MASTER

APPLIED ECONOMETRICS AND FORECASTING

MASTER´S FINAL WORK

DISSERTATION

HOW BANK LENDING AFFECTS FIRMS' LIFECYCLE: A MARKOV CHAIN APPROACH

CLOÉ LEAL DE MAGALHÃES

OCTOBER - 2019

LISBON SCHOOL OF ECONOMICS & MANAGEMENT
UNIVERSIDADE DE LISBOA

MASTER
APPLIED ECONOMETRICS AND FORECASTING

MASTER´S FINAL WORK
DISSERTATION

HOW BANK LENDING AFFECTS FIRMS' LIFECYCLE: A MARKOV CHAIN APPROACH

CLOÉ LEAL DE MAGALHÃES

SUPERVISION:
PIERRE JOSEPH MARIA HOONHOUT

OCTOBER - 2019

GLOSSARY

BPLim – Banco de Portugal's Microdata Research Laboratory

CBSD – Central Balance Sheet Database.

CCR – Central Credit Register.

IES – Informação Empresarial Simplificada.

MNL – Multinomial logit model.

NFC – Non-financial corporations.

ABSTRACT, KEYWORDS AND JEL CODES

In the last decade, a high proportion of inefficient firms has been observed in the most advanced economies, contradicting the firm dynamics theory that states that economically unviable firms exit the market. Recent research claims that survival of ‘zombie’ firms is related to banks’ lending behaviour, as financially stressed firms are allowed to survive due to the credit granted to them.

This dissertation analyses the impact of additional loans over the probability for the Portuguese firms to move across *profitable*, *non-profitable* and *exit* states. In particular, the probabilities for a *non-profitable* firm to remain *non-profitable* or turn to *profitable* or *exit* are analysed. Using firm and bank level information for the 2011-2015 period, an autoregressive multinomial logistic model is estimated to describe the firms’ dynamics as an absorbing Markov chain, conditional on the existence of an additional bank loan, and controlling for other covariates.

The results contradict the hypothesis that banks’ lending behaviour is keeping ‘zombie’ firms from exiting the market. *Non-profitable* firms that were granted an additional loan during the 2011-2015 period were more likely to recover to the *profitable* state, while having a lower probability of remaining *non-profitable* and keeping a similar probability of *exit*. As the firms’ ‘zombieness’, measured by the probability to remain *non-profitable*, was lower for firms with additional bank loans, this points to a low contribution of banks’ lending behaviour for the high proportion of inefficient firms observed in the Portuguese economy.

KEYWORDS: Bank Lending; Productivity; Zombie Firms, Absorbing Markov Chains, Multinomial Logistic Model.

JEL CODES: C25; D22; G21, G33, L25.

TABLE OF CONTENTS

TABLE OF FIGURES

HOW BANK LENDING AFFECTS FIRMS' LIFECYCLE: A MARKOV CHAIN APPROACH

By Cloé Leal de Magalhães

THIS DISSERTATION analyses how additional loans granted to non-profitable firms affect their probability to remain non-profitable, recover to profitable or exit the market. This assessment is carried out through the estimation of a Markov process conditional to the existence of additional bank loans, using the multinomial logit model estimates. Applying this model to Portuguese firm and bank level data from 2011 to 2015, the results point to a positive effect of additional bank loans over survival and recovery rates of non-profitable firms, contradicting some recent research on this topic.

1. INTRODUCTION

The productivity slowdown observed in the advanced economies since the early 2000s, and reinforced by the financial crisis in 2008, has raised several concerns towards the reasons behind the high prevalence of non-productive firms. Firm's entry, exit and growth has been widely analysed through the models for firm dynamics started by Hopenhayn (1992). These models work under the assumption of a strict rationality for firms and lenders to maximize profits, and assume that firms exit the market when its net value is negative; therefore, they are unable to describe the prevalence of firms that are economically unviable.

In recent years, a number of studies have focused in the relation between the financial system and the non-financial corporations to analyse this phenomenon, following the seminal paper by Caballero, Hoshi and Kashyap (2008) on the role of the financial system in the Japanese stagnation during the decade of 1990's. The hypothesis behind these studies is that economically unviable firms were able to survive due to banks' assistance, lowering the aggregate productivity and reducing opportunities for entrants.

This dissertation aims at answering the following question: what is the effect of granting an additional bank loan over non-profitable firms' dynamics? To pursue this objective, the impact of bank lending over the dynamics of Portuguese firms from 2011 to 2015 was estimated. This time span is characterized by the reinforcement of financial regulation over banks, and by a high proportion of financially stressed firms, low profitability and financial deleverage for firms. The macroeconomic context, along with the availability of rich datasets with microdata, offer the possibility to further investigate the relation between a stressed financial system and the high prevalence of inefficient firms.

Making use of firm and bank level microdata, a multinomial logistic model was used to estimate a Markov process where firms walk across states according to their economic status (*profitable*, *non-profitable* and *exit*). The hypothesis that ‘zombie’ firms are maintained by banks’ assistance is evaluated by the effect of an additional loan over the probability for a firm to remain *non-profitable*. The results contradict the hypothesis that banks’ lending behaviour was the cause of the high prevalence of ‘zombie’ firms in the Portuguese economy, as *non-profitable* firms with additional loans from their banks presented a higher probability to survive as *profitable* firms, a lower probability to remain *non-profitable* and a similar probability of *exit*.

This dissertation is organized as follows: section 2 provides a literature review on firm dynamics and the relation between bank lending and ‘zombie’ firms; section 3 describes the Markov process for firms’ dynamics and the multinomial logistic model that was used to estimate the process; and finally section 4 provides the empirical results and discussion.

2. Firm's Dynamics, bank lending and Zombie firms

Economic growth is a major concern among researchers. From a micro perspective, economic growth arises from the firms' capacity to produce goods and services in an efficient way. For that reason, a large number of models was developed to describe firms' dynamics. In particular, the models for firms' endogenous entry, exit and growth were built upon the work of Hopenhayn (1992). According to this models, firms choose the amount of capital and labour (and therefore their size) that maximizes the expected profit, given by their productivity functions and their costs of production. Firms exit the market if the expected profits are negative. The Hopenhayn model also included idiosyncratic shocks that could explain firm's behaviour; later developments included other determinants. Cooley and Quadrini (2002) and Albuquerque and Hopenhayn (2004), in particular, developed models to account for financial frictions considering financial costs for both equity and debt as determinants for the choice of the optimum level of capital.

Cooley and Quadrini (2002) found that financial frictions improved the capacity for these models of endogenous growth to explain empirical regularities observed for firms' dynamics according to their size (smaller firms rely more on debt) and to their age (younger firms rely more on debt). Albuquerque and Hopenhayn (2004) focused on the firm-lender relation to describe firms' decisions to entry, exit and growth. Under this model, the firm needs the lender to fund the initial investment and the early working capital; the lender provides the debt only if the expected values of the project is positive. The lender continues to finance the project as long as the debt can be repaid, and the firm is liquidated by the lender as soon as the expected value of the project is lower than the liquidation value.

Though the models for endogenous growth are able to describe some stylized facts about firm's dynamics, they are unable to explain the high prevalence of unviable firms that has been observed in the most advanced economies, nor the high amount of credit granted to those firms. Adalet McGowan (2017) estimates 'zombie' firms to account for 6% of the firms, on average, in eight OECD countries. Alexandre (2017), using different criteria, estimates that zombies accounted for 35% of Portuguese firms in 2012 and 26% in 2015.

After the financial crisis and the economic slowdown that followed, research turned to the relation between the financial system and firms as a cause for the prevalence of economically unviable firms. The role of the banking system in the Japanese economic stagnation in the 1990s was widely studied as a case of banks' assistance to inefficient firms, creating bank-related exit barriers. In their seminal study, Caballero, Hoshi and Kashyap (2008) found evidence of subsidized loans to insolvent firms in Japan, and measured the negative impact of those subsidized firms ('zombies') on the aggregate productivity and on firms' dynamics. Peek and Rosengren (2005) found evidence of Japanese banks 'evergreening' credit to zombie firms to avoid losses in their balance-sheets, under the benevolent supervision of public authorities, who wished to avoid the financial and political costs of massive firms and banks bankruptcy. Other reasons could be behind this behaviour, namely assisting troubled firms included in the same *keiretsu* (Japanese economic group).

Following the research on the Japanese case, some authors searched for an 'evergreening' behaviour in the banking system as the proportion of less productive firms started to rise in the OECD economies after the financial crisis in 2008. After the crash of Lehman Brothers and the financial crisis that followed, the financial regulation has been reformed in Europe, putting additional pressure on banks. The Basel II Accord gave place to the Basel III Accord, the Equity Ratios were revised and new levels of capital requirements were defined (Haldane and Neumann, 2016). The Capital Requirements Directive IV was introduced in 2013 to implement the Basel III Accord in the European Union, along with the new financial regulation architecture in the European Union introduced by the European Banking Union (Carletti and Leonello, 2016). This Directive raised the European bank's Core Equity Tier 1 (CET 1) from 2.5% to 4% of risk-weighted assets. In addition, the European Banking Authority (EBA) introduced the stress tests in 2010, which were significantly improved in 2011 (Petrella and Resti, 2016). During this process to reinforce the financial regulation, banks were frequently asked to raise capital to meet the new requirements.

While this reform was taking place, the level of overdue loans ratio was increasing as the economic situation was deteriorating, putting additional pressure for banks to meet the capital requirements. Blattner et al. (2018) found evidence of banks' assistance to Portuguese firms after the EBA's 2011 stress test, followed by and obligation for some

Portuguese banks to raise their capital ratios by 2012. The authors point two reasons that could be behind the banks' assistance to zombies: (i) to avoid the recognition of impairments, which would affect banks' capacity to meet the capital requirements, and (ii) risk-shifting behaviour, as stressed banks gamble with the probability for a zombie to recover. Banks might be using soft information to evaluate which firms are more likely to recover and choose to assist those firms. This hypothesis had been rejected by Peek and Rosengren (2005), who found no evidence of this 'cherry picking' the best zombies to improve their portfolio; on the contrary, these authors focused on the attempt to avoid losses as the main cause for zombie lending. According to Blattner et al. (2018), some firms are more likely to be assisted by banks: firms with larger loans, firms with insufficient collateral to cover the non-performing debt and firms with a long relation with their banks.

