Factors Influencing the European Bank's Probability of Default: An Application of SYMBOL Methodology

Purificación Parrado-Martínez[a,*], Pilar Gómez-Fernández-Aguado[a], Antonio Partal-Ureña[a]

[a] Department of Financial Economics and Accounting, Universidad de Jaén, ES-23071 Jaén, Spain

Abstract

This paper analyses European banks' probability of default (PD) by estimating a new measure that is based on the SYstemic Model of Bank Originated Losses (SYMBOL). First, we calculate the individual PD of a sample of European credit institutions during the period of 2011-2016. Then, dynamic panel data models are estimated to analyse the influence of several bank-specific and macroeconomic variables on the PD. We conclude that capital adequacy, liquidity, asset quality and profitability indicators influence the European banks' PD. The macroeconomic scenario, the industry concentration and the size of banks also appear to have an impact on their risk.

Keywords: Probability of default; Basel regulatory framework; CAMEL indicators; SYMBOL; Financial stability.

JEL classification: G21; G28; C15; C23

*Corresponding author at: Department of Financial Economics and Accounting, Universidad de Jaén, Paraje Las Lagunillas, s/n, 23071 Jaén, Spain. Tel.: +34 953212611. E-mail address: pparrado@ujaen.es (P. Parrado-Martínez).

Acknowledgments: We are grateful to the participants of the ACEDE 2018 XXVII Conference, the 23rd EBES Conference, and the II Workshop of the ACEDE's Financial Economics Section (2017). Special thanks are due to the anonymous referee and the editor (Jonathan Batten) for their guidance and very constructive comments.
Funding: This work was supported by the *University of Cantabria Foundation for Education and Research in the Financial Sector (UCEIF)* and the *Santander Financial Institute* (SANFI) [2016 Research Grant]. The usual disclaimer applies.

Factors Influencing the European Bank's Probability of Default: An Application of SYMBOL Methodology

1. Introduction

The consequences caused by the financial crisis of 2008 triggered a serious and complex situation of instability that led us to consider the collapse of the euro. The action of the European Central Bank (ECB) supporting the single currency in the summer of 2012, as well as the decisions taken to build a Banking Union[1] (BU) and get a more transparent, unified and secure banking sector, normalized to some extent the functioning of financial markets. The development of the BU is based on a unique normative code that will improve the effectiveness of the regulation, supervision and governance of the financial sector to favour improved financial stability. It includes rules on capital requirements[2], restructuring and resolution processes[3], and a harmonized set of national deposit guarantee systems[4] that aims to strengthen the financial safety net and will give way to the future European Deposit Insurance Scheme (EDIS).

The European Commission has assessed the effect of these regulatory proposals on bank losses and government contingent liabilities (European Commission, 2012, 2014 and 2016) by carrying out quantitative impact studies based on the SYMBOL (Systemic Model of Bank Originated Losses) Model. This microsimulation model estimates bank losses using the Basel risk assessment framework based on bank balance and regulatory data. It allows for evaluating the individual effects of regulatory measures (higher capital requirements, reinforced deposit insurance and the introduction of resolution funds) and the most effective

---

[1] European Central Bank, 2012. Financial Integration in Europe, Frankfurt.
[2] It is a regulation on capital requirements (CRD IV and CRR, of the European Parliament and of the Council, 2013), which adapts the Basel III agreement to EU legislation and improves the capital and liquidity requirements of financial institutions to deal with unexpected losses.
[3] Directive 2014/59/EU of the European Parliament and of the Council of 15 May 2014 establishes a framework for the recovery and resolution of credit institutions and investment firms.
[4] Directive 2014/49/EU of the European Parliament and of the Council of 16 April 2014 on deposit guarantee schemes.

combinations (Pagano et al., 2012). Likewise, it makes it possible to quantify the impact of a banking crisis on public accounts by considering the tools of the banking security network, which is fundamental for the process of the legislative implementation of the BU. The European Commission carried out a scientific review of the SYMBOL model to evaluate the validity of the methodological approach, concluding that the conceptual framework is based on innovative theoretical knowledge, and the methodology follows solid scientific principles[5] (Hordijk and Kancs, 2018). In this sense, the report suggests the convenience of exploring new extensions of the model to expand it or integrate it with other dynamic models.

While the process towards the consolidation of the BU is contributing to strengthened financial stability, works continue to be oriented towards reducing the risk and reinforcing the resilience of the banking sector[6]. The analysis of financial stability requires the systematic monitoring of all risk sources and vulnerabilities, considering the extent to which shocks can be absorbed by the financial system. The use of risk methodologies based on regulatory proposals to assess the soundness of the banking sector is key to a coherent legislative development with the objective of promoting financial stability.

In this sense, our work proposed a new risk measure (the probability of default) derived from the SYMBOL methodology. In addition, based on this measure, we examine which bank-specific and macroeconomic factors have had an impact on the default risk of the European banking sector after the financial crisis. The probability of bank default is defined as the probability that portfolio losses exceed total capital requirements, which are calculated following the Basel regulatory approach. From this perspective, the probability that banks default reflects the likelihood that losses will fall into the tail of the loss distribution ("tail

[5] The SYMBOL review was carried out by an external Scientific Review Panel closely following Guidelines for the review of models used in support of E.U. policies (European Commission, 2018). The objective of the review was verifying and consolidating the scientific credibility of the model and identifying most relevant areas for a future model development (Hordijk and Kancs, 2018).

[6] Roadmap of the European Council. Bratislava, 16 September 2016.

risk"). This risk occurs with a low probability, but if it occurs, it causes serious banking losses that endanger financial stability.

Considering a representative sample of the European listed banks during the period 2011-2016, we empirically estimate the banks' probabilities of defaults. The Generalized Method of Moments (GMM) estimator for dynamic panel models is used to examine the influence of several factors related to the bank business model (CAMEL indicators) and the macroeconomic context on the PD of European banks. These analyses allow for the identification of the bank distress from a micro-macro perspective. To the best of our knowledge, our work is the first to examine the determinants of European listed bank risk after the global financial crisis, using the SYMBOL methodology to proxy the banks' PD.

The rest of the paper is structured as follows. Section 2 presents the theoretical background. Section 3 describes the variables and methodology employed. Section 4 presents and discusses the results. Finally, Section 5 concludes the study.

2. Theoretical background

Our work contributes to two different strands of the literature: research on new measures for the regulation and supervision of the banking sector, and research on the factors influencing the banking risk in Europe.

2.1. The SYMBOL Methodology

The European Commission's SYMBOL model was developed by the Joint Research Centre in cooperation with the European Commission's former Directorate General for Internal Market and Services and academics (De Lisa et al., 2011). The model estimates the distribution of losses in a banking system, such as for a country or a set of financial institutions that share common characteristics. Losses are generated via Monte Carlo simulations using the Basel Internal Rating Based (IRB) function.

The initial aim of the model was to determine the optimal target level of equity for the Italian Deposit Guarantee System under different scenarios of Loss Given Default (LGD) and the correlation between banks' assets (De Lisa et al., 2011). The European Commission (2012) applied this methodology to assess the legacy needs of the National Deposit Guarantee Systems and based on these results the current size of the funds of each country was established at 0.8% of covered deposits. For the future development of EDIS, the Commission continues working with the results of the stress tests obtained with the SYMBOL model (European Commission, 2016). In addition, it has made it possible to evaluate the impact of the financial reforms adopted by the EU included in the Economic Review of the Financial Regulation Agenda. Specifically, these include the implications of the new capital requirements of Basel III (Marchesi et al., 2012; Pagano et al., 2012; European Commission, 2012 and 2014; Cannas et al., 2013), and the incorporation of the Bank Recovery and Resolution Directive (BRRD) and the Single Resolution Mechanism Regulation (Cariboni et al., 2015; Galliani and Zedda, 2015; Benczur, 2017).

Previous works developed with the SYMBOL methodology have generally used the final output of the model, which is the distribution of losses in the banking system. However, the process allows for obtaining an intermediate output, which is the probability of individual bank default (PD Bank) from the data of public balances and regulatory capital (De Lisa et al., 2011). PD Bank is defined as the probability that unexpected losses associated with the portfolio's debtors (implied obligor probability of default, IOPD) exceed bank capital (represented by the sum of regulatory capital requirements and any excess capital). This measure can be considered as a new bank risk proxy from the Basel regulatory framework that complement the traditional measures based on accounting and market information (Gómez et al., 2018).

2.2. Determinants of European bank risk

Previous studies on the determinants of bank risk have considered accounting information by mainly applying two proxies for bank risk: the NPL ratio and the Z-score. The NPL ratio (non-performing loans[7] to total gross loans) or changes in the ratio have been used as a measurement of the strength of a bank (e.g., Berger and DeYoung, 1997; Salas and Saurina, 2002; Ahmad and Arrif, 2007; Delis and Kouretas, 2011; Festic, Kavler and Repina, 2011) since they reflect the quality of a loan portfolio. Similarly, the Z-score has been widely used to measure bank risk (e.g., Demirgüç-Kunt and Huizinga, 2010; Laeven and Levine, 2009; Agoraki, Delis and Pasiouras, 2011; Lepetit and Strobel, 2015; Khan et al., 2016; Giordana and Schumacher, 2017), and a higher Z-score indicates that a bank is farther from default (Delis and Staikouras, 2011).

Furthermore, research using market information as a complement to accounting indicators has been based on Merton's (1974) approach to model the risk of credit defaults. Some studies have used credit risk spread CDs (Alter and Schüler, 2012; Chiaramonte and Casu, 2013; Annaert et al., 2013; Samaniego-Medina et al., 2016; Drago et. al. 2017) or credit ratings (Calomiris, 2009; Hau et al., 2012; Wang, 2017).

In the European context, many studies on the analysis of bank risk and bank distress has emerged in the last decade, such as Haq and Heany (2012), Betz et al. (2014), Fernandes et al. (2016), Jabra et al. (2017), and Constantine et al, (2018), among others.

Our work contributes to this previous research on the determinants of bank risk in Europe and supports the applicability of the SYMBOL methodology. Our main contributions are the following. First, we empirically estimated the PD for a sample of listed European banks. The majority of these banks occupy important positions in their country and the world rankings. Second, our period of analysis comprises the post-crisis years. This period is characterized by relevant regulatory and supervisory changes in response to the consequences

[7] Impaired loans

of the global financial crisis. Third, different from previous research that used the NPL ratio or Z-score as proxies for bank credit risk (Baselga et al, 2015; Jabra et al., 2017; among others), we study bank risk from a wider perspective by estimating a new variable that takes into account the new focus of regulation and considers the unexpected losses.

