VILNIUS UNIVERSITY

LAURYNAS NARUŠEVIČIUS

ECONOMETRIC ASSESSMENT OF BANK STABILITY

Doctoral dissertation
Physical sciences, mathematics (01 P)

Vilnius, 2017

Doctoral dissertation was written in 2012–2016 at Vilnius University.

Scientific supervisor:

prof. habil. dr. Alfredas Račkauskas (Vilnius university, physical sciences, mathematics – 01 P).

VILNIAUS UNIVERSITETAS

LAURYNAS NARUŠEVIČIUS

EKONOMETRINIS BANKŲ STABILUMO VERTINIMAS

Daktaro disertacija
Fiziniai mokslai, matematika (01 P)

Vilnius, 2017 metai

Disertacija rengta 2012–2016 metais Vilniaus universitete.

Mokslinis vadovas:

prof. habil. dr. Alfredas Račkauskas (Vilniaus universitetas, fiziniai mokslai, matematika – 01 P).

Acknowledgements

I would like to convey my sincerest gratitude to my supervisor prof. habil. dr. Alfredas Račkauskas for his precious academic guidance, continuous support and all the help during the years of my PhD studies at Vilnius University. Without his active involvement and guidance it would be hard to achieve the results of this dissertation as they are now. I was very grateful to have such devoted advisor and mentor for my PhD studies.

I extend my sincere gratitude to my colleague Vygantas Butkus. It was fantastic to have the opportunity to work together.

I would like to thank prof. dr. Audronė Jakaitienė and prof. dr. Marijus Radavičius for reading my dissertation and providing very useful comments that helped to improve this work.

Most of all, I would like to express my deepest gratitude to Giedrė Dzemydaitė and my family for their inspiration, encouragement and manifold support.

Laurynas Naruševičius

Contents

Introduction

The stability of the banking sector and the entire financial system has recently been a hot issue both among the central banks and other financial market players, as well as in the public at large. Although financial crises are not new phenomena, the central banks have only taken concern in the stability of the financial system in the past several decades. Financial crises are rare events. However, in most cases they are followed by an economic downturn, therefore the central banks and supervisors should take measures to ensure the stability of the financial system.

A set of tools available for measuring the financial stability includes macroeconomic stress testing which is one of the methods used to assess the resilience of the banking system to various risks that may arise in the near future. The Global financial crisis of 2007-2008 has provided a strong impetus for the development and application of testing methodologies. Stress tests are carried out by commercial banks, supervisory authorities, and central banks, as they seek to measure the resilience of a specific institution or the entire sector against adverse developments in the economy. In this respect, stress testing approaches can be divided into the following two categories: bottom-up and top-down. The bottom-up stress tests are conducted by the commercial banks, based on their own data and models. The central bank or the supervisory authority may impose certain restrictions on the modelling exercise of commercial banks (e.g. to define general scenarios, mandatory assumptions, and methodological principles) and may use their in-house analytical tools to check or challenge the results, obtained by commercial banks. The purpose of bottom-up stress tests is to measure the resilience of a specific bank against economic shocks, which makes them one of the tools of so called micro-prudential supervision. The top-down tests are carried out by the central banks, in most cases, without direct involvement of commercial banks. The rules, scenarios and modelling assumptions used in such stress tests are uniform for all banks subject to this exercise. Top-down testing is often used to provide a benchmark to compare the results of bottom-up stress tests. This approach helps identify important inconsistencies in the results of the tests carried out by commercial banks. Top-down testing is one of the macro-prudential supervisory instruments, since it has as its aim the measurement of resilience against adverse economic shocks in the entire banking system.

The main purpose of stress testing framework proposed, in this thesis, is to quantify the resilience of the entire Lithuanian banking system and its constituent institutions against adverse economic shocks. Solvency stress testing is focused on the assessment of the banks' capital adequacy under an adverse macroeconomic scenario. The exercise has a two-year time horizon and involves the consistent modelling of items in the banks' profit and loss account on a quarterly basis.

In general, the banking sector plays an important part in the economy. As one of the main sources for financing the economic activity, banks may influence business cycles. On the other hand, bank revenues show fluctuations in time as they depend on the overall economic activity. Bank profitability is a prime determinant of bank stability, lending capacity, and is an important part of the stress testing, because profits can offset a large part of incurred credit losses. A stable banking sector may stimulate the economy and is able to withstand economic shocks. Therefore, it is important to understand the relationship between bank revenue and macroeconomic variables as it could help assess the stability of a banking sector.

One of the purposes of this analysis is to examine the relationship between the profitability of the Lithuanian banking sector and its determinants. The knowledge of the relationship is useful for banks and their supervisors who are responsible for maintaining a stable financial sector. In this thesis we adopt the panel error correction model to assess long-term and short-term internal and external determinants of items from bank income statements (net interest income, net fee and commission income, and operating expenses). The pooled mean group (PMG) estimation technique, developed by [124, 125], allows us to impose homogeneity in the long-term coefficients and enables heterogeneity in the short-term coefficients. Therefore, this study contributes to the sparse literature on the Lithuanian banking sector analysis and introduces an estimation technique, which is new in this field of research.

The majority of studies on the relationship between bank profitability and explanatory variables used return on assets or return on equity as the dependent variable. The studies of [67] and others analyzed cross-country data, meanwhile [11], [36], [47] analyze data of single separate countries. The studies have found that bank profitability is determined by bank-specific, industry-specific, and macroeconomic determinants. Variables used in the studies and the effect on profitability differ as data sets vary across the studies. Few papers ([8], [7]) have studied the determinants of separate items of bank revenue and expenses. However, none of these studies include data on the Lithuanian banking sector in their analysis. Therefore, it is important to investigate whether we can find a similar relationship in a transitional economy such as Lithuania.

In the second part of the thesis, the focus of the analysis goes from the Lithuanian banking sector to European banking sector and cluster analysis in particular. In recent years, the cluster analysis, aiming to discover group structures among a set of observations, gains much popularity in the literature. Partitioning of the time series data helps to detect characteristic patterns, to forecast future performance, etc. The methods used in the cluster analysis can be divided into three categories: the methods based on (1) similarity of raw data; (2) features extracted from raw data, and (3) models built from raw data. Recall that a measure D of dissimilarity (or equivalent similarity) of objects X and Y is symmetric: $D(X, Y) = D(Y, X)$, non-negative $D(X, Y) \geqslant 0$, and such that $D(X, X) = 0$. The similarity measure can be, but not necessarily is, a metric, i.e. $D(X, Z) \leqslant D(X, Y) + D(Y, Z)$. One of our goals of this thesis is to consider various dissimilarity measures and apply them to the data under investigation.

As it has been mentioned before, after the Global financial crisis in 2007-2008, the financial sector and especially banks gained much attention.

Authorities in whole Europe began to use actively macroprudential tools focused on the banking system as a whole. A large number of the macro-level instruments were introduced and are applied to all banks. However, the banking sector is heterogeneous and some tools could be ineffective to some banks. It would be useful to find groups of banks that have similar characteristics and design or calibrate some macroprudential instruments that would become appropriate to that group. Therefore, one of the goals of this thesis is to discuss the clustering of banks.

In this work, distance measures, based on time series as well as on functional data properties, are exploited. In addition to the univariate clustering, where banks are grouped into clusters according to one bank-specific ratio, a multivariate clustering is applied, where banks are clustered, based on their several ratios. Since in the cluster analysis data are unlabelled, a related issue is to find an appropriate number of clusters that are most proper for the data. The resulting clusters should not only have good statistical properties, but also give results that are, in our case, economically explainable.

The next part of the thesis consists of the analysis of functional data methods that would be useful for estimating and forecasting a particular cluster. The question what is better, to forecastin the aggregate quantity directly and then disaggregate, or to forecast the individual components directly and then aggregate them to form the forecast of the total, is important in many applications. This is also known as the top-down versus bottom-up forecasting problem (see, e.g., [112]). However, in any specific practical application usually it is difficult to argue on theoretical grounds what the correct approach should be. Therefore, this question usually is settled empirically by trying both approaches.

In this work, we consider the estimation and forecasting problems of the banking data. We analyze the capital adequacy ratio which determines the capacity of the bank to meet potential losses arising from credit risk, market risk, operational risk, and others. This ratio ensures that the banks do not expand their business without having adequate capital. We make a cluster analysis and divide our sample into a few clusters. We are interested in the future performance of the capital adequacy ratio of the cluster as well as of the each bank individually.

Furthermore, the advantages provided by the functional data analysis methods are explored. Assuming that for each $j = 1, \dots, N$ we have a random function $X_j = (X_j(t), 0 \leq t \leq T)$, we investigate forecasting of the aggregated process

$$\overline{X}_N = \frac{1}{N} \sum_{j=1}^{N} X_j$$

by comparing two approaches. The first is by fitting a model to $\overline{X}_N$ directly and then forecasting $\overline{X}_N(T+h)$ for a time horizon h. By the second approach we fit models to each component X_j and aggregate the forecasts of $X_j(T+h)$, thus obtaining another result for prediction of $\overline{X}_N(T+h)$. In this case, we use the feature of the functional data, which allows us to divide the function into as many data points as it is needed.

At the end of in this thesis, a novel functional regression model and its estimation method are presented. This model can be used to estimate the functional relationship between one realization of the stochastic process and other functional covariates. This model is also used to obtain forecasted values of the mean process $\overline{X}_N(T+h)$.

The remainder of this dissertation is structured as follows. Section 1 describes a top-down stress testing framework for the Lithuanian banking sector and the provides analysis of profitability. Section 2 gives time series and functional data clustering results and introduces the top-down versus bottom-up forecasting problem. Finally, the main findings are summarized.

Objective and tasks

The thesis consists of two main parts, therefore, there are two main objectives, which were achieved. The first one, is to analyze the Lithuanian banking sector and particularly banks that are operating in it and to develop a new framework, which could help to assess their resilience to an adverse economic development. The second main goal is to discuss the clustering of a given set of the European banks into groups, based on their performance, and to compare similarity/dissimilarity measures, based on the time series properties or on the functional data properties. Furthermore, we aimed to consider the top-down versus bottom-up forecasting problem of the mean process of the cluster.

In order to achieve that the following aims were raised:

1. To develop a framework that could be regularly used for the assessment of the resilience of the Lithuanian banking sector.

2. To build an econometric models that establish links between banks' credit losses and the dynamics of macroeconomic variables.

3. To contribute to the literature on the Lithuanian banking sector analysis and to adapt a panel error correction model to assess long-term and short-term internal and external determinants of banks' revenue and expenses.

4. To introduce a new similarity/dissimilarity measure and to compare with some existing measures, based on the time series or on functional data properties.

5. To extend some dissimilarity measures from a univariate case to the multivariate case.

6. To propose theoretical functional data models and to apply them to banking functional data.

Scientific novelty

A new top-down stress testing framework to assess the resilience of the Lithuanian banking system was introduced. This framework is following international recommendations and best practices and is able to quantify how banks are prepared to withstand potential adverse economic developments. The stress testing framework involves the application of econometric models that help to establish links between the dynamics of macroeconomic variables and the developments in banks' credit losses and profitability. Furthermore, as bank profitability is an important determinant of bank stability and lending capacity, it is necessary to understand what might influence banks' revenue and expenses. A panel error correction model was employed to assess long-term and short-term internal and external determinants of items from the bank income statement.

In the second part of the work, the clustering of a given set of the European banks was discussed and the similarity/dissimilarity between banks was estimated, using measures based on time series or functional data properties. When making the cluster analysis, two dissimilarity measures, not commonly used in the literature, were proposed and two measures were extended from the univariate to the multivariate cases. In the last section, the advantages of the functional data analysis methods were explored and applied to the extracted clusters. The theoretical functional data models were proposed and afterwards used for the top-down and bottom-up forecasting problem.

Statements presented for defence

1. It is shown that the new framework for the top-down stress testing of the Lithuanian banking system is useful to assess the resiliance of the entire banking system and its constituent institutions against adverse economic shocks. Proposed stress testing procedure follows international recommendations and best practices.

2. It is revealed that panel error correction model can be used to assess long-term and short-term internal and external determinants of the Lithuanian banks' revenue and expenses.

3. It is shown that functional data analysis methods can be used for cluster analysis of the banking data. Introduced new dissimilarity measures, based on the functional data properties, perform better than the measures, based on the time series properties.

4. It is demonstrated that the application of theoretical models, proposed in the thesis, provides equally good top-down and bottom-up forecasting results in the case of banking data.

Publications by the Author

The results of the doctoral research will appear in 5 research papers. Four of them have already been published, one is submitted for publishing in general or specialized periodical peer reviewed journals.

Published articles:

1. Naruševičius L. (2013). *Modelling banks' profitability using dynamic panel data estimation.* Social technologies, Vol. 3(2), p. 278-287. ISSN 2029-7564 (Online).

2. Butkus V. and L. Naruševičius (2015). *Lietuvos bankų sistemos makroekonominis testavimas nepalankiausiomis sąlygomis.* Pinigų studijos, No. 1, p. 74-93. ISSN 1648-8970 (Online).

3. Naruševičius L. and A. Račkauskas (2016). *Comparing dissimilarity measures: a case of banking ratios.* Informatica, Vol. 27(3), p. 649-672. ISSN 1822-8844 (Online), (ISI WoS).

4. Naruševičius L. (2017). *Bank profitability and macroeconomy: evidence from Lithuania.* Technological and Economic development of Economy. ISSN 2029-4921 (Online), (ISI WoS).

Submitted for publication

1. Naruševičius L. and A. Račkauskas (2017). *Forecasting with functional data: case study.*

The results of the thesis were presented in the following conferences:

1. Naruševičius L. *Modelling banks' profitability using dynamic panel data estimation.* International Conference Social Technologies'13. October 10-11, 2013, Vilnius, Lithuania.

2. Naruševičius L. and A. Račkauskas. *Comparing dissimilarity measures: a case of banking ratios.* X$^{\text{th}}$ Tartu Conference on Multivariate Statistics. 27 June - 1 July, 2016, Tartu, Estonia.

3. A. Račkauskas and L. Naruševičius. *Forecasting with functional data: a case study*. 17$^{\text{th}}$ Applied Stochastic Models and Data Analysis International Conference, 6-9 June, 2017, London, United Kingdom.

Attended conferences, which are related to the topic of the thesis and were useful to have a better understanding of the issues in this field of research:

1. *3rd Annual European Stress testing Forum for Banks*, September 13-14, 2014, Amsterdam, The Netherlands.
2. *Conference on Financial Stability and Stress Testing*, June 3-5, 2015, Basel, Switzerland.

Chapter 1

Stress testing and bank profitability

1.1 Review of the macroeconomic stress testing

Until the Global financial crisis of 2007-2008 the interest in stress testing was mainly restricted to practitioners, i.e. financial supervisors, central bankers, and risk managers. It has been pointed out that the severity of the crisis has been largely due to its unexpected nature and that a more extensive and rigorous use of stress testing methodologies would have probably helped to alleviate the intensity and repercussions of the turmoil. Therefore, the stress testing has become the key topic in policy discussions and a regular subject for news media. In general, stress tests are quantitative tools, used by banking supervisors and central banks for assessing the soundness of banking systems in the event of extreme, but still plausible, shocks (macroeconomic stress tests). They are also an important management instrument for banks, since they provide financial institutions with useful indications on the reliability of the internal systems, designed for the measurement of risks (microeconomic or prudential stress tests) [131]. The system-wide nature of macroeconomic stress tests also reflects the use of a macroeconomic adverse scenario, which can cover several risk factors, unlike a sensitivity analysis where the health of a bank or of the financial system is checked against specific risk factors and in isolation from the other parts of the financial system.

The International Monetary Fund (IMF) in cooperation with World Bank in 1999 launched the regular use of stress testing as part of its Financial Sector Assessment Programs (FSAP), which encouraged the development of comprehensive stress testing frameworks for assessing the resilience of financial systems to adverse disturbances. Since then, the use of stress tests to address systemic risk has deepened, following the recent financial crisis in the US and Europe. Ouro and Schumacher (2012) [116] analyzed different experiences and proposed seven "best practice" principles for stress testing. The principles are:

- Define appropriately the institutional perimeter for the tests.
- Identify all relevant channels of risk propagation.
- Include all material risks and buffers.
- Make use of the investors' viewpoint in the design of stress tests.
- Focus on tail risks.
- When communicating stress test results, speak smarter, not just louder.
- Beware of the "black swan".

The proposed principles emphasize that the success of stress tests cannot be reduced to the choice of a few parameters, but should be seen in a broader context. The survey of central banks and supervisory authorities in 23 countries and stress tests in FSAPs shows that, despite major improvements since the crisis, practices still fall short of these principles.

Next to the prime responsibility for monetary policy, the responsibility for helping to safeguard financial stability features in the mandate of the central banks after the Global financial crisis. This task requires the systematic review of possible sources of risk to the financial systems that are of a potential systemic nature and the assessment of their potential magnitude. Macro-prudential policies and tools aim at limiting the systemic risk or instances of widespread instability in the financial system. This is opposed to micro-prudential oversight, which focuses on banks individually to ensure their soundness as single entities. Macroeconomic stress testing models, that can be employed to assess the impact on the financial sector of the materialisation of identified risks, have become the workhorse of analytical tools for macro-prudential risk assessments and are the backbone of central banks' systemic risk assessment tools. The main goal of macroeconomic stress tests is to identify structural vulnerabilities in the financial (or banking) system and to assess its resilience to shocks. In this respect, aggregate stress tests can usefully enrich the financial stability toolbox, mostly because they provide forward looking information on the impact of possible extreme events. Furthermore, this kind of simulation allows consideration of the interconnections across economic sectors, capturing major risk sources for intermediaries, disentangling interactions across different risks.

