Structural Change

Lecture 2

Piotr Dybka

Applied econometrics

What is structural change?

Suppose we have the model:

$$y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_K x_{Ki} + \epsilon_i$$

The coefficients β_k denote the structural parameters of the model we are looking for, which tell us the effect of a given variable x_{ki} on the dependent variable under study.

Structural change refers to a situation in which the value of a parameter β_k changes in the true data-generating process. It may be caused by changes in the economic system (e.g. market shifts), an intervention (e.g. a change in economic policy), a change in the law, etc.

Illustration: a change in the intercept β_0

Series generated as

$$y_i = 2 + 0.1\,x_i + \epsilon_i$$

and from observation 60 onward:

$$y_i = 5 + 0.1\,x_i + \epsilon_i$$

The **level** of the series changes, but the slope stays the same – a one-off jump is visible.

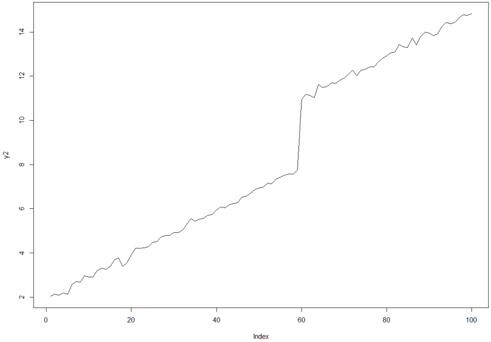


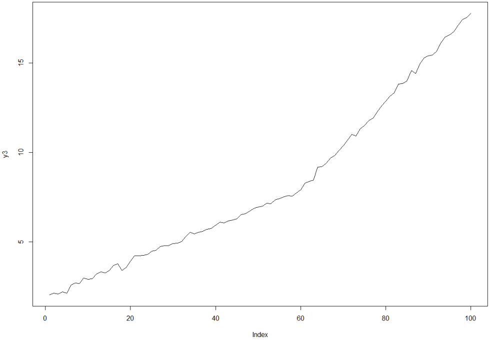
Illustration: a change in the slope β_1

Series generated as

$$y_i = 2 + 0.1\,x_i + \epsilon_i$$

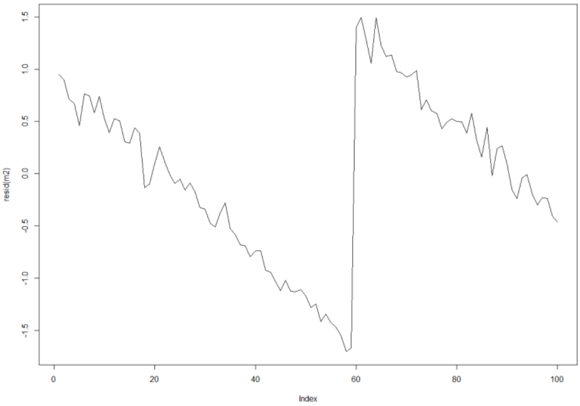
and from observation 60 onward:

$$y_i = -6.9 + 0.25\,x_i + \epsilon_i$$



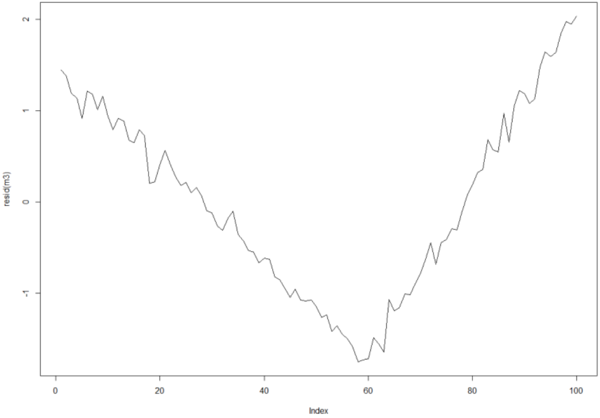
This time the **growth rate** changes. The constant was chosen so that the series stays continuous – the moment of the change is therefore much harder to spot by eye.

