Credit Risk and Business Cycle over Different Regimes

Juri Marcucci and Mario Quagliariello*
Bank of Italy

Second draft: September 2006. Preliminary. Please do not quote

ABSTRACT

In the recent banking literature, the relationships between credit risk and the business cycle have been analyzed for both (macro) financial stability and (micro) risk management purposes. The vast majority of these studies generally neglect the presence of asymmetric effects, i.e., the possibility that the impact is dissimilar over different phases of the business cycle. In this paper, we try to make a step forward and shed some light on these open issues. For our analysis, we employ Threshold Regression models with two or more regimes both at the aggregate and at the individual level, exploiting a unique dataset on Italian bank borrowers' default rates. In particular, we analyze whether the relationship between business cycle and credit risk is subject to regime switches, determining endogenously the thresholds. Furthermore, we test whether the impact of the business cycle is more pronounced when credit risk starts at higher levels, endogenously identifying the risk threshold over/below which such impact is different. Our results suggest that the impact of the business cycle is more pronounced when starting credit risk levels are higher and during downturns.

JEL Classification: C22, C23, G21, G28

Keywords: Threshold Panel Data Models, Credit Risk, Default Rate, Business Cycle, Regime Switching

---

* The views expressed are those of the authors and do not necessarily reflect those of the Bank of Italy.

Contents

1. Introduction

In the recent banking literature, the relationship between credit risk and the business cycle has been analyzed for both (macro) financial stability and (micro) risk management purposes. Indeed, the potential impact of the economic developments on banks' portfolios is relevant for both policy makers, interested in forecasting and preventing banks' instability due to unfavorable economic conditions, and risk managers, who pay attention on the robustness of their capital allocation plans under different scenarios. These different perspectives are not mutually exclusive; the reform of the Basel Accord on banks' capital requirements made it clear the need to match the micro and macro dimensions.

From a macro prudential point of view, many analyses have quantified the effects of macroeconomic conditions on asset quality (for a survey, see Quagliariello, in press). As an example, Pesola (2001) shows that shortfalls of GDP growth below forecast contributed to the banking crises in the Nordic countries, while Salas and Saurina (2002) document that macroeconomic shocks are quickly transmitted to Spanish banks' portfolio riskiness. Similarly, using Italian data, Marcucci and Quagliariello (in press) find that bank borrowers' default rates increase in bad macroeconomic times. Meyer and Yeager (2001) and Gambera (2000) document that a small number of macroeconomic variables are good predictors for the non-performing loan ratio in the US. Similarly, Hoggarth *et al.* (2005) provide evidence of a link between the state of the UK business cycle and banks' write-offs. Analogous evidence is provided in cross-country comparisons by Bikker and Hu (2002), Laeven and Majoni (2003) and Valckx (2003).

However, the vast majority of these studies generally neglect asymmetric effects, i.e., the possibility that the impact on banks' portfolios is different in different phases of the business cycle. An exception is the paper by Gasha and Morales (2004) who apply a SETAR model to country-level data and show that GDP growth affects non-performing loans only below a certain threshold.

By contrast, these asymmetries are somewhat taken into account in a number of studies on credit risk management. In particular, some analyses on the properties of credit rating transition matrices over the cycle have analyzed whether transition probabilities are affected to a larger (smaller) extent by recessionary (expansionary) conditions. Regime switching models are the tool commonly used for this kind of investigations. On the basis of GDP growth, Nickell *et al.* (2000) divide the business cycle into three categories (peaks, normal times and troughs) and find that, in peaks, low-rated bonds are less prone to downgrades. Default probabilities are particularly sensitive to the business cycle. They also note that the effect of the cycle on investment-grade obligors is more to raise volatility than to shift ratings systematically down. The impact of macroeconomic conditions appears therefore to be asymmetric and dependent on the starting creditworthiness of each borrower.

Bangia *et al.* (2002) in their analysis of the linkage between macroeconomic conditions and migration matrices distinguish two states of the economy, expansion and recession, and condition the migration matrix to these states. Their findings suggest that the downgrading probabilities, particularly in the extreme classes, increase significantly in recessions. Pederzoli and Torricelli (2005) adopt a similar framework in order to assess the impact of the business cycle on capital requirements under Basel II. This approach requires the identification of expansions/recessions based on some external sources. Moreover, discrete regime switching models may reveal unsatisfactory for dynamic credit risk management. For example, Lucas and Klaassen (2005) point out that the combination of an insufficient distinction between multiple economic regimes as well as a lacking identification of these regimes may weaken on the ability of these tools to discriminate between default regimes[1]. In particular, they show that implied asset correlations and default rate volatilities are biased towards zero and implausibly low. A further shortcoming of this literature is that the hypothesis that asymmetries depend on the severity of the recession rather than on the dichotomy expansion/recession is completely neglected.

Overall, the existing literature provides an incomplete picture of the evolution of credit quality over the business cycle. In this paper, we try to make a relevant step forward and shed some light on these open issues. For our analysis, we employ Threshold Regression approach and exploit a unique dataset on Italian bank borrowers' default rates. In particular, we analyze whether the relationship between business cycle and credit risk is subject to regime switches and endogenously determine the thresholds at which the system switches from one regime to the other. Furthermore, we test whether the impact of the business cycle is more pronounced when starting credit risk levels are higher identifying endogenously the risk threshold over/below which such impact is different. We also suggest a 4-regime approach which allows us to provide a very comprehensive discussion on the behavior of default rates over changing economic conditions.

---

[1] Most of the work in this field employs the NBER business cycle classifications. Lucas and Klaassen (2005) "cast serious doubts" on their use.

We find that the impact of the business cycle is more pronounced when credit risk levels are high and in unfavorable economic conditions. Under risk sensitive capital requirements, this evidence may provide some guidance to banks and supervisors in the choice of adequate capital buffers over different phases of the business cycle. Furthermore, the methodology we propose may be easily adapted for stress testing banks' portfolios.

The paper is organized as follows. Section 2 presents the threshold regression models adopted both for the aggregate data and for the panel data with 2 or more regimes. Section 3 describes the data on Italian banks' portfolios while in section 4 we comment on the empirical results. Finally, section 5 draws some concluding remarks and directions for further research.

2. Threshold regression models

Threshold regression models are quite popular in the non-linear time-series literature because they are relatively simple to estimate and interpret. The idea of approximating a general non-linear structure by a threshold regression model with a small number of regimes is due to Tong (see Tong, 1983 for an early review and Tong, 1990 for a deeper review). When the discontinuity in the threshold is replaced by a smooth transition function, the model can be generalized into a smooth transition model (see Chan and Tong, 1986, Granger and Teräsvirta, 1993, and Teräsvirta *et al.*, 1994).

In this section we briefly introduce the threshold regression model for the aggregate time series of default rate, the threshold model for panel data with 2 or more regimes defined over the same threshold variable and the threshold model for panel data with 4 regimes identified through two different threshold variables.

Our starting hypothesis is that the default rate is affected by the business cycle and that such an impact is subject to one or more regime-switch.

2.1 Aggregate model with two regimes

If the observed data on the dependent variable (default rate) are $dr_1, \ldots, dr_T$, the simplest model that relates the latter with the macroeconomic conditions (proxied by a measure of the output gap, GAP_t) with no thresholds and 1 regime takes the form

$$dr_t = \beta_{01} + \beta_{11} GAP_{t-1} + e_t \quad (1)$$

This model can be simply estimated by OLS. A more general model allows for the presence of two regimes defined by an observable threshold variable q_t that can be either the dependent or the independent variable.[2] In the former case, the model becomes

$$dr_t = \left(\beta_{01} + \beta_{11} GAP_{t-1}\right) I\left(dr_t \le \gamma\right) + \left(\beta_{02} + \beta_{12} GAP_{t-1}\right) I\left(dr_t > \gamma\right) + e_t \tag{2}$$

where $I(\cdot)$ is the indicator function. In this case we are distinguishing the 2 regimes with respect to the overall financial conditions. The model gives an estimate of γ which can be viewed as the threshold $\hat{\gamma}$ that characterizes good financial conditions $I\left(dr_t \le \hat{\gamma}\right)$ from bad financial conditions $I\left(dr_t > \hat{\gamma}\right)$. In the latter case, the model changes only in the threshold variable taking the form

$$dr_t = \left(\beta_{01} + \beta_{11} GAP_{t-1}\right) I\left(GAP_{t-1} \le \gamma\right) + \left(\beta_{02} + \beta_{12} GAP_{t-1}\right) I\left(GAP_{t-1} > \gamma\right) + e_t \tag{3}$$

Here we characterize each regime depending on the general macroeconomic conditions, distinguishing between recessionary conditions $I\left(GAP_{t-1} \le \hat{\gamma}\right)$ and expansionary phases $I\left(GAP_{t-1} > \hat{\gamma}\right)$. In models (2) and (3) the error e_t is assumed to be a martingale difference sequence with respect to the sigma algebra generated by the past history of the variables. The models can be more compactly represented as

$$dr_t = GAP_{t-1}\left(\gamma\right)'\theta + e_t \tag{4}$$

where $GAP_{t-1}\left(\gamma\right) = \left(GAP_{t-1}^{'} I\left(q_t \le \gamma\right), GAP_{t-1}^{'} I\left(q_t > \gamma\right)\right)$, $GAP_{t-1}^{'} = \left(1, GAP_{t-1}\right)^{'}$ and $\theta = \left(\beta_{01}, \beta_{02}, \beta_{11}, \beta_{12}\right)$. The parameters of interest are the coefficients θ and the threshold γ. Even though the regression equation (4) is non linear in the parameters, it can be estimated through least squares (LS). Under the additional assumption that $e_t \sim N\left(0, \sigma^2\right)$, LS is equivalent to maximum likelihood estimation. These models can be estimated by sequential conditional LS. For any given value of γ, the LS estimate of θ is

$$\hat{\theta}\left(\gamma\right) = \left(\sum_{t=1}^{T} GAP_{t-1}\left(\gamma\right) GAP_{t-1}^{'}\left(\gamma\right)\right)^{-1} \left(\sum_{t=1}^{T} GAP_{t-1}\left(\gamma\right) dr_t\right) \tag{5}$$

