Chapter 3. Panel Threshold Regression Models

School of Economics and Management - University of Geneva

Christophe Hurlin (Université of Orléans)

University of Orléans

May 2018

1. Introduction

In econometrics, **threshold regression models** are a category of **regime-switching** models in which

- The slope parameters vary according to a "regime" switching mecanism that depends on a **threshold variable**,
- The regime is **observable** ex-post, contrary to the Markovian regime switching models.

1. Introduction

Definition (threshold regression model)

A typical threshold regression model is given by

$$y_t = \alpha + \beta_0' x_t + \beta_1' x_t h(q_t; \theta) + \varepsilon_t$$

where β_0 and β_1 are $K \times 1$ vectors, q_t is a **threshold variable**, θ a vector of parameters and $h(q_t; \theta)$ a **transition function.**

1. Introduction

Example (threshold regression model)

If the transition function is a binary function such that

$$h\left(q_t; c\right) = \begin{cases} 1 \text{ if } q_t \geq c \\ 0 \text{ if } q_t < c \end{cases},$$

then the model is simply defined by

$$y_t = \alpha + \beta_0' x_t + \beta_1' x_t \mathbb{I}_{(q_t \geq c)} + \varepsilon_t$$

where $\mathbb{I}_{(.)}$ is the indicator function and c is a location parameter.

1. Introduction

Example (threshold regression model)

If the transition function is a binary function, we have a **regime-switching mechanism** for the slope parameters that depends on the threshold variable and a location parameter c

$$y_t = \begin{cases} \alpha + \left(\beta_0' + \beta_1'\right) x_t + \varepsilon_t \text{ if } q_t \geq c \\ \alpha + \beta_0' x_t + \varepsilon_t \text{ if } q_t < c \end{cases},$$

In this simple case, we have two "regimes" for the slope parameters, i.e. β_0 and $\beta_0 + \beta_1$.

1. Introduction

Remarks

- The transition function $h(q_t; \theta)$ may be smoother than a binary function.
- In general, the transition function $h(q_t; \theta)$ is assumed to verified

$$0 \leq h(q_t; \theta) \leq 1$$

Sometimes, we have

$$\lim_{q_t \to +\infty} h(q_t; \theta) = 1 \qquad \lim_{q_t \to -\infty} h(q_t; \theta) = 0$$

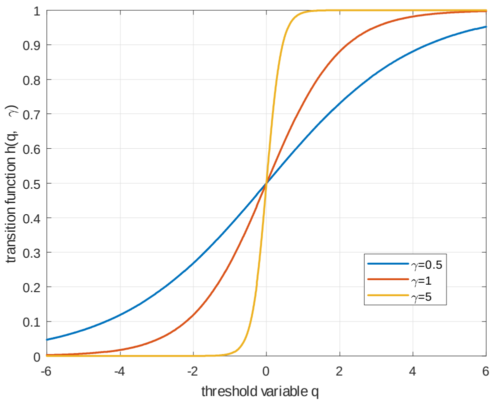
Example (Logistic transition function)

A logistic transition function is defined as:

$$h\left(q_t;\gamma,c\right)=\frac{1}{1+\exp\left(-\gamma\left(q_t-c\right)\right)},\;\;\gamma>0$$

1. Introduction

Figure: Logistic transition function with $c = 0$



1. Introduction

Panel data

- The threshold regressions in panel data models allow to model the **heterogeneity** of the slope parameters.
- These models give a **parametric approach** of the heterogeneity which is associated to an **economic "story"** (interpretation).

Introduction

The outline of this chapter is the following:

Section 2

The Panel Threshold Regression (PTR) Model

2. The panel threshold regression model

Objectives

1. Introduce the **panel threshold regression (PTR)** model.
2. Understand the link with **heterogeneous** panel models.
3. Understand the link with **time varying parameters** panel models.
4. Understand the differences with **random coefficient** models.

2. The panel threshold regression model

- The **Panel Threshold Regression (PTR)** model has been introduced by Hansen (1999).
- In this paper, threshold regression methods are developed for **non-dynamic** panels with individual fixed effects.

Hansen, B. E. 1999. Threshold effects in non-dynamic panels: estimation, testing, and inference, *Journal of Econometrics*, 93, 334-368

2. The panel threshold regression model

Definition (panel threshold regression model)

The panel threshold regression (PTR) model is defined as

$$y_{it} = \alpha_i + \beta_1' x_{it} \mathbb{I}_{(q_{it} \leq c)} + \beta_2' x_{it} \mathbb{I}_{(q_{it} > c)} + \varepsilon_{it}$$

where the dependent variable y_{it} is scalar, α_i is a fixed effect, the threshold variable q_{it} is scalar, the regressor x_{it} is a k vector, and $\mathbb{I}_{(.)}$ is the indicator function and c is a threshold parameter.

2. The panel threshold regression model

Assumptions

- The threshold variable is exogeneous or at least predetermined ($q_{it} = y_{i,t-d}$ with $d \geq 1$).
- For the identification of β_1 and β_3, it is required that the elements of x_{it} are not time invariant.
- The threshold variable q_{it} is not time invariant.
- The error ε_{it} is assumed to be i.i.d. with $\mathbb{E}\left(\varepsilon_{it}\right) = 0$ and $\mathbb{V}\left(\varepsilon_{it}\right) = \sigma_{\varepsilon}^2$.

2. The panel threshold regression model

Panel Threshold Regression model

An alternative specification of the PTR is

$$y_{it} = \alpha_i + \beta_0' x_{it} + \beta_3' x_{it} \mathbb{I}_{(q_{it} \leq c)} + \varepsilon_{it}$$

or equivalently

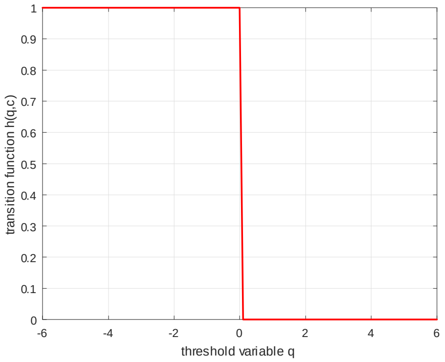
$$y_{it} = \alpha_i + \beta_0' x_{it} + \beta_3' x_{it} h\left(q_{it}; c\right) + \varepsilon_{it}$$

whith $h\left(q_{it}; c\right) = \mathbb{I}_{(q_{it} \leq c)}$ the binary transition function

$$y_{it} = \begin{cases} \alpha_i + \left(\beta_0' + \beta_3'\right) x_{it} + \varepsilon_{it} \text{ if } q_{it} \leq c \\ \alpha_i + \beta_3' x_{it} + \varepsilon_{it} \text{ if } q_{it} > c \end{cases},$$

