# European Banks in the XXI century: are their Business Models sustainable?

Emanuele De Meo*, Annalisa De Nicola°, Giuseppe Lusignani**, Federica Orsini* and Lea Zicchino*

## Abstract

Banks' activity is challenged by the macroeconomic, financial and regulatory environment emerging from the financial crisis. Lower rates of economic growth, low interest rates, more stringent regulation on capital and liquidity, the need to decrease leverage for a number of banks, increased market scrutiny following the new rules on crisis management have put pressure on pre-crisis bank business models. In this paper we introduce a new methodology to identify different business models and group European banks according to them. We then analyse the relationship between business models and a set of bank performance indicators and how this relationship has changed over time.

## Table of contents

1 Introduction .... 2
2 Bank Business Models: review of the literature .... 4
3 Data and methodology .... 5
3.1 Identification of Bank Business Models and Peer Groups .... 5
3.2 The dataset .... 6
4 Results .... 9
4.1 Identified Bank Business Models and Peer Groups .... 9
4.2 The performance of different Bank Business Models and Peer Groups .... 10
5 Conclusions .... 13
References .... 15
Appendix .... 16
A.1 Deterministic clustering .... 16
A.2 Probabilistic clustering .... 17
A.3 Figures and tables .... 19

---

*Prometeia.
°The author was affiliated to Prometeia while working on the paper.
**University of Bologna and Prometeia.

# 1 Introduction

In the years before the financial crisis, expansionary monetary policies and accommodative banking regulation were the main drivers of excessive risk taking by the banking sector. This in turn led to a progressive shift of banks away from their traditional role of intermediaries; declining interest rates and incorrect pricing of risk associated to traditional assets and new complex financial securities fuelled a massive search for yield. Several banks adopted unsustainable business models; their failure triggered the financial crisis and the subsequent overhaul of international banking regulation and supervision.

In response to the financial crisis, banks have made efforts to reshape their business models (see ECB, 2016). Three factors were at play in this process: 1) regulatory reforms requiring banks to hold more capital of higher quality and, indirectly, forcing banks to adapt their operating structures to new requirements; 2) financial market pressures, giving an incentive to managers to exit low-margin activities and boost returns; 3) changes in banks' own preferences of the optimal risk-return profile.

As a result, banks have adopted a number of decisions concerning their business strategy.[1] First, they are strengthening their core business activities, scaling back risky assets and shifting towards retail businesses from trading and wholesale activities. Second, they are reducing their leverage and the share of wholesale funding. Third, a significant number of EU banks are reducing their exposure to peripheral markets and, in general, their international presence. Fourth, banks are diversifying their income sources in response to the low interest rates and credit demand environment, as well as pursuing higher efficiency by means of a reduction in operational costs. Finally, as regulation may diminish the benefits deriving from economies of scale and complexity, they banks are streamlining their processes.

Different Bank Business Models (BBMs) may have different degrees of resilience to the new macro-financial environment. Therefore, we ask ourselves whether the current banking business models are still sustainable.

The banking supervisory authority operating under the Single Supervisory Mechanism (SSM) needs to answer this same question within the new Supervisory Review and Evaluation Process (SREP). Under EBA Guideline,[2] the SREP framework consists of four evaluation areas (business models, governance and control systems, capital risk and liquidity risk). In particular, business model analysis (BMA) aims at assessing a bank's ability to achieve satisfactory short-term profits (within a 12-month horizon), and medium-term sustainability (over a 3-year horizon). Based on the final score assigned to a bank's business model, the supervisory body may require a bank to adjust its capital base, liquidity or take management actions (including the improvement of their risk management functions, the reduction of the risks associated to the products distributed to clients, and changes to their business model and strategies).

In this paper we assess the characteristics of different business models of banks (BBMs) operating in Europe, their economic sustainability based on their ability to achieve satisfactory profits, and the factors affecting their relative performance and risk.

[1] See "Risk assessment of the European banking system", European Banking Authority, June 2015.
[2] See EBA/GL/2014/13.

This paper investigates four main research issues: 1) how can we identify the BBMs operating in Europe? 2) what are the business strategies specific to each BBM? 3) how profitable were the different business models in the last decade? 4) which macro-financial factors and individual banks' characteristics drive profitability and risk the most at the business model level?

We extend the empirical literature on BBMs along several dimensions. First, we use probabilistic, along with deterministic, statistical learning techniques to identify BBMs in Europe. At the best of our knowledge, ours is the first attempt to use both approaches to identify different business strategies in the banking sector. Second, we use *supervised* learning techniques to identify homogeneous peer groups of banks according to the criteria outlined by the EBA in the SREP Guidelines for the categorization of financial institutions. Third, we develop a panel data framework to identify the relevant macro-financial factors and the individual banks' characteristics that explain the heterogeneity of bank profitability at the business model level.

To anticipate some of the results, we get the following evidence.

First, we identify three different BBMs: *Retail banks*, mainly financing loans through deposits; *Investment banks*, with a higher incidence of financial assets and derivatives and a significant exposure to less stable funding sources; *Diversified banks*, which rely on diversified funding sources, have a non-negligible trading activity and a significant exposure to customers (firms and households) in traditional intermediation activities.

Second, we form Peer Groups (PGs) within the three different BBMs based on EBA criteria for identifying systemic relevance, operational complexity, business specialization and cross-border activity. We obtain that Non-Complex Retail Domestic banks represent the largest PG and are mostly located in Italy, followed by Portugal. The second largest PG is the Complex Retail Domestic one, with banks mostly located in Italy and Germany.

Third, coming to the performance of different business models, we show that Retail banks were the best performers prior to the financial crisis. However, at the peak of the Sovereign Debt Crisis they had the worst RoA and experienced the most significant deterioration in credit quality, only partially offset by improved cost efficiency. Moreover, Retail banks that are operationally complex and international performed significantly better before the start of the financial crisis, while Non-Complex Retail banks (both domestic and international), i.e. banks with a less diversified sources of income and therefore more hit by the decline in credit demand and by the deterioration in credit quality, experienced the lowest RoA from 2011 to 2014.

Finally, we identify the relevant risk factors at the business model level. Economic growth, the yield curve level and slope and sovereign default risk are the main factors affecting bank profitability; among bank characteristics, credit quality is a positive driver of Retail banks RoAs, whereas exposure to financial markets, proxied by equity market returns, is a significant driver for non-Retail banks. Finally, at least for non-Retail banks, there is some evidence that holding more capital is not harmful to bank profitability.

The rest of the paper is organised as follows. In section 2 we review the empirical literature on Bank Business Models; Section 3 is dedicated to the identification of BBMs and PGs; in Section 4 we outline the evidence on the performance of different BBMs and PGs; Section 5 contains the Conclusions. The Appendix contains details on our methodological approach, the complete list of banks included in the sample as well as Figures and Tables.

## 2 Bank Business Models: review of the literature

A number of academic studies have analysed the relationship between BBMs' characteristics and performance. In particular, research has focused on the interaction between bank performance and a series of factors: bank capital, funding structure, efficiency, diversification, exposure to financial markets and corporate governance.[3] Altunbas et al. (2011) found that: a) managerial strategies and abilities matter in explaining return and risk at the bank level; b) higher credit growth, lower dependence on customer deposits, larger size and lower capital account for higher bank risk, especially in the run-up to the crisis.

In the existing literature, BBMs have been mainly identified using statistical learning techniques.[4] Previous contributions tend to differ more because of the sample of financial institutions considered and/or the indicators according to which banks are grouped than in their methodological approaches.

Ayadi and De Groen (2015) exploit balance sheet indicators and income statements of a sample of 2.542 European banks and subsidiaries. They use cluster analysis on a set of indicators that should represent a proxy for banks' business strategies (loans to customers and banks, trading assets and derivatives, debt liabilities). They identify five BBMs: *Focused Retail* banks; *Diversified Retail* banks (*Type I* and *Type II*, distinguished by the weight of trading activities); *Wholesale* banks; *Investment* banks. In short, Focused and Diversified Retail banks tend to show higher profitability (RoA) in the years preceding the financial crisis, but substantially lower RoA during the peak of the Sovereign Debt Crisis (higher customers' default rates and contractionary business cycles reduced the profitability of traditional intermediation and determined higher loan loss provisions).

Roengptya et al. (2014) adopt a similar strategy on a sample of 222 individual banks from 34 countries. Clustering techniques are applied to a set of balance sheet ratios (loans, traded securities, deposits, wholesale debt, interbank activity). Banks are grouped in three categories: *Retail-funded*; *Wholesale-funded*; *Trading*. The authors find that the outbreak of the recent crisis marked a steep drop in the RoE of banks in advanced economies across all BBMs and that Trading and Wholesale-funded banks showed a higher income volatility in the aftermath of the financial crisis.

ECB (2016) also adopts clustering techniques to identify BBMs for 113 significant institutions supervised by the SSM based on: risk-weighted assets (or size), net fee and commission income as a share of operating income, customer funding and interbank funding as a share of total liabilities, trading assets and domestic exposure as a share of total assets. They identify seven BBMs and find that larger and more retail-oriented banks are generally associated with lower default risk, however with some notable differences between pre- and post-crisis periods. In particular, income diversification emerges as beneficial in the post-crisis period, whereas during the crisis bank size was no longer an effective shield against increasing default risk.

