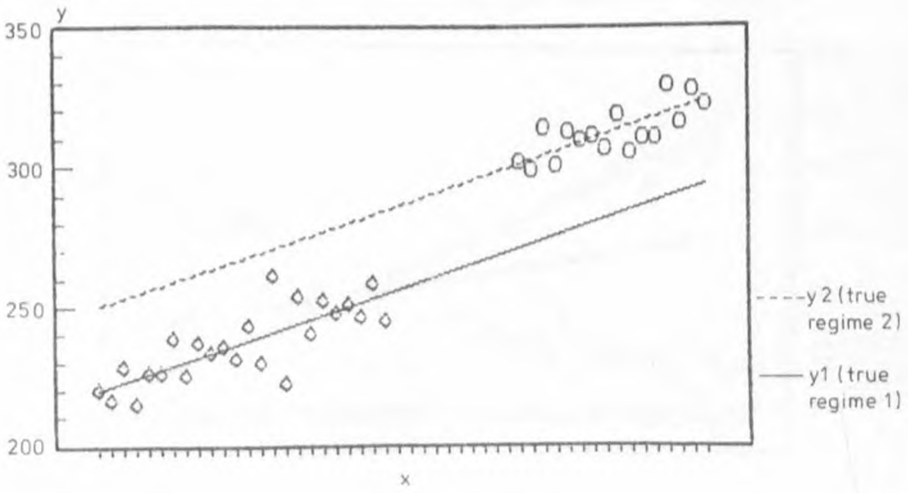

Fig. 3. The case of explicit intercept change discontinuity caused by parallel translation of the function

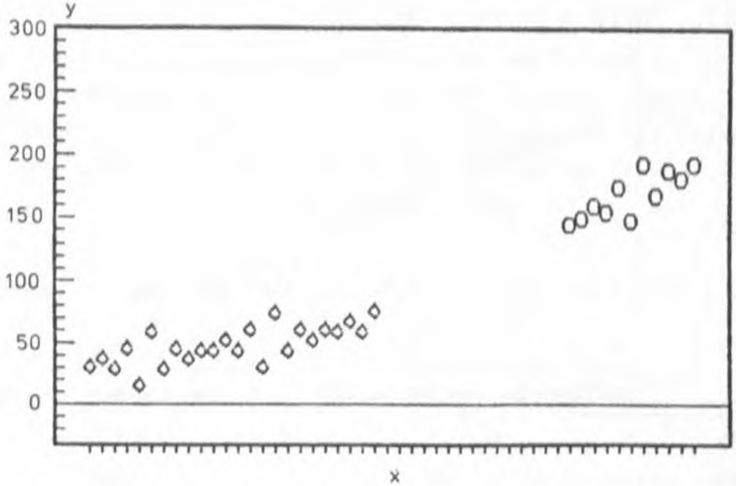

Fig. 4. Observations generated by a discontinuous relationship working in two regimes

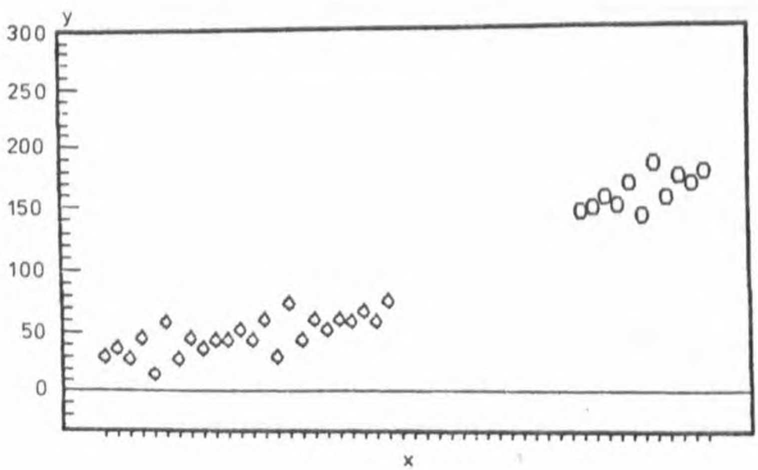

Fig. 5. Observations generated by another discontinuous relationship working in two regimes

Figure 3 shows the case of *explicit* intercept change, leading to discontinuity. The difference of intercepts (the size of parallel translation of the function) represents the *explicit* intercept change $\Delta\mathbb{E}$.