Overall, the 'zombie literature' points to a causal relation between banks' assistance to less productive firms, which acted as an exit barrier, and the productivity slowdown observed in the OECD economies during the last decade. Adler et al. (2017) points to the relation between zombies and the capital misallocation that lowers the aggregate productivity. Adalet McGowan et al. (2017) analyses zombie firms in the OECD countries and suggests that banks' assistance to less productive firms favours the survival of inefficient firms at the cost of diminishing the opportunities for entrants and healthy incumbents. Gouveia and Osterhold (2018) found evidence of the negative effects of a high proportion of zombies on the Portuguese aggregated productivity, as more resources are sunk in less efficient firms and the entrance and exit channels get distorted by a higher mean productivity of entrants and a lower mean productivity of exiters.

To analyse the impact of banks' behaviour over the firms' performance, many approaches have been proposed in the literature. As the recent research on zombie firms assumes this as a state caused by banks' assistance, the zombie definition focuses mainly on the firms' financial costs. Adalet McGowan et al. (2017), Gouveia and Osterhold (2018) and Azevedo et al. (2018) propose a definition of zombie firms based on their capacity to pay the debt cost, defining zombies as firms with an interest coverage ratio (EBITDA over interest expenses) below 1 for three consecutive years; firms with less than 10 years of activity are excluded from this definition. Caballero, Hoshi and Kashyap,

(2008) and Alexandre (2017), on the other hand, define zombies as firms with a debt cost below the theoretical cost, given the risk level.

This study proposes a different approach to analyse the zombie phenomenon. First, this analysis turns to a definition of zombie based on the firms' operating performance, in the spirit of the firm dynamics models. Secondly, it is based on individual transitions across states of profitability and non-profitability, as a stable proportion of zombies in the economy across time may be given by a changing population of zombies. In this sense, the firms' 'zombieness' is measured by the probability for a firm to remain non-profitable.

3. THEORETICAL FRAMEWORK: NON-EFFICIENCY AS A STATE IN FIRMS' LIFECYCLE

As stated in the previous section, this analysis turns to the definition of unviable firms that follows from the models for endogenous firm growth. The focus relies on the firms' capacity to produce goods and services in an efficient way, in the sense that the firm is able to obtain a profit from its operating activity. This definition of efficiency is necessarily linked to its productivity: more productive firms are likely to be profitable. Therefore, *profitable* is a dichotomous variable assuming 1 if the firm is able to obtain a profit from its activity, and 0 otherwise. As the firms may face losses for strategic reasons (investment plan, entering new markets or launching new products), an additional condition was added to capture firms with losses during a time of growth of their activity. Therefore, in this context, a firm is *profitable* if (i) has positive operating profit, or (ii) the sales' annual rate of change is over 2%. Firms are *non-profitable* otherwise.

The model for the firms' lifecycle has four states, as in Figure 1. *New* is the state for firms with 2 years of activity. After the second year of activity, firms either survive into a *profitable* firm at a rate *e*, turn into a *non-profitable* firm at a rate *n*, or *exit* at a rate of $1 - e - n$. The factors that determine the survival of the firms in the first years of activity are out of the scope of this analysis, therefore firms that do not survive until the third year of activity are excluded to avoid the contamination of the results.

Profitable firms can change into *non-profitable* at a rate *z*, exit at a rate *d*, or remain *profitable* with a probability of $1 - z - d$. *Non-profitable* firms may recover to *profitable* at a rate *r*, *exit* at a rate *w* or remain *non-profitable* at a rate $1 - r - w$. The probability to remain *non-profitable* is a measure of the firms' 'zombieness': the higher this probability, the longer the firms are expected to remain in this state, therefore surviving for a longer period without being viable from an economic point of view.

If firms *exit*, they are absorbed in this state, therefore the probability to remain in this state is 1. Firms are in this state if (i) are identified as ceased in the database, or (ii) in the last year they are observed in the database, if they have at least two years afterwards without data.

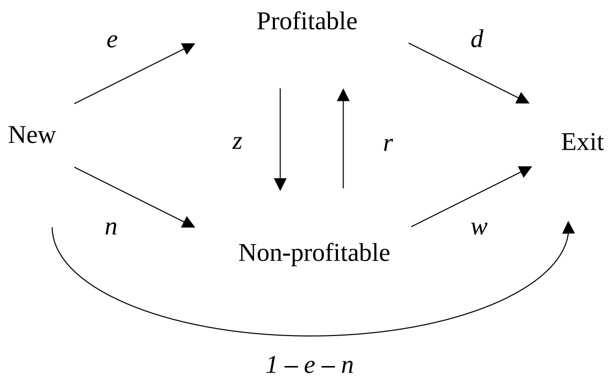


FIGURE 1 – Model for firms' dynamics

3.1. Absorbing Markov chains

Having a set of M states $\{S_1,\ldots,S_M\}$, a *first-order Markov chain* is a process that describes the passage of the observations from one stage to another. Each move is called a *step*. If the process is in state S_i in period *t*, the probability to move to state $S_j, j = 1, \ldots, M$ in *t+1* is called the *transition probability* and is given by:

(3.1.1) $$P(Y_{t+1} = S_j | Y_t = S_i) = p_{ij}$$

By assumption, the transition probability p_{ij} does not depend on where the process was in *t - 1*, *t - 2*, … This implies that

(3.1.2) $$p_{ij} = P(Y_{t+1} = S_j | Y_t = S_i) = P(Y_{t+k+1} = S_j | Y_{t+k} = S_i), \forall\, k$$

Therefore, period *t* is not necessary to describe the process, but only the number of steps. The transition probabilities are elements of the *1-step transition matrix* $P_{(MxM)}$.

All the process can be described as a function of P. Considering $p_{ij}^{(2)}$ as the probability for the process to move from state S_i to S_j in two steps, it corresponds to the probability to move from S_i to any of the M states in step 1 and from here to state S_j in step 2:

(3.1.3) $$p_{ij}^{(2)} = \sum_{k=1}^{M} p_{ik}p_{kj}$$

This is equivalent to the product of the i-th line and the j-th column of P. Therefore, matrix P^2 is the 2-step transition matrix of the process. Following this, P^N is the N-step transition matrix of the process, and the element $p_{ij}^{(N)}$ of this matrix is the probability for the process to start in S_i and be in S_j after N steps.

State S_k is an *absorbing state* if, once the process reaches this state, it is impossible to leave it. This implies that $p_{kk} = 1$. A M-state Markov Chain can have at most M-1 absorbing states; if there is at least one absorbing state, the process is said to be an *absorbing markov chain*. In these processes, the states that are not absorbing are *transient states*.

The canonical form of an M-state absorbing Markov chain with T transient states and A absorbing states (T + A = M) is

(3.1.4) $$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{R} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}$$

Having $P_{(MxM)}$ as the 1-step transition matrix, $Q_{(TxT)}$ as the matrix of the transition probabilities between transient states, $R_{(TxA)}$ as the transition matrix from transient to absorbing states, $0_{(AxT)}$ as a matrix of zeroes (as the transition probabilities from absorbing to transient states are null) and $I_{(AxA)}$ as the identity matrix.

For N steps, the transition matrix is given by

(3.1.5) $$\boldsymbol{P}^{\boldsymbol{N}} = \begin{bmatrix} \boldsymbol{Q}^{\boldsymbol{N}} & \boldsymbol{R}^{*} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}$$

Where $\boldsymbol{R}^{*} = \boldsymbol{f}(\boldsymbol{Q}, \boldsymbol{R})$. The process is absorbed after reaching one of the A absorbing states; thus, as N increases, the probability of being in a transient state gets close to zero:

(3.1.6) $$\lim_{N\to\infty} \boldsymbol{Q}^{\boldsymbol{N}} \to \boldsymbol{0}$$

A variable of interest is the expected number of steps before the process is absorbed, which can provide the 'life expectancy' given the process starts in one of the M states. Assuming the process starts in the transient state S_i, the probability of being in transient state S_j after N steps is q_{ij}^N. Making X_N as a dummy variable that assumes 1 if the process is in state S_j after N steps, then

(3.1.7) $$\boldsymbol{P}(X_N = 1) = E(X_N) = q_{ij}^N$$

From here, it is possible to know the expected number of times the process is in state S_j after N steps, given the process starts on stage S_i:

(3.1.8) $$E(X_0 + X_1 + \cdots + X_N) = \sum_{k=0}^{N} q_{ij}^k$$

Note that $Q^0 = \mathrm{I}$, allowing to include the initial state in the count. When $N \to \infty$ then

(3.1.9) $$E(X_0 + X_1 + \cdots) = \sum_{k=0}^{\infty} q_{ij}^k = n_{ij}$$

having n_{ij} as the element of the *fundamental matrix* $N = I + Q + Q^2 + \cdots$ that gives the expected number of times the process is in state S_j before being absorbed, given it started in state S_i. It can be proven that $N = (I - Q)^{-1}$, by doing:

(3.1.10) $$(I - Q)N = (I - Q)(I + Q + Q^2 + \cdots)$$
$$(I - Q)N = I - Q^N$$

From (3.1.6) it is known that as N tends to infinity, $I - Q^N = I$ and

(3.1.11) $$N = (I - Q)^{-1}$$

Having N, it is possible to obtain the expected number of steps before the process is absorbed, given it starts in the transient state S_i:

(3.1.12) $$t_i = \sum_{k}^{M} n_{ik}$$

Considering **c** as a column vector of 1, the $t_{(Mx1)}$ column with the expected number of steps before the process is absorbed for all the transient states is given by:

(3.1.13) $\boldsymbol{t} = \boldsymbol{N}\boldsymbol{c}$

Markov chains can be estimated with a multinomial logistic regression model, as described in the next section.

3.2. Multinomial logistic model

The *multinomial logistic model* (MNL) applies to the cases where an individual can choose among several unordered outputs. Considering the categorical dependent variable $Y = 0, \dots, J$, the MNL predicts the relative probability of obtaining one outcome *j*, given a set of individual attributes:

(3.2.1) $$P(Y = j|\boldsymbol{x}) = \frac{e^{x\boldsymbol{\beta_j}}}{\sum_{h=0}^{J} e^{x\boldsymbol{\beta_h}}}$$

The MNL allows the inclusion of K regressors, including a constant term and individual characteristics in the form of continuous or categorical variables. The regressor matrix $\boldsymbol{x}_{(\boldsymbol{NxK})}$ is the same for all possible *J* outcomes, but the effect of the regressors, given by the coefficient vectors $\boldsymbol{\beta_j}$, is choice-specific. The more general case where the set of regressors is choice-specific refers to the *conditional logit model*, and lies outside the scope of this study.