Macroeconomic stress tests, used to support macro-prudential oversight, are usually performed in a centralised fashion, i.e. they are called top-down stress tests. Such top-down exercises are different from the supervisory stress tests conducted for micro-prudential oversight, which assess individual banks' ability to withstand shocks, typically using tailor-made scenarios or sensitivity analysis. The tests are conducted under supervisory guidance by the supervised entities and are called bottom-up stress tests. In general, the top-down stress tests have consistent assumptions and applied econometric methods which allows the comparability of results. However, data availability is lower for authorities, as banks have a more detailed internal portfolio information than they disclose. On the other hand, a bottom-up stress tests have different assumptions and methodologies among banks which raise considerable problems for the comparison and aggregation of the results. The middle of those two approaches is represented by coordinated exercises, whereby the same baseline and adverse scenarios are given to all participating banks along with a strict methodological guidance, such as in the case of exercises, conducted by the European Banking Authority (EBA) [54]. Jobst et al. (2013) [91] provided a short classification of solvency stress testing and some examples of its application (see Figure 1.1).

In spite of the role stress tests have in macro-prudential analysis by

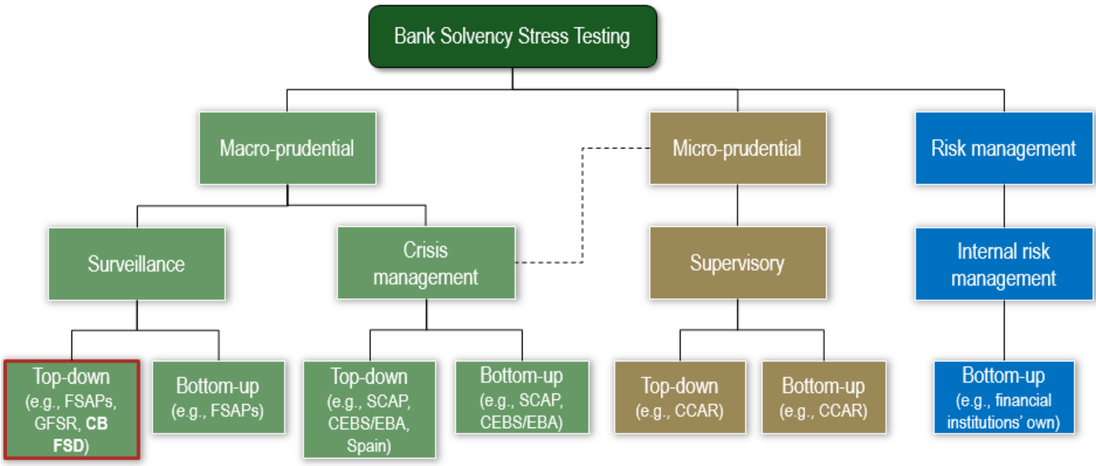

Figure 1.1: Solvency stress testing applications

Source: Jobst et al. (2013) [91]

Notes: Top-down stress tests are either conducted using the data of individual banks and then aggregated or on an aggregated portfolio; bottom-up stress tests are conducted by individual institutions using their own internal risk models and data; FSAP – Financial stability assessment program (IMF); GFSR – Global financial stability report (IMF); CB FSD – Central bank Financial stability departments; SCAP - Supervisory Capital Assessment Program (USA 2009); CEBS/EBA – Committe of European Banking Supervisors/European Banking Authority 2010/2011 EU-wide Stress Testing Exercise; CCAR - Comprehensive Capital Analysis and Review (USA); EBA 2014 - European Banking Authority EU-wide stress testing 2014.

central banks, it must also be acknowledged that macro stress tests have important limitations. As emphasised by Borio et al. (2012) [22] macroeconomic stress tests are not appropriate early warning indicators. While macroeconomic stress tests were an effective crisis management and resolution tool in the recent Global financial crisis, the authors also criticise stress tests for missing the build-up of risks on banks' balance sheets in the run-up to the current crisis. Moreover, most stress testing models have difficulty to capture the typically non-linear nature of systemic risks or macro feedback loops, and they fail to adequately reflect counterparty and liquidity risks. It is worth noting that the relevance and accuracy of any stress testing exercise relies on the underlying data input.

As pointed out by Jones et al. (2004) [92], the macroeconomic stress testing is a complex multi-step process that can be seen as the interaction of different competencies: it is part investigative, part diagnostic, part numerical, and part interpretative. The level of detail that top-down stress tests may pursue critically depends on data availability by national authorities. Reliance on detailed data allows the use of the more sophisticated modelling approaches.

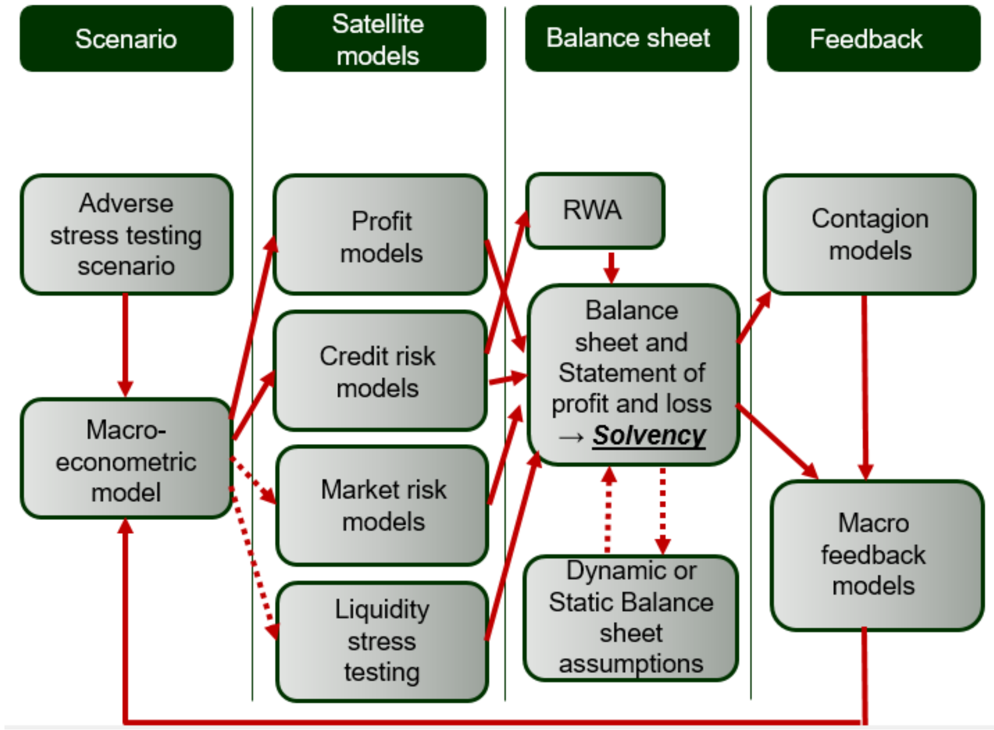

Figure 1.2: The four component structure of the solvency analysis framework

Source: formed by authors based on Henry and Kok (2013)[79].

To sum up, a marcoeconomic top-down stress testing or a forward looking bank solvency analysis consists of a number of different but interconnected analytical steps. This approach is revelant especially when the analysis is made using, individual bank level information and can be described as a modular system with four main component structure (see Fig. 1.2). The first element is the scenario design in which a macroeconomic or financial shock scenario is designed and calibrated. The second element consists of econometric models, or so-called satellite models, which translates the scenario into variables affecting the evolution of bank balance sheet components (credit risk models, market risk models) and bank's loss absorbtion capacity (profit models, liquidity models). The third element (balance sheet part) takes the projected profit and losses, derived from the satellite models, to individual bank balance sheets with the purpose of calculating the resulting impact on bank's solvency positions. Usually the solvency of a bank is evaluated by the capital adequacy ratio. The last element (feedback module) tries to assess what might be the derived second round effects on the initial bank solvency in terms of contagion within the financial system and in terms of feedback effects on the real economy [79].

Scenario

The process of macroeconomic stress testing begins with the identification of the potential macroeconomic or financial shock that could impact the resilience of the banking system. A potencial risk depends on the characteristics of the banking system, on their business models, on the features of financial regulation, and on the overall macroeconomic environment. As an example, for banks that are mainly active in the domestic loan markets, the analysis should focus on th ecredit risk and factors, such as interest rates, unemployment, real estate prices, etc., that may have a negative impact on the business. For large internationally active banks, the global risk factors, such as oil price and other raw material prices, exchange rates, etc., are much more important.

Once the potential risks have been specified, it is important to investigate the events that trigger the shock and determine the level above which the magnitude of the shock leads to the materealisation of the risk, i.e. to a stress scenario. In other words, the stress scenario should ensure that the level of severity is appropriate, i.e. has a sufficiently strong impact on the banks. On the other hand, it still must be possible, it should reflect a material risk. When the shock is severe, but the calculated losses are small, the impact on the banks will be limited, thus requiring a review of the risk assessment. Therefore, the implementation of stress tests ussually is an iterative process, since some originally identified shocks may lead to relatively small impacts, while some risks originally assessed as small may lead to large impacts, if there are substantial exposures [38].

The choice of extreme but plausible events is frequently based on a discretionary assessment of the analyst. However, the shock of the stress scenario can be calibrated in a number of ways:

1. Ad-hoc calibration without referring to any model or historical distribution of the risk factor. Instead, the shock size calibration could take historical movements of the revelant economic variables observed during the past crisis episodes. Though historical scenarios are easier to implement, hypothetical scenarios may be the only available option when structural breaks in the financial system make the past history no longer informative.

2. Shock size calibration, based on historical distributions. This is a hybrid solution, where hypothetical scenarios are based on historical distributions, but they are not necessarily linked to specific events.

3. Shock size calibration, based on shock distributions, where shocks are generated from the dynamic model. Some dynamic model produces the fit and the resulting residuals, which are interpreted as shocks. Those shocks can be calibrated using the size and distribution of the corresponding model residuals.

In practice, the second approach is usually preferred in stress testing exercises. The main reason why this approach, which does not rely on the pre-defined model specification (i.e. it is non-parametric), is applied is that stress scenarios often require shocks to many enocomic or financial variables that are strongly interrelated. Such large-scale multivariate distributions are difficult to treat analytically (i.e. parametrically), so a non-parametric approach is preferred.

Once the revelant potential risk factors are identified and shocks affecting such factors are properly calibrated, they are used as input into the relevant dynamic macro-econometric model. Structural macro-econometric models are the most appropriate tools for understanding how the economic system behaves when the assumed shock materializes. These models are typically developed for forecasting the evolution of key macroeconomic variables, providing a coherent development of economic and financial sectors. The model uses the inputs - calibrated shocks - and returns the values of the macroeconomic variables under the stress scenario. Other types of models, such as VAR ([23]) or global VAR (GVAR) ([127], [29]), could also be used to produce a coherent development of the macroeconomic variables.

However, the presence of a macro-econometric model does not reduce the need for expert judgement. Indeed, the design of the adverse scenario entails a series of decisions that are crucial for the validity of the stress testing exercise and reliability of the results. On the one hand, most macroeconomic models are valid tools for forecasting the dynamics of the economy in normal times, whereas the hypothesis of a linear relationship across macroeconomic and financial variables might be unlikely to be valid in extreme events, when non-linearities can be substantial. On the other hand, the response of market participants to extreme shocks may be difficult to model and forecast, also due to the lack of relevant data.

Satellite models

When the adverse macroeconomic stress testing scenario is prepared, the next step is to translate that scenario into an impact on bank profitability (i.e. loss bearing capacity) and various forms of risks held by banks on their balance sheet, for example, credit risk, interest rate risk, market risk. The adverse scenario in the top-down stress testing is transformd via the so-called satellite model, which is an econometric equation, or a set of equations that relates macroeconomic variables with the bank-specific variables.

For most banks losses, related to the credit portfolio resulting from borrowers' failure, are the major risk component (i.e. credit risk) with a potential to have a significant impact on bank's assets and ultimately on the capital adequacy ratio. For this reason, the modelling and projection of credit risk is the key element in the overall analytical framework, used for conducting a forward-looking solvency assessment. The relationship between aggregate credit risk parameters and macroeconomic variables has been widely analyzed in the literature. Depending on data availability there are several indicators which can reflect credit risk:

- Probability of Default (PD);
- Loss Given Default (LGD);
- Loss Rates, i.e. product of PD and LGD;
- Default rate, i.e. the number of defaulting loans to total outstanding loans;
- Non-performing loans (NPL);
- Loan Loss Reserves (LLR) or Loan Loss Provisions;
- Write-off rates.

NPL and LLR can be expressed as a ratio to gross loans, LLR can also be expressed as a ratio to the outstanding amount of NPL, i.e. coverage ratio. The different measures of credit risk have overlapping definitions, but can be considered to vary in terms of their time perspertive. While PD, that measures the probability of borrowers default x-days ahead, is the most forward looking metric, the write-off rate which shows the point in time when non-performing loans are written off, is the least forward looking metric of the credit risk.

In short, the satellite model for a credit risk estimates a functional relation between macroeconomic variables and credit risk indicators. The choice of the econometric model depends on the data availability, time span of the sample or preference of a modeller. Various econometric techniques are applied in practice. For example, it can be simple OLS regression, static or dynamic panel regression ([98], [24]), ARIMAX equations ([63]), VAR model ([108]) or VECM model ([10]) or some other ecomonetric model. Jiménez and Mencía (2007) [90] argue that micro-contagion effects between economic sectors create an additional channel of default correlation and allow sectoral default rates to depend on macroeconomic variables as well as on latent factors that can capture contagion factors. Foglia (2009) [57] gives a short overview of the methods used by central banks to test the credit risk.

Besides the credit risk, the market risk is the second most important risk factor. Losses from the market risk can come from a financial market downturn that affects a large set of market variables, such as interest rates, exchange rate, equity and commodity prices, sovereign bond yields, and volatilities. The adverse scenario is usually assumed to be instantaneous, one-off shock that has severe impact on bank results. The market risk adverse scenario can be set ad-hoc, on the basis of the historical information (some Value at Risk (VaR) model or ARCH, GARCH model could be used), or depend on the macroeconomic scenario assumptions for the evolution of some variables, such as stock prices, exchange rates or interest rates.

Incorporation of the market risk into the general stress testing framework faces a few challenges. Granular trading book portfolio information is changing constantly and is generally not available. The second problem is the different time duration of the shock. While the credit risk stress testing is developed from 2 to 5 year horizon, the market risk stress is assumed only for one month. To combine different time dimensions is not a straightforward task, therefore, the market risk in the trading book is often not assessed or incorporated with a simplified view.

Another risk type which is even less incorporated into the macroeconomic stress test is liquidity risk. The liquidity risk can emerge due to endogenous behavioural response by banks or other financial market participants. Banks can face two types of liquidity riks: funding liquidity risk and market liquidity risk. In general, banks become illiquid before they are insolvent, thus there are few attempts to incorporate the liquidity risk into the macroeconomic stress test [130]. However, there are a few aspects that make it difficult to achieve empirical progress in this field. First of all, in order to measure the liquidity risk, assets and liabilities together with their maturity information have to be considered. And this fact extends the universe of required data considerably. These data are usually confidential and are changing continuously, especially during the stress period. Second, data on behavioural responses by depositors and other banks in the interbank market are essentially not available. Therefore, liquidity stress tests are based on the rules of thumb rather than on empirical relationships. Moreover, the link between shocks and solvency, modelled during the stress test, and liquidity is even less clear.

Even in the stress period banks continue to generate revenue, that is used as a primary buffer against losses originating from the credit or market risk. Although the first stress testing frameworks usually focused only on credit losses, later some kind of profit models were also incorporated. The main idea is the same as in the credit risk modelling, - translate the macroeconomic stress scenario into banks' revenue. In general, three main components of the profit before credit losses are modelled, i.e. net interest income, non-interest income, and operating expenses. Projections of these items can be made by experts or using simple OLS regresssion ([63]), with an autoregressive distributed lag model (ARDL) ([131]) and with a static or dynamic panel data model ([25]) or some other econometric technique.

Balance sheet

In this module of stress testing, the results of the satellite models are collected once the adverse scenario was used as an input. Forecasts from the satellite models are translated into bank's revenues, expenses, and losses. Before calculating the bank's capital adequacy ratio, i.e. solvency position, an assumption on the balance sheet and risk-weighted assets (RWA) dynamics must be made.

A top-down macroeconomic stress testing can be based either on static or balance sheet assumptions. A static balance sheet assumption means that all items from the bank's balance sheet are kept at the same level over a stress test horizon as it is the last known observation. Under this assumption, banks do not strategically react to shocks by adjusting their business strategy or taking management actions. Although, the static balance sheet assumption might be less realistic, because bank's balance sheet is never static, but it helps to compare different banks and their riskiness at the current level. Under the dynamic balance sheet assumption, banks can adjust their balance sheet either exogenously given or endogenously optimising balance sheet structure. Using the exogenously given dynamic approach, various scenarios of changes in balance sheet structure reflecting anticipated changes in market demand for bank products, funding conditions and bank's reaction to the economic cycle, can be applied. The dynamic balance sheet approach with endogenous balance sheet developments is based on the bank's optimising behaviour. Usually, it is assumed that a bank optimally restructures its assets, following a risk-adjusted return maximisation programme. It means that the bank maximizes return on equity, adjusted by the covariance of risks in its balance sheet.

