PERUVIAN ECONOMIC ASSOCIATION

# Panel Models with Two Threshold Variables: The Case of Financial Constraints

Arturo Lamadrid-Contreras

N.R. Ramírez-Rondán

Working Paper No. 128, October 2018

# Panel Models with Two Threshold Variables: The Case of Financial Constraints*

Arturo Lamadrid-Contreras
Citibanamex

N.R. Ramírez-Rondán
Universidad del Pacífico

October, 2018

**Abstract**

We develop threshold estimation methods for panel data models with two threshold variables and individual fixed specific effects covering short time periods. In the static panel data model, we propose least squares estimation of the threshold and regression slopes using fixed effects transformations; while in the dynamic panel data model, we propose maximum likelihood estimation of the threshold and slope parameters using first difference transformations. In both models, we propose to estimate the threshold parameters sequentially. We apply the methods to a 15-year sample of 565 U.S. firms to test whether financial constraints affect investment decisions.

*JEL Classification:* C13, C23, G11.
*Keywords:* Threshold model, panel data, capital market imperfections.

*We would like to thank Bruce Hansen and Carlos Barrera for their valuable comments and suggestions. We also thank the participants of the 2017 International Panel Data Conference (Thessaloniki, Greece); the 2017 Latin American Meeting of the Econometric Society (Buenos Aires, Argentina); the 25th Annual Meeting of the Midwest Econometrics Group (St. Louis, Missouri); the 2016 Congress of the Peruvian Economic Association (Lima, Peru); the 2016 Annual Meeting of the Economic Society of Chile (Viña del Mar, Chile), the XXXIV Meeting of Economists of the Central Bank of Peru (Lima, Peru); and the Universidad del Pacífico Research Seminar.
Corresponding author: N.R. Ramírez-Rondán, Department of Economics, Universidad del Pacífico. Av. Salaverry 2020, Lima 11, Lima, Peru. E-mail address: nr.ramirezr@up.edu.pe.

# 1 Introduction

One of the most interesting non-linear regression models is the threshold regression model developed by Howell Tong at the end of the 1970s.[1] This model has been enormously influential in economics and remains popular in current applied econometric practice. The model splits the sample into classes based on the value of an observed variable, regardless of whether or not it exceeds a given threshold; that is, it internally sorts the data on the basis of a given threshold determinant into groups of observations, each of which obeys the same model. Hansen (2011) provides an excellent literature review on the use of these models in econometrics and economics.

The literature contains a well-developed least squares estimation and inference theory for threshold models with exogenous regressors, including Chan (1993) and Hansen (1997, 2000), while Chen et al. (2012) extended these model to allow two threshold variables. Another important extension of these works would be to consider two threshold variables in panel data models; that is, cross section and time series data. Thus, in this paper we introduce econometric techniques appropriate for panel data models with two threshold variables. We describe least squares estimation and maximum likelihood estimation methods for the static and dynamic models, respectively.

These methods are used to study the relationship between investment and capital market imperfections or financial constraints. Capital market imperfections may imply that firms are restricted in their access to external finance. The impact of such imperfections depends on the degree of informational asymmetries and on firms' growth opportunity (balance sheet) conditions (Hu and Schiantarelli, 1998). Informational asymmetries between borrowers and lenders generate agency cost, which creates a pecking order; that is, the more severe the information and agency problems, the higher the cost of external finance, and thus the greater the sensitivity to internal finance such as cash flow.

The impact of capital imperfections also depends on the extent of the growth opportunities; Gonzáles et al. (2005) argue that in the case of firms with ample growth opportunities, internal and external finance are substitutes and so the firms' investment decisions are independent of their financial structure. Improving growth prospects help to solve problems of overspending and to decrease the financial constraints of firms suffering from severe information asymmetry; conversely, when firms have limited growth opportunities, a positive relationship between investment and cash flows indicates investment in negative net present value projects (Pawlina and Renneboog, 2005), which may due to agency problems.

[1] See Tong (2007) for the emergence of the threshold model.

Empirical studies divide firms into constrained and unconstrained groups based on a single variable that measures capital market imperfections (see Schiantarelli, 1996; Hubbard, 1998, for a review). Nonetheless, the criteria used to split the sample differ from study to study, since there are multiple factors that influence a firm's financial strength and borrowing ability. In this paper we taken into account two indicators of capital market imperfections: a measure of the degree of informational asymmetry, and a measure of growth opportunities. Thus, we expect heavily indebted firms with limited growth opportunities to face much higher costs of external finance and hence their investment decisions to be more sensitive to internal finance.

The outline of the paper is as follows. In section 2 we develop estimation techniques based on a static panel threshold model with two threshold variables. In section 3 we develop estimation techniques based on a dynamic panel threshold model with two threshold variables. In section 4 we show the performance of the estimators proposed via Monte Carlo experiments. In section 5 we present an application to test whether financial constraints affect investment decisions. Finally, in section 6 we conclude.

# 2 Static Panel Data Model

The observed data are from a balanced panel $\{y_{it}, x_{it} : 1 \leq i \leq n, : 1 \leq t \leq T\}$. The subscript $i$ indexes the individual and the subscript $t$ indexes time. The dependent variable $y_{it}$ is scalar. The threshold variables $q_{it} = (q_{1it}, q_{2it})$ can be elements or functions of the vector $x_{it}$ of exogenous variables and must have continuous distributions. The equation of interest is

$$
\begin{aligned}
y_{it} &= \mu_i + \beta_1' x_{it} 1(q_{1it} \leq \gamma_1, q_{2it} \leq \gamma_2) + \beta_2' x_{it} 1(q_{1it} \leq \gamma_1, q_{2it} > \gamma_2) \\
&\quad + \beta_3' x_{it} 1(q_{1it} > \gamma_1, q_{2it} \leq \gamma_2) + \beta_4' x_{it} 1(q_{1it} > \gamma_1, q_{2it} > \gamma_2) + e_{it}, \qquad (1)
\end{aligned}
$$

where the threshold parameters $\gamma = (\gamma_1, \gamma_2) \in \Gamma$, where $\Gamma = [\underline{\gamma_1}, \overline{\gamma_1}] \times [\underline{\gamma_2}, \overline{\gamma_2}]$ is a strict subsets of the support of $q_{it}$. These parameters are unknown and need to be estimated. $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)'$ are the slope parameters and $\beta_i \neq \beta_j$ for some $i \neq j$; $\mu_i$ is the individual specific effect assumed to be fixed and $e_{it}$ is the error term assumed to be independent and identically distributed (iid). The analysis is asymptotic with fixed $T$ as $n$ grows to infinity.

Another compact representation of (1) is to set

$$x_{it}(\gamma) = \begin{pmatrix} x_{it}1(q_{1it} \leq \gamma_1, q_{2it} \leq \gamma_2) \\ x_{it}1(q_{1it} \leq \gamma_1, q_{2it} > \gamma_2) \\ x_{it}1(q_{1it} > \gamma_1, q_{2it} \leq \gamma_2) \\ x_{it}1(q_{1it} > \gamma_1, q_{2it} > \gamma_2) \end{pmatrix};$$

then, equation (1) equals

$$y_{it} = \mu_i + \beta' x_{it}(\gamma) + e_{it}. \tag{2}$$

## 2.1 Estimation

One traditional method for eliminating the individual fixed specific effect $\mu_i$ is to remove individual-specific means. While straightforward in linear models, threshold specification (2) calls for more careful treatment. Similar to Hansen (1999) in a panel data model with one threshold variable, we take averages over the time index $t$, which produces

$$\overline{y}_i = \mu_i + \beta' \overline{x}_i(\gamma) + \overline{e}_i, \tag{3}$$

where $\overline{y}_i = T^{-1} \sum_{t=1}^{T} y_{it}$, $\overline{e}_i = T^{-1} \sum_{t=1}^{T} e_{it}$ and

$$\begin{aligned} \overline{x}_i(\gamma) &= \frac{1}{T} \sum_{t=1}^{T} x_{it}(\gamma) \\ &= \begin{pmatrix} \frac{1}{T} \sum_{t=1}^{T} x_{it}1(q_{1it} \leq \gamma_1, q_{2it} \leq \gamma_2) \\ \frac{1}{T} \sum_{t=1}^{T} x_{it}1(q_{1it} \leq \gamma_1, q_{2it} > \gamma_2) \\ \frac{1}{T} \sum_{t=1}^{T} x_{it}1(q_{1it} > \gamma_1, q_{2it} \leq \gamma_2) \\ \frac{1}{T} \sum_{t=1}^{T} x_{it}1(q_{1it} > \gamma_1, q_{2it} > \gamma_2) \end{pmatrix}, \end{aligned}$$

and taking the difference between (2) and (3) yields

$$y^*_{it} = \beta' x^*_{it}(\gamma) + e^*_{it}, \tag{4}$$

where $y^*_{it} = y_{it} - \overline{y}_i$, $x^*_{it}(\gamma) = x_{it}(\gamma) - \overline{x}_i(\gamma)$ and $e^*_{it} = e_{it} - \overline{e}_i$.