In the latent variable formulation, the probability of obtaining one outcome (e.g., outcome 1) is given by

(3.2.2) $$P(Y = j|\boldsymbol{x}) = P(Y_1^* > Y_l^*, \forall\, l = 2, \dots, J)$$

Having $Y_j^* = 1, \dots, J$ as the latent variable for output *j* defined as follows:

(3.2.3) $$Y_j^* = \boldsymbol{x\beta_j} + \boldsymbol{u_j}$$

Where $\boldsymbol{u_j}$ follows a standard type-1 extreme value distribution.

To allow the estimation of M-state Markov chains, the MNL must include state dependence through the inclusion of the lagged dependent variable. The *J* - state Markov chain conditional to the set of exogenous variables $\mathbf{z}$ has a 1-step transition matrix with elements i, j given by:

(3.2.4) $$p_{ij} = P(Y_t = j|\boldsymbol{z_t}, Y_{t-1} = i)$$

Having $i, j = 0, \dots, J$. For simplicity, $\boldsymbol{x} = [\boldsymbol{z_t} \quad Y_{t-1}]$ and the subscript *t* is removed, as in (3.1.2), without loss of generality.

The lagged dependent variable is necessarily endogenous, as $E(u_t|\boldsymbol{x}_k) \neq 0$ for $t = k - 1$ because $\boldsymbol{x}_{it+1}$includes y_{it}. Nevertheless, the MNL estimator still produces consistent estimators if y_{it} is sequentially exogenous, i.e., $E(\boldsymbol{u}_t|\boldsymbol{x}_k) = 0$ for $t \geq k$ (see Wooldridge, 2010, pages 482-3 for a discussion of the use of the lagged dependent variable on the pooled probit and logit models). This is equivalent to assume that the probability of outcome *j* conditional to $\boldsymbol{x}_t$ does not depend on any past values of $\boldsymbol{x}_t$, and therefore the model is dynamically complete:

(3.2.5) $$P(Y = j|\boldsymbol{x}_t, \boldsymbol{x}_{t-1}, \dots, \boldsymbol{x}_0) = P(Y = j|\boldsymbol{x}_t)$$

The pooled MNL is used in this analysis instead of the dynamic MNL for two reasons: first, the panel is unbalanced, as the entry and exit of firms needs to be observed in order to estimate transition probabilities for new firms and for firms exiting the market. This limits the estimation of the dynamic MNL, which requires balanced panels to observe each individual though the entire period. Secondly, the purpose of this analysis is not to study firms' individual behaviour itself, but the population as a hole, which changes from year to year. The use of the pooled MNL is equivalent as considering the observations for each year as independent samples.

It follows from (3.2.1) that the model is indetermined. As the sum of the probabilities always adds up to one, setting any $\boldsymbol{\beta_j^*}$, the remaining $\boldsymbol{\beta_k}$ $, \neq j$ will adjust to obtain the same probabilities for the *J* outcomes. The model is solved by defining outcome *J = 0* as the pivot category and making $\boldsymbol{\beta_0} = \boldsymbol{0}$:

(3.2.6) $$\begin{cases} P(Y = 0|\boldsymbol{x}) = \frac{1}{1+\sum_{h=1}^{J} e^{x\boldsymbol{\beta_h}}} \\ P(Y = j|\boldsymbol{x}) = \frac{e^{x\boldsymbol{\beta_j}}}{1+\sum_{h=1}^{J} e^{x\boldsymbol{\beta_h}}}, j = 1, \dots, J \end{cases}$$

Therefore, coefficient vectors $\boldsymbol{\beta_j}, j = 1, \dots, J$ are relative to a base outcome, and have no absolute interpretation. Also, provide little information about the marginal effect of the regressors on the outcome. The marginal effect of x_k over $P(Y = j|\boldsymbol{x})$ is:

(3.2.7) $$\frac{\partial P(Y=j|\boldsymbol{x})}{\partial x_k} = \frac{e^{\boldsymbol{x\beta_j}}}{1+\sum_{h=1}^{J} e^{\boldsymbol{x\beta_h}}} \left[\beta_{jk} - \frac{\sum_h^J \beta_{hk} e^{\boldsymbol{x\beta_h}}}{1+\sum_h^J e^{\boldsymbol{x\beta_h}}}\right]$$

In the MNL, $\boldsymbol{\beta_j}$ provides information on the effect of $\boldsymbol{x}$ on the probability of obtaining $Y = j$, relative to the probability of obtaining any other outcome:

(3.2.8) $$\frac{P(Y=j|\boldsymbol{x})}{P(Y=m|\boldsymbol{x})} = \frac{e^{\boldsymbol{x\beta_j}}}{e^{\boldsymbol{x\beta_m}}} = e^{\boldsymbol{x}(\boldsymbol{\beta_j}-\boldsymbol{\beta_m})}$$

Expression (3.2.8) gives the odds-ratio of obtaining outcome *j* relative to obtain outcome *m*. The odds-ratio does not depend on the remaining outcomes. To obtain a measure linear on the coefficients, the log-odds ratio can be used, as follows:

(3.2.9) $$ln\left[\frac{P(Y=j|\boldsymbol{x})}{P(Y=m|\boldsymbol{x})}\right] = \boldsymbol{x}(\boldsymbol{\beta}_j - \boldsymbol{\beta}_m)$$

If *m* is the base outcome, it follows that the log-odds ratio is equal to $\boldsymbol{x\beta}_j$. Therefore, the signs of $\boldsymbol{\beta}_j$ show the direction of the effect of each x_k on the probability of outcome *j* relative to the probability of the base outcome.

The MNL predicts relative probabilities (odds ratios), and those probabilities are not a linear function of the coefficients. The marginal effect of each regressor on the probability of each outcome, as in (3.2.7), depends on all $x_k, k = 1, \dots, K$ and all $\boldsymbol{\beta_j}, j = 1, \dots, J$. The term inside brackets in (3.2.7) shows that the sign of the marginal effect is not necessarily given by β_{jk}. Therefore, there is little interpretation for the coefficients of the MNL.

In such cases, the *average marginal effect* is often used to access the effect of the regressors over the expected outcome. For a given set of observed (fixed) *K-1* regressors $\boldsymbol{x_0}$, the average partial effect of x_k over outcome *J* is given by

(3.2.10) $$AME = \frac{1}{n}\sum_i^n \frac{\partial P(Y=j|\boldsymbol{x_0},x_k)}{\partial x_k}$$

Expression (3.2.10) corresponds to the marginal effect of x_k over outcome *J*, given in (3.2.7), averaged for the *n* observations in the sample, keeping the remaining regressors as observed. If x_k is a dichotomous variable, the average marginal effect can be estimated as the average difference between $P(Y = j|\boldsymbol{x_0}, x_k = 1)$ and $P(Y = j|\boldsymbol{x_0}, x_k = 0)$ for the observations in the sample.

The MNL model is estimated by maximum likelihood, using the log-likelihood function

$$(3.2.11) \qquad \mathcal{L}(\boldsymbol{\beta}) = \textstyle\sum_i^n \ell_i(\boldsymbol{\beta}) = \sum_i^n \sum_j^J \left[\tilde{y}_{ij} \ln\left(\frac{e^{x_i\beta_j}}{\sum_k e^{x_i\beta_k}}\right)\right]$$

Where $\tilde{y}_{ij}$is 1 if $y_{ij} = j$ and 0 otherwise. The standard errors for $\widehat{\boldsymbol{\beta}}_{\boldsymbol{MNL}}$ are obtained from the partial second derivatives of the likelihood function, as in Hosmer (2000).

3.3. Neglected heterogeneity

When relevant explanatory variables are omitted from the model, the estimation of the MNL is valid if the unobserved variables are uncorrelated with $\boldsymbol{x}$. In the presence of neglected heterogeneity, model (3.2.1.) will be given by

$$(3.3.1) \qquad P(Y = j|\boldsymbol{x}, \boldsymbol{c}) = \frac{e^{\boldsymbol{x\beta_j+c}}}{\sum_{h=1}^J e^{\boldsymbol{x\beta_h+c}}}$$

Where $\boldsymbol{c}$ is the unobserved explanatory variable. In a panel data model, $\boldsymbol{c}$ can account for individual time variant or invariant unobserved variables. If $\boldsymbol{c}$ is independent of $\boldsymbol{x}$, than the parameter estimates $\widehat{\boldsymbol{\beta}}_{\boldsymbol{MNL}}$ are inconsistent, but not the average marginal effect of $\boldsymbol{x}$ on $P(Y = j|\boldsymbol{x}, \boldsymbol{c})$ (see Wooldridge, 2010, pages 470-2 for a discussion on the effect of neglected heterogeneity over the average marginal effects of the probit and logit models).

If there is one x_k not independent of $\boldsymbol{c}$, than x_k is an endogenous regressor and the MNL estimation produces inconsistent estimators for $\boldsymbol{\beta}$ and for the average marginal effects. To test for endogenous regressors, Wooldridge (2010) suggests the Rivers and Vuong 2-step procedure for the probit model: (1) OLS regression of x_k on all other (exogenous) regressors and an instrumental variable v; (2) probit the binary dependent variable y on $\boldsymbol{x}$ and the residuals of the first step regression. The *t*-test on the coefficient of the residuals in the second regression is equivalent to test the null hypothesis of exogenous x_k. Concerning the MNL, the references used in this study do not provide a valid endogeneity test.

3.4. Hypothesis tests

To access the significance of the regressors in the MNL, the standard *z*-test holds, as $\hat{\beta}_{k,MNL}$ follows a normal distribution (Hosmer, 2000). The test statistic for $H_0: \beta_k = 0$ is $z = \hat{\beta}_k / \sqrt{var(\hat{\beta}_k)}$.

Also, the Likelihood ratio test (LR) and the Wald test can be applied. The LR test has the null hypothesis $H_0: \boldsymbol{\beta}^* = \boldsymbol{0}$, having $\boldsymbol{\beta}^*$ as the column vector containing the *Q* coefficients to be tested. The LR test statistic is

$$(3.4.1) \qquad LR = 2\left[\mathcal{L}(\widehat{\boldsymbol{\beta}}) - \mathcal{L}(\widetilde{\boldsymbol{\beta}})\right]$$

Having $\widetilde{\boldsymbol{\beta}}$ as the coefficients' estimates of the restricted model and $\mathcal{L}(*)$ as the log likelihood function as in (3.2.9). Under the null hypothesis, the LR follows a χ^2 distribution with *Q* degrees of freedom.