---

[3] See Altunbas et al. (2011) for an extensive review.
[4] See section 3 for a formal definition of the notion of statistical learning in section 3.

Bonaccorsi di Patti et al. (2016) identify bank business models for 112 European individual banks under the supervision of the SSM. The identification approach relies on the following criteria: specialization; size; core business; share of cross border exposure. They identify eight business models: *Lending banks* (4 groups of banks with loan-to-assets ratios above 50 per cent, divided according to their size and the relevance of cross border activities); *Diversified banks* (in which traditional banking is less relevant, divided in Large and Smaller banks); *Network banks* (acting as hubs for small local banks); *Public* and *Development banks* (banks having a public interest purpose). Their main findings are that small Lending banks have higher revenues and operating costs; Lending banks have a higher exposure to country-specific conditions as proxied by the business cycle; and finally small and medium Lending banks are characterized by higher credit risk density.

# 3 Data and methodology

In this section we describe the methodology behind our identification strategy and the dataset.

## 3.1 Identification of Bank Business Models and Peer Groups

We identify BBMs and PGs with a two-step approach in which we use different statistical learning methods depending on our prior knowledge on banking activity at the individual bank's level.[5]

A key aspect of our analysis is the distinction between *unsupervised* and *supervised* learning.[6] In particular, we use unsupervised learning to identify BBMs, assuming that no prior knowledge on the actual number and composition of business models operating in Europe is available. Among unsupervised learning methods, cluster analysis is a convenient tool as it allows to establish the number of clusters (BBMs) and their composition. The analysis is conducted on balance sheet indicators, which should effectively summarise banks' business strategies. We opt for hierarchical clustering to perform banks' aggregation at this level of analysis. We then extend the existing literature on BBMs by using probabilistic, along with deterministic, statistical grouping techniques, as in principle a bank's business model could be multi-faceted and the deterministic attribution of each bank to a single cluster could be too restrictive. At the best of our knowledge, ours is the first attempt to use both approaches to identify distinct business strategies in the banking sector (see the Appendix A.1 and A.2 for technical details).

In the second step, we group banks following the EBA Guidelines on the supervisory review and evaluation process (SREP). In the Guidelines, EBA groups financial institutions in four categories according to the following characteristics: systemic relevance; dimension; organizational complexity; cross border activity. At this level of analysis, we assume that some prior knowledge about individual banks' complexity,

---

[5] Generally speaking, given a set of characteristics X for a sample of observations, statistical learning involves deriving an output Y = f(X) from a model f(.) given the observed values of X in our sample. In the context of BBMs identification, banks are the observations, X are banks' characteristics, f(.) is some grouping or classification algorithm and Y is a set of group labels (the business models).

[6] Unsupervised learning refers to the situation in which only observations on X are available while Y are not. On the contrary, supervised learning refers to a situation in which observations on Y are partially or completely available (see James et al., 2013).

specialization and cross-border activity is available to the researcher. Therefore, PG identification can be conducted with supervised learning. Among supervised learning methods, *Key nearest neighbor* (KNN) allows prior knowledge about the number of existing peer groups, the label (name) of each group and the membership of a restricted set of banks to each group. This prior knowledge comes from a combination of qualitative (EBA Guidelines, Financial Stability Board's list on systemic relevance) and quantitative (balance sheets, EBA stress tests data) information.

In the KNN jargon, banks on which we have prior knowledge about PG membership are called *training* banks, whereas banks that we need to classify are called *test* banks. For each bank, KNN computes a distance measure according to a set of individual characteristics and classifies each bank as belonging to different PGs according to a majority rule: in other words, given a training set of banks, a bank belongs to the peer containing the majority of training peers which are closest according to some distance measure.

## 3.2 The dataset

We use public consolidated banking data provided by SNL Financials. The sample covers the period from 2006 to 2014, spanning different economic cycles and including the 2007-'08 financial crisis and the Sovereign Debt crisis. We use annual data on 77 banking groups (Table 1), most of which were subject to the 2014 EBA EU-wide stress tests and defined as significant within the SSM. We select 73 banks from 14 countries in the Euro area, based on their relevance and availability of published balance sheets over the sample period, accounting for around 82% of the EMU's total banking assets, plus 4 UK banking groups. Italy and Germany account for the majority of banks in the sample, given the fragmentation of their banking sector.

In order to identify BBMs, we use a set of balance sheet variables that should reflect strategic management choices and summarise banks' business strategies. They relate to the asset and the liability side of banks' balance sheets and are expressed as ratios to total assets. These variables capture BBM's key banking activities, funding strategies and financial exposure.

- Customer loans: the share of total loans to customers, net of reserves for loan losses. High values, typical for traditional commercial banking models, indicate greater reliance on traditional sources of income (net interest income).
- Interbank loans: the share of loans to other financial institutions. It is a proxy of interbank and wholesale activities.
- Financial assets: the sum of debt and equity instruments, derivatives and other financial assets. A high share of financial assets would indicate the prevalence of investment activities.
- Customer deposits: they refer to retail and corporate deposits. The higher the customer deposits over total assets, the more stable the bank funding.
- Interbank liabilities: they include liabilities towards other banks and central banks. It is a measure of the reliance on short-term funding.

- Derivative liabilities: total negative replacement values of hedging and non-hedging derivatives. Given their nature, derivative instruments are usually the riskiest exposures of banks with higher share of trading activities.

- Debt liabilities: the ratio provides a general measure of a bank's exposure to market funding, mostly longer-term funding. It includes senior and subordinated debt and it is obtained by netting customer deposits, bank deposits and derivative exposure from total financial exposure.

The PG analysis is based on the criteria provided by the EBA in the Guidelines on the supervisory review and evaluation process for the categorization of institutions (SREP).

We classify banks in the following PGs:

a) Systemic Investment banks;
b) Non Systemic Investment banks;
c) Diversified and Non-Specialized banks;
d) Diversified and Specialized banks;
e) Systemic Retail banks;
f) Complex and International Retail banks;
g) Complex and Domestic Retail banks;
h) Non-Complex and International Retail banks;
i) Non-Complex and Domestic Retail banks.

Therefore, banks are classified according to the following aspects: Systemic relevance (SR), Cross Border activities (CB), Business and Organisational Complexity (CO) and Specialisation (SP).

Systemically important institutions, as identified by the Financial Stability Board, are ruled out from the automatic classification procedure. The KNN algorithm is applied to banks belonging to the same BBM (Retail or Diversified). Investment banks are also ruled out from the procedure, as only six Investment banks are non-systemically important.

In more detail:

- **Systemic relevance (SR)** is defined by the Financial Stability Board. We use the FSB's list of G-SIBs as of November 2015.

- **Business and Organizational Complexity (CO)** can be measured by a combination of two Herfindahl indices:

  a. the HCBL index measures the concentration of operating income across different business lines. Formally, the HCBL index is computed as the sum of squares of the individual business lines' operating income as a percentage of the total operating income; data refer to the segment performance analysis as reported in the 2014 banks' annual reports:

$$\mathrm{HCBL}_i = \sum_{l=1}^{L_i} k_{i,l}^2$$

where $i$ is the bank, $L_i$ is the number of business lines in the bank and $k_{i,l}$ is the share of business line l's operating income on the total operating income of bank $i$. The more concentrated is a bank 's exposure, the higher the HCBL index. The index approaches a maximum value of 1 when total operating income tends to come from just one business line.

b. the HCCR index measures the concentration of counterparties' credit risk exposure as a percentage of total credit risk. Formally:

$$\mathrm{HCCR}_i = \sum_{c=1}^{C_i} v_{i,c}^2$$

where $i$ is the bank, $C_i$ is the total number of sectors to which the bank is exposed and $v_{i,c}$ is the share of sector $c$'s exposure on total credit risk exposure in bank $i$. The higher the concentration of credit risk exposure across sectors, the higher the HCCR index. The index approaches a maximum value of 1 when a bank tends to be exposed to only one sector.

We describe higher concentration values across business lines and sectors as a situation in which the bank is operationally non complex.

- **Cross border exposure (CB)**, measured by the ratio of credit risk exposure in the home country as a percentage of total credit risk exposure; data are obtained from the EBA/ECB stress test.
- **Specialization (SP)**, measured by the exposure of the bank to sectors other than retail and corporate. Thus, we construct the ratio between credit exposures to non-retail and non-corporate counterparties and total credit risk exposure. Data come from the EBA/ECB stress test.

For Retail banks, given the *training* banks belonging to the Complex and Non-Complex (domestic and international) PGs, we follow a 2-step methodology to classify *test* banks:

- first, KNN classifies banks based on complexity, alternatively as Complex Retail banks - i.e. banks that operate in several business lines (low values of HCBL) and, at the same time, are exposed to various counterparties (low values of HCCR) - or Non-Complex Retail banks – i.e. banks operating in a limited number of business lines, offering credit products to retail and corporate customers;
- second, banks belonging to the Complex and Non-Complex groups are further split according to their cross-border exposure into banks that operate domestically (low value of CB) and banks having sizable cross-border activities (International).