3. Variables and methodological aspects

Dynamic panel data regressions are estimated to explore the main factors influencing the individual probabilities of default (PD) of European credit institutions during the period of 2011-2016.

3.1. Probability of default (PD) by the European Commission's SYMBOL model

SYMBOL model estimates the losses associated with banking defaults using the Basel regulatory framework for capital. The model focuses on the tail risk by considering default when bank losses exceed the regulatory capital available. The modelling rests on the following assumptions (European Commission, 2016): 1) the Basel III regulatory model for credit risk is correct, 2) banks report risks accurately and in agreement with this model, and 3) all risks in the bank can be represented as a single portfolio of credit risks[8]. The model allows one to consider banking contagion channels as sources of financial instability (Campolongo et al., 2010). Individual banks' losses are generated via Monte Carlo simulation using the Basel FIRB loss distribution function (BCBS, 2006). Simulated losses are based on an estimation of the average default probability of the portfolio of assets of any individual bank ($IOPD_i$), which is derived from the bank's minimum capital requirement and total assets. The model is processed in four methodological steps (De Lisa et al., 2011; Cariboni et al., 2015; Muresano and Pagano, 2016; Benczur et al., 2017; Hordijk and Kancs, 2018). The first three stages are necessary to estimate the individual probability of the default of banks (PD

[8] This representation does not indicate that other risks are not considered but simply that they can be "mapped" in credit risk terms and modelled using the same framework.

bank), and the last stage determines the aggregate loss distribution of the system as a whole. Below, we describe the methodological considerations to estimate a bank's PD.

STEP 1: Estimation of the Implied Obligor Probability of Default of the portfolio of each individual bank (IOPD_i).

The IOPD_i is obtained from the Basel IRB formula to establish the minimum capital requirements for credit risk (*FIRB approach)*. For each exposure *l* of bank *i*, the IRB formula establishes the capital requirement $CR_{i,l}$ to cover the unexpected losses in the time horizon of one year at a 99.9% confidence level. Using publicly available data on capital requirements, total assets, and the regulatory values for the other parameters, the IOPD_i is obtained according to equation [1].

$$CR_{i,l}(PD_{i,l})=\left[LGD\cdot N\left(\sqrt{\frac{1}{1-R(PD_{i,l})}}\cdot N^{-1}(PD_{i,l})+\sqrt{\frac{R(PD_{i,l})}{1-R(PD_{i,l})}}\cdot N^{-1}(0.999)\right)-PD_{i,l}\cdot LGD\right]\cdot M(PD_{i,l}) \qquad [1]$$

where:

- $PD_{i,l}$ is the default probability of exposure *l*,
- *R* is the correlation among the exposures in the portfolio, which is defined as

$$R(PD_{i,l}) = 0.12\cdot\frac{1-e^{-50\cdot PD_{i,l}}}{1-e^{-50}}+0.24\cdot\left(1-\frac{1-e^{-50\cdot PD_{i,l}}}{1-e^{-50}}\right)$$

- *LGD is the loss given default (considered to be 45% in the FIRB approach), and*
- *M(PDi,l) is an adjustment term, which is defined as*

$$M(PD_{i,l})=\frac{(1+(M-2.5)\cdot b_{i,l})\cdot 1.06}{1-1.5\cdot b_{i,l}}$$

In this last formula, *M* is the time to maturity (considered to be 2.5 years in the FIRB approach), and *b* is the maturity adjustment, which is computed as $b_{i,l} = \left(0.11856 - 0.05478\cdot\ln(PD_{i,l})\right)^2$.

The minimum capital requirement of the bank (MCR_i) is equal to the sum of the capital requirements of all the exposures:

$$MCR_i = \sum_l CR_{i,l} \cdot A_{i,l} \qquad [2]$$

where $A_{i,l}$ is the amount of exposure *l*.

Because there is no available public information about the different banking exposures, the model considers only one debtor that is equivalent to the total portfolio, and it estimates the $IOPD_i$ by solving the following equation:

$$CR(IOPD_i) \cdot \sum_l A_{i,l} = MCR_i \qquad [3]$$

where $CR(IOPD_i)$ is the minimum capital requirement based on the Basel regulation (equal to 8% of the risk-weighted assets), and $\sum_l A_{i,l}$ is the total assets of the bank.

STEP 2: Simulation of correlated losses for the banks in the system

The model uses the Monte Carlo simulation to generate the bank loss distribution by using the same IRB formula and imposing a correlation structure among banks (with a correlation coefficient set to $\rho = 50\%$)[9]. This correlation exists because of the banks' common exposure, either to the same borrower or, more generally, to a particular common influence of the business cycle. In each simulation run *j*, the losses for bank *i* are simulated as follows:

$$L_{i,j} = LGD \cdot N\left[\sqrt{\frac{1}{1\text{-}R(IOPD_i)}}\, N^{-1}(IOPD_i) + \sqrt{\frac{R(IOPD_i)}{1\text{-}R(IOPD_i)}} \cdot N^{-1}(\alpha_{i,j})\right] \qquad [4]$$

where N is the normal distribution function, and $N^{-1}(\alpha_{n,i})$ are correlated normal random shocks. The $IOPD_i$ is the implied obligor probability of default estimated for each bank in step 1, and LGD is the loss given default, which is set as in the Basel regulation equal to 45%.

STEP 3: Determination of the failure event and estimation of PD bank

The default of a bank is determined by the size of the simulated losses and the regulatory capital available to absorb unexpected shocks. A bank *i* is considered in default

9 SYMBOL is often run imposing an equal correlation factor of 50% among all banks. A discussion and a sensitivity check of this assumption can be found in De Lisa et al. (2011) and Benczur et al. (2017).

when its simulated losses (L_{ij}) exceed the sum of the expected losses (EL_i) and the total actual capital (K_i) given by the sum of its minimum capital requirement plus the bank's excess capital:

$$L_{i,j} - EL_i - K_i > 0 \Rightarrow Bank\ i\ defaults \qquad [5]$$

The probability of the default of a bank (PD bank) is calculated as the number of times that the bank defaults over the total number of simulations (500,000).

3.2. Factors influencing the banks' PD

Factors influencing banks' PD can be classified into two main groups. First, there is a group of determinants that are specific to each bank and that have generally been used to evaluate bank risk, and they are called CAMEL indicators. They include capital adequacy, assets quality, management quality, earnings and liquidity. The second group of determinants includes the variables related to the banking sector structure and to the macroeconomic environment, such as size, sector concentration, economic growth, inflation, interest rates and unemployment.

3.2.1. CAMEL indicators

The CAMEL rating system was developed by the Federal Reserve of the United States and the Federal Deposit Insurance Corporation (FDIC) in 1979. The main objective of this internal supervision tool is the assessment of the soundness of financial institutions. In fact, from a regulatory perspective, bank risk has generally been evaluated using indicators associated with different risk categories such as Capital adequacy, Asset quality, Management Quality, Earnings and Liquidity (CAMEL). Nurazi and Evans (2005) and Olweny and Shipho (2011) demonstrated that CAMEL ratios can be used to predict bank failures. In the same way, Berger et al. (2000), Männasoo and Mayes (2009), Jin et al. (2011), Poghosyan and Cihák (2011), Cole and White (2012), Chiaramonte and Casu (2013), Kerstein and Kozberg (2013), and Vazquez and Federico (2015), among others, have used CAMEL variables to analyse the risk profile of entities around the world while also considering their structural and macroeconomic variables.

Table 1 shows the CAMEL variables included in our models. They include the following: i) concerning capital, the leverage ratio (C1), the Tier 1 capital coverage ratio (which is calculated as the actual own funds divided by the required own funds) (C2) and the CET1 ratio (C3); ii) concerning asset quality, the NPL ratio (AQ1) and the coverage ratio (AQ2); iii) concerning management and earnings, the ROA (ME1), the ROE (ME2) and the efficiency ratio (ME3); and iv) concerning liquidity and funding, the loans-to-deposits ratio (L1), the loans-to-total assets ratio (L2) and the liquidity ratio (L3).

[INSERT TABLE 1 ABOUT HERE]

The choice of these final explanatory variables is driven by availability considerations and multicollinearity issues. Correlation analyses and collinearity diagnostics were performed to assess the extent of multicollinearity among independent variables (See appendix A).

3.2.2. Structural and macroeconomic variables

In addition to bank-specific risk indicators, some factors related to the industry structure and the macroeconomic environment may influence the bank´s PD (see Table 1).

As structural variables, we consider the size of the bank and the sector concentration Herfindahl Index. Previous literature found some evidence concerning the influence of the size of banks on their risk (Klomp and Haan, 2012; Vázquez and Federico, 2015). On the one hand, several authors show the existence of a negative relationship between size and risk because of the managerial abilities and efficiency of large banks (Saunders et al., 1990; Boyd and Prescott, 1986; Salas and Saurina, 2002; Jabra et al., 2017). On the other hand, there is a view that supports that larger banks tend to be riskier due to a moral hazard problem

(González, 2005; De Jonghe, 2010; Uhde and Heimeshoff, 2009). Regarding the banking sector concentration, empirical research shows that their influence on bank risk could be positive, thus corroborating the concentration fragility view (Boyd and De Nicoló, 2005), or negative, thus supporting the concentration stability view (Allen and Gale, 2000).

Finally, regarding the macroeconomic scenario, we consider the economic growth, inflation, interest rates and unemployment rates. Festic et al. (2011), Poghosyan and Cihak (2011), Delis and Kouretas (2011), Baselga-Pascual et al. (2015), and Jabra et al. (2017), among others, found significant effects of some of these variables on bank credit risk and on the likelihood of bank distress.

3.3. Econometric model

We estimate dynamic panel data models to examine the determinants of the European banks' PD. The main benefit of using panel data is that they overcome the unobservable, constant and heterogeneous characteristics of each bank in the sample, e.g., omitted variables such as the ability of the manager or the risk culture of the bank. In addition, the dynamic panel data methodology enables us to correct for a typical problem in analysing determinants of bank risk: endogeneity. For instance, capital, asset quality, liquidity and business management indicators may influence the probability of default of the bank, but the probability of default could cause banks to modify their ratios of capital, asset quality, etc. Finally, dynamic panel data models allow us to consider the persistence of bank risk.

We use the system-GMM estimator developed for dynamic panel models by Arellano and Bover (1995) and Blundell and Bond (1998). Specifically, we employ the two-step estimation procedure included in the xtabond2 Stata routine written by Roodman (2009), with corrected standard errors for small samples proposed by Windmeijer (2005). This estimator may control the correlation of errors over time, the heteroscedasticity across banks and the measurement errors, due to the application of the orthogonality conditions in the variancecovariance matrix (Castro et al., 2015). Moreover, the efficiency of the GMM is improved by adding new nonlinear functions of the exogenous variables to the instruments (Hsiao, 2003).