To test more general hypothesis, both linear and nonlinear, the Wald test has the null hypothesis $H_0: c(\boldsymbol{\beta}^*) = \boldsymbol{q}$, considering $c(\boldsymbol{\beta}^*)$ as the column vector of *Q* continuous functions of $\boldsymbol{\beta}^*$ to be tested and $\boldsymbol{q}$ as the column vector with the values of $c(\boldsymbol{\beta}^*)$ to be tested. The Wald test statistic is

$$(3.4.2) \qquad W = [c(\boldsymbol{\beta}^*) - \boldsymbol{q}]'\left[C(\boldsymbol{\beta}^*)\widehat{\boldsymbol{V}}C(\boldsymbol{\beta}^*)'\right]^{-1}[c(\boldsymbol{\beta}^*) - \boldsymbol{q}]$$

Having $C(\boldsymbol{\beta}^*) = \frac{\partial c(\boldsymbol{\beta}^*)}{\partial \boldsymbol{\beta}^*}$. When the null hypothesis is $H_0: \beta_j = 0$, (3.4.2) reduces to $W = \frac{\widehat{\beta}_j^2}{\widehat{\sigma}_{\beta_j}^2}$. Under the null hypothesis, the Wald test statistic follows a χ^2 distribution with Q degrees of freedom.

As the output of interest of the MNL is the predicted probability for each outcome, it is relevant to test hypothesis for this statistic. Following Long (2009), the predicted probability for outcome *j* is a function of the estimated coefficients $\widehat{\boldsymbol{\beta}}_j$:

(3.4.3) $$P\left(y_i = j|\boldsymbol{x}, \widehat{\boldsymbol{\beta}}_j\right) = G\left(\widehat{\boldsymbol{\beta}}_j\right),$$

Where $G\left(\widehat{\boldsymbol{\beta}}_j\right)$ is the logistic function as in (3.2.1) in the case of the MNL. The delta method states that $G\left(\widehat{\boldsymbol{\beta}}_j\right)$ follows a normal distribution around $G\left(\boldsymbol{\beta}_j\right)$ with asymptotic variance given by

(3.4.4) $$Var\left[G\left(\widehat{\boldsymbol{\beta}}_j\right)\right] = \frac{\partial G\left(\widehat{\boldsymbol{\beta}}_j\right)}{\partial \widehat{\boldsymbol{\beta}}_j{}'} Var\left(\widehat{\boldsymbol{\beta}}_j\right) \frac{\partial G\left(\widehat{\boldsymbol{\beta}}_j\right)}{\partial \widehat{\boldsymbol{\beta}}_j}$$

The difference in the predicted probability of outcome *j* between two independent groups of observations (1) and (2) follows a normal distribution with asymptotic variance given by

(3.4.5) $$Var\left[G\left(\widehat{\boldsymbol{\beta}}_j\right)^1\right] + Var\left[G\left(\widehat{\boldsymbol{\beta}}_j\right)^2\right]$$

3.5. Goodness of fit

The logistic models differ from the linear regression by assuming a categorical output variable. The goodness of fit statistics for the linear regression model are not valid for the logistic models, as the residuals of the latter do not follow a normal distribution.

Several measures have been proposed to assess the goodness of fit of MNL. As the R^2 is not applicable in the case of the MNL, the Pseudo- R^2 proposed by McFadden is mostly used to assess the capacity of the model to fit the data. The Pseudo- R^2 is given by $1 - \mathcal{L}(\boldsymbol{\beta})/\mathcal{L}_0(\boldsymbol{\beta})$, having $\mathcal{L}_0(\boldsymbol{\beta})$ as the log-likelihood for the base model with only an intercept. The LR test can also be used, considering the base model with only an intercept as the restricted model.

Additionally, the Pearson's χ^2 test can be used to evaluate the goodness of fit of the MNL, if at least one $\boldsymbol{x}$ is continuous. The test statistic is given by:

(3.5.1) $$X^2 = \sum_i^n \sum_j^J \left[\frac{(\tilde{y}_{ij}-\hat{p}_{ij})^2}{\hat{p}_{ij}}\right]$$

Considering $\tilde{y}_{ij}$ as a binomial variable that assumes 1 if $y_i = j$ and 0 otherwise, and *n* as the sample size. Under the null hypothesis that $\widehat{\boldsymbol{\beta}}$ are the true coefficients of the model, the Pearson's test statistic follows a χ^2 distribution with $n(J-1)$ degrees of freedom.

Hosmer and Lemeshow (2000) analyse the goodness of fit for the MNL under the presence of categorical regressors, which leads to a number of possible regressor patterns *Q* inferior to the number of observations *N*. A special case is the saturated MNL model, presenting a finite number of possible regressor patterns. This affects the distribution of the residuals, which is given by the residual for each pattern, weighted by the probability for each pattern to occur. Considering $\hat{p}_{qj}$ the expected probability for output *j* in the regressor pattern *q*, the Pearson's χ^2 for the saturated model, which follows a χ^2 distribution with $Q(J-1)$ degrees of freedom, is given by

(3.5.2) $$X_S^2 = \sum_q^Q \sum_j^J \left[\frac{\left(\sum_i^{n_q} \tilde{y}_{qj} - \sum_i^{n_q} \hat{p}_{qj}\right)^2}{\sum_i^{n_q} \hat{p}_{qj}}\right]$$

The authors also propose the Hosmer-Lemeshow goodness of fit test statistic for the binomial logit, corresponding to the Pearson's χ^2 test over a sample of *G* equally distributed groups (percentiles). Fagerland, Hosmer and Bonfin (2008) propose an extension of the Hosmer-Lemeshow test for the multivariate logit, which is based in the division of the sample in *G* percentiles, ordered according to the predicted probability for each observation to be in one of the output categories except the base category $(1-\hat{p}_{i0})$:

(3.5.3) $$HL_G = \sum_g^G \sum_{j=1}^J \left[\frac{\left(\sum_i^{n_g} \tilde{y}_{gj} - \sum_i^{n_g} \hat{p}_{gj}\right)^2}{\sum_i^{n_g} \hat{p}_{gj}}\right]$$

Under the null hypothesis that $\widehat{\boldsymbol{\beta}}$ are the true coefficients of the model, the Hosmer-Lemeshow test statistic follows a χ^2 distribution with $(G-2)(J-1)$ degrees of freedom. It has been pointed out that the multivariate Hosmer-Lemeshow test is not invariant to the choice of the base category ($J=0$); also the definition of the degrees of freedom of the χ^2 distribution is not consensual.

4. EMPIRICAL ANALYSIS

4.1. Database

This analysis uses data from the Central Balance Sheet Database and the Central Credit Register from Banco de Portugal. The Central Balance Sheet Database (CBSD) is an annual database covering all non-financial corporations (NFC) operating in Portugal from 2006 to 2017. This information is collected through IES – Informação Empresarial Simplificada and benefits from the quality control procedures defined by Banco de Portugal to guarantee temporal and vertical consistency of the data.

The Central Credit Register (CCR) is a database comprising monthly data on all credit liabilities vis-à-vis the Portuguese financial corporations above 50€. This database provides detail on the amounts owed, the credit instrument, the original and residual maturity of the liabilities, and their situation (regular/overdue), among other information. This database is available from 2009 onwards.

The CBSD and CCR databases were accessed through Banco de Portugal's Microdata Research Laboratory (BPLim), which is a database with anonymized information at a firm and bank level made available for researchers by Banco de Portugal.

As the CCR database begins in 2009, and to allow the identification of exited firms and the use of lagged variables, the data used in this analysis covers the 2011-2015 period, on an annual basis. Though this dataset does not cover the period immediately after the financial crisis, it allows the observation of firms' dynamics during a period of financial turbulence and adjustment, marked by a strong pressure for banks to meet higher capital requirements.

Firms with no employees or annual sales under 10.000€ were excluded. Firms without bank loans registered on CCR, therefore with no responsibilities towards the Portuguese financial system, are also excluded from the estimation. Variable *State* indicates the state of the firm, as defined in section 3. According to the definition of the *New* state, only the transitions from the third year of activity onwards are observed; therefore firms with less than three years of activity are excluded from the sample. Variable *Bank* is a dummy variable assuming 1 if the firm got an additional loan above 100.000€ or representing more than 5% of total assets, and 0 otherwise. It should be noted that this variable assumes 1 only if there is a pre-existent relation with the bank granting the loan. As this variable intends to capture the loans 'evergreening' to *non-profitable* firms, the thresholds were defined to exclude retail exposures, mostly related to operations concerning the revolving of small amounts granted for liquidity purposes, which are recurrent among Portuguese firms. Tables IX to XI in Appendix provide additional information on the distribution of additional loans, in Euros and in percentage of assets, observed in the sample, along with an analysis of the impact of changing the criteria used to identify and additional bank loan[1].

Some variables were defined to characterize the sample. *Employees* is the number of remunerated persons employed by the firms during the year, used as a proxy for the firms' size. Alternatively, it may be considered *size*, a categorical variable assuming 1 for microfirms, 2 for small firms, 3 for medium firms and 4 for large firms, following the definition of the European Commission (2003). *Industry* is a categorical variable assuming one of 6 economic activities, as in Table XII in Appendix. *Sector depressed* is a dummy variable assuming 1 if the firms' sector of activity (defined at 2-digit level of the NACE code) presented a negative aggregate turnover growth rate, and 0 otherwise. This variable, interacted with *Bank*, allows to infer if banks additional loans have a different impact during depressed periods. Table XIII in Appendix provides the complete list of variables and their definitions.

The database contains 672,143 observations (Table I), corresponding to 205,470 distinct firms with an average of 3.3 observations in the database; only 38% of the firms are observed through the entire period (5 years). The observations are well distributed across the years covered in the database, the proportion of observations in each year rounding 20%, though this proportion is slightly decreasing by year (Table II). This is expected as the number of firms with credit relations with the financial system decreased during this period (Banco de Portugal, 2018). The activities with the largest proportion of observations are "Trade, accommodation and food services" (38.9%) and "Other

[1] Several thresholds for the additional loan, in total amount and in percentage of assets, were tested. The proportion of firms with additional bank loans is more sensible to changes in the threshold for the loan in percentage of total assets than for the threshold for total amount (Table X in Appendix). Nevertheless, the observed transition matrices across states remain similar using other thresholds, as shown in Table XI in Appendix.

services" (21.7%). On the opposite side, "Agriculture and mining" represents 3.3% of the observations.