The calculation of risk-weighted assets supplements, the projected profits or losses of a bank with a conditional forecast of the future capital requirement, is at the end of the stress scenario horizon. RWA dynamics can either be static, i.e. kept at the same level, or it can be adjusted, based on losses. Additionally, if there are projected probability of default (PD) ratios, the Advanced Internal Rating-based (IRB) formula of Basel II can be used. The IRB formulae imply that RWA equals $K \cdot 12.5 \cdot EAD$, where

$$K = \left[LGD \cdot N \left(\sqrt{\frac{1}{1-R}} \cdot G(PD) + \sqrt{\frac{R}{1-R}} \cdot G(0.999) \right) - (LGD \cdot PD) \right] \cdot \\ \cdot \frac{1+(M-2.5)b}{1-1.5b} \tag{1.1}$$

and $N(\cdot)$ denotes a cumulative distribution function, $G(\cdot)$ denotes the inverse of the cumulative distribution function, R is correlation, M is an effective maturity, and b denotes the maturity adjustment. Usually, it is assumed that LGD is constant and the adjustment of the risk-weighted assets takes place through changes in PD.

The capital adequacy ratio (CAR) calculation, which shows bank's solvency position, is the main objective of the top-down macroeconomic stress testing exercise. In most cases, the CAR calculation at the end of testing horizon is performed as the sum of the existing capital stock and profit or loss, accumulated over the stress period in relation to the end-horizon risk-weighted assets. Once the solvency position under the given adverse scenario has been calculated, a useful metric, by which to assess the capital adequacy of a bank under the stressed conditions, is the capital shortfall, given the minimum threshold for the solvency ratio. This benchmark determines the potential need for recapitalisation.

Feedback

Usually a macroeconomic stress testing exercise ends up, once the first-round effect on the stressed banks' capital adequacy ratio is derived. However, in the real world, it is expected that banks would react to adverse situations by adjusting their assets and liabilities in certain ways, which in turn could have implications on the real economy and on other banks in the country. The second round effects could be splitted into two parts: contagion effects to the other institutions in the financial system and feedback effects to the real economy.

A decrease of the capital adequacy ratio of some banks, i.e. deterioration of the solvency situation, under the adverse scenario could give rise to the negative contagion effect on other banks in the system. The contagion effect could spread via direct bilateral linkages or indirectly via confidence effects. For example, the default of one bank, that is active in the interbank market, may lead to the failure of other banks. The global financial crisis in 2007-2008 and a subsequent euro area sovereign debt crisis offers good examples of such phenomena. The macroeconomic stress test gives the number of banks which could not be able to comply with the minimum solvency requirement under the stressed conditions. In order to fill the resulting capital gap, a stressed bank could in turn be not able to repay its creditors (other banks) in the interbank market, in a such way triggering losses at other banks through direct bilateral exposures. Those losses, if large enough, may cause insolvency of interbank creditors, which in turn may not be able to fulfill their own obligations on time, triggering a cascade of defaults in the interbank market. A direct channel (bilateral exposures) is only one form of transmission that the financial contagion can take. Other forms include contagion through protection selling and buying ([77]), through overlapping portfolios, i.e. common exposures ([26]), or through indirect channels, such as information contagion, correlation, behavioural commonalities ([3]).

The research, related to the financial contagion, can be divided into four main groups: static and dynamic statistical models and models of static and dynamic network of flows. Static statistical models are based on the graph theory ([133]), percolation ([82]) or random matrices theory, whereby the contagion potential is identified with a view to determine the topological properties of networks. The second group in the literature uses time series models and is based on market data, i.e. stock prices, credit default swap spreads, interest rates, etc. Some previous studies tried to capture contagion, using event studies to detect impact of bank failures on stock prices of other banks in the system. Polson and Scott (2011) [129] used an explosive volatility model to capture stock market contagion measured by excess cross-sectional correlations. Some other authors have tried to capture the conditional spillover probabilities at the tail of the distribution by using quantile regressions [147]. The third group of models investigates the flow of payments in the system. Cascade models ([59]) analyze sequences of defaults, typically using the interbank clearing payments approach, which foresees the equilibrium (instantaneous) resolution of payments. Some models try to explain the behavioural foundation of linkages through the game theoretical formation of networks [4]. The fourth group in the literature is related to models of flows in dynamic networks. The main differentiating feature of this group is related to the direct modelling of the evolution of financial institutions' balance sheets, taking into account some important behavioural aspects of banking systems [64].

Another type of the second round effects is impact of the stressed bank on the macroeconomic development, which subsequently amplifies the initial shock. A model with well-developed real-financial linkages is needed to capture this kind of feedback. Dynamic stochastic general equilibrium (DSGE) models are one of the tools that can be applied in order to capture the feedback effects between financial variables and macroeconomy ([37], [40]). The linkages between the real sector and the financial sector could also be assessed by reduced-form large-scale macro-econometric models with a tight theoretical structure ([46]). The third class of models employs vector autoregessive models. For example, Gray et al. (2013) [71] developed the Contingent claim analysis global vector autoregressive (CCA-GVAR) model which includes interactions between the macroeconomy, corporate sector, sovereigns, and the banking sector. Meanwhile, Giannone et al. (2012) [65] used large-scale Bayesian VAR (BVAR) models. BVAR models can be seen as an alternative to the GVAR model approach, with the latter compressing the parameter space to overcome the curse of dimensionality by operating with weights.

All second-round effects might be important for stress testing. Neglecting these kinds of reactions may determine a loss of information for the analyst and cause the interpretation of outcomes of the exercise less comprehensive.

Final thoughts from the stress testing literature

Jones et al. (2004) [92] have pointed out that stress testing is not a precise tool that can be used with a scientific accuracy, it is rather an art, which requires quantitative techniques, human judgment and a series of discretionary assumptions. Meanwhile, Drehmann (2008) [50] argues that different objectives of the stress testing can lead to different and sometimes conflicting model requirements. The model accuracy, forecast performance, transparency, the suitability for storytelling and other priorities cannot always be achieved within the same model. Understanding these trade-offs for different model specification is not easy, even though it is essential when building stress testing models. Drehmann (2008) also concludes that two main challenges for stress testing models are data limitations and the endogeneity of risk. Data limitations imply that stress testing models are not always econometrically robust. And finally, the endogeneity of risk is the main issue for standard stress testing models since it challenges the fundamentals of model set-ups, which assume a chain from an exogenous shocks via the data generating process to the impact on banks' balance sheets.

1.2 Stress testing framework for the Lithuanian banking sector

1.2.1 Modelling framework

The purpose of stress testing is to assess whether the capital buffers, built up by the banks, are sufficient to absorb the impairment losses on the loan portfolio, which would arise due to adverse developments in the macroeconomic environment. The macroeconomic stress testing procedure for the Lithuanian banking sector, proposed in this work, consists of the following three main steps (for their components and interrelations see Fig 1.3):

- Step 1 involves the construction of an adverse macroeconomic scenario, that is then used for the assessment of resilience of the banking system. The scenario is developed using the structural macroeconomic model for the Lithuanian economy, the statistical features of official macroeconomic indicators as well as expert judgement;
- Step 2 involves the application of econometric models, that help establish links between the dynamics of macroeconomic variables and the developments in a bank's credit risk and profitability. These models are of two types: the models of credit losses and the models of profitability (the latter are used to model the items of the next profit and loss account);
- Finally, Step 3 involves the aggregation of modelling results, obtained in different building blocks, into a single profit and loss account and the simultaneous assessment of changes in capital and risk-weighted assets. These variables define the target variable of the exercise, i.e. the capital adequacy ratio, used to draw the main conclusions about the resilience of the bank. Other indicators, modelled as part of the exercise, can also provide additional insights about the characteristic aspects of the bank's operations.

Data quality and affordability is a very important component during the development of stress testing procedure. This is one of the reasons why different countries have different methods for their testing methodologies and quite different econometric models used. Macroeconomic indicators are available to everybody, but financial or credit risk-related indicators (PD, LGD) are usually unknown or data lines are short. The proposed macroeconomic stress testing procedure uses multiple data sources, first of all, sampled quarterly macroeconomic indicators, published by Statistics Lithuania. A specific list of variables may depend on the macroeconomic model and

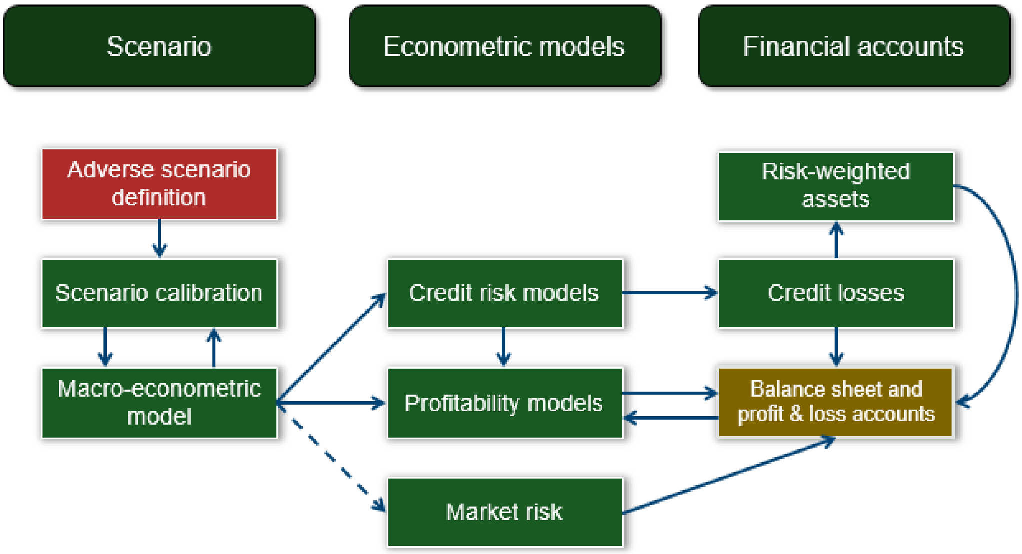

Figure 1.3: Lithuanian stress testing framework

Source: formed by authors.

econometric equations, but the most important variables are GDP and its components, unemployment rate, and inflation. All these data are public and are often referred to in various economic surveys discussing the country's economic situation. They are used to create macroeconomic scenarios and to estimate econometric models. Another source of data is provided by commercial banks profit (loss) and balance sheet accounts. These data were used in the compilation of econometric models for linking macroeconomic indicators to the banks' credit risk and profitability change. The banks provided detailed reports that include a number of key components, such as credit losses, net interest income, net fee and commission income, operating expenses, etc. (all components are indicated in Table 1.1, section 1.2.6). The third source of data is a query for information to commercial banks for the data, which are additionally needed in the solvency stress test. Following the request, the information on the loan portfolio quality indicators (provisions and non-performing loans), calculated basing on sectors of economic activity, wes provided. These data formed the basis of variables in the econometric assessment of credit losses.

The current stress test methodology is based on each bank's data. This allows a more accurate assessment of riskiness of the bank in the case of the adverse economic development as the bank's loan portfolio structure is taken into account. However, the test results are shown collectively in change of bank performance and the entire banking system's capital adequacy ratio.

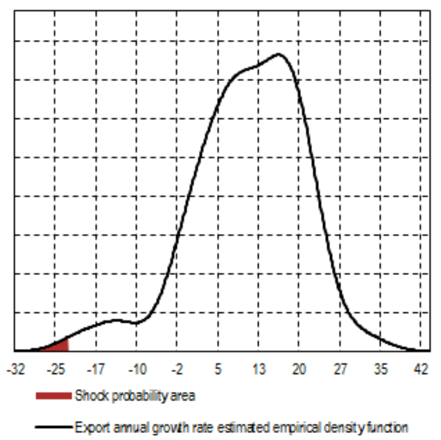

Figure 1.4: Lithuanian real export growth (empirical density function)

Source: Statistics Lithuania and authors' calculation.

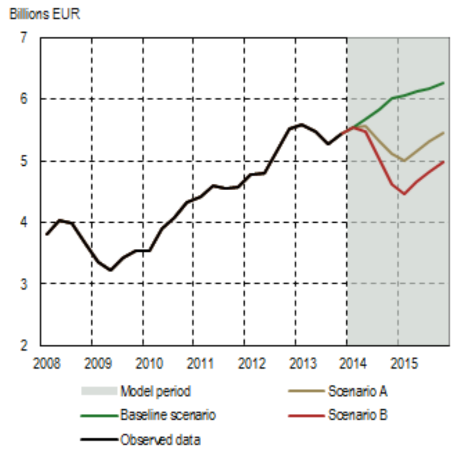

Figure 1.5: Lithuanian real export development under hypothetical stress testing scenarios

Source: Statistics Lithuania and authors' calculation.

The first phase of the banking sector stability assessment, as mentioned above, is creation of a macroeconomic scenario. It consists of two major steps: identification of the actual risk to the economy and assessment of the potential effect on the overall economy. International pratice is not made up of the standard macroeconomic scenario that could be used for stress testing, but there are some existing basic principles of their formation (e.g. see [15], [63], [79]). The most important thing is to draw up a certainly exceptional adverse scenario that would include a significant slowdown in the economic activity. However, the scale of the expected economic shock should depend on the country's economic cycle and the objectives of the stress test procedure. The composed scenario should be relevant to the present economy. On the other hand, the scenario cannot be too extreme, because the larger the shock, the less likely its confirmation is, and to apply a scenario that is completely unable to materialize is not the objective. So, the conclusion of the scenario building is that we need to find a balance between the size of the shock and its plausibility.

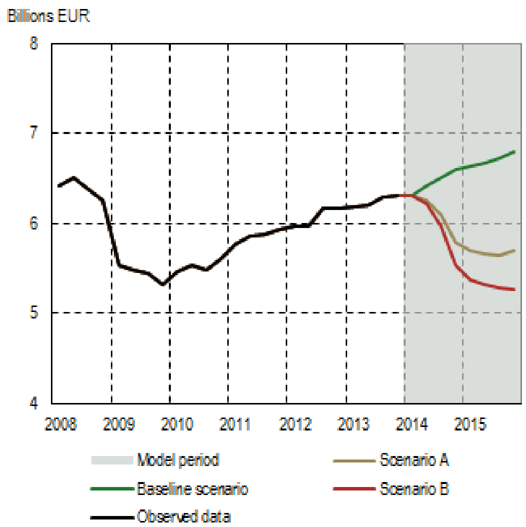

Figure 1.6: Lithuanian real GDP development under hypothetical stress testing scenarios

Source: Statistics Lithuania and authors' calculation.

During the stress test, a series of macroeconomic scenarios are analyzed. In particular, using official macroeconomic forecasts prepared by Bank of Lithuania, the baseline development of the banking system is assessed. The results of this scenario are used to test other scenarios and to compare the results and sustainability of the banking business in the most likely economic development. Later on, other adverse economic scenarios are considered. For example, a historical density function of annual Lithuanian export growth is analyzed (see Figure 1.4). Then the $mean - 3\sigma$ (Scenario A) or $mean - 4\sigma$ (Scenario B) is taken as the initial export shock and future development of Lithuanian export is evaluated (Figure 1.5). Afterwards, using the macro-econometric model of Lithuanian economy ([30]) a coherent development of other economic variables is calculated. However, most often, adverse scenarios are summarized by the GDP change. You can see the change of GDP according to the baseline scenario and several hypothetical adverse scenarios, depicted in Figure 1.6. Such principle of creation of the shock and adverse scanario, i.e. a combination of econometric models, statistical properties of the indicators, and expert judgement, helps us to ensure that the scenario will be extreme, but possible.

1.2.2 Assumptions

Stress testing, like any other economic modelling, is only a simplified representation of reality produced by formalizing the known basic laws and relationships, so this procedure does not cover all aspects of reality. Therefore, the results obtained through stress testing are not forecasts. On the contrary, they represent the analysis of highly unlikely events in order to identify potential problem areas in the banking system. Therefore, the results should be interpreted with caution and with due regard to the assumptions made.

The preparation of adverse scenarios that provide atypical situations, is one of the most prominent stress test procedure features. Under these circumstances, both market participants and government or regulatory authorities may accept non-standard solutions, so many well-known patterns may be void. This fact aggravates and limits modelling, thus the usual assumptions of the status quo are made, allowing calculations to flesh out and distance themselves from the unpredictable matters. During the stress testing, the following well-known static balance sheet assumptions are made:

- the structure of loan portfolio of the banks remains unchanged throughout the time horizon of the test;
- the natural amortization of the loan portfolio is offset by new loans, hence its gross value remains unchanged;
- any profit earned within the period covered by the test is used to increase capital;
- the banks do not pay any dividend and do not resort to any other means to raise capital;
- changes in risk-weighted assets may only result from changes in the quality of the loan portfolio;

- the banking supervisors and public authorities are assumed to take no measures to mitigate the consequences of an economic shock;
- the tests exclude the strategic decisions, which may be made by the banks, and their effects on the capital adequacy ratio.

The assumptions are static, therefore, do not realistically reflect the possible events. On the other hand, such an assumption provides modelling objectivity, clarity, and simplicity, which makes it possible to purify the essential impact channels. In addition, as demonstrated by the recent financial crisis, such an assumption is based on practical results: economic shocks highly reduced banks capability to do some self-change.

1.2.3 Credit loss modelling

The largest impact on the asset quality of commercial banks and, accordingly, on their capital adequacy ratio, arises from losses, which they sustain due to credit risk. Therefore, credit risk modelling is viewed as one of the key elements of macroeconomic stress testing as it helps assess the potential solvency and stability of the banks involved. The exercise includes the modelling of potential credit losses of a specific bank in light of a hypothetical macroeconomic scenario constructed for the test, i.e. the analysis of relationships between the credit risk and macroeconomic variables. All macroeconomic variables are applied as exogenous model variables and, therefore, determine the results of stress testing. Credit losses are directly dependent on the size of the loan portfolio, and, therefore, to ensure the stationarity of variables and comparability between banks, only relative ratios are modelled (credit losses over gross loan portfolio), and the loss of the euro amount is expressed only in the last stage of calculation.

