UPPSALA
UNIVERSITET

Working Paper 2017:3

Department of Statistics

Panel Smooth Transition Regression Models

A. González, T. Teräsvirta, D. van Dijk,
and Y. Yang

Working Paper 2017:3
October 2017
Department of Statistics
Uppsala University
Box 513
SE-751 20 UPPSALA
SWEDEN

Title: Panel Smooth Transition Regression Models
Author: A. Gonzales, T. Teräsvirta, D. van Dijk, Y. Yang
E-mail: yukai.yang@statistik.uu.se

Panel Smooth Transition Regression Models

Andrés González[a], Timo Teräsvirta[bc], Dick van Dijk*[def], and Yukai Yang[gh]

[a] *Banco de la Republica, Bogota, Colombia*
[b] *CREATES, Aarhus University*
[c] *C.A.S.E., Humboldt-Universität zu Berlin*
[d] *Econometric Institute, Erasmus University Rotterdam*
[e] *Erasmus Research Institute of Management (ERIM)*
[f] *Tinbergen Institute*
[g] *Department of Statistics, Uppsala University*
[h] *Center for Economic Statistics, Stockholm School of Economics*

October 2017

Abstract

We introduce the panel smooth transition regression model. This new model is intended for characterizing heterogeneous panels, allowing the regression coefficients to vary both across individuals and over time. Specifically, heterogeneity is allowed for by assuming that these coefficients are bounded continuous functions of an observable variable and fluctuate between a limited number of "extreme regimes". The model can be viewed as a generalization of the threshold panel model of Hansen (1999). We extend the modelling strategy originally designed for univariate smooth transition regression models to the panel context. The strategy consists of model specification based on homogeneity tests, parameter estimation, and model evaluation, including tests of parameter constancy and no remaining heterogeneity. The model is applied to describing firms' investment decisions in the presence of capital market imperfections.

Keywords: financial constraints; heterogeneous panel; investment; misspecification test; nonlinear modelling of panel data; smooth transition model.

JEL Classification Codes: C12, C23, C52, G31, G32.

Acknowledgements: This work has been supported by Jan Wallander's and Tom Hedelius's Foundation, Grants No. J99/37 and J02-35. In addition, Yang acknowledges support from the same foundation, Grant No. P2016-0293:1. Teräsvirta and Yang have received support from the Center for Research in Econometric Analysis of Time Series (CREATES), funded by the

*Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands, e-mail: djvandijk@ese.eur.nl (corresponding author)

Danish National Research Foundation, Grant No. DNRF 78. We also wish to thank Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Fonds de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under Grant No. 2.5020.11, and Swedish National Infrastructure for Computing (SNIC) through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project SNIC 2015/6-117, for providing the necessary computational resources. Responsibility for any errors or shortcomings in this work remains ours.

Note: This is a revised and updated version of the Working Paper No. 604 (2005) in the Working Paper Series of Economics and Finance, Stockholm School of Economics.

1 Introduction

In regression models for panel data it is typically assumed that the heterogeneity in the data can be captured completely by means of (fixed or random) individual effects and time effects, such that the coefficients of the observed explanatory variables are identical for all observations. In many empirical applications, however, this poolability assumption may be violated or at least may be viewed as questionable. For example, there is a sizable literature documenting that, due to capital market imperfections such as information asymmetry between borrowers and lenders, investment decisions of individual firms depend on financial variables such as cash flow. The sensitivity of investment to cash flow often is found to vary across firms according to the severity of the information asymmetry problem or their investment opportunities. In particular, external finance may be limited mainly for firms facing high agency costs due to information asymmetry or for firms with limited profitable investment opportunities. For firms constrained in this manner, investment will depend on the availability of internal finance to a much larger extent than for unconstrained firms. A heterogeneous panel data model is required for modelling investment behaviour of firms in such a situation.

Various panel data models that allow regression coefficients to vary over time and across cross-sectional units (or "individuals") have been developed, see Hsiao (2003, Chapter 6) and Pesaran (2015, Chapter 28) for overviews. These include random coefficients models as surveyed by Hsiao and Pesaran (2008) and models with coefficients that are functions of other exogenous variables. A specific example of the latter type of parameter heterogeneity is the panel threshold regression (PTR) model developed by Hansen (1999). In this model, regression coefficients can take on a small number of different values, depending on the value of another observable variable. Interpreted differently, the observations in the panel are divided into a small number of homogeneous sets or "regimes", with different coefficients in different regimes. A feature that makes the PTR model quite appealing is that individuals are not restricted to remain in the same set for all time periods if the so-called threshold variable that is used for grouping the observations is time-varying. In the aforementioned empirical example of firms' investment decisions it is likely that information asymmetry and investment opportunities change over time, causing firms to switch between constrained and unconstrained regimes.

In this paper we consider a nonlinear panel model we call the panel smooth transition regression (PSTR) model. It generalizes the PTR model by allowing the regression coefficients to change smoothly when moving from one "extreme" regime or state to another. The PTR model separates the observations clearly into several sets or groups based on the value of the threshold variable with sharp "borders" or thresholds. In practice, this may not always be feasible. For example, it seems difficult to argue that there is an exact level of financial constraints defining two groups of firms, each with different sensitivity of investment to cash flow, simultaneously assuming that all firms within these groups are homogeneous. Rather, it seems more realistic to assume that the sensitivity of cash flow changes gradually as a function of the level of financial constraints. The PSTR model is designed to take this possibility into account.

Since the appearance of the working paper version (González, Teräsvirta and van Dijk, 2005) of this article, our PSTR model has been applied to quite a wide variety of economic modelling problems. These include the relationship between pollution and economic growth (Aslanidis and Xepapadeas 2006, 2008), the inflation-growth nexus (Espinoza, Leon, and Prasad 2012, Seleteng, Bittencourt, and van Eyden 2013, Omay, van Eyden, and Gupta 2017), the effects of oil prices on the current account of oil-exporting countries (Allegret, Couharde, Coulibaly, and Mignon 2014), borrowing costs of European countries during the recent financial crisis (Delatte, Gex, and López-Villavicencio 2012, Delatte, Fouquau, and Portes 2017), the behaviour of exchange rates (Béreau, López Villavicencio and Mignon 2010, 2012; Chuluun, Eun, and Kiliç 2011; Cho, 2015), the Feldstein-Horioka puzzle of domestic savings and investment rates (Fouquau, Hurlin, and Rabaud 2008), earnings persistence of firms (Cheng and Wu 2013), the relationship between temperature and electricity consumption (Besseca and Fouquau 2008), and the relationship between patents and market value in the pharmaceutical industry (Chen, Shi and Chang 2014), just to name a few. These studies demonstrate the fact that the PSTR model offers an attractive possibility of capturing heterogeneity in panel data.

In this work we develop and describe a complete model building procedure for PSTR models with empirical applications in mind. The modelling cycle includes different stages of model specification, parameter estimation and model evaluation, and is an extension of the procedure that is available for smooth transition regression models for a single cross-section or time series, see Teräsvirta (1998), van Dijk, Teräsvirta, and Franses (2002), and Teräsvirta, Tjøstheim and Granger (2010, Chapter 16), among others. As part of the specification stage we suggest a novel Lagrange Multiplier (LM) test of parameter homogeneity. Although the test is designed specifically against the PSTR alternative, it has wider applicability as a general test of poolability of panel data, see also Baltagi (2013, Section 4.1). Similarly, as part of the evaluation stage we develop a test of parameter constancy in PSTR models, but also this test can be applied to other panel models. We conduct an extensive set of Monte Carlo simulation experiments to evaluate the performance of the various specification and evaluation tests. There we uncover that the wild cluster bootstrap of Cameron, Gelbach, and Miller (2008) is an extremely useful procedure to obtain satisfactory size and power properties in finite samples.

In our empirical application we take up the problem of individual firms' investment decisions in the presence of credit market imperfections. Using a balanced panel of 565 US firms observed for the years 1973-1987, we find that a two-regime PSTR model with Tobin's Q as transition variable adequately captures the heterogeneity in regression coefficients across firms. The model identifies firms with limited growth opportunities (low Q values) as a separate group that is distinct from firms with moderate or good growth opportunities. The transition from the lower regime associated with small values of Tobin's Q to the upper regime with large values of Q is smooth. On average about 12% of firms switch regimes in a given year, clearly illustrating the relevance of not constraining firms to remain in the same group over time. We find significant negative effects of debt on investment only for low Q firms, showing that leverage matters for investment only for firms with poor growth opportunities or firms with growth opportunities that are not recognized by the market. Similarly, the coefficient estimate of lagged cash flow is positive and significant only for low Q firms, which corroborates previous findings that internal finance is relevant for investment mainly for financially constrained firms.

The paper is organized as follows. Section 2 introduces the panel smooth transition regression model, focusing on interpretation of the model structure and on its relation to the PTR model of Hansen (1999). Section 3 describes the model building procedure for PSTR models. Section 4 considers the small sample properties of the different test statistics involved in the modelling cycle by means of Monte Carlo simulation. Special attention is given to the issue of cross-sectional heteroskedasticity and the consequences thereof for the performance of the tests. Section 5 contains the empirical application, and Section 6 concludes.

2 Panel smooth transition regression model

The Panel Smooth Transition Regression (PSTR) model can be interpreted in two different ways. First, it may be thought of as a linear heterogeneous panel model with coefficients that vary across individuals and over time. Heterogeneity in the regression coefficients is allowed for by assuming that these coefficients are bounded continuous functions of an observable variable, called the transition variable. This makes them fluctuate between a limited number (often two) of "extreme regimes". As the transition variable possibly is individual-specific and time-varying, the regression coefficients are allowed to be different for each of the individuals in the panel and to change over time. Second, the PSTR model can simply be considered as a nonlinear homogeneous panel model. The latter interpretation is in fact common in the context of single-equation smooth transition regression (STR) or univariate smooth transition autoregressive (STAR) models, see Teräsvirta (1994, 1998). Given the current context, we prefer the first interpretation.

The basic PSTR model with two extreme regimes is defined as

$$y_{it} = \mu_i + \lambda_t + \beta_0' x_{it} + \beta_1' x_{it} g(q_{it}; \gamma, c) + u_{it} \tag{1}$$

for $i = 1, \ldots, N$, and $t = 1, \ldots, T$, where N and T denote the cross-sectional and time dimensions of the panel, respectively. The dependent variable y_{it} is a scalar, x_{it} is a k-dimensional vector of time-varying exogenous variables, μ_i and λ_t represent fixed individual effects and time effects, respectively, and u_{it} are the errors. Furthermore, the regressors x_{it} are assumed exogenous. Possible extensions of the model to relax this restriction are discussed in Section 6.

The transition function $g(q_{it}; \gamma, c)$ in (1) is a continuous function of the observable variable q_{it} and is normalized to be bounded between zero and one. These two extreme values are associated with regression coefficients β_0 and $\beta_0 + \beta_1$. More generally, the value of the transition variable q_{it} determines the value of $g(q_{it}; \gamma, c)$ and thus the effective regression coefficients $\beta_0 + \beta_1 g(q_{it}; \gamma, c)$ for individual i at time t. We follow Teräsvirta (1994, 1998) and Jansen and Teräsvirta (1996), see also Teräsvirta *et al.* (2010, Chapter 3), by using the logistic specification

$$g(q_{it}; \gamma, c) = \left(1 + \exp\left(-\gamma \prod_{j=1}^{m} (q_{it} - c_j)\right)\right)^{-1} \quad \text{with } \gamma > 0 \text{ and } c_1 < c_2 < \ldots < c_m \tag{2}$$

where $c = (c_1, \ldots, c_m)'$ is an m-dimensional vector of location parameters and the slope parameter γ determines the smoothness of the transitions. The restrictions

$\gamma > 0$ and $c_1 < \ldots < c_m$ are imposed for identification purposes. In practice it is usually sufficient to consider $m = 1$ or $m = 2$, as these values allow for commonly encountered types of variation in the parameters. For $m = 1$, the model implies that the two extreme regimes are associated with low and high values of q_{it} with a monotonic transition of the coefficients from β_0 to $\beta_0 + \beta_1$ as q_{it} increases, where the change is centred around c_1. When $\gamma \to \infty$, $g(q_{it}; \gamma, c)$ becomes an indicator function $\mathtt{I}[q_{it} > c_1]$, defined as $\mathtt{I}[A] = 1$ when the event A occurs and zero otherwise. In that case the PSTR model in (1) reduces to the two-regime panel threshold model of Hansen (1999). For $m = 2$, the transition function has its minimum at $(c_1 + c_2)/2$ and attains the maximum value one both at low and high values of q_{it}. When $\gamma \to \infty$, the model becomes a three-regime threshold model whose outer regimes are identical and different from the mid-regime. In general, when $m > 1$ and $\gamma \to \infty$, the number of distinct regimes remains two, with the transition function switching back and forth between zero and one at $c_1, \ldots, c_m$. Finally, for any positive integer value m the transition function (2) becomes constant when $\gamma \to 0$, in which case the model collapses into a homogeneous or linear panel regression model with fixed effects.

