TECHNISCHE UNIVERSITÄT WIEN
Vienna | Austria

Master's Thesis

Banking Efficiency in the Eurozone

conducted at the institute of
Statistics and Mathematical Methods in Economics

at the
Vienna University of Technology

under the guidance of
Ao. Univ. Prof. Wolfgang Scherrer

written by
Oliver Leodolter

Vienna, September 2015

Year: 2015
Title: Banking Efficiency in the Eurozone
Author: Oliver Leodolter

Scientific environment

The author has carried out the research reported in this master thesis at the Research Group on Econometrics and System theory of the Institute of Statistics and Mathematical Methods in Economics at the Vienna University of Technology, Austria.

This research was funded by the Jubiläumsfonds of the Austrian National Bank, Project number 15495.

A Project report for the project has already been written in december 2014 together with Prof. Dr. Wolfgang Scherrer and Mag. Ronald Scheucher ([16]). This master thesis is an extension to this report.

Acknowledgements

First of all I would like to express my gratitude to my supervisor Wolfgang Scherrer for the possibility to work on this project. Furthermore for advising me on scientific problems, arising during the research, giving me valuable notions and for helping me writing this master thesis. Next I would like to thank Ronald Scheucher, who was an important assistance on all matters concerning banking balancing, gathering the data and the choice of the DEA factors.
Also a lot of thanks to my colleagues for proofreading the drafts and especially for accompanying me the last five years.
Finally I really need to thank my family for supporting me throughout the entire process, by encouraging me and keeping me motivated.

Abstract

In the last years since 2007 the financial sector and especially the banking sector have often been the target for criticism. In 2013 the European Parliament passed the law for the further implementation of the Single Supervisory Mechanism (SSM) [1]. The aim of the SSM is to centralize the different national banking supervision in the body of the European Central Bank (ECB). In the course of the introduction of the SSM the ECB started the so called Asset Quality Review (AQR) in the fall of 2013 (see [6]). During the AQR the balance sheets of the 130 system-relevant banks in the Eurozone were audited and evaluated by their ability to survive economic shocks. The results were published on the 26th of october 2014. The main target for banking regulation has been the ratio of equity to riskweigthed assets. In this paper we introduced the definition of systemic banking performance by adding measures for the banks performance as financial intermediaries. We thought exclusively assessing the banks by their sustainability is not sufficient to guarantee an efficient working banking industry. We were looking for indicators from the banks balance sheet that are suitable to reflect the banks achievements in serving the economy. Therefore we used a panel of 70 European banks over 10 years, that were rated significant by the ECB and are under SSM-regulation (see [6]). We used a nonparametric approach (Data Envelopment Analysis) to track efficiency scores over 10 consecutive years. We used a slack based model together with the concept of super-efficiency to get highly discriminable efficiency scores. For the analysis of the efficiency growth we used the concept of the Malmquist index for DEA models. Further analyses regarding the sensibility of the efficiency score to changes in the input factors and composition of the sample were made to get a more detailed look on the bahaviour of the efficiency scores. A special focus was to see to what extent the financial crisis shows in the data. First we studied the data using descriptive statistics to illustrate the development in the banking sector over the years. Then we used panel regression models to try to identify which strategic variables are the determinants for banking efficiency in Europe. Due to the particular nature of DEA efficiency scores we used a robust covariance estimator, proposed by Driscoll and Kraay, for interference. To get stable results extreme bound analysis (see [14]) was implemented. This study was conducted during the project 'Banking Performance in Euroland. Efficiency and the Impact of Strategic Variables: 2003-2012' at the Vienna University of Technology. The project was funded by the Jubiläumsfonds of the Austrian National Bank, Project nr 15495.

[1] see http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014R0468&from=EN

Contents

Chapter 1

Introduction

1.1 Introduction

On the 9th of august 2007 the interbank rates increased enormously [1] as a consequence of a loss of confidence in the banking sector. Thus this event is often referred to as the starting point of a huge financial crisis with substantial effects on the world economy. Some say that the following crisis was the effect from a long story of financial deregulation, mismanagement and a non-transparent financial system. The massive lending towards the real estate sector, the securitization of this loans and finally the bursting of a housing bubble in the United States led to the so-called Subprime Crisis, which marks the first step of a long way leading to the bankruptcy of Greece and a heated debate about the future of the European Union. Due to this severe consequences we are all well advised to detect what went wrong and how we can prevent such crises in the future. Undoubtedly, banks, and especially their investment branches, played a crucial role in this story and therefore they are the main part of this investigation. For a further discussion on the triggers of the financial crisis see [17].

Due to the extended globalization on financial markets, soon after august 2007 the first european banks had to deal with the consequences of their risky investments in the United States and the crisis from overseas expanded to Europe (see [3]). Mistrust on the financial markets gave rise to rising interest rates, sinking of stock prices and further to a drop in GDP growth (see Figure 1.1). Governments had to rescue major banks (AIG, UBS, Hypo Real Estate and many more), through direct investments, providing guarantees and granting loans, to prevent a complete collapse of the system (for an extensive list of government policy actions see [13]). These immense investments into the banking sector caused the governments to raise taxes or borrow money, both at the cost of the tax payer. The extended borrowing of money led to an increase in the debt ratio (see Figure 1.1).

The central banks tackled the shrinking economic growth by reducing central bank lending rates to stimulate the economy (see Figure 1.1). In addition new banking regulation were introduced mainly regulating the banks equity ratios and capital requirements (Basel III). As part of this enhanced banking regulation the European Parlia-

[1]The 3-month LIBOR rate for USD jumped from 5.38% on the 8th august to 5.50% on the 9th august. For comparison the average daily change of the 3 month LIBOR from the 1st january 2007 till the 8th august 2007 was 0.0001 percentage points. Source:https://research.stlouisfed.org/

ment passed the law for the further implementation of the Single Supervisory Mechanism (SSM). The aim of the SSM is to centralize the different national banking supervision in the body of the European Central Bank (ECB) [4]. In the course of the introduction of the SSM the ECB started the so called Asset Quality Review (AQR) in the fall of 2013. During the AQR the balance sheets of the 130 system-relevant banks in the Eurozone were audited and evaluated by their ability to survive economic shocks (see [6]). The results were published on the 26th of october 2014 (see [5]). The main target for banking regulation are capital requirements and liquidity requirements, measured by the ratio of equity to risk-weigthed assets and liquid assets to short term liabilities, this also applies for the AQR (see [1]).
We therefore tried to extend the definition of banking performance by adding measures for the banks performance as financial intermediary. So we added the banks ability to allocate money in the economy, their ability to screen the credit market and gather reliable information on credit worthiness, and their ability to refinance. These skills were measured using data from the banks balance sheets, which we extracted from their published annual reports. We used a panel of 70 European banks, that were rated significant by the ECB and are therefore under SSM-regulation. After defining the foundations for our efficiency measure and gathering the data, we used a non-parametric approach (Data Envelopment Analysis) to track efficiency scores over ten consecutive years and to construct a Malmquist Productivity Index. In detail we used a Slack Based Model of (super-)efficiency for the efficiency score and the Malmquist Index of efficiency change. After calculating the efficiency scores I conducted a sensibility analysis. This analysis shows us what the best levers are to boost efficiency and what banks are the most influential in the set.
The longest and last part of this work consists of regression analysis to find out what are the key determinants of systemic banking efficiency growth. Due to the specific dependency structure of efficiency scores retrieved from DEA, we used OLS estimation together with a robust variance estimator proposed by Driscoll and Kraay. Because of the fact that we had a lot of potential variables to test, and therefore a great variety of different linear models, we decided to use extreme bound analysis to find highly stable significant variables. Finally we analysed the residuals of the regression to check for serial correlation and cross-sectional dependence.

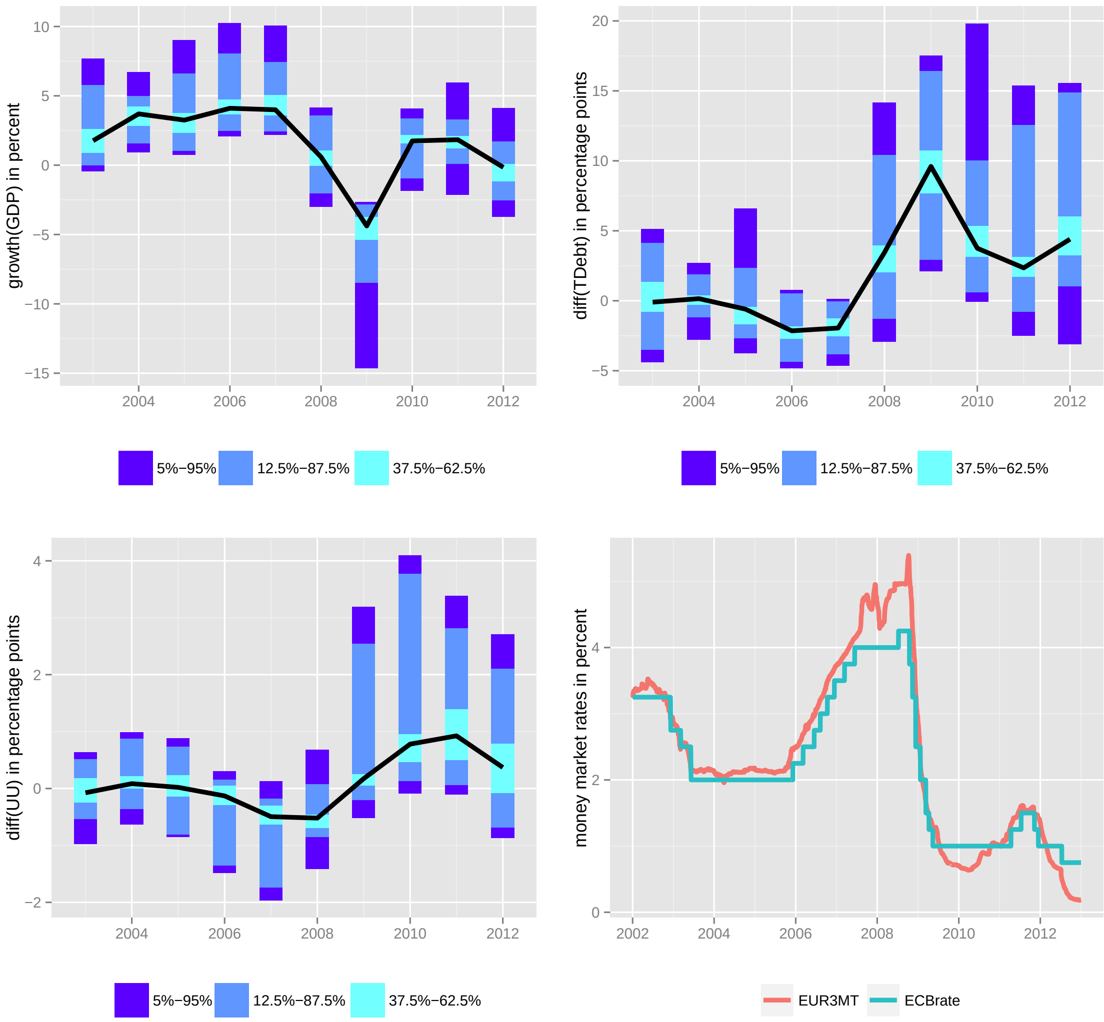

Figure 1.1: The data on GDP, TDebt and UU are from the ECB Statistical Warehouse (see Table A.2). The plots show the quantiles of the variables for the 18 home countries of the observed banks. The continous line is the median value. The daily quotations of the 3 month Euribor are taken from `https://www.bundesbank.de/Navigation/DE/Statistiken/statistiken.html`.

Chapter 2

Data Envelopment Analysis

Data Envelopment Analysis is a non-parametric method to measure the efficiency of Decision Making Units (DMU). A DMU can be any unit or entity, that produces s different outputs, $y_r, r = 1, \ldots, s$, from m different inputs, $x_i, i = 1, \ldots, m$, like schools, enterprises, factories or banks. For the rest of this chapter I will identify a DMU$_k$ with its input-output-combination $(x_{1k}, \ldots, x_{mk}, y_{1k}, \ldots, y_{sk}) = (\boldsymbol{x}_k, \boldsymbol{y}_k)$. The concept of DEA is to measure the efficiency of n given DMUs, relative to the efficient frontier. The efficient frontier is seen as a production function from R^m to R^s, where every point on this production function symbolizes the optimal amount of outputs, that can be produced with the given inputs. In the concept of DEA the shape of this production function isn't a priori given, but is rather the result of the underlying data. The key to calculate the efficiencies of the DMU's is to solve specific optimization problems. Since all of the models presented in this paper can be written as Linear Problems (LP) we can solve them numerically, for example with the SIMPLEX algorithm. I am going to start the explanation of DEA by presenting the radial measures, because they are very suitable to communicate the motivation and main idea of DEA. Afterwards I will proceed with the specific Slack-Based Model (SBM), we used for conducting the banks efficiencies, and the concept of superefficiency in the SBM. Finally the Malmquist index as a possibility to quantify the change in efficiency will be presented.

2.1 Radial Models

When Charnes, Cooper and Rhodes first introduced DEA, they measured the efficiency for a DMU$_k$ by solving the following optimization problem (2.1):

$$\begin{aligned} \theta_k^* = \max \quad & \theta_k = \frac{\sum_{r=1}^{s} u_{rk} y_{rk}}{\sum_{i=1}^{m} v_{ik} x_{ik}} \\ \text{s.t.} \quad & 1 \geq \frac{\sum_{r=1}^{s} u_{rk} y_{rj}}{\sum_{i=1}^{m} v_{ik} x_{ij}} \qquad j = 1, \dots, n \\ & 0 \leq v_{ik}, i = 1, \dots, m \\ & 0 \leq u_{rk}, r = 1, \dots, s \end{aligned} \tag{2.1}$$

The $x_{ij}, i = 1, \dots, m, j = 1, \dots, n$ and the $y_{rj}, r = 1, \dots, s, j = 1, \dots, n$ are the given input/output factors of the DMUs in the sample. The interpretation of this problem is that DEA tries to maximize the ratio of weighted output to weighted input, by choosing the weights (u_{rk}, v_{ik}) optimal. The constraint normalizes this ratio to 1, meaning that every DMU$_k$ with $\theta_k^* = 1$ is efficient in the sense, that you can chose weights such that no other DMU has a better weighted output to weighted input ratio. Rewritting this optimization problem into a linear problem and observing the dual form leads to the following problem (2.2):

$$\begin{aligned} \theta_k^* = \min \quad & \theta_k \\ \text{s.t.} \quad & \theta_k x_{ik} \geq \sum_{j=1}^{n} \lambda_{jk} x_{ij} \qquad i = 1, \dots, m \\ & y_{rk} \leq \sum_{j=1}^{n} \lambda_{jk} y_{rj} \qquad r = 1, \dots, s \\ & 0 \leq \lambda_{jk}, j = 1, \dots, n \end{aligned} \tag{2.2}$$

It holds that $0 < \theta_k^* \leq 1$. By adding the constraint $\sum_{j=1}^{n} \lambda_j = 1$, we get the radial inputoriented model, introduced by Banker, Charnes and Cooper (BCC model). The BCC model accounts for Variable Returns to Scale (VRS). VRS means that the rate of change in output relative to the associated increase in input can be different depending on the level of input.

As mentioned before, during the DEA process the efficient frontier gets constructed. This efficient frontier is the set of optimal input-output-combinations implied by the observed DMUs. The Production Possibility Set (P) contains all the points that could be reached under the given circumstances. Analytically P is defined as:

$$\left\{ (x_1, \dots, x_m, y_1, \dots, y_s) | x_i \geq \sum_{j=1}^{n} \lambda_j x_{ij} \quad \forall i, y_r \leq \sum_{j=1}^{n} \lambda_j y_{rj} \quad \forall r, \lambda_j \geq 0 \quad \forall j, \sum_{j=1}^{n} \lambda_j = 1 \right\}.$$

The efficient frontier is the border of the set P. Figure 2.1 shows the production possibility set, spanned by the DMU$_k, k = 1, \dots, 5$. Given the definition of P the dual problem calculates the fraction θ_k^* to which the inputs can be reduced, such that $(\theta_k^* \boldsymbol{x}_k, \boldsymbol{y}_k) \in P$. Figure 2.1 shows how the input-output-combination of the inefficient DMU$_6$ gets projected on the efficient frontier spanned by the DMU$_k, k = 1, \ldots, 5$, by reducing the input with the factor θ_6^*. This fact is the reason why these models are referred to as 'radial' and inputoriented, because they measure the extend to which all input factors can be reduced radially without leaving P. The discussion on radial, outputoriented measures is similar and can be found in Cooper,Seiford and Tone (see [8]).

The major drawback of radial measures is, that $\theta_k^* = 1$ doesn't ensure that the DMU$_k$ is Pareto-Koopman efficient. [1]

This drawback is due to the occurance of slacks, referred to as "mix inefficiencies". In Figure 2.2 one can see how the radial reduction works in a model with VRS .In this example every DMU$_k, k = 1, \ldots, 6$ uses two different inputs to produce one unit of the output. The inefficient input-output-combination of DMU$_6$ gets projected on the efficient frontier. This specific part of the efficient frontier is a convex combination of the input-output-combinations of DMU$_2$ and DMU$_3$, therefore DMU$_2$ and DMU$_3$ are peers for DMU$_6$. The point $(\theta_6^* x_{16}, \theta_6^* x_{26})$ is Pareto-Koopman efficient, while the radial projection of the point (x_{15}, x_{25}) is not $(x_{14} < \theta_5^* x_{15}$ while $x_{24} = \theta_5^* x_{25})$.

A detailed discussion on slacks as well as the division of technical inefficiency in mix-inefficiency, scale inefficiency and pure technological inefficiency can be found in Cooper, Seiford and Tone [8].

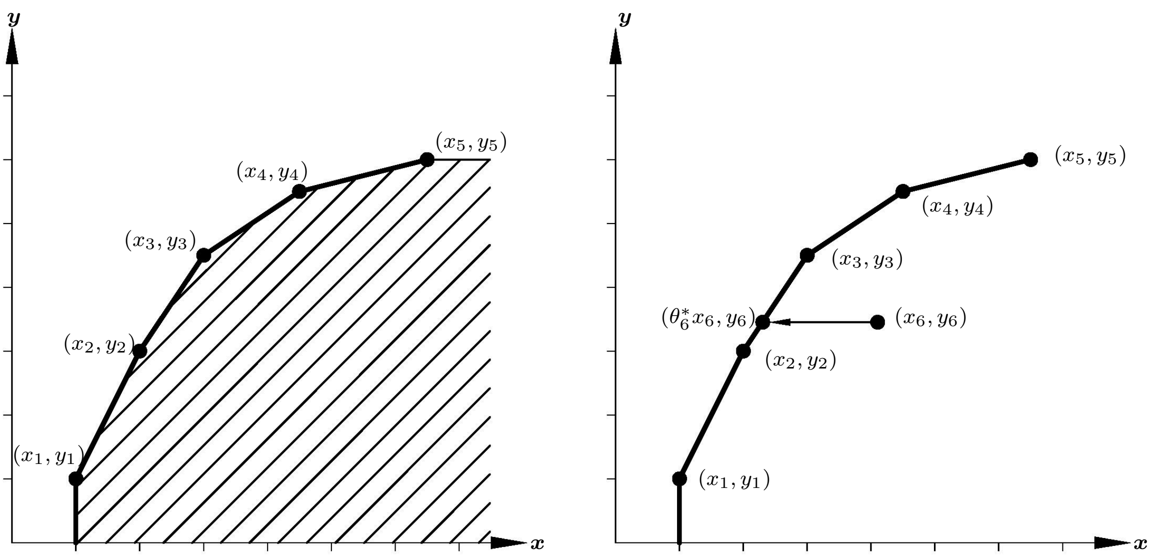

Figure 2.1: **Left**: Production Possibility Set in a single-input-single-output model with variable returns to scale.
Right: Projection of an inefficient input-output combination on the efficient frontier.

[1] Definition: A DMU is Pareto-Koopman efficient if and only if it is not possible to improve any input or output without worsening some other input or output.

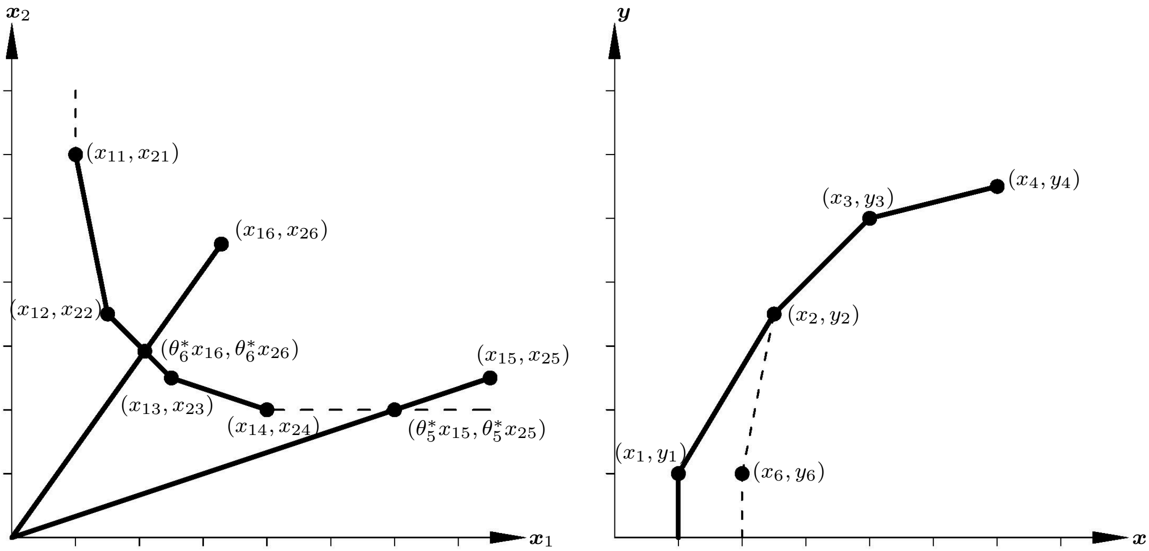

Figure 2.2: **Left**: Radial projection on the frontier in a variable returns to scale model with two inputs and one output. The output is 1 for every DMU.
Right: The efficient frontier changes when excluding the DMU1 from the sample. The plot shows again a single-input-single-output production possibility set.