In order to take into account the different operating characteristics and borrowers econometrically more a accurate assessment of the risk profile of the loan portfolio is broken down by the institutional sector. The bank loan portfolio is divided into seven parts. Loans to non-financial firms constitute five parts: 1) industry; 2) trade; 3) financial intermediation; 4) public sector; 5) other loans. The following loan groups are modelled separately: 6) loans to households for house purchase; 7) consumer and other loans to households.

The credit risk modelling scheme is presented in Figure 1.7. There are three stages in the calculation of the final result. The index i represents the institutional sector ($i = 1, 2, \dots, 7$), t is time period, and b denotes a particular bank.

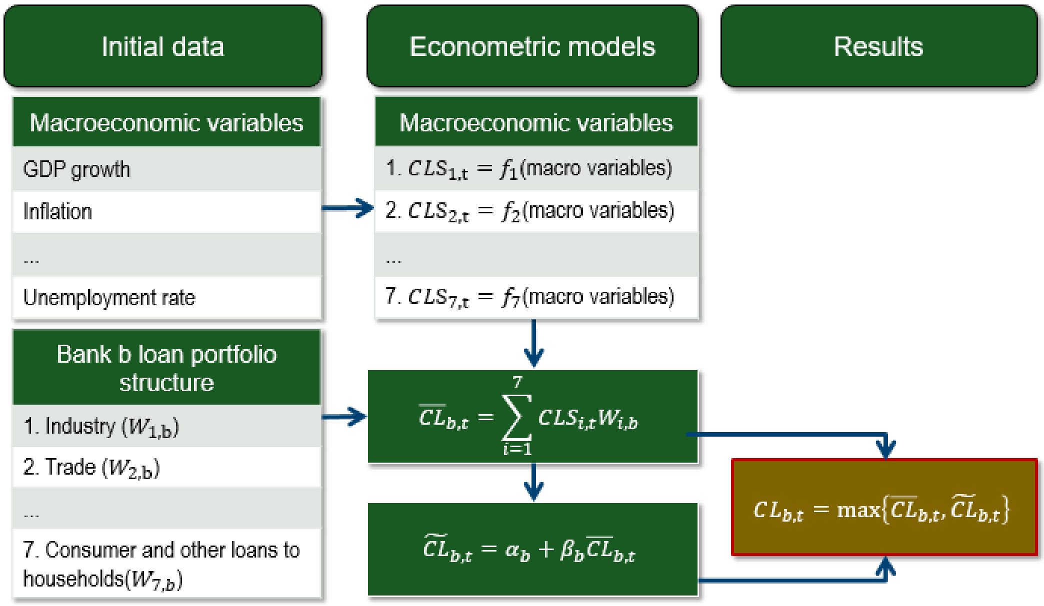

Figure 1.7: Modelling framework for the credit losses

Source: formed by authors.

Credit losses are modelled in several stages. The first stage deals with the examination of the relationships between credit losses in specific economic sectors and macroeconomic variables. Depending on the data quality, the entire loan portfolio of the banks is split into seven parts. The losses assessed reflect the average credit losses of the entire sector ($CLS_{i,t}$). Credit losses are modelled using a linear regression:

$$CLS_{i,t} = \alpha_i + \sum_{j=1}^{k} \beta_{i,j} M_{j,t} + \varepsilon_t, \tag{1.2}$$

where M_j are the macroeconomic variables, ε_t is the error term.

The second stage involves the assessment of credit losses in light of the portfolio structure of a specific bank. The expected credit losses ($\overline{CL}_{b,t}$) are calculated as a weighted sum (where the weights are the proportions of loans to respective sectors in the loan portfolio). Expected credit losses of the bank b are as follows:

$$\overline{CL}_{b,t} = \sum_{i=1}^{7} CLS_{i,t} W_{i,b} \tag{1.3}$$

where CLS_i are modelled losses of the economic sector i, $W_{i,b}$ is the weight of structure of the loan portfolio in the bank b.

The third stage involves the assessment of risk appetite of an individual bank. If the risk appetite of the bank b is not too high, the actual losses ($CL_{X,t}$) should be close to the expected ones ($\overline{CL}_{b,t}$). If the risk appetite is actually present, the exercise continues with the estimation of credit losses of the bank X ($\widetilde{CL}_{b,t}$):

$$\widetilde{CL}_{b,t} = \alpha_b + \beta_b \overline{CL}_{b,t} + \epsilon_t. \tag{1.4}$$

The final result shows the potential credit losses, expressed using the following formula:

$$CL_{b,t} = \max(\overline{CL}_{b,t}, \widetilde{CL}_{b,t}) \tag{1.5}$$

This formula is applied conservatively as it is not clear which of the credit loss estimates ($\overline{CL}_{b,t}$) or ($\widetilde{CL}_{b,t}$) is more suitable for the hypothetical scenario. The purpose of stress testing is to assess the potential losses in the worst-case scenario. Therefore, it is important to make sure that credit losses are not underestimated for the sole reason of differences in the internal provisioning rules applied.

The assessment of credit losses (eq.1.2) is carried out bt the least squares method. The significance of estimates is verified using the Newey-West covariance matrix estimate, that takes into account the residual heteroskedasticity and autocorrelation ([113], [114]).

Recall that, in the standard least squares model, the coefficient variance-covariance matrix is derived as:

$$\begin{aligned}\Sigma &= E(\hat{\beta}-\beta)(\hat{\beta}-\beta)^{'} \\ &= (X^{'}X)^{-1}E(X^{'}\epsilon\epsilon^{'}X)(X^{'}X)^{-1} \\ &= (X^{'}X)^{-1}T\Omega(X^{'}X)^{-1} \\ &= \sigma^2(X^{'}X)^{-1}.\end{aligned} \tag{1.6}$$

This derivation holds because it is assumed that error terms(ϵ) are conditionally homoskedastic, which implies that $\Omega = E(X^{'}\epsilon\epsilon^{'}X/T) = \sigma^2(X^{'}X/T)$. Newey and West (1987) [113] proposed the HAC method to calculate the estimate of $E(X^{'}\epsilon\epsilon^{'}X/T)$. The HAC coefficient covariance estimator is given by:

$$\hat{\Sigma}_{NW} = (X^{'}X)^{-1}T\hat{\Omega}(X^{'}X)^{-1}, \tag{1.7}$$

where $\hat{\Omega}$ is a long-run covariance estimator.

The estimation consists of a sample of the quarterly data from 2000 to 2012. Historical data show that two sectors, i.e. financial intermediation and public sector, do not have any credit losses, so their loss function is set to zero constants. All the other sectors are evaluated separately. The estimated equation coefficients are presented in Table A.1 (Appendix A). Exogenous variables are selected so that they meet the economic logic and well explain the data, i.e. it is aimed that the estimated coefficients have a particular sign and are statistically significant.

In the second stage, the econometric assessment is not applied. Bank's potential losses are calculated as the weighted sum. It is assumed that the loan portfolio structure does not change during the modelling period, and the weights taken from the known starting time. The last stage consist of evaluation of banks' willingness to take risk. The actual loss ($CL_{b,t}$) relationship with the expected losses ($\overline{CL}_{b,t}$) is studied and is estimated using a linear regression.

1.2.4 Profitability modelling

Operating profit generated by banks is a very important part of the overall assessment, because profits can offset a large part of incurred credit losses and thus have a significant impact on the final stress testing result. Due to this reason, the modelling of bank profitability is included in solvency stress testing (see, e.g., [8], [25], [63]).

In this framework, banks' operating profit is divided into the following six components: 1) net interest income, 2) net fee and commission income 3) net investment income 4) net other operating income, 5) operating expenses, 6) amortization. In assessing these lines separately, and not the entire operating profit in general, can accurately determine the relationship to the real economy, in addition, it can be seen which line change has the greatest impact on banks' profitability.

A dynamic panel data model, which helps to assess the relationships, equally affecting the entire banking system, has been chosen for profitability modelling. This is important since the stress tests carried out are top-down, i.e. they involve the comparison of the results of tests run on individual banks. Modelling the profitability for each bank separately, for instance by using a linear regression or time series model, you can get inconsistent results, i.e. that the macroeconomic variable can act positively on one bank and negatively on the another, or patterns can be completely different. It would be difficult to compare the results of stress test among banks. When a dynamic panel data model is applied, macroeconomic variables affect equally all banks, so the results are easily comparable. In addition, the model includes bank-specific indicators and the unobserved individual fixed effect, which allows us to customize the results without sacrificing comparability.

Profitability modelling is described by the equation:

$$Y_{b,t} = \alpha + \eta_b + \beta Y_{b,t-1} + \sum_{j=1}^{k} \gamma_j M_{j,t} + \sum_{s=1}^{l} \delta_s B_{s,b,t} + \varepsilon_{b,t}, \tag{1.8}$$

where $Y_{b,t}$ is a modelled profit (loss) statement line, i.e. net interest income, net fee and commision income or operating expenses, respectively, η_i means the unobserved bank fixed effect, $M_{j,t}$ are macroeconomic variables, $B_{s,b,t}$ are bank-specific variables.

The dynamic panel data model is estimated using instrumental variables (IV) estimation technique. This allows for the estimated coefficient to be consistent and effective. $Y_{b,t-2}$ and exogeous variables are used as intstruments in the model.

The available quarterly data series of bank accounts and macroeconomic variables covers the period from 2004 to 2013. Exogenous variables are selected on the basis of other authors (e.g., [7], [8], [11]) who examined the indicators that have an effect on bank profitability.

The most recent value of 12-quarter moving average was used as a proxy for other operating income and amortisation in the period considered by the stress test. Another operating income has a high volatility, so it is difficult to customize any econometric model. For this reason the 12-quarter moving average is used. Depreciation has a low impact on overall operating profit, so the modelling is the 12-quarter moving average.

1.2.5 Market risk assessment

The assessment of the banks' exposure to the market risk takes into account the volatility of net investment income. The approach is calibrated in such a way that a higher volatility in banks' investment income results in higher losses under the stressed conditions.

Under the baseline scenario, losses due to the market risk are computed as 1 times the standard deviation with respect to investment income of the previous three-year period. These losses are distributed across the stress test horizon in the following way: 50 per cent of losses are attributed to the first year of the test and 30 per cent – to the second year. The value obtained by subtracting the losses thus attributed from the average investment income of the previous three years represents the banks' sensitivity to the market risk under the baseline scenario.

Under the adverse scenario, losses due to the market risk are estimated as 2 times the standard deviation with respect to investment income of the previous five-year period. These losses are distributed across the stress test horizon as follows: 50 per cent of losses are attributed to the first year of the test and 30 per cent – to the second year. The value obtained by subtracting the losses thus attributed from the average investment income of the previous three years shows the banks' sensitivity to the market risk under the adverse scenario.

1.2.6 Aggregation of results

All econometrically-modelled variables are used to calculate banks' revenue and expenses. Using these values a simplified version of banks' profit and loss accout is constructed (see. Table 1.1).

Table 1.1: Summary of the modelled profit and loss account items

Profit and loss account item	**Estimation**
Net interest income (*NII*)	$= f(Y_{b,t-1}, M_{j,t}, B_{s,b,t})$
Net fee and commision income (*NCI*)	$= f(Y_{b,t-1}, M_{j,t}, B_{s,b,t})$
(-) Operating expenses (*OE*)	$= f(Y_{b,t-1}, M_{j,t}, B_{s,b,t})$
(-) Amortization (*D*)	$= MA(12)$
Net other operating income (*NOI*)	$= MA(12)$
Operating profit (*OP*)	$= NII + NCI - OE - D + NOI$
(-) Credit losses (*CL*)	$= f_{CL}(M_{j,t})$
Revenue from non-banking operations (*NBP*)	$= 0$
Profit before taxes (*GP*)	$= OP - CL + NBP$
(-) Taxes (*T*)	$= 0.15 \cdot \max(GP, 0)$
Net profit (*NP*)	$= GP - T$

Source: formed by authors.

This report allows us to analyze the results of stress testing in more detail and to identify the underlying factors. The c constructed profit and loss account report shows how banking income and expenses has changed during the testing period, when credit losses may be highest. However, most important are two variables - credit losses and net profit. Credit losses not only determine bank's net profit, but also affect the change in risk-weighted assets. The final variable of the profit and loss account is the net profit, which acts as a determinant factor of the ensuing changes in the bank's capital, which, in turn, define the value of the capital adequacy ratio. The procedure is described in Table 1.2 and illustrated in Figure 1.8.

The capital adequacy ratio is the main variable, which is used to sum-

Table 1.2: Summary of the solvency position calculation

Variable	Calculation
Capital (C)	$= C_{t-1} + NP$
Risk-weighted assets (RWA)	$= RWA_{t-1} - CL$
Capital adequacy ratio (CAR)	$= C/RWA$

Source: formed by authors.

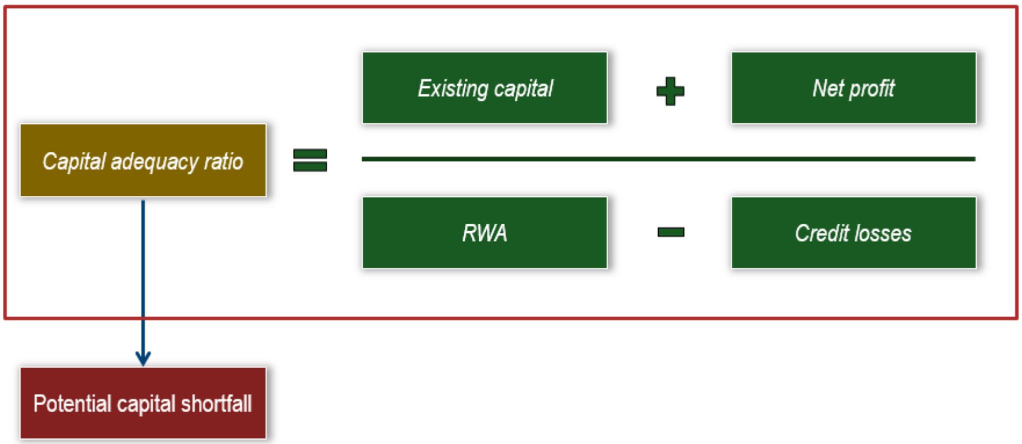

Figure 1.8: Solvency assessment

Source: formed by authors.

marize the results of stress testing and to calculate the potential capital shortfall (see Figure 1.8). This ratio can also be affected by changes in risk-weighted assets. These changes are not modelled directly. Instead, the developments in risk-weighted assets are defined by changes in the loan portfolio quality, taking the assumptions into account.

The capital adequacy ratio is a key variable as the basis for drawing conclusions about the resilience of the banking sector to unexpected economic shocks. Figure 1.9 shows possible dynamics of the capital adequacy ratio with various hypothetical stress testing adverse scenarios. However, while assessing the stability of the banking sector, it is important to take into account not only the final value of this indicator, but also all the factors that led to it and its change.

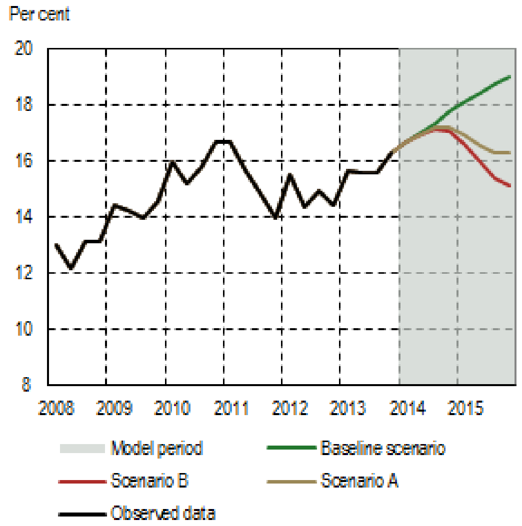

Figure 1.9: Capital adequacy ratio of the banking sector under hypothetical stress testing scenarios

Source: bank data and authors' calculation.

1.2.7 Conclusion of the Lithuanian stress testing framework

The recent global financial crisis has led to significant a increase in macroeconomic stress testing approaches that are used to assess the country's financial sector stability. Various central banks, supervisory authorities and international organizations applied stress testing methods, evaluated them again, and made some changes, that allowed them to improve the general macroeconomic stress testing methodology.

The proposed stress testing procedure enables the assessment of the banks in Lithuania, as well as of the whole banking sector, more precisely, it allows us to determine whether the bank accumulated capital reserves are sufficient to cover loan portfolio losses that could occur in the event of adverse changes in the macroeconomic environment. In this regard, there are two important aspects that are relevant to the banks: changes of the credit risk and profitability. Therefore, this stress testing framework focuses on econometric models that enable linking macroeconomic environment indicators to banks' profit and loss accounts.

The discussed procedure is a "top-down" assessment of the entire banking system. It means that all banks apply the same assumptions and common methodology, so the results are comparable with one another. On the other hand, such a procedure does not take into account the specific activities or bank-specific aspects. This fact is very important to consider when analysing and comparing the results.

Although the methodology applied here is in line with best practices in other countries and provided recommendations for stress testing, however, it is limited due to available credit risk data granularity. Just like any modelling framework, the Lithuanian stress testing procedure also has an important drawback: stress testing models allow us to evaluate the reality only in a simplified and generalized manner. Thus, the results must be assessed carefully. In addition, they are not forecasts. Such results, taking into account the specifics of modelling and assumptions, allow us to identify potential problem areas in the banking system and to focus on them.