A generalization of the PSTR model to allow for more than two different regimes is the additive model

$$y_{it} = \mu_i + \lambda_t + \beta_0' x_{it} + \sum_{j=1}^{r} \beta_j' x_{it} g_j(q_{it}^{(j)}; \gamma_j, c_j) + u_{it} \tag{3}$$

where the transition functions $g_j(q_{it}^{(j)}; \gamma_j, c_j)$, $j = 1, \ldots, r$, are defined by (2) with polynomial degrees m_j. If $m_j = 1$, $q_{it}^{(j)} = q_{it}$, and $\gamma_j \to \infty$ for all $j = 1, \ldots, r$, the model in (3) becomes a PTR model with $r + 1$ regimes. Consequently, the additive PSTR model can be viewed as a generalization of the multiple regime panel threshold model in Hansen (1999). Additionally, when the largest model that one is willing to consider is a two-regime PSTR model (1) with $r = 1$ and $m = 1$ or $m = 2$, model (3) plays a role in the evaluation of the estimated model. More specifically, the multiregime model (3) constitutes a natural alternative hypothesis in diagnostic tests of no remaining heterogeneity, as discussed in Section 3.3.2.

3 Building panel smooth transition regression models

Application of nonlinear models such as the PSTR model requires a careful and systematic modelling strategy. The modelling cycle that is available for smooth transition regression (STR) models for a single time series y_t, $t = 1, \ldots, T$, or potentially also for a single cross-section y_i, $i, \ldots, N$, can be readily extended to panel STR models. The STR model building procedure consists of specification, estimation and evaluation stages. In the panel case, specification includes testing homogeneity, selecting the transition variable q_{it} and, if homogeneity is rejected, determining the appropriate form of the transition function, that is, choosing the proper value of m in (2). Nonlinear least squares is used for parameter estimation. At the evaluation stage the estimated model is subjected to misspecification tests to check whether it provides an adequate description of the data. The null hypotheses to be tested at this stage include parameter constancy, no remaining heterogeneity and no autocorrelation in the errors. Finally, one also has to choose the number of transitions in the panel, which means selecting r in model (3).

In the following subsections we discuss the different elements of the model building procedure in more detail, see also Teräsvirta (1998), van Dijk, Teräsvirta, and Franses (2002) and Teräsvirta *et al.* (2010, Chapter 16), among others. For ease of exposition, throughout this section we focus on the PSTR model with fixed individual effects only, that is, we set $\lambda_t = 0$ for all t in (1).

An R package containing procedures for all aspects of the PSTR model building procedure is available at `https://cran.r-project.org/package=PSTR`.

3.1 Model specification: testing homogeneity

The initial specification stage of the modelling cycle essentially consists of testing homogeneity against the PSTR alternative. This is important for two reasons. First, there is a major statistical issue, namely, the PSTR model is not identified if the data-generating process is homogeneous, and to avoid the estimation of unidentified models homogeneity has to be tested first. Second, a homogeneity test may be useful for testing propositions from economic theory, such as identical sensitivity of investment to cash flow or other variables for all firms in a population.

The PSTR model (1) with (2) can be reduced to a homogeneous model by imposing either $\mathsf{H}_0 : \gamma = 0$ or $\mathsf{H}_0' : \beta_1 = 0$. The associated tests are nonstandard because under either null hypothesis the PSTR model contains unidentified nuisance parameters. In particular, the location parameters c_j are not identified under either null hypothesis, and this is also the case for β_1 under H_0 and for γ under H_0'. The problem of hypothesis testing in the presence of unidentified nuisance parameters was first studied by Davies (1977, 1987). Luukkonen, Saikkonen, and Teräsvirta (1988), Andrews and Ploberger (1994) and Hansen (1996) proposed alternative solutions in the time series context. We follow Luukkonen, Saikkonen, and Teräsvirta (1988) and test homogeneity using the null hypothesis $\mathsf{H}_0 : \gamma = 0$. To circumvent the identification problem we replace $g(q_{it}; \gamma, c)$ in (1) by its first-order Taylor expansion around $\gamma = 0$. After reparameterisation, this leads to the auxiliary regression

$$y_{it} = \mu_i + \beta_0'^* x_{it} + \beta_1'^* x_{it} q_{it} + \ldots + \beta_m'^* x_{it} q_{it}^m + u_{it}^* \tag{4}$$

where the parameter vectors $\beta_1^*, \ldots, \beta_m^*$ are multiples of γ, and $u_{it}^* = u_{it} + R_m \beta_1' x_{it}$, where R_m is the remainder of the Taylor expansion. Consequently, testing $\mathsf{H}_0 : \gamma = 0$ in (1) is equivalent to testing the null hypothesis $\mathsf{H}_0^* : \beta_1^* = \ldots = \beta_m^* = 0$ in (4). Note that under the null hypothesis $\{u_{it}^*\} = \{u_{it}\}$, so the Taylor series approximation does not affect the asymptotic distribution theory when the null hypothesis is tested by an LM test.

In order to define the LM-type statistic, we write (4) in matrix notation as follows:

$$y = D_\mu \mu + X\beta + W\beta^* + u^* \tag{5}$$

where $y = (y_1', \ldots, y_N')'$ with $y_i = (y_{i1}, \ldots, y_{iT})'$, $i = 1, \ldots, N$, $D_\mu = (I_N \otimes \iota_T)$ where I_N is the $(N \times N)$ identity matrix, ι_T a $(T \times 1)$ vector of ones, and $\mu = (\mu_1, \ldots, \mu_N)'$.

Moreover, $X = (X_1', \ldots, X_N')'$ where $X_i = (x_{i1}, \ldots, x_{iT})'$, $W = (W_1', \ldots, W_N')'$ with $W_i = (w_{i1}, \ldots, w_{iT})'$ and $w_{it} = (x_{it}' q_{it}, \ldots, x_{it}' q_{it}^m)'$, $\beta = \beta_0^*$ and $\beta^* = (\beta_1^{*\prime}, \ldots, \beta_m^{*\prime})'$. Finally, $u^* = (u_1'^*, \ldots, u_N'^*)'$ is a $(TN \times 1)$ vector with $u_i^* = (u_{i1}^*, \ldots, u_{iT}^*)'$. The LM test statistic has the form

$$\mathrm{LM}_\chi = \hat{u}^{0\prime} \tilde{W} \hat{\Sigma}^{-1} \tilde{W}' \hat{u}^0 \tag{6}$$

where $\hat{u}^0 = (\hat{u}_1^{0\prime}, \ldots, \hat{u}_N^{0\prime})'$ is the vector of residuals obtained by estimating the model under the null hypothesis and $\tilde{W} = M_\mu W$, where $M_\mu = I_{NT} - D_\mu (D_\mu' D_\mu)^{-1} D_\mu'$, is the standard within-transformation matrix. Furthermore, $\hat{\Sigma}$ is a consistent estimator of the covariance matrix $\Sigma = \mathsf{E}(\widehat{\beta}^* - \beta^*)(\widehat{\beta}^* - \beta^*)'$. When the errors are homoskedastic and identically distributed across time and individuals, the standard covariance matrix estimator

$$\hat{\Sigma}^{\mathrm{ST}} = \hat{\sigma}^2 (\tilde{W}' \tilde{W} - \tilde{W}' \tilde{X} (\tilde{X}' \tilde{X})^{-1} \tilde{X}' \tilde{W}) \tag{7}$$

where $\tilde{X} = M_\mu X$, and $\hat{\sigma}^2$ is the error variance estimated under the null hypothesis, is available. When the errors are heteroskedastic or autocorrelated, an appropriate estimator of Σ is given by

$$\hat{\Sigma}^{\mathrm{HAC}} = [-\tilde{W}' \tilde{X} (\tilde{X}' \tilde{X})^{-1} : I_{km}] \hat{\triangle} [-\tilde{W}' \tilde{X} (\tilde{X}' \tilde{X})^{-1} : I_{km}]' \tag{8}$$

where I_{km} is a $(km \times km)$ identity matrix, and

$$\hat{\triangle} = \sum_{i=1}^{N} \tilde{Z}_i' \hat{u}_i^0 \hat{u}_i^{0\prime} \tilde{Z}_i$$

with $\tilde{Z}_i = (I_T - \iota_T (\iota_T' \iota_T)^{-1} \iota_T') Z_i$, where $Z_i = [X_i, W_i]$, $i = 1, \ldots, N$. The estimator (8) is consistent for fixed T as $N \to \infty$, see Arellano (1987) for details and Hansen (2007) for an analysis of the remaining cases in which N and T approach infinity jointly or $T \to \infty$ with N fixed. Under the null hypothesis, LM_χ is asymptotically distributed as $\chi^2(mk)$, whereas the F-version $\mathrm{LM}_{\mathrm{F}} = \mathrm{LM}_\chi (TN - N - k - mk)/(TNmk)$ has an approximate $\mathrm{F}(mk, TN - N - k - mk)$ distribution.

Two remarks concerning the homogeneity test are in order. First, the test can be used for selecting the appropriate transition variable q_{it} in the PSTR model. In this case, the test by means of the Taylor expansion is carried out for a set of "candidate" transition variables and the variable that gives rise to the strongest rejection of linearity (if any) is chosen as the transition variable. Second, the homogeneity test can also be used for determining the appropriate order m of the logistic transition function in (2). Granger and Teräsvirta (1993) and Teräsvirta (1994, 1998), see also Teräsvirta *et al.* (2010, Chapter 16), proposed a sequence of tests for choosing between $m = 1$ and $m = 2$. Applied to the present situation this testing sequence reads as follows: Using the auxiliary regression (4) with $m = 3$, test the null hypothesis $\mathsf{H}_0^* : \beta_3^* = \beta_2^* = \beta_1^* = 0$. If it is rejected, test $\mathsf{H}_{03}^* : \beta_3^* = 0$, $\mathsf{H}_{02}^* : \beta_2^* = 0|\beta_3^* = 0$ and $\mathsf{H}_{01}^* : \beta_1^* = 0|\beta_3^* = \beta_2^* = 0$. Select $m = 2$ if the rejection of H_{02}^* is the strongest one, otherwise select $m = 1$. For the reasoning behind this simple rule, see Teräsvirta (1994).

3.2 Parameter estimation

Estimating the parameters $\theta = (\beta_0', \beta_1', \gamma, c')'$ in the PSTR model (1) is a relatively straightforward application of the fixed effects estimator and nonlinear least squares (NLS). We first eliminate the individual effects μ_i by removing individual-specific means and then apply NLS to the transformed data.

While eliminating fixed effects using the within transformation is standard in linear panel data models, the PSTR model calls for a more careful treatment. Rewrite model (1) as follows:

$$y_{it} = \mu_i + \beta' x_{it}(\gamma, c) + u_{it} \tag{9}$$

where $x_{it}(\gamma, c) = (x_{it}', x_{it}' g(q_{it}; \gamma, c))'$ and $\beta = (\beta_0', \beta_1')'$. Subtracting individual means from (9) yields

$$\tilde{y}_{it} = \beta' \tilde{x}_{it}(\gamma, c) + \tilde{u}_{it} \tag{10}$$

where $\tilde{y}_{it} = y_{it} - \bar{y}_i$, $\tilde{x}_{it}(\gamma, c) = (x_{it}' - \bar{x}_i', x_{it}' g(q_{it}; \gamma, c) - \bar{w}_i'(\gamma, c))'$, $\tilde{u}_{it} = u_{it} - \bar{u}_i$, and $\bar{y}_i$, $\bar{x}_i$, $\bar{w}_i$ and $\bar{u}_i$ are individual means, with $\bar{w}_i(\gamma, c) \equiv T^{-1} \sum_{t=1}^{T} x_{it} g(q_{it}; \gamma, c)$. Consequently, the transformed vector $\tilde{x}_{it}(\gamma, c)$ in (10) depends on γ and c through both the levels and the individual means. For this reason, $\tilde{x}_{it}(\gamma, c)$ needs to be recomputed at each iteration in the NLS optimization.

From (10) it is seen that the PSTR model is linear in β conditional on γ and c. Thus, we apply NLS to determine the values of these parameters that minimize the concentrated sum of squared errors

$$Q^c(\gamma, c) = \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\tilde{y}_{it} - \hat{\beta}(\gamma, c)' \tilde{x}_{it} (\gamma, c) \right)^2 \tag{11}$$

where $\hat{\beta}(\gamma, c)$ is obtained from (10) by ordinary least squares at each iteration in the nonlinear optimization. In case the errors u_{it} in (9) are normally distributed, this estimation procedure is equivalent to maximum likelihood (ML), where the likelihood function is first concentrated with respect to the fixed effects μ_i.