TABLE I

NUMBER OF OBSERVATIONS AND MEAN VALUES

Variable	Obs.	Mean	Std. Dev.
Assets (€)	672 143	2 819 053	85 801 242
Sales (€)	672 143	1 812 320	35 999 355
Employees	672 143	14	128
Bank	672 143	0,15	0,359
Sector depressed	672 143	0,60	0,490
CCR exposure (€)	672 143	455 564	5 375 450

TABLE II

SAMPLE STRUCTURE

	Freq.	Percentage
by size		
Micro firms	508 623	75,7%
Small firms	136 503	20,3%
Medium firms	23 084	3,4%
Large firms	3 933	0,6%
by industry		
Agriculture and mining	22 090	3,3%
Manufacturing	99 851	14,9%
Utilities	54 120	8,1%
Construction and real estate	88 379	13,1%
Trade, accomm. and food	261 547	38,9%
Other services	146 156	21,7%
by year		
2011	150 008	22,3%
2012	138 000	20,5%
2013	130 404	19,4%
2014	125 942	18,7%
2015	127 789	19,0%

Around 60% of the observations concern to firms operating in a depressed activity; most of them occurred in 2011 and 2012, the years with the worst performance for the NFC in the period covered in this study. Additionally, firms operating in depressed activities are concentrated in the "Trade, accommodation and food services" (42.8%)

"Other services" (24.1%) and "Construction and real estate" (15.5%). Table XIV, in Appendix, provides further information on this variable.

The sample is mostly composed by micro firms (75.7%) and small firms (20.3%). Large firms correspond to 0.6% of the observations. Despite the exclusions based on sales and employees thresholds, this structure is very similar to the one observed for the population of non-financial corporations in Portugal (Banco de Portugal, 2018). The average firm in the sample has €2.8 million of assets and 14 employees, nearly twice the average Portuguese firm.

An additional bank loan was granted to 15% of the observations in the sample. A closer look at the *Bank* variable shows that 67% of the firms that received at least one additional loan between 2011 and 2014 received it only once during this time span; firms that record *Bank* = 1 twice, mostly get an additional loan in two consecutive years (Table III).

TABLE III

PATTERN OF VARIABLE BANK

2011	2012	2013	2014	2015	Obs.	Percent.	Percent cum.	# bank
0	0	0	0	1	11 593	16%	16%	1
1	0	0	0	0	11 545	16%	33%	1
0	0	0	1	0	9 427	13%	46%	1
0	0	1	0	0	7 865	11%	57%	1
0	1	0	0	0	6 967	10%	67%	1
0	0	0	1	1	3 060	4%	71%	2
0	0	1	1	0	2 100	3%	74%	2
0	0	1	0	1	1 977	3%	77%	2
1	1	0	0	0	1 587	2%	79%	2
1	0	1	0	0	1 442	2%	81%	2
0	1	1	0	0	1 386	2%	83%	2
1	0	0	0	1	1 377	2%	85%	2
0	1	0	1	0	1 366	2%	87%	2
1	0	0	1	0	1 364	2%	89%	2
0	1	0	0	1	1 231	2%	91%	2
Other patterns					6 434	9%	100%	
Total					70 721	100%		

Notes: the table includes observations with at least one additional loan (*Bank* = 1) between 2011 and 2015. All the patterns in the sample with one or two additional loans are represented on the table; 'other patterns' (9% of the observations) refer to observations with 3 to 5 additional loans.

Concerning the firms' state, 75.41% are *profitable* firms, 21.57% are *non-profitable* firms and 3.02% are in the *exit* state. Table IV shows the transition matrix observed for the database. Nearly 5% of firms were *new* firms in the previous year. *Non-profitable* firms post the higher death rate in the following year (4.95%), while this proportion is 2.5% among *profitable* firms.

TABLE IV
TRANSITION MATRIX FOR VARIABLE STATE

	State n			
State n-1	*Profitable*	Non-*profitable*	Exit	Total
New	25 082	6 391	1 731	33 204
	(75.54%)	(19.25%)	(5.21%)	(100%)
Profitable	426 125	94 436	13 373	533 934
	(79.81%)	(17.69%)	(2.5%)	(100%)
Non-*profitable*	55 668	44 141	5 196	105 005
	(53.01%)	(42.04%)	(4.95%)	(100%)
Total	506 875	144 968	20 300	672 143
	(75.41%)	(21.57%)	(3.02%)	(100%)

Notes: transition frequencies in two consecutive years observed in the sample. Row percentages in parenthesis.

4.2. Econometric model

The probability of change to state *j* is estimated by a multinomial logit model, as in (3.2.6), defining *profitable* as the pivot category and for *j* = *non-profitable, exit*:

$$\boldsymbol{x\beta}_j = \beta_{1j} + \beta_{2j} state_{t-1} + \beta_{3j} bank_{t-1} + \beta_{4j} sector\ depressed_{t-1} \qquad (4.2.1)$$

The variable $state_{t-1}$ is the firms' state in the previous year, assuming three possible outcomes: *new*, *profitable* and *non-profitable*; $bank_{t-1}$ is the binary variable assuming 1 if the firm got an additional loan from a bank in the previous year, and *sector depressed* assumes 1 if the firm operates in a depressed activity and 0 otherwise. The model in (4.2.1) is a saturated model, having 12 possible patterns for the regressors.

The set of controls may also include $bank_{t-1}$ interacted with $sector\_depressed_{t-1}$, to evaluate if the effect of bank assistance significantly varies for depressed activities.

4.3. Results and discussion

Table V presents the results of the MNL, considering two specifications for the model as in (4.2.1) estimated for the Portuguese data from 2011 to 2015.

In the baseline model (1), only *bank* and the state in the previous year (*profitable* and *non-profitable*, having *new* as the base category) are considered. All regressors are statistically significant (individually and jointly, as the Likelihood Ratio test points). *Bank* has a negative sign, meaning that having an additional bank loan reduced the log-odds ratio of being *non-profitable* and *exit* compared to *profitable*. The model also shows state dependence, as the coefficients for the state in the previous year are statistically significant. The log-odds of being *non-profitable* and *exit* in *t*, relative to being *profitable*, was lower if the firm was *profitable* in *t-1*, compared to being *new* in *t-1*. *Non-profitable* firms in *t-1* showed higher log-odds of being *non-profitable* and *exit* in *t*, relative to being *profitable*, than *new* firms in *t-1*.

TABLE V

ESTIMATION OUTPUTS FOR THE MULTINOMIAL LOGIT

Outcome	Regressor	Baseline (1)	Final (2)
Non-profitable	Bank $_{t-1}$	-0,172 (0,000)	-0,175 (0,000)
	Profitable $_{t-1}$	-0,136 (0,000)	-0,137 (0,000)
	Non-profitable $_{t-1}$	1,133 (0,000)	1,136 (0,000)
	Sector depressed $_{t-1}$		-0,017 (0,01)
	Sector depressed $_{t-1}$* Bank $_{t-1}$		0,004 (0,816)
	Const.	-1,343 (0,000)	-1,333 (0,000)
Exit	Bank $_{t-1}$	-0,073 (0,000)	-0,116 (0,000)
	Profitable $_{t-1}$	-0,787 (0,000)	-0,788 (0,000)
	Non-profitable $_{t-1}$	0,301 (0,000)	0,32 (0,000)
	Sector depressed $_{t-1}$		-0,112 (0,000)
	Sector depressed $_{t-1}$* Bank $_{t-1}$		0,073 (0,066)
	Const.	-2,663 (0,000)	-2,602 (0,000)
	Number of obs	672 143	672 143
	Pseudo R2	0,0361	0,0362
	Log likelihood	-420687,1250	-420660,0625
	LR test	31556,348	31610,496
	Prob > chi1	0,000	0,000
	Pearson's chi2 (sat)		1107,963
	Prob > chi0		0,000
	Hosmer-Lemeshow (G=3)		192,230
	Prob > chi2		0,000

Notes: p-values in parentheses. *Profitable* is the base outcome. Pearson's chi2 and Hosmer-Lemeshow tests are presented for the final model only.

The final model (2) includes *sector depressed*, alone and interacted with *bank*. Variable *sector depressed* is statistically significant and presents a negative coefficient for outcomes *non-profitable* and *exit,* meaning the log-odds of being *non-profitable* and *exit*, relative to being *profitable*, was lower for firms in depressed activities with no additional bank loans. As for the interaction between *bank* and *sector depressed*, it is not statistically significant for outcome 3 (*non-profitable*), but is positive and statistically significant at a 10% significance level for outcome 4 (*exit*), pointing to a lower effect of an additional bank loan over survival rates on depressed periods.

The average marginal effects over the probability of the three possible outcomes confirm the conclusions for the final model (Table VI). On average, an additional bank loan increased the probability of Portuguese firms being *profitable* by 2.7 percentage points (pp), and reduced the probability of being *non-profitable* by 2.7 pp; the effect on *exit* was 0.01 pp and was not statistically significant.

TABLE VI

AVERAGE MARGINAL EFFECTS FOR THE FINAL MODEL

Variable	Profitable	Non-profitable	Exit
Bank	.0275275	-.0265521	-.0009754
	(0.000)	(0.000)	(0.088)
Profitable n-1	.0422005	-.0151458	-.0270546
	(0.000)	(0.000)	(0.000)
Non-profitable n-1	-.2254401	.227217	-.0017769
	(0.000)	(0.000)	(0.205)
Sector depressed	.0048012	-.0019547	-.0028465
	(0.000)	(0.052)	(0.000)

Note: p-values in parentheses. As all the regressors are dummy variables, the average marginal effect for each variable corresponds to the average difference in the probability of each outcome considering that variable assumes 1 and the probability of each outcome considering that variable assumes 0, keeping the original value for all the other variables.

Given the regressor patterns for the final model, the Hosmer-Lemeshow test can only be calculated for G=3. The test rejects the null hypothesis that $\widehat{\boldsymbol{\beta}}$ are the true coefficients of the model. As for the Pearson's χ^2 for the saturated model, the null hypothesis is also rejected. This results point to a miscalibration of the model to predict the outcomes of some groups of observations (Table XVI in Appendix provides additional information on the regressor patterns for this sample and the Pearson's χ^2 test). As the purpose of this study is to access the effect of an additional bank loan for the average firm of the sample, and not to predict individual outcomes, these results do not imply dropping the previous conclusions; nevertheless, this may point to the existence of unobserved heterogeneity that explain firms' dynamics that limit the capacity for this model to predict individual outcomes.