Let

$$y_i^* = \begin{bmatrix} y_{i2}^* \\ \vdots \\ y_{iT}^* \end{bmatrix}, \ x_i^*(\gamma) = \begin{bmatrix} x_{i2}^*(\gamma) \\ \vdots \\ x_{iT}^*(\gamma) \end{bmatrix}, \ e_i^* = \begin{bmatrix} e_{i2}^* \\ \vdots \\ e_{iT}^* \end{bmatrix}$$

denote the stacked data and errors for an individual, with one time period deleted, and let

$$Y^* = \begin{bmatrix} y_1^* \\ \vdots \\ y_i^* \\ \vdots \\ y_n^* \end{bmatrix}, \ X^*(\gamma) = \begin{bmatrix} x_1^*(\gamma) \\ \vdots \\ x_i^*(\gamma) \\ \vdots \\ x_n^*(\gamma) \end{bmatrix}, \ e^* = \begin{bmatrix} e_1^* \\ \vdots \\ e_i^* \\ \vdots \\ e_n^* \end{bmatrix}.$$

Using this notation, (4) is equivalent to

$$Y^* \;=\; X^*(\gamma)\beta + e^*, \tag{5}$$

and given $\gamma$, the conditional least squares (CLS) estimator for $\beta$ is

$$\widehat{\beta}(\gamma) = (X^*(\gamma)'X^*(\gamma))^{-1}X^*(\gamma)'Y^*; \tag{6}$$

thus, the vector of regression residuals is

$$\widehat{e}^*(\gamma) = Y^* - X^*(\gamma)\beta(\gamma), \tag{7}$$

and the sum of squared errors is

$$S(\gamma) = \widehat{e}^*(\gamma)'\widehat{e}^*(\gamma). \tag{8}$$

Chan (1993) and Hansen (1999) recommend the estimation of $\gamma$ by conditional least squares in the context of a model with one threshold variable, and Chen et al. (2012) in the context of a model with two threshold variables. Thus, we define the estimator of $\gamma$ as the value that minimizes

$$\widehat{\gamma} = \underset{\gamma}{\operatorname{argmin}}\, S(\gamma); \tag{9}$$

and once $\widehat{\gamma}$ is obtained, the slope coefficient estimate is $\widehat{\beta} = \widehat{\beta}(\widehat{\gamma})$. The residual vector is $e^*(\widehat{\gamma})$ and the residual variance

$$\widehat{\sigma}^2 = \frac{1}{n(T-1)}\widehat{e}^*(\widehat{\gamma})'\widehat{e}^*(\widehat{\gamma}).$$

The criterion function (8) is not smooth, so conventional gradient algorithms are not suitable for its maximization. Hansen (1999, 2000) suggests using a grid search over the threshold variable space in a model with one threshold variable; that is, constructing an evenly spaced grid on the empirical support of the threshold variable. While these estimates might seem desirable in theory, their implementation might be somewhat cumbersome in practice for the case of the model with two threshold variables. In a model with one threshold variable and two threshold parameters, Hansen (1999) argues that the sequential estimation is found to be consistent in the multiple change-point model (Chong, 2003; Bai, 1997); thus, the same logic would seem to apply to the multiple threshold model.

Following these ideas, we suggest estimating model (5) sequentially. The method works as follows. In the first stage, we can consistently estimate one of the threshold parameters, for example $\gamma_1$, across the whole sample; to that end, initially we assume that $\beta_1 = \beta_2$ and $\beta_3 = \beta_4$, which means that we have two regimes instead of four. Then, we obtain $\widehat{\gamma}_1$ by minimizing

$$\widehat{\gamma}_1 = \underset{\gamma_1}{\operatorname{argmin}}\, S(\gamma_1); \tag{10}$$

fixing the first-stage estimate $\widehat{\gamma}_1$, within four subsamples, the second-stage estimate of $\gamma_2$ is

$$\widehat{\gamma}_2^r = \underset{\gamma_2}{\operatorname{argmin}}\, S(\widehat{\gamma}_1, \gamma_2). \tag{11}$$

In the model with time as a threshold variable and two breaks, Bai (1997) has shown that the second break (threshold parameter) estimate is asymptotically efficient, but the first break estimate is not. Hansen (1999) argues that this is because the estimate $\widehat{\gamma}_1$ was obtained from a concentrated function contaminated by the presence of a neglected regime. Hansen (1999) proposes that the asymptotic efficiency of the first threshold estimate can be improved by a third-stage estimation. Thus, by fixing the second-stage estimate $\widehat{\gamma}_2^r$, the refinement first threshold estimate is

$$\widehat{\gamma}_1^r = \underset{\gamma_1}{\operatorname{argmin}}\, S(\gamma_1, \widehat{\gamma}_2). \tag{12}$$

Bai (1997) shows that the refinement estimator $\widehat{\gamma}_1^r$ is asymptotically efficient in the change-point model; Hansen (1999) argues that a similar result is expected to hold in threshold regression with one threshold variable and two threshold parameters. Likewise, we expect a similar results to hold in the model with two threshold variables. Thus, once $\widehat{\gamma}^r = (\widehat{\gamma}_1^r, \widehat{\gamma}_2^r)$ is obtained, the slope coefficient estimate is $\widehat{\beta} = \widehat{\beta}(\widehat{\gamma}^r)$.

## 2.2 Inference

In the context of threshold autoregression estimation, Chan (1993) establishes that the limiting distribution of the threshold parameter estimator converges to a functional of a compound Poisson process at a rate $n$. However, the distribution is too complicated to be used in practice due to the dependence on the nuisance parameters. Hansen (2000) developed an asymptotic distribution for the threshold parameter estimate based on the small threshold effect assumption, in which the threshold model becomes the linear model asymptotically. The limiting distribution converges to a functional of a two-sided Brownian motion process at a rate $n^{1-2\alpha}$ with $0 < \alpha < 1/2$. The distribution does not depend on the nuisance parameters; thus, the distribution can be available in a simple closed form.

Chen et al. (2012) adopt the approach of Hansen (2000) in a time series model with two threshold variables, finding both consistency and that the joint distribution of the least squares estimator $\widehat{\gamma}$ converges to a functional of a two-sided Brownian motion process. We expect that these results will also hold in the static panel data model with two threshold variables since our assumptions meet the asymptotic theory of Hansen (1999) and Chen et al. (2012).

Hansen (2000) argues that the best way to form confidence intervals for the threshold is to form the non-rejection region using the likelihood ratio statistic for testing on $\widehat{\gamma}$. To test hypothesis $H_0 : \gamma = \gamma_0$, the likelihood ratio test is to reject large values of $LR(\gamma_0)$ where

$$LR(\gamma) = n(T-1)\frac{S(\gamma) - S(\widehat{\gamma})}{S(\widehat{\gamma})}. \tag{13}$$

Hansen (1996) shows the $LR(\gamma)$ converges in distribution to $\xi$ as $n \to \infty$, where $\xi$ is a random variable with distribution function $P(\xi \leq z) = (1 - exp(-z/2))^2$. Thus, the asymptotic distribution of the likelihood ratio statistic is non-standard, yet free of nuisance parameters. Since the asymptotic distribution is pivotal, it may be used to form valid asymptotic confidence intervals. Furthermore, the distribution function $\xi$ has the inverse

$$c(a) = -2\ln\left(1 - \sqrt{1-a}\right), \tag{14}$$

where $a$ is the significance level. To form an asymptotic confidence interval for $\gamma$, the non-rejection region of confidence level $1-a$ is the set of values of $\gamma$, such that $LR(\gamma) \leq c(a)$, where $LR(\gamma)$ is defined in (13) and $c(a)$ is defined in (14). The easiest way to find this is by plotting $LR(\gamma)$ against $\gamma$ and drawing a flat line at $c(a)$.

Bai (1997) shows (for the analogous case of change-point models) that the refinement estimators have the same asymptotic distributions as the threshold estimate in a single threshold model. Based on that finding, in a static panel data model with one threshold variable and two threshold parameters, Hansen (1999) suggests that confidence intervals can be constructed in the same way as the threshold estimate in a single threshold model. We expect the same results to hold in our model with two threshold variables and two threshold parameters.

For $\widehat{\gamma}_2^r$, let

$$LR_2^r(\gamma_2) = n(T-1)\frac{S(\widehat{\gamma}_1^r, \gamma_2) - S(\widehat{\gamma}_1^r, \widehat{\gamma}_2^r)}{S(\widehat{\gamma}_1^r, \widehat{\gamma}_2^r)}, \tag{15}$$

and for $\widehat{\gamma}_1^r$, let

$$LR_1^r(\gamma_1) = n(T-1)\frac{S(\gamma_1, \widehat{\gamma}_2^r) - S(\widehat{\gamma}_1^r, \widehat{\gamma}_2^r)}{S(\widehat{\gamma}_1^r, \widehat{\gamma}_2^r)}; \tag{16}$$

then, the asymptotic $(1-a)$ percent confidence intervals for $\gamma_1$ and $\gamma_2$ are the set of values of $\gamma_1$ and $\gamma_2$ such that $LR_2^r(\gamma_1) \leq c(a)$ and $LR_1^r(\gamma_2) \leq c(a)$, respectively.

In relation to the slope parameters $\beta$, as is standard in threshold models, the model (4) conditional on $\gamma$ is linear. Thus, the least squares estimation of $\beta$ is consistent and asymptotically normally distributed, as $n$ tends to infinity and $T$ is fixed, since we assume that the regressors and the threshold variables are exogenous variables.

## 2.3 Testing for the threshold

Following Chen et al. (2012), in order to determine the number of regimes, we first consider the null hypothesis of no threshold effect:

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4; \tag{17}$$

under $H_0$ the thresholds $\gamma_1$ and $\gamma_2$ are not identified, so classical tests have non-standard distributions. The fixed effect equation (4) belongs to the class of models considered by Hansen (1996), who suggests a bootstrap to simulate the asymptotic distribution of the likelihood ratio test. So, under the null hypothesis of no threshold, the model is

$$y_{it} = \mu_i + \beta_1' x_{it} + e_{it}, \tag{18}$$

and after the fixed effect transformation is made, we have

$$y_{it}^* = \beta_1' x_{it}^* + e_{it}^*. \tag{19}$$

The regression parameter $\beta_1$ is estimated by least squares (LS), yielding estimate $\tilde{\beta}_1$, residuals $\tilde{e}_{it}^*$ and sum of squared errors $S_0 = \tilde{e}_{it}^{*\prime} \tilde{e}_{it}^*$. The likelihood ratio test statistics of $H_0$ is defined as

$$F = n(T-1)(S_0 - S(\widehat{\gamma}))/S(\widehat{\gamma}). \tag{20}$$

Rejection of the null hypothesis suggests the existence of more than one regime. Hansen (1999) argues that the asymptotic distribution of $F$ is non-standard, and strictly dominates the $\chi_k^2$ distribution, which appears to depend in general upon moments of the sample and thus critical values cannot be tabulated. Hansen (1996) shows that a bootstrap procedure attains the first-order asymptotic distribution, so $p$-values constructed from the bootstrap are asymptotically valid. As such, the asymptotic distribution can be approximated by the following bootstrap procedure.