The current testing procedure needs to be developed further, taking into account the international practice. One possible direction for the stress testing procedure development is more sophisticated econometric models for assessing the credit risk. Going in this direction, it would be useful to apply quantile regression, allowing the evaluation of nonlinear relationships between macroeconomic variables and bank indicators, specific to the economic shock period. Advanced modeling methods provide more opportunities, on the other hand, using sophisticated techniques, it is much more difficult to explain the results. In general, further efforts are associated with an attempt to overcome the analytical and modelling challenges and to increase the accuracy of the Lithuanian macroeconomic stress testing procedure.

1.3 Literature related to bank profitability

Several studies have attempted to identify internal and external determinants of bank profitability. Bank-specific or internal determinants of profitability come from balance sheets and profit (loss) accounts. Meanwhile, some industry-specific and macroeconomic variables have been proposed for both internal and external determinants, depending on the objective of the study.

A number of studies have analyzed profitability of either cross-country or individual countries' banking systems. Cross-country panel data sets have been investigated by [111], [43], [2], [67], [117], [7], [66], [97], [141], [60], [48]. Examples of single countries' analysis are studies of [8], [11], [9], [61], [36], [47], [142], [93], [137], [143], [18], [121], [34]. Certainly, the empirical results of the above mentioned studies vary as time periods, data sets, examined environments and countries differ. On the other hand, some internal and external determinants of bank profitability are common across all studies.

The main variables of the profitability measure used in the studies are return on assets (or return on average assets) and return on equity (or return on average equity). Another variable used for the profitability measure is net interest margin, i.e. net interest income divided by total assets.

Most of the studies examined variables such as bank size, capital ratio, operational efficiency, and risk measure for bank-specific determinants of profitability. The relationship between size and profitability is found to be negative by [117] and [36]. These results support evidence that large banks often face scale inefficiencies and small and medium banks encounter economies of scale and scope. Meanwhile, [67] and [11] determined a statistically insignificant relationship between size and profitability. An important determinant of bank profitability is the quality of loan portfolio. The effect of credit risk, i.e. loan losses, is clearly significant and negative ([11]; [36]; [142]; [93]). The authors in [43] and [2] have found that a loan to the total assets ratio, as a proxy of risk, is positive, meaning that a higher risk is rewarded with better profitability. However, this may be true during normal growth periods, but during crises higher risk leads to a higher losses and a lower profitability.

Results on the relationship between bank capital and profitability are rather interesting. In theory, the expected relationship between the capital adequacy ratio and returns should be negative, as a high capital adequacy ratio signalizes that bank is operating overcautiously and ignoring potentially profitable transactions. An empirically negative relationship has been determined by [66] and [137]. On the contrary, [43], [2], [117], [61] took another measure of bank capitalization, i.e. equity over total assets, and determined a strong positive relationship. Authors argue that banks with higher capital ratio indicate higher stability of a bank and its ability to gain profit in the future. This leads to lower costs of funding and, consequently, higher profitability. Furthermore, several authors found empirical evidence that better operational efficiency has a positive influence on bank profits. The authors in [11] used overhead costs over total assets, [66] and [47, 48] used the cost-to-income ratio as a measure of operational efficiency. Another internal determinant of profitability is the ownership of a bank. Demirguc-Kunt and Huizinga (1999) [43] argues that foreign ownership has a positive effect on profitability in developing countries and a negative effect in industrial countries. This result supports the fact that foreign banks have technological edge in developing countries and there is no such advantage in industrial countries. Dietrich and Wanzenried (2011) [47] have also found that foreign-owned banks are less profitable than Swiss banks.

Many studies also include macroeconomic and other external determinants of bank profitability. Authors in [117], [7], [47, 47], [137] among others used the annual GDP growth rate to link business cycle and bank earnings. [11], [9] and [93] used output gap as a measure of business cycle. All authors found a strong positive correlation between business cycle and bank profitability. Their results support the pro-cyclical feature of bank profits.

Empirical results of [43], [2], [11] show a positive impact of inflation on the bank profitability. This finding suggests that with inflation bank income increases more than bank costs. Though Dietrich and Wanzenried (2014) [48] argue that inflation has a positive and significant effect in low- and middle-income countries, it does not affect profitability in high income countries. Furthermore, many authors used some measure of interest rates in their researches. [43], [2], [61] used the short-term interest rate, [7] used the long-term interest rate, [36], [47] and [137] used the interest rate spread, i.e. the difference between long-term and short-term interest rates. The positive impact of higher interest rates reflects the fact that banks are able to increase lending rates quickly. This result may be due to imperfect market competitive conditions, especially in developing countries.

Market concentration is frequently used as an external determinant of profitability, which represents the market structure. This variable is related to the structure-conduct-performance (SCP) hypothesis. The SCP (or market-power) hypothesis states that increased market power yields monopoly profits. Research results of [111], [43] support the SCP hypothesis, i.e. bank concentration is statistically significant, and has a positive influence on profitability. [47] have also found a positive relationship, but the impact was minor. Meanwhile, the empirical results of [35] have shown that the structure-conduct-performance hypothesis is supported in the market of Western European banks, but is not supported in Eastern Europe. Similar results were obtained by [110], who has stated that a greater market share leads to a higher profitability in advanced economies, but the SCP hypothesis is not supported in emerging economies. However, [11] has found no evidence to support the SCP hypothesis, the concentration variable was negative and statistically insignificant.

Studies of Albertazzi and Gambacorta (2009) [7] and Andersen et al. (2008) [8] are most closely related to our research. [7] studied not only the relationship between return on equity and other determinants, but also analyzed different components of income statements. Authors investigated the link between items of income statements (net interest income, non-interest income, operating expenses, provisions and profit before taxes) and internal and external determinants. [7] has defined that GDP has an impact on both net interest income and provisions. Meanwhile, [8] analyzed aggregated Norwegian banking sector data using the error correction framework. Authors found a long-term co-integrating relationship between net interest income and GDP and the real interest rate and a similar co-integrating relationship between fee income and macroeconomic variables. Andersen et al. (2008) [8] have also stated that reversion to the long-term relationship is relatively fast in the net income equation and slower in the fee income equation.

Finally, the available literature makes a comprehensive analysis of internal and external determinants of bank profitability. Nevertheless, the long-term and the short-term relationship between bank income components and bank-specific and macroeconomic variables has not yet been analyzed in detail. Moreover, our study contributes to the relatively sparse amount of literature on the Lithuanian banking sector related to profitability analysis. This new analysis should serve as a relevant addition to the available literature on the determinants of bank profit.

1.4 Profitability analysis of the Lithuanian banking sector

In this thesis, the analysis of data on the Lithuanian banking sector, covering the period from 2004 to 2013, is examined. Therefore, this period includes pre-crisis and post-crisis data. The empirical results show that bank size is an important determinant in the long-term of all three items from the income statement. This result reflects the fact that the Lithuanian banking sector is still developing, therefore, bank size allows banks to generate higher revenue, but also causes higher expenses. Furthermore, the overall economic activity also significantly influences the performance of a bank. We determined a statistically significant long-term relationship between real GDP and net interest income as well as operating expenses. Therefore, prior expectations on long-term relationships between dependent variables and explanatory variables are confirmed by empirical results. Empirical estimation suggests various variables as short-term determinants of income statement items. We found that short-term interest rate and credit losses have an influence on net interest income, real export has an impact on net fee and commission income, and compensation of employees has an effect on operating expenses. The pooled mean group estimation technique and analysis of separate income statement items enables us to have a better insight into the Lithuanian banking sector and determinants of its revenue and expenses.

1.4.1 Determinants of bank income and expenses

In this work, we extract net interest income, net fee and commission income and operating expenses from bank income statements and analyze them separately. These income statement items are the main components of operating profit of different banks. In this section, we describe dependent variables and independent variables (determinants of bank income and expenses) selected for the study.

Dependent variables

Net interest income is the main component of bank revenue. Net interest income is calculated as the difference between interest income and interest expenses. Banks operating in Lithuania are described as traditional banks, i.e. their main business is to provide loans for customers and to collect deposits. Therefore, the main driver of interest income is revenue from loan payments received from customers. Other sources of interest income are less important for banks. Meanwhile, banks finance their activity through deposits and subordinated debts.

The recent financial crisis had a strong effect on net interest income. The pressure came from both sides, i.e. decreasing interest income and increasing interest expenses. Most of the banks experienced significant decline of net interest income during 2009. From 2010 the trend of net interest income is slightly upward.

Net fee and commission income is the second component of revenue included in our research. Net fee and commission income is calculated as the difference between fee and commission income and fee and commission expenses. Fee on payment transaction and currency exchange transactions are the main elements of net interest income. This income statement item also includes other fee charges from a variety of bank services.

Unlike net interest income, net fee and commission income showed an upward trend throughout the period analyzed. The importance of net fee and commission income increased particularly following the crisis as banks tried to compensate for the decline in net interest income by attracting more revenue from fee and commission charges.

Operating expenses include salaries and payments to employees, IT development costs as well as other operating expenses. On average, salaries and payments to employees amount to almost 70 percent of total operating expenses. Although there was a decline of operating expenses in 2009 and 2010, expenses began to grow again soon afterwards.

Independent variables

Real gross domestic product (GDP) is used as a measure of overall economic activity in Lithuania. Many studies have determined a positive relationship between GDP and bank profitability (e.g. [43]; [11]; [7]). There are several reasons why bank earnings may be pro-cyclical. First of all, demand for credit usually increases during the upswing of economic cycle as concerns of risks decrease. Secondly, increased demand for loans allows banks to set a wider interest margin. Therefore, the growth of revenue from lending activities could be more rapid than the growth of cost associated with bank financing. An increased demand for bank transactions and other operations also exists during an economic boom. It may lead to a higher fee and commission income. The relationship between economic activity and commercial bank revenue may be opposite during economic downswings. Hence we expect to find a relationship between GDP and bank revenue, i.e. net interest income and net fee and commission income.

Many companies that export their products abroad require currency exchange operations and other bank transaction services. Therefore, we include a level of real export as a measure of demand for bank services and expect a positive relationship between the real export and net fee and commission income.

Based on the findings of [7] and other authors, our analysis also includes inflation rate. [7] determined a positive and significant relationship between inflation rate and non-interest income as well as operating cost. Commercial banks may react to higher inflation rate by increasing charges on their transactions and operations. On the other hand, a higher inflation rate may put pressure on a bank's operating expenses since prices of different services also increase. Hence, we expect that inflation rate has an effect on net fee and commission income and operating expenses.

Three-month VILIBOR is used as a proxy of short-term interest rate. Banks finance long-term loans by taking short-term deposits. Three-month or six-month interest rate is one of the components used to set the price for loans. Therefore, interest rate is an important element for the business of commercial banks as it determines their ability to earn income from their core banking activities. We expect that short-term interest rate will have an influence on net interest income.

Salaries and payments to employees constitute a significant part of operating expenses. The growth of employee wages may raise operating expenses as well. Compensation per employee is used as a proxy for salaries and other payments to employees.

Unemployment rate may also have an impact on operating expenses. A higher unemployment rate allows banks to postpone the rise of salaries or even to cut them. And, vice versa, lower unemployment rate enables employees to negotiate better working conditions. Therefore, unemployment rate and its changes may also be important for the ability of banks to control their operating expenses.

In the related literature one of the main questions is whether the size of a bank affects its ability to gain more profit. Larger banks usually have a higher number of products to offer and a wider customer service network. This allows more cross-selling opportunities because banks have more clients and can offer more services. On the other hand, the effect of size may be negative because of bureaucratic and other reasons. Total assets are used to estimate the relationship between the bank size and dependent variables.

Loans provided by banks are the primary source of interest income. Therefore, loan stock (net) is included in the research of determinants for net interest income. Loan stock (net) is used instead of loan stock (gross) to reflect the fact that not all customers are repaying their loans.

Credit losses over loans stock (gross) is used as a measure of credit risk. Banks report their credit losses in their income statements. In theory, higher credit losses show that the quality of loan portfolio is deteriorating. Therefore, the negative effect on net interest income is expected from credit losses over the loans stock (gross) ratio.

The level of loan loss provisions is another measure of a bank's loan portfolio credit risk. Other than credit losses, which is a flow variable, loan loss provisions are a stock variable. A higher level of loan loss provisions indicates that a bigger part of loan portfolio does not generate revenue. Furthermore, banks may improve their credit risk monitoring and evaluation and that will have an effect on future decisions, related to portfolio growth and the level of credit risk.

The Herfindahl-Hirschman index (HHI) is used to examine the market structure in the Lithuanian banking sector. HHI is calculated as the sum of squares of the market shares of all the banks operating in Lithuania. The SCP hypothesis states that a highly concentrated market may lead to monopoly profits. Banks may pay lower interest rates on deposits and require higher rates on issued loans. On the contrary, a lower HHI in the banking sector might be a result of greater competition. In that case the relationship between market concentration and bank earnings may be negative. Therefore, the relationship between HHI and bank profit is undefined in literature sources and must be analyzed empirically.

1.4.2 Data and methodology

In this section, we describe our data set in more detail and introduce the methodology used to estimate the long-term and the short-term relationship between dependent variables and macroeconomic and bank-specific variables.

Data

In this study, we use a data set that contains quarterly data of eight banks operating in Lithuania and covers the period from 2004 to 2013, i.e. $N = 8$ and $T = 40$. Bank-specific variables (net interest income (NII), net fee and commission income (NCI), operating expenses (OE), total assets (A), loan stock (net) (NLS), credit losses over loans stock (gross) (CL), loan loss provisions (PRO)) are taken from quarterly income statements and balance sheet reports. All variables are expressed in thousand euros except CL which is calculated as a ratio. Macroeconomic variables are taken from the Lithuanian Department of Statistics (Statistics Lithuania). The levels of real GDP (GDP) and real export (REX) are expressed in million euros, compensation of employees (CPE) is expressed in euros, and unemployment rate (UNR) and inflation (HICP) are expressed in percentages. Three-month VILIBOR (STI) is taken from the database of the Bank of Lithuania. In order to standardize the level of variables and reduce volatility in the further analysis all variables, expressed in euros, are taken in logs. Descriptive statistics of the data are presented in Table B.1, Appendix B. Furthermore, pair-wise correlations of dependent and explanatory variables are presented in Table B.2, Appendix B. Correlation between variables show a potential relationship between them, but the results need to be interpreted with caution since they are estimated between non-stationary variables.

Methodology

Many studies of bank profitability use the dynamic panel approach as profitability or income statement items show persistence in time. In our case, we are also using this approach. However, as banks operating in the Lithuanian banking system are obviously heterogeneous, we would like to assess the long-term relationship that is common, but allow for short-term heterogeneous dynamics.

A large number of dynamic panel estimators is provided in the literature on the topic. In this study, the pooled mean group (PMG) estimator developed by Pesaran et al. (1997, 1999) [124, 125] is chosen because it suits our purpose best. The PMG estimator constrains long-term coefficients across cross-sectional units and at the same time allows intercepts, short-term coefficients and adjustment to the equilibrium relationship do differ. Haque (1999) [73] argues that neglecting cross-sectional heterogeneity in the short-term can lead to misleading inferences about the long-term relationship.

Consider the panel autoregressive distributed lag $ARDL(p, q_1, \ldots, q_n)$ model according to which dependent variables are explained by their own lags and by lags of bank-specific and macroeconomic determinants. Given data on cross-sectional units $i = 1, 2, \ldots, N$ and time periods $t = 1, 2, \ldots, T$, panel $ARDL(p, q_1, \ldots, q_n)$ model can be written as follows:

$$y_{it} = \mu_i + \sum_{j=1}^{p} \lambda_{ij} y_{i,t-j} + \sum_{j=0}^{q_i} \delta_{ij}^{'} B_{i,t-j} + \sum_{j=0}^{q_i} \gamma_{ij}^{'} M_{t-j} + \varepsilon_{it}, \tag{1.9}$$

where y_{it} is an explained variable (net interest income, net fee and commision income, operating expenses), B_{it} are bank-specific variables, M_t are macroeconomic variables, μ_i are unobserved fixed effects.

Using the first differences of $ARDL(p, q_1, \ldots, q_n)$ specification (1.9), we can write the following re-parameterization of the model:

$$\begin{aligned}\Delta y_{it} =& \mu_i + \phi_i y_{i,t-1} + \beta_i^{'} B_{it} + \eta_i^{'} M_t + \\ &+ \sum_{j=1}^{p-1} \lambda_{ij}^{*} \Delta y_{i,t-j} + \sum_{j=0}^{q_i-1} \delta_{ij}^{*'} \Delta B_{i,t-j} + \sum_{j=0}^{q_i-1} \gamma_{ij}^{*'} \Delta M_{t-j} + \varepsilon_{it},\end{aligned} \tag{1.10}$$

where: $\phi_i = -(1 - \sum_{j=1}^{p} \lambda_{ij})$; $\beta_i = \sum_{j=0}^{q_i} \delta_{ij}$; $\eta_i = \sum_{j=0}^{q_i} \gamma_{ij}$;
$\lambda_{ij}^{*} = -\sum_{m=j+1}^{p} \lambda_{im}$, $j = 1, 2, \ldots, p-1$;
$\delta_{ij}^{*} = -\sum_{m=j+1}^{q_i} \delta_{im}$, $j = 1, 2, \ldots, q_i - 1$;
$\gamma_{ij}^{*} = -\sum_{m=j+1}^{q_i} \gamma_{im}$, $j = 1, 2, \ldots, q_i - 1$, i.e. the coefficients are functions of initial coefficients in equation (1.9).