In order to evaluate the final model for specific groups of firms, Table XV in Appendices provides additional outputs. Model (1) is restricted to firms in "Construction and real estate" activities, model (2) to large firms and model (3) to small firms. "Construction and real estate" activities and small firms were the most affected during the financial crisis; large firms, on the other hand, tend to behave differently from the remaining firms. The results show that in fact these subgroups presented a different dynamic. Concerning the effect of an additional bank loan, in particular, large firms did not show a statistically significant effect over the log-odds of *non-profitable* and *exit* over *profitable*, but for the remaining groups this effect was significant and similar to that observed for the final model. This is an interesting result, as it shows that large firms may behave differently from the remaining firms, and that additional bank loans do not seem to impact their recovery and exit rates. Nevertheless, these firms account for 0.6% of the sample and do not change the conclusions for the sample as a hole.

Additionally, considering the database before excluding firms with no credit relations with banks, nearly 40% of the observations concern to firms with no credit relations with the Portuguese financial system. Though these firms are excluded from the estimation, they are relevant to understand firms' dynamics and the difference between those firms who have access to additional lines of credit and those who do not. Model (4) provides and additional analysis by estimating the final model for all firms. Concerning the effect of an additional bank loan, the conclusions are also similar to the final model.

4.4. Markov chain matrices

The transition matrices of the Markov chain process (as defined in Figure 1) allows a more intuitive analysis of the results of the MNL obtained for the final model (2) as in Table V. The 1-step transition matrices of the process as in (3.1.4) are obtained with the expected probability for each outcome, given the state and the regressors' values in the previous step. The results are provided in Table VII.

For non-depressed activities, the effect of having an additional bank loan increased the probability of moving from *non-profitable* to *profitable* in 4.18pp, mostly compensated by a reduction in the probability of remaining *non-profitable* (Table VII-C). The p-values for the differences in the expected probabilities show these differences are statistically significant. The probability of *exit* is similar. The same situation was observed for depressed activities: the probability of moving from *non-profitable* to *profitable* increased by 3.9pp, compensated by a reduction in the probability of remaining *non-profitable* by 4pp. These differences are also statistically significant. The effect of an additional bank loan over *exit* was not statistically significant.

The same was observed for *profitable* firms: for these, an additional bank loan had a similar effect for non-depressed and depressed activities, and increased the probability of remaining *profitable* by 2.6pp, mostly compensated by a reduction in the probability of moving to *non-profitable*. The *exit* probability was slightly lower for firms with additional bank loans.

TABLE VII

1-STEP TRANSITION MATRICES FOR THE FINAL MODEL

Sector depressed = 0
(non-depressed activities)

(A) Bank=1

	New	Efficient	Non-effic.	Exit
New	**0,0000**	**0,7768** **(0,000)**	**0,1719** **(0,000)**	0,0513 (0,000)
Efficient	**0,0000**	**0,8177** **(0,000)**	**0,1578** **(0,000)**	0,0245 (0,000)
Non-effic.	**0,0000**	**0,5619** **(0,000)**	**0,3870** **(0,000)**	0,0511 (0,000)
Exit	0,0000	0,0000	0,0000	1,0000

(B) Bank=0

	New	Efficient	Non-effic.	Exit
New	**0,0000**	**0,7475** **(0,000)**	**0,1971** **(0,000)**	0,0554 (0,000)
Efficient	**0,0000**	**0,7913** **(0,000)**	**0,1820** **(0,000)**	0,0267 (0,000)
Non-effic.	**0,0000**	**0,5201** **(0,000)**	**0,4268** **(0,000)**	0,0531 (0,000)
Exit	0,0000	0,0000	0,0000	1,0000

(C) Difference (A)-(B)

	New	Efficient	Non-effic.	Exit
New	**0,0000**	**0,0293** **(0,000)**	**-0,0252** **(0,000)**	-0,0042 (0,002)
Efficient	**0,0000**	**0,0263** **(0,000)**	**-0,0242** **(0,000)**	-0,0021 (0,001)
Non-effic.	**0,0000**	**0,0418** **(0,000)**	**-0,0398** **(0,000)**	-0,0020 (0,074)
Exit	0,0000	0,0000	0,0000	0,0000

Sector depressed=1
(depressed activities)

(D) Bank=1

	New	Efficient	Non-effic.	Exit
New	**0,0000**	**0,7802** **(0,000)**	**0,1703** **(0,000)**	0,0495 (0,000)
Efficient	**0,0000**	**0,8201** **(0,000)**	**0,1562** **(0,000)**	0,0237 (0,000)
Non-effic.	**0,0000**	**0,5659** **(0,000)**	**0,3846** **(0,000)**	0,0495 (0,000)
Exit	0,0000	0,0000	0,0000	1,0000

(E) Bank=0

	New	Efficient	Non-effic.	Exit
New	**0,0000**	**0,7545** **(0,000)**	**0,1955** **(0,000)**	0,0500 (0,000)
Efficient	**0,0000**	**0,7961** **(0,000)**	**0,1799** **(0,000)**	0,0240 (0,000)
Non-effic.	**0,0000**	**0,5269** **(0,000)**	**0,4250** **(0,000)**	0,0481 (0,000)
Exit	0,0000	0,0000	0,0000	1,0000

(F) Difference (D)-(E)

	New	Efficient	Non-effic.	Exit
New	**0,0000**	**0,0256** **(0,000)**	**-0,0252** **(0,000)**	-0,0005 (0,357)
Efficient	**0,0000**	**0,0241** **(0,000)**	**-0,0237** **(0,000)**	-0,0003 (0,311)
Non-effic.	**0,0000**	**0,0390** **(0,000)**	**-0,0404** **(0,000)**	0,0014 (0,142)
Exit	0,0000	0,0000	0,0000	0,0000

Notes: 1-step transition matrices as in (3.1.4). Lines correspond to *state* in t-1 and columns correspond to *state* in t. Cells in bold correspond to matrix Q in the canonical form. The transition probabilities correspond to the expected probability for the outcome given the regressor pattern (*bank*, *sector depressed* and *state t-1*). P-values in parenthesis using the standard errors as in (3.4.4.) for the expected probabilities and (3.4.5) for the differences in the expected probabilities.

These results point that an additional loan to Portuguese *non-profitable* firms was not the cause for the high proportion of zombies in the economy, in the sense that the probability of remaining *non-profitable* significantly decreased with this variable.

The fundamental matrices of the process allows an analysis of the firms' dynamics, in particular the expected survival rates. As Table VIII shows, firms that were granted an additional loan had higher survival rates and were expected to live longer as *profitable* firms. The increase in the life expectancy for *non-profitable* firms in non-depressed activities with *Bank* = 1 was 3.01 years, most of this value arising from an increase in the expected number of steps as *profitable* (3.49 years, from 20.84 years to 24.33 years).

An increase in the survival rates for *non-profitable* firms with *Bank* = 1 was also observed for depressed activities, though substantially lower than the increase for non-depressing activities. *Non-profitable* firms with additional loans were expected to survive 0.89 years longer than *non-profitable* firms without additional bank loans, of which 1.91 years as *profitable* firms and less 1.02 years as *non-profitable* firms.

TABLE VIII

FUNDAMENTAL MATRICES AND EXPECTED NUMBER OF STEPS BEFORE ABSORPTION FOR THE FINAL MODEL

Sector depressed = 0 (non-depressed activities)

(A) Bank=1

	t	New	Efficient	Non-effic.
New	**32,4670**	1,0000	24,8011	6,6659
Efficient	**33,3756**	0,0000	26,5419	6,8337
Non-effic.	**32,2266**	0,0000	24,3308	7,8958

(B) Bank=0

	t	New	Efficient	Non-effic.
New	**29,3815**	1,0000	21,2804	7,1011
Efficient	**30,2677**	0,0000	22,9730	7,2947
Non-effic.	**29,2085**	0,0000	20,8449	8,3637

(C) Difference (A)-(B)

	t	New	Efficient	Non-effic.
New	**3,0855**	0,0000	3,5207	-0,4352
Efficient	**3,1079**	0,0000	3,5689	-0,4610
Non-effic.	**3,0181**	0,0000	3,4860	-0,4679

Sector depressed=1 (depressed activities)

(D) Bank=1

	t	New	Efficient	Non-effic.
New	**33,7021**	1,0000	25,8611	6,8410
Efficient	**34,6135**	0,0000	27,6063	7,0072
Non-effic.	**33,4560**	0,0000	25,3871	8,0689

(E) Bank=0

	t	New	Efficient	Non-effic.
New	**32,7491**	1,0000	23,9228	7,8263
Efficient	**33,6411**	0,0000	25,6228	8,0183
Non-effic.	**32,5671**	0,0000	23,4802	9,0869

(F) Difference (D)-(E)

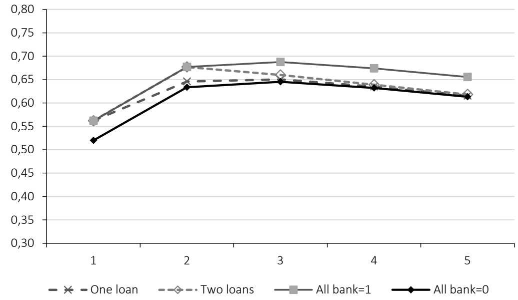
	t	New	Efficient	Non-effic.
New	**0,9530**	0,0000	1,9383	-0,9853
Efficient	**0,9724**	0,0000	1,9835	-1,0111
Non-effic.	**0,8889**	0,0000	1,9069	-1,0180

Note: t (column in bold) gives the expected number of steps before the process is absorbed given it starts in the transient state S_i, as in (3.1.13). The cells not in bold correspond to the fundamental matrix *N* as in (3.1.11).

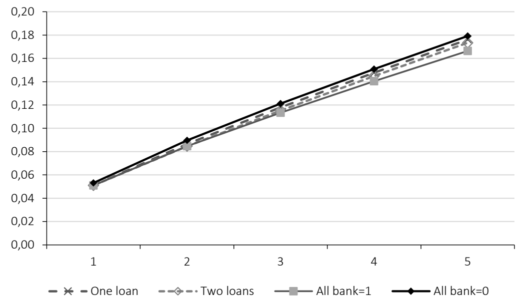
The fundamental matrix of the process provides the expected number of periods before absorption assuming fixed regressors. Nevertheless, it is unlikely that a firm receives an additional loan every year (see Table II). To evaluate the effect of a single additional loan and two consecutive additional loans on the firms' probability of moving from *non-profitable* to *profitable* after 5 years, the 5-step transition matrix of the process can be obtained by multiplying the conditional 1-step transition matrices according to the pattern to be analyzed. *One loan* refers to the case where a firm gets a single additional bank loan in T=1; in this case, *Bank* assumes 1 in T=1 and 0 in T=2, …, 5. *Two loans* refers to the case where a firm receives additional loans in two consecutive years; *Bank* assumes 0 in the remaining periods. These situations are the most common patterns in the database, and are compared with the extreme cases where firms have their loans renewed every year (*All Bank*=1) and no additional loans (*All Bank*=0). The results are presented in Figure II. The results point that an additional loan in the first step or in the two initial steps increased the probability of moving to *profitable* in the short-run; nevertheless, after 5 years the situation was similar to not ever getting an additional loan. Concerning death rates, the conclusions are similar.