Similar to Hansen (1999), take the regression residuals $\widehat{e}_{it}^*$ and group them by individual: $\widehat{e}_i^* = (\widehat{e}_{i1}^*, \widehat{e}_{i2}^*, \ldots, \widehat{e}_{iT}^*)$. Treat the sample $\{\widehat{e}_1^*, \widehat{e}_2^*, \ldots, \widehat{e}_n^*\}$ as the empirical distribution to be used for bootstrapping. Draw (with replacement) a sample of size $n$ from the empirical distribution and use these errors to create a bootstrap sample under $H_0$. Using the bootstrap sample, estimate the model under the null (19) and alternative (4), and calculate the bootstrap value of the likelihood ratio statistic $F$ (20). Repeat this procedure a large number of times to calculate the percentage of draws for which the simulated statistic exceeds the actual value. The null is rejected if this $p$-value is smaller than the desired critical value.

Rejection of the null hypothesis implies the presence of threshold effects. To determine the number of regimes, we follow the general-to-specific approach of Chen et al. (2012).

First, we test a three-regime model against a four-regime model. We test each of the following hypotheses $H_0$: $\beta_a = \beta_b$ for $a \neq b$ and $a, b \in \{1, 2, 3, 4\}$ against the alternative hypothesis $H_1$ corresponding to four regimes. A likelihood ratio test used to test these pairs of hypotheses is

$$F(\widehat{\gamma}) = n(T-1)(S_0(\widehat{\gamma}) - S_1(\widehat{\gamma}))/S_1(\widehat{\gamma}), \tag{21}$$

where $S_0(\widehat{\gamma})$ is the residual sum of squares under $H_0$ and $S_1(\widehat{\gamma})$ is the residual sum of squares under $H_1$. We will draw the bootstrap errors from the residuals calculated under the alternative hypothesis, which should be the residuals from LS estimation under the alternative model (4). The dependent variable should be generated under the null hypothesis which depends on the parameter values $\widehat{\beta}$ and $\widehat{\gamma}$, the LS estimates under the null.

Rejection of each of the null hypotheses $H_0$: $\beta_a = \beta_b$ for $a \neq b$ and $a, b \in \{1, 2, 3, 4\}$ implies the existence of four regimes. If any one of them is accepted, then there are less than four regimes, and following Chen et al. (2012), we proceed to test a two-regime model against a three-regime model. For instance, if $H_0$: $\beta_1 = \beta_2$ is accepted, we test the following three hypotheses $H_0$: $\beta_1 = \beta_2 = \beta_3$, $H_0$: $\beta_1 = \beta_2 = \beta_4$ and $H_0$: $\beta_1 = \beta_2$ and $\beta_3 = \beta_4$. The alternative hypothesis is $H_1$: there are three regimes with $\beta_1 = \beta_2$. If all the above null hypotheses are rejected, we conclude that there are three regimes; otherwise, we conclude that the model has two regimes.

# 3 Dynamic Panel Data Model

Model (1) is a static panel data model. In some applications there may be dynamics; that is, they may allow lagged lagged dependent variables as regressors. The dynamic model with two threshold variables can take the form

$$\begin{aligned} y_{it} \;=\;& \mu_i + \beta_1 y_{it-1} 1(q_{1it} \leq \gamma_1, q_{2it} \leq \gamma_2) + \beta_2 y_{it-1} 1(q_{1it} \leq \gamma_1, q_{2it} > \gamma_2) \\ & + \beta_3 y_{it-1} 1(q_{1it} > \gamma_1, q_{2it} \leq \gamma_2) + \beta_4 y_{it-1} 1(q_{1it} > \gamma_1, q_{2it} > \gamma_2) + e_{it}, \end{aligned} \tag{22}$$

where the threshold variables $q_{it} = (q_{1it}, q_{2it})$ may be elements or functions of the vector $x_{it}$ of exogenous variables and must have continuous distributions. For simplicity, we do not consider other exogenous regressors; the model may be extended to a model with exogenous regressors, but at the cost of a more cumbersome notation.

The previous techniques cannot be used in the dynamic model (22), because any transformation to eliminate the individual fixed specific effect will introduce a correlation between the transformed regressors and the transformed error term in the model. In this context Ramírez-Rondán (2015) develops econometric techniques to estimate a dynamic panel model with one threshold variable; in this section we extend that work using model (22) with two threshold variables.

## 3.1 Estimation

Similar to Ramírez-Rondán (2015), we propose a maximum likelihood approach to estimate the dynamic panel model with two threshold variables; to this end, the error term must be assumed to be independently identically normal distributed with mean 0 and finite variance $\sigma_u^2$. We also assume that the initial values, $y_{i0}$ and $x_{i0}$, are observable.

When the individual specific effects, $\mu_i$, are fixed, the least-squared dummy variable (LSDV) estimator of the linear version of model (22) leads to an inconsistency in the slope parameter estimator as $n$ grows to infinity for a fixed $T$ (Nickell, 1981). If the errors $u_{it}$ are normally distributed, then the LSDV are also the maximum likelihood estimator (MLE), conditional on the initial observation $y_{i0}$, the MLE also leads to an inconsistency in the slope parameter estimator, due to the classical incidental parameter problem in which the number of parameters increases with the number of observations (Lancaster, 2000).

To address the incidental parameter problem we take the first difference to eliminate the individual specific effect in model (22), obtaining

$$\begin{aligned} y_{it} - y_{it-1} \quad = \quad & \beta_1 (y_{it-1} 1(q_{1it} \leq \gamma_1, q_{2it} \leq \gamma_2) - y_{it-2} 1(q_{1it-1} \leq \gamma_1, q_{2it-1} \leq \gamma_2)) \\ & + \beta_2 (y_{it-1} 1(q_{1it} \leq \gamma_1, q_{2it} > \gamma_2) - y_{it-2} 1(q_{1it-1} \leq \gamma_1, q_{2it-1} > \gamma_2)) \\ & + \beta_3 (y_{it-1} 1(q_{1it} > \gamma_1, q_{2it} \leq \gamma_2) - y_{it-2} 1(q_{1it-1} > \gamma_1, q_{2it-1} \leq \gamma_2)) \\ & + \beta_4 (y_{it-1} 1(q_{1it} > \gamma_1, q_{2it} > \gamma_2) - y_{it-2} 1(q_{1it-1} > \gamma_1, q_{2it-1} > \gamma_2)) \\ & + e_{it} - e_{it-1}; \end{aligned} \tag{23}$$

to simplify notation, let $\Delta y_{it} \equiv y_{it} - y_{it-1}$,

$$\Delta y_{it-1}^*(\gamma) = \begin{pmatrix} y_{it-1} 1(q_{1it} \leq \gamma_1, q_{2it} \leq \gamma_2) - y_{it-2} 1(q_{1it-1} \leq \gamma_1, q_{2it-1} \leq \gamma_2) \\ y_{it-1} 1(q_{1it} \leq \gamma_1, q_{2it} > \gamma_2) - y_{it-2} 1(q_{1it-1} \leq \gamma_1, q_{2it-1} > \gamma_2) \\ y_{it-1} 1(q_{1it} > \gamma_1, q_{2it} \leq \gamma_2) - y_{it-2} 1(q_{1it-1} > \gamma_1, q_{2it-1} \leq \gamma_2) \\ y_{it-1} 1(q_{1it} > \gamma_1, q_{2it} > \gamma_2) - y_{it-2} 1(q_{1it-1} > \gamma_1, q_{2it-1} > \gamma_2) \end{pmatrix},$$

and $\Delta e_{it} \equiv e_{it} - e_{it-1}$. Then equation (23) becomes

$$\Delta y_{it} = \beta' \Delta y^*_{it-1}(\gamma) + \Delta e_{it}; \tag{24}$$

also, note that for $t = 2, 3, \ldots, T$, (24) is well defined, but not for $\Delta y_{i1}$ because $\Delta y^*_{i0}(\gamma)$ is missing; that is, $y_{i,-1}$ is not available.

When the time period is fixed, or the panel covers only a short period, the MLE of the dynamic linear panel model depends on the initial condition and the assumption on the initial condition plays a crucial role in devising consistent estimates. Anderson and Hsiao (1981) show the assumptions under which the MLE leads to consistent or inconsistent estimates of the slope parameter. This problem arises because the covariance matrix depends on the initial conditions; if $T$ grows to infinity the initial condition problem disappears.

By continuous substitution of equation (24) for the first period, $\Delta y_{i1}$, the resulting equation has an intractable form and depends on the structural parameters. Moreover, it is clear that equation (24) does not depend on the individual specific fixed effect for all $t$. Thus, to address the initial condition problem, we assume the process has started from a finite period in the past, namely for given values of $y_{i,-1}$ such that

$$E(\Delta y_{i1} | x_i) = \delta,$$

where $x_i = (x_{i0}, x_{i1}, ..., x_{iT})'$. This assumption imposes the restriction that the expected changes in the initial endowments are the same across all individuals, though this does not necessarily require that the process has reached stationary. It is important to note that the "small threshold effect" assumption specifies that the difference in slopes decreases as the sample size increases, i.e. the threshold model becomes the linear model as the sample size grows. Thus, $E(\Delta y_{i1}|x_i) = \delta$ can be seen as an approximation for a large number of individuals.