We treat all the CAMEL indicators as endogenous variables by employing suitable instruments for both the equations in levels and the equations in differences. Arellano and Bover (1995) propose the use of instruments in first differences for equations in levels and instruments in levels for equations in first differences. Blundell and Bond (1998) support the efficiency of Arellano and Bover's (1995) estimator, especially for short sample periods and persistent data.

We verify the validity of the model and the instruments by performing a battery of tests. First, we employ the AR(2) or m_2 statistic tests for the lack of second-order serial correlation in the first-differenced residuals (Arellano and Bond, 1991). Second, we employ the Hansen *J* statistic of over-identifying restrictions tests for the absence of correlation between the instruments and the error term. Finally, we employ several Wald tests for the joint significance of the reported coefficients in the models.

Our baseline equation is the following:

$$PD_{i,j,t} = \alpha + \delta \cdot PD_{i,j,t-1} + \beta \cdot V_{i,j,t} + \sum \varphi_t \cdot D_{i,t} + \varepsilon_{i,j,t} \qquad [6]$$

where $PD_{i,j,t}$ represents the probability of default of the bank *i* in country *j* at year *t*; $PD_{i,j,t-1}$ denotes its lagged value, δ measure the speed of mean reversion, α is the constant term, $V_{i,j,t}$ denotes the explanatory variables (CAMEL indicators, structural and macroeconomic variables), β is the vector of coefficient estimated, and $\sum \varphi_t \cdot D_{i,t}$ represents the year dummies for the period 2011-2016. Finally, $\varepsilon_{i,j,t}$ is the disturbance term, which is divided into two components: the unobserved heterogeneity (η_i) and the idiosyncratic error (ν_{it}).

3.4. Data

Our sample consists of European banks that have available information for all the analysed variables during the period of 2011-2016. The data are obtained from *Orbis Bank Focus Database* by *Bureau Van Dijk*. At the end of 2016, this database contained 5 years' history for listed banks and 3 years for unlisted. Considering this limitation, the requirement of the statistical methodology and the convenience of the homogeneity of the sample, we focus on European listed banks.

After eliminating observations with missing values, extreme values considered outliers and banks for which there are not at least five consecutive years of information available[10], we finally have an unbalanced panel with 571 observations related to 97 listed entities of 20 European countries during the period of 2011-2016. In general, all the selected entities occupy important positions in the world and country rankings (based on the levels of assets). Tables 2 and 3 show the number of observations and institutions by country and the summary descriptive statistics, respectively. As we can see, the majority of institutions in the sample are commercial banks. Countries with more entities in the sample are Denmark, Italy, Poland, Spain and Germany.

[INSERT TABLES 2 AND 3 ABOUT HERE]

4. Results

4.1. Evolution of the European banking sector after the last regulatory reforms

Tables 4 and 5 establish an initial outline of the European banking sector from 2011 to 2016. Table 4 shows the descriptive statistics for the banks' PD and their main components. PD is the probability of default of a bank and represents the probability of a bank's losses

[10] To estimate a system-GMM methodology, it is advisable to use a minimum of 5 consecutive years´ information for each individual in the sample in order to obtain the validity tests of the model.

falling in the tail of its loss distribution not covered by its capital (given by the sum of regulatory capital requirements and any excess capital). IOPD is the average default probability of the obligor portfolio of each bank and is employed to compute asset losses. Excess capital (EC) represents the bank excess capitalisation rate. (It relates the difference between regulatory capital and minimum requirements on assets.)

The mean value of the PD for the period of study is 0.0033%. Half of the entities present values lower than 0.0026%. The values are relatively low because the Basel minimum capital requirements are established for a confidence interval of 99.9%, which supposes a maximum risk tail of 0.1%. The bank risk exhibits a clear improvement throughout the period. In this sense, the PD diminishes by nearly 60%. This behaviour can be explained by two factors: the reduction in the portfolio risk (IOPD) and the increase in the bank capital excess (EC). More than a half of the entities in the sample show a decrease of 21% in their portfolio risk and an increase of 40% in their bank capital surplus.

The leverage, capital quality and liquidity indicators exhibit improvements according to the new requirements of Basel III (see Table 5). In contrast, the asset quality indicators show an increase in the delinquency rate and a decrease in the coverage ratio. Consequently, the return on assets and the efficiency notably improve during the whole period, while the return on equity experienced an increase since 2013.

Additionally, the size of the entities and the concentration of the European banking sector have experienced notable rises because of the mergers that have happened in the last years. Finally, the macroeconomic variables show a post-crisis outlook with lower interest rates and inflation. The evolution of the GDP and the unemployment rates shows the recession from 2011 to 2013 and the notable improvements from then up to 2016, thus highlighting the start of the economic recovery.

[INSERT TABLES 4 AND 5 ABOUT HERE]

4.2. Factors influencing the probability of default of European banks

Table 6 reports the results of the empirical estimation of Equation (6). Model 1 only includes the bank-specific CAMEL indicators while Model 2 also considers a set of structural and macroeconomic variables. The results show the main factors influencing the European banks' PD based on SYMBOL Model. The statistical significance of the lagged dependent variable and the higher values of δ indicate the dynamic nature of the model's specification and its strong persistence, respectively.

[INSERT TABLE 6 ABOUT HERE]

Leverage and capital quality indicators have significantly positive impacts on PD such that an increase in these indicators increases the ability to absorb sudden losses and, therefore, to reduce the probability of default. Our findings agree with the expected relations established by the European Banking Authority (2015) and support previous research, such as Čihák and Schaeck (2010), Baselga-Pascual et al. (2015) and Leung et al. (2015).

Management and earnings indicators appear to be statistically significant as well but they have different impacts on risk. The return on assets (ME1) shows a positive influence on risk (an increase in ME1 causes a decrease in the PD), while the return on equity exhibits negative effects (an increase in ME2 causes a rise in the PD). In this sense, the literature shows contradictory findings for the relationship between profitability and the risks of banks. Several studies have found evidence of a negative relationship between bank performance and risk (Angbazo, 1997; Čihák and Schaeck, 2010; Poghosyan and Čihák, 2011; Baselga-Pascual et al., 2015; among others). Banks with higher earnings are less likely to experience distress. However, there are reports that support the contrary argument. Ahmad and Arrif (2007) explain the positive relationship between the earnings ratio and the credit risk as higher returns on assets may imply a higher proportion of riskier loans with the consequence of a potential increase in the bank credit risk.

Finally, the liquidity indicator L1 also influences our measure of PD such that the higher that the ratio is, the lower the probability of default of the banks. This relation could be due to the interrelation between the liquidity creation (or illiquidity) and the bank capital. Following the Basel III regulatory changes, the study of the bidirectional interactions between regulatory capital and liquidity creation has attracted notable attention, but results are inconclusive. The literature suggests a complex relationship between these two variables such that capital may improve or deteriorate liquidity creation, and vice versa (Distinguin et al., 2013; Tran et al., 2016; Casu, et al., 2018; DeYoung, et al., 2018). Our findings are in line with the "liquidity risk" hypothesis in which as greater liquidity creation increases the risk of illiquidity for banks, banks should strengthen their solvency because capital acts as a buffer against unexpected withdrawals from customers. Therefore, the rises in capital resulting from liquidity creation may favourably affect the risk.

When we introduce the structural and macroeconomic variables in the equation (model 2 in Table 6), the results reveal relevant impacts of inflation, interest rates, unemployment and the concentration of the banking sector on the probability of default of European banks. The findings confirm the "concentration-stability" argument. Concentration may increase the profitability and market power of banks. These profitable firms have the ability to increase more capital so that their losses will be largely covered (Altuntas and Rauch, 2017), and thus, the unexpected losses will be lower.

The size also appears to be statistically significant in our model. The negative sign of this variable indicates that an increase in the size of the bank causes a reduction of the PD. Boyd and Prescott (1986) and Salas and Saurina (2002) state that this relationship could be due to economies of scale and scope through which larger banks may diversify loan portfolio risk more efficiently.

Finally, when we introduce these additional variables, the model provides us with new findings about the CAMEL indicators (see Table 6). One more indicator appears to be statistically significant: the delinquency rate (AQ1). There is a general consensus in the literature that an increase in the percentage of non-performing loans (NPLs) can reduce the quality of the banking sector's assets and increase the probability of default (De Nicoló et al., 2003; Blasco and Sinkey, 2006; Männasoo and Mayes, 2009; Festic et al., 2011; Baselga-Pascual et al., 2015, among others). However, in our model, the ratio shows a different influence in risk such that the higher that the delinquency rate is, the lower the bank´s probability of default. It could be a consequence of the new capital regulation, which ensures that banks with lower quality assets will reinforce their capital, and this higher buffer of capital will lead to a higher coverage of unexpected losses and more stable banks (Laeven and Levine, 2009; Jokipii and Milne, 2011; Ghosh, 2015; Acosta et al., 2017).

4.3. Impact of the financial crisis and the specialization of the entity

Considering our period of study and the previous research about the influence of crises on the determinants of bank risk (see Haq and Heaney, 2012), we introduce a dummy variable to differentiate the years when the consequences of crisis were more critical (2011-2012) and the recovery period (2013-2016)[11].

The results in Table 7 (Model 1) show the statistical significance of this new variable, which demonstrates the notable influence of the financial crisis in the European banks' PD. The positive sign indicates that European banks have reduced their PD after the crisis.

Additionally, following the previous literature on the influence of the banks' ownership concentration in its risk (Shehzad et al., 2010), we introduce in the equation a set of dummies to control for the effects of the specialization of the entity (commercial banks,

[11] The dummy variable takes the value of 1 for crisis years (2011-2012) and it takes the value of 0 for the remaining years (2013-2016).

saving banks and credit cooperatives). However, the results do not prove to be statistically significant for these variables (see Model 2 in Table 7).

[INSERT TABLE 7 ABOUT HERE]

4.4. Robustness test

4.4.1. The European Banking Union

An important project to reduce the financial fragmentation and to restore confidence in the European context is the creation of a banking union (Abascal et al., 2015). In Table 8, we substitute the crisis dummy with a dummy to examine the existence of the European banking union.[12] In the analysis, we differentiate two periods. The first one covers from 2011 to 2013, when the consequences of the crisis were more critical. In this scenario, a common strategy towards an economic and monetary union was announced. By September 2012, the European Commission proposed the Single Supervisory Mechanism (SSM) and established a roadmap to achieve the European banking union.

The second period includes the years from 2014 to 2016. The Single Supervisory Mechanism (SSM) came into effect in November 2014. In addition, a political agreement was achieved to create the second pillar of the banking union: the Single Resolution Mechanism (SRM). The Single Resolution Board (SRB) was established in January 2015, but it did not undertake any action until January 2016 (Abascal et al., 2015).