Non-depressed activities

(a) Non-*profitable* to *profitable*

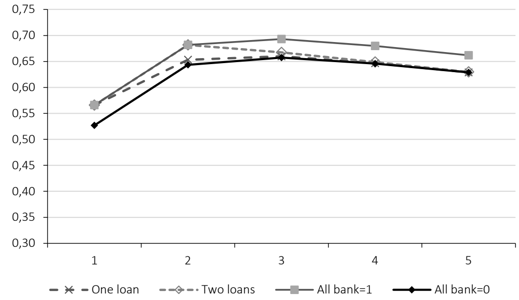


(b) Non-*profitable* to exit



Depressed activities

(c) Non-*profitable* to *profitable*



(d) Non-*profitable* to exit

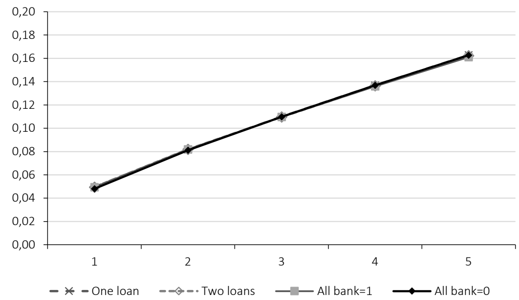


FIGURE 2 – Transition probabilities up to 5 years for the final model

5. CONCLUDING REMARKS

The literature on non-productive (zombie) firms assumes, in general, a causal relation between loans 'evergreening' to distressed firms and the survival of those inefficient firms. Zombie firms are often identified as financially stressed firms assisted by banks, being the financial distress, and not the operational efficiency or productivity, which defines the firms' state as a zombie. This comes from the hypothesis that turbulence in the financial markets during the last decade was one of the causes for the prevalence of a high proportion of zombies and, in consequence, for the productivity slowdown observed in most of the developed economies.

In this study firms are categorized according to their operational performance. *Non-profitable* firms are those who are not able to obtain an operating profit, for reasons that are not related to the firms' growth. This definition is independent from firms' funding decisions and financial pressure to meet obligations towards their creditors.

The role of banks' behavior towards *non-profitable* firms is analyzed through the impact of an additional loan over the probability for these firms to recover, stay *non-profitable* or *exit* the market. Estimating this model for Portuguese firms with a credit relation with banks for the 2011-2015 period, the results obtained point to a positive effect of these additional loans over firms' dynamics, which increased their probability to become *profitable* firms while reducing the probability to remain *non-profitable*. This was observed even for depressed activities, which was the case for most firms during the2011-12 period, when banks were under a higher pressure to meet additional capital requirements and, therefore, were facing higher incentives to avoid recognizing losses in their balance-sheets.

This model was estimated for specific sub-groups ("Construction and real estate", small and large firms), and also for all firms regardless of having a credit relation with the banking system. These groups may have behaved differently during this period. The estimation of the model for these sub-groups show that for all these cases, except for large firms, the effect of an additional bank loan is different in magnitude but the same conclusions can be taken as for the final model. Concerning large firms, the results show that an additional loan does not seem to impact the firms' 'zombieness', as the coefficients for these variable are not statistically significant.

The model presented in this study provides an interesting contribution to the analysis of the effects of the financial crisis over non-financial corporations, by proposing a different framework based on individual transitions across efficiency states. Besides offering a new perspective to analyze the zombie phenomenon, the results seem to contradict the hypothesis that stressed banks impact negatively the aggregate productivity by allowing inefficient firms to survive as zombies.

The results point to significant differences in the recovery rates of Portuguese *non-profitable* firms that were granted additional bank loans during the 2011-2015 period, but the results also suggest the existence of unobserved heterogeneity that may be limiting the capacity for this model to predict individual outcomes. Despite lacking references for a valid endogeneity test for the MNL, the Rivers and Vuong 2-step procedure for the probit model was used to test endogeneity in *bank*, using the total amount of the loan in *t*-1 as instrumental variable. The test does not reject the null hypothesis of exogeneity of *bank* at 1% significance level, but reservations about the validity of the test and the weak instrumental variable available do not allow this test to be conclusive. Further developments to this analysis could include controlling for the unobserved heterogeneity, either by including additional explanatory variables or by IV estimation.

REFERENCES

ADALET MCGOWAN, M., D. Andrews, V. Millot (2017), *The Walking Dead? Zombie firms and productivity performance in OECD countries*, OECD Economic Department working papers, nº1372.

ADLER, G., R. Duval, D. Furceri, S. K. Çelik, K. Koloskova, M. Poplawski-Ribeiro (2017), *Gone with the Headwinds: Global Productivity*, IMF Staff Discussion Note nº 17/04, International Monetary Fund.

ALBUQUERQUE, R., H. Hopenhayn (2004), *Optimal Lending Contracts and Firm Dynamics*, The Review of Economic Studies, vol. 71, nº 2, pp 285-315.

ALEXANDRE, F. (Coord.) (2017), "Investimento empresarial e o crescimento da economia portuguesa", Fundação Calouste Gulbenkian.

AZEVEDO, N., M. Mateus, A. Pina (2018), *Bank credit allocation and productivity: stylized facts for Portugal*, Banco de Portugal Working Paper nº25/2018.

BLATTNER, L., L. Farinha, F. Rebelo (2018), *When losses turn into loans: the cost of undercapitalized banks*, Banco de Portugal Working Paper.

CABALLERO, R., T. Hoshi, A. K. Kashyap (2008), *Zombie lending and depressed restructuring in Japan*, The American Economic Review, nº98, vol. 5, pp. 1943-77.

COOLEY, T. F., V. Quadrini (2002), *Financial Markets and Firm Dynamics*, The American Economic Review, vol.91, nº 5, pp 1286-1340.

CARLETTI, E., A. Leonello (2016), *Regulatory reforms in the European Banking Sector*, in "The Palgrave Handbook of European Banking", T. Beck and B. Casu (ed.), Palgrave Macmillan.

EUROPEAN COMMISSION (2003), *Commission Recommendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises* (2003/361/EC).

FAGERLAND, M. W., D. W. Hosmer (2012), *A generalized Hosmer–Lemeshow goodness-of-fit test for multinomial logistic regression models*, The Stata Journal 12, nº 3, pp. 447–453.

FAGERLAND, M. W., D. W. Hosmer, A. M. Bofin. (2008), *Multinomial goodness-of-fit tests for logistic regression models*, Statistics in Medicine 27: 4238–4253.

GOUVEIA, A.F., C. Osterhold (2018), *Fear the walking dead: zombie firms, spillovers and exit barriers*, OECD Productivity working papers, June 2018, nº13.

GREENE, W. H. (2003), "Econometric analysis", fifth edition, Prentice Hall, New Jersey.

GRINSTEAD, C. M. and J. L. SNELL (2004), "Introduction to probability", American Statistical Society.

HALDANE, A., T Neumann (2016), *Complexity in regulation, in* "The Palgrave Handbook of European Banking", T. Beck and B. Casu (ed.), Palgrave Macmillan.

HOPENHAYN, H. A. (1992), *Entry, exit and firm dynamics in the long run*, Econometrica, nº60, vol. 5, pp 1127-1150.

HOSMER, D. W., S. Lemeshow (2000), "Applied logistic regression", second edition, John Wiley & Sons.

LONG, J. (2009), *Group comparisons in logit and probit using predicted probabilities*, Indiana University, Working Paper, June 25, 2009.

PEEK, J., E. S. Rosengren (2005), *Unnatural selection: perverse incentives and the misallocation of credit in Japan*, American Economic Review, nº 95, vol. 4, pp 1144-1166.

PETRELLA, G., A. Resti (2016), *The interplay between Banks and Markets: supervisory stress test results and investor reaction, in* "The Palgrave Handbook of European Banking", T. Beck and B. Casu (ed.), Palgrave Macmillan.

RAJARSHI, M. B., (2013), "Statistical inference for discrete time series stochastic processes", Springer India.

WOOLDRIDGE, J. M. (2010), "Econometric analysis of cross-section and panel data", second edition, MIT Press, Cambridge.

APPENDICES

TABLE IX

DISTRIBUTION OF VARIABLE ADDITIONAL LOAN (IN EUROS AND PERCENTAGE OF ASSETS)

	Additional loan (€)	*Additional loan / assets*
Obs	672 143	672 143
Mean	44 286,029	0,059
Std. Dev.	1 126 303,107	19,265
P1	0,000	0,000
P5	0,000	0,000
P10	0,000	0,000
P25	0,000	0,000
P50	0,000	0,000
P75	987,000	0,005
P90	28 505,000	0,084
P95	85 396,000	0,161
P99	580 625,000	0,385

TABLE X

PROPORTION OF FIRMS WITH ADDITIONAL LOANS FOR ALTERNATIVE THRESHOLDS

			Additional loan					Difference from the adopted definition	
sector depressed	state t-1	state t	> 500 thd€ or 10% of assets	> 100 thd€ or 10% of assets	**> 100 thd€ or 5% of assets**	other cases / no loans	Total	> 500 thd€ or 10% of assets	> 100 thd€ or 10% of assets
(1)	(2)	(3)	(4)	(5)	**(6)**	(7)	(8)	(9) = (4) - (6)	(10) = (5) - (6)
No	New	Profitable	18%	19%	**24%**	76%	100%	-6%	-6%
No	Profitable	Profitable	11%	13%	**19%**	81%	100%	-7%	-5%
No	Non-profit.	Profitable	6%	7%	**10%**	90%	100%	-4%	-3%
No	New	Non-profit.	14%	15%	**19%**	81%	100%	-4%	-4%
No	Profitable	Non-profit.	9%	10%	**15%**	85%	100%	-5%	-5%
No	Non-profit.	Non-profit.	5%	6%	**10%**	90%	100%	-4%	-3%
No	New	Exit	19%	19%	**24%**	76%	100%	-5%	-5%
No	Profitable	Exit	10%	11%	**14%**	86%	100%	-5%	-4%
No	Non-profit.	Exit	8%	9%	**11%**	89%	100%	-3%	-2%
Yes	New	Profitable	15%	15%	**20%**	80%	100%	-5%	-5%
Yes	Profitable	Profitable	9%	10%	**15%**	85%	100%	-6%	-5%
Yes	Non-profit.	Profitable	5%	6%	**10%**	90%	100%	-4%	-3%
Yes	New	Non-profit.	12%	12%	**16%**	84%	100%	-4%	-4%
Yes	Profitable	Non-profit.	8%	9%	**13%**	87%	100%	-5%	-4%
Yes	Non-profit.	Non-profit.	5%	6%	**9%**	91%	100%	-4%	-3%
Yes	New	Exit	18%	18%	**21%**	79%	100%	-4%	-4%
Yes	Profitable	Exit	10%	11%	**14%**	86%	100%	-4%	-4%
Yes	Non-profit.	Exit	6%	7%	**9%**	91%	100%	-3%	-2%
		Total	9%	11%	**15%**	85%	100%	-6%	-5%