Hence, the marginal distribution of $\Delta y_{i1}$ conditional on $x_i$ can be written as

$$\Delta y_{i1} = \delta + \upsilon_{i1}, \tag{25}$$

where $\upsilon_{i1}$ is the error term in the first period. Under the exogeneity of $x_{it}$ and by construction, $E(\upsilon_{i1}|x_i) = 0$, $E\upsilon_{i1}^2 = \sigma_\upsilon^2$, and we assume $Cov(\upsilon_{i1}, \Delta e_{i2}|x_{it}) = -\sigma_u^2$ and $Cov(\upsilon_{i1}, \Delta e_{it}|x_{it}) = 0$ for $t = 3, ..., T$, $i = 1, ..., n$; that is, we assume homoscedasticity across regimes. The auxiliary external parameter, $\delta$, can be a function of the structural parameters, but similarly to Hsiao et al. (2002) we treat the external parameters as free parameters in the sense they do not depend on the structural parameters.

**Maximum Likelihood Function**

Let $\Delta y_i = (\Delta y_{i1}, \Delta y_{i2}, ..., \Delta y_{iT})'$ and $\Delta e_i = (\upsilon_{i1}, \Delta e_{i2}, ..., \Delta e_{iT})'$. The Jacobian of the transformation from $\Delta e_i$ to $\Delta y_i$ is unity and the joint probability distribution function of $\Delta y_i$ and $\Delta e_i$ are therefore the same. The covariance matrix of $\Delta e_i$ has the form

$$\Omega = \sigma_u^2 \begin{bmatrix} \omega & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & & \\ 0 & -1 & 2 & & \\ \vdots & & & \ddots & -1 \\ 0 & & & -1 & 2 \end{bmatrix} = \sigma_u^2 \Omega^\star, \tag{26}$$

where $\omega = \sigma_\upsilon^2/\sigma_u^2$.

Let $\beta_\delta = (\delta, \beta')'$ and define the matrix $\Delta y_{i,-1}(\gamma)$ as follows

$$\Delta y_{i,-1}(\gamma) = \begin{bmatrix} 1 & 0 \\ 0 & \Delta y_{i1}^*(\gamma) \\ 0 & \Delta y_{i2}^*(\gamma) \\ \vdots & \vdots \\ 0 & \Delta y_{iT-1}^*(\gamma) \end{bmatrix};$$

and, under the assumption that $e_{it}$ is independent normal, the joint probability distribution function of $\Delta y_i$ conditional on $x_i$ is given by

$$\begin{aligned} \ln L(\beta_\delta, \gamma, \sigma_u^2, \omega) = & -\frac{nT}{2}\ln(2\pi) - \frac{n}{2}\ln|\Omega(\gamma)| \\ & - \frac{1}{2}\sum_{i=1}^{n}[(\Delta y_i - \Delta y_{i,-1}(\gamma)\beta_\delta)'\Omega^{-1}(\Delta y_i - \Delta y_{i,-1}(\gamma)\beta_\delta)]. \end{aligned} \tag{27}$$

The likelihood function (27) is well defined, depends on a fixed number of parameters. The only unknown element of $\Omega^\star$ is $\omega$ and it can be shown that $|\Omega| = \sigma_u^{2T}[1 + T(\omega - 1)]$ (see Hsiao et al., 2002). For this maximization, note that $\gamma$ is assumed to be restricted to a bounded set $\Gamma = [\underline{\gamma}; \overline{\gamma}] = [\underline{\gamma_1}, \overline{\gamma_1}] \times [\underline{\gamma_2}, \overline{\gamma_2}]$. Note that since this set is also closed, it is compact on $\mathbb{R}^2$. Then, the MLE $(\widehat{\delta}, \widehat{\beta}, \widehat{\gamma}, \widehat{\sigma}_u^2, \widehat{\omega})$ are the values that globally maximize $\ln L(\delta, \beta, \gamma, \sigma_u^2, \omega)$.

### ML Estimators of $\boldsymbol{\delta}$, $\boldsymbol{\beta}$, $\boldsymbol{\sigma_u^2}$ and $\boldsymbol{\omega}$ for a given $\boldsymbol{\gamma}$

We start the estimation procedure by considering a fixed $\gamma$. Then, for a given $\gamma$, taking the first-order partial derivatives with respect to $\beta_\delta$, $\sigma_u^2$ and $\omega$ and setting the partial derivatives equal to zero gives[2]

$$\widehat{\beta}_\delta(\gamma) = \left( \sum_{i=1}^{n} \Delta y_{i,-1}(\gamma)' \widehat{\Omega}^\star(\gamma)^{-1} \Delta y_{i,-1}(\gamma) \right)^{-1} \left( \sum_{i=1}^{n} \Delta y_{i,-1}(\gamma)' \widehat{\Omega}^\star(\gamma)^{-1} \Delta y_i \right), \tag{28}$$

$$\widehat{\sigma}_u^2(\gamma) = \frac{1}{nT} \sum_{i=1}^{n} [(\Delta y_i - \Delta y_{i,-1}(\gamma)\widehat{\beta}_\delta(\gamma))' \widehat{\Omega}^\star(\gamma)^{-1} (\Delta y_i - \Delta y_{i,-1}(\gamma)\widehat{\beta}_\delta(\gamma))], \tag{29}$$

and

$$\widehat{\omega}(\gamma) = \frac{T-1}{T} + \frac{1}{\widehat{\sigma}_u^2(\gamma) n T^2} \sum_{i=1}^{n} [(\Delta y_i - \Delta y_{i,-1}(\gamma)\widehat{\beta}_\delta(\gamma))' \Phi (\Delta y_i - \Delta y_{i,-1}(\gamma)\widehat{\beta}_\delta(\gamma))], \tag{30}$$

where

$$\Phi = \begin{bmatrix} T^2 & T(T-1) & T(T-2) & \dots & T \\ T(T-1) & (T-1)^2 & (T-1)(T-2) & \dots & (T-1) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ T & (T-1) & (T-2) & \dots & 1 \end{bmatrix}.$$

Note that the ML slope estimators depend on $\sigma_u^2$ and $\omega$, and these in turn depend on the slope parameters; thus, in order to compute the MLE $\widehat{\delta}(\gamma)$, $\widehat{\beta}(\gamma)$, $\widehat{\sigma}_u^2(\gamma)$ and $\widehat{\omega}(\gamma)$ for a given $\gamma$, we can use a grid search procedure whereby the MLE are computed for a number of values of $\omega(\gamma) > 1 - 1/T$ at a given $\gamma$, and then choosing that value of $\omega(\gamma)$, which globally maximizes the log-likelihood function (27).

### ML Estimator for the Threshold Parameter $\boldsymbol{\gamma}$

The ML estimators for a given $\gamma$ are $\widehat{\beta}_\delta(\gamma) = (\widehat{\delta}(\gamma), \widehat{\beta}(\gamma))$, $\widehat{\sigma}_u^2(\gamma)$ and $\widehat{\omega}(\gamma)$. Therefore the threshold parameters, $\gamma = (\gamma_1, \gamma_2)$, are estimated by maximizing the concentrated log-likelihood function (31),

[2] The formulas for the first-order partial derivatives with respect to $\beta_\delta$, $\sigma_u^2$ and $\omega$ are similar to Ramírez-Rondán (2015).

$$
\begin{aligned}
\ln L(\gamma) &= -\frac{nT}{2}\ln(2\pi) - \frac{n}{2}\ln|\widehat{\Omega}(\gamma)| \\
&\quad -\frac{1}{2}\sum_{i=1}^{n}(\Delta y_i - \Delta y_{i,-1}(\gamma)\widehat{\beta}_\delta(\gamma))'\widehat{\Omega}(\gamma)^{-1}(\Delta y_i - \Delta y_{i,-1}(\gamma)\widehat{\beta}_\delta(\gamma)) \\
&= -\frac{nT}{2}\ln(2\pi) - \frac{n}{2}\ln|\widehat{\Omega}(\gamma)| - \frac{1}{2}\sum_{i=1}^{n}\Delta\widehat{e}_i(\gamma)'\widehat{\Omega}(\gamma)^{-1}\Delta\widehat{e}_i(\gamma). \qquad (31)
\end{aligned}
$$

The criterion function (31) is not smooth; therefore, similarly to the static model, we also suggest estimating the model (22) sequentially. The method works as follows. In the first stage, we first consistently estimate one of the threshold parameters, for example $\gamma_1$, across the whole sample; to this end we initially assume that $\beta_1 = \beta_2$ and $\beta_3 = \beta_4$ (or set $\gamma_2$ at any value), which means that we have two regimes instead of four. Thus, we obtain $\widehat{\gamma}_1$ by maximizing

$$
\widehat{\gamma}_1 = \underset{\gamma_1}{\operatorname{argmax}} \ln L(\gamma_1); \qquad (32)
$$

fixing the first-stage estimate $\widehat{\gamma}_1$, within both subsamples, the refinement second-stage estimate of $\gamma_2$ is

$$
\widehat{\gamma}_2^r = \underset{\gamma_2}{\operatorname{argmax}} \ln L(\widehat{\gamma}_1, \gamma_2); \qquad (33)
$$

similarly, fixing the second-stage estimate $\widehat{\gamma}_2^r$, the refinement first threshold estimate is

$$
\widehat{\gamma}_1^r = \underset{\gamma_1}{\operatorname{argmax}} \ln L(\gamma_1, \widehat{\gamma}_2). \qquad (34)
$$

### ML Estimators for the Slope Parameters $\boldsymbol{\beta}$

Once $\widehat{\gamma} = (\widehat{\gamma}_1, \widehat{\gamma}_2)$ are obtained by maximizing (31), the ML estimators of slope parameters are $\widehat{\beta} = \widehat{\beta}(\widehat{\gamma})$. That is, once the refinement estimator $\widehat{\gamma}^r = (\widehat{\gamma}_1^r, \widehat{\gamma}_2^r)$ are obtained, the slope coefficient estimates are $\widehat{\beta} = \widehat{\beta}(\widehat{\gamma}^r)$. Also, the ML estimators of the remaining parameters that involve the estimation method are $\widehat{\delta} = \widehat{\delta}(\widehat{\gamma}^r)$, $\widehat{\sigma}_u^2 = \widehat{\sigma}_u^2(\widehat{\gamma}^r)$ and $\widehat{\omega} = \widehat{\omega}(\widehat{\gamma}^r)$.