A practical issue that deserves special attention in the estimation of PSTR models is the selection of starting values for the NLS optimization. We follow the common practice for STR models to obtain starting values by means of a grid search across the parameters in the transition function $g(q_{it}; \gamma, c)$. This approach is based on the aforementioned fact that (10) is linear in β when γ and c are fixed. Hence, the concentrated sum of squared residuals (11) can be computed easily for an array ("grid") of values for γ and c such that $\gamma > 0$, and $c_{j,\min} > \min_{i,t} \{q_{it}\}$ and $c_{j,\max} < \max_{i,t} \{q_{it}\}$, $j = 1, \ldots, m$, and the values minimizing $Q^c(\gamma, c)$ can be used as starting values of the nonlinear optimization algorithm. An alternative approach to obtain starting values is simulated annealing, as recently considered by Schleer (2015) for STR and vector STR models.

Finally, it should be noted that numerical complications may occur when the slope parameter γ is large. They are due to the fact that in that situation γ is of completely different magnitude from the other parameters, which slows down convergence of any standard derivative-based optimization algorithm. Furthermore, the log-likelihood is typically rather flat in the direction of γ when this parameter is large, which may aggravate the problem. A way of alleviating this difficulty is to apply the transformation $\gamma = \exp\{\eta\}$ (or $\eta = \ln \gamma$) in (2) and estimate η instead of γ. Note that this transformation makes the identifying condition $\gamma > 0$ redundant.

It has been suggested by Goodwin, Holt, and Prestemon (2011), see also Hurn, Silvennoinen, and Teräsvirta (2016).

3.3 Model evaluation

Evaluation of an estimated PSTR model is an essential part of the model building procedure. In this section we consider two misspecification tests for this purpose. Specifically, we adapt the tests of parameter constancy over time and of no remaining nonlinearity developed by Eitrheim and Teräsvirta (1996) for univariate STAR models to fit the present panel framework, where we interpret the latter as a test of no remaining heterogeneity. We also discuss an alternative use of the test of no remaining heterogeneity as a specification test for determining the number of regimes in the PSTR model. We do not consider a panel version of the Eitrheim and Teräsvirta (1996) test of no error autocorrelation, because Baltagi and Li (1995) already derived such a test for panel models.

3.3.1 Testing parameter constancy

Testing parameter constancy in panel data models has not received as much attention as it has in the time series literature. A plausible explanation is that traditionally in many applications the time dimension T was relatively small, which made the assumption of parameter constancy a rather uninteresting hypothesis to test. However, as the number of empirical panel data sets with relatively large T is increasing, testing parameter constancy is becoming more important. Even though we here develop a test specifically for PSTR models, it can after minor modifications be applied to other fixed effects panel data models as well.

Our alternative to parameter constancy is that the parameters in (1) change smoothly over time. The model under the alternative may be called the Time Varying Panel Smooth Transition Regression (TV-PSTR) model, and it is defined as follows:

$$
\begin{aligned}
y_{it} &= \mu_i + (\beta_{10}' x_{it} + \beta_{11}' x_{it} g(q_{it}; \gamma_1, c_1)) \\
&\quad + f(t/T; \gamma_2, c_2)(\beta_{20}' x_{it} + \beta_{21}' x_{it} g(q_{it}; \gamma_1, c_1)) + u_{it} \qquad (12)
\end{aligned}
$$

where $g(q_{it}; \gamma_1, c_1)$ is defined in (2) and $f(t/T; \gamma_2, c_2)$ is another transition function. Model (12) has the same structure as the time-varying smooth transition autoregressive (TV-STAR) model discussed in Lundbergh, Teräsvirta, and van Dijk (2003). We may also write (12) as

$$
\begin{aligned}
y_{it} &= \mu_i + (\beta_{10} + \beta_{20} f(t/T; \gamma_2, c_2))' x_{it} \\
&\quad + (\beta_{11} + \beta_{21} f(t/T \gamma_2, c_2))' x_{it} g(q_{it}; \gamma_1, c_1) + u_{it} \qquad (13)
\end{aligned}
$$

to explicitly show the deterministic character of time-variation in the parameters of the model. It should be noted that the TV-PSTR model Geng (2011) introduced is a special case of ours, as she assumed $\beta_{11} = \beta_{21} = 0$ in (12).

The TV-PSTR model accommodates various alternatives to parameter constancy depending on the definition of $f(t/T; \gamma_2, c_2)$. This function is assumed to have the form

$$
f(t/T; \gamma_2, c_2) = \left(1 + \exp\left(-\gamma_2 \prod_{j=1}^{h} (t/T - c_{2j})\right)\right)^{-1} \qquad (14)
$$

where $c_2 = (c_{21}, \ldots, c_{2h})'$ is an h-dimensional vector of location parameters with $c_{21} < c_{22} < \ldots < c_{2h}$, and $\gamma_2 > 0$ is the slope parameter. This is identical to $g(q_{it}; \gamma, c)$ as defined in (2) with $q_{it} = t/T$. Thus, when setting $h = 1$ the TV-PSTR model allows for a single monotonic change, while the change is symmetric around $(c_{21} + c_{22})/2$ in case $h = 2$. The smoothness of the change is controlled by γ_2. When $\gamma_2 \to \infty$, $f(t/T; \gamma_2, c_2)$ becomes an indicator function $\mathtt{I}[t/T > c_{21}]$ in case $h = 1$ and $1 - \mathtt{I}[c_{21} < t/T \leq c_{22}]$ in case $h = 2$. This means that (14) also accommodates instantaneous structural breaks.

When $\gamma_2 = 0$ in (14), $f(t/T; 0, c_2) \equiv 1/2$, so the model defined in (12) has constant parameters and $\mathsf{H}_0 : \gamma_2 = 0$ can be chosen to be the null hypothesis of parameter constancy. When it holds, the parameters β_{20}, β_{21} and c_2 in (12) are not identified. Our solution to this identification problem is the same as the one used in Section 3.1, namely to replace $f(t/T;\gamma_2,c_2)$ by its first-order Taylor expansion around $\gamma_2=0$. After rearranging terms this yields the auxiliary regression

$$\begin{aligned} y_{it} = \mu_i + \beta_{10}^{*\prime}x_{it} + \beta_1^{*\prime}x_{it}(t/T) + \beta_2^{*\prime}x_{it}(t/T)^2 + \ldots + \beta_h^{*\prime}x_{it}(t/T)^h \\ + \left(\beta_{11}^{*\prime}x_{it} + \beta_{h+1}^{*\prime}x_{it}(t/T) + \ldots + \beta_{2h}^{*\prime}x_{it}(t/T)^h\right) g\left(q_{it};\gamma_1,c_1\right) + u_{it}^* \end{aligned} \quad (15)$$

where $u_{it}^* = u_{it} + R_h\left(t/T,\gamma_2,c_2\right)$ and $R_h\left(t/T,\gamma_2,c_2\right)$ is the remainder term. In (15), the parameter vectors β_j^* for $j=1,2,\ldots,h,h+1,\ldots,2h$ are multiples of γ_2, such that the null hypothesis $\mathsf{H}_0:\gamma_2=0$ in (12) can be reformulated as $\mathsf{H}_0^*:\beta_j^*=0$ for $j=1,2,\ldots,h,h+1,\ldots,2h$ in the auxiliary regression. Under H_0^* $\{u_{it}^*\}=\{u_{it}\}$, so the Taylor series approximation does not affect the asymptotic distribution theory. The χ^2- and F-versions of the LM-type test can be computed as in (6) defining $w_{it}' = (x_{it}', x_{it}'g(q_{it},\hat{\gamma}_1,\hat{c}_1))\otimes s_t'$ with $s_t=((t/T),\ldots,(t/T)^h)'$ and replacing $\tilde{X}$ in (7) and (8) by $\tilde{V}=M_\mu V$, where $V=(V_1',\ldots,V_N')'$ with $V_i=(v_{i1}',\ldots,v_{iT}')'$ and $v_{it}=(x_{it}',x_{it}'g(q_{it},\hat{\gamma}_1,\hat{c}_1),\,(\partial\hat{g}/\partial\gamma_1)x_{it}'\hat{\beta}_2,(\partial\hat{g}/\partial c_1')x_{it}'\hat{\beta}_2)'$. Under the null hypothesis, LM_χ is asymptotically distributed as $\chi^2(2hk)$ and $\mathrm{LM}_\mathrm{F}=\mathrm{LM}_\chi/2hk$ is approximately distributed as $F(2hk,TN-N-2k(h+1)-(m+1))$. When the null model is a homogeneous fixed effects model ($\beta_{11}\equiv\beta_{20}\equiv\beta_{21}\equiv 0$ in (12)), a simplified version of (15) (without the terms $(\beta_{11}^{*\prime}x_{it}+\beta_{h+1}^{*\prime}x_{it}(t/T)+\ldots+\beta_{2h}^{*\prime}x_{it}(t/T)^h)g(q_{it};\gamma_1,c_1)$) renders a parameter constancy test for this model.

Eitrheim and Teräsvirta (1996) pointed out potential numerical problems in the computation of the test of parameter constancy (as well as the test of no remaining heterogeneity to be discussed below). In particular, when the estimate of γ_1 in the model under the null hypothesis is relatively large, such that the transition between regimes is rapid, the partial derivatives of $g(q_{it};\gamma_1,c_1)$ with respect to γ_1 and c_1 evaluated at the estimates under the null are equal to zero for almost all observations. As a result, the moment matrix of $\tilde{V}$ becomes near-singular such that the value of the LM statistic cannot be reliably computed. However, the contribution of the terms involving these partial derivatives to the test statistic is negligible at large values for γ_1. They can simply be omitted from the auxiliary regression without influencing the empirical size (or power) of the test statistic. If this is done, the degrees of freedom in the F-tests have to be modified accordingly.

3.3.2 Testing the hypothesis of no remaining heterogeneity

The assumption that a two-regime PSTR model (1) with (2) adequately captures the heterogeneity in a panel data set can be tested in various ways. In the PSTR framework it is a natural idea to consider an additive PSTR model (3) with two transitions ($r = 2$) as an alternative. Thus,

$$y_{it} = \mu_i + \beta_0' x_{it} + \beta_1' x_{it} g_1(q_{it}^{(1)}; \gamma_1, c_1) + \beta_2' x_{it} g_2(q_{it}^{(2)}; \gamma_2, c_2) + u_{it} \tag{16}$$

where the transition variables $q_{it}^{(1)}$ and $q_{it}^{(2)}$ can be but need not be the same. The null hypothesis of no remaining heterogeneity in an estimated two-regime PSTR model can be formulated as $\mathsf{H}_0 : \gamma_2 = 0$ in (16). This testing problem is again complicated by the presence of unidentified nuisance parameters under the null hypothesis. As before, the identification problem is circumvented by replacing $g_2(q_{it}^{(2)}; \gamma_2, c_2)$ by a Taylor expansion around $\gamma_2 = 0$. This leads to the auxiliary regression

$$y_{it} = \mu_i + \beta_0^{*\prime} x_{it} + \beta_1' x_{it} g_1(q_{it}^{(1)}; \hat{\gamma}_1, \hat{c}_1) + \beta_{21}^{*\prime} x_{it} q_{it}^{(2)} + \ldots + \beta_{2m}^{*\prime} x_{it} q_{it}^{(2)m} + u_{it}^* \tag{17}$$

where $\hat{\gamma}_1$ and $\hat{c}_1$ are estimates under the null hypothesis. Since $\beta_{21}^*, \ldots, \beta_{2m}^*$ are multiples of γ, the hypothesis of no remaining heterogeneity can be restated as $\mathsf{H}_0^* : \beta_{21}^* = \ldots = \beta_{2m}^* = 0$. If $\beta_1 \equiv 0$ in (17), the resulting test collapses into the homogeneity test discussed in Section 3.1.

In order to compute the LM test statistic defined in (6) and its F-version we set $w_{it} = (x_{it}' q_{it}^{(2)}, \ldots, x_{it}' q_{it}^{(2)m})'$ and again replace $\tilde{X}$ in (7) and (8) by $\tilde{V}$, where in this case $v_{it} = (x_{it}', x_{it}' g(q_{it}^{(1)}, \hat{\gamma}, \hat{c}_1), (\partial \hat{g}/\partial \gamma) x_{it}' \hat{\beta}_1, (\partial \hat{g}/\partial c_1') x_{it}' \hat{\beta}_1)'$. When H_0^* holds, the LM_χ statistic has an asymptotic $\chi^2(mk)$ distribution, whereas LM_F has an approximate $F(mk, TN - N - 2 - k(m + 2))$ distribution.

3.3.3 Determining the number of regimes

The tests of parameter constancy and of no remaining heterogeneity can be generalized to serve as misspecification tests in an additive PSTR model of the form (3) with $r > 0$. The purpose of the test of no remaining heterogeneity thus is in fact twofold. It is indeed a misspecification test but also a useful tool for determining the number of transitions in the model. The following sequential procedure may be used for this purpose:

1. Estimate a linear (homogeneous) model and test homogeneity at a predetermined significance level α.

2. If homogeneity is rejected, estimate a two-regime PSTR model.

3. Test the hypothesis of no remaining heterogeneity for this model. If it is rejected at significance level $\tau\alpha$, with $0 < \tau < 1$, estimate an additive PSTR model with $r = 2$. The purpose of reducing the significance level by a factor τ is to avoid excessively large models.