2.2 Slack Based Model and Super Slack Based Model

A measure that ensures Pareto-Koopman's efficiency is the measure calculated by the so called Slack Based Model (SBM). There are several possibilities to implement SBM. In this study we used a non-oriented SBM with VRS for the calculation of the banks' efficiency measures. The efficiency measure for the DMU$_k$ in the slack based model is calculatated by solving the following optimization problem (2.3):

$$
\begin{aligned}
\rho_k^* = \min \quad & \rho_k = \frac{1-(1/m)\sum_{i=1}^m z_i^-/x_{ik}}{1+(1/s)\sum_{r=1}^s z_r^+/y_{rk}} \\
\text{s.t.} \quad & x_{ik} = \sum_{j=1}^n x_{ij}\lambda_{jk} + z_i^-, i = 1,\dots,m \\
& y_{rk} = \sum_{j=1}^n y_{rj}\lambda_j - z_r^+, r = 1,\dots,s \\
& 1 = \sum_{j=1}^n \lambda_{jk} \\
& 0 \le \lambda_{jk}, j = 1,\dots,n \\
& 0 \le z_i^-, i = 1,\dots,m \\
& 0 \le z_r^+, r = 1,\dots,s
\end{aligned}
\tag{2.3}
$$

The variables z_i^- and z_r^+ are called input slacks resp. output slacks. The variables represent the input excesses and the output shortfall to the efficient frontier. For the remaining study I will call a DMU$_k$ with $\rho_k^* = 1$ efficient.
The relationship between the two measures ρ_k^* and θ_k^* can be characterised with the following facts:

1. $\rho_k^* \leq \theta_k^*$
2. If there are no "mix inefficiencies": $\theta_k^* = 1 \Rightarrow \rho_k^* = 1$
3. A DMU_k is Pareto-Koopman efficient if and only if $\rho_k^* = 1$

A method for rewriting the above model into a linear model can be found in [12]. The reference set for a DMU_k is defined as:

$$E_k := \left\{ j \in 1, \ldots, n | \lambda_{jk}^* > 0 \right\},$$

where the λ_{jk}^* are the optimal decisions calculated in the previous optimization problem (2.3).[2] Every element of E_k is efficient and and is further called a peer to the DMU_k. The projection on the efficient frontier $(\hat{x}_k, \hat{y}_k)$ can be written as a convex linear combination of its peers:

$$\hat{x}_{ik} = \sum_{j \in E_k} x_{ij} \lambda_{jk}^* \quad \forall i = 1, \ldots, m \qquad \hat{y}_{rk} = \sum_{j \in E_k} y_{rj} \lambda_{jk}^* \quad \forall r = 1, \ldots, s$$

Another interesting task is to quantify how strong the efficiency of an efficient DMU is. A possibility to do this is the concept of super-efficiency. The idea behind super-efficiency is as follows: if a DMU_k is efficient it is part of the efficient frontier. If we remove this DMU_k from the sample the efficient frontier changes and the DMU_k has to lie outside of this new production possibility set (see Figure 2.2). Then we measure the distance of the DMU_k to this new production possibility set. In Figure 2.2 one can see how the production possibility set changes when excluding the DMU_1 from the sample.

To calculate the super-efficiency of DMU_k we use a two step approach. The first step is to calculate ρ_k^* by solving the optimization problem (2.3). If $\rho_k^* = 1$ the second step is to solve the following optimization problem (2.4):

$$\begin{aligned}
\gamma_k^* = \min \quad & \gamma_k = \frac{1 + (1/m) \sum_{i=1}^m w_i^- / x_{ik}}{1 - (1/s) \sum_{r=1}^s w_r^+ / y_{rk}} \\
\text{s.t.} \quad & x_{ik} = \sum_{j=1, j \neq k}^n x_{ij} \lambda_{jk} - w_i^-, i = 1, \ldots, m \\
& y_{rk} = \sum_{j=1, j \neq k}^n y_{rj} \lambda_{jk} + w_r^+, r = 1, \ldots, s \\
& 1 = \sum_{j=1, j \neq k}^n \lambda_{jk} \\
& 0 \leq \lambda_{jk}, j = 1, \ldots, j-1, j+1, \ldots, n \\
& 0 \leq w_i^-, i = 1, \ldots, m \\
& 0 \leq w_r^+ \leq y_{rk}, r = 1, \ldots, s
\end{aligned} \tag{2.4}$$

This second step, the calculation of the super-efficiency, has a feasible solution if and only if $\rho_k^* = 1$ in the first step. It holds that $\gamma_k^* \geq 1$. The super-efficiency or Super Slack Based Measure (SSBM) for a DMU_k is:

[2]The λ_{jk}^* are not necessarily unique

- ρ_k^* if $\rho_k^* < 1$
- γ_k^* if $\rho_k^* = 1$

Possibilities for numerically calculating superefficiency can be found in [12].

2.3 Malmquist Index

One important task during this project was to analyse if banking efficiency has changed during the crisis and which banks performed better and which performed worse. Therefore we needed a concept to measure change in efficiency. The Malmquist index introduced by Professor Sten Malmquist was originally used to compare the production technology of two economies. In the context of DEA a Malmquist index can be constructed to evaluate the productivity change of a DMU between two time periods. For the construction of the Malmquist Index it is necessary to introduce the following notation:

$$(x_{1k}^t, \ldots, x_{mk}^t, y_{1k}^t, \ldots, y_{sk}^t) = (\boldsymbol{x}_k^t, \boldsymbol{y}_k^t)$$

is the input output combination of the DMU$_k$ for the time period t and $P^t :=$

$$\left\{(x_1, \ldots, x_m, y_1, \ldots, y_s) | x_i \geq \sum_{j=1}^{n} \lambda_j x_{ij}^t \quad \forall i, y_r \leq \sum_{j=1}^{n} \lambda_j y_{rj}^t \quad \forall r, \lambda_j \geq 0 \forall j, \sum_{j=1}^{n} \lambda_j = 1\right\}$$

is the production possibility set for the time period t. If we identify the DMU$_k$ with its input-output-combination $(\boldsymbol{x}_k, \boldsymbol{y}_k)$ we can write $\rho^s(\boldsymbol{x}_k^t, \boldsymbol{y}_k^t)$ for the efficiency of the input-output-combination of DMU$_k$ for the time period t compared to the production possibility set for the time period s, P^s. Using this notation we can define the Malmquist Index, the Catch-Up and the Frontiershift from period t to s:

- Malmquist Index: MQ $:= \left[\frac{\rho^t(\boldsymbol{x}_k^s, \boldsymbol{y}_k^s)}{\rho^t(\boldsymbol{x}_k^t, \boldsymbol{y}_k^t)} \times \frac{\rho^s(\boldsymbol{x}_k^s, \boldsymbol{y}_k^s)}{\rho^s(\boldsymbol{x}_k^t, \boldsymbol{y}_k^t)}\right]^{1/2}$
- Catch-Up: CU $:= \frac{\rho^s(\boldsymbol{x}_k^s, \boldsymbol{y}_k^s)}{\rho^t(\boldsymbol{x}_k^t, \boldsymbol{y}_k^t)}$
- Frontier Shift: FS $:= \left[\frac{\rho^t(\boldsymbol{x}_k^t, \boldsymbol{y}_k^t)}{\rho^s(\boldsymbol{x}_k^t, \boldsymbol{y}_k^t)} \times \frac{\rho^t(\boldsymbol{x}_k^s, \boldsymbol{y}_k^s)}{\rho^s(\boldsymbol{x}_k^s, \boldsymbol{y}_k^s)}\right]^{1/2}$

One can easily see that MI $=$ FS $\times$ CU. Therefore the Malmquist index combines the movement of the efficient frontier and the change in relative efficiency.

For the calculation of the Malmquist index and its components we need the within scores $\rho_k^t(\boldsymbol{x}_k^t, \boldsymbol{y}_k^t)$ and $\rho_k^s(\boldsymbol{x}_k^s, \boldsymbol{y}_k^s)$, as well as the intertemporal scores $\rho_k^t(\boldsymbol{x}_k^s, \boldsymbol{y}_k^s)$ and $\rho_k^s(\boldsymbol{x}_k^t, \boldsymbol{y}_k^t)$. For the within scores ($t = s$) you can use either the SBM (2.3) or the SSBM (2.4) measure from the previous chapter. [8] refers to this distinction as the inclusive scheme, using SBM, and the exclusive scheme, using SSBM. For the task of measuring the efficiency change of the european banks we used the SSBM measure for the within scores.

For the computation of the intertemporal scores we need two different optimization problems:

$$
\begin{aligned}
\rho_k^{s*}(\boldsymbol{x}_k^t, \boldsymbol{y}_k^t) = \min \quad \rho_k^s(\boldsymbol{x}_k^t, \boldsymbol{y}_k^t) &= \frac{1-(1/m)\sum_{i=1}^m z_i^-/x_{ik}^t}{1+(1/s)\sum_{r=1}^s z_r^+/y_{rk}^t} \\
\text{s.t.} \quad x_{ik}^t &= \sum_{j=1}^n x_{ij}^s \lambda_{jk} + z_i^-, i=1,\dots,m \\
y_{rk}^t &= \sum_{j=1}^n y_{rj}^s \lambda_{jk} - z_r^+, r=1,\dots,s \\
1 &= \sum_{j=1}^n \lambda_{jk} \\
0 &\le \lambda_{jk}, j=1,\dots,n \\
0 &\le z_i^-, i=1,\dots,m \\
0 &\le z_r^+, r=1,\dots,s
\end{aligned}
\tag{2.5}
$$

and

$$
\begin{aligned}
\gamma_k^{s*}(\boldsymbol{x}_k^t, \boldsymbol{y}_k^t) = \min \quad \gamma_k^s(\boldsymbol{x}_k^t, \boldsymbol{y}_k^t) &= \frac{1+(1/m)\sum_{i=1}^m w_i^-/x_{ik}^t}{1-(1/s)\sum_{r=1}^s w_r^+/y_{rk}^t} \\
\text{s.t.} \quad x_{ik}^t &= \sum_{j=1}^n x_{ij}^s \lambda_{jk} - w_i^-, i=1,\dots,m \\
y_{rk}^t &= \sum_{j=1}^n y_{rj}^s \lambda_{jk} + w_r^+, r=1,\dots,s \\
1 &= \sum_{j=1}^n \lambda_{jk} \\
0 &\le \lambda_{jk}, j=1,\dots,n \\
0 &\le w_i^-, i=1,\dots,m \\
0 &\le w_r^+ \le y_{rk}, r=1,\dots,s
\end{aligned}
\tag{2.6}
$$

The first optimization problem (2.5) is used if the input-output-combination of $(\boldsymbol{x}_k^t, \boldsymbol{y}_k^t)$ lies inside the production possibility set P^s. The second optimization problem (2.6) is used if $(\boldsymbol{x}_k^t, \boldsymbol{y}_k^t)$ lies outside of P^s. For every $(\boldsymbol{x}_k^t, \boldsymbol{y}_k^t)$ either problem (2.5) or problem (2.6) has a feasible solution. The intertemporal scores are

- ρ_k^{s*} if there is a feasible solution for (2.5)
- γ_k^{s*} if there is a feasible solution for (2.6)

In the case $s = t$ the above problems are exactly the optimization problems, introduced in the previous chapter, for calculating the super-efficiency. For a further discussion of the Malmquist Index see [8].

Chapter 3

Regressionanalysis

3.1 Panel Data Models

Panel data are very common in macroeconomic or microeconomic problems. Panel data consist of observations for N individuals for T periods. Despite the fact, that there are often not enough observations to consistently estimate a linear modell for each individual, which makes pooling the data necessary, there are some advantages of using Panel modells:
One benefit of panel modells is, that they are capable of observing heterogeneity and dependencies between individuals. Additionally it is possible to analyze if there are some unobserved heterogeneities and serial correlation over the time periods. A linear model for panel data has the general form

$$y_{it} = \alpha + X'_{it}\beta + u_{it} \tag{3.1}$$

where $\alpha \in \mathbb{R}$, $X_{it} \in \mathbb{R}^{K\times 1}$ and $\beta \in \mathbb{R}^{K\times 1}$ and α and β are unknown parameters. In the general case the residuals u_{it} take the form

$$u_{it} = \mu_i + \lambda_t + \nu_{it}. \tag{3.2}$$

Depending on the actual characteristics of individual effects, μ_i, and the time effects, λ_t, there are different methods for estimating the parameters β and α. The errors ν_{it} are often referred to as idiosyncratic errors. Further we will assume that the idiosyncratic errors are independent and identically distributed with $\mathbb{E}(\nu_{it}) = 0$, $\mathbb{E}(\nu_{it}^2) = \sigma_\nu^2$ and uncorrelated with the regressors.

The most simple method for estimating α and β is using the simple OLS estimators for 3.1. Two important assumptions to ensure that the OLS estimators are unbiased and consistent are (see [18]):

- $\mathbb{E}(u|X) = 0$ (zero conditional mean assumption)
- $\mathbb{E}(X'u) = 0$ (population orthogonality assumption)

where u is the vector of the errors u_{it} and X is the regressor matrix as in (3.4). The zero conditional mean assumption is not fulfilled if $\exists i : \mathbb{E}(\mu_i) \neq 0$ or $\exists t : \mathbb{E}(\lambda_t) \neq 0$. To use the OLS standard variance estimators we further need the assumption that

$\mathbb{E}(u^2|X) = \sigma^2$, which is not fulfilled if $\exists i : \mathbb{E}(\mu_i^2) \neq 0$ or $\exists t : \mathbb{E}(\lambda_t^2) \neq 0$. In the following sections I am going to introduce methods to solve this problems, namely the one-way fixed effects (FE) estimator and the one-way random effects (RE) estimator for models without time effects ($\lambda_t = 0, \forall t = 1, \ldots, T$). A detailed discussion on two-way error component models can be found in [7]. Furthermore I will only present results for balanced panels, which means that for every individual observations for the same time periods are available.

3.2 With-in and Between Transformation

If you assume the individual error terms μ_i in 3.2 to be fixed, you can write (3.1) in vector notation as:

$$y = X\beta + Z_\mu \mu + \nu \tag{3.3}$$

with

$$\nu = \begin{pmatrix} \nu_{11} \\ \nu_{12} \\ \vdots \\ \nu_{1T} \\ \nu_{21} \\ \vdots \\ \nu_{N1} \\ \vdots \\ \nu_{NT} \end{pmatrix} \qquad y = \begin{pmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1T} \\ y_{21} \\ \vdots \\ y_{N1} \\ \vdots \\ y_{NT} \end{pmatrix} \qquad X = \begin{pmatrix} X'_{11} \\ X'_{12} \\ \vdots \\ X'_{1T} \\ X'_{21} \\ \vdots \\ X'_{N1} \\ \vdots \\ X'_{NT} \end{pmatrix} \tag{3.4}$$

$Z_\mu = I_N \otimes \iota_T$, where I_N is an identity matrix of dimension N, ι_T is a vector of ones with dimension T and $\otimes$ denotes the Kronecker Product. Due to the fact that the vector $\alpha\iota_{NT}$ is linear dependent from the columns of Z_μ, most statistical softwares omit the parameter α and therefore I omit α as well. Using OLS to estimate (3.3) leads to the so called Least Square Dummy Variable (LSDV) estimator. Using the Frisch-Waugh-Theorem it is possible to estimate β by transforming the equation (3.3) to

$$Qy = QX\beta + Q\nu, \tag{3.5}$$

where $Q = I_{NT} - P$ and $P = Z_\mu(Z'_\mu Z_\mu)^{-1}Z'_\mu$. Q is the orthogonal projection on Z_μ. P and Q are both symmetric, idempotent and $PQ = 0$. P can also be written as $P = I_N \otimes \bar{J}_T$, where $\bar{J}_T = \frac{1}{T}J_T$ and J_T is a $T \times T$ matrix of ones.

$$Z_\mu = \begin{pmatrix} \iota_T & 0 & \cdots & \cdots & 0 \\ 0 & \iota_T & 0 & \cdots & 0 \\ \vdots & 0 & \iota_T & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & \ldots & \iota_T \end{pmatrix} \qquad P = \begin{pmatrix} \bar{J}_T & 0 & \cdots & \cdots & 0 \\ 0 & \bar{J}_T & 0 & \cdots & 0 \\ \vdots & 0 & \bar{J}_T & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & \ldots & \bar{J}_T \end{pmatrix} \tag{3.6}$$

The interpretation of P and Q is now very easy once you have the above representation. P takes the individual mean over time for every individual, which means that $Py = (\bar{y}_{1\cdot}, \bar{y}_{2\cdot}, \dots, \bar{y}_{N\cdot})' \otimes \iota_T$, where $\bar{y}_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T} y_{it}$. You can easily see that Qy demeans every observation with the individual mean over time. Therefore the transformation (3.5) is called the with-in transformation and the analogue transformation using P is called between transformation.

3.3 Fixed Effects Model

In the fixed effects model, the individual-specific effects μ_i are assumed to be deterministic (fixed) and the remainder disturbances ν_{it} are independent and identically distributed $IID(0, \sigma_\nu^2)$. We also assume that the X_{it} are independent of the ν_{it} for all i and t. This specification lead to the linear model (3.3). Estimating β using OLS on (3.3) or equivalently using OLS on the transformed model (3.5) results in the fixed effects or with-in estimator

$$\hat{\beta}_{FE} = (X'QX)^{-1}X'Qy.$$

Under the assumption that the true model is the fixed effects model the above estimators for β and μ_i are the best linear unbiased estimators (BLUE). Under standard assumptions for the idiosyncratic errors ν_{it} and the regressors X_{it}, the estimator $\hat{\beta}_{FE}$ is consistent for $NT \to \infty$ while the estimators $\hat{\mu}_i$ are only consistent for $T \to \infty$ (see [18]).

3.3.1 Random Effects Model

Contrary to the fixed effects model the individual errors μ_i are assumed to be stochastic (random) in the random effects model. The individual effects μ_i are assumed to be independent and identically distributed and independent from the idiosyncratic error ν_{it}. Given these specifications the variance-covariance matrix Σ for the $u_{it} = \mu_i + \nu_{it}$ is given as:

$$\Sigma = \mathbb{E}(uu') = Z_\mu \mathbb{E}(\mu\mu')Z_\mu' + \mathbb{E}(\nu\nu') = \sigma_\mu^2(I_N \otimes J_T) + \sigma_\nu^2(I_N \otimes I_T) = \tag{3.7}$$

$$= \begin{pmatrix} B_T & 0 & \cdots & \cdots & 0 \\ 0 & B_T & 0 & \cdots & 0 \\ \vdots & 0 & B_T & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & \dots & B_T \end{pmatrix} \tag{3.8}$$

with

$$B_T = \begin{pmatrix} \sigma_\mu^2 + \sigma_\nu^2 & \sigma_\mu^2 & \cdots & \cdots & \sigma_\mu^2 \\ \sigma_\mu^2 & \sigma_\mu^2 + \sigma_\nu^2 & \sigma_\mu^2 & \cdots & \sigma_\mu^2 \\ \vdots & \sigma_\mu^2 & \sigma_\mu^2 + \sigma_\nu^2 & \dots & \sigma_\mu^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_\mu^2 & \sigma_\mu^2 & \vdots & \dots & \sigma_\mu^2 + \sigma_\nu^2 \end{pmatrix} \in \mathbb{R}^{T\times T},$$

where σ_μ^2 is the variance of the individual effects and σ_ν^2 is the variance of the idiosyncratic errors. This implies that the errors are homoskedastic, serial correlated within one individual and uncorrelated between observations from different individuals:

$$\text{cov}(u_{it}, u_{js}) = \begin{cases} \sigma_\mu^2 + \sigma_\nu^2, & \text{for } i = j, t = s \\ \sigma_\mu^2, & \text{for } i = j, t \neq s \\ 0, & \text{for } i \neq j \end{cases}$$

Using the matrices P and Q we can decompose Σ into:

$$\Sigma = \sigma_1^2 P + \sigma_\nu^2 Q, \tag{3.9}$$

with $\sigma_1^2 = T\sigma_\mu^2 + \sigma_\nu^2$. Using $PP = P$, $QQ = Q$ and $PQ = 0$ it is easy to verify that

$$\Sigma^{-1} = \frac{1}{\sigma_1^2} P + \frac{1}{\sigma_\nu^2} Q \quad \text{and} \quad \Sigma^{-1/2} = \frac{1}{\sigma_1} P + \frac{1}{\sigma_\nu} Q. \tag{3.10}$$

Transforming (3.1) using $\Sigma^{-1/2}$ is called RE transformation. The GLS estimator in the RE model is given as:

$$\hat{\beta}_{RE} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y = (X'QX + \frac{\sigma_\nu^2}{\sigma_1^2}X'PX)^{-1}(X'Qy + \frac{\sigma_\nu^2}{\sigma_1^2}X'Py). \tag{3.11}$$

Given this form one can see that the GLS estimator is a weigthed average of the within estimator $\hat{\beta}_{FE}$ and the between estimator $\hat{\beta}_{Between}$. The between estimator is the OLS estimator for

$$\bar{y}_i = \bar{x}_i\beta + u_i, \tag{3.12}$$

where $\bar{y}_i$ and $\bar{x}_i$ are the individual averages. Due to the fact, that σ_1^2 and σ_ν^2 are typically unknown, we have to consistently estimate these parameters in order to obtain a feasible GLS (FGLS) estimator. There are several methods from Amemiya, Wallace and Hussain, Swamy and Arora and Nerlove for estimating these parameters. The methods for estimating σ_1^2 and σ_ν^2 and a more detailed analysis on this topic can be found in [7].

3.3.2 FGLS Estimator

A third method for estimating β is to use a more general FGLS estimator than the RE Estimator is (see (3.11)). When assuming random individual effects the variance matrix has the structure presented above (see (3.8)). This structure is very special and neglects every correlation between observations of different individuals and also imposes a very peculiar form of autocorrelation. A more general approach is to assume the variance matrix has the following structure:

$$\Sigma_{FGLS} = \mathbb{E}(uu') = \begin{pmatrix} V_T & 0 & \cdots & \cdots & 0 \\ 0 & V_T & 0 & \cdots & 0 \\ \vdots & 0 & V_T & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & \cdots & V_T \end{pmatrix} \tag{3.13}$$

where V_T is an arbitrary matrix in $\mathbb{R}^{T\times T}$. This estimation method allows a more general structure for the autocorrelation within the same individual. A simple estimator for V_T is:

$$\hat{V}_T = \sum_{i=1}^{N} \hat{u}_i \hat{u}_i', \tag{3.14}$$

with $\hat{u}_i \in \mathbb{R}^T$ is the estimated error vector for the individual i. You can use the estimated errors from a simple OLS estimation. The FGLS estimator is then given as:

$$\hat{\beta}_{FGLS} = (X'\hat{\Sigma}_{FGLS}^{-1}X)^{-1}X'\hat{\Sigma}_{FGLS}^{-1}y \tag{3.15}$$

3.3.3 Consistency and assymptotic behaviour Fixed Effects and Random Effects Estimator

As already mentioned the FE estimator $\hat{\beta}_{FE}$ is consistent and asymptotically normal with $NT \to \infty$. However the RE estimator $\hat{\beta}_{RE}$ requires a more strict assumption on the individual effects, μ_i. For a consistent estimation it is necessary, that $\mathbb{E}(\mu_i|X_{i\cdot}) = \mathbb{E}(\mu_i) = 0$ (see [18]). Although the FE estimator is consistent and asymptotically normal under standard assumptions we also need a consistent variance matrix estimator to use the standard t- and Wald-teststatistics.

3.4 Tests on Serial Correlation

The occurance of serial correlation [1] in the errors is a problem for the correct estimation of the variance matrix for the estimated coefficients. Therefore we have to test if serial correlation is present in the idiosyncratic errors, to see if the model is correctly specified. A common test to check for serial correlation is the Breusch-Godfrey-test.