[2] Threshold regression models are regime-switching models where each regime is determined by the value of a particular observable threshold variable. These models differ from the regime-switching models *à la* Hamilton (Hamilton, 1989) because in the latter each regime is governed by an unobserved

with residuals $\hat{e}_t(\gamma) = dr_t - GAP_{t-1}^{'}(\gamma)\theta(\gamma)$ and residual variance $\hat{\sigma}_T^2(\gamma) = T^{-1}\sum_{t=1}^{T}\hat{e}_t(\gamma)^2$. The LS estimate of γ is the value that minimizes the residual variance, i.e.

$$\hat{\gamma} = \arg\min_{\gamma\in\Gamma} \hat{\sigma}_T^2(\gamma) \tag{6}$$

where $\Gamma = \left[\underline{\gamma}, \overline{\gamma}\right]$. The minimization problem in (6) can be solved by direct search. The residual variance $\hat{\sigma}_T^2(\gamma)$ takes on at most T distinct values as γ varies. Thus, to obtain the LS estimates of (6) we can run OLS regression of (4) for $\gamma \in \Gamma$, where the elements of Γ are slightly less than T because we have to take a certain percentage $(\eta\%)$ of observations out to ensure that each regime has a minimum number of observations. Then the value of γ that minimizes the residual variance is the LS estimate of the threshold parameter. The LS estimates of the coefficients θ are then found as $\hat{\theta} = \hat{\theta}(\hat{\gamma})$. Similarly, the LS residuals are $\hat{e} = dr_t - GAP_t(\hat{\gamma})'\hat{\theta}$, with sample variance $\hat{\sigma}_T^2 = \hat{\sigma}_T^2(\hat{\gamma})$.

An important question is whether the model in (2) is statistically significant relative to the linear specification in (1). The relevant null hypothesis is $H_0 : \beta_{j1} = \beta_{j2}$, $j = 0,1$. As it is well known, this testing problem is non standard because there are some parameters that are not identified under the null (the so-called 'Davies' (1977, 1987) problem'). Based on the theories of Davies (1977, 1987) and Andrews and Ploberger (1994), Hansen (1996) shows that if the errors are *iid*, a test with near-optimal power against alternatives distant from the null hypothesis is the standard *F*-statistic

$$F_T = T\left(\frac{\tilde{\sigma}_T^2 - \hat{\sigma}_T^2}{\hat{\sigma}_T^2}\right) \tag{7}$$

where

$$\tilde{\sigma}_T^2 = T^{-1}\sum_{t=1}^{T}\left(dr_t - GAP_t'\tilde{\theta}\right)^2 \tag{8}$$

is the residual variance under the null hypothesis and $\tilde{\theta}$ is the OLS estimate under the null of no threshold, i.e.

$$\tilde{\theta} = \left(\sum_{t=1}^{T} GAP_t GAP_t'\right)^{-1}\left(\sum_{t=1}^{T} GAP_t dr_t\right) \tag{9}$$

state variable which is usually modeled as a first-order Markov chain. For details see Hamilton (1994) or Franses and Van Dijk (2000).

Since F_T is a monotonic function of $\hat{\sigma}_T^2$, it has been shown that $F_T = \sup_{\gamma \in \Gamma} F_T(\gamma)$ where $F_T(\gamma) = T\left(\tilde{\sigma}_T^2 - \hat{\sigma}_T^2(\gamma)\right) / \hat{\sigma}_T^2(\gamma)$ is the pointwise *F*-statistic against the alternative $H_1 : \beta_{j1} \neq \beta_{j2}$ when γ is known. Since γ is not identified, the asymptotic distribution of F_T is not χ^2. Hansen (1996) shows that the asymptotic distribution of F_T can be approximated by a bootstrap procedure. Letting u_t^*, $t = 1, \ldots, T$ be *iid* $N(0,1)$ and setting $dr_t^* = u_t^*$, we can regress dr_t^* on the observations GAP_t to obtain $\tilde{\sigma}_T^{*2}$ and on $GAP_t(\gamma)$ to obtain $\hat{\sigma}_T^{*2}$. Thus we can form the statistic $F_T^*(\gamma) = T\left(\tilde{\sigma}_T^{*2} - \hat{\sigma}_T^{*2}(\gamma)\right) / \hat{\sigma}_T^{*2}(\gamma)$ and $F_T^* = \sup_{\gamma \in \Gamma} F_T^*(\gamma)$. Hansen (1996) shows that the distribution of F_T^* converges weakly in probability to the null distribution of F_T under local alternatives for θ. Therefore, repeated bootstrap draws from F_T^* can be used to approximate the asymptotic null distribution of F_T. The bootstrap approximation to the asymptotic *p*-value of the test is constructed by counting the percentage of bootstrap samples for which F_T^* exceeds the observed F_T. If the errors are conditionally heteroskedastic, it is necessary to replace the *F*-statistic $F_T(\gamma)$ with a heteroskedasticity consistent Wald or Lagrange multiplier test. Setting $R = [I \quad -I]$, $M_T(\gamma) = \sum_{t=1}^{T} GAP_t(\gamma) GAP_t(\gamma)'$ and $V_T(\gamma) = \sum_{t=1}^{T} GAP_t(\gamma) GAP_t(\gamma)' \hat{e}_t^2$, then the pointwise Wald statistic is

$$W_T(\gamma) = \left(R\hat{\theta}(\gamma)\right)' \left[R\left(M_T(\gamma)^{-1} V_T(\gamma) M_T(\gamma)^{-1}\right) R'\right]^{-1} R\hat{\theta}(\gamma) \qquad (10)$$

and the appropriate test for the null is $W_T = \sup_{\gamma \in \Gamma} W_T(\gamma)$. To obtain the bootstrap *p*-value, it is sufficient to repeat the bootstrap procedure as before setting $dr_t^* = \hat{e}_t u_t^*$.

2.2 Panel data model with a single threshold variable and two or more regimes

As in Hansen (1999) we start assuming that the observed data are from a balanced panel $\{dr_{it}, x_{it}, q_{it}\}$ with $1 \leq i \leq N$ and $1 \leq t \leq T$. The scalar dr_{it} is the dependent variable, the *k* vector x_{it} contains the exogenous (or predetermined variables), while q_{it} is an *s* vector $(s \geq 1)$ containing the threshold variables. The subscript *i* indicates the individual bank, while *t* designates the time period (in our case the quarter).

The simplest version of the model is a static panel data model with 2 regimes

$$\begin{aligned} dr_{it} &= \mu_i + \alpha_1 \ln\left(TA_{it}\right) + \alpha_2 \ln\left(TA_{it}\right)^2 + \alpha_3 \ln\left(TA_{it}\right)^3 + \alpha_4 \mathrm{lgr}_{it} + \alpha_5 \mathrm{lgr}_{it}^2 + \alpha_6 \mathrm{lgr}_{it}^3 \\ &+ \alpha_7 \ln\left(TA_{it}\right) \cdot \mathrm{lgr}_{it} + \beta_{11} GAP_{t-1} I\left(dr_{it} \le \gamma_1\right) + \beta_{12} GAP_{t-1} I\left(dr_{it} > \gamma_1\right) + e_t \end{aligned} \tag{11}$$

where μ_i are individual fixed effects, $\ln\left(TA_{it}\right)$ and lgr_{it} are the log of total assets and the loan growth rate of bank *i* at time *t*, respectively, while $I\left(\cdot\right)$ is the indicator function. The logarithm of total assets is included to control for size, while loan growth rate controls for loan dynamics. The non-linear terms are included to reduce the possibility of spurious correlations due to omitted variable bias. Lack of data does not permit the estimation of models with a richer set of bank-specific variables.

In model (11) the observations are divided into 2 regimes depending on whether the default rate of bank *i* at time *t* is smaller or larger than the threshold γ_1. This threshold is endogenously determined by the model and separates good banks from bad banks. Each regime (good or bad bank) is characterized by a different regression slope β_{1j}, $j = 1, 2$ and to identify them it is required that both the regressors and the threshold variables are not time invariant. The errors e_{it} are assumed to be *iid* with zero mean and finite variance σ^2. The asymptotic analysis is performed with fixed *T* and $N \to \infty$.

Model (11) can be generalized in two ways. First, we can identify different business cycle regimes by using a measure of output gap as the threshold variable, i.e.

$$\begin{aligned} dr_{it} &= \mu_i + \alpha_1 \ln\left(TA_{it}\right) + \alpha_2 \ln\left(TA_{it}\right)^2 + \alpha_3 \ln\left(TA_{it}\right)^3 + \alpha_4 \mathrm{lgr}_{it} + \alpha_5 \mathrm{lgr}_{it}^2 + \alpha_6 \mathrm{lgr}_{it}^3 \\ &+ \alpha_7 \ln\left(TA_{it}\right) \cdot \mathrm{lgr}_{it} + \beta_{11} GAP_{t-1} I\left(GAP_t \le \gamma_1\right) + \beta_{12} GAP_{t-1} I\left(GAP_t > \gamma_1\right) + e_t \end{aligned} \tag{12}$$

In this way, the first regime is characterized by recessionary conditions, while in the second one we have booming conditions (the output gap is greater than a certain threshold, γ_1). Secondly, we can generalize model (11) by considering the possibility of more than 2 regimes over the same threshold variables. For example we can consider 3 regimes over the banking variable (that is good, medium and bad bank) with the following model

$$\begin{aligned} dr_{it} &= \mu_i + \alpha_1 \ln\left(TA_{it}\right) + \alpha_2 \ln\left(TA_{it}\right)^2 + \alpha_3 \ln\left(TA_{it}\right)^3 + \alpha_4 \mathrm{lgr}_{it} + \alpha_5 \mathrm{lgr}_{it}^2 + \alpha_6 \mathrm{lgr}_{it}^3 \\ &+ \alpha_7 \ln\left(TA_{it}\right) \cdot \mathrm{lgr}_{it} + \beta_{11} GAP_{t-1} I\left(dr_{it} \le \gamma_1\right) + \beta_{12} GAP_{t-1} I\left(\gamma_1 < dr_{it} \le \gamma_2\right) \\ &+ \beta_{13} GAP_{t-1} I\left(dr_{it} > \gamma_2\right) + e_t \end{aligned} \tag{13}$$

where $\gamma_1 < \gamma_2$.