Our results (Table 8) do not prove the statistical significance of this additional variable, possibly because the effects will happen in the long term, but we can check that the CAMEL indicators, the size of banks and several macroeconomic variables maintain their statistical significance, thus supporting the robustness of the results.

[INSERT TABLE 8 ABOUT HERE]

[12] The variable takes the value of 1 for the period previous to the bank union (2011-2013) and 0 for the period from 2014 to 2016.

4.4.2. Alternative econometric methodologies

As a second robustness test, we report the results using different econometric methods of estimation (the pooled OLS, fixed-effect model and difference GMM). Table 9 shows that the main CAMEL indicators, and some of the structural and macroeconomic variables remain significant, regardless of the methodology; this outcome proves the robustness of the specification.

[INSERT TABLE 9 ABOUT HERE]

4.4.3. Alternative definition of the dependent variable

Finally, we consider an alternative representation of the dependent variable in equation (6). Although the odds ratios and log odds are more typical in a logistic regression context, the conversion of the probabilities of default to log odds ratios may have some benefits in terms of the symmetry of the distribution for statistical analyses (Lipsitz et al., 1991).

The log odds ratio can take any value and has an approximately normal distribution. The results obtained do not differ substantially from those obtained previously; most of the explanatory variables retain both their signs and their statistical significance.

[INSERT TABLE 10 ABOUT HERE]

5. Conclusions

The development of the European Banking Union (BU) is posing a host of regulatory proposals focused on the stability and the sustainability of the financial system. In this scenario, the European Commission designed the SYMBOL methodology, which presents a model that estimates the banking losses using the Basel risk assessment and regulatory framework (De Lisa et al., 2011). Based on this method, our work proposes a new measure of bank risk (the probability of default), which may complement the traditional risk proxies based on accounting and market information. Considering a representative sample of European listed banks during the period from 2011-16, we estimate the banks' PD and we analyse the influence of several CAMEL, structural and macroeconomic indicators in these PD.

Results for the European banking system show a relevant reduction of risk from a regulatory approach during the period of study. This reduction has been motivated to some extent by a decrease in the portfolio risk, but especially by the banking capital excess. The leverage and capital quality indicators positively influence the PD just due to the new bank standards. However, the profitability shows contradictory results, which is the same as in the previous literature. While an increase in the returns on assets causes a decrease in the risk, the returns on equity exhibit the opposite impact. Liquidity creation also has a positive effect on risk, probably due to the increases of capital. Finally, asset quality continues to be one of the challenges for the strengthening of the BU because of the high levels of delinquency in some areas of the banking sector. Regarding the structural and macroeconomic variables, size, concentration index, inflation, interest rate and unemployment are the main determinants of the PD.

Our findings may be important to regulatory and supervisory authorities for several reasons. First, the PD based on the SYMBOL model could be considered as a new proxy for bank risk from a regulatory approach. It could be used jointly with the traditional indicators for a more complete analysis of unexpected bank losses. Second, given the standardized calculation of this measurement, it could enable comparisons across different credit institutions, thus improving market discipline. Finally, our results might be useful to the design of new regulations focused on the key factors that affect the banks' probability of default.

References

Abascal, M., Alonso-Gispert, T., Fernández de Lis, S., GolecKi, W.A., 2015. A banking union for Europe: Making a virtue out of necessity. The Spanish Review of Financial Economic 13, 20-39.

Acosta, J., Grill, M., Hannes, J., 2017. The leverage ratio, risk-taking and bank stability. European Central Bank. Working Paper Series 2079

Agoraki, M.-E. K., Delis, M. D., Pasiouras, F., 2011. Regulations, competition and bank risk-taking in transition countries. Journal of Financial Stability 7, 38–48.

Ahmad, N.H., Arrif, M., 2007. Multi-country study of bank credit risk determinants. International Journal of Banking and Finance 5 (1), 135-152.

Allen, F., Gale, D., 2000. Financial contagion. Journal of Political Economy 108, 1–33.

Alter, A., Schüler, Y.S., 2012. Credit spread interdependencies of European states and banks during the financial crisis. Journal of Banking and Finance 36, 3444–3468.

Altuntas, M. Rauch, J., 2017. Concentration and financial stability in the property-liability insurance sector: global evidence. The Journal of Risk Finance 18 (3), 284-302.

Angbazo, L. 1997., Commercial bank net interest margins, default risks, interest rate risks and off-balance sheet banking. Journal of Banking and Finance 2, 55-87.

Annaert, J., De Ceuster, M., Van Roy, P., Vespro, C., 2013. What determines Euro area bank CDS spreads? Journal of International Money and Finance 32, 444–461.

Arellano, M., Bond., S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies 58, 277-297.

Arellano, M., Bover, O., 1995. Another look at the instrumental variable estimation of error-components models. Journal of Econometrics 68(1), 29-51.

Basel Committee on Banking Supervision (BCBS), 2006. Basel II: International Convergence of Capital Measurement and Capital Standards. A Revised Framework. June.

Baselga-Pascual, L., Trujillo-Ponce, A., Cardone-Riportella, C., 2015. Factors influencing bank risk in Europe: Evidence from the financial crisis. North American Journal of Economics and Finance 34, 138-166.

Benczur, P., Cannas, G., Cariboni, J., Di Girolamo, F., Maccaferri, S., Petracco Giudici, M., 2017. Evaluating the effectiveness of the new EU bank regulatory framework: A farewell to bail-out? Journal of Financial Stability 33, 207-223.

Berger, A. N., DeYoung, R., 1997. Problem loans and cost efficiency in commercial banks. Journal of Banking and Finance 21, 849–870.

Berger, A.N., Davies, S.M., Flannery, M.J., 2000. Comparing Market and Supervisory Assessments of Bank Performance: Who Knows What When? Journal of Money, Credit and Banking 32 (3), 641-667.

Betz, F., Oprica, S., Peltonene, T.A., Sarlin, P., 2014. Predicting Distress in European Banks. Journal of Banking and Finance 45, 225-241.

Blasco, M., Sinkey, J., 2006. Bank asset structure, real-estate lending, and risk-taking. Quarterly Review of Economics and Finance 46, 53–81.

Blundell, R., Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87(1), 115-143.

Boyd, J. H., de Nicoló, G., 2005. The theory of bank risk taking and competition revisited. Journal of Finance 60, 1329–1343.

Boyd, J. H., Prescott, E. C., 1986. Financial intermediary coalitions. Journal of Economic Theory 38, 211–232.

Calomiris, C., 2009. Financial innovation, regulation and reform. Cato Journal 29, 65-91.

Campolongo, F., De Lisa, R., Zedda, S., Vallasca, F., Marchesi, M., 2010. Deposit Insurance Schemes: target fund and risk-based contributions in line with Basel II regulation. JRC Scientific and Technical Reports 57325.

Cannas, G., Cariboni, J., Forys, M., Joensson, H., Langedijk, S., Marchesi, M., Nda-cyayisenga, N., Pagano, A., Petracco-Giudici, M., 2013. Quantitative estimation of a part of the costs and benefits of bank structural separation. JRC Scientific and Technical Report 88531.

Cariboni; J., Fontana, A., Langedijk, S., Maccaferri, S., Pagano, A., Petracco-Giudici, M., Rancan, M., Schich, S., 2015. Reducing and sharing the burden of bank failures. OECD Journal: Financial Market Trends 2015/2, 29-61.

Castro, P., Tascón, M.T., Amor-Tapia, B., 2015. Dynamic analysis of the capital structure in technological firms based on their life cycle stages. Spanish Journal of Finance and Accounting 44 (4), 458-486.

Casu, B., Di Pietro, F., Trujillo-Ponce, A., 2018. Liquidity creation and Bank Capital. Journal of Financial Services Research. https://doi.org/10.1007/s10693-018-0304-y.

Chiaramonte, L., Casu, B., 2013. The determinants of bank CDS spreads: evidence from the financial crisis. European Journal of Finance 19, 861–887.

Čihák, M., Schaeck, K., 2010. How Well Do Aggregate Prudential Ratios Identify Banking System Problems? Journal of Financial Stability 6(3), 130-44.

Constantine, A., Peltonen, T.A., Sarlin, P., 2018. Network linkages to predict bank distress. Journal of Financial Stability 35, 226-241.

Cole, R.A., White, L.J., 2012. Deja vu all over again: The causes of U.S. commercial bank failures this time around. Journal of Financial Services Research 42, 5-29.

De Jonghe, O., 2010. Back to the basis on banking? A micro-analysis of banking system stability. Journal of Financial Intermediation 19, 387–417.

De Lisa, R., Zedda, S., Vallascas, F., Campolongo, F., Marchesi, M., 2011. Modelling deposit insurance scheme losses in a Basel 2 framework. Journal of Financial Services Research 40, 123-141.

De Nicoló, G., Geadah, S., Rozhkov, D., 2003. Bridging the great divide: Poorly developed financial systems in the CIS-7countries may jeopardize their sustained growth. Finance and Development 40, 42–45.

Delis, M. D., Kouretas, G. P., 2011. Interest rates and bank risk-taking. Journal of Banking and Finance 35, 840–855.

Delis, M. D., Staikouras, P. K., 2011. Supervisory effectiveness and bank risk. Review of Finance 15, 511–543.

Demirgüç-Kunt, A., Huizinga, H., 2010. Bank activity and funding strategies: The impact on risk and returns. Journal of Financial Economics 98, 626–650.

DeYoung, R., Distinguinb, I., Tarazic, A., 2018. The joint regulation of bank liquidity and bank capital. Journal of Financial Intermediation 34, 332-46.

Distinguin, I., Roulet, C., Tarazi, A., 2013. Bank regulatory capital and liquidity: evidence from US and European publicly traded banks. Journal of Banking and Finance 37, 3295-3317.

Drago, D., Di Tommaso, C., Thornton, J., 2017. What determines bank CDS spreads? Evidence from European and US banks. Finance Research Letters 22, 140-145.

European Banking Authority (EBA), 2015. Guidelines on methods for calculating contributions to deposit guarantee schemes, EBA/GL/2015/10. https://www.eba.europa.eu/documents/10180/1089322/EBA-GL-2015-10+GL+on+methods+for+calculating+contributions+to+DGS.pdf (accessed 01.08.16).

European Commission, 2012. Directorate-General for Economic and Finan-cial Affairs. Fiscal Sustainability Report. http://ec.europa.eu/economy,_finance/publications/european_economy/2012/pdf/ee-2012-8_en.pdf (accessed 12.08.16).