Furthermore, the equation (1.10) can be rearranged under the form of a panel error correction equation, i.e. changes in dependent variables are explained by gap from the long-term equilibrium and short-term dynamics of other variables:

$$\begin{aligned}\Delta y_{it} =& \phi_i (y_{i,t-1} - \alpha_i - \beta_i^{*'} B_{it} - \eta_i^{*'} M_t) - \sum_{j=1}^{p-1} \lambda_{ij}^{*} \Delta y_{i,t-j} - \\ &- \sum_{j=0}^{q_i-1} \delta_{ij}^{*'} \Delta B_{i,t-j} - \sum_{j=0}^{q_i-1} \gamma_{ij}^{*'} \Delta M_{t-j} + \varepsilon_{it},\end{aligned} \tag{1.11}$$

where ϕ_i denotes the error correction coefficient or the speed of adjustment to equilibrium values; long-term coefficients: $\alpha_i = -\frac{\mu_i}{\phi_i}$; $\beta_i^* = -\frac{\beta_i}{\phi_i}$; $\eta_i^* = -\frac{\eta_i}{\phi_i}$.

The estimation of this model takes the following assumptions:

Assumption 1: the disturbances ε_{it} in (1.9) are independently distributed across i and t, with means 0, variances $\sigma_i^2 > 0$, and finite fourth-order moments. They are also distributed independently of the regressors.

Assumption 2: $ARDL(p, q_1, \ldots, q_n)$ model is stable in that the roots of $\sum_{j=1}^{p} \lambda_{ij} z^j = 1$, $i = 1, 2, \ldots, N$ lie outside the unit circle. This assumption ensures that $\phi_i < 0$ and hence there exists a long-term relationship between dependent variable and regressors. [123] give a framework for testing Assumption 2, irrespective of whether the regressors are $I(0)$ or $I(1)$.

Assumption 3 (Long-term homogeneity): The long-term coefficients on B_{it} and M_t are the same across the groups, i.e. $\beta_i^* = \beta^*$ and $\eta_i^* = \eta^*$, $i = 1, 2, \ldots, N$.

The authors [122] have shown that ARDL approach yields consistent estimates of the long-term coefficients, irrespective of whether the underlying regressors are $I(1)$ or $I(0)$. We will use panel unit root tests to examine stationarity of the data.

Furthermore, we also use the mean group (MG) estimator developed by Pesaran and Smith (1999) [126] which imposes no restriction on long-term coefficients. The MG estimator allows intercepts, short-term coefficients, error correction coefficients and long-term coefficients to differ across cross-sectional units. [126] have shown that the MG estimator will produce consistent estimates of long-term coefficients. However, the MG estimator will be inefficient in case of long-term homogeneity. [125] argue that in case of long-term homogeneity the PMG estimator is consistent and efficient. Therefore, the Hausman type test ([75]) could be applied to the difference between the MG and PMG estimators to test homogeneity of long-term coefficients.

1.4.3 Empirical results

Unit root test

The first step is to analyze statistical properties of the data set. We used unit root tests to investigate stationarity and order of integration of the data. For dependent variables and bank-specific variables we performed panel unit root tests: Im, Pesaran and Shin developed by [86] and Fisher ADF, Fisher PP proposed by [106] and [32]. For macroeconomic variables we applied the traditional augmented Dickey-Fuller test (ADF) introduced in [45].

The null hypothesis for panel unit root tests is that all series contain a unit root. The alternative hypothesis is that the fraction of individual series, that follows stationary processes, is non-zero. In [86]) panel unit root estimation is based on averaging individual augmented Dickey–Fuller unit root tests. Meanwhile, Fisher ADF and Fisher PP tests statistic is derived by combining p-values from individual unit root tests.

The results of the panel unit root tests of dependent variables are presented in Table 1.3. All three tests support the hypothesis of a unit root in the data. Furthermore, analysis of first differences shows that dependent variables are integrated of order one. The stationarity of bank-specific explanatory variables was also tested and the results are reported in Table B.3 (Appendix B). Panel unit root tests show that bank assets, loan stock (net) and provisions are $I(1)$ processes and credit losses are integrated of order $I(0)$.

The null hypothesis of the augmented Dickey-Fuller unit root test is that data have unit root, and an alternative hypothesis is that data are stationary. The null hypothesis of the ADF test is accepted if test statistics are bigger than [104] critical value at 5 percent significant level. ADF unit root test results (Table B.4, Appendix B) show that all macroeconomic variables, except unemployment rate, are integrated of order one. Unit root test results verify that we can use the pooled mean group estimator. However, unemployment rate may not be a suitable regressor.

Table 1.3: Panel unit root tets (dependent variables)

Variable	Im, Pesaran and Shin		Fisher ADF		Fisher PP	
	Level	First differences	Level	First differences	Level	First differences
Net interest income	-0.066	-8.036	18.06	90.70	21.24	479.6
	(0.474)	*(0.000)*	*(0.321)*	*(0.000)*	*(0.170)*	*(0.000)*
Net fee and commision income	-0.023	-11.92	15.22	217.3	34.98	728.7
	(0.491)	*(0.000)*	*(0.508)*	*(0.000)*	*(0.004)*	*(0.000)*
Operating expenses	-0.073	-7.685	14.61	154.6	14.10	670.4
	(0.471)	*(0.000)*	*(0.554)*	*(0.000)*	*(0.592)*	*(0.000)*

Notes: p-values are reported in the parenthesis. For Im, Pesaran and Shin, Fisher ADF panel unit root tests number of lags was selected using the AIC criterion. Panel unit root tests include intercept and trend.
Source: Bank data and author's calculation.

Co-integration tests

Panel unit root tests have showed that most of the dependent and independent variables are integrated of order one. The second step was to test whether there is a long-term relationship between variables. This relationship was tested using heterogeneous panel co-integration tests proposed in [118, 119]. The null hypothesis of no co-integration is tested using residual-based tests. [118, 119] introduced two types of tests: within dimension test and between dimension test. The first type of tests are based on pooling residuals of co-integrating equation along the within dimension. This type includes four statistics: panel ν, panel ρ, panel PP, and panel ADF. The second type of tests are based on pooling residuals of co-integrating equation along the between dimension. This type includes three statistics: group ρ, group PP and group ADF. [118] has stated that all seven appropriately standardized statistics are asymptotically normally distributed.

Panel co-integration tests results, presented in Table B.5 (Appendix B), show bivariate estimation of co-integration between net interest income and independent variables. The null hypothesis of no co-integration is rejected at the 5 per cent significance level. The results show that a strong long-term relationship exists between net interest income and total assets as well as loan stock (net), where null hypothesis of no co-integration was rejected by six out of seven statistics. The results of panel co-integration tests also support the long-term relationship between the dependent variable and macroeconomic variables, i.e. compensation of employees, inflation and short-term interest rate. At the 5 per cent significance level three statistics reject the null hypothesis of no co-integration between net interest income and real GDP, but at the 10 per cent significance level there are four statistics that reject the null hypothesis.

Table B.6 (Appendix B) provides the panel co-integration test results of a long-term relationship between net fee and commission income and explanatory variables. The majority of statistic data rejects the null hypothesis of no co-integration between the dependent variable and bank-specific variables (total assets, loan stock (net) and provisions) as well as macroeconomic variables (real GDP, inflation, short-term interest rate compensation of employees, unemployment rate and HHI). Three out of seven statistics also reject the null hypothesis for real export.

The results of panel co-integration tests between operating expenses and independent variables are presented in Table B.7 (Appendix B). Test results show that operating expenses are less dependent on bank-specific and macroeconomic variables than other dependent variables. The null hypothesis is rejected at the 5 percent significance level for total assets, loan stock (net) compensation of employees and short-term interest rate. Test results supporting the long-term relationship between operating expenses and short-term interest rate are slightly unexpected. At the 10 percent significance level the null hypothesis is rejected for inflation and real GDP. Panel co-integration test results show that there is no long-term relationship between operating expenses and other explanatory variables.

Panel co-integration tests results confirm that items of income statements (net interest income, net fee and commission income and operating expenses) have a long-term relationship with some explanatory variables. Therefore, it is reasonable to apply the panel error correction model.

Net interest income

After a preliminary examination of the data set for stationarity and co-integration we continue our analysis and use the pooled mean group estimation methodology. Therefore, we estimate a separate equation:

$$\Delta y_{it} = \phi_i(y_{i,t-1} - \alpha_i - \beta X_t) - \gamma_i \Delta X_t + \varepsilon_{it} \tag{1.12}$$

where y_{it} is net interest income and X_t is a bank-specific (B_{it}) or macroeconomic variable (M_t). Hence we include one by one determinants of net interest income and estimate equation (1.12) independently of other explanatory variables. Such estimation can give primary information on the importance of variables in the long-term relationship. Of course, their eventual impact may be different, since variables also depend on the correlation with other determinants of a dependent variable.

The pooled mean group estimation results[1]show that bank-specific variables, i.e. total assets and loan stock (net), are statistically significant in the long-term. Furthermore, there are also negative and statistically significant error correction coefficients for those variables. These results indicate that bank size is an important variable determining net interest income. Similarly, PMG estimation results show that there is a statistically significant long-term relationship between net interest income and real GDP as well as compensation of employees. Error correction coefficients were also negative and significantly different from zero. PMG estimation results show that other variables are not statistically significant for net interest income in the long-term.

The final step was estimation of the panel error correction model of net interest income determined by bank-specific and macroeconomic variables. Since we are interested in both long-term and short-term determinants of a dependent variable, we include only two variables in the long-run estimation. Furthermore, we also include two independent variables in the short-term estimation. The variables to be included in the model have been chosen using the following methodology. We begin with the long-term estimation, where we choose variables based on the results of panel co-integration tests and the individual PMG estimation results. Afterwards we include short-term variables selected based on the significance of coeffi-

[1]Pooled mean group estimation results of individual variables are not presented to the conserve space, but are available upon request.

cients and economic meaning of the sign.

Panel co-integration tests and individual PMG estimation show that there is long term-relationship between the size (total assets or loan stock (net)) and net interest income. Previous examination of the data set also shows that macroeconomic variables (real GDP and compensation of employees) are an important determinant in the long-term. We estimated equation (1.11) where one bank-specific and one macroeconomic variable were included in the long-term. Based on the pooled mean group estimation results we include total assets and real GDP in our model as long-term determinants of net interest income. Co-integration of the dependent variable and two independent variables was also tested with panel co-integration tests. The results, presented in Table B.8 (Appendix B), reject the null hypothesis of no co-integration between these variables.

Short-term determinants of net interest income were selected following the examination of contemporaneous and one period lagged variables. Short-term interest rate and one period lagged credit losses were included in the final panel error correction model. The pooled mean group estimation results for the model of net interest income are presented in Table 1.4.

The pooled mean group estimation results show that total assets and real GDP have a positive influence on net interest income in the long run. The error correction coefficient is negative and statistically significant at the 5 per cent significance level. This result supports the estimated long-term relationship as valid. [7] and [8] have also obtained similar results for the long-term determinant of net interest income. The Lithuanian banking sector is still developing, therefore, it may gain from economies of scale. Lithuanian banks are too small to face scale inefficiency determined by other authors ([117]; [36]). Meanwhile, increasing overall economic activity creates greater demand for loans and banks are able to increase their revenue from interest income. A change in three-month VILIBOR has a positive effect on net interest income in the short-term. Similarly to [7] and [61], our finding confirms the fact that banks have market power and are able to increase lending rates quickly. The estimation results also show that dynamics of loan portfolio quality plays an important role in the short-term changes of net interest income. Therefore, signs of the estimated coefficients

Table 1.4: Net interest income estimation results

Long-term coefficients	Estimation method	
	PMG	MG
$\log(GDP)$	1.191 *(0.005)*	1.419 *(0.027)*
$\log(A)$	0.556 *(0.000)*	0.443 *(0.007)*
Short-term coefficients		
Error correction	-0.441 *(0.000)*	-0.513 *(0.000)*
$\Delta(STI)$	0.105 *(0.088)*	0.100 *(0.067)*
$\Delta(CL)_{-1}$	-0.048 *(0.040)*	-0.039 *(0.059)*
Constant	-4.540 *(0.000)*	-6.023 *(0.019)*

Note: p-values are reported in parenthesis.
Source: Statistics Lithuania, bank data and author's calculation.

are in line with our prior expectations.

As a robustness check we performed a mean group estimation of the panel error correction model. MG estimator imposes no restriction on long-term coefficients, i.e. it allows heterogeneity in the long-term. The MG estimates of the coefficients are similar to PMG estimates. This is verified by the Hausman test statistic ([75]) of 0.49, which is $\chi^2(2)$ under the null hypothesis of no difference between the PMG and MG estimators. Therefore, we may conclude that the PMG estimator is efficient and preferred over the MG estimator.

Net fee and commision income

The same steps were taken when estimating the long-term and short-term relationship between net fee and commission income and explanatory variables. The pooled mean group estimation results of an individual determinant of net fee and commission income show that bank-specific (total assets, loan stock (net)) and macroeconomic variables (real GDP, compensation of employees) are significantly different from zero in the long-term. Furthermore, error correction coefficients were negative and statistically significant at the 5 per cent significance level. Although four out of seven panel co-integration test statistics could not reject null hypothesis of no co-integration between net fee and commission income and real export, the PMG estimation results suggest otherwise, i.e. the long-term coefficient and the negative error correction coefficient are statistically significant. Therefore, all these variables may be included in the final equation for the net fee and commission income.

The pooled mean group estimation results of net fee and commission income are presented in Table 1.5. Similarly to the net interest income equation, total assets are also an important determinant of this income statement item in the long-term. Bigger banks are able to offer more products to their customers and, therefore, have more cross-selling opportunities. This result differs from that of [7], who found that the size of a bank has a negative impact on net fee and commission income. As for the second explanatory variable, we found that real export influences dynamics of NCI. A higher trade activity requires more currency exchange and other banking operations and, therefore, generates fee and commission income for banks. Error correction coefficient is equal to -0.321, i.e. less than in NII, but it is also significantly negative. Therefore, net interest income is closely related to the size of a bank and economic activity. The PMG estimation results were also supported by panel co-integration tests (Table B.9, Appendix B) where most of the test statistics reject the null hypothesis.

Short-term dynamics are explained by the lagged value of change in net fee and commission income itself. Moreover, lagged value of change in real export has a positive impact on change in NCI. Other examined variables were not statistically significant in the short-term.

The mean group estimation of the error correction model for net fee and commission income gives similar coefficient estimates. However, real export coefficients are less significant than PMG estimates The Hausman

Table 1.5: Net fee and commission income estimation results

Long-term coefficients	Estimation method	
	PMG	MG
$\log(REX)$	0.551	0.403
	(0.000)	*(0.072)*
$\log(A)$	0.321	0.444
	(0.000)	*(0.001)*
Short-term coefficients		
Error correction	-0.321	-0.537
	(0.000)	*(0.000)*
$\Delta\log(NCI)_{-1}$	-0.191	-0.094
	(0.000)	*(0.008)*
$\Delta\log(REX)_{-1}$	0.5131	0.440
	(0.061)	*(0.142)*
Constant	-0.344	-1.190
	(0.001)	*(0.032)*

Note: p-values are reported in parenthesis.
Source: Statistics Lithuania, bank data and author's calculation.

test statistic is equal to 0.68, and, therefore, supports the assumption that homogeneity could be imposed in the long run.

Operating expenses

The last dependent variable for which we estimated long-term and short-term determinants is operating expenses. Similarly to other dependent variables, bank size (total assets or loan stock (net)) is an important determinant for operating expenses. Individual PMG estimation shows that long-term coefficients and negative error correction coefficients are significantly different from zero. Moreover, macroeconomic variables (real GDP and compensation of employees) are also significant determinants for the dependent variable. Contrary to panel co-integration test results, PMG estimation does not support the long-term relationship between operating expenses and short-term interest rate. Other variables were also not significant in the long-term.

Table 1.6 presents the results of the pooled mean group estimation of operating expenses equation. We determined a long-term relationship between OE, total assets and the real GDP. The error correction coefficient is highest among the estimated equations and significantly negative. This result suggests that although bank size and overall economic activity is important for the income of banks (NII and NCI), but it increases expenses as well. Bigger banks have more employees and larger customer service chains. Furthermore, real GDP growth creates initiatives for employees to require salary increase and prices for services supporting banking operations may also rise (panel co-integration tests are given in Table B.10, Appendix B). These findings are in line with [7], who also found a positive impact of total assets and real GDP on operating expenses.

Table 1.6: Operating income estimation results

Long-term coefficients	Estimation method	
	PMG	MG
$\log(GDP)$	1.035 *(0.000)*	1.240 *(0.029)*
$\log(A)$	0.462 *(0.000)*	0.525 *(0.000)*
Short-term coefficients		
Error correction	-0.535 *(0.000)*	-0.674 *(0.000)*
$\Delta \log(OE)_{-1}$	-0.203 *(0.001)*	-0.173 *(0.005)*
$\Delta \log(CPE)$	0.351 *(0.014)*	0.265 *(0.084)*
Constant	-4.029 *(0.000)*	-5.070 *(0.005)*

Note: p-values are reported in parenthesis.
Source: Statistics Lithuania, bank data and author's calculation.

Short-term dynamics of operating expenses are explained by the lagged value of change in operating expenses and by change in compensation of employees. In line with our expectations, CPE is an important determinant of operating expenses and has a positive effect on it because expenses to employees constitute a significant part of operating expenses.

Similarly to the previous equations, the mean group estimation of error correction model for operating expenses gives close values of coefficients. The estimated Hausman test statistic is equal to 1.83. Therefore, we also conclude that long-term homogeneity could be imposed on operating expenses.