The estimated covariance matrix for the ML slope estimators $\widehat{\beta}_\delta = (\widehat{\delta}, \widehat{\beta}')'$ is

$$\begin{aligned}\operatorname{Cov}\left[\begin{array}{c}\widehat{\delta}\\ \widehat{\beta}\end{array}\right] &= \left(\sum_{i=1}^{n}\Delta y_{i,-1}(\widehat{\gamma})'\widehat{\Omega}^{-1}\Delta y_{i,-1}(\widehat{\gamma})\right)^{-1}\\ &= \widehat{\sigma}_u^2\left(\sum_{i=1}^{n}\Delta y_{i,-1}(\widehat{\gamma})'\widehat{\Omega}^{\star-1}\Delta y_{i,-1}(\widehat{\gamma})\right)^{-1};\end{aligned}$$

or, under a suitable partition of the matrix $\Delta y_{i,-1}(\widehat{\gamma})$, the estimated covariance matrix for the ML slope estimators $\widehat{\beta}=(\widehat{\beta}_1,\widehat{\beta}_2,\widehat{\beta}_3,\widehat{\beta}_4)'$ is

$$\operatorname{Cov}\left(\ \widehat{\beta}\ \right)=\widehat{\sigma}_u^2\left(\sum_{i=1}^{n}\Delta y^{\circ}_{i,-1}(\widehat{\gamma})'\widehat{\Omega}^{\star-1}\Delta y^{\circ}_{i,-1}(\widehat{\gamma})\right)^{-1},$$

where

$$\Delta y^{\circ}_{i,-1}(\gamma)=\left[\begin{array}{c}0\\ \Delta y^{*}_{i1}(\gamma)\\ \Delta y^{*}_{i2}(\gamma)\\ \vdots\\ \Delta y^{*}_{iT-1}(\gamma)\end{array}\right].$$

## 3.2 Inference

Similar to the static model, in the dynamic panel data model we can construct confidence interval for $\gamma_1$ and $\gamma_2$ as follows. For $\widehat{\gamma}_2^r$ let

$$LR_2^r(\gamma_2)=nT\frac{S(\widehat{\gamma}_1^r,\gamma_2)-S(\widehat{\gamma}_1^r,\widehat{\gamma}_2^r)}{S(\widehat{\gamma}_1^r,\widehat{\gamma}_2^r)}, \tag{35}$$

where $S(\gamma)=\sum_{i=1}^{n}\Delta\widehat{e}_i(\gamma)'\widehat{\Omega}^{\star-1}\Delta\widehat{e}_i(\gamma)$, that is called the minimum distance estimator. For $\widehat{\gamma}_1^r$ let

$$LR_1^r(\gamma_1)=nT\frac{S(\gamma_1,\widehat{\gamma}_2^r)-S(\widehat{\gamma}_1^r,\widehat{\gamma}_2^r)}{S(\widehat{\gamma}_1^r,\widehat{\gamma}_2^r)}, \tag{36}$$

and the asymptotic $(1-a)$ percent confidence intervals for $\gamma_1$ and $\gamma_2$ are the set of values of $\gamma_1$ and $\gamma_2$ such that $LR_2^r(\gamma_1)\leq c(a)$ and $LR_1^r(\gamma_2)\leq c(a)$, respectively.

In relation to the slope parameters, $\beta$, the likelihood function (31) is well defined; it depends on a fixed number of parameters, and satisfies the usual regularity conditions conditional on $\gamma$. Therefore, the MLE of $\beta$ is consistent and asymptotically normally distributed, as $n$ tends to infinity and $T$ is fixed.

As regards consistency of the maximum likelihood threshold estimates, the proposed dynamic panel model meets the assumptions of the asymptotic theory developed by Hansen (2000), Chen et al. (2012) and Ramírez-Rondán (2015); thus, by following similar arguments of aforementioned studies, the consistency of the threshold estimates can be proven.

### 3.3 Testing for the threshold

Due to the time dependence of the dynamic model, following Kapetanios (2008) we consider resampling with replacement only in the cross-sectional dimension, as in the static model case. The parametric bootstrap can be implemented by resampling the "pairs" $(\Delta y_i, \Delta y_{i-})$, where $\Delta y_i = (\Delta y_{i1}, \Delta y_{i2}, ..., \Delta y_{iT})'$ and $\Delta y_{i-} = (0, \Delta y_{i1}, ..., \Delta y_{iT-1})'$, which can be implemented similarly with the residuals of $\Delta e_i = (\upsilon_{i1}, \Delta e_{i2}, ..., \Delta e_{iT})'$.

Thus, similar to test for the threshold in the static model, we first consider the null hypothesis of no threshold effect:

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4, \tag{37}$$

under $H_0$ the thresholds $\gamma_1$ and $\gamma_2$ are not identified, so classical tests have non-standard distributions. The first difference equation (24) fall in the class of models considered by Hansen (1996) who suggested a bootstrap to simulate the asymptotic distribution of the likelihood ratio test. Under the null hypothesis of no threshold, the model is

$$y_{it} = \mu_i + \beta_1 y_{it-1} + e_{it}, \tag{38}$$

after the first difference transformation is made, we have

$$\Delta y_{it} = \beta_1 \Delta y_{it-1} + \Delta e_{it}, \tag{39}$$

and the model in the first period is $\Delta y_1 = \delta + \upsilon_{i1}$.[3]

The regression parameters $\beta_1$ and $\delta$ are estimated by maximum likelihood (ML), yielding estimate $\tilde{\beta}_1$, $\tilde{\delta}$, $\tilde{\omega}$, residuals $\Delta\tilde{e}_i = (\tilde{\upsilon}_{i1}, \Delta\tilde{e}_{i2}, ..., \Delta\tilde{e}_{iT})'$ and the minimum distance estimator $S_0 = \Delta\tilde{e}_i^{*\prime}\tilde{\Omega}^{\star-1}\Delta\tilde{e}_i^{*}$. Hence, the likelihood ratio test statistics of $H_0$ is defined as

$$F = n(T-1)(S_0 - S(\widehat{\gamma}))/S(\widehat{\gamma}). \tag{40}$$

[3]Note that the equation in the first period is needed for consistency of the ML estimators (for the dynamic linear panel model estimation see Hsiao et al., 2002).

Rejection of the null hypothesis suggests the existence of more than one regime. Similar to the static model, take the regression residuals $\widehat{v}_{i1}$ and $\Delta\widehat{e}_{it}$ and group them by individual: $\Delta\widehat{e}_i = (\widehat{v}_{i1}, \Delta\widehat{e}_{i2}, \ldots, \Delta\widehat{e}_{iT})$. Treat the sample $\{\Delta\widehat{e}_1, \Delta\widehat{e}_2, \ldots, \Delta\widehat{e}_n\}$ as the empirical distribution to be used for bootstrapping. Draw (with replacement) a sample of size $n$ from the empirical distribution and use these errors to create a bootstrap sample under $H_0$. Using the bootstrap sample, estimate the model under the null (39) and alternative (24) and calculate the bootstrap value of the likelihood ratio statistic $F$ (40).

Rejection of the null hypothesis implies the presence of threshold effects. To determine the number of regimes, we follow again the general-to-specific approach of Chen et al. (2012). First, we test a three-regime model against a four-regime model. We test each of the following hypotheses $H_0$: $\beta_a = \beta_b$ for $a \neq b$ and $a, b \in \{1, 2, 3, 4\}$ against the alternative hypothesis $H_1$: there are four regimes. A likelihood ratio test used to test these pairs of hypotheses is

$$F(\widehat{\gamma}) = n(T-1)(S_0(\widehat{\gamma}) - S_1(\widehat{\gamma}))/S_1(\widehat{\gamma}), \tag{41}$$

where $S_0(\widehat{\gamma})$ is the minimum distance estimator under $H_0$ and $S_1(\widehat{\gamma})$ is the minimum distance estimator under $H_1$.

We will draw the bootstrap errors from the residuals calculated under the alternative hypothesis, which should be the residuals from ML estimation under the alternative model (24). The dependent variable should be generated under the null hypothesis which depends on the parameter values $\widehat{\beta}$, $\widehat{\delta}$, $\widehat{\omega}$ and $\widehat{\gamma}$, the ML estimates under the null. The next steps are the same as the procedure in the static panel model.

# 4 Monte Carlo Experiments

## 4.1 Models

We use two models: the first model is a static panel data and the second model is a dynamic panel data. Moreover, we construct an exogenous variable as a function of the two threshold variables. Then we use the following models to generate $y_{it}$,

$$\begin{aligned} y_{it} &= \mu_i + \beta_1 x_{it} 1(q_{1it} \leq \gamma_1, q_{2it} \leq \gamma_2) + \beta_2 x_{it} 1(q_{1it} \leq \gamma_1, q_{2it} > \gamma_2) \\ &\quad + \beta_3 x_{it} 1(q_{1it} > \gamma_1, q_{2it} \leq \gamma_2) + \beta_4 x_{it} 1(q_{1it} > \gamma_1, q_{2it} > \gamma_2) + e_{it}, \end{aligned} \tag{42}$$

and

$$
\begin{aligned}
y_{it} &= \mu_i + \beta_1 y_{it-1} 1(q_{1it} \le \gamma_1, q_{2it} \le \gamma_2) + \beta_2 y_{it-1} 1(q_{1it} \le \gamma_1, q_{2it} > \gamma_2) \\
&\quad + \beta_3 y_{it-1} 1(q_{1it} > \gamma_1, q_{2it} \le \gamma_2) + \beta_4 y_{it-1} 1(q_{1it} > \gamma_1, q_{2it} > \gamma_2) + e_{it}. \qquad (43)
\end{aligned}
$$

We generate the threshold variables as $q_{1it} \sim N(1/2, 1)$ and $q_{2it} \sim N(3/2, 1) + 0.3q_{1it}$, the exogenous variable as $x_{it} \sim N(-1/2, 1) + 0.4q_{1it} - 0.3q_{2it}$, and the error term as $e_{it} \sim N(0, 1)$. The variable $y_{it}$ in the static model follows equation (42); while in the dynamic models, $y_{it}$ follows equation (43), and it is generated from $t = -10$ to $t = T$, and then we discard the first 10 observations by using the observations $t = 0$ through $T$ for estimation (we also set $y_{i,-10} = 0$).