4. Continue until the null hypothesis of no remaining heterogeneity can no longer be rejected (using significance level $\tau^{r-1}\alpha$ when the additive PSTR model under the null includes r transition functions).

4 Size and power simulations

4.1 Design of experiment

We study the small sample properties of the different LM tests developed in Section 3 by means of Monte Carlo experiments. In the simulations we do not only consider different cross-sectional and time dimensions of the panel (N and T), but also investigate the effect of cross-sectional heteroskedasticity on the size and power of the tests. All experiments reported in this section are replicable using the R code available at `https://github.com/yukai-yang/PSTR_Experiments`.

The design of the Monte Carlo experiments is as follows. The number of replications equals 10,000 throughout. Each experiment is carried out for all possible combinations of $N = 20, 40, 80, 160$ and $T = 5, 10, 20$. In the various data generating processes (DGPs) we fix the number of regressors k in x_{it} at 2. The $(2+r) \times 1$ vector of exogenous regressors and transition variables $\left(x_{it}', q_{it}^{(1)}, \ldots, q_{it}^{(r)},\right)'$ is generated independently for each individual from the following VAR(1) model:

$$\begin{pmatrix} x_{it} \\ q_{it}^{(1)} \\ \vdots \\ q_{it}^{(r)} \end{pmatrix} = \kappa + \Theta \begin{pmatrix} x_{i,t-1} \\ q_{i,t-1}^{(1)} \\ \vdots \\ q_{i,t-1}^{(r)} \end{pmatrix} + \varepsilon_{it} \tag{18}$$

where $\kappa = (0.2, 0.2, 2.45, \ldots, 2.45)'$ and $\Theta = \text{diag}(0.5, 0.4, 0.3, \ldots, 0.3)$. The vector of shocks ε_{it} is drawn from a $N(0, \Sigma_\varepsilon)$ distribution where $\Sigma_\varepsilon = DRD$, $D = \sqrt{0.3} I_{2+r}$ and $R = [r_{ij}]$ with $r_{ii} = 1$ and $r_{ij} = 1/3$, $i \neq j$, $i, j = 1, \ldots, 2+r$. This generates both serial and contemporaneous correlation between the regressors and the transition variables. Values of the endogenous variable y_{it} are generated from the additive PSTR model

$$y_{it} = \mu_i + \beta_{i0}' x_{it} + \sum_{j=1}^{r} \beta_j' x_{it} g(q_{it}^{(j)}; \gamma_j, c_j) + u_{it} \tag{19}$$

where $\mu_i = \sigma_\mu e_i$ with $\sigma_\mu = 10$, and both e_i and u_{it} are i.i.d. standard normal. The values of r, m, and $(\gamma_j, c_j')'$ vary from one experiment to another. We consider two definitions of β_{i0}. In the first one, referred to as homoskedasticity, $\beta_{i0} = \beta_0 = (1, 1)'$ for all individuals i. The second one consists of defining $\beta_{i0} = \beta_0 + \nu_i$, where $\nu_i \sim N(0, I_2)$. This results in heteroskedastic errors in the auxiliary regressions such that the degree of heteroskedasticity is positively related to the regressors x_{it}.

The simulations are carried out using both the tests based on the asymptotically relevant χ^2-distribution and the approximative F-test. As the small sample properties of the latter are superior to those of the former, the results reported here are mainly based on the F-test. The next subsection constitutes the only exception. Results are available for significance levels 0.01, 0.05 and 0.10. For space reasons, only results for the significance level 0.05 are reported here.

4.2 Testing homogeneity

4.2.1 Size

In order to investigate the empirical size of the homogeneity test developed in Section 3.1 we generate samples from a homogeneous panel model with fixed effects ($r = 0$ in (19)). Results can be found in Table 1. The table contains rejection frequencies of the null hypothesis for both the standard χ^2-test and its F-approximation (indicated by F), and their robust versions (indicated by HAC). We compute the test statistics for $m_a = 1, 2, 3$, where m_a is the order of the auxiliary regression (4).

- insert Table 1 about here -

Table 1 has two panels. Panel (a) contains results of simulations in which the errors are homoskedastic, whereas the results in Panel (b) are based on designs with heteroskedastic errors. The results in Panel (a) demonstrate the well-known fact that the LM test based on the asymptotic null distribution is oversized. This is the case for all values of N, although the empirical size does improve somewhat with increasing T. The size distortion becomes worse with increasing m_a. The F-version corrects the size, but in fact the test is now undersized especially for small N and T. For the heteroskedasticity-robust (HAC) version of the test, the F-distribution based test is heavily undersized for all combinations of N and T. The χ^2-test is less undersized but has acceptable empirical size only for $T \geq 10$, $N = 160$, and $m_a = 1$. Panel (b) shows that the standard χ^2- and F-based test statistics are both substantially oversized, with the size distortion becoming larger when the cross-sectional dimension N of the panel increases. Furthermore, for the F-based version of the test, the empirical size also increases with the time dimension T of the panel, while it remains fairly constant for the χ^2-test.

In light of the size problems noted above, we examine whether bootstrapping the LM statistic can be used to correct these, see also Becker and Osborn (2010). For this purpose we first use the residual-based wild bootstrap (WB) based on the Rademacher distribution. It is seen from Table 1 that this does a decent job when the errors are homoskedastic but not when they are heteroskedastic. In the latter case the WB-LM test remains oversized, although it performs markedly better than the standard and HAC versions of the test. The remaining size distortion is caused by the fact that the coefficients β_{i0} are random and positively correlated with x_{it} in (19).

To remedy the situation the wild bootstrap is replaced by the wild cluster bootstrap (WCB) proposed by Cameron, Gelbach, and Miller (2008), which is designed to account for within-group dependence in panel data by resampling entire clusters of observations. In our setting the observations of the same individual over time form a cluster (such that the number of clusters in our simulations is equal to the number of individuals N). In Table 1 we observe that the empirical size of the WCB-LM test is very close to the nominal one for all combinations of N and T, both when the errors are homoskedastic and when they are heteroskedastic. Hence, we recommend the WCB approach when testing homogeneity, and in fact for this reason we concentrate on the WCB-LM test in the remainder of the simulation experiments.

It should be mentioned here that in bootstrapping the LM statistic we have made use of the warp-speed method proposed by Giacomini, Politis, and White (2013). This has been done to save computation time that otherwise, given the extent of the simulations, would have been rather excessive. All power experiments in the following sections also rely on the warp-speed approach.

4.2.2 Power

In this power experiment, we generate samples from the PSTR model (19) with $r = 1$ and with either a monotonically increasing ($m = 1$) or symmetric ($m = 2$) transition function (2). In both cases, we set $\beta_1 = (0.7, 0.7)'$. The parameters in the transition function are set equal to $c_1 = 3.5$ when $m = 1$ and $c_1 = (3.0, 4.0)$ when $m = 2$. The slope parameter $\gamma_1 = 4$ in both cases. The results for the WCB-LM test can be found in Table 2. When $m = 1$, the first-order auxiliary regression ($m_a = 1$) is sufficient. Increasing the order ($m_a = 2, 3$) weakens the power, although this effect diminishes for larger dimensions N and T of the panel. Heteroskedasticity has an adverse effect on the power. Not surprisingly, when $m = 2$ (so that the transition is symmetric around $(c_1 + c_2)/2 = 3.5$) the test based on the first-order auxiliary regression has very low power. Using an auxiliary regression with $m_a = 2$ results in a strong increase in power. Hence, (at least) a second-order auxiliary regression is required to capture a nonmonotonic transition. Choosing $m_a = 3$ lowers the power slightly but not by much. This effect is best seen in simulations with heteroskedastic errors.

- insert Table 2 about here -

4.3 Model evaluation

In this section we consider the case in which a PSTR model with a single transition function has been fitted to the data, and we want to evaluate the model by misspecification tests. We consider the test of no remaining heterogeneity (where the null model has one transition and it is tested against a model with two transitions) and the test of parameter constancy (using the same null model).

4.3.1 Size

We investigate the size properties of the misspecification tests using samples generated from the PSTR model (19) with $r = 1$, $\beta_1 = (1, 1)'$, $m = 1$, $\gamma_1 = 3$ and $c_1 = 3.5$. Table 3 contains the results, for samples with homoskedastic and heteroskedastic errors in panels (a) and (b) as before. For the test of no remaining heterogeneity results are reported for the case where the second transition function is assumed to be governed by the same transition variable $q_{it}^{(1)}$ as the transition function included in the null model. (Results for this test with a different variable governing the second transition are similar.) Overall, both misspecification tests perform satisfactorily, in the sense that the empirical size is quite close to the nominal significance level of 0.05. Apparently the wild-cluster bootstrap procedure is slightly affected by heteroskedastic errors, in the sense that this increases the empirical size but only slightly. The largest size distortion occurs for the smallest panels, with $N = 20$ and $T = 5$, and when the tests are based on a high-order auxiliary regression (15) or (17), with $m_a = 3$. In that case, the tests become quite substantially undersized. This effect quickly disappears when either of the panel dimensions increases and also when a lower-order auxiliary regression is used.

- insert Table 3 about here -

4.3.2 Power

Table 4 reports the empirical power of the WCB-LM test of parameter constancy. We generate artificial panel data sets using the TV-PSTR model (12) either with monotonic change centred in the middle ($h = 1$ and $c = 0.5T$ in (14)), or with nonmonotonic change ($h = 2$ and $c_1 = 0.3T$ and $c_2 = 0.7T$ in (14)), except that β_{10} is replaced by β_{i0} as defined in (19) in order to consider both homoskedasticity and heteroskedasticity. We follow the parameter settings in section 4.3.1 for the corresponding parameters in (12). For the parameters in the time-varying component in (12), we set $\gamma_2 = 4$, $\beta_{20} = 0.7\beta_{10}$, $\beta_{21} = 0.7\beta_{11}$. Since the heteroskedastic errors have been produced by a random β_{10}, we do not consider the random β_{20}.

Results for the two cases ($h = 1$ and $h = 2$) are shown in the left and right panels of the table. The left panel of Table 4 displays the same pattern as Table 3: when the parameter change is monotonic, the power is weaker in case a higher-order auxiliary regression (15) with $h_a = 2$ or 3 is used. This is to be expected as a first-order auxiliary regression is enough to detect monotonic change. When the change is nonmonotonic, the test based on $h_a = 1$ does have some power but, as before, the WCB-LM test based on the second-order auxiliary regression performs considerably better. We also observe that the presence of heteroskedastic errors substantially lowers the power of the test for panels with small or moderate dimensions N and T.

- insert Table 4 about here -

We examine the power properties of the test of no remaining heterogeneity by generating panels from (19) with $r = 2$, $q_{it}^{(1)} = q_{it}^{(2)}$, $m_1 = m_2 = 1$, $\gamma_1 = \gamma_2 = 8$,

$c_1 = 3$, and $c_2 = 4$. We consider two scenarios. In both cases, we use $\beta_0 = (1, 1)'$, and $\beta_1 = (0.7, 0.7)$. In the first scenario, we set $\beta_2 = \beta_1$, such that in this DGP heterogeneity is monotonic in $q_{it}^{(1)}$. In other words, the 'effective' regression coefficients are monotonically increasing functions of the transition variable, changing from β_0 to $\beta_0 + \beta_1$ to $\beta_0 + 2\beta_1$ with increasing $q_{it}^{(1)}$. In the second scenario, we set $\beta_2 = -\beta_1$, implying that that the coefficients in the lower ($q_{it}^{(1)} \ll c_1$) and the upper regimes ($q_{it}^{(1)} \gg c_2$) are identical. In both cases, we estimate a PSTR model with $r = 1$ and $m = 1$ and then apply the test of no remaining heterogeneity using the correct transition variable. Note that the second DGP resembles a PSTR model with $r = 1$ and $m = 2$. Nevertheless, even in this case we estimate a PSTR model with $r = 1$ and $m = 1$ in order to find out whether the test of no remaining heterogeneity is able to detect misspecification of the form of the heterogeneity (that is, of the order of the logistic function).

- insert Table 5 about here -

Table 5 contains results for both scenarios. When the combined transition is monotonic in $q_{it}^{(1)}$ (left panel), we find that the test of no remaining heterogeneity has only moderate empirical power, especially in the presence of heteroskedastic errors. In the case of non-monotonic heterogeneity (right panel), empirical power is substantially higher, especially when a first-order auxiliary regression ($m_a = 1$) is used. These results are not completely unexpected, in the sense that a model with a single transition function (with $m = 1$) might already capture the first type of heterogeneity quite accurately, but should have much more difficulty describing the second type.

5 Investment and capital market imperfections

In the presence of capital market imperfections, firms' investment decisions are not independent of financial factors such as cash flow and leverage. First, asymmetric information between borrowers and lenders concerning the quality of available investment opportunities generates agency costs that result in outside investors demanding a premium on newly issued debt or equity. This creates a "pecking order" or "financing hierarchy" with internal funds having a cost advantage relative to external capital. Hence, investment will be positively related to the availability of internal sources of finance, measured for example by cash flow. Second, high leverage reduces firms' ability to finance growth, such that firms with valuable investment opportunities should aim for lower leverage. One may therefore expect a negative relationship between future investment and leverage or "debt overhang".