The situation complicates, when both - the translation and the slope change are present, since the change in the intercept is influenced by both effects $\Delta\mathbb{I}$ and $\Delta\mathbb{E}$. These effects may have opposite directions. After inspecting Figures 4 and 5 careful researcher may be inclined to allow for structural change in both - the slope (change in the response parameter) and intercept (parallel translation of the whole function), i.e. to estimate the parameters of the following function:

$$\hat{y} = b_{oo}u_1 + b_{o1}u_2 + b_{11} \times (x\ u_1) + b_{12} \times (x\ u_2) + e \qquad (4.5)$$

or in the alternative parameterization:

$$y = b_{oo} + \Delta b_{o1}u_2 + b_1 \times x + \Delta b_1 \times (x\ u_2) + e \qquad (4.6)$$

In the second parameterization the resulting change of intercept Δb_{01} (being measured in the estimation process) contains both - the *implicit* and the *explicit* effects: $\Delta b_{01} = \Delta\mathbb{I} + \Delta\mathbb{E}$.

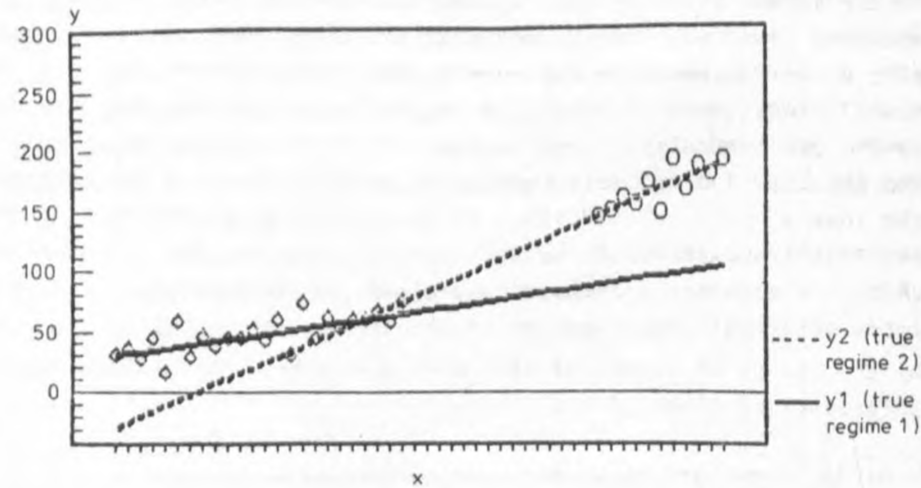

Fig. 6. The case of Figure 4 with the true regimes marked (The change of intercept seen)

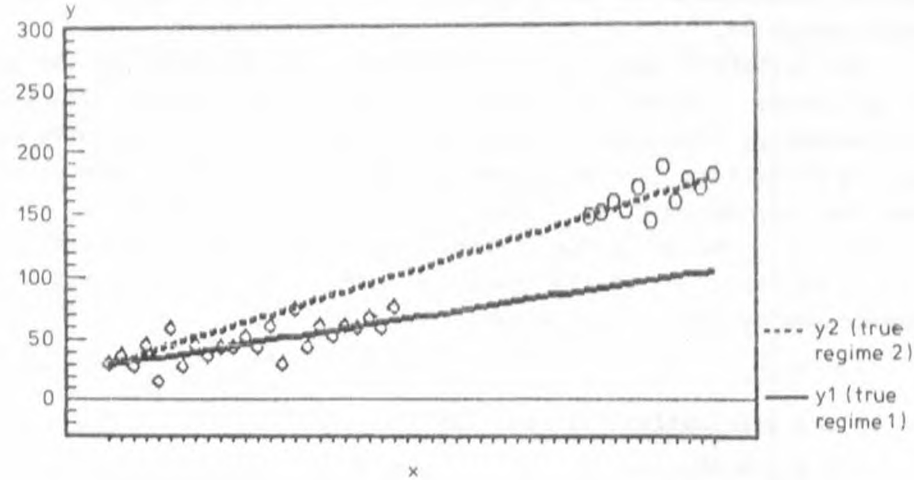

Fig. 7. The case of Figure 5 with the true regimes marked
(The intercept changes cancel out)

The true regimes are shown on Figures 6 and 7. The results of estimation confirm the change of intercept in the case shown on Figure 4, and its lack on Figure 5[12], although brief visual inspection does not suggest any special differences in the data plotted on figures. In the case 5 the two effects $\Delta\mathbb{I}$ and $\Delta\mathbb{E}$ simply cancel out. However, if one wants to measure the jump due to the discontinuity, there exists the possibility to calculate the explicit and implicit changes $\Delta\mathbb{I}$ and $\Delta\mathbb{E}$. Figure 8 demonstrates the idea of such calculation. $\Delta\mathbb{E}$ is identified as the size of the parallel translation of the first regime line so, that it passes through the center of the second cloud of data points (i.e. the point $(\bar{x}_2, \bar{y}_2)$. $\Delta\mathbb{I}$ equals to the change of intercept induced by the change of slope of the *translated line* (the difference of intercepts of lines y2 and y4)[13].