TABLE XI

TRANSITION MATRICES FOR VARIABLE STATE CONDITIONAL ON HAVING AN ADDITIONAL BANK LOAN FOR ALTERNATIVE THRESHOLDS

Non-depressed activities | Depressed activities

Additional loan: > €500.000 or 10% of assets

	State n					State n			
State n-1	Profitable	Non-profit.	Exit	Total	State n-1	Profitable	Non-profit.	Exit	Total
New	74%	18%	7%	100%	New	83%	13%	4%	100%
Profitable	48%	46%	6%	100%	Profitable	59%	34%	7%	100%
Non-profit.	78%	19%	3%	100%	Non-profit.	86%	12%	2%	100%

Additional loan: > €100.000 or 10% of assets

	State n					State n			
State n-1	Profitable	Non-profit.	Exit	Total	State n-1	Profitable	Non-profit.	Exit	Total
New	75%	18%	7%	100%	New	83%	12%	4%	100%
Profitable	48%	46%	6%	100%	Profitable	59%	35%	6%	100%
Non-profit.	79%	18%	3%	100%	Non-profit.	87%	11%	2%	100%

Additional loan: > €100.000 or 5% of assets

	State n					State n			
State n-1	Profitable	Non-profit.	Exit	Total	State n-1	Profitable	Non-profit.	Exit	Total
New	75%	18%	7%	100%	New	83%	12%	4%	100%
Profitable	50%	45%	5%	100%	Profitable	59%	35%	5%	100%
Non-profit.	79%	18%	3%	100%	Non-profit.	87%	11%	2%	100%

TABLE XII

INDUSTRIES

Industry	Name	NACE 2 digit codes
1	Agriculture & Mining	01 to 09
2	Manufacturing	10 to 33
3	Utilities	35 to 39 (Electricity), 49 to 53 (Transport), 58 to 63 (Communications)
4	Construction & Real Estate	41 to 43 (Construction), 68 (Real estate)
5	Trade & Accommodation and Food	45 to 47 (Trade), 55 to 56 (A&F)
6	Other services	69 to 96

TABLE XIII

LIST OF VARIABLES AND DEFINITIONS

Variable	Reference	Definition
Zombie firm (interest rate definition)	Caballero et al. (2008) Alexandre (2017)	Firms with a cost of debt below the theoretical interest rate
Zombie firm (interest coverage ratio definition)	Adalet McGowan et al. (2017) Gouveia and Osterhold (2018) Azevedo et al. (2018)	Firms with EBITDA over interest expenses ratio below 1 for 3 consecutive years and more than 10 years of activity
New firm	This study	Firms with less than 3 years of activity
Profitable firm	This study	Firms with 3 or more years of activity and one of the following conditions: (i) positive operating result or (ii) sales' annual growth rate above 2%
Non-profitable firm	This study	Firms with 3 or more years of activity and negative operating result and sales' annual growth rate below 2%
Exit firm	This study	Firms that ceased activity during the economic year.
Bank	This study	Dichotomous variable assuming 1 if the firm recorded an increase in the loans from a bank already a creditor in the previous year above €100.000 or above 5% of its total assets, and 0 otherwise.
Sector depressed	This study	Dichotomous variable assuming 1 if the firm is in a 2-digit level activity (NACE Rev.2) that recorded a negative sales' annual growth rate.

TABLE XIV

OBSERVATIONS WITH SECTOR DEPRESSED=1 BY INDUSTRY AND YEAR

	2011	2012	2013	2014	2015	Total
Frequencies						
Agriculture and mining	4 026	966	4 271	3 891	677	13 831
Manufacturing	9 513	12 351	4 337	5 695	2 119	34 015
Utilities	4 027	10 857	4 153	792	3 452	23 281
Construction and real estate	22 136	18 812	13 666	6 738	910	62 262
Trade, accomm. and food	58 331	53 804	45 109	0	15 089	172 333
Other services	31 212	27 188	24 110	4 306	10 193	97 009
Total	**129 245**	**123 978**	**95 646**	**21 422**	**32 440**	**402 731**
Percentage						
Agriculture and mining	1,0%	0,2%	1,1%	1,0%	0,2%	3,4%
Manufacturing	2,4%	3,1%	1,1%	1,4%	0,5%	8,4%
Utilities	1,0%	2,7%	1,0%	0,2%	0,9%	5,8%
Construction and real estate	5,5%	4,7%	3,4%	1,7%	0,2%	15,5%
Trade, accomm. and food	14,5%	13,4%	11,2%	0,0%	3,7%	42,8%
Other services	7,8%	6,8%	6,0%	1,1%	2,5%	24,1%
Total	**32,1%**	**30,8%**	**23,7%**	**5,3%**	**8,1%**	**100,0%**

TABLE XV

ESTIMATION OUTPUTS FOR ADDITIONAL SPECIFICATIONS

Outcome	Regressor	Construction and real estate (1)	Large firms (2)	Small firms (3)	All firms (4)
Non-profitable	Bank $_{t-1}$	-0,16	0,099	-0,159	-0,245
		(0,000)	(0,521)	(0,000)	(0,000)
	Profitable $_{t-1}$	0,083	1,231	0,059	-0,174
		(0,059)	(0,226)	(0,284)	(0,000)
	Non-profitable $_{t-1}$	0,922	3,18	1,691	1,002
		(0,000)	(0,002)	(0,000)	(0,000)
	Sector depressed $_{t-1}$	-0,094	-0,186	-0,034	0,021
		(0,000)	(0,137)	(0,051)	(0,000)
	Sector depressed $_{t-1}$* Bank $_{t-1}$	-0,067	0,156	-0,03	-0,02
		(0,164)	(0,469)	(0,457)	(0,224)
	Const	-1,399	-3,435	-1,957	-1,209
		(0,000)	(0,001)	(0,000)	(0,000)
Exit	Bank $_{t-1}$	-0,026	0,886	-0,085	-0,217
		(0,717)	(0,057)	(0,162)	(0,000)
	Profitable $_{t-1}$	-0,621	-2,403	-0,926	-0,948
		(0,000)	(0,000)	(0,000)	(0,000)
	Non-profitable $_{t-1}$	0,159	-0,954	0,477	0,005
		(0,046)	(0,124)	(0,000)	(0,764)
	Sector depressed $_{t-1}$	-0,205	0,778	-0,264	-0,068
		(0,000)	(0,048)	(0,000)	(0,000)
	Sector depressed $_{t-1}$* Bank $_{t-1}$	-0,021	-1,471	0,193	0,044
		(0,829)	(0,03)	(0,027)	(0,233)
	Const	-2,4	-2,62	-2,682	-2,256
		(0,000)	(0,000)	(0,000)	(0,000)
	Number of obs	88 379	3 933	136 503	1 112 084
	Pseudo R2	0,0144	0,0758	0,0495	0,0336
	Log likelihood	-58145,5078	-1613,3492	-69854,6250	-742851,5625
	LR test	1694,216	264,712	7278,148	51634,246
	Prob > chi1	0,000	0,000	0,000	0,000

Notes: p-values in parentheses. *Profitable* is the base outcome.

TABLE XVI

PEARSONS' χ^2 TEST – OBSERVED AND EXPECTED OBSERVATIONS BY GROUP

Regressor pattern				Outcome probabilities			Profitable		Non-profitable		Exit		Total			
State n-1	Bank	Sector depressed	Obs.	Profitable	Non-profitable	Exit	Observed	Expected	Observed	Expected	Observed	Expected	Observed	Expected	Pearsons' χ^2 statistic	p(chi>0)
1	0	0	12 735	0,75	0,20	0,06	9 881	9 519	2 249	2 510	605	706	12 735	12 735	55	0,000
1	0	1	15 545	0,75	0,20	0,05	11 586	11 729	3 149	3 039	810	777	15 545	15 545	7	0,029
2	0	0	206 533	0,79	0,18	0,03	161 452	163 436	39 376	37 589	5 705	5 508	206 533	206 533	116	0,000
2	0	1	237 165	0,80	0,18	0,02	190 581	188 803	41 108	42 674	5 476	5 687	237 165	237 165	82	0,000
3	0	0	24 409	0,52	0,43	0,05	14 317	12 694	8 892	10 419	1 200	1 296	24 409	24 409	438	0,000
3	0	1	67 203	0,53	0,43	0,05	33 775	35 410	30 018	28 562	3 410	3 231	67 203	67 203	160	0,000
1	1	0	2 359	0,78	0,17	0,05	1 762	1 833	460	405	137	121	2 359	2 359	12	0,002
1	1	1	2 565	0,78	0,17	0,05	1 853	2 001	533	437	179	127	2 565	2 565	53	0,000
2	1	0	48 325	0,82	0,16	0,02	39 285	39 513	7 834	7 627	1 206	1 185	48 325	48 325	7	0,026
2	1	1	41 911	0,82	0,16	0,02	34 807	34 373	6 118	6 546	986	992	41 911	41 911	34	0,000
3	1	0	3 891	0,56	0,39	0,05	2 485	2 186	1 244	1 506	162	199	3 891	3 891	93	0,000
3	1	1	9 502	0,57	0,38	0,05	5 091	5 377	3 987	3 655	424	470	9 502	9 502	50	0,000
Total			672 143				506 875	506 875	144 968	144 968	20 300	20 300	672 143	672 143	1 108	0,000

Note: observed and expected number of observations by regressor pattern. The outcome probabilities provide the expected probability of the outcomes given the regressors patterns. The expected observations are given by the product of the observations in the regressor pattern and the expected probability of the outcome. Pearsons' χ^2 statistic is calculated as in (3.5.2). P(chi>0) is the p-value, considering 24 degrees of freedom for the total statistic and 2 degrees of freedom for the partial statistics