The Breusch-Godfrey-test is a test for autocorrelation in regression models. For any linear model:

$$y_t = \alpha + \boldsymbol{x}_t\beta + u_t, \quad u_t = \rho_1 u_{t-1} + \ldots + \rho_p u_{t-p} + \varepsilon_t \tag{3.16}$$

the Breusch-Godfrey-test tests the null hypothesis, that there is no serial correlation ($\rho_1 = \ldots = \rho_p = 0$), against the alternative, that at least one $\rho_i \neq 0$. The test statistic is given as nR^2, where R^2 is the coefficient of determination for the linear model:

$$\hat{u}_t = \alpha + \boldsymbol{x}_t\beta + \rho_1\hat{u}_{t-1} + \ldots + \rho_p\hat{u}_{t-p} + \varepsilon_t, \tag{3.17}$$

$n = T - p$ is the number of available observations and $\hat{u}_i$ are the OLS residuals of the regression 3.16. Breusch and Godfrey have shown that $nR^2 \sim \chi^2_p$ (see [19]).
To test for serial correlation in the idiosyncratic errors of panel data one has to first transform the data using a with-in or RE transformation and then run the Breusch-Godfrey-test on the transformed modell. In the case of an RE model this procedure is

[1] The residuals u_{it} of a panel regression model are called serial correlated if $\exists i,t,s : \mathrm{cov}(u_{it}, u_{is}) \neq 0$

appropriate under the assumption, that the idiosyncratic errors are heteroscedastic and serial uncorrelated (see [18], section 10.7.2). In the FE case the within residuals are serial correlated with $\text{cov}(u_{it}, u_{is}) = -1/(T-1)$ (see [18], section 10.5.4). While in the case of large T this correlation dies out, the test procedure is not applicable in panels with a short time dimension. For short FE panel models Wooldridge recommends another test. The Wooldridge-test is based on estimating

$$\hat{u}_{it} = \alpha + \delta \hat{u}_{i(t-1)}, \tag{3.18}$$

and testing if $\delta = -1/(T-1)$, where $\hat{u}_{it}$ are the within residuals. Under standard assumptions the standard t-statistic is asymptotically normal distributed, and therefore the t-statistic can be used for testing the null of $\delta = -1/(T-1)$. If the null hypothesis is rejected there is probably serial correlation that doesn't die out with $T \to \infty$. Both the Breusch-Godfrey-test for panel data as well as the Wooldridge-test for short FE panels are implemented in the plm-package [9].

3.5 Tests on Cross Sectional Dependence

Cross-sectional dependence [2] is a serious issue in panel data models, especially when using efficiency score from a previous DEA calculation. Efficiency scores from DEA are per definition interdependent with a very complicated dependence structure (see 5.1.2). We tested the regression errors for cross-sectional dependence using the Lagrange Multiplier (LM) statistic proposed by Breusch and Pagan and the test statistics introduced by Pesaran. Breusch and Pagans LM statistic is given as

$$CD_{LM} = T \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij}^2, \tag{3.19}$$

where $\hat{\rho}_{ij} = \left(\sum_{t=1}^{T} \hat{u}_{it}\hat{u}_{jt}\right)\left(\sum_{t=1}^{T} \hat{u}_{it}^2 \sum_{t=1}^{T} \hat{u}_{jt}^2\right)^{-1/2}$ is the pair-wise correlation of the residuals. Depending on the underlying model either the residuals from an OLS, within or RE estimation are used. Under the null hypothesis of no cross-sectional dependence ($\text{cov}(u_{it}, u_{jt}) = 0, \forall t, i \neq j$), the statistic is standard normal distributed, for the asymptotics when first $T \to \infty$ and then $N \to \infty$. However this isn't very suitable in cases where N is relatively large compared to T. Therefore Pesaran proposed a different test. In contrast to Breusch and Pagan, Pesaran doesn't use the squared pair-wise correlation. Pesarans test statistic is given as:

$$CD_{glob} = \sqrt{\frac{2}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \sqrt{T_{ij}}\hat{\rho}_{ij} \right). \tag{3.20}$$

T_{ij} is the quantity of observations to calculate $\hat{\rho}_{ij}$. In the case of a balanced panel $T_{ij} = T, \forall i, j$. Under the null hypothesis of no cross sectional dependence this test-statistic is normal distributed for N and T tending to infinity (see [15]) . Contrary to

[2] The residuals u_{it} of a panel regression model are called cross-sectional dependent if $\exists i, j, t, s$: $\text{cov}(u_{it}, u_{js}) \neq 0$

the CD_{LM} statistic the CD statistic tends to have good small sample properties even for small T. However the CD statistics inherits the problem, that it is likely to accept the null hypothesis if there are positive correlations as well as negativ correlation, which is quiet possible using DEA efficiency scores. A possible solution is to use the local version of the CD statistic, where you only use the correlations according to an a priori defined weight matrix, comparable to the weight matrices used for Spatial Dependence. The local version of Pesarans test is given as:

$$CD_{loc} = \sqrt{\frac{1}{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\omega_{ij}}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \omega_{ij} \sqrt{T_{ij}} \hat{\rho}_{ij} \right), \tag{3.21}$$

where $\omega_{ij} = 1$, if the individuals i and j are neighbours, and 0, if not.

3.6 Robust Covariance Estimation

We could observe in the data, that there is serial correlation as well as cross sectional dependence. Due to the fact, that neither the FGLS estimator nor the RE estimator can account for cross sectional dependence, we decided to use OLS estimates, together with a robust variance matrix estimation proposed by Driscoll and Kraay (see [11] and [10]). The main idea from Driscoll and Kraay is the following: given a bivariate (X_{it} is a scalar so $K = 1$) linear panel regression model:

$$y_{it} = X_{it}'\beta + u_{it}, \tag{3.22}$$

where the errors u_{it} can be serial correlated, as well as cross-sectional dependent. The standard OLS-estimator can be written as

$$\sqrt{T}(\hat{\beta}_{OLS} - \beta) = \frac{\sqrt{T} \sum_{t=1}^{T} \sum_{i=1}^{N} X_{it} u_{it}}{\left(\frac{1}{NT} \sum_{t=1}^{T} \sum_{i=1}^{N} X_{it}^2 \right) NT} = \frac{1}{Q_T} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} h_t, \tag{3.23}$$

with $h_t = \frac{1}{N} \sum_{i=1}^{N} X_{it} u_{it}$ and $Q_T = \frac{1}{NT} \sum_{t=1}^{T} \sum_{i=1}^{N} X_{it}^2$. The variance can now be written as:

$$V_T = \frac{1}{Q_T^2} \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} \mathbb{E}[h_t h_s] = \frac{S_T}{Q_T^2} \tag{3.24}$$

By defining the variable h_t Driscoll and Kraay reduced the original problem to a time-series estimation problem. Consistently estimating S_T will lead to an consistent estimator of V_T. In the general case ($K \geq 1$) Driscoll and Kraay proposed the following estimator for S_T:

$$\hat{S}_T = \hat{\Omega}_0 + \sum_{j=1}^{m(T)} (1 - \frac{j}{m(T)+1})[\hat{\Omega}_j + \hat{\Omega}_j'], \tag{3.25}$$

where $\hat{\Omega}_j = T^{-1}\sum_{t=j+1}^{T} h_t h'_{t-j}$, $m(T) = \mathcal{O}(T)$, $h'_t = N^{-1}\sum_{i=1}^{N} X_{it}\hat{u}_{it}$, where $\hat{u}_{it}$ are the OLS residuals from the regression $y_{it} = X_{it}\beta + u_{it}$. Driscoll and Kraay have shown that under general assumptions for Generalized Methods of Moments (GMM) estimators, this estimate for S_T is consistent. The resulting variance matrix estimate

$$\hat{V}_T = Q_T^{-1}\hat{S}_T Q_T^{-1} \text{ with } Q_T = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} X_{it}X'_{it}$$

for the OLS estimator $\hat{\beta}$ is robust against serial correlation and cross sectional dependence. This variance matrix estimate is implemented in R in the lmtest-package [20].

3.7 Sensitivity Analysis

"A fragile inference is not worth taking seriously" is the first sentence of Leamers working paper about Sensitivity Analysis (see [14]). Leamer deals with the problem that the estimate and significance of a coefficient in a regression model depends on the set of regressors, i.e. adding or removing regressors can alter the conclusions drawn on the impact of the regressor under consideration. Therefore Leamer suggested an extreme bound analysis, where you consider a set of test models. The test models are constructed by adding regressor variables out of a set of candidate variables (test variables) to your base model, which consists of the regressors under consideration (base variables). The aim is to find base variables that are highly robust against alteration of the test variables. The procedure works as follows:

1. Choose base variables $\boldsymbol{x}_{it} = (x_{it}^1, \ldots, x_{it}^k)$ and test variables $\boldsymbol{z}_{it} = (z_{it}^1, \ldots, z_{it}^m)$
2. For every combination of testvariables $\boldsymbol{z}_{it}^* = (z_{it}^{j_1}, \ldots, z_{it}^{j_l}), l \leq m, 1 \leq j_1 < j_2 < \ldots < j_l \leq m$ one estimates the regressionmodel $y_{it} = \boldsymbol{x}_{it}\beta + \boldsymbol{z}_{it}^*\gamma + u_{it}$, and conducts a two-sided t-test for the components of $\hat{\beta}$. For m test variables it is possible to construct $2^m - 1$ different test models.
3. If the sign of the estimate is definite and the t-test for a base variable is significant in all the test models, the variable is called robust

Chapter 4

Data

4.1 Data Acquisition

We obtained the data for the 70 banks in the sample by studying the published annual reports for the years 2003-2012 [1]. All of the 70 banks were listed on the ECB list of significant financial institutions chosen to be under control of the SSM (see [6]). Some of the banks on the ECB list had to be removed, because we didn't consider them suitable for our analysis (for example the RCI Banque or the Volkswagenbank). For the rest of the banks we were not able to get their annual reports. Before the introduction of the IFRS standard the big banking groups published balance sheet positions according to the different national laws. While there are some balance sheet positions that are rather easy to obtain and that are pretty similar for the different national regulations, like Total Assets, Cash and Balances at Central Banks, Customer Loans etc., there are a lot of key performance indicators, that are not normalized, like the provision for Loan Losses, the maturity distribution of Claims/Liabilities or the Debt securities from public entities. We were often forced to use proxy performance indicators, because the banks didn't publish the ones we were looking for. Due to these more or less exotic indicators we were not able to use commercial databases like Bankscope or CaptalIQ, which lack of this information. Also the portrayal and splitting of the balance sheet and income statement positions varied extremely for the different banks, which further complicated the data acquisition. For the collection of the macroeconomic variables we used public databases from the ECB, the Worldbank, the IMF and Eurostat.

4.2 DEA Dimensions

The motivation for the selection of the DEA factors was to combine the banks ability to fulfill their economic tasks, like maturity transformation, lot size transformation and risk transformation, on the one side and to be secured against economic shocks on the other side. For the DEA we used three input variables as well as three output variables. The three input variables are:

[1] Most of the annual reports can be found on the homepages of the respective institute. For some of the reports we had to directly contact the Investors Relations division of the banks.

- Total Assets (TA): The Total Assets are meant to reflect the size of the bank
- Risky Assets to Total Assets (RA.TA): The ratio of the Risky Assets to the Total Assets represents the risk load of the bank's business. We extract the Risky Assets from the Balance sheet by substracting positions with little or no risk potential like the Cash and Cash Equivalents (e.g. titles eligible for refinancing with central banks) and all debt securities issued by governmental institutions.
- Liabilities to Banks to Total Assets (LtB.TA): The ratio of the banks liabilities to other banks to the Total Assets is used to distinguish between banks, that rely heavily on refinancing through other banks and banks that are able to gain enough money through customer deposits and other long term loans.

On the output side we used the following three variables:

- Liquidity Production (LP): This variable combines the banks' loans and receivables from private customers and other banks as well as the banks holding of debt securities.This index should measure to which extent the banks fulfill their macroeconomic task of providing the economy with money.
- Information Production (IP): Banks play a very vital role in the economy by screening the market and gathering information on creditworthiness and the financial market situation. To get a grip on the performance of the banks in producing this information, we used an indirect measure. (Varying) annual allowances and direct write-downs in the loan category are interpreted as a malfunction of the banks screening process. Direct write-downs and an increase in allowances are clarified as miscalculated risk (at the time of the grant of the loan). We calculate the information production by substracting the income statement relevant devaluations of loans (write-downs and new allowances, without netting) from the amount of outstanding loans to customers and banks. This difference can be interpreted as the amount of performing loans. As a result the information production is the ratio of performing loans to total loans.
- Stability Index (SI): Here, we try to explore the sustainablity of the banks‘ capital. In that context, the leverage ratio Equity (Eq) to Risky Assets (RA) serves as a proxy for the contribution of the individual unit to (some kind of technical) stability of the financial system which is at the center of the Basel Accords.

4.3 Descriptive Statistics

We were able to collect balance sheet positions for 70 banking groups for the period ranging from 2003 to 2012 (see Table A.1), 14 country specific macroeconomic variables and 13 different global macroeconomic variables (see Table A.2). The period 2003-2012 is mainly characterized by the big financial crisis, that started in 2007 and the subsequent reactions of the financial and banking markets in 2008 and 2009 (see Figure 4.2 and Figure 1.1). Due to the fact, that most of the repercussions on the banks balance sheet positions can be seen in the years 2008 and 2009, we split the time range into three regimes (pre crisis: before 2007, crisis: 2008-2009 and post crisis: after 2009). We will later use this distinction for the regression analysis.

4.3.1 Enormous growth prior the crisis

In the years before the crisis (2003-2007) the banking sector grew enormously, with a median cumulative growth from 2003 to 2007 of 75%, this is equivalent to an yearly growth of 12% (see Figure 4.1). And about 36% of all the banking groups at least doubled their TA (see Table A.8). Only about 3% of the considered banks reduced their TA in this period (see Table A.8) This changes dramatically in the years 2008-2012, where the median yearly growth shrunk to about 1% and nearly 46% of the banks reduced their TA during this period (see Table A.8). However the variation between the banks his huge and there are some extreme (upward) outliers, that can mostly be explained by merging and acquisition activities. For example in 2007 the Banca Lombarda and the BPU Banca merged together into the Gruppo UBI, resulting in TA growth rates of 200% for the Banca Lombarda and 60% for the BPU Banca. The liquidity production, customer loans and risky assets behaved very similar to the total assets (see Table 4.1 and Figure 4.1). This behaviour is not very surprising considering that the median proportions for LP, CL and RA to TA are 0.87, 0.57 and 0.91 (see Table A.7). This boom in the banking sector was accompanied by an increase in stock prices. The Stoxx Euro 600 Banks Index increased by 58% from the end of 2003 to the end of 2006. The financial markets reacted immediately to the beginning of the financial crisis in 2007, with a decline of european bank stock prices of 17% in the year 2007 and another 64% in 2008, while in the same time the Euro Stoxx 50 gained nearly 7% in 2007 and lost 'only' 44% in 2008 (see Figure 4.2). The growth in liquidity production can also be seen in the growth rates of the monetary aggregate M1[2], which grew by nearly 18% in 2005 and then declined. The shares, included in the aggregates M2 and M3 (time deposits, money market funds, repurchase agreements and issued debt securities), grew about 19% in 2007 and 12% in 2008 and then decreased. Especially the components of M3 (repurchase agreements, money market fund shares and issued debt securities) reacted severly to the collapse of the interbanking market. The annual growth rates of this aggregate were -16% in 2009 and -23% in 2010.

4.3.2 Value adjustments and write-offs followed by recapitalization

In the years before the crisis the SI showed no significant movement, with median growth rates around 0, except for the year 2004 were the median value of the yearly difference of the SI was significant negativ with a change of -0.2 percentage points. In 2008 the SI showed an significant downward movement with an median change of -0.7 percentage points (see Table A.9). This drop is probably caused by massive devaluations and write-offs that primarily reduced the banks equity. This downward movement was encountered by an median increase of the SI in 2009 of 0.9 percentage

[2]For an exact definition of the monetary aggregates please visit https://www.ecb.europa.eu/stats/money/aggregates/aggr/html/index.en.html.

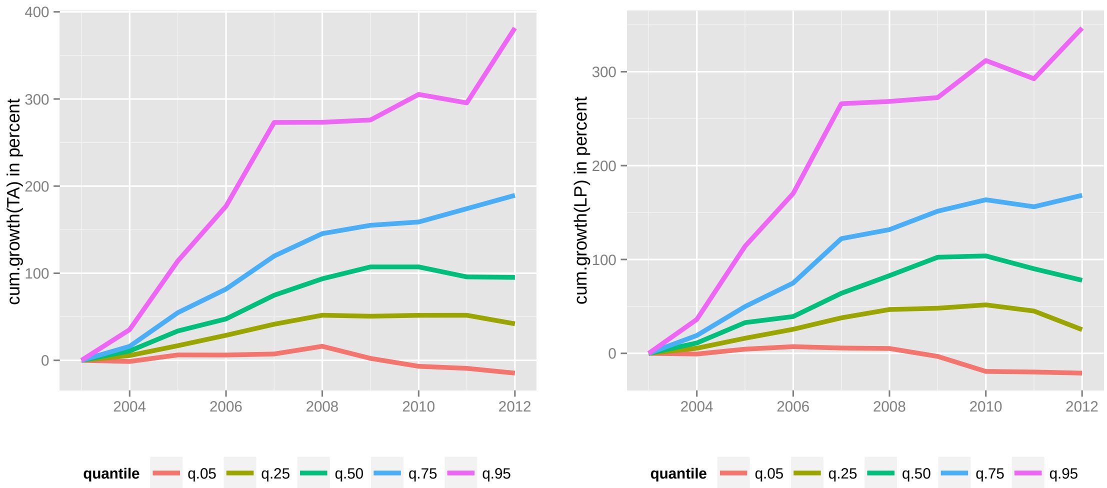

Figure 4.1: Cumulative growth of total assets(TA) and liquidity production(LP) since 2003. The plots show the quantiles of the variables for the 70 observed banks.

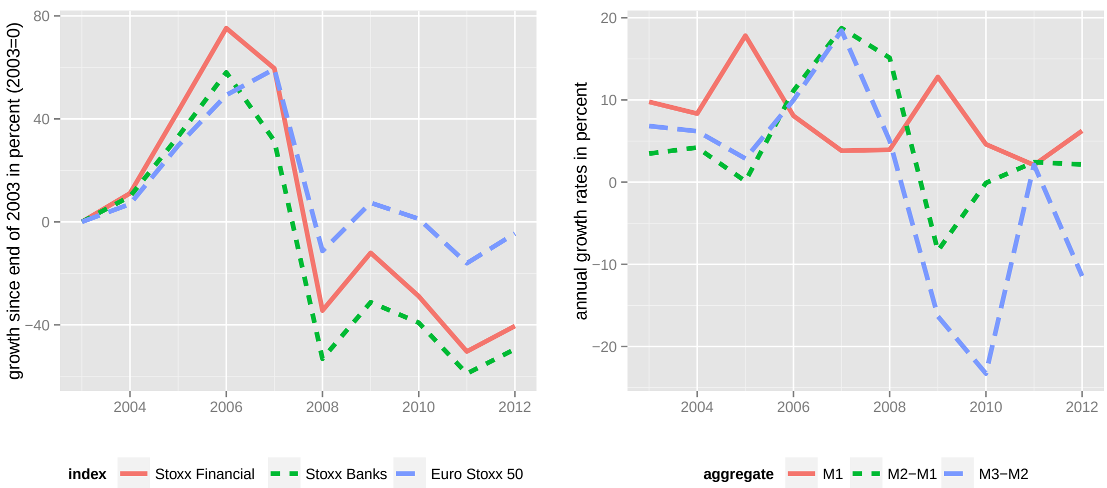

Figure 4.2: Movement of the Stoxx Euro 600 Financial Service, Stoxx Euro 600 Banks and Euro Stoxx 50 indices normalized at 2003. Growth rates of the monetary aggregates M1, M2 and M3.

	TA		LP		CL	
quantile	2003-2007	2007-2012	2003-2007	2007-2012	2003-2007	2007-2012
0%	-0.12	-0.65	-0.13	-0.71	-0.39	-0.74
5%	0.07	-0.39	0.06	-0.52	0.07	-0.38
25%	0.41	-0.03	0.38	-0.08	0.40	-0.04
50%	0.75	0.14	0.64	0.14	0.79	0.18
75%	1.20	0.43	1.22	0.37	1.20	0.45
95%	2.73	0.91	2.66	0.91	2.65	0.98
100%	3.38	2.29	3.45	2.00	9.61	11.28

Table 4.1: Total Assets, Liquidity Production and Customer Loans cumulative growth before and after the Crisis

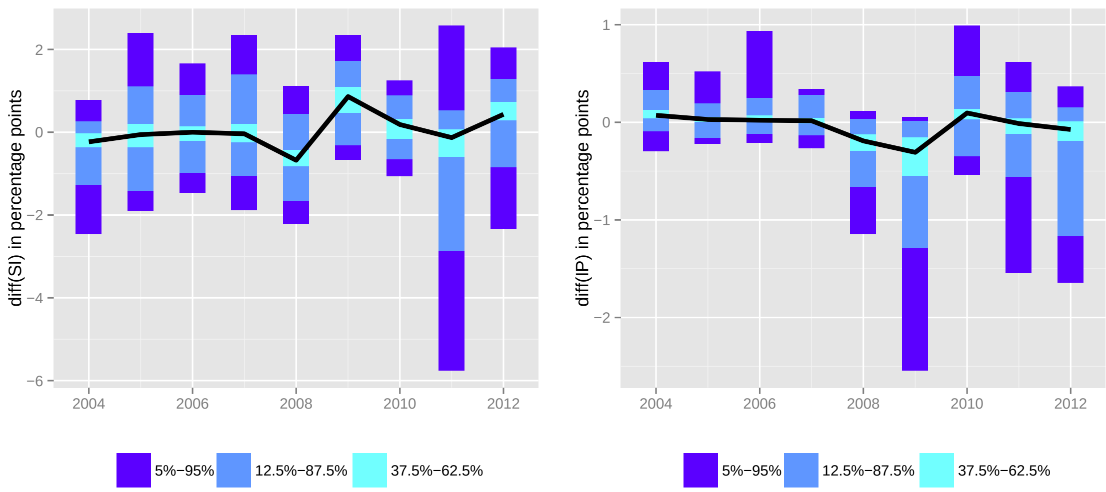

Figure 4.3: Yearly percentage point changes for the stability index (SI) and the information production (IP). The plots show the quantiles of the variables for the 70 observed banks. The continous line is the median value.

points (see Figure 4.3). This fast reaction in terms of an increase in capital was probably due to increased requirements on equity ratio and illustrates that the equity ratio has been the prime instrument for banking regulation. Similar to the stability index the information production was also quiet stable prior to the crises with yearly percentage changes around 0, except for the year 2004. This changes in the years 2008 and 2009 where the banks had to write-off many loans resp. evaluate existing loans. This led to an median drop in IP of -0.2 percentage points in 2008 and -0.3 percentage points in 2009 (see Figure 4.3 and Table A.9). Generally the banks reacted very differently to the events of the financial crises and there are some extreme downward outliers (see Figure 4.4). While some banks managed to get rid of most of their bad loans in the years 2008 and 2009 some other banks sustained further losses in their information production due to further write-offs (see Figure4.4). For example the Bank of Cyprus worsened their IP by -8 percentage points from 2007 to 2012.