[12] The tests for structural change are relevant in this case, see H a c k l (ed.) (1989) for an extensive review. The identification of structural change in the case of Figures 5, 7 on statistical ground (i.e. through testing) may raise problems.

[13] For the second cloud of data shown on the Figures 1 and 2, and the estimates of equation (4.4) we calculate: $\bar{x}_2 = 109.5$, $\bar{y}_2 = 134.5$, $y_1(\bar{x}_2) = 29.67 +$

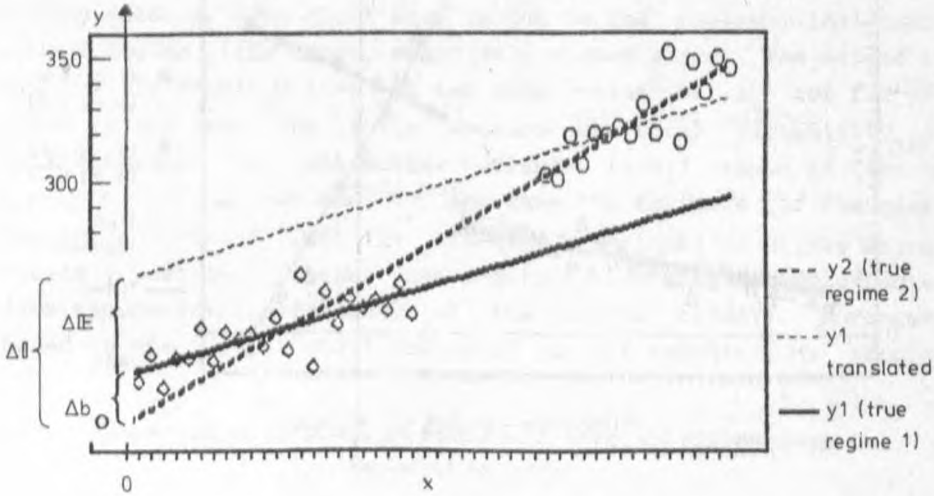

Fig. 8. Decomposition of intercept change when both - implicit $\Delta\mathbb{I}$ and explicit $\Delta\mathbb{E}$ changes of intercept are present

"Does a shift in the marginal propensity to consume cause (induce) a shift in autonomous consumption" asks a referee. As we have seen only values of the change in propensity from some specific neighborhood (the one illustrated on Figure 7, producing new line with the same intercept) will generate data from which no significant intercept change will be estimated, the other values would generate *implied* change. Is it change the researcher wants to interpret? Calculation of $\Delta\mathbb{I}$ and $\Delta\mathbb{E}$ may add some new flavour to such analyses.

The above discussion stresses that, in the estimation process, whenever one allows for structural breaks in slopes, one should take into account the possibility of the *induced* break in the intercept.

$+ 0.52 \times \bar{x}_2 \cong 86.6$. The *explicit* change $\Delta\mathbb{E}$ (measuring vertically the distance between the center $(\bar{x}_2, \bar{y}_2)$ of the second cloud and the first regime line equation $\bar{y}_2 - y_1(\bar{x}_2) = 134.5 - 86.6 = 47.9$. The *implicit* change $\Delta\mathbb{I}$ can be found as the difference $\Delta b - \Delta\mathbb{E} = -62.5 - 47.9 = -110.4$.