Diversified banks are classified based on their specialization. Given our set of training banks, we separate banks mainly operating with the public sector and with other non-retail customers (with high values of SP) from banks mainly dealing with retail and corporate sectors.

In order to assess the ability of different BBMs and PGs to achieve satisfactory profitability levels, we use a set of indicators from banks' income statements. We focus on the main drivers of banks' profitability, such as earnings, cost efficiency and credit risk exposure. For each of these dimensions, we construct a Key Performance Indicator (KPI): the return on assets (RoA - net income to total assets), the operating cost-to-total assets (CTA - operating expenses to total assets) and the cost of credit risk (CRK - customer loan impairment charges to total net customer loans).

# 4 Results

In this section we describe the results of the statistical analysis conducted on our sample of banks. In particular, we give evidence on BBMs and PGs identified by statistical learning techniques. Then we discuss the descriptive and econometric evidence of banks' performance both at the BBM and PG levels.

## 4.1 Identified Bank Business Models and Peer Groups

In this work, the optimal number of clusters was chosen based on hypothesis testing conducted on 30 different test-statistics. In order to avoid excessive instability, we select the median number of clusters across test-statistics in every year of the analysis as optimal.[7] This criterion returns a stable optimal number of three BBMs over our full sample period (2006-2014).

BBMs characteristics can be easily summarised with a radar graph (Figure 1), as in Ayadi and De Groen (2015). For each BBM, the radar plots the median value of each balance sheet indicator across all banks belonging to that cluster, standardised across different BBMs (i.e. in Z-score). The interpretation of the three models is straightforward. One model mainly finances loans through deposits. We call this cluster Retail banks. A second cluster includes banks with a higher incidence of trading assets and derivatives but also a significant exposure to less stable sources of funding. We label these banks Investment banks. The third cluster is more of a hybrid: it has the benefits of diversified funding sources (long-term debt, interbank funding and, to a lesser extent, deposits), it has a non negligible trading activity, but also a significant exposure to customers (firms and households) in traditional intermediation activities. We call this cluster Diversified banks. In 2014 we identify 48 Retail banks, 16 Diversified banks and 13 Investment banks.

In order to overcome the limits of hard clustering (in which each bank belongs to a cluster with probability 1), and assuming the same optimal number of clusters (three), we admit uncertainty in the BBM attribution at the individual bank's level. Therefore, we estimate the probabilities that each bank belongs to a specific BBM (probabilistic, or "fuzzy", clustering). The interpretation of degrees (probabilities) of membership is twofold: first, the probabilities account for uncertainty concerning the "real" attribution of banks to a BBM; second, they can be interpreted as the weight of a particular activity (considered as the core business in each cluster; for example, trading for Investment banks) in the overall business strategy of the bank.

Identification obtained from fuzzy clustering returns almost the same picture as hard clustering: on average, from 2006 and 2014, the two methods tend to give different attributions in only 11 per cent of the cases (for fuzzy clustering, we assume that the bank belongs to the BBM that has the highest probability). The identification appears relatively more unstable during the Sovereign Debt crisis (about 20 per cent of banks from 2010 to 2012 belong to different BBMs in the two methods) and, on the contrary, more stable in 2013 and 2014 (only 8 per cent and 6 per cent of banks belong to different BBMs).

[7] The R package Nbclust provides a user-friendly interface to conduct hypothesis testing on the optimal stopping rule.

Therefore, for the purpose of BBM and PG identification and performance evaluation, we identify BBMs with fuzzy clustering by assigning each bank to its most likely business model. PG identification reflects the criteria outlined in the EBA Guidelines, as shown for non-systemic banks in Figures 2-3. We selected a restricted set of training banks and identified the PGs with the KNN algorithm. PGs are described in short in Table 2.

At the European level, among Retail banks, Non-Complex and Domestic banks represent the largest PG; about 40% of banks belonging to this PG are located in Italy, followed by Portugal. This PG also contains the largest number of banks in Italy. The second most numerous PG is the Complex and Domestic banks, mostly located in Italy and Germany. Therefore, Italian banks are mainly Retail and oriented to domestic markets, with a relevant degree of business complexity because of their credit risk exposures and dispersion of income across different business lines. Finally, our sample also includes three systemic Retail banks (one of them located in Italy).

Diversified banks are about equally distributed between Specialized and Non-Specialized banks. Germany has the highest number of Diversified banks; Diversified banks also represent the most relevant BBM in Germany.

Finally, six out of thirteen Investment banks are Non systemic.

## 4.2 The performance of different Bank Business Models and Peer Groups

In this section we present descriptive and econometric evidence on the performance of different BBMs and PGs.

Retail banks show the highest RoAs before the financial crisis but the worst performance at the peak of the Sovereign Debt crisis, as higher sovereign default risk determined lower credit demand, lower interest rates on loans, higher funding costs, higher capital requirements and lower profitability for traditional banks (Figure 4). Furthermore, during the Euro Area crisis Retail banks experienced, as one can expect, the most significant deterioration in credit quality, measured by the Cost of Risk (Figure 5). These factors were partially offset by higher cost efficiency (i.e. a declining operating expenses as a ratio of total assets), thanks to the effort made by Retail banks to sustain profitability by reducing inefficiencies and cutting structural costs (Figure 6).

Nevertheless, fuzzy clustering gives a slightly different picture in terms of KPIs across different BBMs. That is, probabilistic clustering tends to make performances of different BBMs more similar (see Appendix A.2 for the details on the computation of expected KPIs in the presence of uncertainty about the correct attribution to specific BBMs).

Turning to evidence on PGs' profitability, in 2015 Non-Complex International Retail banks experienced the highest median RoA, followed by Complex International Banks (Figure 7). That is, cross-border activities seemed to be advantageous for Retail banks at least in 2014. International Retail Banks also maintained higher credit quality. Diversified and Investment peers tended to be more efficient than Retail banks, with the lowest Operating expenses-to-assets ratio across different PGs. Finally, Non-Complex International Retail banks had higher capital ratios, meaning that they appeared healthier and with significant excess capital with respect to minimum regulatory requirements. Said differently, these banks had a relatively low exposure to regulatory risk (the risk that capital requirements could be increased in the future) and to the volatility of the cost at which the bank raise capital (the cost of capital).

We now carry out several econometric exercises on banks' profitability with two goals in mind:

a) to investigate the relative performances of different BBMs and PGs. We do this using cross-section and pooling panel data regressions;

b) to identify relevant risk factors at the BBM level. To do this we use dynamic panel regressions.

In the step a) we first run cross-section regressions to assess the relative performances, measured by RoAs, of different BBMs:

$$\mathrm{RoA}_{\mathrm{i}}^{(\mathrm{t})} = \alpha^{(\mathrm{t})} + \sum_{\mathrm{g}} \beta_{\mathrm{g}}^{(\mathrm{t})} \mathrm{D}_{\mathrm{i,g}}^{(\mathrm{t})} + \gamma^{(\mathrm{t})} \mathrm{GDP}_{\mathrm{c}}^{(\mathrm{t})} + \epsilon_{\mathrm{i}}^{(\mathrm{t})}$$

where $\mathrm{RoA}_{\mathrm{i}}^{(\mathrm{t})}$ is bank $\mathrm{i}'s$ RoA in year $\mathrm{t}$; $\mathrm{D}_{\mathrm{i,g}}^{(\mathrm{t})}$ is a vector of dummy variables assuming value equal to 1 when bank $\mathrm{i}$ belongs to group g (that is, to a specific BBM/PG); $\mathrm{GDP}_{\mathrm{c}}^{(\mathrm{t})}$ is country $c$'s real GDP annual growth.

Second, we also exploit the time-series dimension of the dataset to run pooling panel data regressions on the same type of equation:

$$\mathrm{RoA}_{\mathrm{i,t}} = \alpha + \sum_{\mathrm{g}} \beta_{\mathrm{g}} \, \mathrm{D}_{\mathrm{i,t,g}} + \gamma \mathrm{GDP}_{\mathrm{c,t}} + \epsilon_{\mathrm{i,t}}.$$

We run the regressions on two different sub-samples: the first sub-sample refers to the period preceding the the Sovereign Debt Crisis (from 2006 to 2010), while the second sub-sample covers the years in which the effects of the Euro Area crisis are more evident on banks' balance sheets. The cross-section and pooling regressions are estimated with Ordinary Least Squares (OLS).