## 4.2 Individual Fixed Effect Construction

For each model, we consider two designs to construct the individual fixed effect correlated with the exogenous threshold variable; each design considers different sets of the structural parameters. The two designs for generating $\mu_i$ ensures that the random effects slope estimates are inconsistent due to the correlation that exists between the individual specific effects and the explanatory variables $q_{it}$.

### Design 1

The individual effects, $\mu_i$, are generated for the static panel model as

$$\mu_i = u_i + T^{-1} \sum_{t=1}^{T} q_{2it}, \qquad u_i \sim N(2, 3),$$

while for the dynamic panel model as

$$\mu_i = u_i + (T + 11)^{-1} \sum_{t=-10}^{T} q_{2it}, \qquad u_i \sim N(2, 3),$$

and we consider these structural parameters ($\gamma_1$,$\gamma_2$, $\beta_1$, $\beta_2$, $\beta_3$, $\beta_4$)=(0, 2, 0.5, 1.5, -0.8,-2).

### Design 2

The individual effects, $\mu_i$, are generated for the static panel model as

$$\mu_i = u_i + T^{-1} \sum_{t=1}^{T} [-0.7x_{it} 1(q_{1it} \le \gamma_1) + 0.4x_{it} 1(q_{2it} > \gamma_2)], \qquad u_i \sim N(2, 3),$$

while for the dynamic panel model as

$$\mu_i = u_i + (T+11)^{-1} \sum_{t=-10}^{T} [-0.7x_{it}1(q_{1it} \leq \gamma_1) + 0.4x_{it}1(q_{2it} > \gamma_2)], \quad u_i \sim N(2,3),$$

and we consider these structural parameters ($\gamma_1$,$\gamma_2$, $\beta_1$, $\beta_2$, $\beta_3$, $\beta_4$)=(-0.5, 1, -0.3, -1.2, -0.7, 1.2).

### 4.3 Simulation results

Table 1 presents the performance of the estimators. The table shows the bias and root mean square error of the estimators $\widehat{\gamma}_1$, $\widehat{\gamma}_2$, $\widehat{\beta}_1$, $\widehat{\beta}_2$, $\widehat{\beta}_3$ and $\widehat{\beta}_4$ for different choices of numbers of individuals $n$ and a fixed time period $T = 5$. It can be seen that as $n$ increases, the bias of the threshold parameter estimates $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$ decreases quickly; this is consistent with what is found in threshold models, whereby the threshold parameter converges faster (at rate $n$) than the slope parameters (at rate $n^{1/2}$) to the true parameters. Moreover, the bias of the slope parameters $\widehat{\beta}_1$, $\widehat{\beta}_2$, $\widehat{\beta}_3$ and $\widehat{\beta}_4$ decreases.

Similarly, this table shows in general that as the number of individuals $n$ increases, the root mean square error (MRSE) of all parameter estimates decreases. Note that this measure considers the second moments of the data; that is, the RMSE combines bias and efficiency.

## 5 Investment and financing constraints

We apply the methodology to study the relationship between investment and financial constraints. We follow the tradition of papers that investigate this relationship in a context of panel data threshold models (Hansen, 1999; Gonzáles et al., 2005; Seo and Shin, 2016). This literature states that If there are financial constraints, firm investment decisions are not independent of fluctuations in internal finance such as cash flow.

Empirical studies divide firms into constrained and unconstrained groups based on a variable that measures capital market imperfections or financial constraints. Fazzari et al. (1988) argue that cash flow and investment are positively related only when a firm faces constraints on external financing. They use the dividend to income ratio to divide constrained and unconstrained firms. Since the variable dividend payments is treated as a decision variable in Fazzari et al. (1988), it is an endogenous variable. The threshold variable should be an exogenous indicator of a firm's access to external financing. Hansen

Table 1: Performance of estimators

| | | Bias of estimators | | Root mean square error | |
|---|---|---|---|---|---|
| | | $T=5$ | | $T=5$ | |
| Model | Coefficient | $n=50$ | $n=500$ | $n=50$ | $n=500$ |
| Model (42) | $\gamma_1 =$ 0.0 | -0.009 | -0.006 | 0.031 | 0.002 |
| Design 1 | $\gamma_2 =$ 2.0 | -0.006 | 0.000 | 0.060 | 0.005 |
| | $\beta_1 =$ 0.5 | 0.001 | 0.001 | 0.120 | 0.040 |
| | $\beta_2 =$ 1.5 | 0.005 | 0.001 | 0.177 | 0.054 |
| | $\beta_3 =$ -0.8 | -0.007 | 0.002 | 0.120 | 0.038 |
| | $\beta_4 =$ -2.0 | -0.017 | -0.001 | 0.121 | 0.037 |
| Model (42) | $\gamma_1 =$ -0.5 | -0.012 | 0.007 | 0.046 | 0.006 |
| Design 2 | $\gamma_2 =$ 1.0 | -0.006 | -0.001 | 0.044 | 0.005 |
| | $\beta_1 =$ -0.3 | 0.014 | 0.013 | 0.236 | 0.067 |
| | $\beta_2 =$ -1.2 | -0.004 | 0.002 | 0.161 | 0.048 |
| | $\beta_3 =$ -0.7 | 0.017 | -0.005 | 0.162 | 0.050 |
| | $\beta_4 =$ 1.2 | -0.004 | -0.001 | 0.086 | 0.029 |
| Model (43) | $\gamma_1 =$ 0.0 | 0.006 | 0.002 | 0.209 | 0.078 |
| Design 1 | $\gamma_2 =$ 2.0 | -0.012 | 0.005 | 0.254 | 0.127 |
| | $\beta_1 =$ 0.5 | -0.016 | -0.005 | 0.165 | 0.066 |
| | $\beta_2 =$ 1.5 | -0.126 | -0.027 | 0.570 | 0.194 |
| | $\beta_3 =$ -0.8 | -0.005 | -0.005 | 0.332 | 0.071 |
| | $\beta_4 =$ -2.0 | 0.013 | 0.005 | 0.161 | 0.081 |
| Model (43) | $\gamma_1 =$ -0.5 | 0.132 | -0.002 | 0.517 | 0.080 |
| Design 2 | $\gamma_2 =$ 1.0 | 0.074 | 0.002 | 0.420 | 0.051 |
| | $\beta_1 =$ -0.3 | -0.039 | -0.007 | 0.273 | 0.037 |
| | $\beta_2 =$ -1.2 | 0.342 | 0.023 | 0.777 | 0.112 |
| | $\beta_3 =$ -0.7 | 0.173 | 0.015 | 0.477 | 0.072 |
| | $\beta_4 =$ 1.2 | -0.064 | -0.015 | 0.195 | 0.044 |

Note: All results are based on 1000 replications.

(1999) and Hu and Schiantarelli (1998) argue that a natural candidate is debt level, since banks will be reluctant to lend money to debt-heavy firms.

Balance sheet conditions or growth opportunities is another important variable in distinguishing between constrained and unconstrained firms; Gonzáles et al. (2005) argue that in firms with ample growth opportunities, internal and external finance are substitutes, so their investment decisions are independent of their financial structure. However, firms with high information cost and limited growth opportunities face much higher costs of external finance, so their investment decisions are more sensitive to cash flow. A variable that relates growth opportunities is Tobin's Q; Audretsch and Elston (2002) and Kadapakkam et al. (1998) use firm size as an approximation; Gilchrist and Himmelberg (1995) construct a proxy called fundamental Q; Gonzáles et al. (2005) use total market value to assets; and Hubbard (1998) and Lang et al. (1996) use a measure of expected present value of future profits.

Thus, the distinction between constrained and unconstrained firms is based on a variable that measures a degree of information asymmetry or on a variable that measures growth opportunities. In this paper we consider both of these financial constraints on the basis of which firms are divided into four regimes; thus, we consider the debt to assets ratio as a measure of the degree of information, and the total market value to assets ratio as a measure of growth opportunities.[4]

Thus, we estimate a variation of Hansen (1999)'s model with four regimes,

$$\begin{aligned} I_{it} \;=\;& \mu_i + \theta_1 Q_{it} + \theta_2 Q_{it}^2 + \theta_3 Q_{it}^3 + \theta_4 D_{it} + \theta_5 Q_{it} D_{it} \\ &+\beta_1 CF_{it} 1(D_{it} \leq \gamma_1, Q_{it} \leq \gamma_2) + \beta_2 CF_{it} 1(D_{it} \leq \gamma_1, Q_{it} > \gamma_2) \\ &+\beta_3 CF_{it} 1(D_{it} > \gamma_1, Q_{it} \geq \gamma_2) + \beta_4 CF_{it} 1(D_{it} > \gamma_1, Q_{it} > \gamma_2) + e_{it}, \end{aligned} \tag{44}$$

where $I_{it}$ is the investment to assets ratio, $Q_{it}$ is the total market value to assets ratio, $D_{it}$ is the long-term debt to assets ratio; $CF_{it}$ is the cash flow to assets ratio, $i$ indexes firms and $t$ indexes time. Following Hansen (1999), we use a balance panel data of 565 U.S. firms from 1973 to 1987.