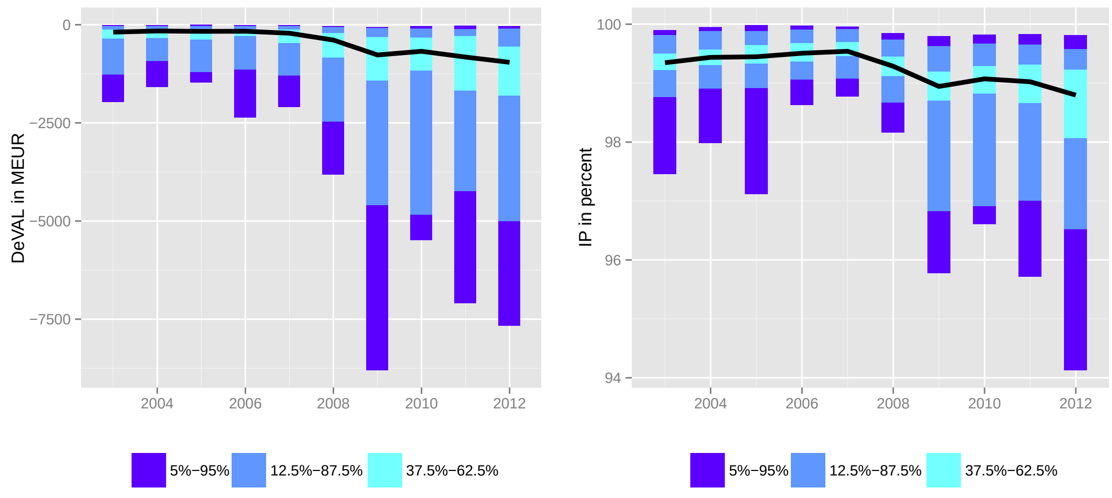

Figure 4.4: Development of the gross loan devaluations (DeVAL) and the information production (IP). The plots show the quantiles of the variables for the 70 observed banks. The continous line is the median value.

4.3.3 Rise of risky assets before the crisis

The median value for ratio of risky assets to total assets (RA.TA) is 0.91 for the whole period, but there are some extreme downward outliers. Prior to the crisis the risky asset ratio was rising with median percentage point changes higher 0. In the year 2009 there was a significant drop with a median change of -0.02 percentage points (see Table A.9). Apparently, this drop was caused by the massive loan write-downs in 2009 (loans account for an high proportion of RA). Another cause was that banks increased the amount of government securities in their balance sheets in 2009 (see Figure 4.6). Over the period the median ratio declined from 0.91 in 2003 to 0.87 in 2012 with a peak of 0.94 in 2006 (see Figure 4.5)
The interbank debt ratio LtB.TA steadily decreases from 19% in 2003 to 12% in 2010 and then jumps to 15% for 2011 and 2012. Figure 4.5 shows that the median growth rates of the interbank debt ratio are negative for all years except for 2005 and 2011.

4.3.4 Unsustainable large growth led to greater value adjustments

As we have seen in the previous section, the banking sector grew rapidly prior the crisis and had to make huge value adjustments in the years 2008 and 2009. The hypothesis, that those banks, that grew unsustainable fast in the years before the crisis, had to make larger value adjustments and write-offs than other banks, seems natural. To test this hypothesis we regressed the total loss in information production in the years 2008 and 2009 on the annualized growth of customer loans from 2003 to 2007. The result is highly significant (see Figure 4.7). This high growth of bad loans was fostered, by the underestimation of the risk of customer loans prior the crisis. The risk premium households and non-financial corporations had to pay for credits, symbolized by the spread of interest rates on loans and the fixed rate on an 3-month-euribor interest rate swap, was steadily decreasing prior the crisis (see Figure 4.8).

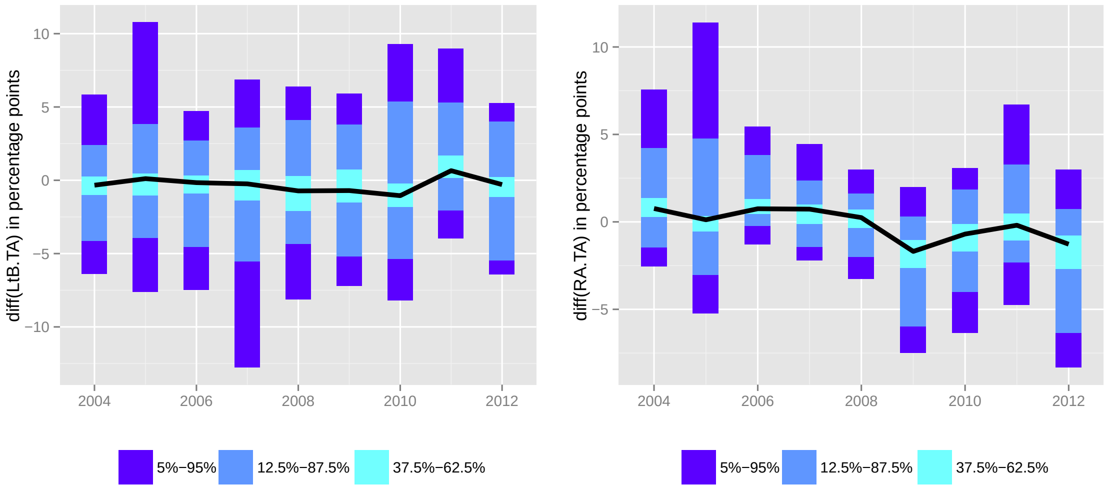

Figure 4.5: Yearly percentage point changes for the ratio of liabilities to banks to total assets (LtB.TA) and risky assets to total assets (RA.TA). The plots show the quantiles of the variables for the 70 observed banks. The continous line is the median value.

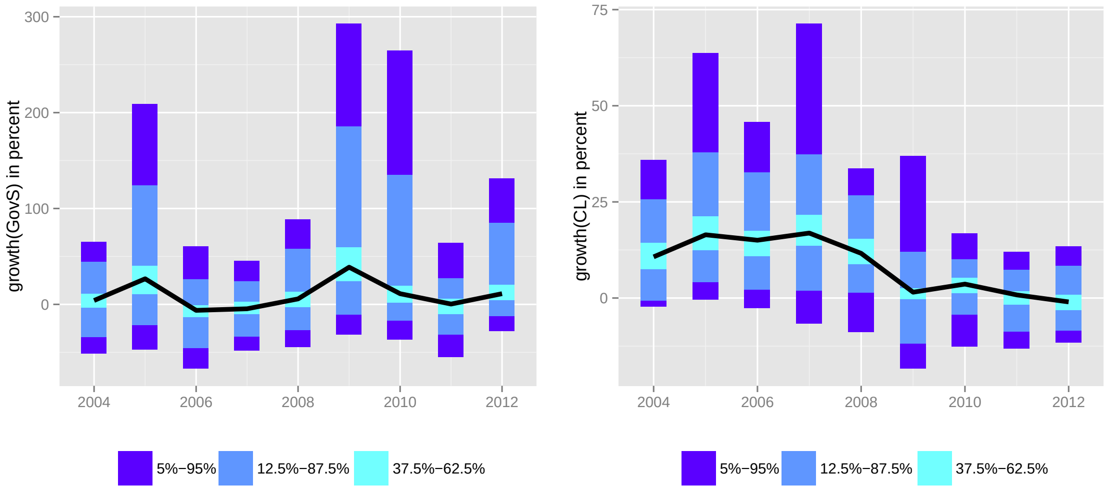

Figure 4.6: Yearly growth rates of GovS and CL for the years 2004 to 2012. The plots show the quantiles of the variables for the 70 observed banks. The continous line is the median value.

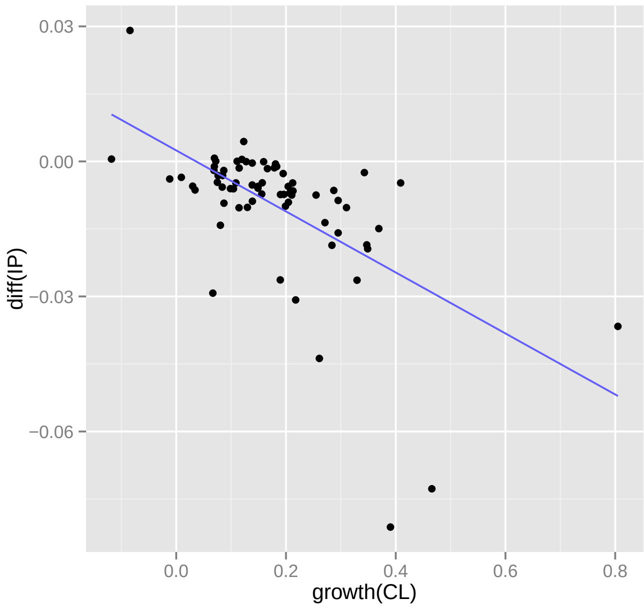

Figure 4.7: The difference of the information production of 2009 and 2007 regressed upon the annualized growth rate of customer loans from 2003 to 2007. The p-value for the coefficient is 7.6e-08 and the adjusted r-squared is 0.34.

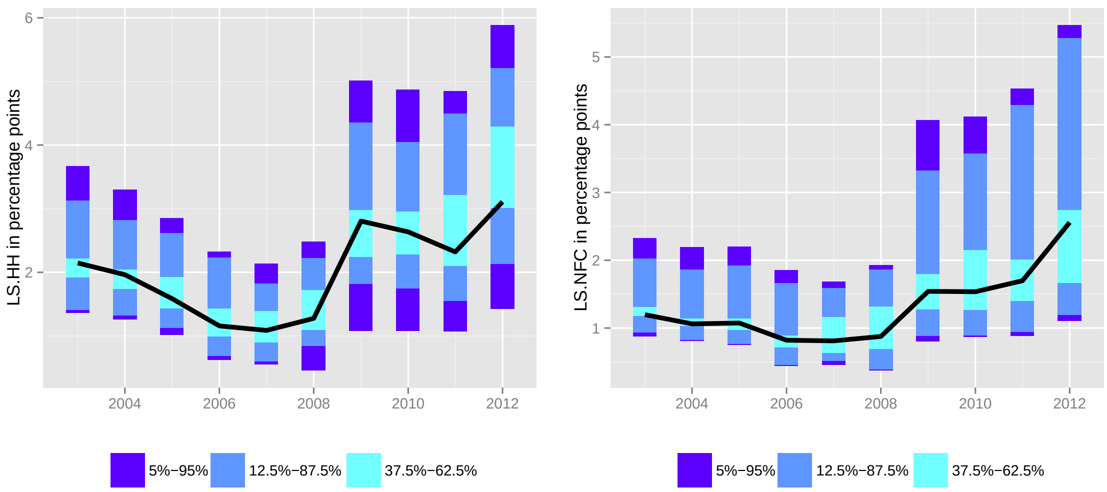

Figure 4.8: Development of the lending spreads for interest rates for households and non-financial corporations to euribor swaprates. The lending spread can either be seen as the profit banks can make by lending their money to households and non-financial customers or as the riskpremium customers have to pay. The swap rate is often used as a proxy for the risk free interest rate in financial mathematics. The plots show the quantiles of the variables for the 18 home countries of the observed banks. The continous line is the median value.

Chapter 5

Empirical Findings

This chapter includes the main part of this thesis: the results of the DEA and the Regressionanalysis. The chapter starts with a descriptive analysis of the DEA results and continues with a discussion on the sensibility of the DEA scores. Afterwards the results of the regression analysis will be discussed.

5.1 Result of DEA

The DEA scores were calculated using the statistical software R, together with the LPsolve-package. We were able to conduct DEA scores for 63 banks over the whole period, and for 5 more banks we had enough data to retrieve DEA scores for some years. Only for 2 banks we weren't able to calculate any efficiency measure, due to the absence of essential data. All in all we have 656 observation of DEA scores and 587 observations of frontier shift, Malmquist index and efficiency change (or catch-up) values. I will start the discussion of the DEA results by presenting the overall output of the efficiency analysis and I will then conclude with a Sensibility analysis of the DEA derived scores.

5.1.1 Results

I want to begin with discussing the SSBM (Super Slack Based Measure) scores. The distribution of the SSBM scores is bimodal and discontinous at 1, the border between efficient and non efficient (see Figure 5.1). But the distribution is very different over the years (see Figure 5.2). In the year 2003 the mean SSBM is 0.74, the median SSBM is 0.65 and there are only 18 banks with an SSBM score higher or equal 1. This changes completely in the years 2007 and 2008. E.g. in 2007 the mean SSBM is 0.86, the median SSBM is 1 and there are 35 banks with an SSBM score higher or equal 1 (see Table 5.1). This phenomen is probably due to the movement of the Frontiershift in this year.

There are eight banks that are efficient in all the years: ABLV Bank, Credit Agricole Group, HSBC Malta Group, ING Bank Group, La Banque Postale, Pohjola Bank, Tatra Bank. Therefore those banks are often used as reference for calculating the efficiency measure (see definition of peers in section 2.2). For example the Credit Agri-

	SSBM		MQ		FS		EC	
	mean	≥ 1(nobs)	mean	≥ 1(nobs)	mean	≥ 1(nobs)	mean	≥ 1(nobs)
2003	0.74	18(67)						
2004	0.80	21(67)	1.05	41(67)	0.95	22(67)	1.12	47(67)
2005	0.77	21(67)	1.03	35(66)	1.10	47(66)	0.95	27(66)
2006	0.79	23(67)	1.08	40(67)	1.03	35(67)	1.07	34(67)
2007	0.86	33(64)	1.04	35(64)	0.94	7(64)	1.13	43(64)
2008	0.84	31(64)	0.92	27(64)	0.96	21(64)	0.97	32(64)
2009	0.79	22(65)	1.12	42(64)	1.16	58(64)	1.01	23(64)
2010	0.77	23(65)	1.05	43(65)	1.06	51(65)	1.00	28(65)
2011	0.66	17(65)	0.91	17(65)	1.03	42(65)	0.89	17(65)
2012	0.69	21(65)	1.10	40(65)	1.07	55(65)	1.05	31(65)

Table 5.1: Result of DEA

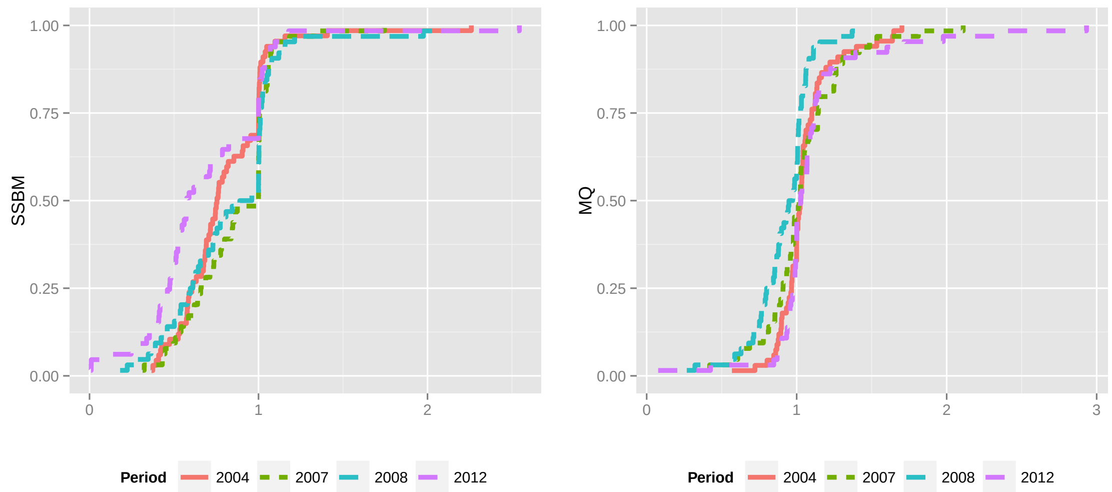

Figure 5.1: The cumulative distribution function for the SSBM scores and the Malmquist index for the years 2004, 2007, 2008 and 2012.

cole Group is the main peer bank for the giant banking groups Deutsche Bank AG, BNP Paribas Group and Banco Santander Group. With an average weight of 0.37, 0.78 and 0.42 over the whole period (see Figure 5.5). The weights of the peers can be interpreted as measures for the influence of the bank on the other banks efficiency measure. According to this measure the Tatra Bank, ABLV Bank and the HSBC Malta Group are the most pivotal banks over the 10 years, with weigth sums of 127, 75 and 60 (see Figure 5.5).
Also there are 20 banks that are inefficient in all the years, for example: RZB Konzern, Piraeus Bank Group and BBVA Group. I tried to split the banks into different efficiency regimes using a kmeans clustering algorithm, but the resulting groups overlapped to often.

The distribution of the Malmquist index is different compared to that of the SSBM scores, in that it is unimodal and there is no discontinuity point at 1. However the

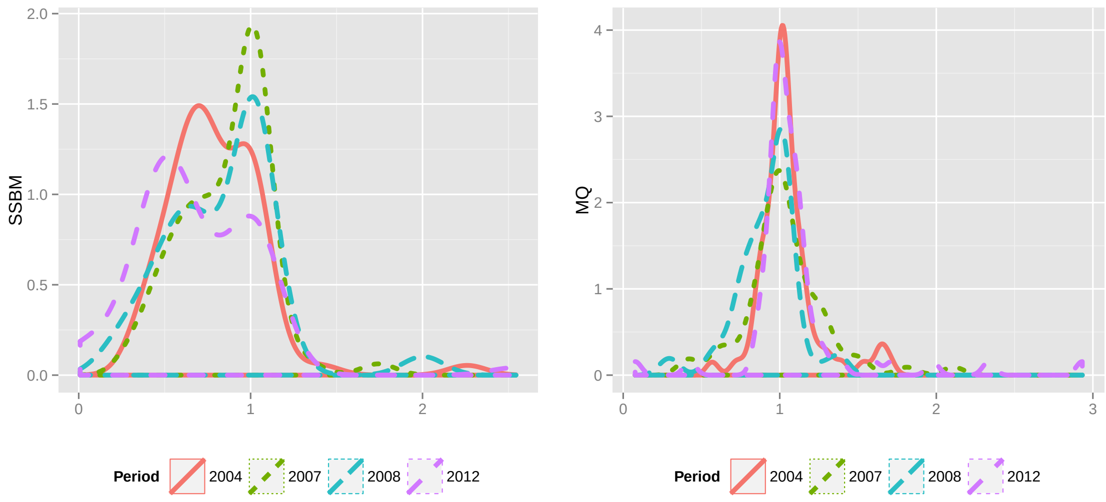

Figure 5.2: The density function for the SSBM scores and the Malmquist index for the years 2004, 2007, 2008 and 2012.

Malmquist index doesn't show any general movement. A two-sided exact binomial test rejects the null hypothesis of a median Malmquist index not equal to one, in all the years, except the year 2011 (see Table A.9). Still in the year 2008 the average Malmquist index is 0.92 and 37 out of 64 banks got less efficient (see Figure 5.3). The low values of the Malmquist index in the year 2008 are a result of a drop of the stability index and the information production in this year (see Figure 4.3). This drop was caused by the already mentioned disturbances on the banking sector starting in the end of 2007. The significant low values in the year 2011 are harder to explain.

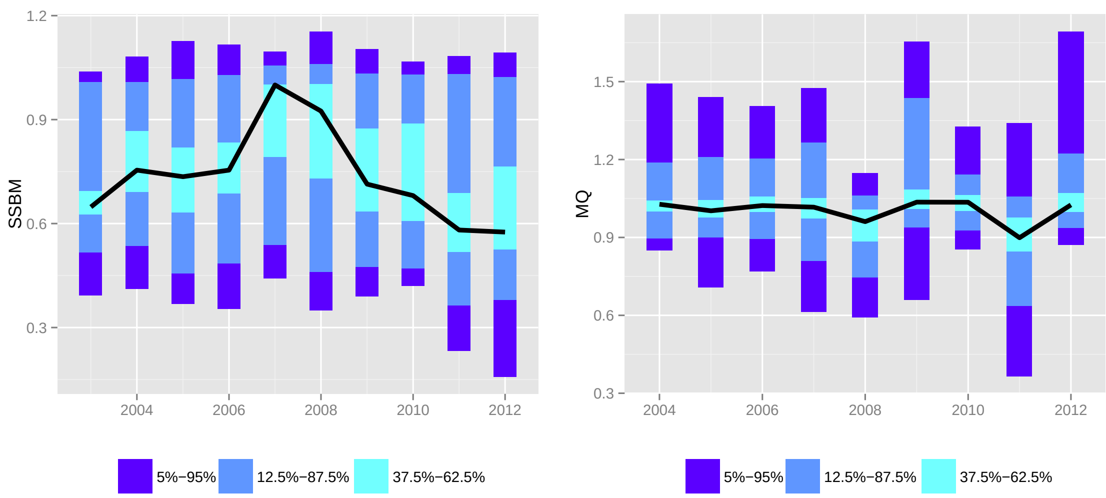

Figure 5.3: Development of the Super Slack Based Measure and the Malmquist Index. The plots show the quantiles of the variables for the observed banks, for which we could calculate the efficiency scores. The continous line is the median value.

In contrast to the Malmquist Index the values for the Frontier Shift show more movements. In the years 2005, 2009, 2010 and 2012 the median value for the Frontier Shift is significantly higher than 1, indicating an extension in the production possibilities of the industry in servicing the real side of the economy (see Table A.9). In the years 2004, 2007 and 2008 the Frontier Shift is significantly lower than 1 (see Figure 5.4). The movement of the Frontier Shift in 2007 and 2008 explains partly why the there are so many efficient banks in the years 2007 and 2008. This drop in production possibility enabled the inefficient banks to catch up to the efficient banks stretching the efficient frontier.

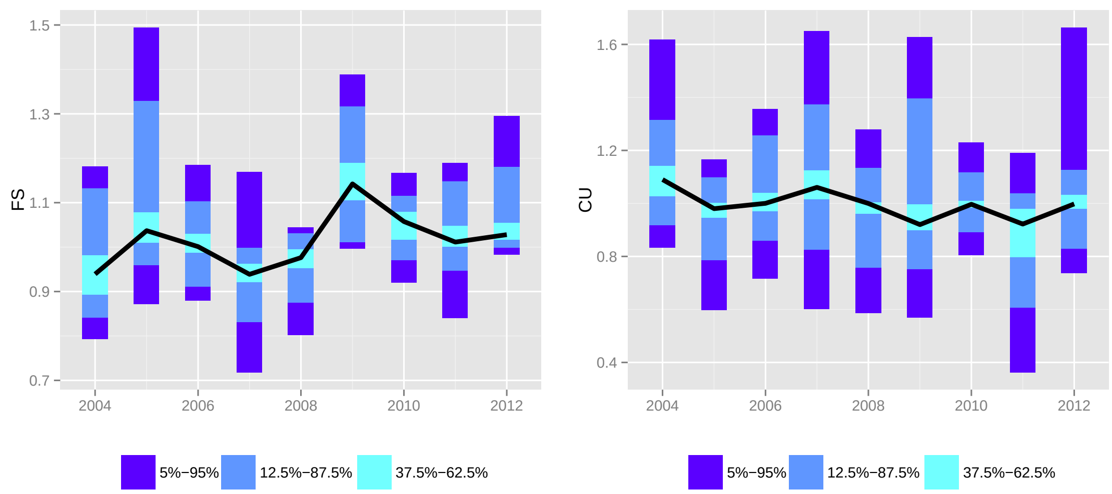

Figure 5.4: Development of the Catch-Up and the Frontier Shift. The plots show the quantiles of the variables for the observed banks, for which we could calculate the efficiency scores. The continous line is the median value.