In step b) we estimate a dynamic panel of the following form:

$$\mathrm{RoA}_{\mathrm{i,t}} = \alpha + \mu_{\mathrm{i}} + \phi \mathrm{RoA}_{\mathrm{i,t-1}} + \beta \mathrm{X}_{\mathrm{i,t}} + \gamma \mathrm{K}_{\mathrm{c,t}} + \delta \mathrm{Z}_{\mathrm{t}} + \epsilon_{\mathrm{i,t}}$$

where $\mu_{\mathrm{i}}$ is a bank fixed effect; $\mathrm{X}_{\mathrm{i,t}}$ is a vector of bank characteristics; $\mathrm{K}_{\mathrm{c,t}}$ is a vector of country-specific factors; $\mathrm{Z}_{\mathrm{t}}$ is a vector of risk factors which are common to all banks. In particular, the vector $\mathrm{X}_{\mathrm{i,t}}$ represents bank i's total assets (a proxy for the bank's size, in logs), Cost of Risk, Cost efficiency and Tier 1 capital ratio. The vector $\mathrm{K}_{\mathrm{c,t}}$ includes a country's real GDP annual growth and the 10-year sovereign interest rate spread (10-year government interest rate minus the 10-year euro swap interest rate). Finally, the vector $\mathrm{Z}_{\mathrm{t}}$ includes the annual return on the Eurostoxx equity index, the 3-month Euribor interest rate (a proxy for the level of the yield curve) and the difference between the 10-year euro interest rate swap and the 3-month Euribor (a proxy for the slope of the interest rate curve). An autoregressive component is added to the regression in order to control for profitability persistence at the bank level. The coefficients are estimated with Generalized Method of Moments (GMM) [8] in order to properly take into account the possibility of endogeneous regressors.[9]

After controlling for macroeconomic factors (i.e. GDP growth), the cross-section regressions reveal that Retail banks performed significantly better than Investment and Diversified banks before the Sovereign Debt crisis (from 2006 to 2010; see Table 3). From 2011, the BBMs do not show significant differences in their RoAs, with the exception of 2012, when Investment banks performed significantly better than Retail banks.

In addition to RoA, we performed cross-section regressions with the Cost of Risk and operating costs/total assets as dependent variables. As expected, for the Cost of Risk the results show a significantly worse performance of the Retail model in 2012 (one of the worst years in terms of deterioration of credit quality). The ratio of operating costs to total assets was significantly higher for Retail banks across the entire period, confirming that more market-oriented models are more efficient.

The pooling panel data regressions estimated on the two sub-samples roughly confirm the previous results (Table 4). Retail banks were the best performers before the effects of the financial crisis became evident on banks' balance sheets; after 2010, different business strategies, when considered at an aggregate level, hardly account for any significant difference in banks' profitability, whereas country factors (proxied by economic growth) appear sufficient in explaining banks' heterogeneous RoAs.

However, when we split BBMs into homogeneous PGs, we get a more precise picture of relative performances of different bank business strategies (Table 5).[10] In particular, Retail banks that are operationally complex and international performed significantly better in periods of lower financial stress, not only with respect to Diversified banks, but also compared to other Non-Complex and/or Domestic Retail banks.

Moreover, and most importantly, the classification in PGs allows a clearer identification of low performers after 2010 (Table 5). In particular, Non-Complex Retail banks (both domestic and international) showed a lower RoA from 2011 to 2014. The "Non-complex Retail banks" include relatively small banks with traditional intermediation activity and limited cross-border exposure, i.e. with a lower degree of diversification in their sources of income and therefore more hit by the decline in credit demand and the deterioration in credit quality. In other words, these banks were not able to effectively substitute their "traditional" banking activity with alternative and more profitable sources of income.

As a final econometric exercise, as mentioned in b), we estimate a dynamic panel to identify relevant risk factors at the BBM level. We run separate regressions for Retail and non-Retail banks. The main results of the analysis are the following (Table 6):

[8] See Wooldridge (2010) for an overview of GMM estimators in panel data models.

[9] All the inference is conducted with robust standard errors.

[10] PGs are identified using 2014 data. For banks' attribution to PGs in the years before 2014, we make the following assumptions: for the years in which the banks are grouped in the same BBM as in 2014, we leave the PG unchanged; when the transition is from Retail to Diversified, we assume that banks belong to the Diversified Non-Specialized PG; when the transition is from Diversified Specialized to Retail, we assume that banks belong to the Retail Non-Complex Domestic PG; finally, when the transition is from Diversified Non-Specialized to Retail, we assume that banks belong to the Retail Complex International PG.

- economic growth, the term structure of interest rates and sovereign risk are the most relevant risk factors for Retail banks; in particular, Retail banks' RoAs depend positively on contemporaneous GDP growth; it is also positively affected by a higher level and slope of the yield curve;[11] on the other hand, RoA is negatively related to sovereign risk, as government bonds represent a significant portion of banks' assets and a higher probability of sovereign default tends to be associated with higher costs of wholesale funding and capital for banks;
- GDP growth remains an important risk factors for non-Retail banks, but Investment and Diversified banks show also a significant exposure to financial markets, as proxied by the return on the Eurostoxx; for this group of banks, the level and slope of the yield curve, the sovereign interest rate spread and bank total assets are not significant regressors;
- when Cost of Risk, Cost efficiency and Tier 1 ratio are added to the econometric specification, credit quality remains the only significant bank-specific factor (higher credit quality implies higher bank RoAs), while the sovereign interest rate spread and equity returns are the only significant macro-financial drivers for Retail banks profitability;
- the opposite is true for non-Retail banks, as the inclusion of a larger set of bank-specific characteristics adds significance to the level and slope of the yield curve (which were non-significant in the regression that includes only bank size);
- there is evidence that capital is not harmful for bank's profitability, at least for non-Retail banks: a 1-percent higher Tier-1 ratio translates into a 5-bps higher RoA.

The same model applied to the Cost of Risk and Cost efficiency as endogenous variables roughly confirms our prior. In particular:

- The Cost of Risk for Retail banks shows a strong negative relationship with GDP growth, which is less significant for non-Retail banks. Moreover, for Retail banks the Cost of Risk is also positively related to the sovereign interest rate spread, which instead is not a significant driver for non-Retail banks;
- Cost efficiency is not significantly related to macroeconomic variables, confirming the "structural" nature of this indicator. For both Retail and non-Retail banks, there is a negative link between Cost efficiency and total assets (that is stronger for the Retail model). This result may indicate the presence of economies of scale.

# 5 Conclusions

In this paper we extended the existing literature on Bank Business Models along several dimensions. First, we used probabilistic, as well as deterministic, statistical learning techniques to identify BBMs in Europe. Second, we used supervised learning techniques to identify homogeneous peer groups of banks according to the criteria outlined by the EBA in the SREP Guidelines. Third, we developed a panel data framework to identify the relevant macro-financial factors and the individual banks characteristics that explain the heterogeneity in profitability among different business models.

[11] See Borio et al. (2015) for a review of the literature on the relationship between bank profitability and the yield curve.

We obtained the following results. First, we identified three different BBMs: Retail, Investment and Diversified banks. When with fuzzy clustering we admit uncertainty in BBM attribution at the individual bank's level (i.e. estimating for each bank the so-called degrees of membership to a specific BBM), the identification provides almost the same result as hard clustering, with a higher risk of incorrect identification during the Sovereign Debt crisis (from 2010 to 2012). Therefore, we assign each bank to its most likely business model to identify and evaluate the performance of the PGs.

Second, given our qualitative and quantitative prior knowledge of individual banks' characteristics, we used supervised learning to classify banks in nine distinct Peer Groups, based on systemic relevance, operational complexity, business specialization and cross-border activity. At the European level, Non-Complex Retail Domestic banks represent the largest PG, with banks mostly located in Italy and Portugal. The second most numerous PG is the Complex Retail Domestic one, with banks mostly located in Italy and Germany.

Third, coming to the performance of different business strategies, we showed that Retail banks were the best performers in the years preceding the financial crisis. However, they had the worst RoAs at the peak of the Sovereign Debt Crisis, as higher government interest rate spreads determined lower credit demand, higher funding costs, higher capital requirements and lower profitability. Furthermore, during the Euro Area crisis, Retail banks experienced the most significant deterioration in credit quality, only partially offset by higher cost efficiency.

Moreover, running pooling panel data regressions at the PG level, we get a more precise picture of relative performances of different bank business strategies. Retail banks that are operationally complex and international performed significantly better in the run-up to the financial crisis. Also, and most importantly, Non-Complex Retail banks (both domestic and international), that is banks with less diversified sources of income and therefore more hit by the decline in credit demand and the deterioration in credit quality, had lower RoAs from 2011 to 2014.

Finally, we identified the relevant risk factors at the business model level. Economic growth, yield curve level and slope and sovereign default risk are the main factors affecting bank profitability; among bank characteristics, credit quality is a positive driver of Retail banks' RoA, whereas exposure to financial markets, proxied by equity market returns, is a significant driver for non-Retail (Investment and Diversified) banks. Finally, at least for Non-Retail banks, there is some evidence that holding more capital is not harmful to bank profitability.

The methodology developed in this paper has numerous applications. First, it can be used by banks to develop benchmarking analysis, i.e. to compare their performance with that of the other members of their peer group and extract useful information on their areas of strength and weakness in terms of profitability, risk, efficiency, asset and liability composition, etc. Moreover, a natural extension of our work would be to exploit our identification strategy to analyse the sustainability of bank business models in a relevant forecasting horizon. In particular, it can help to assess a bank's ability to achieve satisfactory short-term profits (within a 12-month horizon), and medium-term sustainability (over a 3-year horizon), as outlined in the SREP framework. We leave these extensions as areas of future research.