While model (44) is a static panel data model, most economic models also exhibit dynamics; lagged investment captures the accelerator effect of investment, whereby past investments have a positive effect on future investments (Aivazian et al., 2005). Thus, we use a dynamic model by adding the lagged dependent variable as a regressor in model (44) as follows

$$\begin{aligned} I_{it} \;=\;& \mu_i + \theta_1 Q_{it} + \theta_2 Q_{it}^2 + \theta_3 Q_{it}^3 + \theta_4 D_{it} + \theta_5 Q_{it} D_{it} + \theta_6 I_{it-1} \\ &+\beta_1 CF_{it} 1(D_{it} \leq \gamma_1, Q_{it} \leq \gamma_2) + \beta_2 CF_{it} 1(D_{it} \leq \gamma_1, Q_{it} > \gamma_2) \\ &+\beta_3 CF_{it} 1(D_{it} > \gamma_1, Q_{it} \geq \gamma_2) + \beta_4 CF_{it} 1(D_{it} > \gamma_1, Q_{it} > \gamma_2) + e_{it}. \end{aligned} \tag{45}$$

[4] Other variables used in the literature to divide constrained and unconstrained firms are dividends, inventory investment, capital intensity, interest payment to income ratio, and liquid assets to capital stock ratio, among others.

First, in order to determine the number of regimes, we perform several tests, allowing for one, two, three, and four regimes. We find that the test that rejects the null for one regime or no threshold effects $F$ is highly significant, with a bootstrap $p$-value of 0.007 and 0.003 for the static and dynamic panel threshold models, respectively.

The test statistics $F(\widehat{\gamma})$ for different null hypothesis along with their bootstrap[5] $p$-values are shown in Table 2. When we perform the null for three regimes, we reject the null in all cases, except when $\beta_1 = \beta_4$ since this statistic has a $p$-value larger than 0.1. Next, we proceed to test the null of two regimes against the alternative for three regimes with $\beta_1 = \beta_4$; the results indicate that we cannot reject the null in all cases, which means that the investment can be approximated using a three-regime model.

Table 2: Tests for number of regimes

| $H_1$: 4 regimes<br>$H_0$: 3 regimes | | $\beta_1 \neq \beta_2 \neq \beta_3 \neq \beta_4$<br>$\beta_1 = \beta_2$ | <br>$\beta_1 = \beta_3$ | <br>$\beta_1 = \beta_4$ | <br>$\beta_2 = \beta_3$ | <br>$\beta_2 = \beta_4$ | <br>$\beta_3 = \beta_4$ |
|---|---|---|---|---|---|---|---|
| Static model | $F(\widehat{\gamma})$ | 31.3 | 18.5 | 0.6 | 42.2 | 19.5 | 24.0 |
| | $p$-value | 0.000 | 0.007 | 0.503 | 0.000 | 0.013 | 0.007 |
| Dynamic model | $F(\widehat{\gamma})$ | 25.8 | 15.3 | 1.2 | 35.1 | 13.8 | 26.0 |
| | $p$-value | 0.000 | 0.013 | 0.280 | 0.000 | 0.010 | 0.010 |

| $H_1$: 3 regimes<br>$H_0$: 2 regimes | | $\beta_1 = \beta_4 \neq \beta_2 \neq \beta_3$<br>$\beta_1 = \beta_4 = \beta_2$ | <br>$\beta_1 = \beta_4 = \beta_3$ | <br>$\beta_1 = \beta_4$ and $\beta_2 = \beta_3$ |
|---|---|---|---|---|
| Static model | $F(\widehat{\gamma})$ | 36.1 | 36.4 | 43.3 |
| | $p$-value | 0.000 | 0.003 | 0.000 |
| Dynamic model | $F(\widehat{\gamma})$ | 27.1 | 34.0 | 34.6 |
| | $p$-value | 0.000 | 0.007 | 0.003 |

The point estimate of the thresholds and their asymptotic 95 percent confidence intervals are reported in Table 3. The estimate of the threshold level of debt and threshold level of market value are 0.012 and 3.035 in the static panel model, and the corresponding estimates for the dynamic panel model are 0.012 and 2.968.

Thus, the four classes of regimes indicated by the point estimates are those with "low debt and limited growth opportunities" for debt lower than 0.012 percent and market value lower than 3.035 (or 2.968); a "low debt and ample growth opportunities" for debt lower that 0.012 and market value higher than 3.035 (or 2.968); a "high debt and limited growth opportunities" for debt higher that 0.012 and market value lower than 3.035 (or 2.968), and a "high debt and ample growth opportunities" for debt higher that 0.012 and market value higher than 3.035 (or 2.968). The asymptotic confidence intervals for the

[5] 300 bootstrap replications were used for each of the bootstrap tests.

threshold level of market value are not tight, indicating considerable uncertainty about the nature of this division.

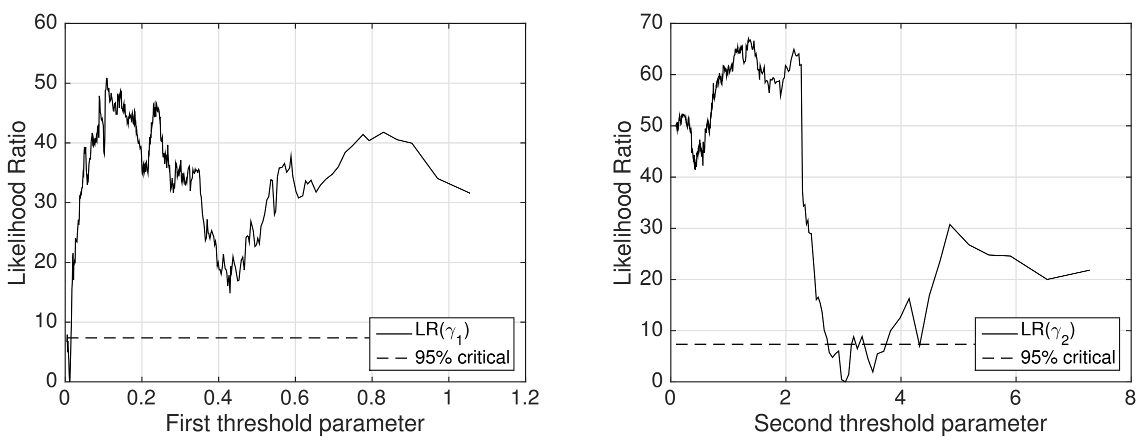
More information can be discerned about the threshold estimates from plots of the concentrated likelihood ratio function $LR(\gamma)$. Figures 1 and 2 show the likelihood ratio function, which is computed when estimating a threshold model. The threshold estimates are the point where the $LR(\gamma)$ equals zero, which occur at $\widehat{\gamma}_1 = 0.012$ and $\widehat{\gamma}_2 = 3.035$, and at $\widehat{\gamma}_1 = 0.012$ and $\widehat{\gamma}_2 = 2.968$ in the static and dynamic panel models, respectively.

Table 3: Asymptotic confidence interval in threshold models

| | Static model | | Dynamic model | |
|---|---|---|---|---|
| | Threshold estimate (%) | 95% confidence interval | Threshold estimate (%) | 95% confidence interval |
| $\widehat{\gamma}_1$ | 0.012 | [0.004, 0.014] | 0.012 | [0.004, 0.016] |
| $\widehat{\gamma}_2$ | 3.035 | [2.754, 4.324] | 2.968 | [2.754, 3.590] |

Note: Asymptotic critical values are reported in Hansen (2000).

Figure 1: Confidence interval construction for thresholds in static model




The coefficients of primary interest are those expressing the interaction between investment and cash flow. The point estimates in Table 4 suggest that the investment decisions of firms under the "high debt and limited growth opportunities" - that is, firms with a high degree of financial constraints - are the most sensitive to internal finance in comparison with the other regimes. In contrast, for firms in a regime of "low debt and ample growth opportunities" - that is, firms that face no financial restrictions - cash flow is not related to investment decisions. Note that cash flow coefficient is significant in the other regime ("high debt and ample growth opportunities" or "low debt and limited growth opportunities") but lower than in the "high debt and limited growth opportunities" regime.

Figure 2: Confidence interval construction for thresholds in dynamic model

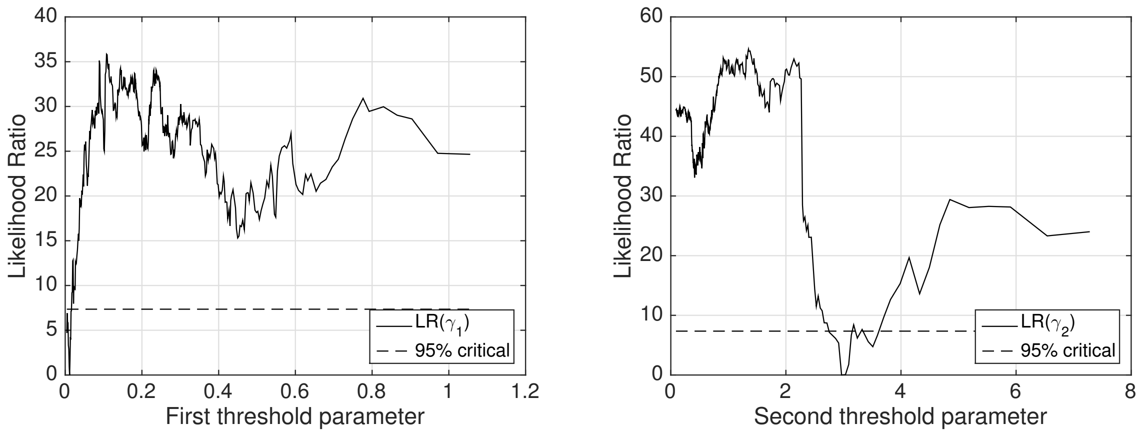



Table 4: Estimation results - Dependent variable: Investment ($I_{it}$)

| | Static model | | Dynamic model | |
|---|---|---|---|---|
| Regressor | OLS estimate | OLS SE | ML estimate | ML SE |
| $Q_{it}$ | 0.013 | 0.002 | 0.012 | 0.001 |
| $Q_{it}^2/10^3$ | -0.372 | 0.060 | -0.387 | 0.046 |
| $Q_{it}^3/10^6$ | 2.684 | 0.581 | 3.176 | 0.500 |
| $D_{it}$ | 0.041 | 0.006 | 0.031 | 0.004 |
| $Q_{it}D_{it}$ | -0.003 | 0.001 | -0.002 | 0.001 |
| $I_{it-1}$ | – | – | 0.237 | 0.011 |
| $CF_{it}1(D_{it} \leq \widehat{\gamma}_1, Q_{it} > \widehat{\gamma}_2)$ | -0.010 | 0.012 | -0.007 | 0.007 |
| $CF_{it}[1(D_{it} > \widehat{\gamma}_1, Q_{it} > \widehat{\gamma}_2)$ or $1(D_{it} \leq \widehat{\gamma}_1, Q_{it} \leq \widehat{\gamma}_2)]$ | 0.035 | 0.010 | 0.031 | 0.006 |
| $CF_{it}1(D_{it} > \widehat{\gamma}_1, Q_{it} \leq \widehat{\gamma}_2)$ | 0.069 | 0.009 | 0.063 | 0.006 |

Note: $\widehat{\gamma}_1 = 0.012$ and $\widehat{\gamma}_2 = 3.035$ are the estimated values for the static model, while $\widehat{\gamma}_1 = 0.012$ and $\widehat{\gamma}_2 = 2.968$ are the estimated values for the dynamic model.