2. The panel threshold regression model

Figure: Transition function $h\left(q_{it}; c\right)$ with $c = 0$



2. The panel threshold regression model

Definition (heterogeneous panel data model)

The PTR model can be viewed as an **heterogeneous** and **time-varying parameters** panel data model

$$y_{it} = \alpha_i + \beta_{it}' x_{it} + \varepsilon_{it} = \begin{cases} \alpha_i + \beta_1' x_{it} + \varepsilon_{it} \text{ if } q_{it} \leq c \\ \alpha_i + \beta_2' x_{it} + \varepsilon_{it} \text{ if } q_{it} > c \end{cases},$$

where the marginal effect (slope parameters) satisfy

$$\frac{\partial y_{it}}{\partial x_{it}} = \beta_{it} = \begin{cases} \beta_1 \text{ if } q_{it} \leq c \\ \beta_2 \text{ if } q_{it} > c \end{cases},$$

2. The panel threshold regression model

Heterogeneous panel data model

$$y_{it} = \begin{cases} \alpha_i + \beta_1' x_{it} + \varepsilon_{it} \text{ if } q_{it} \leq c \\ \alpha_i + \beta_2' x_{it} + \varepsilon_{it} \text{ if } q_{it} > c \end{cases},$$

At a given time t, two cross-section units i and j may have two different slope parameters

$$\frac{\partial y_{it}}{\partial x_{it}} = \beta_1 \neq \frac{\partial y_{jt}}{\partial x_{jt}} = \beta_2 \quad \text{if } q_{it} \leq c \text{ and } q_{jt} > c.$$

2. The panel threshold regression model

Time-varying parameter panel data model

$$y_{it} = \begin{cases} \alpha_i + \beta_1' x_{it} + \varepsilon_{it} \text{ if } q_{it} \leq c \\ \alpha_i + \beta_2' x_{it} + \varepsilon_{it} \text{ if } q_{it} > c \end{cases},$$

A given cross-section unit i may have different slope parameters at different dates t and s

$$\frac{\partial y_{it}}{\partial x_{it}} = \beta_1 \neq \frac{\partial y_{is}}{\partial x_{is}} = \beta_2 \quad \text{if } q_{it} \leq c \text{ and } q_{is} > c.$$

2. The panel threshold regression model

Example (Marginal effect and PTR)

Consider a PTR model with $K = 1$ regressor such that

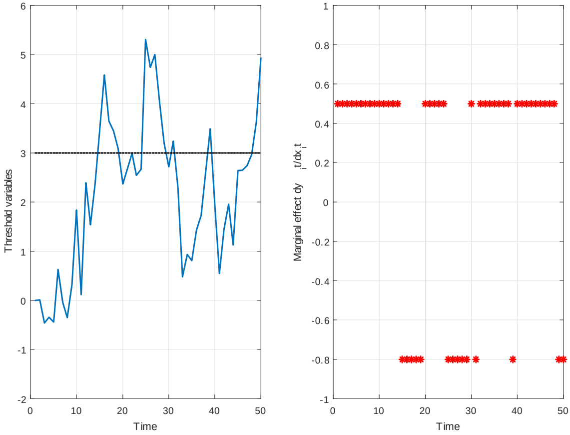
$$y_{it} = \alpha_i + 0.5 x_{it} \mathbb{I}_{(q_{it} \leq 3)} - 0.8 x_{it} \mathbb{I}_{(q_{it} > 3)} + \varepsilon_{it}$$

Then, the marginal effect of the regressor x_{it} on y_{it} is equal to

$$\frac{\partial y_{it}}{\partial x_{it}} = 0.5 \mathbb{I}_{(q_{it} \leq 3)} - 0.8 \mathbb{I}_{(q_{it} > 3)} = \begin{cases} 0.5 \text{ if } q_{it} \leq 3 \\ -0.8 \text{ if } q_{it} > 3 \end{cases},$$

2. The panel threshold regression model

Figure: Marginal effect with $K = 1$, $\beta_1 = 0.5, \beta_2 = -0.8$ and $c = 3$



2. The panel threshold regression model

Panel Threshold Regression model

The PTR model is an alternative to the random coefficient model

Panel Threshold Regression	$\beta_{it} = f\left(q_{it}; c\right)$	β_{it} is a constant term
	Economic interpretation (through q_{it})	
Random Coefficient Model	β_i i.i.d. $\left(\overline{\beta}, \Delta\right)$	β_i is a random variable
	No economic interpretation	

2. The panel threshold regression model

Estimation by NLS

The PTR model can be rewritten as:

$$y_{it} = \alpha_i + \beta' x_{it}(c) + \varepsilon_{it}$$

$$\underset{(2K,1)}{x_{it}(c)} = \begin{pmatrix} x_{it}\mathbb{I}_{(q_{it} \leq c)} \\ x_{it}\mathbb{I}_{(q_{it} > c)} \end{pmatrix} \qquad \underset{(2K,1)}{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

2. The panel threshold regression model

Estimation by NLS

In order to eliminate the individual effects, we apply the Within transformation

$$y_{it}^{*} = \beta' x_{it}^{*}(c) + \varepsilon_{it}^{*}$$

$$y_{it}^{*} = y_{it} - \overline{y}_{i} \quad \overline{y}_{i} = T^{-1} \sum_{t=1}^{T} y_{it}$$

$$x_{it}^{*}(c) = x_{it}(c) - \overline{x}_{i}(c) \quad \overline{x}_{i}(c) = T^{-1} \sum_{t=1}^{T} x_{it}(c)$$

$$\varepsilon_{it}^{*} = \varepsilon_{it} - \overline{\varepsilon}_{i} \quad \overline{\varepsilon}_{i} = T^{-1} \sum_{t=1}^{T} \varepsilon_{it}$$