We can further generalize the model by including a third threshold, so that we can characterize 4 banking regimes. The same strategy can be adopted for model (12) with different business cycle regimes.

A more compact way to represent models (11), (12) and their generalizations is

$$dr_{it} = \mu_i + \theta' x_{it}(\gamma) + e_{it} \tag{14}$$

where $\theta = (\alpha_1,...,\alpha_7,\beta_{11},\beta_{12},...)$ and

$$x_{it}(\gamma) = \left(\ln(TA_{it}), \ln(TA_{it})^2, \ln(TA_{it})^3, \lg r_{it}, \lg r_{it}^2, \lg r_{it}^3, \lg r_{it} \cdot \ln(TA_{it}), GAP_t \cdot I(dr_{it} \le \gamma_1), GAP_t \cdot I(dr_{it} > \gamma_1)\right)' \tag{15}$$

in case of model (11). To estimate this class of models we can employ a fixed effects transformation by removing the individual specific effects. We can take the averages over time of (14) getting

$$\overline{dr}_i = \mu_i + \theta' \overline{x}_i(\gamma) + \overline{e}_i \tag{16}$$

where $\overline{dr}_i = T^{-1}\sum_t dr_{it}$, $\overline{e}_i = T^{-1}\sum_t e_{it}$ and $\overline{x}_i(\gamma) = T^{-1}\sum_t x_{it}(\gamma)$. Taking the differences between (14) and (16) yields

$$dr_{it}^* = \theta' x_{it}^*(\gamma) + e_{it}^* \tag{17}$$

where $dr_{it}^* = dr_{it} - \overline{dr}_i$, $x_{it}^*(\gamma) = x_{it}(\gamma) - \overline{x}_i(\gamma)$ and $e_{it}^* = e_{it} - \overline{e}_i$. Stacking data and errors for each individual i with the first time period deleted and then stacking what results over individuals we get the vector of the dependent variable Y^*, of the independent variables $X^*(\gamma)$ and that of the errors e^*. With this notation, (14) is equivalent to

$$Y^* = X^*(\gamma)\theta + e^* \tag{18}$$

and for any given value of the threshold γ this model can be estimated by OLS, i.e.

$$\hat{\theta}(\gamma) = \left[X^*(\gamma) X^*(\gamma)'\right]^{-1} X^*(\gamma)' Y^* \tag{19}$$

with regression residuals $\hat{e}^*(\gamma) = Y^* - X^*(\gamma)\hat{\theta}(\gamma)$ and sum of squared errors (SSE)

$$S(\gamma) = \hat{e}^*(\gamma)' \hat{e}^*(\gamma) = Y^{*'}\left[I - X^*(\gamma)\left[X^*(\gamma)' X^*(\gamma)\right]^{-1} X^*(\gamma)'\right] Y^* \tag{20}$$

Chan (1993) and Hansen (1999) recommend to estimate γ by LS minimizing the concentrated SSE (20). Thus the LS estimator of γ becomes

$$\hat{\gamma} = \arg\min_{\gamma=[\underline{\gamma},\bar{\gamma}]} S(\gamma) \tag{21}$$

Since it is undesirable for a threshold $\hat{\gamma}$ to be selected when it sorts too few observations in one regime, we can exclude this by restricting the minimization in (21) to values of γ such that a minimal percentage of observations (e.g. 1 or 5%) lie in each regime. Once $\hat{\gamma}$ is obtained the slope estimate becomes $\hat{\theta} = \hat{\theta}(\hat{\gamma})$, the residual vector is $\hat{e}^* = \hat{e}^*(\hat{\gamma})$ and the residual variance is $\hat{\sigma}^2 = \left[n(T-1)\right]^{-1} \hat{e}^{*\prime}\hat{e}^* = \left[n(T-1)\right]^{-1} S(\hat{\gamma})$. Since the SSE $S(\gamma)$ depends on the threshold only through the indicator functions, the SSE is a step function with at most *NT* steps. The minimization in (21) can thus be reduced to a search over at most *NT* different values of the threshold variable. This is achieved by sorting the threshold eliminating the smallest and the largest δ% to ensure a minimum number of observations in each regime. However, since this procedure might be numerically intensive with long time spans and large panels, Hansen (1999) suggests using 393 quantiles, reducing the grid search over $\{1.00\%, 1.25\%, 1.50\%, \ldots, 98.75\%, 99.00\%\}$. For other details on the estimation of these models, see Hansen (1999) and Marcucci and Lotti (2006).

As in the aggregate case, it is important to test whether the models in (11) and (12) are statistically significant relative to their linear specifications in which there are no thresholds. The relevant null hypothesis of no threshold (or one regime) can be represented as $H_0 : \beta_{11} = \beta_{12}$. Again, under the null the thresholds γ are not identified leading to the 'Davies' problem'. This implies that classical tests have non-standard distributions. We can again adopt the bootstrap procedure suggested by Hansen (1996) to simulate the asymptotic distribution of the test under the null hypothesis of no thresholds. Under the null, the model can be compactly represented as

$$dr_{it} = \mu_i + \theta' x_{it} + e_{it} \tag{22}$$

or as in (17) under the fixed effects transformation. The regression parameter θ can be estimated by OLS, yielding the restricted estimate $\tilde{\theta}$, residuals $\tilde{e}_{it}^*$ and sum of squared errors $S_0 = \tilde{e}_{it}^{*\prime}\tilde{e}_{it}^*$. Thus, the likelihood ratio test of H_0 against the alternative of a threshold is based on

$$F_{10} = \frac{S_0 - S_1(\hat{\gamma}_1)}{\hat{\sigma}^2} \tag{23}$$

where $\hat{\sigma}^2 = \left[n(T-1)\right]^{-1} S_1(\hat{\gamma}_1)$.

The asymptotic distribution of this test is non-standard because of the Davies' problem. Following Hansen (1999), we can get asymptotically valid tests by using the bootstrap. Keeping the regressors and threshold variable fixed in repeated bootstrap samples, we adopt the following procedure. First we estimate the model grouping the regression residuals $\tilde{e}_{it}^{*}$ by individual, i.e. for each i we construct $\tilde{e}_{i}^{*} = \left(\tilde{e}_{i1}^{*}, \ldots, \tilde{e}_{iT}^{*}\right)$ treating the sample $\tilde{e}^{*} = \left\{\tilde{e}_{1}^{*}, \ldots, \tilde{e}_{N}^{*}\right\}$ as the empirical distribution for the bootstrap. Then we draw with replacement a sample of size N from the empirical distribution and use these errors $\tilde{e}^{*(b)}$, where the superscript indicates the *b*-th bootstrap replication, to create a bootstrap sample under the null, i.e. $\tilde{dr}_{it}^{*(b)} = \tilde{\theta}' x_{it}^{*} + \tilde{e}_{it}^{*(b)}$. Using the bootstrap sample $\tilde{dr}_{it}^{*(b)}$, estimate the model under the null, i.e. $\tilde{dr}_{it}^{*(b)} = \theta' x_{it}^{*} + u_{it}$ and under the alternative of a threshold, i.e. $\hat{dr}_{it}^{*(b)} = \theta' x_{it}^{*}\left(\gamma_1\right) + v_{it}$. Therefore, we compute the bootstrap value of the likelihood ratio statistic F_{10} as in (23), repeating this procedure a large number (say B) of times. We finally calculate the bootstrap *p*-value as the percentage of draws for which the simulated statistic exceeds the actual value.

2.3 Panel data model with two threshold variables and four regimes

As an extension of Hansen's (1999) model, we can use a 4-regime panel data model where the regimes are determined by two different threshold variables, as suggested by Marcucci and Lotti (2006). The simplest version of the model takes the form

$$\begin{aligned} dr_{it} = {} & \mu_i + \alpha_1 \ln\left(TA_{it}\right) + \alpha_2 \ln\left(TA_{it}\right)^2 + \alpha_3 \ln\left(TA_{it}\right)^3 + \alpha_4 \mathrm{lgr}_{it} + \alpha_5 \mathrm{lgr}_{it}^2 + \alpha_6 \mathrm{lgr}_{it}^3 \\ & + \alpha_7 \ln\left(TA_{it}\right) \cdot \mathrm{lgr}_{it} + \beta_{11} GAP_{t-1} I\left(dr_{it} \le \gamma_1\right) I\left(GAP_{t-1} \le \gamma_2\right) \\ & + \beta_{12} GAP_{t-1} I\left(dr_{it} \le \gamma_1\right) I\left(GAP_{t-1} > \gamma_2\right) + \beta_{13} GAP_{t-1} I\left(dr_{it} > \gamma_1\right) I\left(GAP_{t-1} \le \gamma_2\right) \\ & + \beta_{14} GAP_{t-1} I\left(dr_{it} > \gamma_1\right) I\left(GAP_{t-1} > \gamma_2\right) + e_t \end{aligned} \tag{24}$$

where μ_i are individual fixed effects, $\ln\left(TA_{it}\right)$ and lgr_{it} are the log of total assets and the loan growth rate of bank *i* at time *t*, respectively, while $I\left(\cdot\right)$ is the indicator function.

In model (24) the observations are divided into 4 regimes depending on both the default rate of each bank and the output gap. With this model we can discriminate good and bad banks in both booming and recessionary conditions. In this way we can look at their different behavior over different phases of the business cycle. As before, each regime is characterized by a different slope ($\beta_{1j}, j = 1,...,4$) and to identify them it is required that both the regressors and the threshold variables are not time invariant. The errors are assumed to be *iid* with zero mean and finite variance while the asymptotic analysis is again performed with fixed T and $N \to \infty$.