Other empirical applications of two threshold variables can be found in Chen et al. (2012), who use the past information of price and market turnover as threshold variables to study the stock market in Hong Kong. Meanwhile, Chong and Yan (2014) consider several currency crisis indictors in a framework of currency crisis models. In turn, Donayre and Panovska (2018) study the role of inflation and unemployment in U.S. wage growth (wage Phillips curve), using a VAR framework and providing evidence of unemployment and inflation threshold effects on U.S. wage growth dynamics.

# 6 Conclusion

In this paper we introduce econometric techniques for static and dynamic panel data threshold models with two threshold variables. The setup is for short time period panels and considering individual fixed effects. In short, we extend that works of Chen et al. (2012), Hansen (1999) and Ramírez-Rondán (2015) to panel data models.

In the static panel data model, like Hansen (1999) with a model with one threshold variable, we propose a least squares estimation of the threshold and slope parameters using fixed effects transformations; while in the dynamic panel data model, like Ramírez-Rondán (2015) with a model with one threshold variable, we propose a maximum likelihood estimation of the threshold and slope parameters using first difference transformations.

In both panel data models, we propose to estimate the two threshold parameters sequentially. We also propose a method to construct confidence intervals for the threshold estimates, similar to Hansen (2000); and a test to determine the number of regimes, similar to Hansen (1996) and Chen et al. (2012).

In addition, we evaluate the performance of the estimators in a Monte Carlo experiment for 1000 replications; for a small sample size of number of individuals $n = 50$ and time periods $T = 5$. The threshold parameters estimated sequentially show a relatively small bias in both the static and dynamic models. But when we increase the number of individuals to $n = 500$ for the same time periods $T = 5$, these biases decrease quickly. The RMSE also decreases quickly as the number of individuals increases for a fixed time period.

The methods are applied to a 15-year sample of 565 U.S. firms to test whether financial constraints affect investment decisions. The threshold estimates in both the static and dynamic panel models indicate that for firms under the worst scenario, investment decisions are the most sensitive to cash flow; while for firms under the best scenario, investment decisions and cash flow are not related.

Several extensions to this paper would be desirable. These include allowing for heteroskedasticity, endogenous variables, random effects, non-balanced data, testing random against fixed effects, including time specific effects, considering more than two threshold variables, and comparing the results with alternative approximations such as regression kink.

# References

Aivazian, V.A., Y. Ge and J. Qiu (2005). “The Impact of Leverage on Firm Investment: Canadian Evidence.” *Journal of Corporate Finance* 11(1-2), 277-291.

Anderson, T. and C. Hsiao (1981): “Estimation of Dynamic Models With Error Components.” *Journal of the American Statistical Association* 76(375), 598-606.

Anderson, T. and C. Hsiao (1982): “Formulation and Estimation of Dynamics Models Using Panel Data.” *Journal of Econometrics* 18(1), 47-82.

Audretsch, D.B. and J.A. Elston (2002): “Does Firm Size Matter? Evidence on the Impact of Liquidity Constraints on Firm Investment Behavior in Germany.” *International Journal of Industrial Organization* 20(1), 7-17.

Bai, J. (1997): “Estimating Multiple Breaks One at a Time.” *Econometric Theory* 13(3), 315-352.

Blundell, R. and R.J. Smith (1991): “Conditions Initiales et Estimation Efficace dans les Modeles Dynamiques sur Donnees de Panel: une Application au Comportement d’Invetissement des Enterprises.” *Annales d’Economie et de Statistique* 20/21, 109-123.

Caner, M. and B.E. Hansen (2001): “Threshold Autoregression with a Unit Root.” *Econometrica* 69(6), 1555-1596.

Caner, M. and B.E. Hansen (2004): “Instrumental Variable Estimation of a Threshold Model.” *Econometric Theory* 20(5), 813-843.

Chan, K.S. (1993): “Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model.” *Annals of Statistics* 21(1), 520-533.

Chen, H., T. Chong and J. Bai (2012): “Theory and Applications of TAR Model with Two Threshold Variables.” *Econometric Reviews* 31(2), 142-170.

Chong, T. (2003): “Generic Consistency of the Break-Point Estimator under Specification Errors.” *Econometrics Journal* 6(1), 167-192.

Chong, T. and Yan, I. (2014): “Estimating and Testing Threshold Regression Models with Multiple Threshold Variables.” MPRA Working Paper 54732.

Donayre, L. and Panovska, I. (2018): “U.S. Wage Growth and Nonlinearities: The Roles of Inflation and Unemployment.” *Economic Modelling* 68, 273-292.

Enders, W. and C.W.J. Granger (1998): "Unit-Root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates." *Journal of Business and Economic Statistics* 16(13), 304-311.

Fazzari, S.M., R. Glenn Hubbard and B.C. Petersen (1988): "Financing Constraints and Corporate Investment." *Brookings Papers on Economic Activity* 1988(1), 141-195.

Gilchrist, S. and C.P. Himmelberg (1995): "Evidence on the Role of Cash Flow for Investment." *Journal of Monetary Economics* 36(3), 541-572.

Gonzáles, A., T. Terasvirta and D. van Dijk (2005): "Panel Smooth Transition Regression Models." Research Paper 165, Quantitative Finance Research Centre.

Hansen, B.E. (1996): "Inference When a Nuisance Parameter is not Identified Under the Null Hypothesis." *Econometrica* 64(2), 413-430.

Hansen, B.E. (1997): "Inference in TAR Models." *Studies in Nonlinear Dynamics and Econometrics* 2(1), 1-14.

Hansen, B.E. (1999): "Threshold Effects in Non-Dynamic Panels: Estimation, Testing and Inference." *Journal of Econometrics* 93(2), 345-368.

Hansen, B.E. (2000): "Sample Splitting and Threshold Estimation." *Econometrica* 68(3), 575-603.

Hansen, B.E. (2011): "Threshold Autoregression in Economics." *Statistics and Its Interface* 4(2), 123-127.

Hansen, B.E. and B. Seo (2002): "Testing for Two-Regime Threshold Cointegration in Vector Error-Correction Models." *Journal of Econometrics* 110(2), 293-318.

Hsiao, C, M. Pesaran and K. Tahmiscioglu (2002): "Maximum Likelihood Estimation of Fixed Effects Dynamic Panel Data Models Covering Short Time Periods." *Journal of Econometrics* 109(1), 107-150.

Hu, X. and Schiantarelli, F (1998): "Investment and Capital Market Imperfections: A Switching Regression Approach Using U.S. Firm Panel Data." *Review of Econometrics and Statistics* 80(3), 466-479.

Hubbard, H. (1998): "Capital-Market Imperfections and Investment." *Journal of Economic Literature* 36(1), 193-225.

Kadapakkam, P-R., P.C. Kumar and L.A. Riddick (1998): "The Impact of Cash Flows and Firm Size on Investment: The International Evidence." *Journal of Banking & Finance* 22(3), 293-320.

Kapetanios, G. (1996): "A Bootstrap Procedure for Panel Data Sets with Many Cross-Sectional Units." *Econometrics Journal* 11(2), 377-395.

Lang, L., E. Ofek and R.M. Stulz (1996): "Leverage, Investment, and Firm Growth." *Journal of Financial Economics* 40(1), 3-29.

Lancaster, T. (2000): "The Incidental Parameter Problem Since 1948." *Journal of Econometrics* 95(2), 391-413.

Newey, W. and D. McFadden (1994): "Large Sample Estimation and Hypothesis Testing." *Handbook of Econometrics* IV, Engle, R. and D. McFadden (eds). New York: Elsevier, 2113-2245.

Nickell, S.J. (1981): "Biases in Dynamic Models with Fixed Effects." *Econometrica* 49(6), 1417-1426.

Pawlina, G. and Renneboog, L. (2005): "Is Investment Cash Flow Sensitivity Caused by Agency Costs or Asymmetric Information? Evidence from the UK." *European Financial Management* 11(4), 483-513.

Ramírez-Rondán, N.R. (2015): "Maximum Likelihood Estimation of Dynamic Panel Threshold Models." Working Paper 35, Peruvian Economic Association.

Seo, M.H. and Y. Shin (2016): "Dynamic Panels with Threshold Effect and Endogeneity." *Journal of Econometrics* 195(2), 169-186..

Schiantarelli, F. (1996): "Financial Constraints and Investment: Methodological Issues and International Evidence." *Oxford Review of Economic Policy* 12(2), 70-89.

Tong, H. (2007): "Birth of the Threshold Time Series Model." *Statistica Sinica* 17(1), 8-14.