2. The panel threshold regression model

Let us define

$$\underset{(T,1)}{y_i^*} = \begin{pmatrix} y_{i,1}^* \\ y_{i,2}^* \\ \dots \\ y_{it}^* \end{pmatrix} \qquad \underset{(T,2K)}{X_i^*(c)} = \begin{pmatrix} x_{1,i,1}(c)' \\ x_{1,i,2}(c)' \\ \dots \\ x_{1,it}(c)' \end{pmatrix} \qquad \underset{(T,1)}{\varepsilon_i^*} = \begin{pmatrix} \varepsilon_{i,1}^* \\ \varepsilon_{i,2}^* \\ \dots \\ \varepsilon_{it}^* \end{pmatrix}$$

$$\underset{(Tn,1)}{Y^*} = \begin{pmatrix} y_1^* \\ y_2^* \\ \dots \\ y_n^* \end{pmatrix} \qquad \underset{(Tn,2K)}{X^*(c)} = \begin{pmatrix} X_1^*(c) \\ X_2^*(c) \\ \dots \\ X_n^*(c) \end{pmatrix} \qquad \underset{(Tn,1)}{\varepsilon^*} = \begin{pmatrix} \varepsilon_1^* \\ \varepsilon_2^* \\ \dots \\ \varepsilon_n^* \end{pmatrix}$$

2. The panel threshold regression model

Definition (given threshold)

For any **given threshold** c, the slope coeffcient β can be estimated by ordinary least squares (OLS).

$$\widehat{\beta}(c) = \left(X^{*}(c)' X^{*}(c)\right)^{-1} X^{*}(c)' Y$$

The vector or residuals is given by

$$\widehat{\varepsilon}^{*}(c) = Y^{*} - X^{*}(c)\,\widehat{\beta}(c)$$

and the sum of squared errors is

$$SSR(c) = \widehat{\varepsilon}^{*}(c)'\,\widehat{\varepsilon}^{*}(c)$$

2. The panel threshold regression model

Definition (Threshold estimation)

The estimation if the threshold parameter c is obtained by minimization of the concentrated sum of squared

$$\widehat{c} = \underset{c \in \Theta}{\arg\min}\ SSR\left(c\right)$$

Remarks

- It is undesirable for a threshold c to be selected which sorts too few observations into one or the other regime.
- This possibility can be excluded by restricting the search in to values of c such that a minimal percentage of the observations (say, 1% or 5%) lie in each regime.

$$c \in \Theta = \left[\text{quantile} \left(\{ q_{it} \}_{i=1,t=1}^{N,T}, 0.05 \right), \text{quantile} \left(\{ q_{it} \}_{i=1,t=1}^{N,T}, 0.95 \right) \right]$$

- Given $\widehat{c}$, we can compute the estimates for β as

$$\widehat{\beta} = \beta \left(\widehat{c} \right) = \begin{pmatrix} \beta_1 \left(\widehat{c} \right) \\ \beta_2 \left(\widehat{c} \right) \end{pmatrix}$$

2. The panel threshold regression model

runmycode

Boosting idea dissemination and research reproducibility

2. The panel threshold regression model

2. The panel threshold regression model

```
PURPOSE: Hansen (1999) "Threshold effect in non dynamic panels: Estimation, Testing, and Inference ",
Journal of Econometrics 93, 345-368

-------------------------------------------------------
Usage:   Estimate a two regimes Threshold Panel, and test the effect of threshold

-------------------------------------------------------
Function : Hansen_2reg(Y,Q,X,T,min_reg,CL,simulation,affi)

Where:
           Y : the dependant variable,   size : vector (N*T,1)
           Q : the transition variable, size : vector (N*T,1)
           X : the regressor,           size : matrix (N*T,K)

T          : Time Dimension
min_reg    : Trimming parameter, if is empty : minreg = 0.05
CL         : confidence level for gamma, if is empty : CL= 0.95
simulation : number of simulation for test the threshold effect, if is empty : simulation = 0
affi = 1 if the user want to display the results
-------------------------------------------------------------------------------------------------
RETURNS:
```

2. The panel threshold regression model

```
------------------------------------------------------------------------------------------------
RETURNS:

res.gam           :  estimated threshold parameter gamma;
res.IC_gam        : confidence intervals of the threshold
res.effets_fixes  : fixed effects ;
res.coef_cent     : estimated slope parameters (in each regimes);
res.coef_stat     : give the coefficients, the statistic and a second statistic robust to heteroskedasti
res.residus       : estimated residuals;
res.rss           : sum of squared resid;
res.F1            : F1 Statistic of the test H0: no threshold against H1 one threshold (only if simulato
res.pvalue        : pvalue of F1 statistic (only if simulaton>0)

------------------------------------------------------------------------------------------------
```

2. The panel threshold regression model

```
=================================
==== Model with one threshold ===
=================================

Cross-unit dimension N =  76

Time dimension T =  35

Estimated threshold parameter (gamma) = -2.7731

Confidence Interval (95%) on gamma = [-3.1245 , -2.7278]

Estimated slope parameters (first column) and t-stats (second column) for the FIRST regime : q(t)<=-2.773

  0.5067   30.8901
  0.1641   10.1197
  0.0139    0.7037

Estimated slope parameters (first column) and t-stats (second column) for the SECOND regime : q(t)>-2.773

  0.5572   40.3585
  0.1046    5.8240
  0.0955    9.2412
```

2. The panel threshold regression model

```
SSR = 44.4338

Estimated residual variance = 0.0172

Number of simulation (for F1 statistic pvalue) =  100

F1 test statistic = 112.1812

pvalue = 0.0000
```

2. The panel threshold regression model

Definition (extension to $r + 1$ regimes)

A general specification of the PTR can be proposed with $r + 1 > 2$ regimes or r threshold parameters $c_1, ..c_r$.

$$y_{it} = \alpha_i + \sum_{j=1}^{r} \beta_j' x_{it} \mathbb{I}_{(c_{j-1} < q_{it} \leq c_j)} + \varepsilon_{it}$$

with $c_0 = -\infty$ and $c_{r+1} = +\infty$.

2. The panel threshold regression model

Example (PTR with three regimes)

For instance, a three-regimes model ($r = 3$) with 2 threshold parameters c_1 and c_2 is given by

$$\begin{aligned} y_{it} &= \alpha_i + \beta_1' x_{it} \mathbb{I}_{(q_{it} \leq c_1)} + \beta_2' x_{it} \mathbb{I}_{(c_1 < q_{it} \leq c_2)} \\ &\quad + \beta_3' x_{it} \mathbb{I}_{(q_{it} > c_2)} + \varepsilon_{it} \end{aligned}$$

2. The panel threshold regression model

Example (PTR with four regimes)

For instance, a three-regimes model ($r = 4$) with 3 threshold parameters c_1, c_2 and c_3 is given by

$$\begin{aligned} y_{it} &= \alpha_i + \beta_1' x_{it} \mathbb{I}_{(q_{it} \leq c_1)} + \beta_2' x_{it} \mathbb{I}_{(c_1 < q_{it} \leq c_2)} \\ &\quad + \beta_3' x_{it} \mathbb{I}_{(c_2 < q_{it} \leq c_3)} + \beta_4' x_{it} \mathbb{I}_{(q_{it} > c_3)} + \varepsilon_{it} \end{aligned}$$

2. The panel threshold regression model

Inference

- If one comes to test whether the threshold effect is statically significant in the model with two regimes, the null hypothesis is

$$H_0 : \beta_1 = \beta_2$$

- This null hypothesis corresponds to the hypothesis of no threshold effect.