To estimate this model we can employ the fixed effects transformation as with the 2-regime panel data model discussed before. We can then apply conditional LS minimizing the concentrated SSE as in (21). As before, it is fundamental to test whether the model (24) is statistically significant relative to the simplest models with only one threshold. The null hypothesis in this case is that of one threshold. Thus, we again have the problem of some parameters not identified under the null, implying a non-standard testing problem. We can therefore adopt the bootstrap procedure suggested by Marcucci and Lotti (2006) which is similar to that one discussed before in the case of only one threshold. An approximate likelihood ratio test of one threshold against two thresholds can be based on the statistic

$$F_{21} = \frac{S_1(\tilde{\gamma}_1) - S_2(\hat{\gamma}_1, \hat{\gamma}_2)}{\hat{\sigma}^2} \tag{25}$$

where $\hat{\sigma}^2 = \left[n(T-1)\right]^{-1} S_2(\hat{\gamma}_1, \hat{\gamma}_2)$. The null hypothesis of one threshold is rejected for large values of F_{21}. We can use a similar bootstrap procedure as in the 2-regime panel model to obtain the approximate asymptotic distribution of the test. To generate the bootstrap samples we hold both the regressors and thresholds fixed in repeated bootstrap samples. Then we follow the same steps discussed before to obtain the bootstrap *p*-value of the test.

3. Data on the Italian banks' portfolios

In this section, we apply the methodology described above to a large panel of Italian intermediaries. The starting point for our analysis is the choice of an adequate measure of the default rate. In Italy, banks must value loans in their portfolios at their estimated realizable value. In particular, the exposures to insolvent borrowers, regardless of any collateral received, are classified as bad loans. Since Italian banks tend to classify their exposures correctly and with appropriate timing (Moody's, 2003), bad loans are a good indicator of the riskiness of banks' debtors. Therefore, we compute our riskiness indicator as the ratio of the flow of loans classified as bad debts in the reference period to the performing loans outstanding at the end of the previous period. The ratio can be interpreted as the default rate of banks' borrowers. With respect to other riskiness indicators, which are based on stock measures, such as the non-performing loan ratio, the default rate is a more precise and timely proxy for banks' portfolio riskiness. Furthermore, in order to improve the reliability and timeliness of the indicator, we use the "adjusted" bad debts as signaled by the Italian Central Credit Register.[3]

Since default rates from the Central Credit Register are available at bank level, we can exploit the cross-sectional dimension and work at different levels of aggregation – from the banking system to each single bank's portfolio – in order to check for the robustness of our results. Regarding the proxy for economic conditions, we focus on output gap, which is the difference between the actual and the potential domestic product. We compute the output gap from the Hodrick-Prescott filtered series. For robustness, as alternative proxies, we employ the deviations of GDP series from its trend and the GDP growth rate.

Accounting ratios for the individual institutions are built up using the statistics that intermediaries are required to report to the Bank of Italy; the macroeconomic variables are drawn from the OECD statistics. The resulting dataset includes 220 banks and span from 1990Q1 to 2005Q2. All data are quarterly.

Table 1 reports the summary statistics of the data. We report both the micro bank data and the macro data on different measures of output gap.

[Table 1 about here]

4. Empirical results

4.1 Model with a single threshold variable and two regimes: aggregate data

4.1.1 The default rate as the threshold variable

Table 2 reports the results for the 2-regime threshold model estimated using aggregate data. Given the scarcity of data, in our preferred parsimonious model, the default rate *dr* depends on 1-quarter lag of GAP. The results for model 1 confirm the

[3] Adjusted bad loans are those outstanding when a borrower is reported to the Central Credit Register: a) as a bad debt by the only bank that disbursed credit; b) as a bad debt by one bank and as having an overshoot by the only other bank exposed; c) as a bad debt by one bank and the amount of the bad debt is at least 70% of its exposure towards the banking system or as having overshoots equal to or more than 10% of its total loans outstanding; d) as a bad debt by at least two banks for amounts equal to or more than 10% of its total loans outstanding.

well-known negative relationship between default rates and business cycle, with no regime changes.

Our second set of results shows the impact of the business cycle depending on the starting level of banks' riskiness, proxied by either contemporaneous (model 2) or lagged *dr* (model 3). We note that there is a switch of regime when bank borrowers' default rates are above an endogenous threshold. In particular, the statistical significance and magnitude of β_{12} suggest that for banks with lower asset quality, economic conditions have a statistically significant impact on their loan portfolios. By contrast, for good banks (i.e., when *dr* is below the threshold) the impact of the business cycle on default rates is almost nil and not significant. The LR test for the null of no regime switch is significant at any conventional level, suggesting that our approach is appropriate.

To check the robustness of our results we estimated the same models with different proxies of the business cycle and greater lags for the threshold variable. Our results are not affected by the choice of a greater lag for the threshold variable and are robust to the use of different proxies for the phase of the business cycle (deviation with respect to a linear trend, GDP growth).

As we mentioned above, an advantage of our methodology is that it allows us to obtain an endogenous estimate of the threshold. Looking at Table 2, we observe that the value of the threshold is very similar across models and specifications, ranging between .54 and .58%. Taken at its face value, this means that, when the aggregate default rate is above these figures, the banking system tends to be more prone to macroeconomic turbulences. In a macroprudential perspective, this advises to reinforce monitoring activities in these periods.

[Table 2 about here]

4.1.2 The output gap as the threshold variable

With model 4, we try to assess whether the impact of the business cycle on banks is also subject to a second kind of regime switch, which depends on the phase of the business cycle itself. Our results suggest that in recessions the impact of the business cycle on credit risk is statistically significant and more pronounced than in expansionary phases when the impact is less intense and not significant. Again, these results are generally robust to the use of different business cycle indicators.

4.2 Model with a single threshold variable and two regimes: institutional breakdowns

In order to check the robustness of our results, we re-estimate the models at different levels of aggregation. In particular, we classify banks into 3 different institutional categories: limited companies, cooperative banks and mutual banks. Table 3 shows that the econometric results are substantially unchanged with respect to those presented above for the aggregate case.

[Table 3 about here]

The impact of macroeconomic conditions on credit risk is negative and statistically significant when the quality of banks' portfolios is lower (i.e. for bad banks), while it is not for banks with better portfolios - models 2, 6, 7, 10, and 11. β_{11} is significant at the 10 per cent level for limited banks; however, its magnitude is considerably lower than that of β_{12}. The LR tests reject the null of no regime switch at any conventional level. The estimated thresholds are very close across the 3 categories and similar to those obtained with the aggregate models. When GAP is the threshold variable, the results for models 4, 8 and 12 confirm that the relationship between business cycle and credit risk is significantly negative only during recessions.

4.3 Panel data model with a single threshold variable and multiple regimes

The results presented so far are very supportive of the hypothesis that credit risk is cyclical. However, they also seem to suggest that the issue of cyclicality, as described by the empirical literature, has been somehow misinterpreted. Indeed, according to our evidence, the negative relationship between banks' portfolios riskiness and the business cycle holds only in either unfavorable economic conditions or when average credit quality is already unsatisfactory. By contrast, it does not seem to be statistically significant in good times (i.e., when either economic conditions improve or loan riskiness is low).

However, the small sample sizes advice us to interpret these preliminary results with caution and to improve the analysis, exploiting the cross-sectional dimension of our rich dataset. Using the information on borrowers' default rates available on a bank-by-bank basis and panel data techniques, we are able to identify whether the impact of the business cycle on "bad" and "good" banks is different. At first glance, we expect bad banks to have more cyclical portfolios than good ones.

The sample size also allows us to introduce a more articulated econometric representation, with multiple thresholds over the same variable. We start with a simple model 1, which includes only a single threshold, and move on to models 2 and 3, which include 2 and 3 thresholds respectively. The use of multiple thresholds makes it possible to identify up to 4 regimes (or risk categories), in which we can classify banks (for simplicity we can refer to them as very good, good, bad, and very bad bank regime).

4.3.1 The default rate as the threshold variable

The results of the estimated panel data models with 2 or more regimes over the same threshold variable are provided in Table 4. When a single threshold is introduced, we find that both good and bad banks are significantly affected by the business cycle, but the magnitude of the coefficient on GAP is larger for the bank with lower asset quality. In particular, the increase of *dr* as the result of 1 percentage point decrease of GAP is almost 7 times higher for bad banks.

[Table 4 about here]

Model 2 and 3 provide further strong support to the application of regime-switching models for analyzing credit risk. We find that the magnitude of the coefficients on GAP monotonically increases as we move from low-risk to high-risk regimes. In particular, looking at the results for model 3, we note that β is equal to -0.02 and -0.09 for very good and good banks respectively, while it goes to -0.21 and -0.51 for bad and very bad ones. These are quite substantial differences.

This evidence is confirmed when lagged *dr* is employed as a threshold - model 5 and 6. As a robustness check, we also estimate the models for different institutional categories of banks. The results for these specifications are consistent with those presented above and are therefore not reported for the sake of brevity.

For models 1 and 2 the LR test is significant at any conventional level, while it is significant at the 5 per cent level for model 3. This indicates that the model with 4 regimes (3 thresholds) is adequate. However, when we use lagged *dr*, the LR test rejects the third threshold and we can only estimate a 3-regime model.

The endogenously determined thresholds can be used to assess the evolution of banks' riskiness over time. Table 5 shows the percentage of banks in each regime (or risk-category) and in each quarter for the 4-regime model (3). Figure 1 depicts the percentage and the number of banks in each regime for each quarter for models from (1) through (4).

We note that banks tended to migrate towards riskier regimes between 1992-end and 1993-end, in 1995-Q4 and 1996-Q2; these periods identify recessionary phases in

Italy. However, no significant migration is apparent in 2001 and 2002, notwithstanding the very negative economic conditions. A possible explanation is that banks have improved borrower selection criteria in the last years; furthermore, the very low level of interest rates and the limited level of indebtedness may have helped firms and households to honor their obligations even in unfavorable times.

[Table 5 and Figure 1 about here]

We also note that the share of banks under the fourth regime is almost unaffected by the economic environment. This extreme risk-category (*dr*>1.5) is probably more affected by idiosyncratic factors than systemic ones.

4.3.2 The output gap as the threshold variable

When GAP is used as the threshold variable (model 4), we find further evidence that the relationship between *dr* and GAP is stronger in recessionary conditions than in booms. This is consistent with our previous results. Unfortunately, the LR test fails to reject the null hypothesis of one regime (i.e., no threshold).