5.1.2 Sensibility of DEA scores

Due to the fact, that the values calculated during the DEA are per definition interdependent, peculiar econometric methods will be needed for the Regressionanalysis. In this section I am going to analyze the dependency of the DEA values from the input factors. First I am going to observe how sensitive the DEA scores are to an alteration of the DEA factors. In detail I modified the DEA factors (TA, LP, SI, IP, RA, LtB.TA) for one bank and then I recalculated the DEA factors. For example:

1. I calculate the efficiency score for the Unicredit for the Period 2003 and denote it with $SSBM^{2003}_{Unicredit}$
2. I raise the liquidity production for the Unicredit by 15%, but all the other factors for the unicredit stay the same. The DEA factors of all the other banks don't change either.
3. I calculate the efficiency score for the Unicredit using the altered factor and denote it with $SS\hat{B}M^{2003}_{Unicredit}$

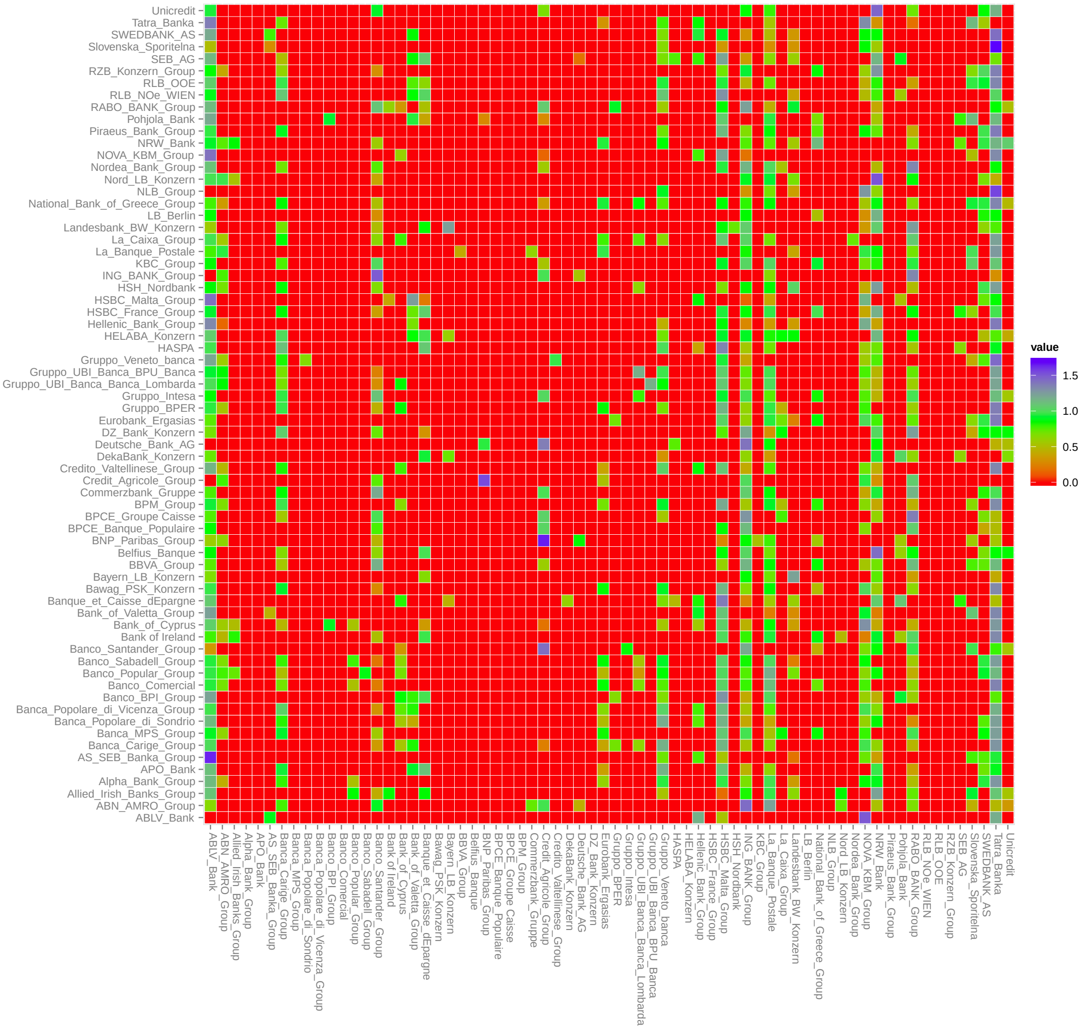

Figure 5.5: Sum of the peer matrices of all ten years. We calculated the peer weigths for every bank in each year (see section Slack Based Model and Super Slack Based Model). Each line shows for each bank what their peers are. The rowsum is 1 for every bank in every year. The columns show for which banks the bank is a peer. This means a bank with many entries in their column is a reference for a lot of banks. For graphical purpose I transformed the entries using the fourth root.

4. Finally I calculate the difference between the two efficiency scores $\frac{S\hat{SBM}^{2003}_{Unicredit} - SSBM^{2003}_{Unicredit}}{SSBM^{2003}_{Unicredit}}$

I did this simulation for every bank and every year. The results of this analysis are quiet interesting. While a reduction of an input factor can have a big influence on the DEA result (a reduction of the TA by 15% leads to an increase in efficiency of 86%), raising an input factor has rather small influences on the DEA result (raising TA by 15% leads to an decrease in efficiency of 8%). The analogue result is true for output factors. (see fig. 5.6). The effect is higher for those factors with a lower standard deviation. This means a bank that can improve in this area can "run away" from the sample. In Figure 5.7 you can see that the observations for the IP are located very densely around the mean, therefore an increase in IP can have huge effects on the SSBM score. This result

also reveals that the DEA scores can be very sensible to data errors.
Further I studied what impact the occurance of a bank in the sample has on the SSBM values of the other banks. For that reason I always excluded one Pivot-bank from the sample and then recalculated the SSBM values and compared them to the SSBM values, where the Pivot-banks are included in the sample. For example:

1. I calculate the efficiency scores for all the banks for the Period 2003 and denote them with $SSBM^{2003}_{.}$
2. Now I exclude the Tatra Bank from the sample
3. I recalculate the efficiency scores for all the banks for the Period 2003 and denote them with $\tilde{SSBM}^{2003}_{.}$
4. Finally, I can calculate the effect the Tatra Bank has on the efficiency scores of the other banks, by calculating $\frac{\tilde{SSBM}^{2003}_{.} - SSBM^{2003}_{.}}{SSBM^{2003}_{.}}$. The mean over all the changes can be interpreted as the overall effect the Tatra Bank has on the sample in the year 2003.

The result of this investigation is, that there are some banks that have massive influences on the DEA scores. For example excluding the Tatra Bank from the sample will raise the average SSBM score in the year 2011 by 16.6% and the median change is 14%. I did this for all the years and for all the banks and calculated the overall effect (measured by the mean change) the Pivot bank has on the sample for every year. Afterwards I regressed the mean change onto the sum of the peerweigths of this bank (the columnsum of the bank, see Figure 5.5). The coefficient of this regression is highly significant and the adjusted R-squared is 0.66 (see fig 5.8).

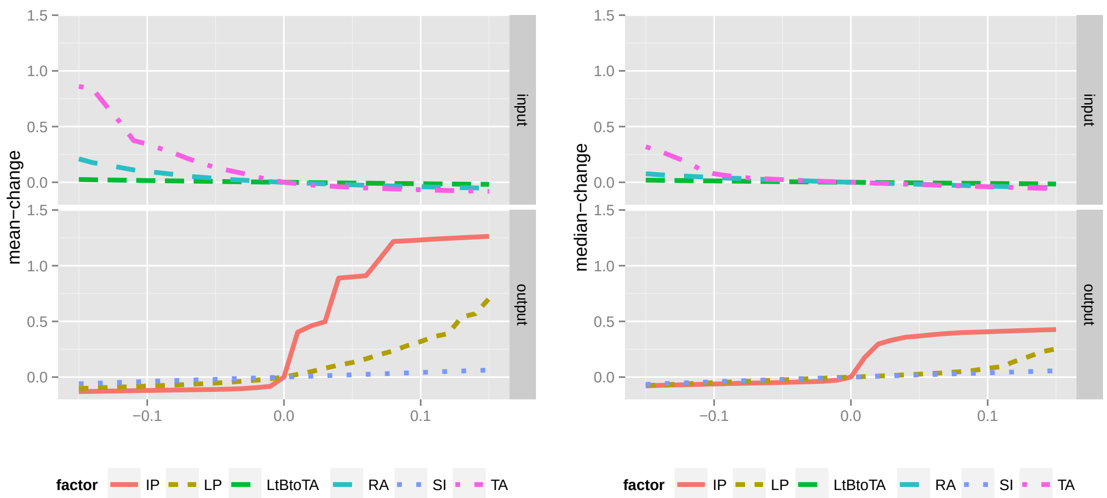

Figure 5.6: Sensitivity of the DEA values to an alteration of the DEA factors. The plots show how the SSBM scores change, when the DEA factors are altered.

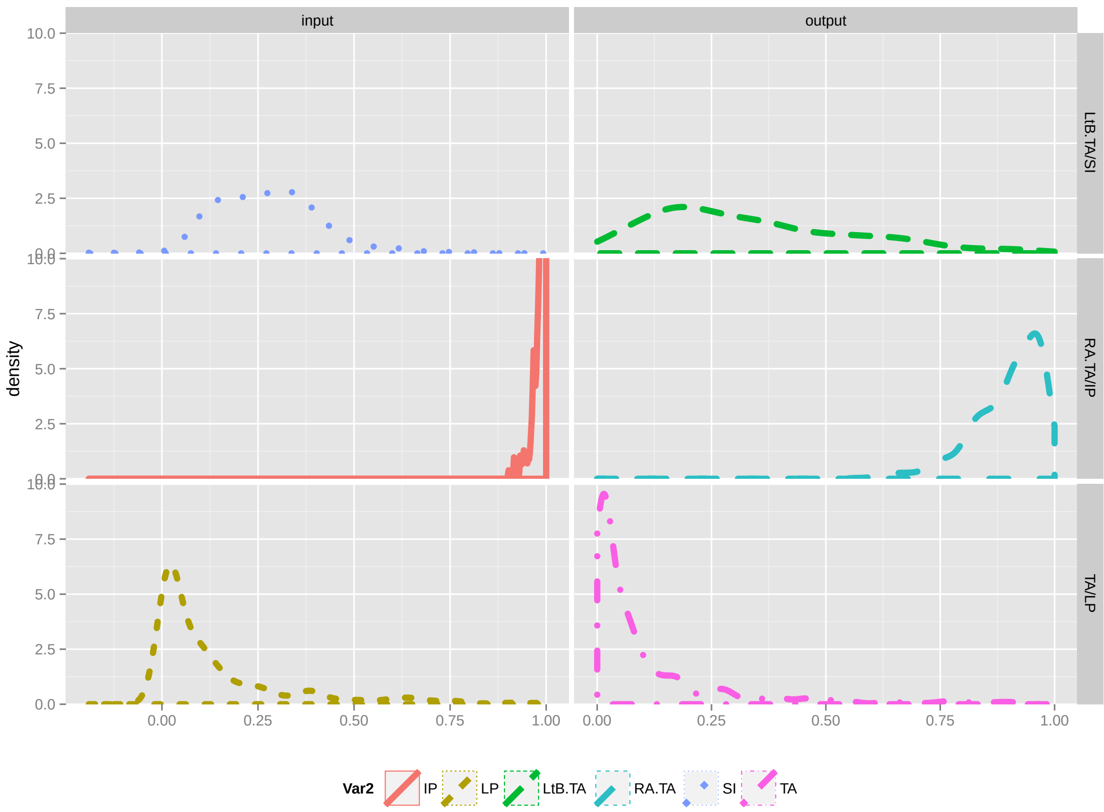

Figure 5.7: Density distribution of the normalized DEA factors. The DEA factors are normalized to 1 by dividing through the maximum.

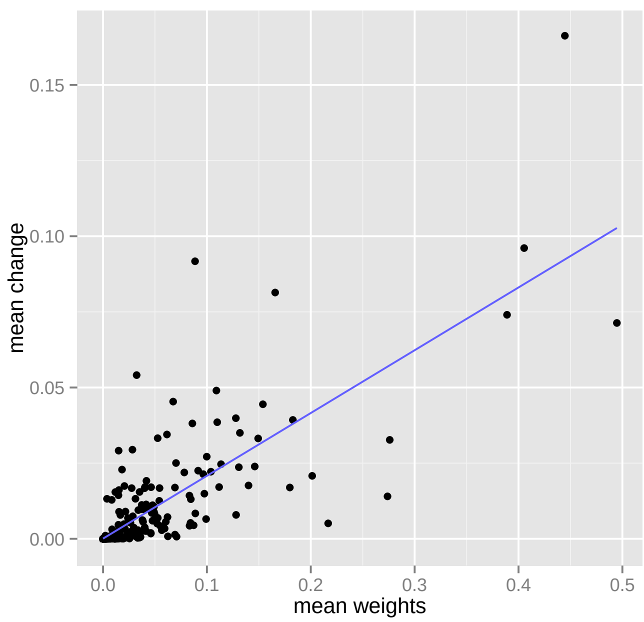

Figure 5.8: The average change on SSBM, when the Pivot bank is excluded, regressed on the mean peerweight of the Pivot bank. The results of the regression are very significant: the estimate is 0.21 with a p-value < 2e-16 and an adjusted r-squared of 0.66.

5.2 Result of Regressionanalysis

Aim of this section is to find the main drivers of the Malmquist index. We used regression analysis to find models which try to explain the development of banking efficiency. We used a great variety of different bank specific and macroeconomic indicators to find a suitable model. In total we considered about 25 candidate regressors like GDP, unemployment rate (UU), inflation rate ($HICP$), public debt ($TDebt$, $FDebt$), an Herfindahl index ($Herfindahl$) for the banking sector and several other indicators for the financial and monetary markets. For a list of considered variables see Table A.2.

We tried a lot of different models, but the problem was that the significance of most of the regressor variables was very fragile for the considered models. Therefore we decided to use an extreme bound analysis together with a robust variance matrix estimator, as described in chapter 3. For the analysis we started using a two-ways fixed effects model. The basic model, resulting from this analysis is given in Table 5.2. This model explains the Malmquist index by the efficiency of the bank in the previous period ($lag(SSBM)$), the first differences of the cost-income ratio of the bank ($diff(CIR)$), the lending spreads for household credits ($LS.HH$) and credits for non-financial corporations ($LS.NFC$) and the squared differences of the government interest rates ($diff(GovInt)^2$). These 5 variables showed up to be stable according to the extrem bound analysis. The estimates and the results of the extreme bound analysis are given in table 5.2.

In compliance with the signs of these five variables we may derive the following relationsships:

- less efficient banks tend to improve their relative efficiency while more efficient banks show a tendency to lose
- banks that improve their operational efficiency in terms of the cost-income ratio, which is often a key performance indicator in banking management, also tend to improve their efficiency according to the Malmquist index
- banks from countries, that are exposed to large changes in the interest rate their governments have to pay, tend to suffer a loss in their relative efficiency
- The lending spreads have a significant influence on the Malmquist index. The systemic efficiency increases with the spread for household credit and decreases with the spread for loans to non-financial corporations. This issue will be discussed in detail in the next sections.

We were surprised that we couldn't prove a relationship between the change in systemic efficiency and the economic development. Variables like GDP growth rates ($growth(GDP)$), changes in unemployment and inflation ($diff(UU)$, $diff(HICP)$), appeared to be insignificant or fragile. As can be seen in table 5.3, the GDP growth rate is in 26% of the testregressions insignificant and the changes in the unemployment rate is insignificant in 90% of the testregressions.

So far we have not added variables that reflect characteristics of the national banking markets. The results in Table 5.4 imply that the variables, $Top5$, a measure for the

	Estimate	Std. Error	t-value	Pr(>|t|)	lower	upper	#
$lag(SSBM)$	-0.81	0.10	-7.71	0.00	-1.05	-0.54	0
$LS.HH$	0.09	0.01	6.43	0.00	-0.00	0.16	1
$LS.NFC$	-0.14	0.04	-3.56	0.00	-0.25	0.00	1
$diff(GovInt)^2$	-0.01	0.00	-8.38	0.00	-0.02	-0.00	0
$diff(CIR)$	-0.02	0.00	-5.17	0.00	-0.03	-0.01	0

Table 5.2: Results of the basic regression model. The columns contain the estimate, the standard error, the t-statistic and the p-value of the coefficient. Further the table includes the minimum of the lower boundary and the maximum of the upper boundary of the 0.95-confidence interval of all the testregressions. The last column is the amount of testregressions in which the coefficient wasn't significant. For the testregression combinations of 4 variables out of 15 (see Table A.10) were used, thus resulting in 1365 testregressions. The minimal adj. r-squared is 0.32 and the maximal adj. r-squared is 0.36

	Estimate	Std. Error	t-value	Pr(>|t|)	lower	upper	#
$growth(GDP)$	1.30	0.77	1.69	0.09	-3.45	5.59	431
$diff(HICP)$	0.01	0.01	0.82	0.41	-0.05	0.08	792
$diff(UU)$	0.01	0.02	0.58	0.56	-0.07	0.08	905
$diff(PriDef)$	-0.01	0.00	-2.11	0.04	-0.03	0.02	658
$diff(TDebt)$	-0.01	0.00	-2.72	0.01	-0.03	0.01	561
$diff(FDebt)$	0.01	0.01	1.67	0.10	-0.02	0.02	961

Table 5.3: Results of regressing the Malmquist Index on Macroeconomical variables. The columns contain the estimate, the standard error, the t-statistic and the p-value of the coefficient. Further the table includes the minimum of the lower boundary and the maximum of the upper boundary of the 0.95-confidence interval of all the testregressions. The last column is the amount of testregressions in which the coefficient wasn't significant. For the testregression combinations of 3 variables out of 14 (see Table A.10) were used, thus resulting in 1001 testregressions. The minimal adj. r-squared is 0.16 and the maximal adj. r-squared is 0.35

concentration of the banking sector, $MCap$ and $stocks\,trade$, two measure for the size of the stock markets, don't have stable significant impact on the banking efficiency.

5.2.1 Adding product terms

So far we have only analyzed linear models and surprisingly neither macroeconomical variables nor market characteristical variables have a significant impact on the development of banking efficiency. Hence we added product terms of market characteristics to our basic regression model to see if the magnitude of the influence of our stable variables is dependent on the market structure.

At first we added productterms with the variable $Top5$. The results given in Table 5.5 are quiet insightful. All the product terms are significant, except the productterm

	Estimate	Std. Error	t value	Pr(>|t|)	lower	upper	#
$Top5$	-0.79	0.23	-3.42	0.00	-1.71	0.61	726
$MCap$	0.00	0.00	2.68	0.01	-0.00	0.00	242
$stocks\ trade$	-0.00	0.00	-0.94	0.35	-0.00	0.00	1818

Table 5.4: Results of regressing the Malmquist Index on variables that reflect the marktcharacteristics on the national banking sectors. The columns contain the estimate, the standard error, the t-statistic and the p-value of the coefficient. Further the table includes the minimum of the lower boundary and the maximum of the upper boundary of the 0.95-confidence interval of all the testregressions. The last column is the amount of testregressions in which the coefficient wasn't significant. For the testregression combinations of 3 variables out of 16 (see Table A.10) were used, thus resulting in 1820 testregressions. The minimal adj. r-squared is 0.12 and the maximal adj. r-squared is 0.34

for the lagged efficiency value ($lag(SSBM)$), and have the opposite sign as the corresponding linear coefficient. This means that a high proportion of the biggest 5 banks in the market weakens the effects caused by the five basic variables. Hence the efficiency of banks from a country with a highly competetive banking market react stronger to changes in the lending spreads ($LS.HH$, $LS.NFC$) or changes in the operative efficiency ($diff(CIR)$). In contrast to Table 5.2 the sign changed for the square of the first difference in government interest rates ($diff(GovInt)^2$).

	Estimate	Std. Error	t-value	Pr(>|t|)	lower	upper	#
$lag(SSBM)$	-1.02	0.20	-5.05	0.00	-1.61	-0.56	0
$LS.HH$	0.36	0.08	4.32	0.00	0.11	0.61	0
$LS.NFC$	-0.81	0.15	-5.28	0.00	-1.20	-0.37	0
$diff(GovInt)^2$	0.12	0.04	3.06	0.00	-0.00	0.25	1
$diff(CIR)$	-0.06	0.02	-3.08	0.00	-0.11	-0.02	0
$lag(SSBM) : Top5$	0.49	0.30	1.66	0.10	-0.25	1.49	547
$LS.HH : Top5$	-0.51	0.13	-3.91	0.00	-0.85	-0.12	0
$LS.NFC : Top5$	1.20	0.25	4.88	0.00	0.56	1.85	0
$diff(GovInt)^2 : Top5$	-0.19	0.06	-3.50	0.00	-0.37	-0.03	0
$diff(CIR) : Top5$	0.06	0.03	2.26	0.02	-0.00	0.13	6

Table 5.5: Results of regressing the Malmquist Index on Macroeconomical variables. The columns contain the estimate, the standard error, the t-statistic and the p-value of the coefficient. Further the table includes the minimum of the lower boundary and the maximum of the upper boundary of the 0.95-confidence interval of all the testregressions. The last column is the amount of testregressions in which the coefficient wasn't significant. For the testregression combinations of 4 variables out of 15 (see Table A.10) were used, thus resulting in 1365 testregressions. The minimal adj. r-squared is 0.36 and the maximal adj. r-squared is 0.39

Next we added productterms with the variable marketcapitalization ($MCap$), to see if banks from countries with greater stock markets react differently to changes of our basic variables, than banks from countries with rather small stock markets. The results in Table 5.6 imply that the size of the stock market has no significant influence on the reaction of the Malmquist index to changes in the basic variables, except for the squared first difference of the government interest rates ($diff(GovInt)^2$). The coefficient for this product term is slightly positive. In general systemic banking efficiency suffers, if there are changes in the level of the interest rate which governments have to pay. But the positive sign for the coefficient of the product term suggest, that this effect is weakened for banks from countries with a high affinity to stock markets. A possible reason for this could be, that in countries with larger stock markets, banks are more independent from governments and thus are not so affected by changes in the governments ability to refinance. E.g.: in the year 2011 the interest rates greek had to pay for new debts went up to 15.7% from 9.1% and the government of Portugal dealt with an increase from 5.4% to 10.2%. In this year Greek banks had an average Malmquist index of 0.21 and Portuguese banks had an average Malmquist index of 0.718. At this time Portuguese listed enterprises had a market capitalization of 26% of the GDP while Greek enterprises only had a market capitalization of 12% of the GDP.