The impact of these capital market imperfections and severity of the resulting problems varies across firms and over time, depending on the degree of informational asymmetry and growth opportunities, among others. For firms with low information costs or ample growth opportunities, internal and external finance are almost perfect substitutes and investment decisions are nearly independent of their financial structure. In contrast, firms with high information costs and limited growth opportunities face much higher costs of external finance or may even be rationed in their access to external funds. This in turn results in greater sensitivity of investment to cash flow. Similarly, capital structure theory suggests a disciplinary role for debt in the sense that leverage restricts managers of firms with poor growth opportunities from investing when they should not. Thus, leverage should mainly affect such firms and have much less effect on investment for firms with valuable growth opportunities recognized by the market.

A substantial number of empirical studies examine the effects of capital market imperfections on investment, see Fazzari, Hubbard, and Petersen (1988), Whited (1992), Bond and Meghir (1994), Carpenter, Fazzari, and Petersen (1994), Gilchrist and Himmelberg (1995), Lang, Ofek, and Stulz (1996), Hsiao and Tahmiscioglu (1997), Hu and Schiantarelli (1998), Moyen (2004), Hovakimian and Titman (2006), Almeida and Campello (2007), Hennessy, Levy, and Whited (2007), Fee, Hadlock, and Pierce (2009), and Hovakimian (2009), among others. Most studies are conducted in the context of the Q theory of investment, adding measures of cash flow or leverage to empirical models that relate investment to Tobin's Q. In perfect capital and output markets, Tobin's Q, defined as the market valuation of capital relative to its replacement value, is a sufficient statistic for investment. A significant positive coefficient on cash flow, for example, can then be interpreted as evidence in favour of the relevance of financing constraints.

In order to examine the question whether or not the effects of financing constraints or other capital market imperfections depend on financial factors, firms are typically divided into groups of "constrained" and "unconstrained" firms. This division is often based on a variable that measures the degree of information asymmetry such as the dividend pay-out ratio, size, age, the presence of a bond rating, and the debt ratio, or on a variable that measures growth opportunities such as Tobin's Q. This approach contains several potential limitations. First, the distinction between "constrained" and "unconstrained" firms is often based on an arbitrary threshold level of the variable that is used to split the sample. Second, in most studies, the composition of these two sets is fixed for the complete sample period in the sense that firms are not allowed to change sets over time. The PSTR model is designed to alleviate these shortcomings.

Following Hansen (1999), we use a balanced panel of 565 US firms observed for the years 1973–1987, extracted from the data set used by Hall and Hall (1993). For each firm i and year t, we obtain the ratios of investment to assets (I_{it}), Tobin's Q or total market value to assets (Q_{it}), long-term debt to assets (D_{it}), cash flow to assets (CF_{it}) and sales to assets (S_{it}). We delete five firms from the original sample because they have aberrant values for some of these variables. Summary statistics are provided in Table 6.

- insert Table 6 about here -

We begin by estimating a homogeneous panel data model for investmentI_{it} with lagged Tobin's Q, sales, debt and cash flow as regressors. The lagged sales to assets ratio can be interpreted as a proxy for future demand for a firm's output and, following Hsiao and Tahmiscioglu (1997) and Hu and Schiantarelli (1998), is included as an additional control for future profit opportunities of the firm. In addition to fixed individual effects, we include fixed time effects to capture macroeconomic effects on investment.

We then apply the LM test of homogeneity developed in Section 3.1, using each of the regressor variables (lagged Q, debt, cash flow and sales) as "candidate" transition variables, again following Hu and Schiantarelli (1998). We only test homogeneity of the coefficients of these firm-specific variables, which implies assuming that macroeconomic effects on investment do not differ across firms. Restricting coefficients of some variables to be constant in the PSTR model has no effect on the distribution theory.

Table 7 shows that homogeneity is strongly rejected for all four candidate transition variables for m_a=1, 2, and 3. Neither the LM-type tests based on the asymptotic χ^2 distributions, their F-versions, nor the wild bootstrap tests are able to identify a specific transition variable as almost all p-values are practically equal to zero. However, the results from the HAC version are quite informative. The p-values of the tests with Tobin's Q as the transition variable are considerably smaller than the others, suggesting that Q is the most appropriate choice of a transition variable.

Next we apply the sequence of tests discussed at the end of Section 3.1 to determine the order m of the logistic transition function. In Table 8, we report the results of the specification test sequence for all four candidate transition variables. It can be seen that $m = 1$ is the best choice for Q according to the HAC test, while the wild bootstrap tests are not informative as their p-values are all equal to zero. Interestingly, significance of the values of the WB and WCB test statistics for debt, cash flow and sales in Table 8 diminishes with increasing m. This lends further support to the choice of Q as the transition variable. Thus we proceed with estimating the following PSTR model:

$$\begin{aligned} I_{it} = \mu_i + \lambda_t + \beta_{01} Q_{i,t-1} + \beta_{02} S_{i,t-1} + \beta_{03} D_{i,t-1} + \beta_{04} CF_{i,t-1} \\ + (\beta_{11} Q_{i,t-1} + \beta_{12} S_{i,t-1} + \beta_{13} D_{i,t-1} + \beta_{14} CF_{i,t-1})\, g(Q_{i,t-1}; \gamma, c) + u_{it} \end{aligned} \quad (20)$$

where λ_t denotes the fixed time effects, and

$$g(Q_{i,t-1}; \gamma, c) = \left(1 + \exp\left(-\gamma(Q_{i,t-1} - c)\right)\right)^{-1}, \quad \text{with } \gamma > 0. \quad (21)$$

- insert Tables 7 and 8 about here -

Before discussing the estimation results in detail, we examine the adequacy of the two-regime PSTR model by applying the misspecification tests of parameter constancy and of no remaining heterogeneity. The results are reported in Table 9. Results from the WB and WCB tests that take both heteroskedasticity and possible within-cluster dependence into account suggest that the estimated model with one transition is adequate.

- insert Table 9 about here -

Parameter estimates appear in Table 11, together with cluster-robust and heteroskedasticity consistent standard errors, see Cameron, Gelbach, and Miller (2011). To facilitate interpretation, we report estimates of β_{0j} and $\beta_{0j} + \beta_{1j}$, for $j = 1, \ldots, 4$, corresponding to regression coefficients in the regimes associated with $g(Q_{i,t-1}; \gamma, c) =$ 0 and 1, respectively. The estimate of γ is such that the transition from the lower regime associated with small values of Tobin's Q to the upper regime with large values of Q is smooth. This is seen from Figure 1, in which the transition function is plotted against Tobin's Q with each circle representing an observation. Also note that the point estimate $\hat{c} = 0.49$ is in between the 25th and 50th percentiles of the empirical distribution of $Q_{i,t-1}$, see Table 6. Hence, the model identifies firms with limited growth opportunities, signalled by their rather low Q values, as a separate group that is distinct from firms with moderate or good growth opportunities.

- insert Table 11 and Figure 1 about here -

One of the key characteristics of the PSTR model is that it allows for time-varying heterogeneity, in the sense that regression coefficients for a particular individual in the panel are not fixed over time. The importance of this feature is illustrated in Table 10, providing various summary statistics of the transition function $g(Q_{i,t-1}; \gamma, c)$, which directly determines the "effective" regression coefficients $\beta_{0j} + \beta_{1j} g(Q_{i,t-1}; \gamma, c)$ for firm i at time t. First, we observe that for all years except 1975 more than half of the firms are classified into the upper regime, in the sense that the value of the transition function exceeds 0.5. It is interesting to note that during the last four years of the sample period, the upper regime contains a substantially higher percentage of firms than during the first decade, at 77.5 percent on average for the years 1984-1987 compared to 58 percent during the years 1974-1983. Similarly, the mean and median values of $g(Q_{i,t-1}; \gamma, c)$ are larger than one half, again except in 1975, and also these take substantially larger values during the years 1984-1987. The 25th and 75th percentiles of the yearly transition function values show that this level shift is mostly due to a change in the lower part of the cross-sectional distribution, in the sense that the 25th percentile jumped from around 0.35 in 1983 to 0.50 in 1984-5 and even 0.60 in 1986-1987, whereas the 75th remained fairly constant. Finally, the rightmost column of Table 10 shows that the year-on-year changes of the transition function are quite substantial for individual firms as well, averaging 0.10 in absolute value. The importance of time-varying regression coefficients for individual firms is even better illustrated by the difference between the maximum and minimum values of $g(Q_{i,t-1}; \gamma, c)$ during the sample period. On average this equals 0.51; and for 109 out of the 560 firms in the panel it even exceeds 0.75. This clearly demonstrates that a fixed clustering of firms into constrained and unconstrained groups would be inappropriate.

- insert Table 10 about here -

Turning to the estimated regression coefficients, it is seen that the estimate of the coefficient on lagged debt is negative and significant for low Q firms, while it is insignificantly different from zero for high Q firms. This is consistent with the findings of Lang, Ofek, and Stulz (1996) that leverage matters for investment only for firms with poor growth opportunities or firms with growth opportunities that are not recognized by the market. The coefficient on lagged cash flow is positive and significant for both groups of firms, although it is considerably larger for low Q firms. This corroborates previous findings that internal finance is relevant for investment mainly for financially constrained firms. We also find that the coefficient on Tobin's Q is negative but insignificant for low Q firms and positive and highly significant for high Q firms. Hence, only firms with good growth prospects respond to changes in their investment opportunities, which is in line with the results of Hu and Schiantarelli (1998). Finally, for sales we find the same pattern as for Tobin's Q.

- insert Figure 2 about here -

Figure 2 shows the estimates of the time effects in the PSTR model, together with the lower and upper bounds of the 90% confidence interval based on cluster-robust and heteroskedasticity-consistent standard errors. These estimates are to be interpreted relative to the value of zero for 1974, the first year in the effective sample period. It is seen that there remains some variation in investment over time beyond what is explained by the included regressors. In particular, the estimates strongly suggest the presence of macroeconomic effects, as they closely follow the business cycle and growth cycle, being lower in 1975-1977, in 1982-1983, and in 1987 than during the remaining years, just like the average level of investment itself.

Finally, we acknowledge that our analysis is subject to caveats, including the possibility that cash flow and leverage contain useful information about growth opportunities not captured by Tobin's Q and the possibility of measurement error in Q. Both of these may lead to spurious effects of cash flow and leverage on investment, as discussed at length in Gilchrist and Himmelberg (1995), Erickson and Whited

(2000), Gomes (2001), and Hennessy (2004), among others. A thorough analysis of these issues, however, is beyond the scope and purpose of this paper.

6 Conclusions

Our panel smooth transition regression model incorporates heterogeneity by allowing regression coefficients to vary as a function of an exogenous variable and fluctuate between a limited number (often two) of extreme regimes. As the transition variable is individual-specific and time-varying, the regression coefficients for each of the individuals in the panel are changing over time. Our approach includes a modelling cycle for the PSTR model, containing tests of homogeneity, parameter constancy and no remaining heterogeneity. Monte Carlo experiments demonstrate that these statistics behave satisfactorily even in panels with small N and T, although the standard tests should be applied with caution given that they are considerably affected by cross-sectional heteroskedasticity. An application to firms' investment behaviour aptly demonstrates the usefulness of the model.

The PSTR model as considered in this paper has fixed effects and exogenous regressors. In some applications, these assumptions might seem unnecessarily restrictive. We therefore point out that our main motivation for choosing this set-up has been to highlight the incorporation of the smooth transition mechanism in a panel data context. Building a PSTR model under weaker assumptions is possible, and some generalizations are in fact fairly straightforward. First, assuming that the individual effects μ_i in (1) are random and independent of the exogenous variables x_{it}, the parameters of interest can be estimated using a feasible GLS procedure as proposed by Wooldridge (2002), which basically uses Arellano's estimator for the covariance matrix. This type of GLS estimator should be consistent and asymptotically normal independent of how T and N tend to infinity, see Hansen (2006) for details.

Second, the model can be extended to handle dynamic panel data by including lagged values of the dependent variable as regressors. This, of course, brings with it the usual complications related to estimation of dynamic panel data models in case T is fixed and finite, see Baltagi (2013, Chapter 8) and Pesaran (2015, Chapter 27).

Third, a model allowing for multiple variables entering the transition function might be relevant in practice and hence worthwhile considering. In Hu and Schiantarelli (1998), for example, several factors including Q, firm size and leverage jointly determine the classification of firms into regimes with different characteristics of investment behaviour. Such an extension is straightforward if the ensuing linear combination has fixed weights. Relaxing this restriction may lead to numerical complexities in estimating the parameters of the model.