Based on the estimated thresholds for GAP, we can identify periods in which economic conditions tend to affect banks' portfolios to a larger extent than in normal times. Looking at the plot in Fig. 1, panel D, we clearly identify 4 "high-impact" periods: between 1993-Q1 and 1994-Q2, in 1997-Q1/Q2, 1999-Q1 and 2005-Q2. We note that these periods do not necessarily overlap with recessionary business cycle phases. For the sake of comparison, we rely on the studies by Altissimo *et al.* (2000) and Bruno and Otranto (2004), who provide a very consistent description of the evolution of the business cycle in Italy. During the period 1987-2002, they identify three main recessions: the first from March 1992 to July 1993, the second from November 1995 to November 1996, and the third at the end of 2001. According to our evidence, only 1993 recession is also a "high-impact" period.

4.4 Panel data model with two threshold variables and 4 regimes

The final set of results is obtained estimating a model with two different threshold variables: *dr* and GAP. In this way, we try to depict a comprehensive picture of the evolution of credit risk across banks and through the business cycle. It is not straightforward to guess the impact of different business cycle regimes on banks with different portfolio regimes; however, our a priori belief is that good banks are less affected by economic conditions than bad ones. In addition, for the latter intermediaries, the impact should be stronger in unfavorable times.

Table 6 shows the results for the 4-regime panel data models with two different threshold variables. We focus only on model 1 and 2.

In model 1, we note that the coefficient on GAP is negative, as expected, only for good banks in expansionary periods and for bad banks in recessions. The magnitude of the coefficient is higher for bad banks, confirming the results we provided in the previous paragraphs. However, it is puzzling that the coefficient turns out to be positive and statistically significant for good banks in recession and bad banks in expansion.

[Table 6 about here]

Model 2 provides more clear-cut results, even though some counter-intuitive evidence remains. In particular, β_{14} (bad banks in expansion) is positive. Further research is needed for these models.

5. Concluding remarks

In this paper we try to make a step forward in explaining the macroeconomic determinants of credit risk and its evolution over the business cycle. With respect to most of the existing studies, which neglect asymmetric effects, we analyze whether the relationship between business cycle and credit risk is subject to regime switches and endogenously determine the thresholds at which regimes switch from one to another.

Using Threshold Regression approach and exploiting a unique dataset on Italian bank borrowers' default rates, we find that the impact of the business cycle is more pronounced when starting credit risk levels are higher and during unfavorable economic times. Moreover, our methodology allows us to identify the risk threshold(s) over/below which the impact is different, providing a powerful tool for financial stability monitoring.

As an example, in the two-regime model, we find that both good and bad banks are significantly affected by the business cycle, but the impact is stronger for the latter. In particular, the increase of the default rate as the result of 1 percentage point decrease of the output gap is almost 7 times higher for bad banks. This evidence is robust to the use of different proxies for the overall economic conditions and holds at various levels of aggregation. By contrast, the evidence arising from our panel data model with two threshold variables is less definite and leaves room for future research.

Overall, our results may provide some guidance to banks and supervisors in the choice of adequate capital buffers in the various phases of the business cycle and for different intermediaries, particularly under risk sensitive capital requirements.

Furthermore, the methodology we propose may be easily employed for stress testing credit risk in banks' books.

References

Altissimo F., Marchetti D. J. and Oneto G. P. (2001), "The Italian Business Cycle: Coincident and Leading Indicators and Some Stylized Facts", *Banca d'Italia - Temi di Discussione,* 377, October.

Bangia A., Diebold F. X., Kronimus A., Schagen C. and Schuermann T. (2002), "Ratings migration and the business cycle, with application to credit risk portfolio stress testing", Journal of Banking and Finance, 26, 445-474.

Bikker J. A. and Hu H. (2002), "Cyclical Patterns in Profits, Provisioning and Lending of Banks", DNB Staff Reports, 86.

Bruno G. and Otranto E. (2004), "Dating the Italian Business Cycle: A Comparison of Procedures", *ISAE Working Paper*, 41, January.

Chan K. S. and H. Tong (1986), On estimating thresholds in autoregressive models. Journal of Time Series Analysis, 7, 179-194.

Chan K.S. (1993), Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. 21, 520-33.

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

Davies R.B. (1987), Hypothesis testing when a nuisance parameter is present only under the alternative. Biometrika, 74, 33-43.

Franses P.H. and D. Van Dijk (2000), "Nonlinear Time Series Models in Empirical Finance", Cambridge: Cambridge University Press.

Gambera M. (2000), "Simple Forecasts of Bank Loan Quality in the Business Cycle", Federal Reserve Bank of Chicago, Supervision and Regulation Department, Emerging Issues Series, S&R-2000-3.

Gasha J. G. and Morales R. A. (2004), "Identifying Threshold Effects in Credit Risk Stress Testing", IMF Working Paper, 150.

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

Hamilton J. D. (1989), A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, Econometrica, 57, 357-384.

Hamilton J. D. (1994), Time Series Analysis, Princeton: Princeton University Press.

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

Hansen B.E. (1999), Threshold effects in non-dynamic panels: Estimation, testing, and inference. Journal of Econometrics, 93, 345-368.

Hoggarth, G., Sorensen S., Zicchino, L., 2005, "Stress tests of UK banks using a VAR approach", Bank of England, working paper n. 282.

Jones M. T. (2005), "Estimating Markov Transition Matrices Using Proportions Data: An Application to Credit Risk", IMF Working Paper, 219.

Laeven L. and Majoni G. (2003), "Loan Loss Provisioning and Economic Slowdowns: Too Much, Too Late?", Journal of Financial Intermediation, 12, 2, 178-197.

Laviola S., Marcucci J. and Quagliariello M. (in press), "Stress testing credit risk: Experience from the Italian FSAP", BNL Quarterly Review.

Lucas A. and Klaassen P. (2005), "Discrete versus continuous state switching models for portfolio credit risk", Journal of Banking and Finance, 30(1), 23-35.

Marcucci J. and F. Lotti (2006), Firms' investments over the business cycle. Bank of Italy, *mimeo*.

Marcucci J. and Quagliariello M. (in press), "Is Bank Portfolio Riskiness Procyclical? Evidence from Italy using a Vector Autoregression", Journal of International Financial Markets, Institutions and Money.

Marotta G., Pederzoli C. and Torricelli C. (2005), "Forward-looking estimation of default probabilities with Italian data", mimeo.

Meyer A. P. and Yeager T. J. (2001), "Are small Rural Banks Vulnerable to Local Economic Downturns?", Federal Reserve Bank of St. Louis Review, 25-38.

Nickell P., Perraudin W. and Varotto S. (2000), "Stability of rating transitions", Journal of Banking and Finance, 24, 203-227.

Pederzoli C. and Torricelli C. (2005), "Capital requirements and business cycle regimes: Forward-looking modelling of default probabilities", Journal of Banking and Finance, 29, 3121-3140.

Pesola, J., 2001, "The role of macroeconomic shocks in banking crises", Bank of Finland, discussion paper n. 6.

Quagliariello M. (in press), "Does Macroeconomy Affect Bank Stability? A Review of the Empirical Evidence", Journal of Banking Regulation.

Salas, V., Saurina, J., 2002, "Credit risk in two institutional settings: Spanish commercial and saving banks", Journal of Financial Services Research, 22(3), December, 203-224.

Teräsvirta T., D. Tjostheim and C. W. J. Granger (1994), Aspects of modeling nonlinear time series. In R. F. Engle and D. L. McFadden (eds.), Handbook of Econometrics, vol. IV. Amsterdam: Elsevier Science, 2917-2957.

Tong H. (1983), Threshold Models in Non-linear Time Series Analysis: Lecture Notes in Statistics 21. Berlin: Springer-Verlag

Tong H. (1990), Non-Linear Time Series: A Dynamical System Approach. New York: Oxford University Press.

Valckx N. (2003), "What Determines Loan Loss provisioning in the EU?", ECB, mimeo.

Table 1: Summary statistics

	Description	Min	25% Percentile	Median	75% Percentile	Max	Mean	Std. Dev.	Skewness	Kurtosis
Micro variables										
dr	*Default rate*	0.000	0.000	0.263	0.459	1.495	0.304	0.283	1.038	4.047
ln(TA)	*Total Assets in logs*	1.792	5.043	6.009	7.333	12.308	6.311	1.776	0.714	3.345
lgr	*Loan growth rate*	-1.033	0.010	0.028	0.047	1.023	0.021	0.078	-5.464	46.698
Macro variables										
GDPG	*GDP growth rate*	-1.840	0.402	1.169	2.528	3.818	1.363	1.338	-0.115	2.368
GAP	*Output gap from HP-filtered GDP*	-2.126	-0.513	0.116	0.639	1.779	0.017	0.829	-0.313	2.588
GAPT	*Output gap from linear trend*	-3.001	-0.734	0.168	0.920	2.539	-0.037	1.298	-0.222	2.316

Note: Quarterly data from 1990:Q1 to 2005:Q2. Total of 13,420 bank-quarters. The default rate is computed as the flow of new 'adjusted' bad debts in each quarter over the outstanding non-performing loans in the previous quarter.

Table 2: Estimates for 2-regime threshold model at the aggregate level

Notes: This table reports the conditional LS estimates for the following threshold model with 2 regimes. ***, ** and * indicate significance at 1, 5 and 10% respectively.