- Under H_0 the model is then equivalent to a linear model.

$$y_{it} = \alpha_i + \beta' x_{it} + \varepsilon_{it}$$

with $\beta = \beta_1 = \beta_2$.

2. The panel threshold regression model

Inference

- The null hypothesis $H_0 : \beta_1 = \beta_2$ can be tested by a standard test.
- If we note S_0 the sum of squared of the linear model, the approximate likelihood ratio test of H_0 is based on:

$$F_1 = \frac{SSR_0 - SRR_1 \left(\widehat{c} \right)}{\widehat{\sigma}^2}$$

where $\widehat{\sigma}^2$ denotes a convergent estimate of σ^2.

2. The panel threshold regression model

Inference

- The main problem is that under the null, the threshold parameter c is not identified (**nuisance parameter**).
- Consequently, the asymptotic distribution of F_1 is not standard and, in particular, does not correspond to a chi-squared distribution.
- This issue has been largely studied in literature devoted to threshold models, notably since the seminal paper by Davies (1977, 1987).
- One solution is to use a bootstrap procedure to determine the asymptotic distribution of the statistic F_1.

2. The panel threshold regression model

Number of regimes (thresholds)

- The same kind of inference procedure can be applied in order to determining the **number of thresholds**.

- A likelihood ratio test of one threshold versus two thresholds is based on the statistic

$$F_2 = \frac{SSR_1\left(\widehat{c}\right) - SSR_2\left(\widehat{c}_1, \widehat{c}_2\right)}{\widehat{\sigma}^2}$$

  where $\widehat{c}_1$ and $\widehat{c}_2$ denote the threshold estimates of the model with three regimes , and $SSR_2\left(\widehat{c}_1, \widehat{c}_2\right)$ denotes the corresponding residual sum of squares.

2. The panel threshold regression model

Number of regimes (thresholds)

- The hypothesis of one threshold is rejected in favor of two thresholds if F_2 is larger than its critical value.
- The corresponding asymptotic p-value can be approximated by bootstrap simulations (Hansen, 1999).
- If the model with two thresholds (three regimes) is not rejected accepted, we to test the hypothesis of two thresholds (three regimes) against the alternative of three thresholds (four regimes).

2. The panel threshold regression model

Number of regimes (thresholds)

- The corresponding likelihood ratio statistic, denoted $F_{3.}$, is defined as:

$$F_3 = \frac{SSR_2\left(\widehat{c}_1, \widehat{c}_2\right) - SSR_3\left(\widehat{c}_1, \widehat{c}_2, \widehat{c}_3\right)}{\widehat{\sigma}^2}$$

  where $SSR_3\left(\widehat{c}_1, \widehat{c}_2, \widehat{c}_3\right)$ denotes the residual sum of squares of the model with four regimes and three threshold parameters.

- Thus, a sequential procedure based on F_1, F_2, F_3, etc. allows to determining the number of regimes

2. The panel threshold regression model

Example

Candelon, Colletaz, Hurlin (2012) investigate the threshold effects in the productivity of infrastructure investment in developing countries.

$$y_{it} = \begin{cases} a_i + \alpha_1 k_{it} + \beta_1 h_{it} + \gamma_1 x_{it} + \varepsilon_{it} & \text{if } q_{it} \leq \lambda \\ a_i + \alpha_2 k_{it} + \beta_2 h_{it} + \gamma_2 x_{it} + \varepsilon_{it} & \text{if } q_{it} > \lambda. \end{cases}$$

where y_{it} is the aggregate added value, k_{it} is physical capital, h_{it} is human capital, x_{it} is infrastructure stock.

Candelon B., Colletaz G., Hurlin C. (2013), Network Effects and Infrastructure Productivity in Developing Countries, *Oxford Bulletin of Economics and Statistics*, 75(6), 887-913.

2. The panel threshold regression model

Table 3. Tests for Threshold Effects: Model A, $q_{it} = x_{it}$ [12]

	Roads	Electricity	Telephones	Railways
Test for single threshold				
F_1	90.2	112.1	59.5	128.4
P-value	0.00	0.00	0.00	0.00
1% Critical Values	13.7	14.3	14.3	14.9
5% Critical Values	15.3	15.5	16.8	16.5
10% Critical Values	19.3	20.6	22.0	21.3
Test for double threshold				
F_2	84.7	74.3	82.6	122.6
P-value	0.00	0.00	0.00	0.00
1% Critical Values	50.4	19.8	17.2	30.7
5% Critical Values	55.2	22.7	19.2	36.9
10% Critical Values	68.2	26.2	25.1	42.2
Test for triple threshold				
F_3	41.7	43.8	55.7	85.9
P-value	0.00	0.00	0.00	0.00
1% Critical Values	13.6	13.6	13.6	13.2
5% Critical Values	15.2	15.6	15.7	15.5
10% Critical Values	18.0	21.2	19.2	18.4

Notes: P-values and critical values are computed from 300 simulations. F1 denotes the Fisher type statistic associated to the test of the null of no threshold against one threshold. F2 corresponds to the test one threshold against two thresholds and F3 corresponds to the test of two thresholds against three thresholds.