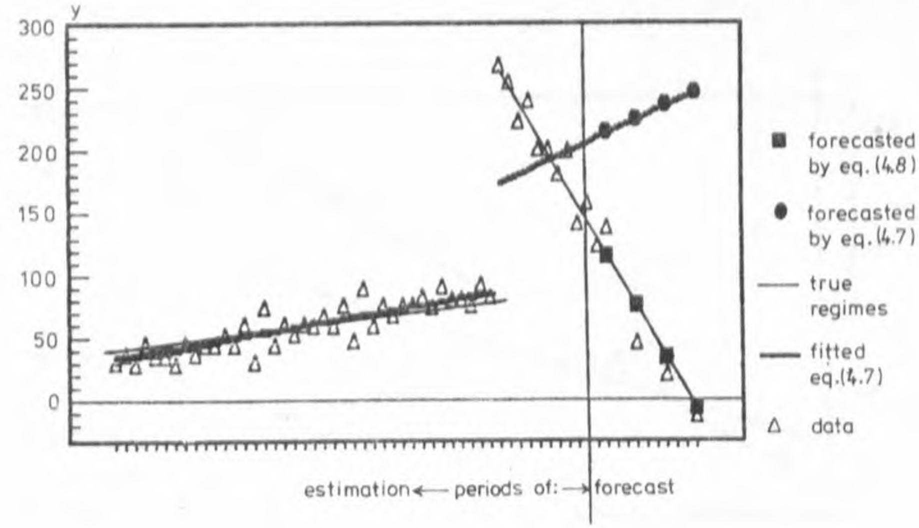

Fig. 9. An extreme example of forecast error due to lack of dummy (interactive variable works for absent dummy)

The adverse effects of neglecting the *induced* change of intercept is shown on Figure 9, based on the following data generation process:

$$z_t = \begin{cases} 30 + 0.5x_t + e_t & \text{for } t = 1, \ldots, 39 \\ 790 - 4.5x_t + e_t & \text{for } t = 40, \ldots, 50 \end{cases}$$

$$e_t \sim N(0,15), \quad x_t = (t - 1) \times 3.$$

Figure 9 shows the fitted values calculated from the estimated model (4.7), that (incorrectly) lacks the dummy u_2, along with the simple forecasting *ex ante,* to show the deterioration of the forecasts outside the sample. The estimated model is shown below:

$$\underset{(t)}{\hat{z}} = \underset{(3.5)}{34.5} + \underset{(2.95)}{0.44} \times x + \underset{(6.4)}{0.75} \times (x\, u_2) \qquad (4.7)$$

$R^2 = 0.7749$, MAPE = 21.49, SE = 30.56, sample size 1-49, which we may compare with the estimates of model that allows for the *implied* change:

$$\underset{(t)}{z} = \underset{(5.4)}{30.26} + \underset{(5.1)}{0.495} \times x + \underset{(10.2)}{758.5} \times u_2 - \underset{(8.8)}{4.982} \times (x\, u_2) \qquad (4.8)$$

$R^2 = 0.9308$, MAPE = 16.55, SE = 16.95.

In the case estimated without a change in intercept the parameter estimate $\Delta b_1 = 0.75$ represents a compromise of the two following effects. The first one is due to the *dispersion* inside the second regime (the upper, negatively sloped cloud). The second is due to the dramatic jump in the mean value of z_t for $t > 39$. Since in our case the latter *dominates* the total variability of this subsample, the interactive variable $(x \cdot u_2)$ *serves* in fact *as a proxy for the (omitted) dummy* u_2, and thus the estimate of the coefficient associated with the interactive variable is highly significantly positive (rather than negative, one would have expected from the decreasing tendency of the second cloud)). Forecasts based on the correct model obviously do not exhibit such errors.

Table 2

Comparison of estimates of models with dummy and without dummy
Explained variable: z

Parameter	Estimate with dummy present	Estimate with dummy absent
$b_0 = 30$	30.255	34.464
$\Delta b_0 = 760$	758.545	none
$b_1 = 0.5$	0.495	0.440
$\Delta b_1 = -5.0$	-4.986	0.754

Source: The author's calculation.

5. FINAL COMMENTS

We discussed the use of *interactive variables* not much discussed in the literature (the model with changing intercept being the notable exception).

Due to its flexibility the model with the interactive variables is a powerful extension of the constant parameter model, with the following distinctive properties:

(i) it can be reparameterized into some constant parameter model;

(ii) the reparameterized model being nonlinear in variables but linear in parameters can easily be estimated with standard estimation techniques such as least squares or its generalized version;

(iii) it allows for simultaneous presence in the equation both - discrete and continuous structural changes.

We proved also that the interactive variables based on binary dummies should be used with care. To avoid the risk of inconsistency of estimators both - the interactive based on dummy, and the respective dummy variables should be included in the initial version of an equation. Finally - even after finding the intercept change nonsignificant one - one may identify the *implicit* and *explicit* changes of intercept, (being in this case roughly of the same size, but opposite signs).

REFERENCES

Abraham B., Ledolter J. (1983), *Statistical Methods for Forecasting*, Wiley.

Almon C. (1988), *The Craft of Econometric Modelling*, University of Maryland Press, College Park.

Amemiya T. (1985), *Advanced Econometrics*, Cambridge University Press, Harvard.

Barten A., Bronsard W. (1970), *Two Stage Least Squares Estimation with Shifts in the Structural form*, "Econometrica", Vol. 38, No. 6.