Fourth and finally, it is also possible to consider a PSTR model in which not only the intercept but all the parameters are random across the cross-section units. In the static case, assuming that the parameters vary randomly and are independent of the regressors, several procedures are available for obtaining consistent estimates of the coefficient means, see Hsiao and Pesaran (2008). In the dynamic case, it is possible to use the group mean estimator, which involves estimating a PSTR model for each individual separately and averaging the coefficients, see Pesaran and Smith (1995). While this procedure gives consistent estimates, there is an important drawback in that estimating an STR model for each cross-sectional unit can be numerically complicated when T is small. One way to overcome this problem is to make the strong assumption that the parameters of the transition function (γ,c) are common across units and time. In this case, the pool mean group estimator by Pesaran, Shin, and Smith (1999) can be used. Detailed investigation of these and other possible extensions of the PSTR model is left for future research.

References

Allegret, J.-P., C. Couharde, D. Coulibaly, and V. Mignon (2014): "Current accounts and oil price fluctuations in oil-exporting countries: The role of financial development," *Journal of International Money and Finance*, 47, 185–201.

Almeida, H., and M. Campello (2007): "Financial constraints, asset tangibility, and corporate investment," *Review of Financial Studies*, 20, 1429–1460.

Andrews, D. W. K., and W. Ploberger (1994): "Optimal tests when a nuisance parameter is present only under the alternative," *Econometrica*, 62, 1383–1414.

Arellano, M. (1987): "Computing robust standard errors for within-group estimators," *Oxford Bulletin of Economics and Statistics*, 49, 431–434.

Aslanidis, N., and A. Xepapadeas (2006): "Smooth transition pollution-income paths," *Ecological Economics*, 57, 182–189.

——— (2008): "Regime switching and the shape of the emission-income relationship," *Economic Modelling*, 25, 731–739.

Baltagi, B. H. (2013): *Econometric Analysis of Panel Data, 5th edition*. Chichester: Wiley.

Baltagi, B. H., and Q. Li (1995): "Testing AR(1) against MA(1) disturbances in an error component model," *Journal of Econometrics*, 68, 133–151.

Becker, R., and D. R. Osborn (2010): "Weighted smooth transition regressions," *Journal of Applied Econometrics*, 27, 795–811.

Béreau, S., A. López-Villavicencio, and V. Mignon (2010): "Nonlinear adjustment of the real exchange rate towards its equilibrium value: A panel smooth transition error correction modelling," *Economic Modelling*, 27, 404–416.

——— (2012): "Currency misalignments and growth: A new look using nonlinear panel data methods," *Applied Economics*, 44, 3503–3511.

Besseca, M., and J. Fouquau (2008): "The non-linear link between electricity consumption and temperature in Europe: A threshold panel approach," *Energy Economics*, 30, 2705–2721.

Bond, S. R., and C. Meghir (1994): "Dynamic investment models and the firm's financial policy," *Review of Economic Studies*, 61, 197–222.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008): "Bootstrap-Based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics*, 90, 414–427.

——— (2011): "Robust Inference With Multiway Clustering," *Journal of Business & Economic Statistics*, 29, 238–249.

Carpenter, R. E., S. M. Fazzari, and B. C. Petersen (1994): "Inventory investment, internal-finance fluctuations, and the business cycle," *Brookings Papers on Economic Activity*, 2, 75–138.

Chen, Y.-S., C.-Y. Shih, and C.-S. Chang (2014): “Explore the new relationship between patents and market value: a panel smooth transition regression (PSTR) approach,” *Scientometrics*, 98, 1145–1159.

Cheng, C.-H., and P.-C. Wu (2013): “Nonlinear earnings persistence,” *International Review of Economics and Finance*, 25, 156–168.

Cho, D. (2015): “The role of covered interest parity in explaining the forward premium anomaly within a nonlinear panel framework,” *Journal of Empirical Finance*, 34, 229–238.

Chuluun, T., C. S. Eun, and R. Kiliç (2011): “Investment intensity of currencies and the random walk hypothesis: Cross-currency evidence,” *Journal of Banking and Finance*, 35, 372–387.

Davies, R. B. (1977): “Hypothesis testing when a nuisance parameter is present only under the alternative,” *Biometrika*, 64, 247–254.

——— (1987): “Hypothesis testing when a nuisance parameter is present only under the alternative,” *Biometrika*, 74, 33–43.

Delatte, A.-L., J. Fouquau, and R. Portes (2017): “Regime-dependent sovereign risk pricing during the Euro crisis,” *Review of Finance*, 21, 363–385.

Delatte, A.-L., M. Gex, and A. López-Villavicencio (2012): “Has the CDS market influenced the borrowing cost of European countries during the sovereign crisis?,” *Journal of International Money and Finance*, 31, 481–497.

Eitrheim, Ø., and T. Teräsvirta (1996): “Testing the adequacy of smooth transition autoregressive models,” *Journal of Econometrics*, 74, 59–75.

Erickson, T., and T. M. Whited (2000): “Measurement error and the relationship between investment and Q,” *Journal of Political Economy*, 108, 1027–1057.

Espinoza, R., H. Leon, and A. Prasad (2012): “When should we worry about inflation?,” *World Bank Economic Review*, 26, 100–127.

Fazzari, S. M., R. G. Hubbard, and B. C. Petersen (1988): “Financing constraints and corporate investment,” *Brookings Papers on Economic Activity*, 1, 141–195.

Fee, C. E., C. J. Hadlock, and J. R. Pierce (2009): “Investment, financing constraints, and internal capital markets: Evidence from the advertising expenditures of multinational firms,” *Review of Financial Studies*, 22, 2361–2392.

Fouquau, J., C. Hurlin, and I. Rabaud (2008): “The Feldstein-Horioka puzzle: A panel smooth transition regression approach,” *Economic Modelling*, 25, 284–299.

Geng, N. (2011): “The dynamics of market structure and firm-level adjustment to India’s pro-market economic liberalization reforms, 1988-2006: A time-varying panel smooth transition regression (TV-PSTR) approach,” *International Review of Economics and Finance*, 20, 506–519.

Giacomini, R., D. N. Politis, and H. White (2013): “A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators,” *Econometric Theory*, 29, 567–589.

Gilchrist, S., and C. P. Himmelberg (1995): "Evidence on the role of cash flow for investment," *Journal of Monetary Economics*, 36, 541–572.

Gomes, J. F. (2001): "Financing investment," *American Economic Review*, 91, 1263–1285.

Goodwin, B. K., M. T. Holt, and J. P. Prestemon (2011): "North American oriented strand board markets, arbitrage activity, and market price dynamics: A smooth transition approach," *American Journal of Agricultural Economics*, 93, 993–1014.

Granger, C. W. J., and T. Teräsvirta (1993): *Modelling Nonlinear Economic Relationships*. Oxford: Oxford University Press.

Hall, B. H., and R. E. Hall (1993): "The value and performance of U.S. corporations.," *Brookings Papers on Economic Activity*, pp. 1–34.

Hansen, B. E. (1996): "Inference when a nuisance parameter is not identified under the null hypothesis," *Econometrica*, 64, 413–430.

——— (1999): "Threshold effects in non-dynamic panels: Estimation, testing and inference," *Journal of Econometrics*, 93, 345–368.

Hansen, C. (2006): "Generalized least squares inference in panel and multilevel models with serial correlation and fixed effects," *Journal of Econometrics*, 140, 670–694.

——— (2007): "Asymptotic properties of a robust variance matrix estimator for panel data when T is large," *Journal of Econometrics*, 141, 597–620.

Hennessy, C. A. (2004): "Tobin's Q, debt overhang, and investment," *Journal of Finance*, 59, 1717–1742.

Hennessy, C. A., A. Levy, and T. M. Whited (2007): "Testing Q theory with financing frictions," *Journal of Financial Economics*, 83, 691–717.

Hovakimian, G. (2009): "Determinants of investment cash flow sensitivity," *Financial Management*, 38, 161–183.

Hovakimian, G., and S. Titman (2006): "Corporate investment with financial constraints: Sensitivity of investment to funds from voluntary asset sales," *Journal of Money, Credit and Banking*, 38, 357–374.

Hsiao, C. (2003): *Analysis of Panel Data, 2nd edition*. Cambridge: Cambridge University Press.

Hsiao, C., and M. H. Pesaran (2008): "Random coefficient models," in *The Econometrics of Panel Data (Third Edition)*, ed. by L. Matyas, and P. Sevestre, pp. 185–213. Berlin: Springer Verlag.

Hsiao, C., and A. K. Tahmiscioglu (1997): "A panel analysis of liquidity constraints and firm investment," *Journal of the American Statistical Association*, 92, 455–465.

Hu, X., and F. Schiantarelli (1998): "Investment and capital market imperfections: A switching regression approach using U.S. firm panel data," *Review of Economics and Statistics*, 80, 466–479.

Hurn, A. S., A. Silvennoinen, and T. Teräsvirta (2016): "A Smooth Transition Logit Model of the Effects of Deregulation in the Electricity Market," *Journal of Applied Econometrics*, 31, 707–733.

Jansen, E. S., and T. Teräsvirta (1996): "Testing parameter constancy and super exogeneity in econometric equations," *Oxford Bulletin of Economics and Statistics*, 58, 735–763.

Lang, L., E. Ofek, and R. M. Stulz (1996): "Leverage, investment, and firm growth," *Journal of Financial Economics*, 40, 3–29.

Lundbergh, S., T. Teräsvirta, and D. van Dijk (2003): "Time-varying smooth transition autoregressive models," *Journal of Business and Economic Statistics*, 21, 104–121.

Luukkonen, R., P. Saikkonen, and T. Teräsvirta (1988): "Testing linearity against smooth transition autoregressive models," *Biometrika*, 75, 491–499.

Moyen, N. (2004): "Investment-cash flow sensitivities: Constrained versus unconstrained firms," *Journal of Finance*, 59, 2061–2092.

Omay, T., R. van Eyden, and R. Gupta (2017): "Inflation–growth nexus: evidence from a pooled CCE multiple-regime panel smooth transition model," *Empirical Economics*, pp. 1–32.

Pesaran, M. H. (2015): *Time Series and Panel Data Econometrics*. Oxford: Oxford University Press.

Pesaran, M. H., Y. Shin, and R. Smith (1999): "Pooled mean group estimation of dynamic heterogenous panels," *Journal of the American Statistical Association*, 94, 621–634.

Pesaran, M. H., and R. Smith (1995): "Estimating long-run relationships from dynamic heterogenous panels," *Journal of Econometrics*, 68, 79–113.

Schleer, F. (2015): "Finding Starting-Values for the Estimation of Vector STAR Models," *Econometrics*, 3, 65–90.

Seleteng, M., M. Bittencourt, and R. van Eyden (2013): "Non-linearities in inflation-growth nexus in the SADC region: A panel smooth transition regression approach," *Economic Modelling*, 30, 149–156.

Teräsvirta, T. (1994): "Specification, estimation, and evaluation of smooth transition autoregressive models," *Journal of the American Statistical Association*, 89, 208–218.

Teräsvirta, T. (1998): "Modelling economic relationships with smooth transition regressions," in *Handbook of applied economic statistics*, ed. by A. Ullah, and D. E. A. Giles, pp. 507–552. New York: Marcel Dekker.

Teräsvirta, T., D. Tjøstheim, and C. W. J. Granger (2010): *Modelling Nonlinear Economic Relationships*. Oxford: Oxford University Press.

van Dijk, D., T. Teräsvirta, and P. H. Franses (2002): "Smooth transition autoregressive models - a survey of recent developments," *Econometric Reviews*, 21, 1–47.

Whited, T. M. (1992): "Debt, liquidity constraints, and corporate investment: Evidence from panel data," *Journal of Finance*, 47, 1425–1459.