2. The panel threshold regression model

Table 4. Four Regimes Panel Models[13]. Model A: $q_{it} = x_{it}$

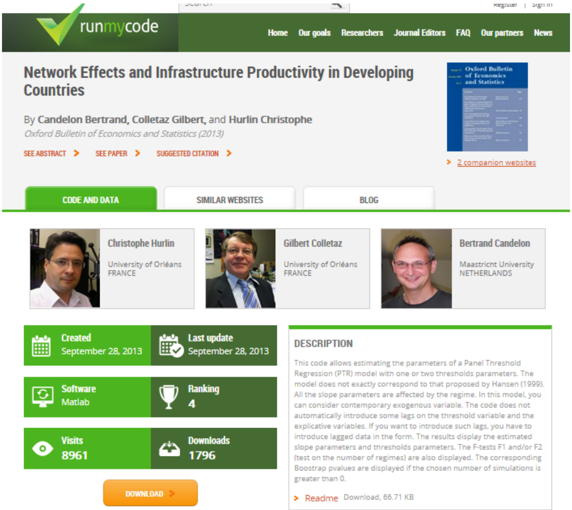
	Roads	Electricity	Telephones	Railways
Regime 1: $q_{it} \leq \lambda_1$				
Physical Capital per Worker	0.575 (32.30)	0.519 (26.78)	0.431 (21.43)	0.613 (43.70)
Human Capital per Worker	0.053 (2.67)	0.145 (8.57)	0.102 (3.24)	0.565 (15.62)
Infrastructure per Worker	0.122 (3.967)	0.022 (1.08)	−0.099 (−2.36)	0.184 (6.54)
Regime 2: $\lambda_1 < q_{it} \leq \lambda_2$				
Physical Capital per Worker	0.403 (12.34)	0.754 (22.06)	0.389 (18.48)	0.597 (42.53)
Human Capital per Worker	0.077 (1.89)	0.185 (5.75)	0.108 (6.18)	0.400 (14.25)
Infrastructure per Worker	−1.682 (−5.67)	0.655 (7.68)	0.220 (14.44)	−0.096 (−1.68)
Regime 3: $\lambda_2 < q_{it} \leq \lambda_3$				
Physical Capital per Worker	0.567 (34.93)	0.584 (38.39)	0.465 (14.51)	0.625 (42.21)
Human Capital per Worker	0.208 (10.62)	0.067 (3.39)	−0.168 (−4.17)	0.228 (9.48)
Infrastructure per Worker	0.166 (8.79)	0.135 (12.09)	0.132 (2.45)	−0.157 (−6.06)

2. The panel threshold regression model

Regime 4: $q_{it} > \lambda_3$				
Physical Capital per Worker	0.574 (32.28)	0.548 (34.94)	0.380 (14.03)	0.526 (28.97)
Human Capital per Worker	0.104 (2.53)	0.252 (7.36)	0.126 (3.27)	0.716 (12.35)
Infrastructure per Worker	0.129 (5.03)	0.041 (2.50)	0.210 (11.13)	−0.014 (−0.54)
Threshold Estimates				
First Threshold $\widehat{\lambda}_1$	−0.917	−3.317	1.337	−1.427
Second Threshold $\widehat{\lambda}_2$	−0.642	−2.954	4.314	−0.882
Third Threshold $\widehat{\lambda}_3$	1.248	−0.519	4.808	0.975
Residual Sum of Squares	12.11	42.01	23.27	26.61

Notes: The t-statistics in parenthesis are computed with an estimator of the covariance matrix robust to heteroskedasticity. The confidence intervals for the threshold parameters are not reported. See Appendix A1, for the confidence intervals in a model with three regimes.

2. The panel threshold regression model



2. The panel threshold regression model

Key Concepts Section 2

1. Panel Threshold Regression Model
2. Heterogeneous and time-varying parameters
3. Non Linear Least Squares
4. Inference for the number of regimes
5. Davies problem

Section 2

The Panel Smooth Threshold Regression (PSTR) Model

3. The panel smooth threshold regression model

Objectives

1. Introduce the **panel smooth threshold regression (PSTR)** model.
2. Understand the link with **heterogeneous** panel models.
3. Understand the link with **time varying parameters** panel models.

3. The panel smooth threshold regression model

- The **Panel Smooth Threshold Regression (PSTR)** model has been introduced by Gonzalez, Teräsvirta and van Dijk (2005).
- This model can be viewed as a generalization of the PTR model.

González, A., Teräsvirta, T., van Dijk, D., 2005. Panel smooth transition regression model. *Working Paper Series in Economics and Finance*, vol. 604.

3. The panel smooth threshold regression model

Definition (PSTR model)

The PSTR with two extreme regimes can be defined as

$$y_{it} = \alpha_i + \beta_0' x_{it} + \beta_1' x_{it} g\left(q_{it}; \gamma, c\right) + \varepsilon_{it}$$

where $g\left(q_{it}; \gamma, c\right)$ is a transition function, q_{it} a threshold variable, c a location parameter and γ a slope parameter.

3. The panel smooth threshold regression model

Transition function

- The transition function $g\left(q_{it};\gamma,c\right)$ is a continuous function of the observed variable q_{it}
- The transition function is normalized to be bounded between 0 and 1.

$$0 \leq g\left(q_{it};\gamma,c\right) \leq 1$$

3. The panel smooth threshold regression model

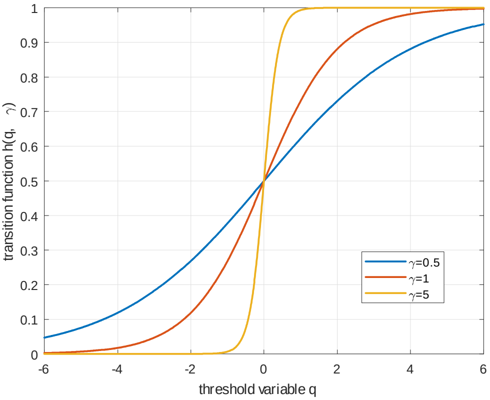
Definition

Gonzalez, Teräsvirta and van Dijk (2005) consider a logistic transition function

$$g\left(q_{it};\gamma,c\right)=\frac{1}{1+\exp\left(-\gamma\left(q_{it}-c\right)\right)},\ \ \gamma>0$$

3. The panel smooth threshold regression model

Figure: Logistic transition function with $c = 0$



3. The panel smooth threshold regression model

Remark 1

If $\gamma \rightarrow \infty$, the logistic function tends to an indicator function and the PSTR corresponds to a **PTR model**

$$\lim_{\gamma \rightarrow \infty} g\left(q_{it}; \gamma, c\right) = \mathbb{I}_{\left(q_{it} \geq c\right)} = \left\{ \begin{array}{l} 1 \text{ if } q_{it} \geq c \\ 0 \text{ if } q_{it} < c \end{array} \right. ,$$

$$y_{it} = \alpha_i + \beta_0' x_{it} + \beta_1' x_{it} \mathbb{I}_{\left(q_{it} \geq c\right)} + \varepsilon_{it}$$