	Estimate	Std. Error	t value	Pr(>|t|)	lower	upper	#
$lag(SSBM)$	-0.72	0.11	-6.30	0.00	-1.04	-0.37	0
$LS.HH$	0.09	0.02	4.49	0.00	-0.03	0.20	24
$LS.NFC$	-0.13	0.03	-4.61	0.00	-0.25	-0.01	0
$diff(GovInt)^2$	-0.03	0.00	-9.85	0.00	-0.04	-0.01	0
$diff(CIR)$	-0.04	0.03	-1.57	0.12	-0.11	0.06	1365
$lag(SSBM) : MCap$	-0.00	0.00	-1.38	0.17	-0.00	0.00	1152
$LS.HH : MCap$	-0.00	0.00	-0.75	0.46	-0.00	0.00	1307
$LS.NFC : MCap$	0.00	0.00	0.45	0.65	-0.00	0.00	1339
$diff(GovInt)^2 : MCap$	0.00	0.00	3.73	0.00	-0.00	0.00	7
$diff(CIR) : MCap$	0.00	0.00	0.95	0.34	-0.00	0.00	1365

Table 5.6: Results of regressing the Malmquist Index on Macroeconomical variables. The columns contain the estimate, the standard error, the t-statistic and the p-value of the coefficient. Further the table includes the minimum of the lower boundary and the maximum of the upper boundary of the 0.95-confidence interval of all the testregressions. The last column is the amount of testregressions in which the coefficient wasn't significant. For the testregression combinations of 4 variables out of 15 (see Table A.10) were used, thus resulting in 1365 testregressions. The minimal adj. r-squared is 0.34 and the maximal adj. r-squared is 0.36

5.2.2 Differences in bank size and time of crisis

Now we split the banks according to their size, measured by their TA. For the classification we used the k-means algorithm on the natural logarithm of TA. We divided the banks into four groups: small(8 banks), medium(23 banks), big(23 banks), large(16 banks) (see Figure 5.9). As can be seen in Figure 5.9 the four cluster separates the banks excatly. There are only a few cases in the ten-year period in which two clusters overlap. This also means that in our sample the relative size of the banks stayed more or less the same.

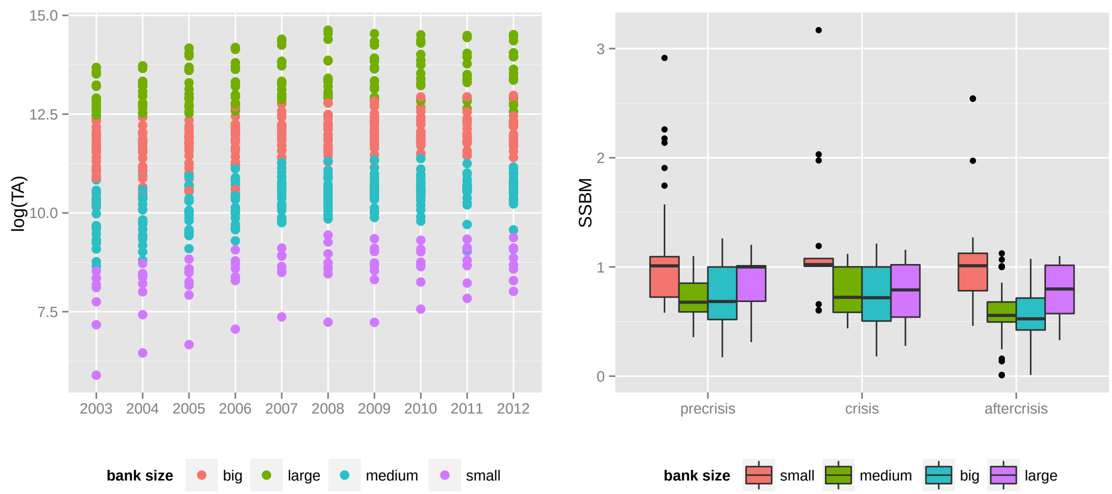

Figure 5.9: The left picture shows the banking cluster for the ten years. We used the k-means algorithm for the clustering of the banks on the vectors of log(TA) for the Period from 2005 to 2012. The right picture shows the distribution of the efficiency scores for the banking groups and the time of the crisis.

Further we checked if the relationship between the Malmquist index and the covariates are "stable" in the sense that, there are no significant differences in the estimates for the different bank clusters. We tested if the estimates for the different bank sizes are equal with a Wald-test, using the robust covariance matrix. The results are given in Table 5.7.

- Although the Wald test accepts the null, the coefficients for the lagged efficiency scores get smaller for larger banks. This implies that the effect that less efficient banks tend to have a higher efficiency growth is heavier if the bank is larger measured by their TA.
- The Wald test rejects the null for the estimates of both lending spreads. Again it seems as the effect on the Malmquist index caused by the extent of the lending spreads are more intensive if the banks are larger.
- The coefficients for the squared first differences of the government interest rates are significantly different regarding the Wald test. Interestingly the estimates are negative except for the cluster of the large banks. Furthermore the coefficients for the big and large banks are not significant with reference to the t-test. A possible explanation could be that bigger banks are more independent from governments and are therefore not prone to changes in government interest rates.

- The estimates for the first differences of the cost-income ratio are smaller for smaller banks, but they are also insignificant for small and medium banks. This implies that an operative efficiency growth leads to an systemic efficiency growth only for big banks.

	Estimate	Std. Error	t value	Pr(>|t|)	Wald
$lag(SSBM) : small$	-0.48	0.05	-8.73	0.00	0.15
$lag(SSBM) : medium$	-0.65	0.19	-3.41	0.00	
$lag(SSBM) : big$	-0.94	0.19	-4.88	0.00	
$lag(SSBM) : large$	-1.01	0.28	-3.62	0.00	
$LS.HH : small$	0.10	0.02	5.07	0.00	0.00
$LS.HH : medium$	0.05	0.04	1.16	0.25	
$LS.HH : big$	0.11	0.05	2.22	0.03	
$LS.HH : large$	0.18	0.03	6.57	0.00	
$LS.NFC : small$	-0.12	0.02	-5.00	0.00	0.01
$LS.NFC : medium$	-0.04	0.05	-0.80	0.43	
$LS.NFC : big$	-0.27	0.07	-3.69	0.00	
$LS.NFC : large$	-0.31	0.07	-4.21	0.00	
$diff(GovInt)^2 : small$	-0.01	0.03	-0.49	0.63	0.00
$diff(GovInt)^2 : medium$	-0.02	0.00	-16.23	0.00	
$diff(GovInt)^2 : big$	-0.00	0.00	-0.64	0.52	
$diff(GovInt)^2 : large$	0.05	0.04	1.26	0.21	
$diff(CIR) : small$	-0.21	0.23	-0.90	0.37	0.00
$diff(CIR) : medium$	-0.25	0.13	-1.95	0.05	
$diff(CIR) : big$	-0.02	0.00	-9.26	0.00	
$diff(CIR) : large$	-0.04	0.01	-3.08	0.00	

Table 5.7: Results of regressing the Malmquist index on our five basis variables with different coefficients for the different banking groups. In the last column is the p-value of the Wald test statistic.

I the next step we split the time period into three regimes: precrisis (2003-2007), crisis(2008-2009) and aftercrisis(2010-2012). We split the period into these three regimes because in 2008 and 2009 the crisis had the biggest impact on the DEA factors. This can be seen best in the movement of the information production and the stability index(see Figure 4.3). The stability index declined enormously in the years 2008 and 2009 and stabilized afterwards except for a few banks were the information production further declined.

Again we checked if there are differences in the estimates for the different regimes. The results are given in Table 5.8.

- The influence of the lending spreads and the lagged efficiency score are more or less the same for the different phases of the crisis.

- There is a significant difference in the estimates for the squared differences of the government interest rates. Especially during the crisis in the years 2008 and

2009 the impact on the Malmquist index was extremely high. This means that in the years 2008 and 2009 the systemic banking efficiency declined in particular in those countries where the yields on government debts altered more.

- In the years 2008 and 2009 there has been a higher correlation between operative efficiency and systemic efficiency.

	Estimate	Std. Error	t value	Pr(>|t|)	Wald
$lag(SSBM) : precrisis$	-0.82	0.10	-7.79	0.00	0.06
$lag(SSBM) : crisis$	-0.66	0.06	-11.38	0.00	
$lag(SSBM) : aftercrisis$	-0.90	0.16	-5.79	0.00	
$LS.HH : precrisis$	0.08	0.05	1.59	0.11	0.71
$LS.HH : crisis$	0.11	0.06	2.03	0.04	
$LS.HH : aftercrisis$	0.08	0.02	4.03	0.00	
$LS.NFC : precrisis$	-0.30	0.13	-2.37	0.02	0.11
$LS.NFC : crisis$	-0.23	0.06	-3.55	0.00	
$LS.NFC : aftercrisis$	-0.15	0.04	-3.80	0.00	
$diff(GovInt)^2 : precrisis$	-0.07	0.02	-2.85	0.00	0.00
$diff(GovInt)^2 : crisis$	-0.46	0.17	-2.73	0.01	
$diff(GovInt)^2 : aftercrisis$	-0.01	0.00	-8.23	0.00	
$diff(CIR) : precrisis$	-0.16	0.13	-1.21	0.23	0.00
$diff(CIR) : crisis$	-0.07	0.01	-10.72	0.00	
$diff(CIR) : aftercrisis$	-0.02	0.00	-14.56	0.00	

Table 5.8: Results of regressing the Malmquist index on our five basis variables with different coefficients for the different phases of the crisis. In the last column is the p-value of the Wald test statistic.

Finally we did the same procedure for a dummyvariable, that indicates whether the country of the corresponding bank is currently exposed to an ongoing systemic banking crisis or not. The start point of an systemic banking crisis is a year with significant signs of financial distress in the banking system and significant policy interventions (see [13]). For the start and end dates of the banking crises we used the results from [13] and extended it accordingly until the year 2012. The results of the regression are given in Table 5.9.

5.2.3 A model without time effects

So far we haven't checked what influence global economic indicators have on the change of efficiency. To test whether variables like the main refinancing rate of the ECB ($ECBrate$), the growth rate of the monetary aggregates ($M1, M2, M3$) or european stock indices ($StoxxF$, $StoxxB$, $Stoxx50$), bias efficiency growth or not we will omit the fixed time effects and consider one-way fixed effects models with individual effects. Basically we take the same five variables that turned out to be stable and add global time dependent variables. The results are (see Table 5.10):

	Estimate	Std. Error	t value	Pr(>|t|)	Wald
$lag(SSBM) : noCrisis$	-0.79	0.09	-8.94	0.00	0.80
$lag(SSBM) : Crisis$	-0.82	0.14	-6.06	0.00	
$LS.HH : noCrisis$	0.09	0.03	2.66	0.01	0.68
$LS.HH : Crisis$	0.10	0.02	6.41	0.00	
$LS.NFC : noCrisis$	-0.15	0.08	-1.75	0.08	0.96
$LS.NFC : Crisis$	-0.15	0.03	-4.53	0.00	
$diff(GovInt)^2 : noCrisis$	-0.06	0.02	-3.41	0.00	0.01
$diff(GovInt)^2 : Crisis$	-0.01	0.00	-10.52	0.00	
$diff(CIR) : noCrisis$	-0.07	0.10	-0.65	0.51	0.66
$diff(CIR) : Crisis$	-0.02	0.00	-5.08	0.00	

Table 5.9: Results of regressing the Malmquist index on our five basis variables with different coefficients for banks where the country has a banking crisis. In the last column is the p-value of the Wald test statistic.

- the increase in systemic banking efficiency tends to be higher wenn the prices of stocks for financial institutions are higher
- surprisingly the ECB main refinancing rate and its first difference are positively correlated with the Malmquist index. We would have suspected that there is a negative correlation, having in mind that a low interest rate makes it easier for banks to produce liquidity for the markets
- efficiency growth is negatively correlated with the growth rate of the monetary aggregate M3.

	Estimate	Std. Error	t value	Pr(>|t|)	lower	upper	ratio
$lag(SSBM)$	-0.80	0.11	-7.43	0.00	-1.04	-0.54	0
$LS.HH$	0.09	0.02	5.70	0.00	0.00	0.16	0
$LS.NFC$	-0.13	0.04	-3.68	0.00	-0.24	0.00	1
$diff(GovInt)^2$	-0.01	0.00	-9.20	0.00	-0.02	-0.00	0
$diff(CIR)$	-0.02	0.00	-5.26	0.00	-0.03	-0.01	0
$growth(StoxxF)$	0.31	0.03	11.39	0.00	0.08	0.75	0
$growth(M3)$	-3.12	0.43	-7.25	0.00	-8.57	0.53	19
$ECBrate$	0.11	0.01	7.43	0.00	-0.00	0.30	1
$diff(ECBrate)$	0.05	0.01	7.09	0.00	-0.07	0.14	305

Table 5.10: Results of regressing the Malmquist Index on basis variables and time dependent global variables. The columns contain the estimate, the standard error, the t-statistic and the p-value of the coefficient. Further the table includes the minimum of the lower boundary and the maximum of the upper boundary of the 0.95-confidence interval of all the testregressions. The last column is the amount of testregressions in which the coefficient wasn't significant. For the testregression combinations of 4 variables out of 15 (see Table A.10) were used, thus resulting in 1365 testregressions. The minimal adj. r-squared is 0.32 and the maximal adj. r-squared is 0.36

We further conducted a Wald test, whether these four time dependent variables are sufficient or if fixed time effects are still necessary. The Wald test compares the above model with the twoways fixed effects model and clearly rejects the null hypothesis (Wald test statistics $W = 261.8, df = 4, Pr(> Chisq) < 2.2e - 16$) and therefore we continue with the twoways fixed effects model.

5.2.4 Non-linear specifications

As we have already seen in the previous subsection, there are some non-linear dependencies between the Malmquist index and the regressor variables. Therefore we performed a RESET test, where we regress the Malmquist index onto the fitted values from our basis model, and onto powers of the fitted values (see [2]). The RESET test checks whether the coefficients of the powered terms are significant or not. For our basis model the RESET test clearly rejects the null (RESET = 8.4002, df1 = 5, df2 = 446, p-value = 1.315e-07) and therefore we assume that there are some undetected non-linear dependencies.
Looking at the outcome of our regression model we have to deal with the problem that the coefficients for the two different lending spreads have distinct signs. This effect is rather hard to explain. Therefore we were looking for transformations of these two variables where the explanation is more suitable to the real-world events. We used two simple transformations for the lending spreads:

- $LS.D = LS.HH – LS.NFC$: this represents the risk premium that households have to pay more for their loans than enterprises. This difference is mostly positive, which means that normally firms are more creditworthy than households.
- $LS.S = LS.HH + LS.NFC$: the sum of the two lending spreads can be interpreted as the overall credit risk in an economy.

If we replace the variables $LS.HH$ and $LS.NFC$ in the regression with these two transformations and if we add the product of these two transformation, we get the following results: The difference of the lending spreads is highly significant and robust with a positive sign. The product of the two transformed variables is highly significant and robust with a negative sign. The sum of the lending spreads is not significant (see Table 5.11). This means that banks in countries, where the market conditions had a bias towards lending money to corporations rather than households, performed better. In contrast to those banks, that operated in markets, where households were evaluated as creditworthy as enterprises or even better. In Figure 5.10 we can see that prior to the crisis the median risk premium households had to pay for loans was sinking towards the risk premium enterprises hat to pay, reaching its minimum in 2007. This highly coincides with the triggers of the subprime crisis, were more and more cheap credits were granted to “normally” credit-un-worthy households. So banks in countries with too optimistic evaluated household creditworthiness had to make higher value adjustments due to too many bad loans. The positive effect of a clear distinction between the creditworthiness of enterprises and that of households gets weakened by the negative sign for the coefficient of the product term, difference between risk premiums and overall risk premium. Meaning that in markets where the overall risk premium is high banks generally grant lesser loans and therefore don't produce as much liquidity as banks in countries with lesser credit risk. As a conclusion one can say that in years 2004-2012 the banks were getting more efficient in markets, where the market situation preferred loans to corporations over loans to households and therefore pushing banks towards lending money to corporations, which seemed to be more secure in the last decade. However, in addition the overall credit risk premium should be low to produce even more "good" liquidity.

	Estimate	Std. Error	t value	Pr(>|t|)	lower	upper	#
$lag(SSBM)$	-0.80	0.10	-7.81	0.00	-1.05	-0.54	0
$LS.S$	-0.01	0.01	-0.80	0.42	-0.06	0.03	1357
$LS.D$	0.24	0.06	4.17	0.00	0.06	0.37	0
$diff(GovInt)^2$	-0.01	0.00	-7.09	0.00	-0.02	-0.00	0
$diff(CIR)$	-0.02	0.00	-5.47	0.00	-0.03	-0.01	0
$LS.S : LS.D$	-0.02	0.01	-4.05	0.00	-0.04	-0.00	0

Table 5.11: Results of regressing the Malmquist Index on our basis variables with the transformed lending spreads. The columns contain the estimate, the standard error, the t-statistic and the p-value of the coefficient. Further the table includes the minimum of the lower boundary and the maximum of the upper boundary of the 0.95-confidence interval of all the testregressions. The last column is the amount of testregressions in which the coefficient wasn't significant. For the testregression combinations of 4 variables out of 15 (see Table A.10) were used, thus resulting in 1365 testregressions. The minimal adj. r-squared is 0.32 and the maximal adj. r-squared is 0.37

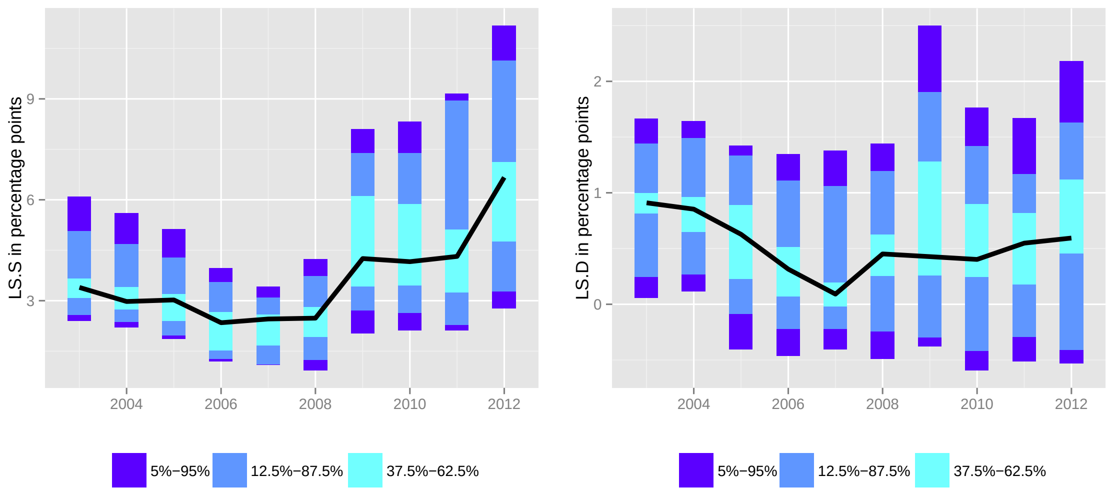

Figure 5.10: Development of the sum, LS.S, and the difference,LS.D, of the lending spreads. The plots show the quantiles of the variables for the 18 home countries of the observed banks. The continous line is the median value.

5.2.5 Residualanalysis

At last we are going to analyse the residuals of our basis model. As a reminder our basis model is a two-ways fixed effects model with our five stable variables: the lagged efficiency score, the lending spread on household credits, the lending spread on corporate credits, the squared first differences of the government interest and the first differences of the cost-income ratio.
In the first instance we tested if the residuals are serially correlated both with a Breusch-Godfrey test for Panel Data and the Wooldridge Test for Panel Data.
The Breusch-Godfrey Test rejects the null of serial uncorrelated idiosyncratic errors ($chisq = 12.57, p - value = 0.01$), while the Wooldridge Test accepts the null ($chisq = 0.01, p - value = 0.92$). As mentioned in chapter 3 the Wooldridge Test is more suitable for small panels, since the errors of a with-in estimation are correlated with $-\frac{1}{T}$ (see [18] for further details).
The next task was to test for cross sectional dependence in the residuals of the regression. At first we conducted a scaled LM test, which produced a highly significant result($CD = 15.47, p - value < 2.2e - 16$). But the test is based on a large T asymptotics (see [9] and [15]). Alternatively we considered the CD test by Pesaran (see [15]) which is also suitable for small T panels. A global version of the test accepts the null, that the residuals are uncorrelated between the banks ($CD = -0.67, p-value = 0.50$).
Due to the fact that the test statistic is an weigthed average of all the pairwise correlations, this test has a poor power if positive as well as negative correlations prevail.
Thus we used a localized version of the CD test were we use only pairwise correlations between neighbours. We define neighbours in two different methods:

- banks are neighbours if they are from the same country
- bank A and bank B are neighbours if in any of the years bank A was a peer to bank B or bank B was a peer to bank A

In the case were neighbourhood is defined by the peer-relationship the local PCD-test accepts the null ($CD = 0.89, p - value = 0.37$), but in the case were neighbourhood is defined by the common country the local PCD-test rejects the null ($CD = 3.44, p - value = 0.00$).
These results indicate that there is indeed correlation between the banks. Also the problem whether there is serial correlation in the errors or not isn't clearly solved.
This is the reason why we finally decided to use the robust covariance estimator from Driscoll and Kraay, that accounts for cross sectional dependence as well as for serial correlation.

5.2.6 Further investigations on GDP growth

Although we couldn't detect a stable significant influence of the GDP growth on the systemic efficiency growth of the banking sector, doesn't mean that there is no dependence between the banking sector and the real side of the economy. So far we have regressed the Malmquist index on the GDP growth, but couldn't find any significant dependence, thus we tried to regress the GDP growth on an weighted average

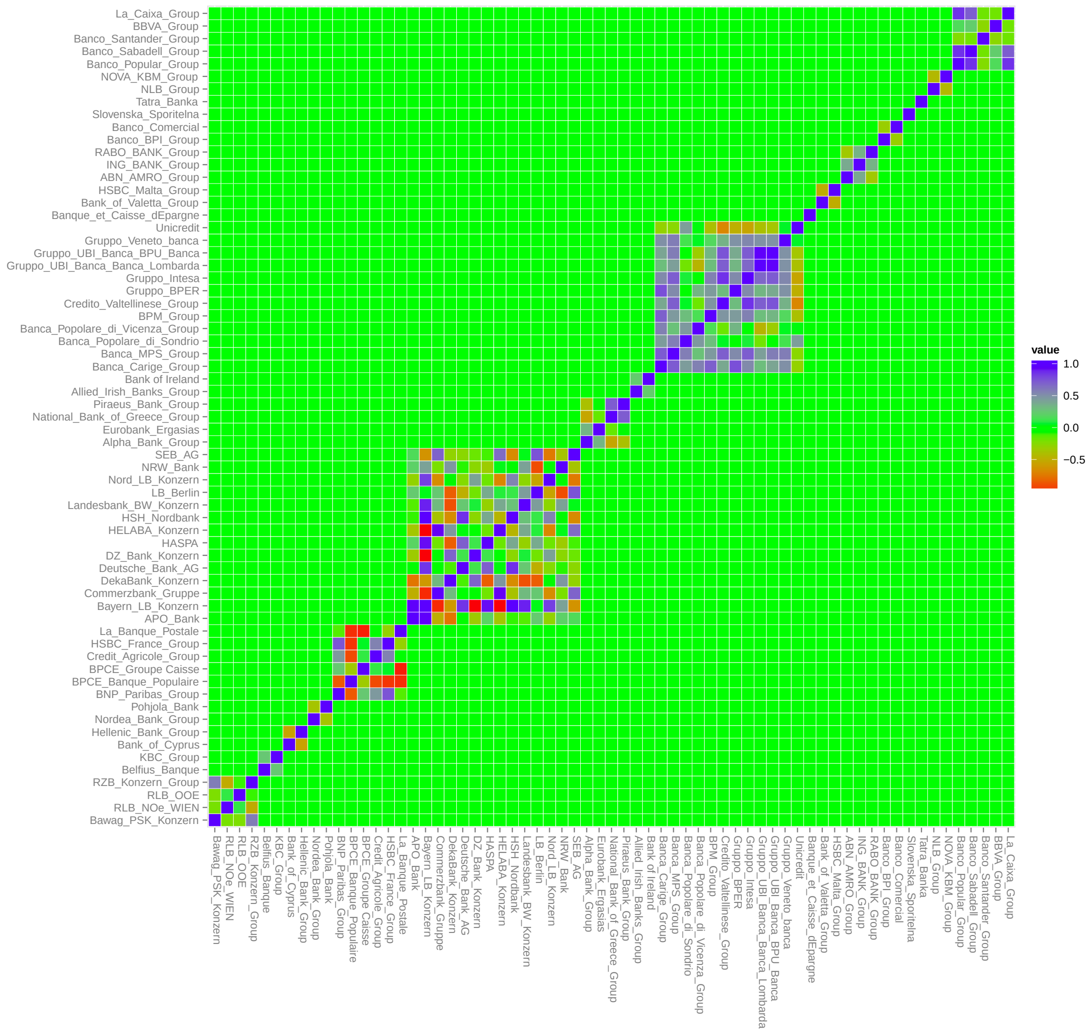

Figure 5.11: Correlationmatrix of the banks. The correlation of banks from different countries is set to 0.

of the Malmquist index (we weighted the Malmquist index with the total assets of the corresponding banks). The results in Table 5.12 show that there is indeed a positive correlation between GDP growth and banking efficiency growth.