Wooldridge, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Table 1: Empirical size of tests of homogeneity

		T=5						T=10						T=20					
		LM		HAC		Bootstrap		LM		HAC		Bootstrap		LM		HAC		Bootstrap	
N	m_a	χ^2	F	χ^2	F	WB	WCB	χ^2	F	χ^2	F	WB	WCB	χ^2	F	χ^2	F	WB	WCB
Panel (a): Homoskedasticity																			
20	1	10.07	3.90	3.56	0.62	4.55	4.94	6.54	4.22	3.33	1.73	5.24	5.12	5.87	4.68	3.97	2.89	5.01	4.72
	2	11.70	2.87	1.67	0.01	4.46	4.83	7.49	4.07	1.65	0.42	5.19	4.95	5.96	4.48	2.03	0.95	5.13	5.12
	3	13.37	1.92	0.68	0.00	4.15	4.76	8.22	3.32	0.95	0.09	5.05	5.19	6.06	4.06	0.89	0.26	4.90	5.08
40	1	9.77	4.69	4.50	1.39	5.11	5.29	6.63	4.52	4.31	2.65	5.34	5.51	5.71	4.76	4.46	3.52	5.31	5.00
	2	11.36	4.29	3.01	0.27	4.67	5.37	7.29	4.45	3.45	1.58	5.06	5.55	6.09	4.69	3.56	2.35	5.47	4.99
	3	12.40	3.59	2.03	0.04	4.59	5.46	7.71	4.26	2.31	0.76	4.92	5.52	5.99	4.47	2.32	1.48	5.51	5.43
80	1	8.51	4.26	4.18	1.61	5.43	5.37	6.83	5.17	4.74	3.31	5.20	4.99	5.75	4.83	4.57	3.79	5.31	4.98
	2	10.85	4.58	3.75	0.81	4.60	5.19	7.07	4.81	3.87	1.95	5.12	5.27	5.86	4.71	4.05	3.08	5.21	5.19
	3	12.16	4.10	2.97	0.41	4.63	5.24	7.88	4.46	2.83	1.26	4.83	5.22	6.09	4.72	3.39	2.32	5.37	5.30
160	1	8.89	4.40	4.36	1.80	5.18	5.47	6.74	4.93	5.10	3.55	5.07	4.85	5.57	4.67	4.62	3.88	5.17	5.31
	2	10.58	4.58	4.21	1.19	4.94	5.59	7.52	5.03	4.22	2.62	4.98	4.96	6.43	5.30	4.76	3.69	5.04	5.49
	3	11.95	4.43	3.55	0.53	4.65	5.23	7.91	4.79	3.81	2.00	4.85	5.10	6.36	5.00	4.33	3.12	5.10	5.42
Panel (b): Heteroskedasticity																			
20	1	15.66	7.51	3.21	0.46	6.42	5.54	15.94	11.86	3.99	1.97	7.52	5.17	15.54	13.55	3.78	2.73	7.86	5.19
	2	19.05	6.84	1.45	0.03	6.18	5.68	18.36	12.34	1.85	0.38	7.46	5.38	18.79	15.57	2.10	1.14	7.90	5.11
	3	21.65	5.08	0.56	0.00	5.83	5.60	20.44	11.69	0.85	0.02	7.05	4.93	20.86	16.69	0.90	0.29	7.27	5.25
40	1	16.08	9.08	4.17	1.19	6.56	5.86	15.88	12.51	4.70	2.84	8.20	5.16	16.91	15.20	4.54	3.53	8.42	5.54
	2	20.38	9.66	2.38	0.26	6.03	5.48	20.51	14.74	3.19	1.34	7.66	4.95	20.78	18.37	3.33	2.15	8.59	5.54
	3	23.61	9.67	1.75	0.05	6.19	5.42	23.02	16.00	2.42	0.71	7.30	5.13	23.95	19.94	2.23	1.35	7.64	5.16
80	1	16.36	9.50	4.44	1.78	6.55	5.32	16.92	13.40	5.00	3.25	7.35	5.11	16.66	15.14	4.55	3.82	8.67	5.09
	2	21.41	10.99	3.83	0.73	6.63	5.50	21.16	15.47	3.67	2.09	7.34	5.10	21.29	18.93	3.77	2.75	8.49	5.20
	3	25.37	12.08	3.21	0.41	6.54	5.36	24.23	17.38	3.03	1.38	7.24	5.23	24.94	21.58	3.14	2.20	8.08	5.19
160	1	16.67	10.38	5.05	2.24	7.11	4.97	16.46	13.33	4.95	3.51	8.21	5.39	17.68	16.06	4.73	4.07	8.37	5.46
	2	22.47	12.76	4.30	1.15	6.83	5.36	21.18	16.41	4.28	2.58	7.75	5.70	23.01	20.81	4.42	3.35	8.10	5.22
	3	27.03	14.16	3.87	0.78	7.12	5.16	25.39	19.17	3.69	1.85	7.31	5.56	26.68	23.50	3.80	2.64	7.91	5.23

Note: The table presents the empirical size of the tests of homogeneity: the standard LM test (LM) and robust (HAC) test based on the asymptotic χ^2 and F distributions, the wild bootstrap (WB) LM test and wild cluster bootstrap (WCB) test, based on a significance level of 0.05. The DGP is a homogeneous panel model with fixed effects and with homoskedastic (Panel (a)) or heteroskedastic (Panel (b)) errors, as described in Section 4.1. The bootstrap procedures make use of the warp-speed method, and rejection frequencies are based on 10,000 replications.

Table 2: Empirical power of the wild cluster bootstrap (WCB) test of homogeneity

			$m = 1$			$m = 2$		
N	m_a	$T =$	5	10	20	5	10	20
Panel (a): Homoskedasticity								
20	1		39.40	74.40	96.70	6.05	5.91	11.89
	2		31.34	67.55	94.86	15.28	34.52	66.61
	3		27.29	64.71	94.65	15.09	31.36	62.60
40	1		72.63	98.05	100.00	6.52	6.27	20.19
	2		63.12	96.17	99.98	29.94	65.36	95.64
	3		59.23	95.86	100.00	29.94	61.23	94.31
80	1		97.28	100.00	100.00	7.37	7.81	37.39
	2		94.71	100.00	100.00	58.80	94.80	100.00
	3		93.97	99.98	100.00	59.78	93.66	99.99
160	1		100.00	100.00	100.00	8.68	10.87	66.54
	2		99.97	100.00	100.00	90.40	99.95	100.00
	3		99.98	100.00	100.00	91.22	99.94	100.00
Panel (b): Heteroskedasticity								
20	1		24.74	44.73	71.09	5.98	6.14	8.21
	2		20.06	36.87	63.78	11.22	19.48	34.06
	3		18.33	34.21	62.67	10.80	17.42	30.94
40	1		47.05	76.31	95.77	6.05	6.23	11.43
	2		37.41	67.52	93.27	18.31	34.02	63.23
	3		33.95	65.72	93.58	17.26	30.80	58.28
80	1		78.20	97.28	99.95	6.32	6.53	18.83
	2		68.82	94.81	99.92	32.24	63.28	92.53
	3		65.92	94.80	99.93	33.04	58.83	90.29
160	1		97.95	99.99	100.00	6.84	7.39	32.66
	2		95.64	99.95	100.00	60.70	92.29	99.88
	3		95.24	99.98	100.00	61.37	90.59	99.81

Note: The table presents presents the empirical power of the wild cluster bootstrap (WCB) test of homogeneity, based on the auxiliary regression (4) with order m_a. Panels with cross-sectional dimension N and time dimension T are generated from the PSTR model (19) with $r = 1$ and with either a monotonically increasing ($m = 1$) or symmetric ($m = 2$) transition function (2), and with homoskedastic (Panel (a)) or heteroskedastic (Panel (b)) errors. The bootstrap makes use of the warp-speed method, and rejection frequencies are reported for a significance level of 0.05, based on 10,000 replications.

Table 3: Empirical size of the wild cluster bootstrap (WCB) test of parameter constancy (PC) and no remaining heterogeneity (NRH)

			PC			NRH		
N	h_a/m_a	$T =$	5	10	20	5	10	20
Panel (a): Homoskedasticity								
20	1		5.85	5.13	5.74	5.24	5.34	5.44
	2		4.93	5.54	5.48	3.97	5.64	5.29
	3		3.53	5.66	5.71	2.48	4.94	5.24
40	1		5.88	5.34	5.40	5.45	5.42	5.53
	2		5.66	5.13	4.98	5.30	5.30	5.22
	3		5.21	5.37	5.42	4.61	5.07	5.45
80	1		5.38	5.42	5.06	5.67	5.08	5.20
	2		5.36	5.43	5.15	5.98	4.91	5.28
	3		5.53	5.17	5.12	5.47	5.03	5.38
160	1		5.09	4.93	4.85	5.04	4.98	4.55
	2		4.91	4.76	5.02	4.97	4.81	4.67
	3		4.80	5.20	5.22	5.16	5.28	4.85
Panel (b): Heteroskedasticity								
20	1		6.94	6.46	6.26	6.06	6.11	6.41
	2		5.79	6.79	6.65	4.76	5.63	5.99
	3		3.75	6.29	6.33	2.83	5.12	5.32
40	1		6.39	6.17	5.94	6.05	5.97	5.89
	2		5.35	5.98	6.11	5.64	6.20	6.18
	3		4.94	5.98	6.23	4.53	5.35	5.54
80	1		5.68	5.88	5.58	5.70	5.50	5.40
	2		5.78	5.92	5.77	5.76	5.48	5.35
	3		5.18	5.76	5.84	5.26	5.17	5.63
160	1		5.65	5.81	5.58	5.43	4.91	5.45
	2		5.21	5.01	5.64	5.68	5.25	5.60
	3		5.05	5.27	5.42	5.14	5.40	5.60

Note: The table presents the empirical size of the wild cluster bootstrap (WCB) test of parameter constancy (PC), based on the auxiliary regression (15) with order h_a, and the test of no remaining heterogeneity (NRH) based on the auxiliary regression (17) with order m_a. Panels with cross-sectional dimension N and time dimension T are generated from the PSTR model (19) with $r = 1$ and a monotonically increasing ($m = 1$) transition function (2), and with homoskedastic (Panel (a)) or heteroskedastic (Panel (b)) errors. The bootstrap makes use of the warp-speed method, and rejection frequencies are reported for a significance level of 0.05, based on 10,000 replications.

Table 4: Empirical power of the wild cluster bootstrap (WCB) test of parameter constancy

		$h = 1$ at $0.5T$			$h = 2$ at $0.3T$ and $0.7T$		
N	h_a $T=$	5	10	20	5	10	20
Panel (a): Homoskedasticity							
20	1	46.45	97.23	100.00	47.98	50.10	27.12
	2	41.94	97.00	100.00	54.05	92.86	99.99
	3	34.37	96.32	100.00	53.89	96.12	99.98
40	1	85.12	100.00	100.00	85.73	80.75	48.56
	2	85.39	100.00	100.00	94.23	99.98	100.00
	3	81.14	100.00	100.00	97.92	100.00	100.00
80	1	99.59	100.00	100.00	99.60	99.02	79.94
	2	99.69	100.00	100.00	99.98	100.00	100.00
	3	99.55	100.00	100.00	100.00	100.00	100.00
160	1	100.00	100.00	100.00	100.00	100.00	98.34
	2	100.00	100.00	100.00	100.00	100.00	100.00
	3	100.00	100.00	100.00	100.00	100.00	100.00
Panel (b): Heteroskedasticity							
20	1	34.00	85.58	99.56	36.61	34.69	18.73
	2	30.47	83.73	99.37	38.85	74.94	97.65
	3	22.68	81.68	99.39	35.14	79.66	98.02
40	1	66.55	99.61	100.00	70.51	60.54	31.46
	2	66.79	99.54	100.00	81.25	98.37	100.00
	3	60.21	99.53	100.00	88.77	99.44	100.00
80	1	94.90	100.00	100.00	96.50	89.73	56.36
	2	95.76	100.00	100.00	99.43	100.00	100.00
	3	94.69	100.00	100.00	99.95	100.00	100.00
160	1	99.97	100.00	100.00	99.97	99.72	86.51
	2	99.98	100.00	100.00	100.00	100.00	100.00
	3	99.98	100.00	100.00	100.00	100.00	100.00

Note: The table presents presents the empirical power of the wild cluster bootstrap (WCB) test of parameter constancy, based on the auxiliary regression (15) with order h_a. Panels with cross-sectional dimension N and time dimension T are generated from the TV-PSTR model (12) either with monotonic change centred in the middle ($h = 1$ and $c = 0.5T$ in (14)) or with nonmonotonic change ($h = 2$ and $c_1 = 0.3T$ and $c_2 = 0.7T$ in (14)), and with homoskedastic (Panel (a)) or heteroskedastic (Panel (b)) errors. The bootstrap makes use of the warp-speed method, and rejection frequencies are reported for a significance level of 0.05, based on 10,000 replications.

Table 5: Empirical power of the wild cluster bootstrap (WCB) test of no remaining heterogeneity

N	m_a	$T=$	$\beta_2 = \beta_1$			$\beta_2 = -\beta_1$		
			5	10	20	5	10	20
Panel (a): Homoskedasticity								
20	1		7.08	13.97	17.91	14.37	46.18	84.50
	2		3.66	11.31	20.02	0.58	22.42	77.70
	3		3.54	10.10	18.26	0.85	9.24	74.80
40	1		11.54	16.46	23.04	36.92	84.71	99.25
	2		9.07	17.26	31.52	18.83	76.55	99.10
	3		7.54	16.07	29.97	9.55	73.20	99.24
80	1		15.49	20.93	37.57	73.90	99.33	100.00
	2		14.46	28.13	53.98	61.20	99.17	99.99
	3		13.32	27.94	56.81	55.92	99.31	100.00
160	1		17.75	32.98	62.29	97.39	100.00	100.00
	2		21.60	48.56	83.86	95.54	100.00	100.00
	3		21.24	52.23	89.54	96.08	100.00	100.00
Panel (b): Heteroskedasticity								
20	1		5.31	9.74	14.77	10.14	25.12	53.11
	2		2.85	7.45	12.08	0.85	11.63	40.50
	3		3.02	6.07	11.11	1.11	4.31	35.06
40	1		8.70	13.60	17.43	22.15	54.01	85.66
	2		6.92	11.49	16.97	9.28	41.80	80.56
	3		5.94	9.46	15.71	3.62	36.17	80.75
80	1		12.94	16.48	22.40	46.04	87.30	99.20
	2		9.18	14.82	27.86	33.40	80.75	98.94
	3		8.78	13.09	25.70	28.26	79.75	99.16
160	1		15.02	20.84	31.13	80.63	99.47	100.00
	2		12.92	25.18	45.32	71.58	99.02	99.99
	3		10.76	24.59	47.00	68.85	99.28	100.00

Note: The table presents the empirical power of the wild cluster bootstrap (WCB) test of no remaining heterogeneity (NRH) based on the auxiliary regression (17) with order m_a. Panels with cross-sectional dimension N and time dimension T are generated from the PSTR model (19) with $r = 2$ and monotonic ($\beta_2 = \beta_1$) or nonmonotonic ($\beta_2 = -\beta_1$) heterogeneity, and with homoskedastic (Panel (a)) or heteroskedastic (Panel (b)) errors. The bootstrap makes use of the warp-speed method, and rejection frequencies are reported for a significance level of 0.05, based on 10,000 replications.