3. The panel smooth threshold regression model

Remark 2

If $\gamma \rightarrow 0$, the logistic function tends to an indicator function and the PSTR corresponds to a **linear panel model**

$$\lim_{\gamma \rightarrow \infty} g\left(q_{it}; \gamma, c\right) = \frac{1}{2} \qquad \forall q_{it}$$

$$y_{it} = \alpha_i + \beta' x_{it} + \varepsilon_{it} \qquad \beta = \frac{\beta_0 + \beta_1}{2}$$

3. The panel smooth threshold regression model

Definition (heterogeneous panel data model)

The PSTR model can be viewed as an **heterogeneous** and **time-varying parameters** panel data mode since the marginal effect (slope parameters) satisfy

$$\frac{\partial y_{it}}{\partial x_{it}} = \beta_{it} = \beta_0 + \beta_1 g\left(q_{it}; \gamma, c\right)$$

as soon as q_{it} does not belong to x_{it}, with by convention

$$\beta_0 \leq \beta_{it} \leq \beta_0 + \beta_1$$

3. The panel smooth threshold regression model

Remark 3

Even there are two "extreme" regimes/values for the slope parameters, in fact there is an **infinity of regimes** (possible values) for the slope parameters β_{it} given the observed values for q_{it}:

$$\beta_{it} = \beta_0 + \beta_1 g\left(q_{it}; \gamma, c\right)$$

$$\beta_0 \leq \beta_{it} \leq \beta_0 + \beta_1$$

3. The panel smooth threshold regression model

Definition (generalization to r+1 regimes)

The PSTR model can be generalised to $r+1$ extreme regimes as follows:

$$y_{it} = \alpha_i + \beta_0' x_{it} + \sum_{j=1}^{r} \beta_j' x_{it} \; g_j(q_{it}; \gamma_j, c_j) + \varepsilon_{it}$$

where the r transition functions $g_j(q_{it}; \gamma_j, c_j)$ depend on the slope parameters γ_j and on the location parameters c_j.

3. The panel smooth threshold regression model

Marginal effects

If the transition variable does not belong to the set of regressors, the slope parameters become:

$$\beta_{it} = \frac{\delta y_{it}}{\delta x_{it}} = \beta_0 + \sum_{j=1}^{r} \beta_j \; g_j(q_{it}; \gamma_j, c_j)$$

3. The panel smooth threshold regression model

Marginal effects

If the transition variable does not belong to the set of regressors, the slope parameters become:

$$\beta_{it} = \frac{\delta y_{it}}{\delta x_{it}} = \beta_0 + \sum_{j=1}^{r} \beta_j \; g_j(q_{it}; \gamma_j, c_j) + \sum_{j=1}^{r} \beta_j \; \frac{\delta g_j(q_{it}; \gamma_j, c_j)}{\delta q_{it}} q_{it}$$

3. The panel smooth threshold regression model

Estimation

- The estimation of the parameters of the PSTR model consists of eliminating the individual effects α_i and then in applying NLS to the transformed model
- See Gonzàlez et al. (2005) or Colletaz and Hurlin (2006), for more details.

González, A., Teräsvirta, T., van Dijk, D., 2005. Panel smooth transition regression model. *Working Paper Series in Economics and Finance*, vol. 604.

Colletaz, G., Hurlin, C., 2006. Threshold effects in the public capital productivity: an international panel smooth transition approach. *Document de Recherche du Laboratoire d'Economie d'Orléans*. 2006-1.

3. The panel smooth threshold regression model

Inference

- Gonzàlez et al. (2005) propose a testing procedure in order (i) to test the linearity against the PSTR model and (ii) to determine the number, r, of transition functions, i.e. the number of extreme regimes which is equal to $r+1$.

- Testing the linearity in a PSTR model can be done by testing

$$H_0 : \gamma = 0 \quad \text{or} \quad H_0 : \beta_0 = \beta_1$$

- But in both cases, the test will be non standard since under H_0 the PSTR model contains unidentified nuisance parameters.

3. The panel smooth threshold regression model

Inference

- A possible solution is to replace the transition function $g_j(q_{it}, \gamma_j, c_j)$ by its **first-order Taylor expansion** around $\gamma = 0$ and to test an equivalent hypothesis in the auxiliary regression

$$y_{it} = \alpha_i + \theta_0 \; x_{it} + \theta_1 \; x_{it} \; q_{it} + \varepsilon_{it}$$

- The parameters θ_i are proportional to the slope parameter γ of the transition function.

- Thus, testing the linearity of the model against the PSTR simply consists of testing

$$H_0 : \; \theta_1 = 0$$

3. The panel smooth threshold regression model

Definition (linearity/homogeneity test)

The F-statistic for the homogeneity assumption $H_0 : \theta_1 = 0$ is then defined by:

$$LM_F = (SSR_0 - SSR_1) / [SSR_0 / (TN - N - 1)]$$

with SSR_0 the panel sum of squared residuals under H_0 (linear panel model with individual effects) and SSR_1 the panel sum of squared residuals under H_1 (PSTR model with two regimes). Under the null hypothesis, the F-statistic has an approximate $F(1, TN - N - 1)$ distribution.

Choice of number of transitions

- The logic is similar when it comes to testing the number of transition functions in the model or equivalently the number of extreme regimes.
- We use a sequential approach by testing the null hypothesis of no remaining nonlinearity in the transition function.

3. The panel smooth threshold regression model

Choice of number of transitions

- We want to test whether there is one transition function ($H_0: \ r = 1$) or whether there are at least two transition functions ($H_0: \ r = 2$).
- Let us assume that the model with $r = 2$ is defined as:

$$y_{it} = \alpha_i + \beta_0 x_{it} + \beta_1 \ x_{it} \ g_1(q_{it}; \gamma_1, c_1) + \beta_2 \, x_{it} \, g_2(q_{it}; \gamma_2, c_2) + \varepsilon_{it}$$

- The logic of the test consists in replacing the second transition function by its first-order Taylor expansion around $\gamma_2 = 0$ and then in testing linear constraints on the parameters in

$$y_{it} = \alpha_i + \beta_0 x_{it} + \beta_1 \ x_{it} \ g_1(q_{it}; \gamma_1, c_1) + \theta_1 \ x_{it} \ q_{it} + \epsilon_{it}^*$$

  and the test of no remaining nonlinearity is simply defined by $H_0: \ \theta_1 = 0$.