European Commission, 2014. Directorate-General for Internal Market and Services. In: Commission Staff Working Document – Economic Review of the Financial Regulation Agenda. http://ec.europa.eu/internal market/finances/docs/general/20140515-erfra-working-document en.pdf.

European Commission, 2016. Effects analysis (EA) on the European Deposit Insurance Scheme (EDIS). https://ec.europa.eu/info/publications/effects-analysis-european-deposit-insurance-scheme-edis_en (accessed 27.09.17).

European Commission, 2018. Guidelines for the review of models used in support of EU policies. Directorate General Joint Research Centre.

Fernandes, C., Farinha, J., Vitorino Martins, F., Mateus, C., 2016. Determinants of European Banks' Bailouts Following the 2007-2008 Financial Crisis. Journal of International Economic Law 19, 707-742.

Festic, M., Kavkler, A., Repina, S., 2011. The macroeconomic sources of systemic risk in the banking sectors of five new EUmember states. Journal of Banking and Finance 35, 310–322.

Galliani, C., Zedda, S., 2015. Will the Bail-in Break the Vicious Circle Between Banks and their Sovereign? Computational Economics 45 (4), 597–614.

Giordana, G.A.; Schumacher, I., 2017 An Empirical Study on the Impact of Basel III Standards on Banks' Default Risk: The Case of Luxembourg. Journal of Risk and Financial Management 10, 8.

Ghosh, A., 2015. Banking-industry specific and regional economic determinants of non-performing loans: Evidence from US states. Journal of Financial Stability 20, 93-104.

Gómez, P., Parrado-Martínez, P., Partal-Ureña, A., 2018. Risk Profile Indicators and Spanish Banks' Probability of Default from a Regulatory Approach. Sustainability 10, 1259, 1-16. http://www.mdpi.com/2071-1050/10/4/1259.

González, F., 2005. Bank regulation and risk-taking incentives: An international comparison of bank risk. Journal of Banking and Finance 29, 1153–1184.

Haq, M., Heaney, R., 2012. Factors determining European bank risk. Journal of International Financial Markets, Institutions & Money 22, 696-718.

Hau, H., Langfield, S., Marques-Ibanez, D., 2012. Bank ratings. What determines their quality? European Central Bank. Working paper series 1484 / october 2012.

Hordijk, L., Kancs, D., editor(s), Kok, C., Ongena, S., Pelizzon, L., Cariboni, J., Heynderickx, W., Maccaferri, S., Pagano, A., and Petracco Giudici, M., 2018. Review of the SYMBOL model, EUR 29233 EN. Publications Office of the European Union, Luxembourg, doi: 10.2760/607271 (online), JCR111667.

Hsiao, C., 2003. Analysis of panel data. Cambridge: Cambridge University Press.

Jabra, W.B., Mighri, Z., Mansouri, F., 2017. Determinants of European Bank risk during financial crisis. Cogent Economics & Finance 5, 1-20.

Jin, J., Kanagaretnam, K., Lobo, G., 2011. Ability of accounting and audit quality variables to predict bank failure during the financial crisis. Journal of Banking and Finance 35(11), 2811-2819.

Jokipii, T., Milne, A., 2011. Bank capital buffer and risk adjustment decisions. Journal of Financial Stability 7, 165-178.

Khan, M.S, Scheule, H., Wu, E., 2016. Funding liquidity and bank risk taking. Journal of Banking and Finance 82, 203-216.

Kerstein, J., Kozberg, A., 2013. Using Accounting Proxies of Proprietary FDIC Ratings to Predict Bank Failures and Enforcement Actions During the Recent Financial Crisis. Journal of Accounting, Auditing and Finance 28(2), 128-151.

Klomp, J., Haan, J., 2012. Banking risk and regulation: Does one size fit all? Journal of Banking and Finance 36, 3197–3212.

Laeven, L., Levine, R., 2009. Bank governance, regulation and risk taking. Journal of Financial Economics 93, 259–275.

Leung, W. S., Taylor, N., Evans, K. P., 2015. The determinants of bank risks: Evidence from the recent financial crisis. Journal of International Financial Markets, Institutions and Money 34, 277-293.

Lepetit, L., Strobel, F., 2015. Bank insolvency risk and Z-score measures: A refinement. Finance Research Letters 13, 214-224.

Lipsitz, S.R., Laird, N.M, Harrington, D.P., 1991. Generalized estimating equations for correlated binary data: Using the odds ratio as a measure of association. Biometrika 78 (1), 153-160.

Männasoo, K., Mayes, D. G., 2009. Explaining bank distress in Eastern European transition economies. Journal of Banking and Finance 33, 244–253.

Marchesi, M., Petracco Giudici, M., Cariboni, J., Zedda, S., Campolongo, F., 2012. Macroeconomic Cost-benefit Analysis of Basel III Minimum Capital Requirements and of Introducing Deposit Guarantee Schemes and Resolution Funds. European Commission JRC Scientific and Policy Report 24603.

Merton, R.C., 1974. On the pricing of corporate debt: the risk structure of interest rates. Journal of Finance 29 (2). 449–470.

Muresano, R., Pagano, A., 2016. Adapting and Optimizing the Systemic Model of Banking Originated Losses (SYMBOL) Tool to the Multi-Core Architecture. Computational Economics 48, 253-280.

Nurazi, R., Evans, M., 2005. An Indonesian Study of the Use of CAMEL(S) Ratios as Predictors of Bank Failure. Journal of Economic and Social Policy 10 (1), 1-23.

Olweny, T., Shipho, T.M., 2011. Effects of Banking Sectoral Factors on the Profitability of Commercial Banks in Kenya. Economics and Finance Review 1(5), 1-30.

Pagano, A., Cariboni, J., Petracco-Giudici, M., 2012. Symbol model database and analyses for public finance sustainability. JRC Technical Report 77925.

Poghosyan, T., Cihak, M., 2011. Determinants of bank distress in Europe: Evidence from a new data set. Journal of Financial Services Research 40, 163–184.

Roodman, D., 2009. How to do xtabond2: An introduction to difference and system GMM in Stata. The Stata Journal 9, 86-136.

Salas, V., Saurina, J., 2002. Credit Risk in Two Institutional Regimes: Spanish Commercial and Savings Banks. Journal of Financial Services Research 22 (3), 203-224.

Samaniego-Medina, R., Trujillo-Ponce, A., Parrado-Martínez, P., di Pietro, F., 2016. Determinants of bank CDS spreads in Europe. Journal of Economics and Business. 86, 1–15.

Saunders, A., Strock, E., Travlos, N. G., 1990. Ownership structure, deregulation, and bank risk taking. The Journal of Finance 45, 643–654.

Shehzad, C.T., Haan, J., Scholtens, B., 2010. The impact of bank ownership concentration on impaired loans and capital adequacy. Journal of Banking and Finance 34(2), 399-408.

Tran, V.T., Lin, T., Nguyen, H., 2016. Liquidity creation, regulatory capital, and bank profitability. International Review of Financial Analysis 48, 98–109.

Uhde, A., Heimeshoff, U., 2009. Consolidation in banking and financial stability in Europe: Empirical evidence. Journal of Banking and Finance 33, 1299–1311.

Vazquez, F., Federico, P., 2015. Bank funding structures and risk: Evidence from the global financial crisis. Journal of Banking and Finance 61, 1-14.

Wang, L., 2017. Bank Rating Gaps as Proxies for Systemic Risk. University of Alberta. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2966413.

Windmeijer, F., 2005. A finite sample correction for the variance of linear efficient GMM estimators. Journal of Econometrics 126, 25-51.

Tables

TABLE 1. Variables definition

Category	Variable and definition	Notation	Data Source
Dependent variable			
Probability of default	Probability of default of a bank based on simulated losses (SYMBOL Methodology)	PD	Own elaboration
CAMEL variables			
Capital	Leverage ratio (Tier 1 capital/Total assets)	C1	Orbis Bank Focus
	Capital coverage ratio (Actual common equity Tier 1 ratio/Required common equity Tier 1 ratio)	C2	Orbis Bank Focus
	CET1 ratio (Common equity Tier 1 capital/Risk-weighted assets)	C3	Orbis Bank Focus
Asset Quality	Non-performing loans ratio (Non-performing loans/Total loans and debt instruments)	AQ1	Orbis Bank Focus
	Coverage ratio (Loans loss reserves to impaired loans)	AQ2	Orbis Bank Focus
Management Quality and Earnings	Return on asset (Net income/Total assets)	ME1	Orbis Bank Focus
	Return on equity (Net income/Shareholders' equity)	ME2	Orbis Bank Focus
	Efficiency rate (Cost to income ratio)	ME3	Orbis Bank Focus
Liquidity and Funding	Loans to deposit ratio	L1	Orbis Bank Focus
	Loans to total assets ratio	L2	Orbis Bank Focus
	Liquidity ratio (Liquid assets/Total assets)	L3	Orbis Bank Focus
Structural variables	Size (Natural log of total assets)	Size	Orbis Bank Focus
	Sector concentration (Herfindahl index for credit institutions to total assets)	HI	European Central Bank
Macroeconomic variables	Economic growth- Real GDP growth rate (percentage change on previous year)	GDP	Eurostat
	Inflation- Inflation rate (annual average rate of change %)	Inflation	Eurostat
	Interest rate- Annual interest rate % (EMU convergence criterion series)	Interest	Eurostat
	Unemployment- Unemployment rate (% of active population)	Unemploy	Eurostat

TABLE 2. Number of institutions and observations by country

Country	Commercial banks	Saving banks	Credit cooperatives	Observations
Germany	5	0	0	28
Austria	3	0	1	24
Bulgaria	1	0	0	6
Cyprus	1	0	0	6
Croatia	4	0	0	22
Denmark	20	1	0	126
Slovakia	4	0	0	24
Spain	7	1	0	48
Finland	2	0	0	12
France	4	0	1	30
Greece	3	0	0	17
Ireland	2	0	0	12
Italy	14	0	4	106
Malta	1	0	0	6
Netherland	1	0	0	6
Poland	8	1	1	57
Portugal	0	1	0	5
United Kingdom	1	1	0	12
Czech Republic	1	0	0	6
Sweden	2	1	0	18
Total	84	6	7	571

TABLE 3. Summary descriptive statistics (2011-2016)

Variable	N	Mean	Sd	Median	Max.	Min.
PD	571	0.003	0.004	0.003	0.047	0
C1	571	7.643	3.566	6.900	23.030	1.370
C2	571	313.723	188.016	275.660	2905.400	-171.170
C3	571	14.190	4.340	13.230	34.140	5.710
AQ1	571	11.059	10.066	7.620	57.530	0.200
AQ2	571	63.673	36.090	57.030	386.320	2.470
ME1	571	0.390	1.190	0.430	9.970	-7.150
ME2	571	4.098	42.786	6.270	670.240	-596.310
ME3	571	62.926	16.850	61.580	233.450	24.050
L1	571	104.847	44.054	97.410	316.850	4.510
L2	571	57.533	16.672	61.820	90.250	3.030
L3	571	19.340	14.038	16.410	94.480	1.370
Size	571	16.721	2.418	16.680	21.500	11.710
HI	571	0.089	0.053	0.070	0.370	0.030
GDP	571	1.195	2.537	1.300	25.600	-9.100
Inflation	571	1.153	1.366	0.600	4.500	-1.600
Interest	571	2.956	2.556	2.370	22.500	0.090
Unemploy	571	10.702	5.458	9.200	27.500	4

Notes: This table reports the main descriptive statistics for the variables used in the study. The sample comprises 97 bank institutions. See Table 1 for a description of the variables.