Consequences of structural change – example 1 (β_0)



Part of the "effect" of the change in β_0 has been absorbed by the estimates of the β_k parameters – hence the clear downward trend in the residuals, "broken" exactly at the point where the structural change occurs. This means the estimator is **biased** (an omitted variable – one that would capture the structural change).

Consequences of structural change – example 2 (β_1)



The change in the direction of the trend visible in the residuals is a result of "averaging" the value of the coefficient β_1: before the structural change the coefficient is overstated (and the residuals fall), and after the structural change it is understated (and the residuals rise).

Structural change – formal notation

We can formally represent structural change as a situation where:

$$\begin{cases} y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_K x_{Ki} + \epsilon_i & \text{for } S_i = 0 \\ y_i = \beta_0' + \beta_1' x_{1i} + \cdots + \beta_K' x_{Ki} + \epsilon_i & \text{for } S_i = 1 \end{cases}$$

- Where S_i is a variable indicating whether structural change occurred for a given observation, with $\beta_0 \neq \beta_0'$, $\beta_1 \neq \beta_1'$, and so on.
- This way of writing the model gives us the simplest way to test for and account for structural change. The model above can also be written as:

$$y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_K x_{Ki} + \gamma_0 S_i + \gamma_1 (S_i x_{1i}) + \cdots + \gamma_K (S_i x_{Ki}) + \epsilon_i$$

- Which means that $\beta_0' = \beta_0 + \gamma_0$, $\beta_1' = \beta_1 + \gamma_1$, and so on.

The Chow test

- To check parameter stability (the absence of structural change) we can use the Chow test.
- Running the test requires **specifying the point** at which structural change occurs – e.g. by analysing the model's residuals.
- The Chow test can also be used on cross-sectional data! In that case identifying the break point may require reordering the data (by the characteristic along which the change could have occurred).

1. We split the sample into two subsamples (according to where we expect structural change might have occurred).
2. We create a variable S_i, which takes the values 0 and 1 in the two subsamples, respectively.
3. We estimate the extended model:

$$\begin{aligned} y_i = {} & \beta_0 + \beta_1 x_{1i} + \cdots + \beta_K x_{Ki} \\ & + \gamma_0 S_i + \gamma_1 (S_i x_{1i}) + \cdots + \gamma_K (S_i x_{Ki}) + \epsilon_i \end{aligned}$$

4. We test the hypothesis $H_0: \ \gamma_0 = \gamma_1 = \cdots = \gamma_K = 0$ using the statistic

$$F = \frac{(SSE_R - SSE_U)/(K+1)}{SSE_U/(N - 2(K+1))}$$

5. Distribution $F_{(K+1,\, N-2(K+1))}$; we reject H_0 if $F \geq F_c$

where SSE_R, SSE_U are residual sums of squares with and without the restrictions ($\gamma_k = 0$).

The Chow test – an alternative version

1. We split the sample into two subsamples (according to where we expect structural change might have occurred).
2. We estimate two models based on the two subsamples (to obtain the parameter vectors $\boldsymbol{\beta}$ and $\boldsymbol{\beta}'$) and compute the residuals from these models, e_{Ui}.
3. We estimate the model on the full sample and compute the residuals from that model, e_{Ri}.
4. We compute $SSE_R = \sum_i e_{Ri}^2$ and $SSE_U = \sum_i e_{Ui}^2$ (here we sum the residuals from the two models estimated on the subsamples).
5. We test the hypothesis $H_0: \ \boldsymbol{\beta} = \boldsymbol{\beta}'$ using the statistic

$$F = \frac{(SSE_R - SSE_U)/(K+1)}{SSE_U/(N-2(K+1))}$$

6. Distribution $F_{(K+1,\, N-2(K+1))}$; we reject H_0 if $F \geq F_c$

Both versions give an **identical** value of the F statistic – they differ only in how it is computed.