3. The panel smooth threshold regression model

Choice of number of transitions

- Let us denote SSR_0 the panel sum of squared residuals under H_0, i.e. in a PSTR model with one transition function.

$$y_{it} = \alpha_i + \beta_0 x_{it} + \beta_1 \; x_{it} \; g_1(q_{it}; \gamma_1, c_1) + \varepsilon_{it}$$

- Let us denote SSR_1 the sum of squared residuals of the transformed model.

$$y_{it} = \alpha_i + \beta_0 x_{it} + \beta_1 \; x_{it} \; g_1(q_{it}; \gamma_1, c_1) + \theta_1 \; x_{it} \; q_{it} + \epsilon_{it}^*$$

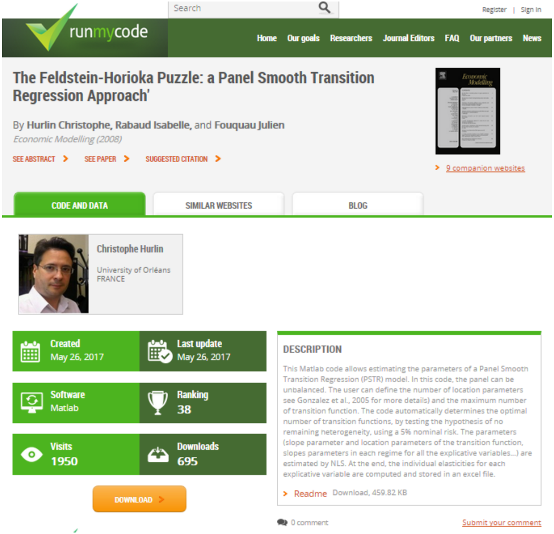
- The F-statistic LM_f can be calculated in the same way by adjusting the number of degrees of freedom.

3. The panel smooth threshold regression model

Testing procedure

1. Given a PSTR model with $r = r^*$, we test the null $H_0 : r = r^*$ against $H_1 : r = r^* + 1$.
2. If H_0 is not rejected the procedure ends.
3. Otherwise, the null hypothesis $H_0 : r = r^* + 1$ is tested against $H_1 : r = r^* + 2$.
4. The testing procedure continues until the first acceptance of H_0.
5. Given the sequential aspect of this testing procedure, at each step of the procedure the significance level must be reduced by a constant factor $0 < \rho < 1$ in order to avoid excessively large models. Gonzàlez et al. (2005) suggest $\rho = 0.5$.

3. The panel smooth threshold regression model



3. The panel smooth threshold regression model

```
PURPOSE: Gonzalez, Terascirta and Van Dijk (2004), "Panel Smooth Transition Regression Model
and An Application to Investment Under Credit Constraint", Stockholm School of Economics

-------------------------------------------------------
Usage:   Estimate a Panel Smooth Threshold Panel

-------------------------------------------------------
 Function : res=STAR_Panel(Y,Q,X,N0,m0,rmax,condini_user)

Where:
           Y               : the dependant variable,  size, vector (N*T,1)
           Q               : the transition variable, size, vector (N*T,1)
           X               : the regressor, size, matrix (N*T,K)
           N               : number of individuals
           m               : number of location parameters
           rmax            : maximum number of transition functions authorised (r<=rmax)
           condini_user    : initial conditions given by the user

-----------------------------------------------------------------------------------------------
```

3. The panel smooth threshold regression model

```
RETURNS:

        res.balanced       : the message 'Balanced Sample' indicates a balanced sample
        res.N              : cross-section dimension
        res.T              : time dimension. For unbalanced panel, give the T dimension for each c
        res.m              : number of location parameters for each transition (fixed by the user)
        res.r              : Optimal number of transition function
        res.gam            : Estimated slope parameter (for each transition function)
        res.c              : Estimated location parameter (for each transition function)
        res.beta           : Estimated slope parameters for the explicative variables (by column f
        res.beta_std       : Standard errors (corrected for heteroskedasticity, by column for each
        res.beta_tstat     : t-statistics (by column for each transition)
        res.beta_std_nc:   : Standard errors (not corrected for heteroskedasticity, by column for
        res.rss            : RSS
        res.fixed          : Estimated fixed effects
        res.resid          : Residuals
        res.g              : Values of the transition function (N*T,r)matrix, where r is the numbe
        res.exitflag       : Equal to 1 if there is convergence
        res.output         : Informations on the convergence
        res.coef_indi      : Individual coefficients for each explicative variable (NT,K) matrix
        res.nbparam        : Number of parameters
        res.AIC            : Akaike criteria
        res.BIC            : BIC criteria
```

3. The panel smooth threshold regression model

```
**********************
*** LINEARITY Tests ***
**********************

H0: Linear Model H1: PSTR model with at least one Threshold Variable (r=1)

    Wald Tests (LM):          W = 84.104    pvalue = 0.000

    Fisher Tests (LMF):       F = 45.876    pvalue = 0.000

    LRT Tests (LRT):          LRT = 89.161    pvalue = 0.000
```

3. The panel smooth threshold regression model

```
***************************************************************************
*** TESTING THE NUMBER OF REGIMES: TESTS OF NO REMAINING NON-LINEARITY ***
***************************************************************************

 Initial Conditions : Assumed Number of Thresholds r = 1   Number of Regressions = 270
Initial Conditions on (c,gamma)
   5.0000   -1.4357

 Estimation of the Model with r = 1 and m = 1 : Convergence = 1   RSS = 3.982
 RSS under H1 = 3.956

H0: PSTR with r = 1  against  H1: PSTR with at least r = 2

    Wald Tests (LM):         W = 4.872    pvalue = 0.088

    Fisher Tests (LMF):      F = 2.364    pvalue = 0.095

    LRT Tests (LRT):         LRT = 4.888    pvalue = 0.087


Given the choices of rmax = 2 and m = 1, the OPTIMAL (LMF criterion) NUMBER OF THRESHOLD FUNCTIONS is r = 1
```

3. The panel smooth threshold regression model

```
**************************************
*** FINAL ESTIMATION OF PSTR MODEL ***
**************************************

  Final Estimation of the Model with r = 1 and m = 1 by NLS ***

  Initial Conditions on (gamma,c) :
   4.8632   -1.4664

 RSS = 3.982      Convergence = 1

 AIC = -5.221      BIC = -5.184

Estimated slope parameter of the transition function (one for for each transition function)
   4.8632

Estimated location parameters (per column for each transition function)
  -1.4664
```

3. The panel smooth threshold regression model

```
Estimated location parameters (per column for each transition function)
  -1.4664

Estimated slope parameters (per column for each transition function)
   0.3460   -0.0974
   0.4430   -0.3796

Standard Errors of estimated slope parameters corrected fo heteroskedasticity (per column f
   0.0234    0.0199
   0.0352    0.0342

t-statistics based on corrected standard errors (per column for each transition function)
  14.8021   -4.9051
  12.5958  -11.1053

OUPUT: Individual Elasticities for each explicative variable (first column is the cross sec
```

3. The panel smooth threshold regression model

Example

Hurlin, Rabaud and Fouquau (2008) considers a PSTR model for determining the relative influence of five factors on the Feldstein and Horioka result for OECD countries.

$$I_{it} = \alpha_i + \beta_0 S_{it} + \beta_1 S_{it} \; g(q_{it}; c) + \epsilon_{it}$$

They consider five main factors (as potential threshold variable) generally considered in this literature: (i) economic growth, (ii) demography, (iii) degree of openness, (iv) country size and (v) current account balance.