TABLE 4. Evolution of the European banks' PD

		2011	2012	2013	2014	2015	2016	Total
PD	Mean	0.0049	0.0037	0.0032	0.0032	0.0025	0.0021	0.0033
	Sd	0.0064	0.0040	0.0030	0.0029	0.0020	0.0017	0.0037
	Q1	0.0014	0.0014	0.0014	0.0014	0.0014	0.0008	0.0014
	Q2	0.0030	0.0026	0.0026	0.0026	0.0022	0.0014	0.0026
	Q3	0.0062	0.0054	0.0044	0.0040	0.0035	0.0028	0.0040
IOPD	Mean	0.3556	0.3176	0.3129	0.3061	0.3076	0.2821	0.3135
	Sd	0.2200	0.2623	0.2145	0.1938	0.2230	0.1744	0.2168
	Q1	0.1852	0.1482	0.1473	0.1506	0.1523	0.1464	0.1557
	Q2	0.3268	0.2986	0.2799	0.2799	0.2824	0.2574	0.2848
	Q3	0.5061	0.4405	0.4529	0.4376	0.4130	0.4066	0.4393
EC	Mean	3.9392	4.1651	4.1730	3.9572	4.2955	4.5973	4.1874
	Sd	2.9659	2.7076	2.3582	2.0850	1.8310	1.9731	2.3511
	Q1	1.9095	2.2225	2.4765	2.3407	2.9354	3.1342	2.4765
	Q2	3.1059	3.5302	3.5686	3.6399	4.0489	4.3477	3.6697
	Q3	5.1371	5.3694	5.5103	4.9930	5.3837	5.9774	5.4283

Notes: This table shows the main descriptive statistics (mean, standard deviation, first, second and third quartiles) for the PD and their main components.

TABLE 5. Evolution of the explanatory variables

	2011	2012	2013	2014	2015	2016	Total
C1	7.550 (4.180)	7.608 (3.926)	7.580 (3.603)	7.456 (3.339)	7.764 (3.087)	7.906 (3.235)	7.643 (3.566)
C2	275.126 (95.015)	308.419 (163.094)	323.513 (273.766)	325.638 (284.101)	315.468 (91.266)	332.768 (96.376)	313.723 (188.016)
C3	12.704 (4.351)	13.771 (4.417)	13.967 (4.394)	14.097 (4.039)	14.890 (4.030)	15.699 (4.324)	14.190 (4.340)
AQ1	8.232 (5.951)	10.241 (8.049)	11.436 (9.327)	12.131 (11.012)	12.082 (11.780)	12.111 (12.363)	11.059 (10.066)
AQ2	67.643 (53.237)	62.153 (35.305)	63.537 (32.974)	66.065 (40.725)	61.851 (26.137)	60.890 (19.683)	63.673 (36.090)
ME1	0.134 (1.348)	0.195 (1.384)	0.363 (1.168)	0.381 (1.026)	0.568 (0.802)	0.699 (1.256)	0.390 (1.189)
ME2	7.792 (76.544)	-3.099 (64.943)	3.148 (23.716)	3.285 (16.995)	6.832 (9.157)	7.010 (12.220)	4.098 (42.786)
ME3	66.652 (22.630)	64.776 (19.771)	61.969 (15.618)	60.283 (13.482)	61.939 (13.809)	62.128 (13.477)	62.926 (16.850)
L1	115.379 (48.098)	108.683 (46.374)	103.226 (39.023)	101.783 (41.673)	100.686 (44.293)	99.713 (43.668)	104.847 (44.054)
L2	58.929 (17.111)	57.195 (16.706)	57.372 (16.857)	56.492 (16.619)	57.402 (16.750)	57.911 (16.328)	57.532 (16.672)
L3	19.036 (13.007)	19.585 (14.091)	19.175 (14.262)	19.485 (14.333)	19.225 (14.450)	19.522 (14.341)	19.340 (14.037)
Size	16.587 (2.563)	16.672 (2.455)	16.681 (2.406)	16.734 (2.405)	16.763 (2.380)	16.891 (2.349)	16.721 (2.418)
HI	0.086 (0.058)	0.084 (0.049)	0.089 (0.055)	0.092 (0.057)	0.092 (0.053)	0.093 (0.046)	0.089 (0.053)
GDP	1.374 (2.240)	-0.145 (2.852)	-0.068 (1.570)	1.526 (1.656)	2.502 (3.626)	2.036 (1.057)	1.195 (2.537)
Inflation	2.910 (0.664)	2.738 (0.717)	1.079 (0.697)	0.272 (0.546)	-0.039 (0.486)	0.008 (0.443)	1.153 (1.366)
Interest	4.593 (2.376)	4.327 (3.833)	3.344 (1.747)	2.438 (1.242)	1.675 (1.675)	1.379 (1.619)	2.956 (2.556)
Unemploy	9.905 (4.437)	11.190 (5.561)	11.712 (6.165)	11.280 (5.890)	10.389 (5.382)	9.633 (4.865)	10.702 (5.458)

Notes: The table shows the mean and standard deviation (in parentheses) of each indicator for all the observations in the sample, by year. The sample comprises 97 bank institutions (571 observations). See Table 1 for a description of the variables.

TABLE 6. Factors influencing European bank´s PD

Variables	PD (1)	PD (2)
Lagged dependent	0.32764***	0.30338***
	(0.03962)	(0.05051)
C1	-0.00011**	-0.00018***
	(0.00005)	(0.00006)
C2	-1.27^{e-06}***	$-1.56e^{-06}$***
	(1.94^{e-07})	$(1.81e^{-07})$
C3	-0.00018***	-0.00017***
	(0.00004)	(0.00004)
AQ1	-0.00002	-0.00003**
	(0.00001)	(0.00001)
AQ2	1.44^{e-06}	$-6.73e^{-07}$
	(4.19^{e-06})	$(4.44e^{-06})$
ME1	-0.00077***	-0.00070***
	(0.00029)	(0.00023)
ME2	0.00002***	0.00001***
	(5.52^{e-06})	$(5.48e^{-06})$
ME3	-1.99^{e-06}	$8.63e^{-08}$
	(0.00001)	(0.00001)
L1	-6.53^{e-06}**	$8.24e^{-07}$
	(2.62^{e-06})	$(3.13e^{-06})$
L2	8.57^{e-06}	0.00001
	(0.00001)	$(9.96e^{-06})$
L3	5.99^{e-06}	0.00002*
	(0.00001)	(0.00001)
Size		-0.00032***
		(0.00010)
HI		-0.00749*
		(0.00240)
GDP		$8.94e^{-06}$
		(0.00005)
Inflation		-0.00019*
		(0.00010)
Interest		0.00018**
		(0.00008)
Unemploy		0.00005**
		(0.00003)
Constant		0.01081***
		(0.00207)
Year dummies	Yes	No
z_1	35.43 (12, 96)	30.30 (13, 96)
z_2	7.05 (5, 96)	4.03 (5, 96)
m_1	-2.82	- 2.74
m_2	-0.66	- 0.64
Hansen	79.62 (200)	78.75 (222)
Number of Obs.	474	474
Number of Banks	97	97

Notes: This table shows the two-step system-GMM estimator developed by Arellano and Bover (1995) and Blundell and Bond (1998). The dependent variable is the bank´s PD. See Table 1 for a complete description of the explanatory variables. All variables are considered endogenous except for the year dummies and the macroeconomic variables. The heteroskedasticity-consistent asymptotic standard errors are reported in parentheses. z_1 and z_2 are Wald tests of the joint significance of the reported coefficients of the explanatory variables and the year dummies (or macroeconomic variables), respectively. These statistics are asymptotically distributed as *F* under the null hypothesis of no significance with the degrees of freedom in parentheses. m_i is a serial correlation test of order *i* using residuals in first differences, which are asymptotically distributed as N(0,1) under the null hypothesis of no serial correlation. Hansen is a test of the over-identifying restrictions, which are asymptotically distributed as χ^2 under the null hypothesis of no correlation between the instruments and the error term, with the degrees of freedom in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

TABLE 7. Impact of the financial crisis and the specialization on the banks' PD

Variables	PD (1)	PD (2)
Lagged dependent	0.30428***	0.29959***
	(0.05044)	(0.04923)
C1	-0.00018***	-0.00016**
	(0.00006)	(0.00007)
C2	-1.56e^{-06}***	-1.50e^{-06}***
	(1.82e^{-07})	(1.88e^{-07})
C3	-0.00017***	-0.00019***
	(0.00004)	(0.00004)
AQ1	-0.00003**	-0.00003**
	(0.00001)	(0.00002)
AQ2	-6.41e^{-07}	-8.72e^{-07}
	(4.73e^{-06})	(4.57e^{-06})
ME1	-0.00070**	-0.00066***
	(0.00025)	(0.00024)
ME2	0.00001**	0.00001**
	(5.56e^{-06})	(5.51e^{-06})
ME3	3.99e^{-07}	-1.80e^{-07}
	(0.00001)	(0.00001)
L1	6.38e^{-07}	1.11e^{-06}
	(3.07e^{-06})	(3.14e^{-06})
L2	0.00001	0.00001
	(0.00001)	(0.00001)
L3	0.00002	0.00002*
	(0.00001)	(0.00001)
Size	-0.00032***	-0.00028***
	(0.00009)	(0.00009)
H-I	-0.00746**	-0.00656**
	(0.00240)	(0.00262)
GDP	8.45e^{-06}	-4.28e^{-06}
	(0.00005)	(0.00006)
Inflation	-0.00019*	-0.00016
	(0.00010)	(0.00011)
Interest	0.00018**	0.00019**
	(0.00008)	(0.00009)
Unemploy	0.00005**	0.00005**
	(0.00002)	(0.00002)
Crisis	0.00532***	0.00496***
	(0.00103)	(0.00112)
Commercial Banks		0.00006
		(0.00005)
Saving Banks		0.00023
		(0.00019)
z_1	35.12 *(14, 96)*	34.37 *(14, 96)*
z_2	4.37 *(5, 96)*	3.98 *(5, 96)*
z_3		1.58 *(2, 96)*
m_1	-2.74	-2.73
m_2	-0.64	-0.58
Hansen	78.96 *(222)*	83.40 (*222*)
Number of Obs.	474	474
Number of Banks	97	97