1.4.4 Conclusion of the profitability analysis

In this work, we examined the long-term and the short-term relationship between bank profitability and explanatory variables, i.e. we analyzed which bank-specific and macroeconomic variables influence income statement items (net interest income, net fee and commission income, and operating expenses). We used the data set from the Lithuanian banking sector covering the period from 2004 to 2013, and applied the pool mean group estimator to investigate determinants of bank revenue and expenses.

Empirical results show that the size of a bank expressed as total assets is an important long-term determinant of revenue and expenses. As the Lithuanian banking sector is still developing, banks are not that big that could face scale inefficiencies found in other researches ([117]; [36]). Lithuanian banks could be attributed to small and medium size banks and, therefore, they can exploit economies of scale and scope. As expected, economic activity is an important macroeconomic determinant of income statement items. Increasing GDP creates initiatives to borrow and invest more in the economy, leading to a higher net interest income. On the other hand, increasing economic activity requires banks to meet higher demand for transactions and loan portfolio maintenance, i.e. banks must raise operating expenses. This finding is in line with the conclusions in [7], [47] and many other authors who also found a pro-cyclical feature related to bank profits.

Our estimation shows that change in interest rate and change in credit losses has an impact on net interest income in the short-term. A positive influence of the interest rate reflects the fact that banks have the market power to increase lending rate quickly. Similarly to [11] and [36], we found that the decreasing quality of loan portfolio lowers the ability of banks to generate revenue, therefore, credit losses have a negative effect on net interest income. Empirical results also show that real export is an important determinant of net fee and commission income in the long-term and in the short-term. Change in demand for currency exchange and other bank operation influences bank revenue from fees and commissions. The short-term relationship between operating expenses and compensation of employees shows that employees' wages constitute a significant part of banks' expenses. All short-term relationships are in line with theoretical expectations.

Other bank-specific and macroeconomic variables were considered to be less important or insignificant determinants of bank profitability. SCP hypothesis is not supported as HHI was found to be an insignificant determinant of all three income statement items and, therefore, not included in the final equations. This result is in line with the findings of [35], who found no evidence to support the SCP hypothesis in Eastern Europe. However, some of these determinants would be important if we could include them in the models.

The approach, used in this study, allows us to analyze long-term and short-term determinants of bank revenue and expenses. These results may be used by supervisors of banks as part of the stress testing exercise used to assess the stability of the banking sector. However, the pooled mean group estimator requires the data set to be quite large. Therefore, a larger data set on Lithuanian banks would help us to include more determinants into the models and have a better understanding of long-term and short-term relationships. This issue could be addressed in future analyses.

Chapter 2

Cluster analysis and forecasting

2.1 Literature related to the time series and functional data cluster analysis

There has been an increased interest in time series clustering after more time dependent data in various fields became available. The results of cluster analysis depend on many choices that must be fixed during the clustering process. In general, the cluster analysis consists of a few basic steps [72]. The first step is to select the features of time series on which clustering is going to be performed. The features should contain all possible information related to the task of interest. The second step is to define a dissimilarity measure between time series. A dissimilarity or distance measure quantifies and compares similarities of two time series. The next step is to choose a clustering algorithm which groups data into clusters. Since the precise number of clusters is not known a priori, clustering results must be evaluated using the appropriate criterion. The final step is the interpretation of results. The expert judgement is also important when drawing the conclusion of cluster analysis.

One of the key elements in cluster analysis is determination of an appropriate dissimilarity/similarity measure between two time series. Since time series has dynamic character the concept of similarity is complex. The two most widely used dissimilarity measures work with raw data. The conventional Euclidean distance measures the distance between two time series at each point in time. According to a dynamic time warping distance ([19]), two time series are close, if there exists a mapping, expressing a time distortion by a deceleration or acceleration so that the maximum length between all the coupled observations is minimized. However, these two dissimilarity measures do not take into account the growth behaviour of the time series. Chouakria and Nagabhushan (2007) [33] have proposed a dissimilarity measure that accounts for both closeness of values and behaviour of time series.

Given that time series are usually high dimensional data that could be noisy, various methods are used to extract some features of data. Dissimilarity is then measured based on these features. Some distance measure takes into account the properties of time series such as correlation [69], autocorrelation [21] or partial autocorrelation. Other distance measures, proposed in the literature, transform raw data and then estimate closeness based on the transformed data. Chan and Fu (1999) [31] among others used a discrete wavelet transform, Faloutsos et al. (1994) [55] employed the discrete Fourier transform, Koegh et al. (2001) [95] proposed piecewise aggregate approximation, Lin et al. [102] introduced a symbolic aggregate approximation. Many other representations are also used in the literature.

A different approach, used in the time series clustering literature, is to assume that time series are generated from a particular parametric model. For example, Piccolo (1990) [128] defined a distance measure in the class of invertible ARIMA processes as the Euclidean distance between the AR(∞) operators approximating the corresponding ARIMA structures. For a class of invertable and stationary ARMA processes, Maharaj (1996) [107] has proposed measure, based on hypothesis testing to determine whether data generating processes significantly differ between two time series. Another group of dissimilarity measures is based on comparing levels of complexity of time series. This category of distances includes the normalized compression distance, proposed by Li et al. (2004) [99] and the complexity-invariant dissimilarity measure introduced by Batista et al. (2011, 2014) [16, 17].

Many authors, working on time series clustering or comparing different dissimilarity measures, make an assumption that time is discrete (e.g. [49]). However, there is another field of research in which the dissimilarity is measured using functional data or their properties. There are several approaches used in the literature. Filtering approach consists of the first step in which curves are expanded into some finite basis of functions and the second step in which clustering is performed using the basis expansion coefficients. For example, Abraham et al (2003) [1] considered B-splines and Peng and Müller (2008) [120] used the principal component scores. Adaptive methods perform simultaneously the dimensionality reduction and clustering, because they consider that the functional form of data depends on clusters. James and Sugar (2003) [89] assumed that the basis expansion coefficients of the curves into a spline basis were distributed according to the mixture of Gaussian distribution with a different mean for each cluster and common variance. Samé et al (2011) [138] have assumed that the curves come from a mixture of regressions on a basis of polynomial functions, with possible changes in regime. Another approach considers dissimilarity or distance between curves. The examples of this method could be found in [56] and [85]. Meanwhile, Jacques and Preda (2014) [87] provided a good survey of methods, used for the functional data clustering.

Once the initial distance matrix is computed, a clustering algorithm can be used to divide data into clusters. There are many different clustering algorithms that are used to cluster time series. The literature provides several categories of algorithms and methods of each category. The clustering is crisp, if each element belongs to only one cluster, or partition is fuzzy, if one element could be in more than one cluster to a different degree. A popular category of crisp clustering is partitional algorithms. This category includes the methods such as *k-means* [105], where the mean of elements in the cluster represents each cluster, and partitioning around metoids (PAM) [115], where the most centrally located element in a cluster represents each cluster. Similar methods of fuzzy clustering are the fuzzy *c-means* [20], modified fuzzy *c-means* [78] and fuzzy *c-metoids* [96]. A second commonly used category is hierarchical clustering. There are two types of hierarchical algorithms: agglomerative, where each element is placed in its own cluster and then elements are merged to form larger clusters until there is one cluster, and a divisive method, that works in the opposite direction. Another category of clustering algorithms is density-based algorithms, which include methods like $DBSCAN$ [53]. In this clustering algorithm, a cluster is extended as long as the density (number of elements) in the neighbourhood exceeds some threshold. The main idea of grid-based clustering algorithms (methods like $STING$ [144]) is to quantize the element space into a finite number of cells that form a grid stucture on which clustering operations are performed. In recent years some new clustering algorithms (e.g. [103]) have been proposed in the literature.

Since the clustering algorithms divide unlabelled data into significant groups, it is important to evaluate the clustering results and find partitioning that fits data best. There are three basic criteria by which the clustering evaluation is usually performed. The first one is compactness of a cluster which should be minimized, i.e. the members of each cluster should be as close to each other as possible. The second one is connectedness of the cluster, i.e. to what extent the elements are placed in the same cluster as their nearest neighbours. The third criterion is separation of clusters which should be maximized, i.e. the clusters should be widely spaced. There are many validity assessment methods provided in the literarure, that combine the measures mentioned above. An example of valididy indixes are the Dunn index [52], Davies and Bouldin Index [41] or Silhouette Width [136, 94].

Time series clustering problems arise in a wide range of fields, including business and economics, physics, medicine, meteorology, and many others. Liao (2005) [100] provided a good survey on time series clustering that includes many references. Meanwhile, an interesting overview of recent time series data mining methods and algorithms can be seen in [58]. As it has been pointed out by Liao (2005) [100], there are not so many studies comparing different time series dissimilarity measures. Few examples of papers that compared several distance measures are works by [49] and [44].

To summarize, the existing literature provides many choices of methods or algorithms at each step of the cluster analysis. Some methods may produce good clustering results in one instance while in other cases, different methods would be better. So the choice of a particular method may influence the final results significantly. Therefore, expert judgement is also an important step in conducting the clustering of a given dataset.

2.2 Functional data

2.2.1 Smoothing

In this section, a short overview of functional data and their main characteristics is presented. Assuming that the functional data $y_i(t)$ for replication i arrive as a finite set of measured values, we have to convert these values to a function $x_i(t)$ with the values computable for any moment t. If our model is expressed as $y_i(t) = x_i(t) + \varepsilon_i(t)$, where the residuals $\varepsilon_i(t)$ are independent of $x_i(t)$, then we can get the orginal signal $x_i(t)$ using a linear smoother:

$$\hat{x} = \sum_{i=1}^{n} s_{ij} y_i \Rightarrow \hat{x} = Sy,$$

where s_{ij} is the weight that the point t_j gives to the point t_i.

One way to represent a curve is to use a set of functional building blocks ϕ_k, $k = 1, \dots, K$ called *basis functions*, which are combined linearly. That is, the function $x(t)$ is expressed in the mathematical notation as:

$$x(t) = \sum_{k \in \mathbb{N}} c_k \phi_k(t) \approx \sum_{k=1}^{K} c_k \phi_k(t) = \mathbf{c}^{'} \mathbf{\Phi},$$

and is called a basis function expansion. The parameters $c_1, c_2, \dots, c_K$ are the coefficients of expansion.

There are several types of bases - Fourier, B-spline, Wavelets, Exponential, Power, Polynomial, etc. Probably two of the most used bases are Fourier and B-spline. These two bases often need to be supplemented with *constant* and *monomial* basis functions. These four basis functions can deal with most of the applied problems in practice.

Many functions are required to repeat themselves over a certain period T, as would be required for expressing the seasonal trend in a long time series. Fourier basis is a periodic basis, composed by the following orthonormal functions:

$$
\begin{aligned}
\phi_1(t) &= 1 \\
\phi_2(t) &= \sin(\omega t) \\
\phi_3(t) &= \cos(\omega t) \\
\phi_4(t) &= \sin(2\omega t) \\
\phi_5(t) &= \cos(2\omega t) \\
&\vdots
\end{aligned}
$$

where the constant ω is related to the period T by the relation: $\omega = 2\pi/T$.

It can be noted that, after the first constant basis function, Fourier basis functions are arranged in successive sin/cos pairs, with both arguments within any pair being multiplied by on of the integers $1, 2, \ldots$ up to some upper limit m. If the series contains both elements of each pair, as is usual, the number of basis functions is $K = 1 + 2m$.

The most common bases for non-periodic functional data are spline functions. Splines are piecewise polynomials. Spline bases are more flexible and therefore more complicated than finite Fourier series. They are defined by a range of validity, knots, and order. There are many different kinds of splines. However in this work we consider only B-splines.

Splines are constructed by dividing the interval of observation into subintervals, with boundaries at points, called break points or simply breaks. Over any subinterval, the spline function is a polynomial of fixed degree or order, but the nature of the polynomial changes as one passes into the next subinterval. The term degree is used to refer the highest power in the polynomial. The order of a polynomial is one higher than its degree. For example, a straight line is defined by a polynomial of degree one since its highest power is one, but is of order two because it also has a constant term.

A spline basis is actually defined in terms of a set of knots. These are related to the break points in the sense that each knot has the same value as a break point, but there may be multiple knots at certain break points.

At each break point, neighbouring polynomials are constrained to have a certain number of matching derivatives. The number of derivatives that must match is determined by the number of knots, positioned at that break point. If only one knot is positioned at a break point, the number of matching derivatives (including the function value itself) is twice less than its order, which ensures that for splines of more than order two the join will be seen to be smooth. This is because a function, composed of straight line segments of order two will have only the function value (the derivative or order 0) matching, so the function is continuous but its slope is not; it means that the joins would not be seen as smooth by most standards.

Order four splines are often used, consisting of cubic polynomial segments (degree three), and a single knot per break point makes the function values and first and second derivative values match. In the large majority of applications, there will be only a single knot at every break point except for the boundary values at each end of the whole range of t. The end points, however, are assigned as many knots as the order of the spline, implying that the function value will, typically, drop to zero outside of the interval over which the function is defined. For an in depth overview of splines see, e.g. de Boor (2001) [42].

Among the different types of spline the B-spline basis is used because they are fast in computation of polynomials and flexible in representation. The B-spline basis can be expressed as:

$$\text{B-spline} = \sum_{k=1}^{L-1+m} c_k B_k(t, \tau)$$

where $L-1$ is the number of interior and m is the order of the polynomial.

The B-spline basis system has a property that is often useful: the sum of the B-spline basis function values at any point t is equal to one. This is because all the other basis functions go to zero at these end points. Also, since each basis function peaks at a single point, it follows that the value of a coefficient multiplying any basis function is approximately equal to the value of the spline function near where that function peaks. Indeed, this is exactly true at the boundaries.

2.2.2 Validation criterion

The choice of the basis function as well as the choise of the parameter number of basis for the data is very important. However, there are no general principals that would enable an optimal choice. The objective of the study and the data usually determine the decision of the basis. It is common practice to use the Fourier basis for periodic data and the B-spline basis for non-recurrent data. Cross-validation (CV) and Generalized Cross-validation are two among several selection criteria that may be used to select the parameter $\nu = (K, \lambda)$. Those two criteria can help us select a suitable number of basis $\nu_1 = K$ and also include the penalty parameter $\nu_2 = \lambda$ in the selection process. Cross-validation criteria is defined as follows ([145]):

$$CV(\nu) = \frac{1}{n} \sum_{i=1}^{n} \frac{(y_i - \hat{r}^{\nu}_{-i}(t_i))^2}{1 - S_{ii}} w(t_i)$$

where $\hat{r}^{\nu}_{-i}(x_i)$ is prediction at the point t_i obtained by omitting i pair (t_i, y_i) and $w(t_i)$ is the weight at the point t_i. S_{ii} is the i diagonal element of the smoothing matrix S ($\nu = trace(S)$).

The Generalized Cross-validation is defined by:

$$GCV(\nu) = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{r}^{\nu}_{i}(t_i))^2 w(t_i) \Xi(\nu)$$

where Ξ denotes the type of the penalizing function. The folowwing types of the Ξ function may be used ([74]):

- Generalized cross-validation (GCV): $\Xi(\nu) = (1 - tr(S)n^{-1})^{-2}$
- Akaike's Information Criteria (AIC): $\Xi(\nu) = exp(2tr(S)n^{-1})$
- Finite Prediction Error (FPE): $\Xi(\nu) = (1 + tr(S)n^{-1})/(1 - tr(S)n^{-1})$
- Shibata' model selector (Shibata): $\Xi(\nu) = (1 + 2tr(S)n^{-1})$
- Rice's bandwidth selector (Rice): $\Xi(\nu) = (1 - 2tr(S)n^{-1})^{-1}$

2.2.3 Descriptive statistics

First steps of any data analysis consists of estimating means and standard deviations. As in the univariate case, there are functional versions of these statistics. Let $x_i(t)$ $i = 1, \ldots, N$ be a sample of curves or functions fit to data. The sample mean and variance for functional data are calculated as follows:

$$\bar{x}(t) = \frac{1}{N} \sum_{i=1}^{N} x_i(t), \tag{2.1}$$

$$s(t) = \frac{1}{N-1} \sum_{i=1}^{N} [x_i(t) - \bar{x}(t)]^2. \tag{2.2}$$

Another useful metric is the bivariate covariance function $\nu(t, s)$ which specifies the covariance between curve values $x_i(t)$ and $x_i(s)$ at times t and s, respectively:

$$\nu(t, s) = \frac{1}{N-1} \sum_{i=1}^{N} [x_i(t) - \bar{x}(t)][x_i(s) - \bar{x}(s)]. \tag{2.3}$$

Different measures can also be used to summarize the functional data that are known as depth measures. The depth is a concept which measures how deep a data point is in the sample. In the univariate case, the median would tipically be the deepest point of clouds of points. However, in the functional data there is more depth measures.

One of the depth measures (also known as Integrated depth) is based on the meadian (see Fraiman and Muniz (2001)). For every $t \in [0, 1]$, let $F_{N,t}$ be the empirical distribution of the sample $x_i(t)$, $i = 1, \ldots, N$ and let $z_i(t)$ denote the univariate depth of the data $x_i(t)$ in this sample, given by $D_i(t) = 1 - |1/2 - F_{N,t}(x_i(t))|$. Then, calculation is made for $i = 1, \ldots, N$:

$$I_i = \int_0^1 D_i(t) \mathrm{d}t \tag{2.4}$$

and observations $x_i(t)$ are ranked according to the values of I_i.