Hurlin C., Rabaud I., and Fouquau J. (2008) The Feldstein-Horioka Puzzle: a PSTR Approach. *Economic Modelling*, 25, 284-299.

3. The panel smooth threshold regression model

Table 2
LM_F tests for remaining nonlinearity

Model	Model A	Model B	Model C
Threshold variable	Growth	Openness	Size
H_0: $r=0$ vs. H_1: $r=1$	73.08(0.00)	293.1(0.00)	6.54(0.01)
H_0: $r=1$ vs. H_1: $r=2$	0.637(0.42)	0.071(0.79)	0.48(0.49)
H_0: $r=2$ vs. H_1: $r=3$	–	–	–
Model	Model D	Model E	Model F
Threshold variable	$\text{pop}_{<15\text{ years}}$	$\text{pop}_{>64\text{ years}}$	Cur. account
H_0: $r=0$ vs. H_1: $r=1$	186.5(0.00)	211.1(0.00)	141.3(0.00)
H_0: r=1 vs. H_1: $r=2$	3.21(0.07)	0.075(0.78)	0.006(0.94)
H_0: $r=2$ vs. H_1: $r=3$	–	–	–

Notes: For each model (i.e. for each threshold variable), the testing procedure works as follows. First, test a linear model ($r=0$) against a model with one threshold ($r=1$). If the null hypothesis is rejected, test the single threshold model against a double threshold model ($r=2$). The procedure goes on until the hypothesis no additional threshold is not rejected. The corresponding LM_F statistic has an asymptotic $F[1, TN-N-(r+1)]$ distribution under H_0. The corresponding p-values are reported in parentheses.

3. The panel smooth threshold regression model

Table 5
Parameter estimates for the final PSTR models

Specification	Model A	Model B	Model C	Model D	Model E	Model F
Threshold variable	Growth	Openness	Size	$\text{pop}_{<15 \text{ years}}$	$\text{pop}_{>64 \text{ years}}$	Cur. account
r^*	1	1	1	1	1	1
Parameter β_0	0.487(0.04)	0.885(0.02)	0.292(0.05)	0.378(0.03)	0.600(0.04)	0.893(0.08)
Parameter β_1	0.126(0.01)	−0.678(0.03)	0.492(0.05)	0.215(0.01)	−0.371(0.02)	−0.676(0.07)
Location parameters c	2.74(1.81)	87.2(16.1)	0.52(0.04)	20.7(0.48)	14.2(0.848)	−2.63(3.34)
Slopes parameters γ	0.774(1.70)	0.037(0.01)	73.6(116)	0.579(1.38)	0.547(0.04)	0.109(0.10)
AIC criterion	2.102	1.907	2.068	1.957	1.939	1.700
Schwarz criterion	2.123	1.927	2.068	1.978	1.959	1.730
Number of obs.	936	960	960	960	960	639

Notes: The standard errors for coefficients in parentheses are corrected for heteroskedasticity. The standard errors for the smooth parameter γ and the threshold parameter c are computed by evaluating the Hessian matrix except for two transition variables: the share of under 15s in total population, and the growth. In this two cases, we use the *outer product of gradients*. For each model, the number of transition functions r is determined by a sequential testing procedure (see Table 1). For each transition function, the estimated location parameters c and the corresponding estimated slope parameter γ are reported. The PSTR parameters can not be directly interpreted as elasticities.

3. The panel smooth threshold regression model

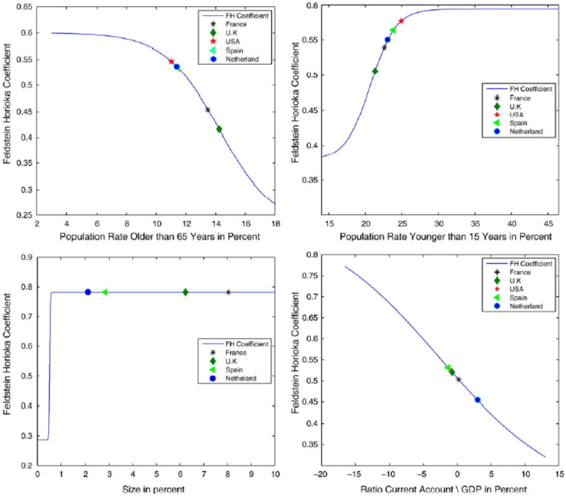


Fig. 2. Estimated FH coefficients PSTR models.

3. The panel smooth threshold regression model

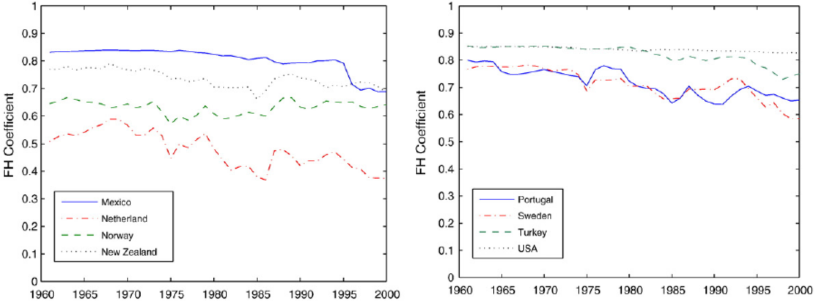


Fig. 3. Estimated FH individual coefficients: PSTR Model B.

3. The panel smooth threshold regression model

Key Concepts Section 3

1. Panel Smooth Threshold Regression Model
2. Marginal effects (slope parameters)
3. Non Linear Least Squares
4. Inference for the number of regimes

End of Chapter 3

Christophe Hurlin (University of Orléans)