	Estimate	Std. Error	t value	Pr(>|t|)	lower	upper	#
$lag(growth(GDP))$	0.62	0.13	4.58	0.00	-0.03	1.03	3
MQ	0.02	0.01	2.90	0.00	-0.00	0.10	7
$TDebt$	-0.00	0.00	-4.00	0.00	-0.00	0.00	16
$HICP$	-0.01	0.00	-2.73	0.01	-0.02	-0.00	0

Table 5.12: Results of regressing the GDP growth on the lagged GDP growth, the Malmquist index, the first difference in the Debt ratio and the change of the harmonised consumer price index. For the testregression combinations of 3 variables out of 10 (see Table A.10) were used, thus resulting in 120 testregressions. The minimal adj. r-squared is 0.61 and the maximal adj. r-squared is 0.68.

Chapter 6

Conclusion

6.1 Summary

In the last years the banking sector has been subject to a lot of criticism from those who demand stricter supervision as well as from those who favor more freedom on the financial markets. Aim of this work wasn't to speak in behalf of any of those sides. Rather we wanted to analyze the current situation of banking supervision and its impact on systemic banking efficiency. At the moment banking regulation only focusses on the sustainability of the sector against economic shocks. Our aim was to find new possibilities to measure the efficiency of banks that also consider whether or not banks fulfilled their macroeconomic tasks like transformation of risk, lot size transformation and maturity transformation. Therefore we used Data Envelopment Analysis to construct an efficiency measure, where we tried to consider the performance in accomplishing these functions. To conduct the efficiency analysis we gathered balance sheet data of 70 (out of 130) system relevant banks that have been under ECB supervision since November 2014. As expected you can clearly see the repercussions of the financial crisis in the balance sheet positions as well as in the efficiency scores. Afterwards we tried to figure out which economic environment is favorable for the systemic efficiency of banks. We conducted panel regression analysis to identify the strategic determinants of banking efficiency. In detail we regressed the Malmquist index, as an indicator for the development of the systemic efficiency, on a multitude of macroeconomic variables, bank specific variables and variables illustrating the conditions on the financial markets. The results of this study are indecisive:

The relationship between the banking efficiency and the growth of the economy is very tricky and our analysis implies that the relationship is unilateral, in that economic growth has no influence on banking efficiency but an increase in the systemic banking efficiency boosts economic growth. This characteristic can be seen completely positive because it implicates that an efficient banking sector can be used to enhance economic expansion.

Unfortunately, we couldn't detect any potential relationships between current political interventions and systemic efficiency. Especially the influence of the ECB interventions on the interest rate level seems to have a different effect than expected.

Further there wasn't any influence from variables like the debt ratio or the primary deficit, on which government could have a direct impact. Solely the change in the interest rate governments have to pay has an impact. However this parameter cannot be directly controlled by governments.

But on the other hand we observed influences of variables that reflect the behavior of the banking sector regarding risk estimation, operative efficiency and the competition on the banking sector. These variables are not exogenous, but nevertheless show how management decisions of banks affect their systemic efficiency. But probably systemic efficiency is not the main driver for these decisions. It would be further interesting if there are possibilities to influence these decisions and thus helping banks serve society.

Whilst studying the efficiency several problems occurred that made the work difficult and demanding:

Collecting balance sheet data was a very sophisticated task, because balance sheet positions are merely harmonized which made it very hard to find comparable statistics. Due to the fact that regulation in the last years didn't really pay attention on the amount of loans that had to be written off or the maturities of claims and liabilities, it was hard or even impossible to get suitable statistics for these amounts. In the course of the asset quality review the ECB worked out guidelines for assessing bank balance sheets and to value them in terms of sustainability. Maybe there will be consistent key performance indicators for the riskiness of a bank's balance sheet, the short term liabilities and the non-performing loans, in the future.

Also gathering data on national banking markets wasn't sufficient enough to analyze every aspect of market characteristics. It would have been nice to have variables for the ratio of domestic to foreign banks in a country, the size of the banking sector compared to the financial industry sector and information on non-performing loans. Although the ECB provides these data now, at the time of the study the time series were unfortunately not long enough.

At last a critical look has to be taken on the method of DEA for assessing bank efficiency and the choice of the used factors. A feature of DEA that makes it vulnerable to criticism is the complexity and opaqueness of the scores. It isn't desirable that every DMU that uses the least amount of any input factor is automatically efficient (when using variable returns to scale), regardless of the output it produces. The same is true for a DMU that is able to produce the highest amount (in the sample) of any output. They are efficient regardless of the amount of inputs they consume. When further using super-efficiency this can lead to extremely high efficiency you didn't expect and you can't explain economically. Therefore results of DEA analysis should be examined with caution and the choice of the factors should be well considered.

6.2 Author's opinion

Conclusively I feel the urge to express my personal opinion on the topic and the study in hand. Although I am no advocate of governmental control of the economy, I consider banking supervision necessary in the current system. As has been mentioned a few times during this work banks fulfill very essential tasks in our society. Therefore there are a lot of liberties and rights exclusively granted to the banking sector. If it is legitimate to grant those benefits to banks is a question for philosophers and politician. But currently those benefits are granted and according to the motto "no rights without duties", I am of the opinion that society has the right and the obligation to monitor if banks do their duties. Personally, I think that current banking supervision is not proper to ensure that banks do their duties. Because I don't think that ensuring sustainability achieves that money arrives were it should arrive, therefore extending or introducing the term banking efficiency is a step in the right direction. Using DEA to retrieve efficiency score seems like a plausible first step and can also be very helpful in analyzing the developments on the banking sector but ultimately I am pessimistic that DEA can provide guidelines for banking supervision. I think proper banking regulation should be based on a variety of indicators, not just one.

Nevertheless I am glad that I spent nearly the last two years studying the European banking sector and I hope that one or another can use this study to conduct further research on this topic.

Appendix A

Tables

abbreviation	description	DEA
TA	total assets	input
CaB	cash and balances with central banks	
$DebtS$	debt securities	
$GovS$	government or public issued debt securities	
LtB	liabilities to other banks	
CoB	claims on other banks	
Eq	equity	
CIR	cost-income ratio	
CL	customer loans	
$DeVAL$	(gross) provision for loan losses and write-offs for devaluation of loans	
$Eq.TA = \frac{Eq}{TA}$	equity ratio	
$RA = TA - CaB - GovS$	risky assets	
$RA.TA = \frac{RA}{TA}$	risk load of the bank's (on-balance-sheet) business	input
$LtB.TA = \frac{LtB}{TA}$	interbank debt ratio	input
$LP = CL + CoB + DebtS$	liquidity production	output
$IP = \frac{CL+CoB}{CL+CoB+DeVAL}$	information production	output
$SI = \frac{Eq}{RA}$	stability index	output
$SSBM$	super slack based measure of systemic efficiency	
CU	catch up	
FS	frontier shift	
MQ	Malmquist index	

Table A.1: This table contains the balance sheet items we extracted from the banks annual reports. The column DEA indicates wheter factor is used for the estimation of the efficiency index or not. If not explicitly mentioned, all the indicators that represent factions like Eq.TA, RA.TA, LtB.TA, IP and SI, are given as values between 0 and 1 (not as percentage).

abbreviation	source	description
GDP	ECB	real gross domestic product at market prices, reference year 2005
$HICP$	ECB	inflation rate measured as change in the harmonised index of consumer prices
UU	ECB	unemployment rate
$PriDef$	ECB	government primary deficit (-) or surplus (+) (as % of GDP)
$TDebt/FDebt$	ECB	general government gross debt as defined in Council Regulation (EC) No 479/2009: total debt (as % of GDP) and foreign debt (as % of GDP)
$GovInt$	ECB	secondary market yields of government bonds with a remaining maturity close to ten years
$Herfindahl$	ECB	Herfindahl index for credit institutions total assets (TA): $\sum_{i=1}^{N} a_i^2$, where $a_i = TA_i \left(\sum_{j=1}^{N} TA_j\right)^{-1}$
$Top5$	ECB	The market share of the five biggest banks measured by the total assets
$MCap$	WB	Market capitalization of all listed stocks in the corresponding country (as % of GDP)
$LS.HH/LS.NFC$	ECB	weigthed spread between the MFI interest rate for new loans to households/non-financial corporations and the swap rate with a maturity corresponding to the loan category initial period of fixation
$Credit\ by\ banks$	WB	Domestic credit provided by the banking sector (as % of GDP) (includes all credit to various sectors on a gross basis, with the exception of credit to the central government, which is net)
$stocks\ trade$	WB	The total value of shares traded during the period (as % of GDP)
$GovDep$	ECB	The amount of deposits from central governments at monetary financial institutions at the end of the year
$GDPEuro17$	ECB	Real gross domestic product at market prices for the Euro 17 area, reference year 2005
$ECBrate$	ECB	The lending rate of the ECB main refinancing rate (period average)
$Euribor$ $1m/3m/6m/1y$	ECB	Euribor rates with a maturity of 1,3 and 6 months and 1 year (period average)
$StoxxF/B/50$	Stoxx	closing-values of the Stoxx 600 Euro Financial Service Industry Index, Stoxx Euro 600 Banks Index and the Euro Stoxx 50 Index
$M1, M2, M3$	ECB	Monetary aggregates M1, M2 and M3

Table A.2: This table contains the macroeconomic variables, their description and the data source. The Sources are: ECB: European Central Bank (http:/sdw.ecb.europa.eu), WB: World Bank (http://data.worldbank.org/), Stoxx: STOXX Limited (http://www.stoxx.com/)

country	crisis years
Austria	2008-2012
Belgium	2008-2012
Cyprus	2012
Estonia	
Finland	
France	2008-2012
Germany	2008-2012
Greece	2008-2012
Ireland	2008-2012
Italy	2008-2012
Latvia	2008-2012
Luxembourg	2008-2012
Malta	
Netherlands	2008-2012
Portugal	2008-2012
Slovakia	2002
Slovenia	2008-2012
Spain	2008-2012

Table A.3: Periods of systemic banking crisis. The data are taken from [13]

institute	cc	TA	CL	LtB	Eq	group
ABLV_Bank	LV	1382	708	12	98	small
ABN_AMRO_Group	NL	584530	288786	102754	16942	large
Allied_Irish_Banks_Group	IE	140937	85791	29859	8772	big
Alpha_Bank_Group	EL	56521	41284	8546	3368	medium
APO_Bank	DE	37479	23295	9226	1665	medium
AS_SEB_Banka_Group	LV	3902	2698	1782	315	small
Banca_Carige_Group	IT	29725	18967	1506	2869	medium
Banca_MPS_Group	IT	187936	124202	19035	8212	big
Banca_Popolare_di_Sondrio	IT	19432	13669	1949	1542	medium
Banca_Popolare_di_Vicenza	IT	28094	21798	3178	2829	medium
Banco_BPI_Group	PT	41751	27288	4266	1816	medium
Banco_Comercial	PT	88955	64134	11312	4984	big
Banco_Popular_Group	ES	108773	89375	11954	6851	big
Banco_Sabadell_Group	ES	78577	62503	6568	4526	big
Banco_Santander_Group	ES	981274	598994	134995	58780	large
Bank of Ireland	IE	158617	100280	20635	6480	big
Bank_of_Cyprus	CY	31398	21648	842	1785	medium
Bank_of_Valetta_Group	MT	5952	2830	735	397	small
Banque_et_Caisse_dEpargne	LU	38627	11230	6512	2146	medium
Bawag_PSK_Konzern	AT	43213	22756	3898	1712	medium
Bayern_LB_Konzern	DE	335960	153013	105935	12079	large
BBVA_Group	ES	518396	318310	67493	27324	large
Belfius_Banque	BE	240206	86516	68244	5372	big
BNP_Paribas_Group	FR	1800872	469752	169084	59181	large
BPCE_Banque_Populaire	FR	376252	172586	68538	20525	large
BPCE_Groupe Caisse	FR	625605	283025	96505	21440	large
BPM_Group	IT	43954	31310	5500	3539	medium
Commerzbank_Gruppe	DE	620835	284142	125473	18018	large
Credit_Agricole_Group	FR	1617313	536267	126010	69501	large
Credito_Valtellinese_Group	IT	20396	16060	1220	1890	medium
DekaBank_Konzern	DE	128831	24283	30694	3268	big
Deutsche_Bank_AG	DE	1579394	337363	346180	22365	large
DZ_Bank_Konzern	DE	403777	109552	105538	9794	large
Erste_Bank_Group	AT	200980	120071	27423	11249	big
Eurobank_Ergasias	EL	68021	44405	5244	3512	medium
Gruppo_BPER	IT	50658	37802	2178	3944	medium
Gruppo_Intesa	IT	598873	354653	47557	48906	large
Gruppo_UBI_Banca_Lombarda	IT	121724	92915	4652	10208	big
Gruppo_UBI_BPU_Banca	IT	121724	92915	5831	10208	big
Gruppo_Veneto_banca	IT	19425	15302	1224	2106	medium

institute	cc	TA	CL	LtB	Eq	group
HASPA	DE	35341	22538	4872	1597	medium
HELABA_Konzern	DE	166961	82196	36620	4906	big
Hellenic_Bank_Group	CY	7592	4143	330	444	small
HSBC_France_Group	FR	204732	49067	32665	5102	big
HSBC_Malta_Group	MT	5006	2967	201	307	small
HSH_Nordbank	DE	173072	94039	45920	5183	big
ING_BANK_Group	NL	888552	533935	107456	31452	large
KBC_Group	BE	322527	135342	43950	16975	big
La_Banque_Postale	FR	146627	30936	7134	4182	big
La_Caixa_Group	ES	254663	165032	13841	19937	big
Landesbank_BW_Konzern	DE	389664	120698	114442	10210	large
LB_Berlin	DE	141891	48201	41690	2698	big
National_Bank_of_Greece	EL	96113	61914	12607	6292	big
NLB_Group	SI	15427	10108	3305	1015	medium
Nord_LB_Konzern	DE	214322	100263	62197	5822	big
Nordea_Bank_Group	FI	431564	254891	36512	17482	large
NOVA_KBM_Group	SI	5210	3277	152	346	small
NRW_Bank	DE	150048	56131	40415	18568	big
Piraeus_Bank_Group	EL	47890	32148	12414	1720	medium
Pohjola_Bank	FI	26765	9720	4677	1421	medium
RABO_BANK_Group	NL	588987	399626	36696	32434	large
RLB_NOe_WIEN	AT	23541	7698	9410	1908	medium
RLB_OOE	AT	29071	15536	10006	1956	medium
RZB_Konzern_Group	AT	136950	73963	41645	5002	big
SEB_AG	DE	50395	21603	16148	2238	medium
Slovenska_Sporitelna	SK	10039	5190	1183	720	small
SWEDBANK_AS	EE	16066	11886	4331	1689	medium
Tatra_Banka	SK	8148	4708	191	597	small
Unicredit	IT	322946	28881	59891	50125	big
Volksbanken_AG_Konzern	AT	47291	23892	12516	2090	medium
median		94987	47244	14403	4134	
mean		238628	108886	39162	10733	

Table A.4: This table contains for balance sheet data for each observed bank. The column cc shows the country code, where the headquarter of the banking group is located, the columns TA, CL, LtB and Eq show the median values for the banks in MEUR. The column group shows our classification of the bank, after clustering the banks by their total assets.

	LP	IP	SI	RA.TA	LtB.TA	SSBM	nobs
ABLV_Bank	1290	0.99	0.08	0.90	0.02	2.22	10
ABN_AMRO_Group	466479	1.00	0.03	0.92	0.17	1.01	10
Allied_Irish_Banks_Group	121468	0.99	0.06	0.91	0.22	0.78	10
Alpha_Bank_Group	50183	0.99	0.08	0.89	0.16	0.68	10
APO_Bank	31768	1.00	0.05	0.99	0.24	0.50	10
AS_SEB_Banka_Group	3376	0.99	0.09	0.90	0.47	1.01	10
Banca_Carige_Group	23587	0.99	0.11	0.94	0.07	0.82	10
Banca_MPS_Group	148834	0.99	0.05	0.93	0.12	0.58	10
Banca_Popolare_di_Sondrio	16142	0.99	0.09	0.90	0.10	0.72	10
Banca_Popolare_di_Vicenza	24583	0.99	0.10	0.96	0.13	0.70	10
Banco_BPI_Group	36690	1.00	0.05	0.91	0.12	0.69	10
Banco_Comercial	77858	0.99	0.07	0.94	0.15	0.61	10
Banco_Popular_Group	100804	0.99	0.07	0.97	0.14	0.68	10
Banco_Sabadell_Group	70878	0.99	0.06	0.95	0.09	0.66	10
Banco_Santander_Group	776466	0.98	0.07	0.90	0.13	1.04	10
Bank of Ireland	133933	1.00	0.04	0.94	0.15	0.51	10
Bank_of_Cyprus	28504	0.99	0.07	0.90	0.03	0.87	10
Bank_of_Valetta_Group	5489	0.99	0.09	0.80	0.15	0.79	10
Banque_et_Caisse_dEpargne	35475	1.00	0.06	0.89	0.17	1.00	10
Bawag_PSK_Konzern	37877	0.99	0.04	0.92	0.09	0.49	10
Bayern_LB_Konzern	303412	1.00	0.03	0.97	0.30	0.82	4
BBVA_Group	430433	0.99	0.07	0.85	0.16	0.77	10
Belfius_Banque	196643	1.00	0.03	0.92	0.30	0.35	10
BNP_Paribas_Group	855586	0.99	0.04	0.86	0.10	1.00	10
BPCE_Banque_Populaire	279433	0.99	0.05	0.95	0.16	0.76	6
BPCE_Groupe Caisse	485891	1.00	0.04	0.96	0.13	0.74	10
BPM_Group	37374	0.99	0.09	0.95	0.12	0.64	10
Commerzbank_Gruppe	495351	0.99	0.03	0.89	0.20	0.52	10
Credit_Agricole_Group	955016	0.99	0.05	0.91	0.08	1.10	10
Credito_Valtellinese_Group	17773	0.99	0.08	0.97	0.07	0.68	10
DekaBank_Konzern	102414	1.00	0.02	0.94	0.26	0.51	4
Deutsche_Bank_AG	686235	1.00	0.02	0.93	0.25	1.00	10
DZ_Bank_Konzern	298367	1.00	0.03	0.94	0.27	0.38	10
Erste_Bank_Group	170790	0.99			0.17		0
Eurobank_Ergasias	59213	0.98	0.09	0.81	0.18	0.81	10
Gruppo_BPER	43504	0.99	0.08	0.94	0.05	0.68	10
Gruppo_Intesa	444873	0.99	0.08	0.96	0.12	0.75	10
Gruppo_UBI_Banca_Lombarda	103266	0.99	0.09	0.97	0.07	0.75	10
Gruppo_UBI_BPU_Banca	103266	0.99	0.09	0.94	0.07	0.77	10
Gruppo_Veneto_banca	16976	0.99	0.11	0.95	0.07	0.96	10

	LP	IP	SI	RA	LtBtoTA	SSBM	nobs
HASPA	30282	0.99	0.05	0.97	0.13	0.53	10
HELABA_Konzern	148492	1.00	0.03	0.98	0.20	0.47	10
Hellenic_Bank_Group	7118	0.98	0.07	0.82	0.06	0.71	10
HSBC_France_Group	127087	1.00	0.03	0.83	0.18	0.43	10
HSBC_Malta_Group	4380	1.00	0.08	0.85	0.04	1.00	10
HSH_Nordbank	158973	0.99	0.04	0.95	0.27	0.42	10
ING_BANK_Group	727923	1.00	0.04	0.91	0.14	1.02	10
KBC_Group	254097	0.99	0.07	0.82	0.17	0.70	10
La_Banque_Postale	133074	1.00	0.04	0.70	0.06	1.05	8
La_Caixa_Group	213245	1.00	0.08	0.93	0.08	0.75	10
Landesbank_BW_Konzern	335299	1.00	0.03	0.97	0.31	0.73	10
LB_Berlin	120693	1.00	0.02	0.96	0.29	0.27	4
National_Bank_of_Greece_Group	80437	0.99	0.09	0.82	0.16	0.78	10
NLB_Group	13709	0.96	0.08	0.82	0.22	0.58	10
Nord_LB_Konzern	197208	1.00	0.03	0.95	0.29	0.52	10
Nordea_Bank_Group	332950	1.00	0.05	0.96	0.09	0.64	10
NOVA_KBM_Group	4576	0.95	0.09	0.84	0.03	1.10	10
NRW_Bank	143854	1.00	0.14	0.82	0.31	1.10	10
Piraeus_Bank_Group	38660	0.99	0.08	0.85	0.24	0.60	10
Pohjola_Bank	21218	1.00	0.06	0.78	0.15	1.00	10
RABO_BANK_Group	500887	1.00	0.06	0.91	0.06	1.02	10
RLB_NOe_WIEN	19070	0.99	0.07	0.96	0.40	0.58	10
RLB_OOE	24965	0.99	0.07	0.97	0.35	0.51	10
RZB_Konzern_Group	116626	0.99	0.04	0.89	0.35	0.42	10
SEB_AG	44110	1.00	0.05	0.91	0.33	0.81	10
Slovenska_Sporitelna	9414	0.97	0.10	0.70	0.13	1.03	10
SWEDBANK_AS	14972	0.99	0.10	0.93	0.25	0.69	10
Tatra_Banka	7612	0.99	0.12	0.74	0.03	1.06	10
Unicredit	219047	1.00	0.15	0.98	0.36	1.00	10
Volksbanken_AG_Konzern	39205	0.99			0.30		0
median	84650	0.99	0.07	0.91	0.15	0.73	10
mean	176453	0.99	0.07	0.89	0.18	0.77	9.37

Table A.5: The table shows statistics for the DEA factors of the banks. The values are the median values over the 10 years. The column nobs shows the amount of years, for which all the relevant data for our DEA model were available.