# References


[1] Ayadi R., De Groen W. P. (2015), “Banking Bank Business Models Monitor”, IRCCF.

[2] Altunbas Y., Manganelli S., Marques-Ibanez (2011), “Bank risks during the Financial Crisis”, ECB Working Paper No 1394.

[3] Bezdek J.C. (1981), “Pattern Recognition with Fuzzy Objective Function Algorithms”, Springer.

[4] Bonaccorsi di Patti E., Felici R., Signoretti F. M. (2016). “Euro Area significant banks: main differences and recent performance”, Questioni di Economia e Finanza (Occasional papers), Banca d’Italia.

[5] Borio C., Gambacorta L., Hoffman B., (2015), “The influence of monetary policy on bank profitability”, BIS Working papers.

[6] ECB (2016), “Recent trends in euro area banks’ Bank Business Models and implications for banking sector stability”, Financial Stability Review, May.

[7] James G., Witten D., Hastie T., Tibshirani R. (2013), “An Introduction to Statistical Learning”, Springer.

[8] Roengptya R., Tarashev N., Tsatsaronis K. (2014), “Bank Bank Business Models”, BIS Quarterly Review, December.

[9] Ward, J. H. (1963), "Hierarchical Grouping to Optimize an Objective Function", Journal of the American Statistical Association 58 (301): 236–244.

[10] Wooldridge J. M., “Econometric Analysis of Cross Section and Panel Data”, MIT Press, 2nd edition.

# Appendix

## A.1 Deterministic clustering

Given a matrix $X$ of $p$ characteristics (balance sheet indicators) for $n$ statistical units (banks), the goal of cluster analysis is to form groups of banks that are homogeneous *within* the group and heterogeneous *between* the groups. At the basis of clustering there is the choice of a distance measure (Euclidean distance). Given two points $x_i$ and $x_j$ in the $p$-dimensional Euclidean space, the Euclidean distance $d(x_i, x_j)$ is given by:

$$d(i,j) = \sqrt{\sum_{s=1}^{p} (x_{is} - x_{js})^2}.$$

The Euclidean distance is computed on standardized variables. The distance matrix $D$ is a $p \cdot p$ matrix containing at the position $i, j$ the distance $d(i, j)$.

A typical hierarchical clustering algorithm is performed through the following steps:

1. initialization: every unit (bank) represents a cluster;
2. selection: the two nearest clusters (based on a distance measure) are grouped in a new cluster;
3. update: the distance matrix $D$ is updated on the units corresponding to the new number of clusters (the number of clusters at step 1 *minus* 1);
4. iteration: repeat steps (2)-(3) $n - 1$ times.

The procedure stops when all banks are grouped in the same cluster.

In this paper we adopt the Ward's minimum variance method (see Ward, 1963), which minimizes an objective function corresponding to the total within-cluster variance $W = \sum_{k=1}^{g} W_k$, $W_k = \sum_{s=1}^{p} \sum_{i=1}^{n_k} (x_{is} - \bar{x}_{sk})^2$. The Ward's method is implemented by the Lance-Williams algorithm; at step 2, the distance between a cluster $k$ and a new group of clusters $i$ and $j$ next to be merged can be estimated as:

$$d(i \cup j, k) = \frac{n_i + n_k}{n_i + n_j + n_k} d(i,k) + \frac{n_j + n_k}{n_i + n_j + n_k} d(j,k) - \frac{n_k}{n_i + n_j + n_k} d(i,j)$$

where $n_i, n_j, n_k$ are clusters' sizes.

Therefore, hierarchical clustering aims at progressively forming homogeneous groups of banks, starting from the most disaggregated level. This process continues until an optimal number of clusters is reached. The *stopping rule* (the optimal number of clusters) is typically the output of hypothesis testing. Given the following statistics

$$W_q = \sum_{k=1}^{q} \sum_{i \in C_k} (x_i - c_k)(x_i - c_k)^T,$$

$$B_q = \sum_{k=1}^{q} n_k (c_k - \bar{x})(c_k - \bar{x})^T,$$

$$S_w = \sum_{k=1}^{q} \sum_{i,j \in C_k} d(x_i, x_j),$$

$$S_b = \sum_{k=1}^{q-1} \sum_{l=k+1}^{q} \sum_{i \in C_k, j \in C_l} d(x_i, x_j),$$

where $q$ is the number of clusters, $c_k$ the centroid of cluster $C_k$, $x_i$ a p-dimensional vector of observations of the $i$th object in cluster $C_k$ and $d$ is the distance function. Hypothesis testing on the optimal number of clusters can be of two groups:

- a first group assesses the homogeneity of data points within the cluster (by minimizing $W_q$), or the heterogeneity between the clusters (by maximizing $B_q$), or a combination of the two measures into a single score; the Calinski and Harabasz index, the Hartigan index, the Ratkowsky index, the Ball index belong to this group;
- a second group measures the distance between two different clusters (by using $S_b$), or the inter-cluster distance (by using $S_w$), or a combination of the two statistics; the Dunn index, the Silhouette index, the DB index belong to this second group.

In our identification procedure, the optimal stopping rule derives from hypothesis testing conducted on 30 test-statistics of both groups.

## A.2 Probabilistic clustering

In principle, we could admit that a bank's attribution to the BBM is subject to a significant degree of uncertainty. In previous research on BBMs identification, "hard" (i.e. deterministic) clustering procedures were mainly used to group banks. However, the main drawback of hard clustering is that each bank belongs to a single BBM. In principle, the bank's BBM could be multi-faceted. For example, one bank could present the characteristics of commercial banks and, at the same time, rely on wholesale funding. In this case, the attribution of the bank to one specific BBM could be ambiguous and highly sensitive to the composition of the sample. To overcome this drawback, we adopt a fuzzy clustering method (Bezdek, 1981).

Fuzzy clustering is based on the minimization of the following objective function:

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} p_{ij}^m \left\| x_i - c_j \right\|^2, m \in (1, \infty)$$

where $m$ determines the cluster *fuzziness*, $\|.\|$ is the norm operator expressing the data, $x_j$ is the observed data, $c_j$ is the cluster center and $p_{ij}$ is the membership degree of observation $i$ to cluster $j$. The larger the fuzzifier $m$, the smaller the memberships, the fuzzier (i.e. the more uncertain) the estimated clusters. In practical applications, the fuzzifier is chosen to be sufficiently small ($m = 2$) in order to control for a limited degree of fuzziness.

The solution to the optimization problem is achieved through numerical methods. The fuzzy c-mean algorithm is performed through the following steps:

1. initialize the matrix $P^{(0)} = (p_{ij})$;
2. calculate the center vectors $C^{(s)}, c_j = \frac{\sum_{i=1}^{N} p_{ij}^m x_i}{\sum_{i=1}^{N} p_{ij}^m}, P = P^{(s)}$;
3. update $P^{(s+1)}, p_{ij}^{(s+1)} = \frac{1}{\sum_{k=1}^{C} \left( \frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)}^{\frac{2}{m-1}}$;
4. repeat 2. and 3. until $\|P^{(s+1)} - P^{(s)}\| > \epsilon$.

Fuzzy clustering maps banks into different BBMs; hence, for each bank, fuzzy clustering returns the so-called membership degrees, i.e. the conditional probability that bank $\mathrm{B_j}$ belongs to BBM $\mathrm{BM_i}$. In symbols, at any point in time, with fuzzy clustering we estimate $\mathrm{P(BM_i|B_j)}$. In order to estimate "expected" key-performance indicators (KPI) at the BBM level, we need to estimate the probability of bank $\mathrm{j}$ conditional to the realization of the BBM $\mathrm{i}$. To obtain these probabilities, we exploit the Bayes theorem:

$$\mathrm{P(B_j|BM_i)} = \frac{\mathrm{P(BM_i|B_j)P(B_j)}}{\mathrm{P(BM_i)}}$$

where $\mathrm{P(B_j)} = \frac{1}{\mathrm{N}}$ , $\mathrm{N}$ is the number of banks in the sample and $\mathrm{P(BM_i)} = \sum_j \mathrm{P\,(BM_i|B_j)P(B_j)}$.