Table 6: Summary statistics

	Mean	St.Dev.	10%	25%	50%	75%	90%
I_{it}	0.088	0.059	0.031	0.049	0.076	0.112	0.158
$Q_{i,t-1}$	1.053	1.201	0.224	0.37	0.671	1.286	2.282
$S_{i,t-1}$	1.843	0.949	0.899	1.271	1.696	2.225	2.835
$CF_{i,t-1}$	0.241	0.197	0.056	0.124	0.215	0.319	0.447
$D_{i,t-1}$	0.233	0.207	0.008	0.09	0.206	0.319	0.471

Note: The table presents the mean, standard deviation (St.Dev.) and selected percentiles of the ratios of investment to assets (I_{it}), Tobin's Q or total market value to assets (Q_{it}), sales to assets (S_{it}), cash flow to assets (CF_{it}) and long-term debt to assets (D_{it}), for a balanced panel of 560 US firms for the period 1973–1987.

Table 7: Homogeneity tests

	LM_χ		LM_F		HAC_χ		HAC_F		WB	WCB
m	test	p-val	test	p-val	test	p-val	test	p-val	p-val	p-val
Transition variable $Q_{i,t-1}$										
1	125.25	0	28.99	0	30.03	$4.8e{-}06$	6.95	$1.4e{-}05$	0	0
2	217.43	0	25.15	0	55.01	$4.4e{-}09$	6.36	$2.9e{-}08$	0	0
3	290.84	0	22.42	0	76.52	$1.9e{-}11$	5.90	$2.6e{-}10$	0	0
Transition variable $D_{i,t-1}$										
1	37.81	$1.2e{-}07$	8.75	$4.8e{-}07$	13.72	0.0082	3.18	0.013	0.0028	0.012
2	86.87	$2.0e{-}15$	10.05	$4.9e{-}14$	21.87	0.0052	2.53	0.0095	0	0
3	89.71	$5.6e{-}14$	6.91	$1.3e{-}12$	24.22	0.019	1.87	0.033	0	$8.0e{-}04$
Transition variable $CF_{i,t-1}$										
1	128.87	0	29.83	0	20.41	0.00041	4.72	0.00083	0	0
2	142.32	0	16.46	0	35.77	$1.9e{-}05$	4.14	$6.1e{-}05$	0	0
3	205.97	0	15.88	0	47.18	$4.3e{-}06$	3.64	$1.8e{-}05$	0	0
Transition variable $S_{i,t-1}$										
1	94.83	0	21.95	0	15.10	0.0045	3.49	0.0074	0	0
2	116.46	0	13.47	0	29.91	0.00022	3.46	0.00055	0	0
3	136.27	0	10.50	0	31.59	0.0016	2.43	0.0037	0	$6.0e{-}04$

Note: The table presents LM-type tests of homogeneity and corresponding p-values in the panel regression of investment on lagged Tobin's Q ($Q_{i,t-1}$), sales to assets ($S_{i,t-1}$), cash flow to assets ($CF_{i,t-1}$) and long-term debt to assets ($D_{i,t-1}$), for a balanced panel of 560 US firms for the period 1973–1987.

Table 8: Sequence of homogeneity tests for selecting order m of transition function

	LM_{χ}		LM_F		HAC_{χ}		HAC_F		WB	WCB
m	test	p-val	test	p-val	test	p-val	test	p-val	p-val	p-val
Transition variable $Q_{i,t-1}$										
H_{03}^*	75.50	$1.6e{-15}$	17.46	$2.9e{-14}$	24.62	$6.0e{-05}$	5.69	0.00014	0	0
H_{02}^*	93.68	0	21.67	0	22.12	0.00019	5.12	0.00041	0	0
H_{01}^*	125.25	0	28.99	0	30.03	$4.8e{-06}$	6.95	$1.4e{-05}$	0	0
Transition variable $D_{i,t-1}$										
H_{03}^*	2.87	0.58	0.66	0.62	1.32	0.86	0.30	0.88	0.0028	0.012
H_{02}^*	49.30	$5.0e{-10}$	11.41	$3.1e{-09}$	10.96	0.027	2.54	0.038	$2.0e{-04}$	0.0024
H_{01}^*	37.81	$1.2e{-07}$	8.75	$4.8e{-07}$	13.72	0.0082	3.18	0.013	0.8762	0.8954
Transition variable $CF_{i,t-1}$										
H_{03}^*	64.83	$2.8e{-13}$	14.99	$3.3e{-12}$	6.68	0.15	1.55	0.19	0	0
H_{02}^*	13.67	0.0084	3.16	0.013	1.89	0.76	0.44	0.78	0.716	0.704
H_{01}^*	128.87	0	29.83	0	20.41	0.00041	4.72	0.00083	0.0012	0.0124
Transition variable $S_{i,t-1}$										
H_{03}^*	20.11	0.00047	4.65	0.00095	2.91	0.57	0.67	0.61	0	0
H_{02}^*	21.89	0.00021	5.07	0.00045	5.26	0.26	1.22	0.30	0.1636	0.2072
H_{01}^*	94.83	0	21.95	0	15.10	0.0045	3.49	0.0074	0.286	0.4184

Note: The table presents the sequence of homogeneity tests for selecting the order m of the logistic transition function in a PSTR model for investment with lagged Tobin's Q ($Q_{i,t-1}$), sales to assets ($S_{i,t-1}$), cash flow to assets ($CF_{i,t-1}$) and long-term debt to assets ($D_{i,t-1}$), for a balanced panel of 560 US firms for the period 1973–1987. The listed null hypotheses have the implications H_{01}^*: $\beta_1^* = 0|\beta_3^* = \beta_2^* = 0$, H_{02}^*: $\beta_2^* = 0|\beta_3^* = 0$, and H_{03}^*: $\beta_3^* = 0$, respectively, in the auxiliary regression (4) with $m = 3$.

Table 9: Misspecification tests

m	LM_χ test	LM_χ p-val	LM_F test	LM_F p-val	HAC_χ test	HAC_χ p-val	HAC_F test	HAC_F p-val	WB p-val	WCB p-val
No remaining heterogeneity										
Transition variable $Q_{i,t-1}$										
$m_a = 1$	118.37	$1.4e{-15}$	5.20	$8.4e{-14}$	48.15	0.00066	2.12	0.0021	0.61	0.75
Transition variable $D_{i,t-1}$										
$m_a = 1$	131.40	0	5.78	$6.7e{-16}$	38.19	0.012	1.68	0.027	0.42	0.60
Transition variable $CF_{i,t-1}$										
$m_a = 1$	109.01	$7.1e{-14}$	4.79	$2.8e{-12}$	35.46	0.025	1.56	0.05	0.56	0.70
Transition variable $S_{i,t-1}$										
$m_a = 1$	162.56	0	7.15	0	33.35	0.042	1.47	0.078	0.75	0.82
Parameter constancy										
$h_a = 1$	54.45	$8.5e{-05}$	2.39	0.00035	132.28	0	5.82	$4.4e{-16}$	1.00	1.00

Note: The table presents misspecification tests in a PSTR model for investment with lagged Tobin's Q ($Q_{i,t-1}$), sales to assets ($S_{i,t-1}$), cash flow to assets ($CF_{i,t-1}$) and long-term debt to assets ($D_{i,t-1}$) as regressors and $Q_{i,t-1}$ as transition variable, for a balanced panel of 560 US firms for the period 1973–1987.

Table 10: Transition function statistics

Year	Perc. of firms with $g(Q_{i,t-1};\gamma,c) > 0.5$	Value of $g(Q_{i,t-1};\gamma,c)$ Mean	Median	25%	75%	Mean abs. change of $g(Q_{i,t-1};\gamma,c)$
1974	60.71	0.63	0.63	0.35	0.98	
1975	38.04	0.48	0.35	0.21	0.76	0.17
1976	51.61	0.57	0.51	0.27	0.93	0.12
1977	60.54	0.62	0.60	0.34	0.95	0.09
1978	59.82	0.63	0.61	0.34	0.95	0.09
1979	60.36	0.62	0.62	0.33	0.95	0.09
1980	60.89	0.62	0.62	0.33	0.94	0.09
1981	65.36	0.66	0.69	0.36	0.97	0.09
1982	59.64	0.63	0.65	0.31	0.97	0.10
1983	63.21	0.66	0.72	0.35	0.98	0.10
1984	76.96	0.75	0.90	0.52	1.00	0.12
1985	73.21	0.73	0.86	0.48	0.99	0.07
1986	78.75	0.77	0.93	0.57	1.00	0.08
1987	80.89	0.79	0.94	0.62	1.00	0.07
Avg	63.57	0.65	0.69	0.38	0.95	0.10

Note: The table presents summary statistics for the value of the transition function $g(Q_{i,t-1};\gamma,c)$ in a PSTR model for investment with lagged Tobin's Q ($Q_{i,t-1}$), sales to assets ($S_{i,t-1}$), cash flow to assets ($CF_{i,t-1}$) and long-term debt to assets ($D_{i,t-1}$) as regressors and $Q_{i,t-1}$ as transition variable, for a balanced panel of 560 US firms for the period 1973–1987.

Table 11: Estimation results of two-regime PSTR model, where the standard errors are obtained by using the cluster-robust and heteroskedasticity-consistent covariance estimator allowing for error dependency within individual firms.

	estimates	s.e.
Low Q firms		
$\beta_{0j} \times 100$		
$Q_{i,t-1}$	-1.72	6.77
$S_{i,t-1}$	-0.05	0.41
$D_{i,t-1}$	-6.39	2.04
$CF_{i,t-1}$	8.60	3.12
High Q firms		
$(\beta_{0j} + \beta_{0j}) \times 100$		
$Q_{i,t-1}$	0.68	0.12
$S_{i,t-1}$	0.99	0.31
$D_{i,t-1}$	0.38	1.17
$CF_{i,t-1}$	3.93	1.24
γ	4.95	1.21
c	0.49	0.25

Figure 1: Estimated transition function (21) of the PSTR model (20). Each circle represents an observation.

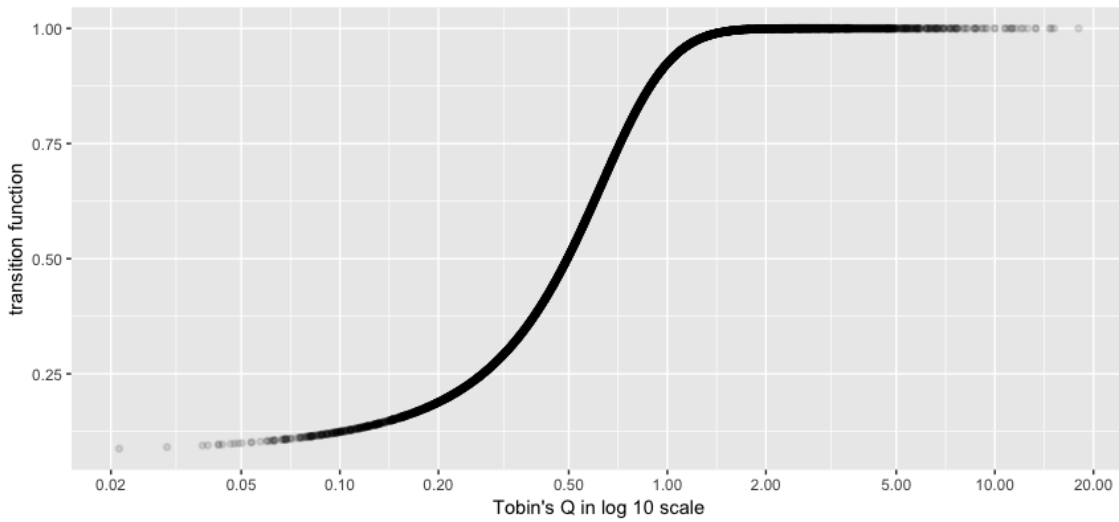

Figure 2: Coefficient estimates of yearly dummies (solid line) with lower and upper bound of 90% confidence intervals based on cluster-robust and heteroskedasticity-consistent standard errors.