Model (1): $dr_t = \beta_{01} + \beta_{11} GAP_{t-1} + e_t$

Model (2): $dr_t = \left(\beta_{01} + \beta_{11} GAP_{t-1}\right) I\left\{dr_t \le \gamma\right\} + \left(\beta_{02} + \beta_{12} GAP_{t-1}\right) I\left\{dr_t > \gamma\right\} + e_t$

Model (3): $dr_t = \left(\beta_{01} + \beta_{11} GAP_{t-1}\right) I\left\{dr_{t-1} \le \gamma\right\} + \left(\beta_{02} + \beta_{12} GAP_{t-1}\right) I\left\{dr_{t-1} > \gamma\right\} + e_t$

Model (4): $dr_t = \left(\beta_{01} + \beta_{11} GAP_{t-1}\right) I\left\{GAP_{t-1} \le \gamma\right\} + \left(\beta_{02} + \beta_{12} GAP_{t-1}\right) I\left\{GAP_{t-1} > \gamma\right\} + e_t$

Model	(1)		(2)		(3)		(4)	
β_{01}	0.5536	***	0.4022	***	0.4074	***	0.4449	***
	(0.0226)		(0.0127)		(0.0151)		(0.0259)	
β_{02}	-		0.7118	***	0.7020	***	0.5559	***
	-		(0.0157)		(0.0166)		(0.0653)	
β_{11}	-0.0752	**	-0.0210		1.50E-05		-0.2183	***
	(0.0303)		(0.0161)		(0.0277)		(0.0356)	
β_{12}	-		-0.0676	***	-0.0747	***	-0.0084	
	-		(0.0191)		(0.0203)		(0.0776)	
γ	-		0.5387		0.5793		0.3005	
R^2	0.1171		0.8214		0.7638		0.3200	
N_1	58		30		30		36	
N_2	-		28		28		22	
LR Test	-		44.59	***	40.00	***	13.02	***
p-value	-		0.000		0.000		0.001	
N. bootstrap	-		1000		1000		1000	
trimming %	-		0.15		0.15		0.15	

Table 3: Estimates for 2-regime threshold model at the aggregate level for bank categories

Notes: This table reports the conditional LS estimates for the following threshold model with 2 regimes. ***, ** and * indicate significance at 1, 5 and 10% respectively.

Models (1), (5) and (9): $dr_t = \beta_{01} + \beta_{11} GAP_{t-1} + e_t$, for limited, cooperative and mutual banks respectively.

Models (2), (6) and (10): $dr_t = (\beta_{01} + \beta_{11} GAP_{t-1}) I\{dr_t \leq \gamma\} + (\beta_{02} + \beta_{12} GAP_{t-1}) I\{dr_t > \gamma\} + e_t$ for limited, cooperative and mutual banks respectively.

Models (3), (7) and (11): $dr_t = (\beta_{01} + \beta_{11} GAP_{t-1}) I\{dr_{t-1} \leq \gamma\} + (\beta_{02} + \beta_{12} GAP_{t-1}) I\{dr_{t-1} > \gamma\} + e_t$, for limited, cooperative and mutual banks respectively.

Models (4), (8) and (12): $dr_t = (\beta_{01} + \beta_{11} GAP_{t-1}) I\{GAP_{t-1} \leq \gamma\} + (\beta_{02} + \beta_{12} GAP_{t-1}) I\{GAP_{t-1} > \gamma\} + e_t$, for limited, cooperative and mutual banks respectively.

	Limited Banks				Cooperative Banks				Mutual Banks			
Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
β_{01}	0.5765 ***	0.4206 ***	0.4257 ***	0.4685 ***	0.4894 ***	0.3603 ***	0.3589 ***	0.3998 ***	0.3661 ***	0.2930 ***	0.2756 ***	0.3023 ***
	(0.0228)	(0.0130)	(0.0154)	(0.0265)	(0.0198)	(0.0090)	(0.0085)	(0.0224)	(0.0142)	(0.0083)	(0.0068)	(0.0144)
β_{02}	-	0.7284 ***	0.7193 ***	0.5825 ***	-	0.6217 ***	0.6215 ***	0.5071 ***	-	0.4921 ***	0.4586 ***	0.3304 ***
	-	(0.0170)	(0.0176)	(0.0674)	-	(0.0168)	(0.0165)	(0.0540)	-	(0.0104)	(0.0136)	(0.0420)
β_{11}	-0.0780 **	-0.0223	-0.0018 *	-0.2198 ***	-0.0725 ***	-0.0084	-0.0042	-0.1888 ***	-0.0482 **	-0.0194	-0.0053	-0.1356 ***
	(0.0304)	(0.0174)	(0.0283)	(0.0371)	(0.0276)	(0.0121)	(0.0136)	(0.0363)	(0.0205)	(0.0124)	(0.0118)	(0.0225)
β_{12}	-	-0.0723 ***	-0.0791 ***	-0.0157	-	-0.0742 ***	-0.0726 **	-0.0354	-	-0.0517 ***	-0.0460 ***	0.0329
	-	(0.0205)	(0.0217)	(0.0799)	-	(0.0219)	(0.0222)	(0.0658)	-	(0.0098)	(0.0160)	(0.0519)
γ	-	0.5369	0.5387	0.3005	-	0.5387	0.5793	0.3005	-	0.6191	0.5793	0.3005
R^2	0.1227	0.8025	0.7501	0.3168	0.1375	0.7846	0.7880	0.3079	0.1216	0.8210	0.7473	0.3199
N_1	58	29	29	36	58	30	30	36	58	37	30	36
N_2	-	29	29	22	-	28	28	22	-	21	28	22
LR Test	-	43.76 ***	39.12 ***	12.83 ***	-	42.83 ***	42.21 ***	11.83 ***	-	41.92 ***	38.08 ***	12.22 ***
p-value	-	0.000	0.000	0.001	-	0.000	0.000	0.008	-	0.000	0.000	0.008
N. bootstrap	-	1000	1000	1000	-	1000	1000	1000	-	1000	1000	1000
trimming %	-	0.15	0.15	0.15	-	0.15	0.15	0.15	-	0.15	0.15	0.15

Table 4: Estimates for threshold regression panel data models with 2 or more regimes over the same threshold variable

Notes: This table reports the conditional LS estimates for the following threshold panel data model with 2 or more regimes (same threshold variable). ***, ** and * indicate significance at 1, 5 and 10% respectively.

Model (1): $dr_{it} = \alpha_1 \ln(TA_{it}) + \alpha_2 \ln(TA_{it})^2 + \alpha_3 \text{lgr}_{it} + \alpha_4 \text{lgr}_{it}^2 + \beta_{11} GAP_{t-1} I(dr_{it} \le \gamma_1) + \beta_{12} GAP_{t-1} I(dr_{it} > \gamma_1) + e_t$

Model (2): $$dr_{it} = \alpha_1 \ln(TA_{it}) + \alpha_2 \ln(TA_{it})^2 + \alpha_3 \text{lgr}_{it} + \alpha_4 \text{lgr}_{it}^2 + \beta_{11} GAP_{t-1} I(dr_{it} \le \gamma_1) + \beta_{12} GAP_{t-1} I(\gamma_1 < dr_{it} \le \gamma_2) + \beta_{13} GAP_{t-1} I(dr_{it} > \gamma_2) + e_t$$

Model (3): $$dr_{it} = \alpha_1 \ln(TA_{it}) + \alpha_2 \ln(TA_{it})^2 + \alpha_3 \text{lgr}_{it} + \alpha_4 \text{lgr}_{it}^2 + \beta_{11} GAP_{t-1} I(dr_{it} \le \gamma_1) + \beta_{12} GAP_{t-1} I(\gamma_1 < dr_{it} \le \gamma_2) + \beta_{13} GAP_{t-1} I(\gamma_2 < dr_{it} \le \gamma_3) + \beta_{14} GAP_{t-1} I(dr_{it} > \gamma_3) + e_t$$

Model (4): $dr_{it} = \alpha_1 \ln(TA_{it}) + \alpha_2 \ln(TA_{it})^2 + \alpha_3 \text{lgr}_{it} + \alpha_4 \text{lgr}_{it}^2 + \beta_{11} GAP_{t-1} I(GAP_{t-1} \le \gamma_1) + \beta_{12} GAP_{t-1} I(GAP_{t-1} > \gamma_1) + e_t$

Model (5) and (6): same as (1) and (2) respectively with threshold dr_{it} replaced by dr_{it-1}.

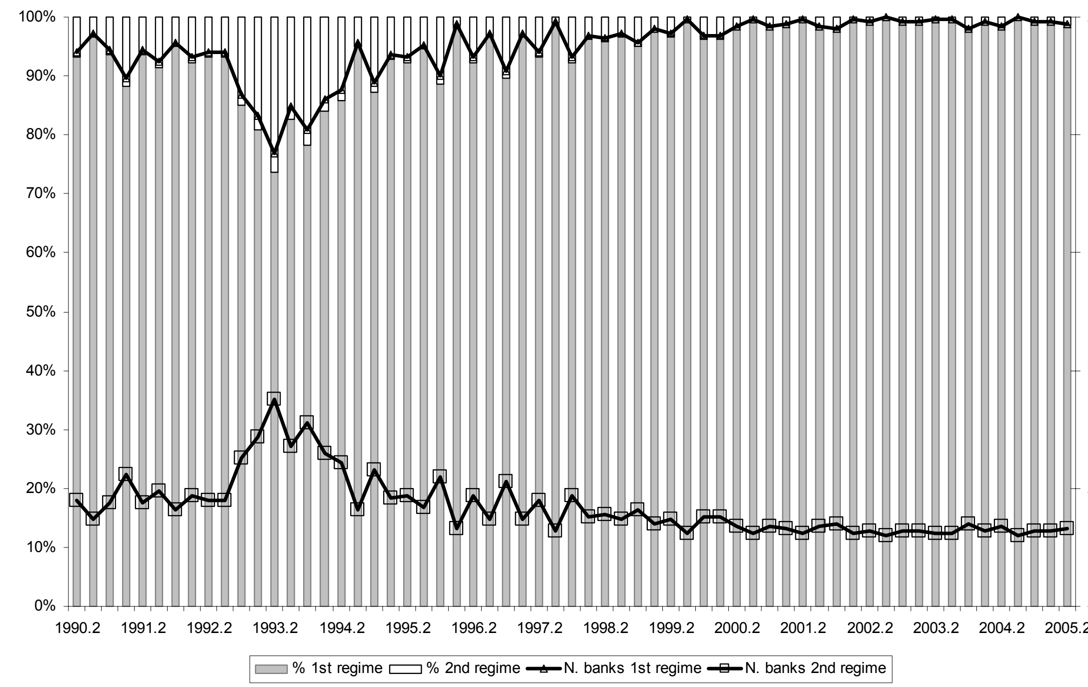
Model	(1)		(2)		(3)		(4)		(5)		(6)	
α_1	0.1564	***	0.1551	***	0.1552	***	0.1761	***	0.1587	***	0.1609	***
	(0.0182)		(0.0182)		(0.0181)		(0.0181)		(0.0181)		(0.0181)	
α_2	-0.0213	***	-0.0212	***	-0.0212	***	-0.0223	***	-0.0217	***	-0.0218	***
	(0.0013)		(0.0013)		(0.0013)		(0.0013)		(0.0013)		(0.0013)	
α_3	-0.3431	***	-0.3414	***	-0.3408	***	-0.3529	***	-0.3538	***	-0.3514	***
	(0.0876)		(0.0877)		(0.0875)		(0.0898)		(0.0891)		(0.0885)	
α_4	-0.8052	***	-0.8021	***	-0.8009	***	-0.8110	***	-0.8067	***	-0.8010	***
	(0.1889)		(0.1888)		(0.1885)		(0.1936)		(0.1920)		(0.1907)	
β_{11}	-0.0244	***	-0.0221	***	-0.0221	***	-0.0666	***	-0.0284	***	-0.0244	***
	(0.0021)		(0.0020)		(0.0020)		(0.0046)		(0.0027)		(0.0029)	
β_{12}	-0.1656	***	-0.0949	***	-0.0947	***	-0.0121	***	-0.0981	***	-0.0478	***
	(0.0184)		(0.0131)		(0.0131)		(0.0036)		(0.0087)		(0.0070)	
β_{13}	-		-0.2340	***	-0.2148	***	-		-		-0.0980	***
			(0.0329)		(0.0340)						(0.0087)	
β_{14}	-		-		-0.5056	***	-		-		-	
					(0.0855)							
γ_1	0.827		0.725		0.725		-0.996		0.742		0.514	
γ_2	-		1.025		1.025		-		-		0.742	
γ_3	-		-		1.456		-		-		-	
N. of banks	220		220		220		220		220		220	
N. of quarters	61		61		61		61		61		61	
N. of quantiles	393		393		393		393		393		393	
Trimming %	0.01		0.01		0.05		0.01		0.01		0.01	
LR Test	531.59	***	54.92	***	36.18	**	329.50		336.63	*	12.08	
p-value	0.000		0.000		0.037		0.397		0.053		0.263	
N. bootstrap	300		300		300		300		300		300	