	SSBM 03	MQ 03-07	SSBM 07.	MQ 07-12	SSBM 12
ABLV_Bank	2.91	0.34	1.74	2.27	2.54
ABN_AMRO_Group	1.01	0.91	0.80	1.14	1.01
Allied_Irish_Banks_Group	0.78	1.36	1.00	1.01	1.00
Alpha_Bank_Group	0.61	1.13	0.86	0.21	0.14
APO_Bank	0.65	0.78	0.51	0.98	0.46
AS_SEB_Banka_Group	0.60	1.19	1.00	1.09	1.07
Banca_Carige_Group	0.67	0.90	0.85	0.82	0.53
Banca_MPS_Group	0.51	1.15	0.64	0.74	0.34
Banca_Popolare_di_Sondrio	0.63	0.96	0.74	0.89	0.51
Banca_Popolare_di_Vicenza	0.63	1.01	0.78	0.90	0.56
Banco_BPI_Group	0.64	1.46	0.86	1.06	1.00
Banco_Comercial	0.53	1.06	0.66	0.85	0.41
Banco_Popular_Group	0.65	1.42	1.01	0.54	0.48
Banco_Sabadell_Group	0.66	1.35	1.00	0.64	0.45
Banco_Santander_Group	0.69	1.53	1.06	1.08	1.05
Bank of Ireland	1.00	1.21	1.00	0.56	0.48
Bank_of_Cyprus	0.65	1.35	1.00	0.57	0.25
Bank_of_Valetta_Group	0.60	0.92	0.87	1.24	0.79
Banque_et_Caisse_dEpargne	1.00	1.00	1.00	1.00	1.00
Bawag_PSK_Konzern	0.36	1.16	0.54	1.38	0.55
BBVA_Group	0.75	1.12	0.84	1.15	0.71
Belfius_Banque	0.73	0.67	0.31	0.68	0.35
BNP_Paribas_Group	0.67	1.64	1.01	1.10	1.00
BPCE_Groupe Caisse	0.64	1.36	0.69	1.42	0.82
BPM_Group	0.52	1.18	0.73	0.96	0.57
Commerzbank_Gruppe	0.39	1.56	0.62	1.07	0.50
Credit_Agricole_Group	1.20	1.00	1.06	1.43	1.10
Credito_Valtellinese_Group	0.67	1.10	1.00	0.71	0.49
Deutsche_Bank_AG	1.01	1.48	1.06	0.58	1.00
DZ_Bank_Konzern	0.42	2.02	0.52	0.84	0.38
Eurobank_Ergasias	0.85	1.17	1.10	0.01	0.01
Gruppo_BPER	0.63	1.14	0.84	0.91	0.57
Gruppo_Intesa	0.66	1.53	1.00	0.94	0.69
Gruppo_UBI_Banca_Lombarda	0.59	1.44	1.00	0.88	0.71
Gruppo_UBI_BPU_Banca	0.62	1.42	1.00	0.88	0.71
Gruppo_Veneto_banca	0.64	1.05	1.02	0.67	0.53
HASPA	0.44	1.34	0.59	1.74	1.00
HELABA_Konzern	0.62	1.03	0.44	1.23	0.52
Hellenic_Bank_Group	0.58	1.30	1.00	1.12	0.69
HSBC_France_Group	0.77	0.63	0.43	0.71	0.29

	SSBM 03	MQ 03-07	SSBM 07.	MQ 07-12	SSBM 12
HSBC_Malta_Group	1.00	1.02	1.32	0.87	1.00
HSH_Nordbank	0.63	0.50	0.32	1.49	0.41
ING_BANK_Group	1.04	1.01	1.01	1.12	1.04
KBC_Group	0.60	1.11	0.71	1.17	0.62
La_Caixa_Group	0.71	1.38	1.00	0.68	0.59
Landesbank_BW_Konzern	1.01	0.99	1.00	0.43	0.41
National_Bank_of_Greece	1.00	0.87	1.00	0.02	0.01
NLB_Group	0.58	0.74	0.62	1.11	0.53
Nord_LB_Konzern	0.39	2.34	1.00	0.51	0.51
Nordea_Bank_Group	0.62	1.45	0.78	0.78	0.57
NOVA_KBM_Group	1.15	0.57	1.05	0.87	0.46
NRW_Bank	0.58	1.93	1.18	0.92	1.02
Piraeus_Bank_Group	0.65	0.84	0.74	0.02	0.01
Pohjola_Bank	1.00	1.00	1.00	1.00	1.00
RABO_BANK_Group	1.00	1.08	1.03	1.12	1.01
RLB_NOe_WIEN	0.53	1.15	0.66	0.96	0.58
RLB_OOE	0.42	1.18	0.56	1.11	0.47
RZB_Konzern_Group	0.32	1.26	0.45	1.21	0.42
SEB_AG	1.00	0.58	0.50	1.20	1.00
Slovenska_Sporitelna	1.03	0.88	1.08	1.34	1.04
SWEDBANK_AS	0.71	0.92	0.65	1.46	1.12
Tatra_Banka	1.02	0.75	1.05	1.52	1.18
Unicredit	1.00	1.09	1.05	0.89	0.76
median	0.65	1.12	1.00	0.96	0.57
mean	0.75	1.14	0.86	0.95	0.68

Table A.6: The table shows the development of the systemic efficiency for every bank, for which we could calculate the efficiency score for each of the ten years. The Malmquist indices are calculated between the period 03-07 and 07-12.

	min	q.125	median	q.875	max	mean	sd	IQR	#miss
TA (M€)	362	11045	94987	569234	2250665	238628	373885	233712	2
CL (M€)	93	6053	47244	282201	799005	108886	153076	106297	2
LP (M€)	343	9834	84650	454028	1242841	176453	236212	190677	2
RA (M€)	298	7779	73208	533583	2151354	219732	347424	215384	44
Eq (M€)	-2316	664	4134	24054	94422	10733	16052	10583	2
CaB (M€)	14	154	1079	8286	193189	5215	15635	3171	2
GovS (M€)	17	550	4719	41748	272205	17662	32636	17560	44
DebtS (M€)	44	1518	13121	95623	377532	39284	60530	46965	2
LtB (M€)	5	1160	14403	102058	495532	39162	60030	48040	2
CoB (M€)	59	919	7418	69926	259894	28283	43796	34765	2
DeVAL (M€)	-34618	-2556	-335	-47	0	-1230	2763	1180	2
LP.TA	0.26	0.76	0.87	0.93	0.98	0.84	0.09	0.09	2
CL.TA	0.07	0.33	0.57	0.73	0.86	0.54	0.17	0.25	2
RA	0.57	0.81	0.91	0.97	1.00	0.89	0.08	0.10	44
IP	0.90	0.98	0.99	1.00	1.00	0.99	0.01	0.01	2
SI	-0.05	0.03	0.07	0.10	0.24	0.07	0.03	0.04	44
LtBtoTA	0.00	0.06	0.15	0.33	0.52	0.18	0.11	0.16	2
Eq.TA	-0.04	0.03	0.06	0.09	0.20	0.06	0.03	0.04	2
CaB.TA	0.00	0.00	0.01	0.05	0.22	0.02	0.02	0.02	2
GovS.TA	0.00	0.02	0.06	0.16	0.41	0.08	0.07	0.08	44
DebtS.TA	0.01	0.06	0.17	0.28	0.56	0.17	0.10	0.13	2
CoB.TA	0.01	0.04	0.09	0.25	0.64	0.13	0.10	0.12	2
DeVAL.TA	-0.08	-0.01	-0.00	-0.00	0.00	-0.01	0.01	0.01	2
CIR	0.18	0.45	0.60	0.75	24.39	0.65	0.95	0.17	5
SSBM	0.01	0.45	0.73	1.03	3.17	0.77	0.33	0.45	44
FS	0.41	0.89	1.01	1.17	2.05	1.03	0.16	0.13	113
EC	0.02	0.78	1.00	1.23	3.17	1.02	0.32	0.18	113
MQ	0.02	0.82	1.01	1.21	3.23	1.03	0.29	0.16	113

Table A.7: This table contains statistics of choosen bank indicators. Absolute values are given in Million Euros. The column sd shows the standard deviation and the column IQR the Interquartile Range (q.75 - q.25).

Var	Period	min	q.125	median	q.875	max	IQR	d(%)	s(%)
TA	2003-2007	-0.03	0.06	0.15	0.27	0.45	0.13	0.36	0.03
TA	2008-2012	-0.23	-0.07	0.01	0.10	0.30	0.09	0.04	0.46
TA	2003-2012	-0.09	-0.02	0.04	0.10	0.18	0.06	0.24	0.17
LP	2003-2007	-0.03	0.04	0.13	0.27	0.45	0.14	0.33	0.03
LP	2008-2012	-0.26	-0.08	0.01	0.09	0.31	0.08	0.03	0.47
LP	2003-2012	-0.08	-0.03	0.04	0.10	0.18	0.07	0.19	0.21
CL	2003-2007	-0.12	0.07	0.16	0.32	0.80	0.13	0.42	0.04
CL	2008-2012	-0.26	-0.07	0.01	0.08	0.63	0.08	0.03	0.43
CL	2003-2012	-0.11	-0.00	0.06	0.11	0.41	0.05	0.30	0.13
CL.TA	2003-2007	-0.11	-0.05	0.01	0.06	0.25	0.04	0.01	0.36
CL.TA	2008-2012	-0.22	-0.04	-0.01	0.05	0.56	0.05	0.01	0.57
CL.TA	2003-2012	-0.11	-0.01	0.01	0.04	0.25	0.04	0.01	0.39
LP.TA	2003-2007	-0.08	-0.03	-0.00	0.01	0.02	0.02	0.00	0.67
LP.TA	2008-2012	-0.06	-0.02	-0.01	0.02	0.08	0.02	0.00	0.63
LP.TA	2003-2012	-0.08	-0.02	-0.00	0.00	0.02	0.01	0.00	0.67
RA.TA	2003-2007	-0.01	-0.00	0.01	0.03	0.08	0.02	0.00	0.25
RA.TA	2008-2012	-0.08	-0.03	-0.01	0.00	0.07	0.02	0.00	0.81
RA.TA	2003-2012	-0.04	-0.01	-0.00	0.01	0.03	0.01	0.00	0.66
CaB.TA	2003-2007	-0.50	-0.13	0.00	0.27	1.38	0.20	0.17	0.49
CaB.TA	2008-2012	-0.32	-0.13	0.05	0.50	1.73	0.31	0.34	0.40
CaB.TA	2003-2012	-0.10	-0.02	0.08	0.26	0.39	0.17	0.49	0.20
CoB.TA	2003-2007	-0.33	-0.16	-0.03	0.14	0.39	0.17	0.07	0.58
CoB.TA	2008-2012	-0.49	-0.15	-0.05	0.08	0.40	0.16	0.04	0.64
CoB.TA	2003-2012	-0.27	-0.12	-0.05	0.02	0.13	0.08	0.04	0.77
LtB.TA	2003-2007	-0.37	-0.14	-0.04	0.09	0.72	0.12	0.03	0.62
LtB.TA	2008-2012	-0.66	-0.19	-0.01	0.25	0.93	0.24	0.17	0.51
LtB.TA	2003-2012	-0.23	-0.09	-0.02	0.10	0.50	0.09	0.16	0.66
GovS.TA	2003-2007	-0.50	-0.28	-0.10	0.05	0.63	0.15	0.05	0.78
GovS.TA	2008-2012	-0.53	-0.09	0.20	0.59	1.59	0.38	0.52	0.25
GovS.TA	2003-2012	-0.32	-0.06	0.01	0.16	0.40	0.15	0.31	0.47
Eq.TA	2003-2007	-0.17	-0.09	0.00	0.12	0.29	0.09	0.04	0.48
Eq.TA	2008-2012	-0.34	-0.07	0.04	0.17	0.41	0.12	0.07	0.31
Eq.TA	2003-2012	-0.17	-0.04	0.01	0.06	0.10	0.06	0.04	0.37
SI	2003-2007	-0.19	-0.10	-0.00	0.11	0.29	0.10	0.03	0.54
SI	2008-2012	-0.34	-0.05	0.05	0.18	0.43	0.12	0.11	0.28
SI	2003-2012	-0.18	-0.04	0.01	0.06	0.11	0.05	0.05	0.36
DeVAL.TA	2003-2007	-0.64	-0.24	-0.09	0.07	0.56	0.15	0.03	0.79
DeVAL.TA	2008-2012	-1.00	-0.05	0.15	0.47	1.55	0.22	0.37	0.20
DeVAL.TA	2003-2012	-1.00	-0.00	0.12	0.23	1.24	0.15	0.64	0.13
IP	2003-2007	-0.01	-0.00	0.00	0.00	0.02	0.00	0.00	0.23
IP	2008-2012	-0.02	-0.01	-0.00	0.00	0.00	0.00	0.00	0.77
IP	2003-2012	-0.01	-0.00	-0.00	0.00	0.00	0.00	0.00	0.86

Table A.8: The table contains the annualized growth rates for selected balance sheet items. IQR is the interquartile range of the growth rates, which is the 75%-quantile-25%q-quantile. d(%) is the percentage of institutes that at least doubled the concerned balance sheet item in the period and s(%) is the percentage of institutes for which the concerned item declined.

	2004	2005	2006	2007	2008	2009	2010	2011	2012
TA median	10.73	17.94	11.50	12.22	7.20	-0.12	1.36	0.45	0.92
TA p-value	0.00	0.00	0.00	0.00	0.00	0.90	0.40	0.72	0.28
LP median	11.02	18.83	10.81	11.96	5.99	1.87	1.25	0.10	0.49
LP p-value	0.00	0.00	0.00	0.00	0.00	0.07	0.55	1.00	0.90
CL median	10.71	16.44	15.03	16.89	11.62	1.52	3.62	0.81	-1.01
CL p-value	0.00	0.00	0.00	0.00	0.00	0.28	0.00	0.55	0.55
CL.TA median	0.36	-0.59	1.98	1.43	2.10	1.06	0.79	-0.66	-0.65
CL.TA p-value	0.34	0.63	0.00	0.04	0.00	0.55	0.04	0.19	0.12
LP.TA median	0.13	0.01	0.19	-0.27	-0.76	0.73	-0.08	-1.23	-0.56
LP.TA p-value	0.34	1.00	0.40	0.55	0.12	0.00	0.72	0.00	0.28
RA median	0.76	0.12	0.75	0.73	0.25	-1.69	-0.70	-0.19	-1.28
RA p-value	0.00	0.90	0.00	0.10	0.53	0.00	0.02	1.00	0.00
CaB.TA median	-0.08	-0.09	0.03	0.08	0.05	-0.01	-0.12	0.02	0.41
CaB.TA p-value	0.15	0.02	0.55	0.01	0.40	0.90	0.02	0.72	0.00
CoB.TA median	0.19	-0.87	-0.55	-1.01	-1.62	-0.26	-0.83	0.01	-0.22
CoB.TA p-value	0.34	0.09	0.12	0.07	0.00	0.72	0.00	0.90	0.02
LtBtoTA median	-0.34	0.11	-0.16	-0.24	-0.72	-0.70	-1.05	0.66	-0.29
LtBtoTA p-value	0.63	1.00	0.55	0.55	0.19	0.28	0.01	0.01	0.55
GovS.TA median	-0.32	0.05	-1.02	-0.80	-0.28	2.03	0.74	-0.17	0.84
GovS.TA p-value	0.14	0.71	0.00	0.00	0.06	0.00	0.01	0.80	0.00
Eq.TA median	-0.08	0.01	0.08	0.03	-0.59	0.55	0.08	-0.17	0.28
Eq.TA p-value	0.15	1.00	0.28	0.72	0.00	0.00	0.19	0.02	0.00
SI median	-0.23	-0.06	-0.00	-0.04	-0.68	0.86	0.19	-0.13	0.43
SI p-value	0.01	0.54	1.00	0.90	0.00	0.00	0.32	0.14	0.00
DeVAL.TA median	0.04	0.02	0.01	0.00	-0.15	-0.24	0.08	-0.00	-0.03
DeVAL.TA p-value	0.00	0.02	0.19	0.72	0.00	0.00	0.01	0.90	0.19
IP median	0.07	0.03	0.02	0.02	-0.19	-0.31	0.10	-0.01	-0.07
IP p-value	0.00	0.03	0.02	0.55	0.00	0.00	0.01	0.72	0.19
MQ median	1.03	1.00	1.02	1.02	0.96	1.04	1.04	0.90	1.03
MQ binom.test	0.09	0.71	0.14	0.53	0.26	0.02	0.01	0.00	0.08
FS median	0.94	1.04	1.00	0.94	0.98	1.14	1.06	1.01	1.03
FS binom.test	0.01	0.00	0.81	0.00	0.01	0.00	0.00	0.02	0.00
CU median	1.09	0.98	1.00	1.06	1.00	0.92	1.00	0.92	1.00
CU binom.test	0.00	0.18	1.00	0.01	0.71	0.03	0.32	0.00	0.80

Table A.9: This table contains the median growth rates (in percent) resp. the median difference (in percentage points). The p-value refers to the p-value of an exact binomial test, which is used to test whether the median value is different to 0 or not. MQ, FS and EC are given in absolut values (not percentage), and the p-value refers to a binomial test, used to test whether the median value is different to 1 or not.

Table 5.2	$HICP$, $diff(HICP)$, $diff(TDebt)$, $diff(FDebt)$, $diff(Herfindahl)$, $growth(Credit\ by\ banks)$, $growth(stocks\ trade)$, $PriDef$, $MCap$, $TDebt$, $FDebt$, UU, $diff(UU)$, $growth(GDP)$, $GovDep$
Table 5.3	$HICP$, $diff(Herfindahl)$, $growth(Credit\ by\ banks)$, $MCap$, $TDebt$, $FDebt$, UU, $GovDep$, $lag(SSBM)$, $LS.HH$, $LS.NFC$, $diff(GovInterest)^2$, $diff(CIR)$, $growth(stocks\ trade)$
Table 5.4	$HICP$, $diff(HICP)$, $diff(TDebt)$, $diff(FDebt)$, $diff(CIR)$ $PriDef$, $lag(SSBM)$, $LS.HH$, $LS.NFC$, $diff(GovInterest)^2$ $TDebt$, $FDebt$, UU, $diff(UU)$, $growth(GDP)$, $GovDep$
Table 5.5	$HICP$, $diff(HICP)$, $diff(TDebt)$, $diff(FDebt)$, $diff(Herfindahl)$, $growth(Credit\ by\ banks)$, $growth(stocks\ trade)$, $PriDef$, $MCap$, $TDebt$, $FDebt$, UU, $diff(UU)$, $growth(GDP)$, $GovDep$
Table 5.6	$HICP$, $diff(HICP)$, $diff(TDebt)$, $diff(FDebt)$, $diff(Herfindahl)$, $growth(Credit\ by\ banks)$, $growth(stocks\ trade)$, $PriDef$, $MCap$, $TDebt$, $FDebt$, UU, $diff(UU)$, $growth(GDP)$, $GovDep$
Table 5.10	$HICP$, $diff(HICP)$, $diff(TDebt)$, $diff(FDebt)$, $diff(Herfindahl)$, $growth(Credit\ by\ banks)$, $growth(stocks\ trade)$, $PriDef$, $MCap$, $TDebt$, $FDebt$, UU, $diff(UU)$, $growth(GDP)$, $GovDep$
Table 5.11	$HICP$, $diff(HICP)$, $diff(TDebt)$, $diff(FDebt)$, $diff(Herfindahl)$, $growth(Credit\ by\ banks)$, $growth(stocks\ trade)$, $PriDef$, $MCap$, $TDebt$, $FDebt$, UU, $diff(UU)$, $growth(GDP)$, $GovDep$
Table 5.12	$lag(MQ)$, $SSBM$, $lag(SSBM)$, $diff(UU)$, $diff(HICP)$, $PriDef$, $diff(PriDef)$, $TDebt$, $FDebt$, $diff(FDebt)$

Table A.10: This table contains for each linear model the list of testregressors used for the extreme bound analysis.

Bibliography

[1] *Wikipedia: Basel III,* https://en.wikipedia.org/wiki/Basel_III. 16th august 2015. 2

[2] *Wikipedia: Ramsey RESET test,* https://en.wikipedia.org/wiki/Ramsey_RESET_test. 16th august 2015. 44

[3] *Wikipedia: Financial Crisis of 2007-08,* https://en.wikipedia.org/wiki/Financial_crisis_of_2007-08, 16th august 2015. 1

[4] European central bank: Guide to banking supervision, https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssmguidebankingsupervision201411.en.pdf, 2014. 2

[5] European central bank: Aggregate report on the comprehensive assessment, https://www.bankingsupervision.europa.eu/ecb/pub/pdf/aggregatereportonthecomprehensiveassessment201410.en.pdf, 26th october 2014. 2

[6] European central bank: Note comprehensive assessment october 2013, https://www.ecb.europa.eu/pub/pdf/other/notecomprehensiveassessment201310en.pdf, 31st october 2013. vii, 2, 21

[7] Badi H. Baltagi. *Econometric Analysis of Panel Data*. John Wiley & Sons Ltd, 3rd edition, 2005. 14, 16

[8] William W. Cooper, Lawrence M. Seiford, and Kaoru Tone. *Data Envelopment Analysis: A Comprehensive Text with Models,Applications, References and DEA-Solver Software*. Springer US, 2nd edition, 2007. 7, 10, 11

[9] Yves Croissant and Giovanni Millo. Panel data econometrics in R: The plm package. *Journal of Statistical Software*, 27(2), 2008. 18, 46

[10] John C. Driscoll and Aart C. Kraay. Spatial correlation in panel data. 1995. 19

[11] John C. Driscoll and Aart C. Kraay. Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics*, 80, 1998. 19

[12] Hsin-Hsiung Fang, Hsuan-Shih Lee, Shiuh-Nan Hwang, and Cheng-Chi Chung. A slack-based measure of super-efficiency in data envelopment analysis: An alternative approach. *Omega*, 41, 2012. 9, 10

[13] Luc Laeven and Fabián Valencia. Systemic banking crises database: An update. *IMF Working Paper*, 2012. 1, 42, 55

[14] Edward E. Leamer. Sensitivity analyses would help. *The American Economic Review*, 75, 1985. vii, 20

[15] M. Hashem Pesaran. General diagnostic tests for cross section dependence in panels. *Cambridge Working Papers in Economics*, 2004. 18, 46

[16] Wolfgang Scherrer, Ronald Scheucher, and Oliver Leodolter. Banking performance in euroland. efficiency and the impact of strategic variables: 2003-2012. 2014. iii

[17] Gunther Tichy. War die finanzkrise vorhersehbar? *Perspektiven der Wirtschaftspolitik*, 11:356–382, 2010. 1

[18] Jeffrey M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, 2005. 13, 15, 17, 18, 46

[19] Jeffrey M. Wooldridge. *Introductory Econometrics: a modern approach*. South-Western Cengage Learning, Ohio, 4 edition, 2009. 17

[20] Achim Zeileis and Torsten Hothorn. Diagnostic checking in regression relationships. *R News*, 2(3):7–10, 2002. 20