Notes: This table shows the two-step system-GMM estimator developed by Arellano and Bover (1995) and Blundell and Bond (1998). The dependent variable is the bank´s PD. See Table 1 for a complete description of the explanatory variables. In addition, we add a set of dummy variables to control for the crisis period and the specialization of the entities (commercial bank, saving bank and credit cooperative). All variables are considered endogenous except for the macroeconomic variables and the dummy variables. The heteroskedasticity-consistent asymptotic standard errors are reported in parentheses. z_1, z_2 and z_3 are the Wald tests of the joint significance of the reported coefficients of the explanatory variables, the macroeconomic variables and the specialization dummy variables, respectively. These statistics are asymptotically distributed as F under the null hypothesis of no significance with the degrees of freedom in parentheses. m_i is a serial correlation test of order i using residuals in first differences, which are asymptotically distributed as N(0,1) under the null hypothesis of no serial correlation. Hansen is a test of the over-identifying restrictions, asymptotically distributed as χ^2 under the null hypothesis of no correlation between the instruments and the error term, with degrees of freedom in parentheses. *, **, *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

TABLE 8. Robustness test: Controlling for the European Bank Union

Variables	PD
Lagged dependent	0.29871***
	(0.04784)
C1	-0.00018***
	(0.00007)
C2	-1.49e^{-06}***
	(1.95e^{-07})
C3	-0.00019***
	(0.00004)
AQ1	-0.00003*
	(0.00001)
AQ2	-7.44e^{-07}
	(4.77e^{-06})
ME1	-0.00067***
	(0.00025)
ME2	0.00001**
	(5.55e^{-06})
ME3	-2.31e^{-06}
	(0.00001)
L1	-3.66e^{-07}
	(4.04e^{-06})
L2	0.00002
	(0.00001)
L3	0.00003*
	(0.00002)
Bank Union	0.00002
	(0.00012)
Size	-0.00028**
	(0.00011)
HI	-0.00760***
	(0.00255)
GDP	3.74e^{-06}
	(0.00005)
Inflation	-0.00016
	(0.00012)
Interest	0.00020**
	(0.00008)
Unemploy	0.00005*
	(0.00003)
Commercial Banks	0.00004
	(0.00008)
Saving Banks	0.00016
	(0.00017)
Constant	0.00992***
	(0.00236)
z_1	25. 47 *(14, 96)*
z_2	4.32 *(5, 96)*
z_3	0.44 *(2, 96)*
m_1	-2.73
m_2	-0.67
Hansen	77.20 *(223)*
Number of Obs.	474
Number of Banks	97

Notes: This table shows the two-step system-GMM estimator developed by Arellano and Bover (1995) and Blundell and Bond (1998). The dependent variable is the bank´s PD. See Table 1 for a complete description of the explanatory variables. In addition, we add a set of dummy variables to control for the bank union (2014-2016 period) and the specialization of the entities (commercial bank, saving bank and credit cooperative). All variables are considered endogenous except for the macroeconomic variables and the dummy variables. The heteroskedasticity-consistent asymptotic standard errors are reported in parentheses. z_1, z_2 and z_3 are the Wald tests of the joint significance of the reported coefficients of the explanatory variables, the macroeconomic variables and the specialization dummy variables, respectively. These statistics are asymptotically distributed as F under the null hypothesis of no significance with the degrees of freedom in parentheses. m_i is a serial correlation test of order i using residuals in first differences, which are asymptotically distributed as N(0,1) under the null hypothesis of no serial correlation. Hansen is a test of the over-identifying restrictions, which is asymptotically distributed as χ^2 under the null hypothesis of no correlation between the instruments and the error term, with degrees of freedom in parentheses. *, **, *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

TABLE 9. Robustness test: Alternative econometric methodologies

Variables	**(1) Pooled OLS PD**	**(1) Fixed-effects PD**	**(1) Difference GMM PD**	**(2) Pooled OLS PD**	**(2) Fixed-effects PD**	**(2) Difference GMM PD**
Lagged dependent	-	-	0.16115***	-	-	0.14578*
			(0.06377)			*(0.07711)*
C1	-0.00024***	-0.00050***	-0.00016***	-0.0003***	-0.00048***	-0.00011
	(0.00007)	*(0.00012)*	*(0.00026)*	*(0.00008)*	*(0.00012)*	*(0.00028)*
C2	$-1.87e^{-06}$***	$-1.70e^{-06}$**	-1.47^{e-06}***	$-2.34e^{-06}$***	$-1.66e^{-06}$**	$-1.38e^{-06}$***
	($2.51e^{-07}$)	*($7.34e^{-07}$)*	*(5.19^{e-07})*	*($3.65e^{-07}$)*	*(7.13e-07)*	*($4.03e^{-07}$)*
C3	-0.00030***	-0.00044***	-0.00050**	-0.00029***	-0.00047***	-0.00054***
	(0.00006)	*(0.00007)*	*(0.00020)*	*(0.00006)*	*(0.00007)*	*(0.00017)*
AQ1	$1.65e^{-06}$	-0.00019***	-0.00023**	-0.00003*	-0.00014***	-0.00015*
	(0.00001)	*(0.00003)*	*(0.00010)*	*(0.00001)*	*(0.00003)*	*(0.00008)*
AQ2	$-3.60e^{-06}$	$-8.00e^{-06}$	-0.00002	$-5.40e^{-06}$	$-3.29e^{-06}$	-0.00001
	($4.24e^{-06}$)	*($4.94e^{-06}$)*	*(0.00001)*	*($3.98e^{-06}$)*	*($4.89e^{-06}$)*	*(0.00001)*
ME1	-0.00054***	0.00005	-0.00037	-0.00040***	0.00014	-0.00032
	(0.00017)	*(0.00014)*	*(0.00033)*	*(0.00015)*	*(0.00014)*	*(0.00024)*
ME2	$-2.17e^{-06}$	$-2.78e^{-06}$	5.42^{e-06}	$-3.52e^{-06}$	$-5.11e^{-06}$*	$3.05e^{-06}$
	($2.97e^{-06}$)	*($2.77e^{-06}$)*	*(4.91^{e-06})*	*($3.88e^{-06}$)*	*($2.73e^{-06}$)*	*($6.66e^{-06}$)*
ME3	-0.00003***	-0.00001	-7.31^{e-06}	-0.00002***	$-9.74e^{-06}$	-0.00001
	($7.91e^{-06}$)	*($9.78e^{-06}$)*	*(0.00002)*	*($7.29e^{-06}$)*	*($9.49e^{-06}$)*	*(0.00003)*
L1	-0.00001*	-0.00001	-2.65^{e-06}	$-2.96e^{-06}$	$-3.93e^{-06}$	-0.00001
	($5.59e^{-06}$)	*($8.73e^{-06}$)*	*(0.00002)*	*($5.72e^{-06}$)*	*($9.01e^{-06}$)*	*(0.00002)*
L2	$4.96e^{-06}$	0.00005	0.00001	0.00001	0.00002	$1.74e^{-06}$
	(0.00001)	*(0.00004)*	*(0.00005)*	*(0.00001)*	*(0.00004)*	*(0.00006)*
L3	$4.69e^{-06}$	0.00001	-4.04^{e-06}	0.00003**	$6.29e^{-07}$	-0.00004
	(0.00002)	*(0.00003)*	*(0.00008)*	*(0.00001)*	*(0.00003)*	*(0.00008)*
Size				-0.00030**	-0.00083	-0.00003
				(0.00012)	*(0.00073)*	*(0.00238)*
GDP				0.00003	0.00002	0.00004
				(0.00005)	*(0.00006)*	*(0.00007)*
Inflation				0.00006	0.00027**	-0.00003
				(0.00015)	*(0.00012)*	*(0.00014)*

Interest				0.00034*** *(0.00008)*	-0.00010 *(0.00011)*	0.00011 *(0.00020)*
Unemploy				0.00007* *(0.00004)*	-0.00020** *(0.00008)*	-0.00013 *(0.00013)*
HI				-0.00763*** *(0.00265)*	-0.04064*** *(0.00945)*	-0.03690 *(0.03133)*
Constant	0.01441*** *(0.00158)*	0.01494*** *(0.00271)*	-	0.01556*** *(0.00302)*	0.03519*** *(0.01288)*	-
Year Dummies	Yes	Yes	Yes	No	No	No
R^2	0.3930	0.4214 *(within)*	-	0.4387	0.4553 *(within)*	-
z_1	9.46 *(11, 96)*	23.27 *(11, 458)*	14.39 *(12, 97)*	9.24 *(12, 96)*	17.88 *(12, 457)*	13.19 *(13, 97)*
z_2	3.34 *(5, 96)*	1.12 *(5, 458)*	0.97 *(4, 97)*	13.98 *(5, 96)*	5.81 *(5, 457)*	1.21 *(5.97)*
Hausman	-	97.96 *(16)*	-	-	125.84 *(15)*	-
m_1	-	-	-2.26	-	-	- 2.19
m_2	-	-	0.79	-	-	- 1.14
Hansen	-	-	39.64 *(34)*	-	-	42.31 *(39)*
Number of Obs.	571	571	377	571	571	377
Number of Banks	97	97	97	97	97	97

Notes: This table shows different methods of estimation of our main equation. The dependent variable is the bank´s PD. See Table 1 for a complete description of the explanatory variables. Models (1) include as determinants only the CAMEL indicators while Models (2) include also the structural and the macroeconomic variables. The first estimation employs Ordinal Least Squares (OLS) with robust standard errors clustered by banks (in parentheses). The second estimation uses a fixed-effects (within) regression, the standard errors are shown in parentheses. R2 is the proportion of variation in the dependent variable explained by the model. Hausman test compares the fixed versus random effects, asymptotically distributed as χ^2 under the null hypothesis that the individual effects are uncorrelated with the other regressors in the model, with degrees of freedom in parentheses. The third estimation employs the two step difference-GMM developed by Arellano and Bond (1991). Collapse option is used in case a reduction of the number of instruments is necessary. All variables are considered as endogenous except for the structural and macroeconomic variables. Robust standard errors are reported in parentheses. The z_1 y z_2 are Wald tests of the join significance of the reported coefficients of the explanatory variables and the year dummies (or macroeconomic variables), respectively. These statistics are asymptotically distributed as *F* under the null hypothesis of no significance with degrees of freedom in parentheses. m_i is a serial correlation test of order *i* using residuals in first differences, asymptotically distributed as N(0,1) under the null hypothesis of no serial correlation. Hansen is a test of the over-identifying restrictions, which is asymptotically distributed as χ^2 under the null hypothesis of no correlation between the instruments and the error term, with degrees of freedom in parentheses. *****, ****** and ******* indicate statistical significance at the 10%, 5% and 1% levels, respectively.