Belsley D. A. (1972a), *Test for the Systematic Variation in Regression Coefficients*, "Annals of Economic and Social Studies", No. 2, p. 494--499.

Belsley D. A. (1972b), *On the Determination of the Systematic Parameter Variation in the Linear Regression Model*, "Annals of Economic and Social Studies", No. 2, p. 487-494.

Chow G. C. (1960), *Tests of Equality Between, Sets of Coefficients in Two Linear Regressions*, "Econometrica", Vol. 67, p. 815-821.

Dziechciarz J. (1989), *Changing and Random Parameter Models*, [in:] *Statistical Analysis and Forecasting of Economic Structural Change*, ed. P. Hackl, Springer.

Gajda J. B. (1985), *Dummies and Interactive Variables in Identification of Structural Change*, presented to the IIASA Workshop on Identification of Structural Changes, May, 1985, Łódź.

Goldberger A. (1963), *Econometric Theory*, Wiley.

Hackl P. (ed.) (1989), *Statistical Analysis and Forecasting of Economic Structural Change*, Springer.

H a c k l P., W e s t l u n d A. (1985), *Statistical Analysis of Structural Change: an Annotated Bibliography*, Collaborative Paper CP-85-31, IIASA, Laxenburg.

H a r v e y A. C., P h i l l i p s G. D. A. (1988), *The Estimation of Regression Models with Time-Varying Parameters*, "Games, Economic Dynamics, and Time Series Analysis", p. 306-321.

J u d g e G. G., H i l l R. C., G r i f f i t h W. E., L e e T. C. (1980), *The Theory and Practice of Econometrics*, Wiley.

J u d g e G. G., H i l l R. C., G r i f f i t h W. E., L u e t k e p o h l H., L e e T. C. (1985), *Introduction to the Theory and Practice of Econometrics*, Wiley.

K r ä m e r W., S o n n b e r g e r H. (1986), *The Linear Regression Model under Test*, Physica.

K l e i n L. R. (1972), *Textbook of Econometrics*, Prentice Hall.

K r e l l e W. (ed.) (1989), *The Future of the World Economy: Economic Growth and Structural Change*, Springer.

L e d o l t e r J. (1986), *Adaptive Estimation and Structural Change in Regression and Time Series Models*, presented to the IIASA Workshop on Economic Growth and Structural Change, IIASA, November 1986, Laxenburg.

R a o C. R. (1973), *Linear Statistical Inference and Its Applications*, Wiley.

T h e i l H. (1971), *Principles of Econometrics*, North Holland.

W e l f e W. (1980), *Modele ekonometryczne ze zmiennymi zerojedynkowymi - zastosowanie do analizy szeregów czasowych, prognozowania i symulacji* [Econometric models with binary variables applications in time series analysis, prediction and simulation], written under the grant PR.III.9, Institute of Econometrics and Statistics, University of Łódź, Łódź.

W e l f e W. (1985), *Econometric Macromodels of Unbalanced Growth*, "Prace Instytutu Ekonometrii i Statystyki", Institute of Econometrics and Statistics, University of Łódź, Łódź.

Jan B. Gajda

ZMIANY STRUKTURALNE PARAMETRÓW REGRESJI A ZMIENNE INTERAKCYJNE

Zmienne interakcyjne wykorzystywane są często do modelowania zmian strukturalnych. Do najpopularniejszych należą zmienne utworzone jako iloczyn dwóch innych zmiennych. Modelując skok wartości parametru związanego z pewną zmienną

wprowadzamy w charakterze dodatkowej zmiennej objaśniającej iloczyn takiej zmiennej i zmiennej zerojedynkowej. Ze względów czysto formalnych zmiana parametru strukturalnego *implikuje* zmianę wyrazu wolnego. Sprawę komplikuje możliwość wystąpienia *explicite* zmiany strukturalnej wyrazu wolnego. Te dwie zmiany mogą się wzajemnie znosić lub wzmacniać. Standardowe testy zmian strukturalnych skupiają uwagę na łącznym efekcie takich zmian, podczas gdy interesujące mogłoby być rozważenie ich w oddzielności. W artykule zawarta jest przestroga o konieczności włączania do estymowanego równania ze zmienną interakcyjną także stowarzyszonej zmiennej zerojedynkowej.

Ponadto artykuł zawiera dyskusję na temat możliwości wykorzystania szerzej rozumianych zmiennych interakcyjnych do modelowania zarówno dyskretnych, jak ciągłych zmian strukturalnych, zwłaszcza w przypadku niejednoczesnych zmian różnych parametrów w różnych momentach czasu.