What if we do not know the moment of change?

- The Chow test requires the break point to be specified *in advance*. If we choose it after looking at the data, the true significance level of the test is higher than the stated one.
- The **fluctuation test** checks parameter stability over the whole sample without specifying a point: in R, `sctest(model)`.
- The **Bai-Perron procedure** treats the moment of change as a *parameter to be estimated* – we look for the split of the sample that minimises the residual sum of squares:

$$(\hat{\tau}_1, \ldots, \hat{\tau}_m) = \arg \min_{\tau_1, \ldots, \tau_m} \sum_{j=1}^{m+1} \sum_{i=\tau_{j-1}+1}^{\tau_j} \left(y_i - \boldsymbol{x}_i' \hat{\boldsymbol{\beta}}_j \right)^2$$

- The number of breaks m is chosen using an information criterion (BIC).
- In R: `breakpoints(y~x)` from the `strucchange` package; `confint()` gives a confidence interval for the *date* of the change, and `breakdates()` gives the estimated moment.

Modelling structural change

Accounting for structural change in a model can be done in two ways. The first is a situation where our goal is the best possible fit of the model to the data (e.g. for forecasting purposes):

1. Adding binary variables (and interactions), as in the Chow test.
2. Threshold models (e.g. threshold VAR, STAR).
3. Switching models (e.g. Markov switching, or for variance – (MS)GARCH).
4. Models with time-varying parameters (e.g. estimated using the Kalman filter).

In the second case, however, our goal may be to precisely measure the effect of the structural change. This applies in particular to situations where the change is caused by a specific action – e.g. a change in economic policy, in a product's packaging, or in how a service is delivered:

1. The difference-in-differences estimator
2. Regression discontinuity design
3. Methods of experimental economics

Adding binary variables

- This is the simplest way to account for structural change. Such a model can be written as:

$$y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_K x_{Ki} + \gamma_0 S_i + \gamma_1 (S_i x_{1i}) + \cdots + \gamma_K (S_i x_{Ki}) + \epsilon_i$$

- The drawback of this approach is that we must determine for which observations and which variables structural change occurred, in order to correctly construct the binary variables S_i.
- If only the intercept changes, the variable S_i alone is enough; if the slope changes, we need the interactions $S_i x_{ki}$.

A threshold model: the AR model example

- As an **example** of a threshold model, let us discuss a modification of the autoregressive model:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} + \epsilon_t$$

- Let us extend our autoregressive model to a threshold TAR model (*Threshold Autoregressive model*):

$$y_t = \begin{cases} \beta_{10} + \beta_{11} y_{t-1} + \cdots + \beta_{1p} y_{t-p} + \epsilon_{1t}, & \text{for } S_{t-d} < r \\ \beta_{20} + \beta_{21} y_{t-1} + \cdots + \beta_{2p} y_{t-p} + \epsilon_{2t}, & \text{for } S_{t-d} \geq r \end{cases}$$

- Where S is a series that determines which regime an observation is in, d is the delay parameter, and r is the threshold value at which the regime changes.
- S can be any series, but if we take y_t itself as S, we obtain the SETAR model (*Self-Exciting Threshold Autoregressive model*):

$$y_t = \begin{cases} \beta_{10} + \beta_{11} y_{t-1} + \cdots + \beta_{1p} y_{t-p} + \epsilon_{1t}, & \text{for } y_{t-d} < r \\ \beta_{20} + \beta_{21} y_{t-1} + \cdots + \beta_{2p} y_{t-p} + \epsilon_{2t}, & \text{for } y_{t-d} \geq r \end{cases}$$