Table 5: Percentage and number of banks in each regime by quarter. Model (3): 4 regimes with panel data: threshold dr_{it}

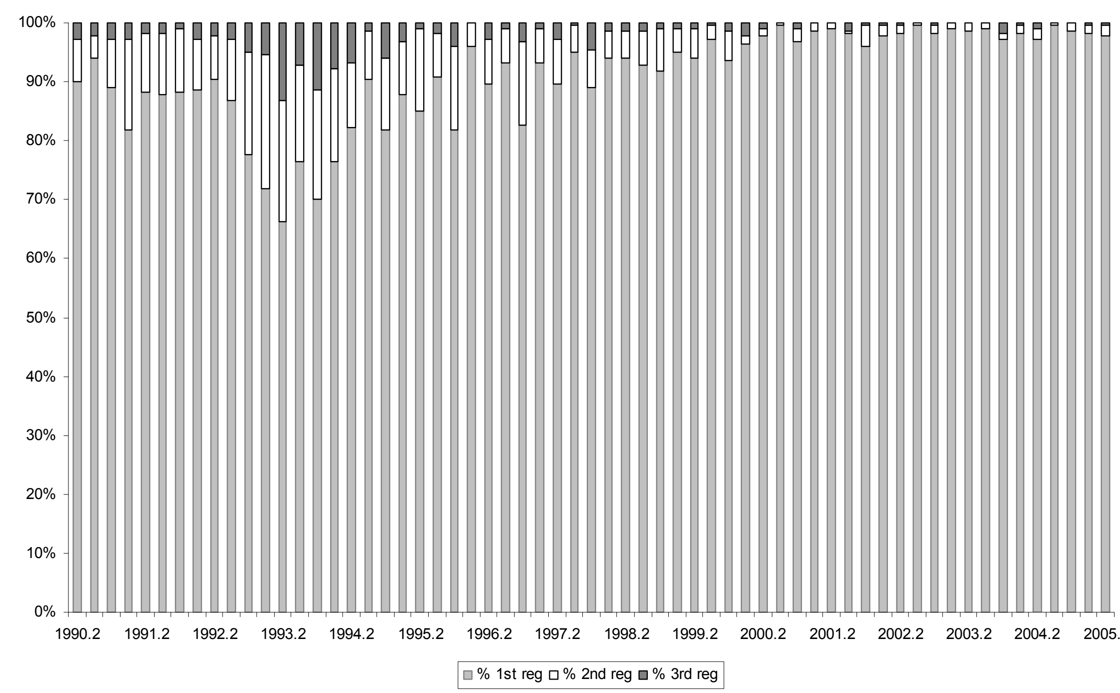
Quarter	% 1st reg.	% 2nd reg.	% 3rd reg.	% 4th reg.	N. banks 1st reg.	N. banks 2nd reg.	N. banks 3rd reg.	N. banks 4th reg.	Quarter	% 1st reg.	% 2nd reg.	% 3rd reg.	% 4th reg.	N. banks 1st reg.	N. banks 2nd reg.	N. banks 3rd reg.	N. banks 4th reg.
1990.2	90.00	7.27	2.73	0.00	198	16	6	0	1997.4	89.09	6.36	4.55	0.00	196	14	10	0
1990.3	94.09	3.64	2.27	0.00	207	8	5	0	1998.1	94.09	4.55	1.36	0.00	207	10	3	0
1990.4	89.09	8.18	2.73	0.00	196	18	6	0	1998.2	94.09	4.55	0.91	0.45	207	10	2	1
1991.1	81.82	15.45	2.27	0.45	180	34	5	1	1998.3	92.73	5.91	1.36	0.00	204	13	3	0
1991.2	88.18	10.00	1.82	0.00	194	22	4	0	1998.4	91.82	7.27	0.91	0.00	202	16	2	0
1991.3	87.73	10.45	1.82	0.00	193	23	4	0	1999.1	95.00	4.09	0.91	0.00	209	9	2	0
1991.4	88.18	10.91	0.91	0.00	194	24	2	0	1999.2	94.09	5.00	0.91	0.00	207	11	2	0
1992.1	88.64	8.64	2.73	0.00	195	19	6	0	1999.3	97.27	2.27	0.45	0.00	214	5	1	0
1992.2	90.45	7.27	2.27	0.00	199	16	5	0	1999.4	93.64	5.00	1.36	0.00	206	11	3	0
1992.3	86.82	10.45	2.73	0.00	191	23	6	0	2000.1	96.36	1.36	2.27	0.00	212	3	5	0
1992.4	77.73	17.27	4.55	0.45	171	38	10	1	2000.2	97.73	1.36	0.45	0.45	215	3	1	1
1993.1	71.82	22.73	5.45	0.00	158	50	12	0	2000.3	99.55	0.45	0.00	0.00	219	1	0	0
1993.2	66.36	20.45	11.82	1.36	146	45	26	3	2000.4	96.82	2.27	0.91	0.00	213	5	2	0
1993.3	76.36	16.36	7.27	0.00	168	36	16	0	2001.1	98.64	1.36	0.00	0.00	217	3	0	0
1993.4	70.00	18.64	10.00	1.36	154	41	22	3	2001.2	99.09	0.91	0.00	0.00	218	2	0	0
1994.1	76.36	15.91	7.27	0.45	168	35	16	1	2001.3	98.18	0.45	1.36	0.00	216	1	3	0
1994.2	82.27	10.91	6.82	0.00	181	24	15	0	2001.4	95.91	3.64	0.45	0.00	211	8	1	0
1994.3	90.45	8.18	1.36	0.00	199	18	3	0	2002.1	97.73	1.82	0.00	0.45	215	4	0	1
1994.4	81.82	12.27	5.91	0.00	180	27	13	0	2002.2	98.18	1.36	0.45	0.00	216	3	1	0
1995.1	87.73	9.09	2.73	0.45	193	20	6	1	2002.3	99.55	0.45	0.00	0.00	219	1	0	0
1995.2	85.00	14.09	0.91	0.00	187	31	2	0	2002.4	98.18	1.36	0.45	0.00	216	3	1	0
1995.3	90.91	7.27	1.82	0.00	200	16	4	0	2003.1	99.09	0.91	0.00	0.00	218	2	0	0
1995.4	81.82	14.09	3.64	0.45	180	31	8	1	2003.2	98.64	1.36	0.00	0.00	217	3	0	0
1996.1	95.91	4.09	0.00	0.00	211	9	0	0	2003.3	99.09	0.91	0.00	0.00	218	2	0	0
1996.2	89.55	7.73	2.73	0.00	197	17	6	0	2003.4	97.27	0.91	1.82	0.00	214	2	4	0
1996.3	93.18	5.91	0.91	0.00	205	13	2	0	2004.1	98.18	1.36	0.45	0.00	216	3	1	0
1996.4	82.73	14.09	3.18	0.00	182	31	7	0	2004.2	97.27	1.82	0.91	0.00	214	4	2	0
1997.1	93.18	5.91	0.91	0.00	205	13	2	0	2004.3	99.55	0.45	0.00	0.00	219	1	0	0
1997.2	89.55	7.73	2.27	0.45	197	17	5	1	2004.4	98.64	1.36	0.00	0.00	217	3	0	0
1997.3	95.00	4.55	0.45	0.00	209	10	1	0	2005.1	98.18	1.36	0.45	0.00	216	3	1	0
									2005.2	97.73	1.82	0.45	0.00	215	4	1	0

Figure 1: Percentage and number of banks in each regime by quarter.