We perform fuzzy clustering on a yearly basis, obtaining time-varying membership degrees. Therefore, the expected $\mathrm{K\tilde{P}I_{i,t}}$of BBM $\mathrm{i}$ at time $\mathrm{t}$ is given by:

$$\mathrm{K\tilde{P}I_{i,t}} = \sum_j \mathrm{K\,PI_{j,t}P(B_j|BM_i)_t}.$$

## A.3 Figures and tables

**Table 1: List of banks**

| Bank | Country | Bank | Country |
|---|---|---|---|
| UniCredit SpA | Italy | Banco de Sabadell, SA | Spain |
| Intesa Sanpaolo SpA | Italy | Banco Popular Español SA | Spain |
| Banca Monte dei Paschi di Siena SpA | Italy | Bankinter SA | Spain |
| Unione di Banche Italiane SpA | Italy | ING Bank NV | Netherlands |
| Banco Popolare Società Cooperativa | Italy | Coöperatieve Rabobank U.A. | Netherlands |
| Banca popolare dell'Emilia Romagna SC | Italy | SNS REAAL NV | Netherlands |
| Banca Popolare di Milano Scarl | Italy | Erste Group Bank AG | Austria |
| Banca Carige SpA | Italy | Raiffeisenlandesbank Oberösterreich AG | Austria |
| Banca Popolare di Vicenza SpA | Italy | Raiffeisenlandesbank Niederösterreich-Wien AG | Austria |
| Veneto Banca SCpA | Italy | Raiffeisen Zentralbank Österreich AG | Austria |
| Banca Popolare di Sondrio SCpA | Italy | Österreichische Volksbanken-AG | Austria |
| Credito Emiliano SpA | Italy | Sberbank Europe AG | Austria |
| Credito Valtellinese Società Cooperativa | Italy | Belfius Banque SA | Belgium |
| Iccrea Holding SpA | Italy | Dexia SA | Belgium |
| Mediobanca - Banca di Credito Finanziario SpA | Italy | KBC Group NV | Belgium |
| Deutsche Bank AG | Germany | Bank of Cyprus Public Company Limited | Cyprus |
| Commerzbank AG | Germany | Hellenic Bank Public Company Ltd. | Cyprus |
| Deutsche Zentral-Genossenschaftsbank AG | Germany | Danske Bank Oyj | Finland |
| Landesbank Baden-Württemberg | Germany | Nordea Pankki Suomi Oyj | Finland |
| Bayerische Landesbank | Germany | OP Financial Group | Finland |
| NORD/LB Norddeutsche Landesbank Girozentrale | Germany | Alpha Bank AE | Greece |
| Landesbank Hessen-Thüringen Girozentrale | Germany | Eurobank Ergasias SA | Greece |
| Hypo Real Estate Holding AG | Germany | National Bank of Greece SA | Greece |
| HSH Nordbank AG | Germany | Piraeus Bank SA | Greece |
| DekaBank Deutsche Girozentrale | Germany | Allied Irish Banks, Plc | Ireland |
| Westdeutsche Genossenschafts-Zentralbank AG | Germany | Permanent TSB Group Holdings Plc | Ireland |
| Landwirtschaftliche Rentenbank | Germany | Bank of Ireland | Ireland |
| Volkswagen Financial Services AG | Germany | Ulster Bank Ireland Designated Activity Company | Ireland |
| Aareal Bank AG | Germany | Banque et Caisse d'Epargne de l'Etat | Luxembourg |
| SEB AG | Germany | Banco BPI SA | Portugal |
| BNP Paribas SA | France | Banco Comercial Português, SA | Portugal |
| Crédit Agricole SA | France | Caixa Geral de Depósitos SA | Portugal |
| Société Générale SA | France | Nova Kreditna banka Maribor d.d. | Slovenia |
| Crédit Mutuel Group | France | Nova Ljubljanska Banka d.d. | Slovenia |
| HSBC France | France | Barclays Plc | United Kingdom |
| La Banque Postale | France | Lloyds Banking Group Plc | United Kingdom |
| Banco Santander, SA | Spain | Royal Bank of Scotland Group Plc | United Kingdom |
| Banco Bilbao Vizcaya Argentaria, SA | Spain | HSBC Holdings Plc | United Kingdom |
| Grupo Cooperativo Cajamar | Spain | | |

**Figure 1: Radar Plot**

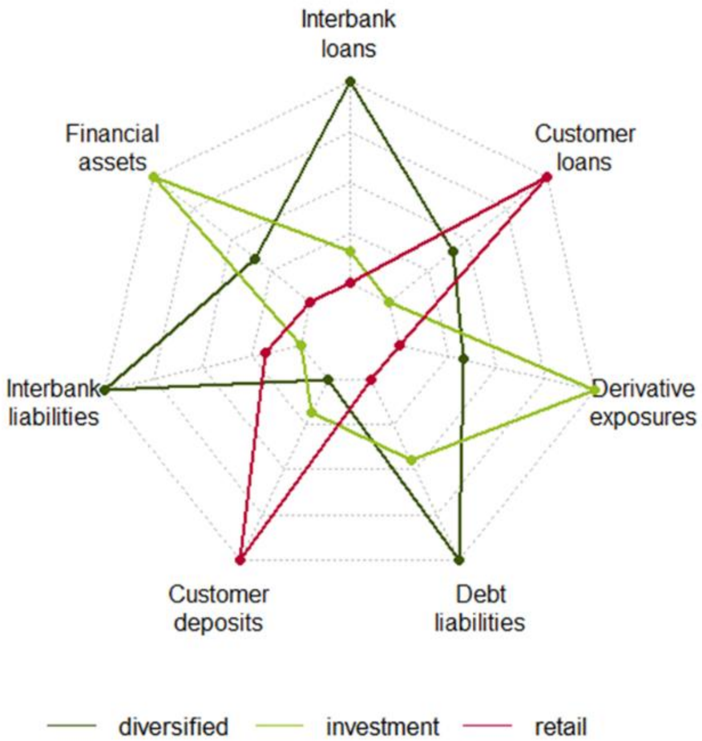
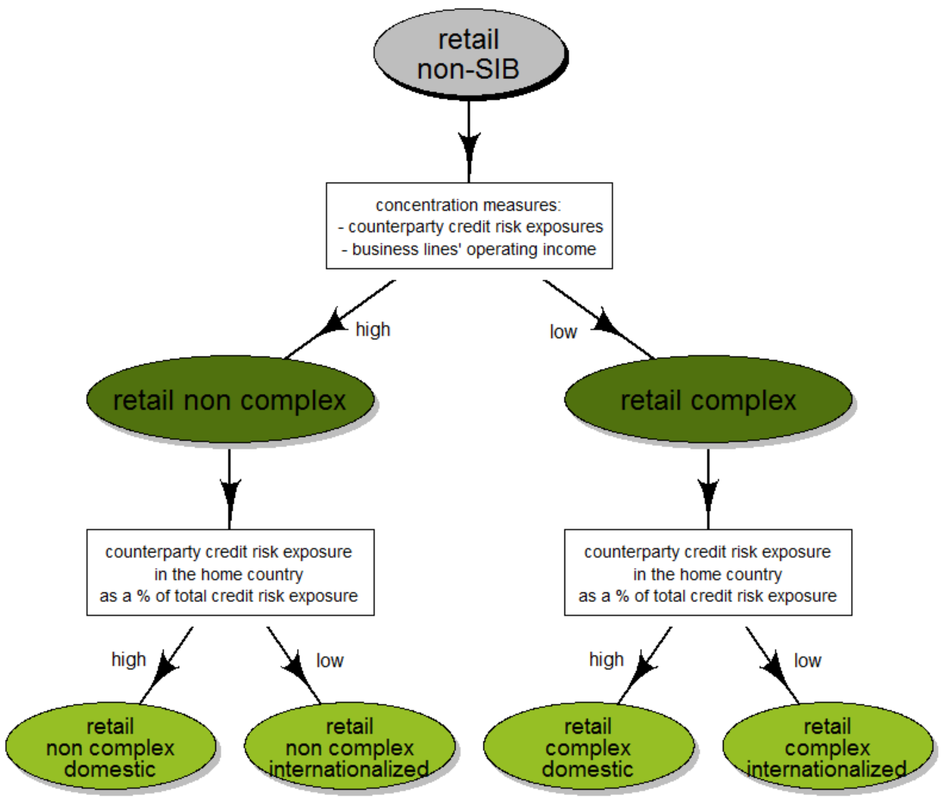
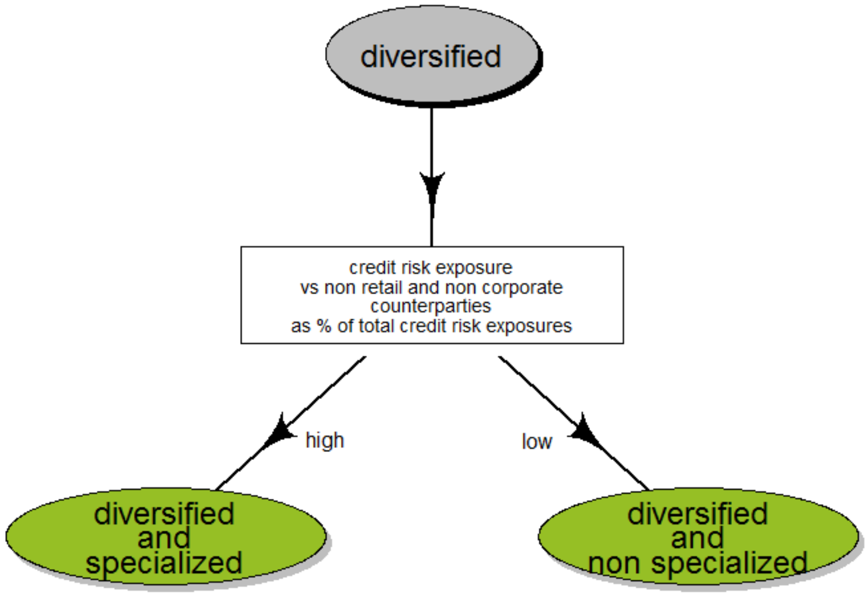