The STAR model

- There are further extensions of the SETAR model. The autoregressive model can be written in vector form:
$$y_t = \boldsymbol{X}_t \boldsymbol{\beta} + \epsilon_t$$
- Where $\boldsymbol{X}_t$ is a vector of successive lags of the dependent variable $(1, y_{t-1}, \ldots, y_{t-p})$, which we can then extend to the STAR model (*Smooth Transition Autoregression*):
$$y_t = \boldsymbol{X}_t \boldsymbol{\beta}_1 + \boldsymbol{X}_t \boldsymbol{\beta}_2 \, F(y_{t-d}, \lambda, r) + \epsilon_t$$
- Where $F(y_{t-d}, \lambda, r)$ is the transition function (from one regime to the other), taking values between 0 and 1. The parameter λ governs the "smoothness" of the transition between regimes, and r is the threshold value. For $F = 0$, $\boldsymbol{\beta}_1$ applies; for $F = 1$, $\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2$ applies.

Possible variants:

- Logistic smooth transition autoregression (LSTAR)
$$F(y_{t-d}, \lambda, r) = \left(1 + e^{-\lambda(y_{t-d} - r)}\right)^{-1}, \quad \lambda > 0$$
- Exponential smooth transition autoregression (ESTAR)
$$F(y_{t-d}, \lambda, r) = 1 - e^{-\lambda(y_{t-d} - r)^2}, \quad \lambda > 0$$

Switching models

- Switching models differ from threshold models in that the regime change does not depend on crossing a threshold value, but is instead described by a **probability**. An example of such models are Markov switching models, based on Markov chains.
- A Markov chain is a stochastic process y_t that satisfies the Markov condition:

$$P(y_t = j \mid y_0 = i_0, y_1 = i_1, \ldots, y_{t-1} = i_{t-1}) = P(y_t = j \mid y_{t-1} = i_{t-1})$$

- For every $t \in N$ and for any $i, j \in S$, where $S = \{1, 2, \ldots, r\}$ is the set of all phase states of the process.
- In other words: the behaviour of the process in the current period depends solely on its state in the previous period.
- The conditional probability of transitioning from state i to state j at a given moment is:

$$p_{ij}(t) = P(y_t = j \mid y_{t-1} = i)$$

- The probabilities $p_{ij}(t)$ form a matrix $\boldsymbol{P}(t) = [p_{ij}(t)]_{r \times r}$ satisfying the conditions:

$$\forall_{t \in N}\ \forall_{i,j \in S} \quad p_{ij}(t) \geq 0 \qquad\qquad \forall_{t \in N}\ \forall_{i \in S} \quad \sum_{j} p_{ij}(t) = 1$$

Switching models

- If the transition probabilities do not depend on t ($\forall_{t \in N}\ p_{ij}(t) = p_{ij}$), we have a homogeneous Markov chain, where $\boldsymbol{P} = [p_{ij}]_{r \times r}$ is the matrix of transition probabilities. In that case the following relationship holds:

$$\boldsymbol{D}_t = \boldsymbol{D}_{t-1}\boldsymbol{P}, \qquad \boldsymbol{D}_t = [d_{1t}, d_{2t}, \ldots, d_{rt}], \qquad d_{it} = P(y_t = i)$$

- Which can be represented as an econometric model of the form:

$$y_t = \begin{cases} y_{1t}, & \text{for } s_t = 1 \\ y_{2t}, & \text{for } s_t = 2 \\ \ldots & \\ y_{rt}, & \text{for } s_t = r \end{cases}$$

- Where s_t is the **unobserved** state of a homogeneous Markov chain with r states (regimes) and transition matrix $\boldsymbol{P}$.
- **Example.** If we assume that the component processes y_{it} that make up the whole process y_t are autoregressive processes with r states and lag order p:

$$y_t = \mu(s_t) + \beta_1 (y_{t-1} - \mu(s_{t-1})) + \cdots + \beta_p (y_{t-p} - \mu(s_{t-p})) + \epsilon_t, \qquad \epsilon_t \sim N\left(0, \sigma^2(s_t)\right)$$

- The "switching" process within the model can be the expected value $\mu(s_t)$, the autoregressive parameters β_j, or the variance of the error term $\sigma^2(s_t)$.