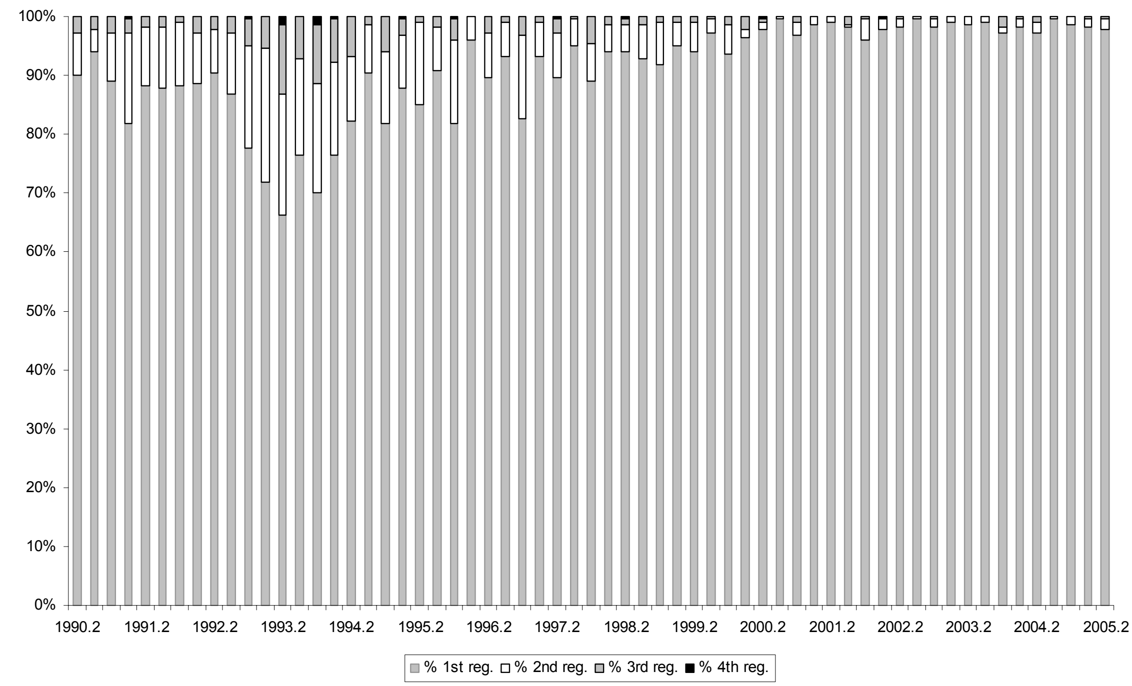
Panel A: Model (1) 2 regimes with panel data: threshold dr_{it}



Panel B: Model (2) 3 regimes with panel data: threshold dr_{it}



Panel C: Model (3) 4 regimes with panel data: threshold dr_{it}



Panel D: Model (4) 2 regimes with panel data: threshold GAP_{t-1}

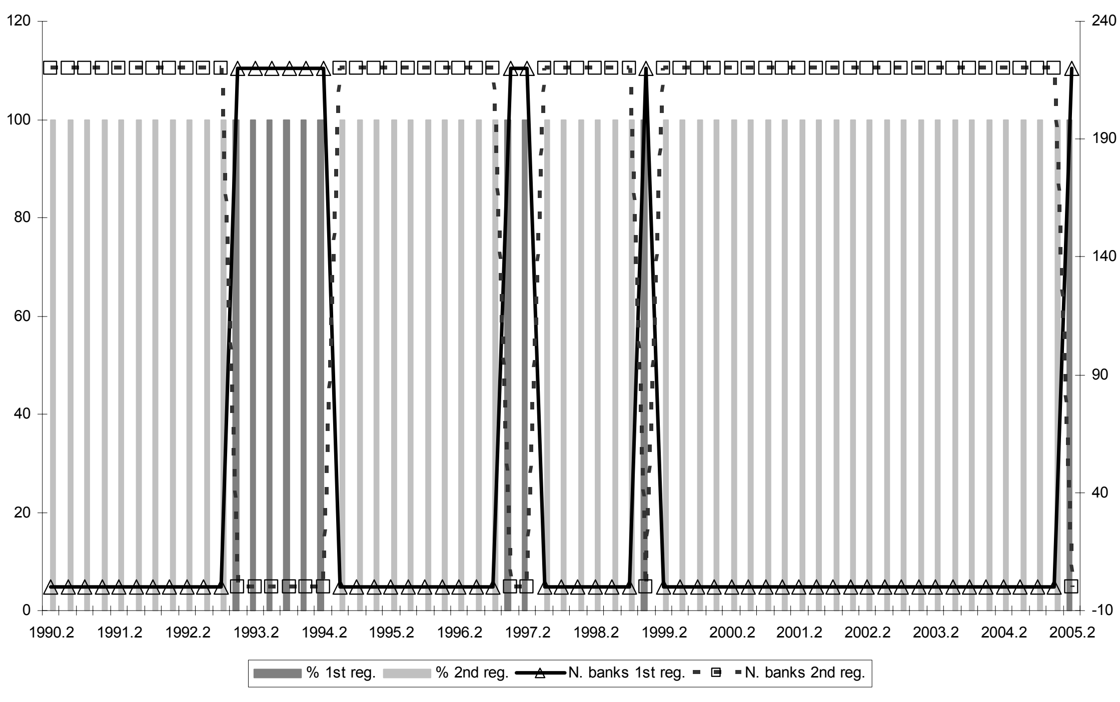


Table 6: Estimates for threshold regression panel data models with 4 regimes over different threshold variables

Notes: This table reports the conditional LS estimates for the following threshold panel data model with 4 regimes (different threshold variable: one micro and one macro). ***, ** and * indicate significance at 1, 5 and 10% respectively.

Model (1):

$$dr_{it} = \alpha_1 \ln\left(TA_{it}\right)^2 + \alpha_2 \text{lgr}_{it} + \alpha_3 (\text{lgr}_{it})^2 + \beta_{11} GAP_{t-1} I\left(dr_{it} \le \gamma_1\right) I\left(GAP_{t-1} \le \gamma_2\right) + \beta_{12} GAP_{t-1} I\left(dr_{it} \le \gamma_1\right) I\left(GAP_{t-1} > \gamma_2\right)$$
$$+ \beta_{13} GAP_{t-1} I\left(dr_{it} > \gamma_1\right) I\left(GAP_{t-1} \le \gamma_2\right) + \beta_{14} GAP_{t-1} I\left(dr_{it} > \gamma_1\right) I\left(GAP_{t-1} > \gamma_2\right) + e_t$$

Model (2):

$$dr_{it} = \alpha_1 \ln\left(TA_{it}\right)^2 + \alpha_2 \text{lgr}_{it} + \alpha_3 (\text{lgr}_{it})^2 + \beta_{11} GAP_{t-1} I\left(dr_{it-1} \le \gamma_1\right) I\left(GAP_{t-1} \le \gamma_2\right) + \beta_{12} GAP_{t-1} I\left(dr_{it-1} \le \gamma_1\right) I\left(GAP_{t-1} > \gamma_2\right)$$
$$+ \beta_{13} GAP_{t-1} I\left(dr_{it-1} > \gamma_1\right) I\left(GAP_{t-1} \le \gamma_2\right) + \beta_{14} GAP_{t-1} I\left(dr_{it-1} > \gamma_1\right) I\left(GAP_{t-1} > \gamma_2\right) + e_t$$

Model (3), (4) and (5): same as (1) with dr_{it} replaced by dr_{it}^{LTD}, dr_{it}^{CO} and dr_{it}^{MU} respectively. Model (6), (7) and (8): same as (2) with dr_{it-j} replaced by dr_{it-j}^{LTD}, dr_{it-j}^{CO} and dr_{it-j}^{MU} respectively ($j = 0,1$).

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
α_1	-0.0034 ***	-0.0077 ***	-0.0068 ***	-0.0057 ***	-0.0016 ***	-0.0106 ***	-0.0081 ***	-0.0059 ***
	(0.0003)	(0.0004)	(0.0003)	(0.0002)	(0.0003)	(0.0003)	(0.0002)	(0.0004)
α_2	-0.2108 ***	-0.3431	-0.1320 ***	-0.2387 ***	-0.2261 ***	-0.1559 ***	-0.4362 ***	-0.4798 ***
	(0.0529)	(0.0883)	(0.0336)	(0.0441)	(0.0300)	(0.0494)	(0.0545)	(0.0431)
α_3	-0.4888 ***	-0.7988 ***	-0.3422 ***	-0.6762 ***	-0.5628 ***	-0.4484 ***	-1.1553 ***	-1.0904 ***
	(0.1127)	(0.1905)	(0.0656)	(0.1130)	(0.0595)	(0.0995)	(0.1377)	(0.0870)
β_{11}	0.0956 ***	-0.0273 ***	0.0689 ***	0.0652 ***	0.0978 ***	-0.0240 ***	-0.0090 **	-0.0344 ***
	(0.0039)	(0.0057)	(0.0038)	(0.0035)	(0.0038)	(0.0051)	(0.0040)	(0.0085)
β_{12}	-0.1254 ***	-0.0200 ***	-0.1135 ***	-0.1007 ***	-0.1258 ***	-0.0230 ***	-0.0508 ***	-0.0214 ***
	(0.0045)	(0.0055)	(0.0041)	(0.0034)	(0.0046)	(0.0049)	(0.0036)	(0.0031)
β_{13}	-0.2929 ***	-0.1403 ***	-0.2592 ***	-0.2684 ***	-0.3457 ***	-0.1765 ***	-0.2422 ***	-0.1388 ***
	(0.0066)	(0.0082)	(0.0052)	(0.0058)	(0.0087)	(0.0058)	(0.0093)	(0.0222)
β_{14}	0.2960 ***	0.0521 ***	0.2195 ***	0.2244 ***	0.4054 ***	0.0804 ***	0.1405 ***	0.0047
	(0.0087)	(0.0112)	(0.0066)	(0.0072)	(0.0114)	(0.0082)	(0.0174)	(0.0117)
γ_1	0.417	0.549	0.506	0.561	0.373	0.546	0.865	0.644
γ_2	-0.074	0.034	-0.074	-0.074	-0.074	-0.131	-0.385	-1.647
N. of banks	220	220	80	18	122	80	18	122
N. of quarters	61	61	61	61	61	61	61	61
N. of quantiles (micro)	393	393	393	393	393	393	393	393
N. of quantiles (macro)	60	60	60	60	60	60	60	60
Trimming %	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
LR Test	8367.10 ***	222.48 ***	6965.86 ***	6492.77 ***	11206.80 ***	756.25 ***	620.70 ***	39.60 ***
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
N. bootstrap	300	300	100	100	100	100	100	100