**Figure 2: Peer group identification for retail banks**



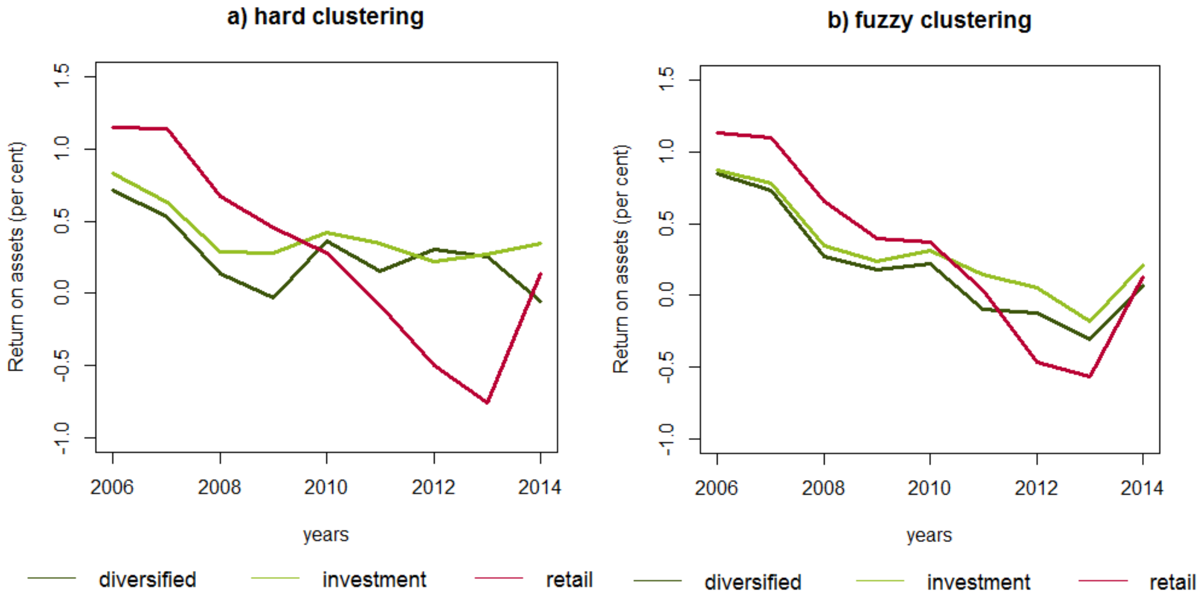
## Figure 3: Peer group identification for diversified banks



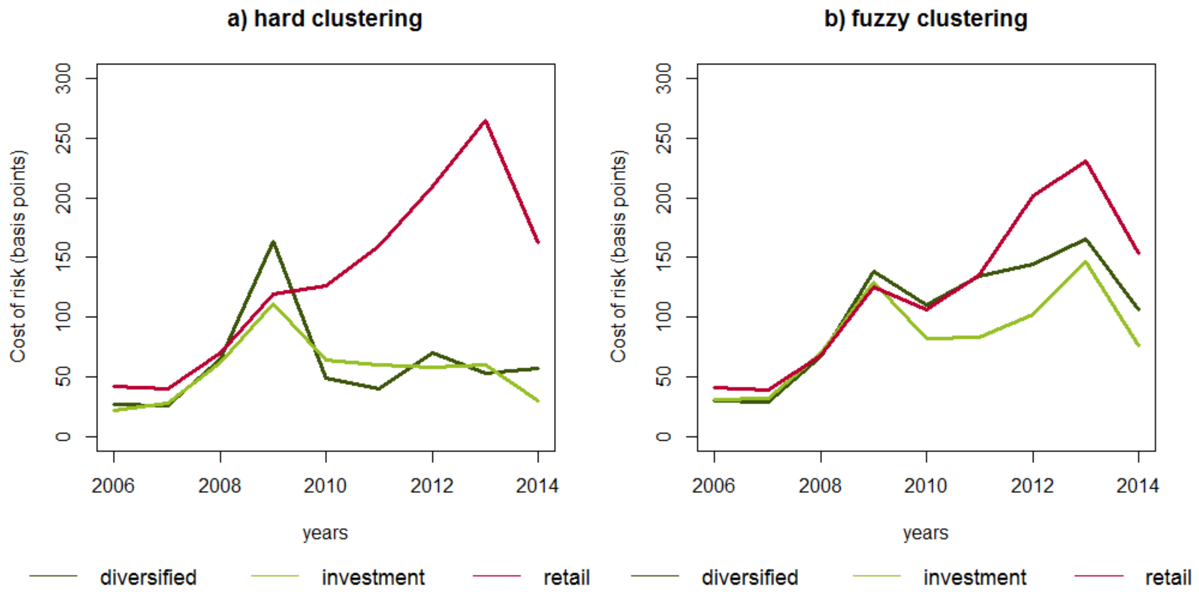

## Table 2: Identified peer groups (2014)

| | N. banks per country per peer | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Country | Systemic investment banks | Non systemic investment banks | Systemic retail banks | Complex and international retail banks | Complex and domestic banks | Non complex and international retail banks | Non complex and domestic retail banks | Diversified and specialized | Diversified and non specialized banks | **N. banks per country** |
| | **inv_SIB** | **inv_NSIB** | **ret_SIB** | **ret_CO_INT** | **ret_CO_DOM** | **ret_NC_INT** | **ret_NC_DOM** | **div_SP** | **div_NS** | |
| Austria | | | | 3 | | | | 1 | 2 | **6** |
| Belgium | | 1 | | 1 | | | | 1 | | **3** |
| Cyprus | | | | | | | 2 | | | **2** |
| Germany | 1 | 3 | | | | 2 | | 5 | 4 | **15** |
| Spain | | | 1 | 1 | 2 | | 2 | | | **6** |
| Finland | | 1 | | | 1 | | 1 | | | **3** |
| France | 3 | 1 | | | 1 | | | 1 | | **6** |
| United Kingdom | 3 | | | | 1 | | | | | **4** |
| Greece | | | | 1 | | 1 | 2 | | | **4** |
| Ireland | | | | 2 | | | 2 | | | **4** |
| Italy | | | 1 | 1 | 3 | | 8 | 1 | 1 | **15** |
| Luxembourg | | | | | | 1 | | | | **1** |
| Netherlands | | | 1 | | 2 | | | | | **3** |
| Portugal | | | | | | | 3 | | | **3** |
| Slovenia | | | | | | 1 | 1 | | | **2** |
| **N. banks per peer** | **7** | **6** | **3** | **9** | **10** | **5** | **21** | **9** | **7** | **77** |