TABLE 10. Robustness test: Alternative definition of the dependent variable

Variables	**Log (PD/(1-PD)) (1)**	**Log (PD/(1-PD)) (2)**
Lagged dependent	0.14261***	0.09928**
	(0.04579)	(0.04250)
C1	-0.08322***	-0.10247***
	(0.02053)	(0.02008)
C2	-0.00010	-0.00019
	(0.00069)	(0.00043)
C3	-0.13382***	-0.14690***
	(0.01783)	(0.01718)
AQ1	0.00574	-0.00245
	(0.00356)	(0.00394)
AQ2	0.00084	-0.00041
	(0.00091)	(0.00109)
ME1	-0.08181**	-0.07141**
	(0.03817)	(0.02974)
ME2	0.00112*	0.00100***
	(0.00062)	(0.00040)
ME3	-0.00164	-0.00252
	(0.00201)	(0.00214)
L1	-0.00132	0.00083
	(0.00097)	(0.00111)
L2	0.00118	0.00021
	(0.00297)	(0.00318)
L3	0.00218	0.00594*
	(0.00306)	(0.00356)
Size		-0.09077***
		(0.03245)
H-I		-1.73755*
		(1.01074)
GDP		0.01310
		(0.00995)
Inflation		-0.06746**
		(0.02750)
Interest		0.03811***
		(0.01207)
Unemploy		0.02261***
		(0.00814)
Constant		-1.23269*
		(0.70881)
Year dummies	Yes	No
z_1	36.67 *(12, 95)*	51.14 *(13, 95)*
z_2	14.55 *(5, 95)*	5.10 *(5, 95)*
m_1	-2.50	-2.30
m_2	0.61	0.44
Hansen	82.37 *(201)*	79.18 *(222)*
Number of Obs.	460	460
Number of Banks	96	96

Notes: This table shows the two-step system-GMM estimator developed by Arellano and Bover (1995) and Blundell and Bond (1998). The dependent variable is the odds ratio log (PD/(1-PD)). See Table 1 for a complete description of the explanatory variables. Model (1) presents the baseline model. Model (2) includes macroeconomic variables as explanatory variables. All the variables are considered endogenous except for the macroeconomic variables and the dummy variables. The heteroskedasticity-consistent asymptotic standard errors are reported in parentheses. z_1 and z_2 are the Wald tests of the joint significance of the reported coefficients of the explanatory variables and the year dummies (or macroeconomic variables), respectively. These statistics are asymptotically distributed as F under the null hypothesis of no significance with the degrees of freedom in parentheses. m_i is a serial correlation test of order i using residuals in first differences, which are asymptotically distributed as N(0,1) under the null hypothesis of no serial correlation. Hansen is a test of the over-identifying restrictions, asymptotically distributed as χ^2 under the null hypothesis of no correlation between the instruments and the error term, with degrees of freedom in parentheses. *, **, *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Appendix A. Correlation and collinearity analyses

Table A.1. Correlation matrix

	PD	C1	C2	C3	L1	L2	L3	AQ1	AQ2
PD	1.0000								
C1	-0.4203 *(0.0000)*	1.0000							
C2	-0.3316 *(0.0000)*	0.2680 *(0.0000)*	1.0000						
C3	-0.5465 *(0.0000)*	0.5257 *(0.0000)*	0.4402 *(0.0000)*	1.0000					
L1	0.0795 *(0.0575)*	-0.2945 *(0.0000)*	-0.1221 *(0.0035)*	-0.1795 *(0.0000)*	1.0000				
L2	0.0532 *(0.2042)*	0.1521 *(0.0003)*	-0.0828 *(0.0480)*	-0.2375 *(0.0000)*	0.3519 *(0.0000)*	1.0000			
L3	-0.1178 *(0.0048)*	0.1066 *(0.0108)*	0.0659 *(0.1159)*	0.3029 *(0.0000)*	-0.2732 *(0.0000)*	-0.5922 *(0.0000)*	1.0000		
AQ1	0.0015 *(0.9716)*	0.2817 *(0.0000)*	-0.0700 *(0.0947)*	0.0037 *(0.9292)*	-0.1324 *(0.0015)*	0.1134 *(0.0067)*	-0.0279 *(0.5054)*	1.0000	
AQ2	-0.0146 *(0.7277)*	0.0300 *(0.4742)*	-0.0327 *(0.4355)*	0.0227 *(0.5891)*	-0.0825 *(0.0488)*	-0.0993 *(0.0177)*	0.1524 *(0.0003)*	-0.2584 *(0.0000)*	1.0000
M1	-0.2951 *(0.0000)*	0.2306 *(0.0000)*	0.1783 *(0.0000)*	0.2926 *(0.0000)*	-0.1872 *(0.0000)*	-0.0166 *(0.6916)*	0.0645 *(0.1236)*	-0.3257 *(0.0000)*	-0.0143 *(0.7332)*
M2	-0.0848 *(0.0429)*	0.0347 *(0.4084)*	0.0418 *(0.3188)*	0.0736 *(0.0788)*	-0.0444 *(0.2893)*	-0.0252 *(0.5486)*	0.0413 *(0.3244)*	-0.1383 *(0.0009)*	-0.0151 *(0.7195)*
M3	0.0288 *(0.4920)*	-0.1073 *(0.0103)*	-0.0407 *(0.3316)*	-0.1394 *(0.0008)*	-0.0377 *(0.3691)*	-0.1343 *(0.0013)*	0.0248 *(0.5538)*	0.0530 *(0.2064)*	-0.0797 *(0.0571)*
Size	0.1765 *(0.0000)*	-0.6540 *(0.0000)*	-0.1697 *(0.0000)*	-0.2492 *(0.0000)*	0.4856 *(0.0000)*	-0.1543 *(0.0002)*	-0.1597 *(0.0001)*	-0.1857 *(0.0000)*	-0.1012 *(0.0156)*
GDP	-0.1802 *(0.0000)*	0.0857 *(0.0407)*	0.0843 *(0.0442)*	0.0744 *(0.0758)*	-0.1262 *(0.0025)*	0.0733 *(0.0803)*	-0.0150 *(0.7206)*	-0.1107 *(0.0081)*	0.0524 *(0.0081)*
Inflation	0.1885 *(0.0000)*	-0.0402 *(0.3382)*	-0.0584 *(0.1631)*	-0.1740 *(0.0000)*	0.0874 *(0.0369)*	0.0285 *(0.4969)*	-0.0173 *(0.6803)*	-0.2351 *(0.0000)*	0.0373 *(0.0000)*
Interest	0.3195 *(0.0000)*	-0.0534 *(0.2028)*	-0.0539 *(0.1986)*	-0.1727 *(0.0000)*	0.0914 *(0.0289)*	0.0917 *(0.0285)*	-0.3064 *(0.0000)*	0.2129 *(0.0000)*	-0.0265 *(0.0000)*
Unemployment	0.2179 *(0.0000)*	-0.1639 *(0.0001)*	-0.0658 *(0.1163)*	-0.1113 *(0.0077)*	0.1080 *(0.0098)*	0.0670 *(0.1100)*	-0.3335 *(0.0000)*	0.2573 *(0.0000)*	-0.0483 *(0.0000)*
H-I	-0.1224 *(0.0034)*	0.1843 *(0.0000)*	0.0078 *(0.8519)*	0.1290 *(0.0020)*	0.0525 *(0.2101)*	0.2532 *(0.0000)*	-0.0349 *(0.4047)*	0.2163 *(0.0000)*	0.0649 *(0.0000)*

Note: This table presents the Pearson´s correlation coefficient for each pair of variable and its statistical significance (p-value in parentheses).

	M1	M2	M3	Size	GDP	Inflation	Interest	Unemployment	H-I
M1	1.0000								
M2	0.3115 *(0.0000)*	1.0000							
M3	-0.3860 *(0.0000)*	-0.1801 *(0.0000)*	1.0000						
Size	-0.1602 *(0.0001)*	-0.0354 *(0.3986)*	-0.0137 *(0.7445)*	1.0000					
GDP	0.3035 *(0.0000)*	0.0298 *(0.4776)*	-0.0769 *(0.0663)*	-0.0500 *(0.2325)*	1.0000				
Inflation	-0.0544 *(0.1939)*	0.0016 *(0.9701)*	0.1377 *(0.0010)*	-0.0624 *(0.1362)*	-0.1794 *(0.0000)*	1.0000			
Interest	-0.2863 *(0.0000)*	0.0107 *(0.7984)*	-0.0363 *(0.3865)*	0.1236 *(0.0031)*	-0.4187 *(0.0000)*	0.3343 *(0.0000)*	1.0000		
Unemployment	-0.2329 *(0.0000)*	-0.1000 *(0.0169)*	-0.0557 *(0.1840)*	0.3030 *(0.0000)*	-0.2983 *(0.0000)*	-0.1020 *(0.0147)*	0.5425 *(0.0000)*	1.0000	
H-I	-0.0694 *(0.0976)*	-0.0048 *(0.9080)*	-0.0358 *(0.3935)*	-0.2757 *(0.0000)*	-0.0206 *(0.6241)*	-0.1192 *(0.0043)*	0.0747 *(0.0746)*	0.1223 *(0.0034)*	1.0000

Note: This table presents the Pearson´s correlation coefficient for each pair of variable and its statistical significance (p-value in parentheses).

Table A.2. VIF Analysis

Variable	**VIF**	**1/VIF**
Size	3.09	0.3232
C1	3.00	0.3333
M1	2.88	0.3475
L2	2.75	0.3631
Interest	2.31	0.4324
L3	2.22	0.4514
Unemployment	2.15	0.4650
AQ1	2.13	0.4690
L1	2.10	0.4757
C3	2.06	0.4857
M2	1.88	0.5312
Lagged PD	1.62	0.6191
GDP	1.45	0.6884
M3	1.44	0.6921
Inflation	1.44	0.6956
H-I	1.38	0.7244
AQ2	1.32	0.7574
C2	1.26	0.7913
Mean VIF	**2.03**	

Note: This table presents the Variance Inflation Factor (VIF) for the estimated coefficients.