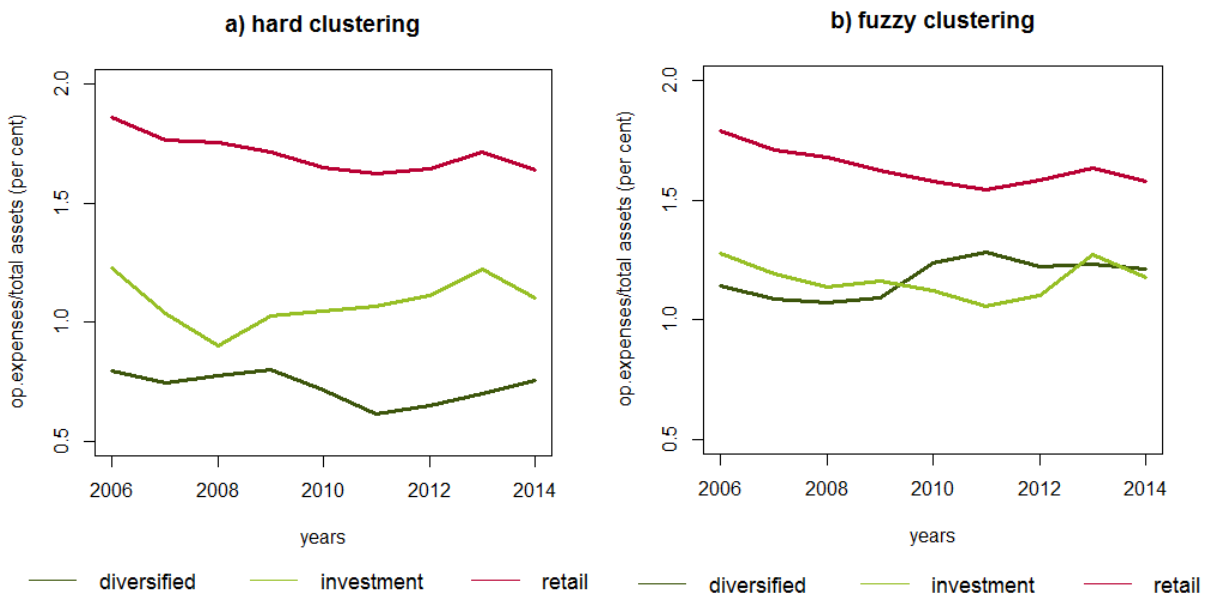
## Figure 4: ROA



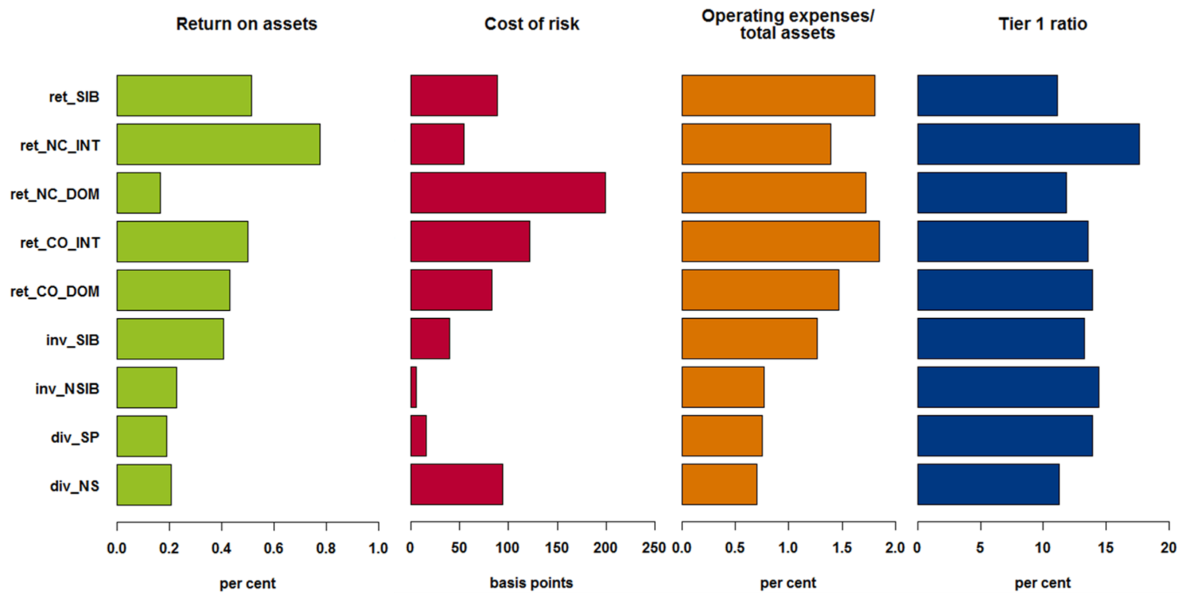

## Figure 5: Cost of risk



## Figure 6: Operating expenses/total assets




## Figure 7: Peer groups' KPI (2014)



## Table 3: Cross-section regressions

**Table 3a) cross section regression of bank ROA on dummy variables for retail and diversified banks**

| | 2006 | | 2007 | | 2008 | | 2009 | | 2010 | | 2011 | | 2012 | | 2013 | | 2014 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| dip.var: ROA | coeff. | | coeff. | | coeff. | | coeff. | | coeff. | | coeff. | | coeff. | | coeff. | | coeff. | |
| Intercept | 0.779 | *** | 0.378 | * | 0.331 | * | 0.870 | *** | -0.040 | | 0.237 | | 0.012 | | 0.184 | | -0.317 | |
| Diversified banks | -0.171 | | -0.299 | * | -0.244 | | -0.437 | * | -0.084 | | -0.656 | | 0.248 | | -0.064 | | -0.119 | |
| Retail banks | 0.342 | * | 0.414 | *** | 0.353 | . | 0.036 | | 0.238 | | -0.189 | | -0.581 | | -0.992 | | 0.185 | |
| GDP_growth | 0.010 | | 0.072 | * | 0.017 | | 0.108 | *** | 0.099 | * | 0.027 | | 0.055 | | 0.055 | | 0.175 | ** |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

**Table 3b) cross section regression of bank ROA on dummy variables for retail and investment banks**

| | 2006 | | 2007 | | 2008 | | 2009 | | 2010 | | 2011 | | 2012 | | 2013 | | 2014 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| dip.var: ROA | coeff. | | coeff. | | coeff. | | coeff. | | coeff. | | coeff. | | coeff. | | coeff. | | coeff. | |
| Intercept | 0.608 | ** | 0.078 | | 0.086 | | 0.433 | ** | -0.124 | | -0.419 | | 0.260 | | 0.120 | | -0.436 | |
| Investment banks | 0.171 | | 0.299 | * | 0.244 | | 0.437 | * | 0.084 | | 0.656 | | -0.248 | | 0.064 | | 0.119 | |
| Retail banks | 0.512 | *** | 0.713 | *** | 0.597 | *** | 0.473 | ** | 0.323 | | 0.467 | | -0.829 | * | -0.927 | | 0.304 | |
| GDP_growth | 0.010 | | 0.072 | * | 0.017 | | 0.108 | *** | 0.099 | * | 0.027 | | 0.055 | | 0.055 | | 0.175 | ** |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Table 4: Pooled regression on business models ROA

**Table 4a) pooled regression of bank ROA on dummy variables for retail and diversified banks**

| | 2006-2010 | | 2011-2014 | |
|---|---|---|---|---|
| dip.var: ROA | coeff. | | coeff. | |
| Intercept | 0.423 | *** | 0.020 | |
| Diversified banks | -0.291 | ** | -0.178 | |
| Retail banks | 0.224 | * | -0.369 | |
| GDP_growth | 0.062 | *** | 0.087 | ** |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

**Table 4b) pooled regression of bank ROA on dummy variables for retail and investment banks**

| | 2006 | | 2007 | |
|---|---|---|---|---|
| dip.var: ROA | coeff. | | coeff. | |
| Intercept | 0.132 | * | -0.158 | |
| Investment banks | 0.291 | ** | 0.178 | |
| Retail banks | 0.515 | *** | -0.191 | |
| GDP_growth | 0.062 | *** | 0.087 | ** |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Table 5: Pooled regression on peer groups ROA

**Table 5) pooled regression of bank ROA on dummy variables for retail peers and diversified banks**

| | 2006-2010 | | 2011-2014 | |
|---|---|---|---|---|
| dip.var: ROA | coeff. | | coeff. | |
| Intercept | 0.453 | *** | 0.121 | |
| Non complex and international retail banks | 0.141 | | -0.744 | . |
| Non complex and domestic retail banks | 0.185 | . | -0.751 | ** |
| Complex and international retail banks | 0.341 | * | -0.276 | |
| Complex and domestic retail banks | 0.146 | | -0.134 | |
| Diversified banks | -0.322 | *** | -0.250 | |
| GDP_growth | 0.063 | *** | 0.074 | * |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Table 6: dynamic panel regressions on business models ROA

**Table 6a) dynamic panel regressions of retail banks' ROA**

| | 2006-2014[1] | | 2006-2014[2] | |
|---|---|---|---|---|
| dip.var: ROA | coeff. | | coeff. | |
| ROA (t-1) | -0.436 | . | -0.099 | * |
| Cost of risk | | | -0.006 | *** |
| Op. expenses/total assets | | | 0.027 | |
| Tier 1 ratio | | | 0.036 | |
| log_total assets | 2.448 | | -0.218 | |
| GDP | 0.090 | *** | -0.002 | |
| euribor 3m | 0.489 | ** | 0.048 | |
| IRS 10y - euribor 3m | 0.543 | ** | -0.014 | |
| 10y gov. bond rate - IRS 10y | -0.313 | ** | -0.118 | *** |
| equity | 0.004 | | 0.003 | * |
| Number of Observations | 277 | | 277 | |

[1]Autocorrelation test (1): normal = -2.280212 (p.value=0.022595)

[2]Autocorrelation test (1): normal = -2.103194 (p.value=0.035449)

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

**Table 6b) dynamic panel regressions of non retail banks' ROA**

| | 2006-2014[1] | | 2006-2014[2] | |
|---|---|---|---|---|
| dip.var: ROA | coeff. | | coeff. | |
| ROA (t-1) | -0.011 | | 0.019 | |
| Cost of risk | | | -0.002 | * |
| Op. expenses/total assets | | | -0.354 | |
| Tier 1 ratio | | | 0.045 | . |
| log_total assets | 0.119 | | -0.035 | |
| GDP | 0.018 | * | 0.012 | |
| euribor 3m | 0.081 | | 0.169 | * |
| IRS 10y - euribor 3m | 0.013 | | 0.149 | * |
| 10y gov. bond rate - IRS 10y | -0.009 | | -0.016 | |
| equity | 0.004 | ** | 0.005 | ** |
| Number of Observations | 219 | | 219 | |

1Autocorrelation test (1): normal = -3.126807 (p.value=0.0017672)

2Autocorrelation test (1): normal = -2.103194 (p.value=0.035449)

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1