Another depth measure is based on how the surrounded curves are in respect to a metric or semi-metric distance, selecting the trajectory most densely surrounded by other trajectories of the process (Cuevas et al. (2007)). The population h-depth of a datum z is given by the function:

$$f_h(z) = E(K_h(\|z - x\|)) \tag{2.5}$$

where x is the random element describing the polulation, $\|.\|$ is a suitable norm and $K_h(t)$ is a re-scaled kernel with the tuning parameter h. Given a random sample $x_i(t)$ $i = 1, \ldots, N$, the empirical h-depth is defined as:

$$\hat{f}_h(z) = \frac{1}{N} \sum_{i=1}^{N} K_h(\|z - x_i\|) \tag{2.6}$$

Cuevas et al. (2002) [39] analyzed depth measure which is calculated through random projections based on Tukey depth. Given a sample $x_i(t)$, $i = 1, \ldots, N$ and a random direction a (independent from the x_i), the projected data along this direction must be calculated. Then, the sample depth of a datum x_i is defined as the univariate depth of the corresponding one-dimensional projection (expressed in terms of order statistics so that the median is the deepest point). When the sample is made of functional data, it is assumed that x_i belongs to the Hilbert space $\mathcal{L}_2[0, 1]$, so that the projection of a datum x is given by the standard inner product $\langle a, x\rangle = \int_0^1 a(t)x(t)dt$. In the finite-dimensional case, the projection of $x = (\xi_1, \ldots, \xi_d)$ along the direction a is evaluated by the usual Euclidean inner product $\langle a, x\rangle = a_1\xi_1 + \cdots + a_d\xi_d$.

Another version is calculated via random projections of the curves and their derivatives. The basic idea is to use the method of random projections simultaneously (for the functions and their derivatives), thus incorporating the information on the function smoothness provided.

Here only a few depth measures are provided that can be used to calculate the depth of the functional data.

2.2.4 Functional principal component analysis

Principal component analysis (PCA) is often used after the analysis of descriptive statistics. PCA shows what are primary modes of variation in the data and how many of them are significant. Eigenvalues of the bivariate variance-covariance function $\nu(t,s)$ are indicators of the importance of these principal components. In principal, they are the same as in the multivariate case. Plotting eigenvalues may help to choose how many principal components are required to produce a reasonable summary of the data. A functional principal component analysis has an eigenfunction associated with each eigenvalue rather than eigenvector. These eigenfunctions describe major variational components.

Functional principal component analysis is designed to explain the functional data through a combination of orthonormal variables that satisfy the property to maximize their variance. The task is to find the function $\xi(t)$:

$$\rho_\xi(x_i) = \int \xi(t) x_i(t) \mathrm{d}t \tag{2.7}$$

that has the largest variation. Function $\xi(t)$ has a size restriction, i.e. it is required that $\int \xi^2(t)\mathrm{d}t = 1$.

The score of the principal component associated with the weigth ξ is the value:

$$\mu = \max_\xi \Big\{ \sum_i \rho_\xi^2(x_i) \Big\} \qquad \text{subject to} \qquad \int \xi^2(t)\mathrm{d}t = 1. \tag{2.8}$$

In the standard terminology, μ and ξ are called as the largest eigenvalue and eigenfunction, respectively, of the estimated variance-covariance function $\nu(t,s)$.

In the functional PCA, as in the multivariate case, a nonincreasing sequence of eigenvalues $\mu_1 \geq \mu_2 \geq \dots \mu_k$ can be constructed stepwise. It can be done by requiring each new eigenfunction to be orthogonal to those computed on previous steps:

$$\int \xi_m(t)\xi_l(t)\mathrm{d}t = 0, \quad j = 1, \dots, l-1 \quad \text{and} \quad \xi_l^2(t)\mathrm{d}t = 1. \tag{2.9}$$

The functional data can be rewritten as a decomposition in the finite orthonormal basis:

$$\hat{x}_i(t) = \sum_{i=1}^{k} \mu_{ik}\xi_k(t). \tag{2.10}$$

2.3 Cluster analysis: a case of banking ratios

2.3.1 Time series clustering methodology

In this section, the time series data under consideration and some measures of dissimilarities of time series, used to cluster the data, are presented.

Data

Six bank-specific variables were taken in our clustering exercise. We included three profitability measures: return on average assets (ROAA), return on average equity (ROAE), net interest margin (NIM); operational efficiency measure - cost to income ratio (CIR); credit quality measure - loan loss provisions over total gross loans (LLP), and bank riskiness measure - total capital ratio (CAR). These bank-specific measures are the main variables, which the describe situation in the banking sector.

In this study, we used annual unconsolidated bank accounts data, covering the time period from 1999 to 2013. A dataset is obtained from Bureau van Dijk *Bankscope* database and includes all commercial, savings, and cooperative banks from the European Union countries. These institutional bank types are mainly focused on financial intermediation. Therefore, we do not include the data from investment banks or other bank types as their business model is essentially different from commercial, savings, and cooperative banks. The preliminary sample consists of six bank-specific variables from 2800 banks.

We needed to edit our data in the following ways. First of all, we excluded all banks with the missing data entries, i.e. we left only those banks that had complete data for the years 1999-2013. Secondly, we excluded banks that had extreme values or large unexplained shifts in the values of variables. The final dataset varied from 260 banks for the capital adequacy ratio to 1332 banks for the ROAA variable. For example, Fig. 2.1 shows our sample of the capital adequacy ratio.

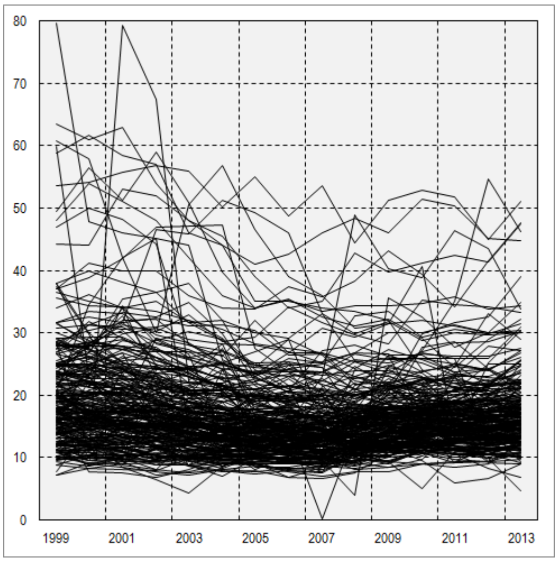

Figure 2.1: Capital adequacy ratio

Source: Bankscope data and authors' calculation.

The return on average assets is usually used as the main bank profitability variable. The ROAA is calculated as the ratio of net profits over average total assets. This ratio shows bank's ability to generate profits from all the activities related to their assets. Average assets are used to calculate the ratio, because they help to capture any changes in assets that occured during the fiscal year. Golin (2001) [70] describes ROAA as the key measure to evaluate bank's profitability.

The second measure of profitability is the return on average equity. ROAE is the net profits expressed as the percentage of average equity. This ratio gives information about the return to shareholders on the equity. Banks usually report both ROAA and ROAE ratios to indicate their profitability. The main difference between these two ratios is that ROAE does not take into account the risk that is associated with a higher leverage. Thus, banks with a higher equity (lower leverage ratio) generally report a lower ROAE, but a higher ROAA.

The third measure of profitability is net interest margin. The NIM ratio is defined as interest income minus interest expense divided by average interest-bearing assets. This ratio is narrower than the other two profitability measures as it focuses only on the profit earned from interest rate related activities.

In our study, we also include cost to income ratio, which shows the efficiency of the bank performance. The CIR is calculated as operating costs divided by the total generated revenue. This ratio is a measure of bank's ability to turn resources into revenue. Changes in CIR can highlight potential problems: if costs are rising at a higher rate than income, CIR will rise from one period to the next.

Credit portfolio quality is an important aspect of overall bank performance. Therefore we analyze loan loss provisions over the total gross loans ratio. The loan loss provisions are taken from a bank's income statement. A higher LLP ratio indicates problems in the credit portfolio and also potential problems on bank's stability.

The capital adequacy ratio is defined as bank's total capital, expressed as a percentage of its risk-weighted assets. CAR determines the capacity of the bank to meet potential losses from credit risk, market risk, operational risk, and others. This ratio ensures that the banks do not expand their business without having adequate capital. The capital adequacy ratio helps us to measure the riskiness of the banking sector. A higher CAR implies a more stable banking system.

Table 2.1: Descriptive statistics

	Mean	**Median**	**Standard deviation**	**Min**	**Max**
ROAA	0.41	0.29	0.76	-18.8	11.21
ROAE	5.10	4.46	6.77	-91.70	95.91
Net interest margin	2.74	2.65	0.98	-0.70	15.68
Cost to income ratio	67.80	68.25	12.43	5.26	186.36
Loan loss provisions/Gross loans	0.71	0.60	1.09	-7.69	17.62
Capital adequacy ratio	17.62	15.67	7.88	0.13	79.60

Source: Bankscope data and authors' calculation.

Descriptive statistics of the data are presented in Table 2.1. The descriptive statistics indicates the need to cluster banks, because data show high standard deviations compared to mean values, e.g. ROAA, ROAE. Cost to income ratio has min and max values greater than 3σ, which also motivates to separate banks into several groups.

The data under investigation have the form:

$$\boldsymbol{x}_t^{(i)} = \left(x_{1,t}^{(i)}, \ldots, x_{d,t}^{(i)}\right), \quad t = 1, \ldots, T; \quad i = 1, \ldots, N.$$

Here the index $i = 1, \ldots, N$ corresponds to a bank, whereas the index t corresponds to time (years in our case), and d correspods to a bank-specific ratio. Since not all banks in our data set have all six ratios, we will mostly consider univariate clustering, i.e. we will cluster banks according to each ratio separately. In addition, we will take profitability and efficiency ratios and consider multivariate clustering, based on four ratios.

Dissimilarity based time series clustering

As pointed out by Liao (2005) [100] and Batista et al. (2014) [17] the dissimilarity measure between two time series is one of the key choices in clustering to be made. The choice of the distance measure is more important than the choice of the clustering algorithm. In this section, we review six dissimilarity measures, used in time series clustering studies.

Euclidean distance
In general, any metric of the finite dimensional Euclidean space could be used as a measure of dissimilarities of two time series. In this research, we used the conventional Euclidean distance based measure as a starting method for clustering. Ding et al. (2008) [49] have showed that a simple Euclidean distance could outperform other dissimilarity measures in many cases.

Suppose that $x = (x_1, \ldots, x_T)$ represents values of some ratio of the bank i and $y = (y_1, \ldots, y_T)$ represents values of some ratio of the bank j $(i, j = 1, \ldots, N$ and $i \neq j)$. In our study $t = 1, \ldots, 15$. Euclidean distance is then described as follows:

$$D_{\lambda,\mathrm{EUCL}}(x,y) = \Big(\sum_{t=1}^{T}((x_t - \lambda x_{t-1}) - (y_t - \lambda y_{t-1}))^2\Big)^{1/2}, \tag{2.11}$$

where λ is a weighting parameter. The classical approach is to take $\lambda = 0$. Then the proximity depends on the closeness of values at the corresponding point of time. However, the distance $D_{\mathrm{EUCL}}(x,y) := D_{0,\mathrm{EUCL}}(x,y)$ does not take into account the growth rates of the vectors x and y. Therefore, we also considered the dissimilarity with $\lambda = 1$. To be more precise we applied $D_{\Delta,\mathrm{EUCL}} = D_{0,\mathrm{EUCL}}(x,y) + D_{1,\mathrm{EUCL}}(x,y)$ in this study.

Adaptive dissimilarity index
Chouakria and Nagabhushan (2007) [33] introduced a dissimilarity index, which is based on an adaptive tuning function and addressed to cover both the behaviour and values proximity measures. They used the first order temporal correlation coefficient to evaluate the proximity between the dynamic behaviour of the series. This coefficient is defined as follows:

$$CORT(x,y) = \frac{\sum_{t=1}^{T-1}(x_{t+1}-x_t)(y_{t+1}-y_t)}{\sqrt{\sum_{t=1}^{T-1}(x_{t+1}-x_t)^2}\sqrt{\sum_{t=1}^{T-1}(y_{t+1}-y_t)^2}}.$$

Temporal correlation coefficient belongs to the interval $[-1,1]$. The value $CORT(x,y) = 1$ means that the series x and y at any time point show a similar dynamic behaviour, i.e. series decrease or increase with a similar growth rate and direction (similar behaviour). The value $CORT(x,y) = -1$ means that both series have a similar growth rate, but direction is opposite (opposite behaviour). The value $CORT(x,y) = 0$ implies that growth rates are stochastically linearly independent and there is no monotonicity between series x and y (different behaviour). The proximity of the values of two time series $D_{\mathrm{EUCL}}(x,y)$ is estimated using the Euclidean distance.

Dissimilarity index, proposed in [33], automatically modulates the proximity of the values according to the proximity of the behaviour. This index is defined by:

$$D_{\mathrm{CORT}}(x,y) = \phi_k[CORT(x,y)] \cdot D_{\mathrm{EUCL}}(x,y), \tag{2.12}$$

where $\phi(u)$ is the exponential adaptive tuning function:

$$\phi_k(u) = \frac{2}{1 + e^{ku}}, \quad k \geqslant 0.$$

The adative tuning function decreases the weight of the proximity between values when the temporal correlation increases from 0 to 1. And it works vice versa when the correlation decreases from 0 to -1. In the case $CORT(x, y) = 0$, i.e. time series show a different behaviour, the dissimilarity index is approximately equal to the value of $D_{\mathrm{EUCL}}(x, y)$. The parameter k modulates the contribution of the temporal correlation and the Euclidean distance to the dissimilarity index $D_{\mathrm{CORT}}(x, y)$.

A complexity-invariant distance measure
Batista et al. (2011, 2014) [16, 17] proposed a dissimilarity index that uses information about the complexity difference between time series x and y. The authors argued that many dissimilarity measures tend to place more complex pairs of time series further apart than the pairs of simple series. The complexity-invariant dissimilarity measure $D_{\mathrm{CID}}(x, y)$ is defined as follows:

$$D_{\mathrm{CID}}(x, y) = CF(x, y) \cdot D_{\mathrm{EUCL}}(x, y), \tag{2.13}$$

where $CF(x, y)$ is a complexity correction factor:

$$CF(x, y) = \frac{\max\{CE(x), CE(y)\}}{\min\{CE(x), CE(y)\}}, \quad CE(x) = \sqrt{\sum_{t=1}^{T-1} (x_{t+1} - x_t)^2}.$$

The complexity correction factor increases the distance between two time series, if there is a complexity difference between them. Furthermore, if time series have a similar complexity then the distance is approximatelly equal to $D_{\mathrm{EUCL}}(x, y)$.

The main idea of [16, 17] is that, if a time series is stretched to become a straight line, then a more complex time series would result in a longer line. Dissimilarity index $D_{\mathrm{CID}}(x, y)$ is parameter-free, simple and increased accuracy of clustering in several experiments accomplished in [16].

Autocorrelation based distance

Bohte et al. (1980) [21], Geleano and Peña (2000) [62] and several other authors used the estimated autocorrelation function to measure the distance between two time series. Suppose that $\hat{\rho}_x = (\hat{\rho}_{1,x}, \ldots, \hat{\rho}_{L,x})'$ and $\hat{\rho}_y = (\hat{\rho}_{1,y}, \ldots, \hat{\rho}_{L,y})'$ are the estimated autocorrelation vectors of x and y. Here L is such that $\hat{\rho}_{i,x} \approx 0$ and $\hat{\rho}_{i,y} \approx 0$ when $i > L$. The dissimilarity between two univariate time series can be measured by:

$$D_{\mathrm{ACF}}(x, y) = \sqrt{(\hat{\rho}_x - \hat{\rho}_y)' \Omega (\hat{\rho}_x - \hat{\rho}_y)},$$

where Ω is a weighting matrix.

If we take $\Omega = I$, i.e. uniform weights, then $D_{\mathrm{ACF}}(x, y)$ is the Euclidean distance between the estimated autocorrelation functions:

$$D_{\mathrm{ACFE}}(x, y) = \sqrt{\sum_{i=1}^{L} (\hat{\rho}_{i,x} - \hat{\rho}_{i,y})^2}. \tag{2.14}$$

Dynamic time warping distance

Berndt and Clifford (1994) [19] proposed dynamic time warping (DTW) to find patterns in time series. This distance measure is popular and widely used in the time series clustering literature. Let N be the set of all possible sequences of n pairs preserving the observation order in the form:

$$r = ((x_{a_1}, y_{b_1}), \ldots, (x_{a_n}, y_{b_n})), \tag{2.15}$$

where $a_i, b_j \in \{1, \ldots, T\}$ such that $a_1 = b_1 = 1$ and $a_n = b_n = T$, and $a_{i+1} = a_i$ or $a_i + 1$ and $b_{i+1} = b_i$ or $b_i + 1$, for $i \in \{1, \ldots, n-1\}$. Then dynamic time warping distance is defined by:

$$D_{\mathrm{DTW}}(x, y) = \min_{r \in N} \left(\sum_{i=1,\ldots,n} |x_{a_i} - y_{b_i}| \right). \tag{2.16}$$

Dynamic time warping dissimilarity measure allows time series to be stretched or compressed to recognize similar shapes.

2.3.2 Functional data clustering methodology

In this section, the data are considered as observations of curves, i.e. the random variables underlying data are countinuous time stochastic processes. To cluster the curves, a non-parametric method, using a specific distance or dissimilarities between functions, is applied. Besides widely used dissimilarity measures for the functions such as Hausdorff distance, L_2-distance or distance, based on functional principal components, we also consider a class of Hölder distances that take into account a certain type of growth rates of the curves. It is shown that this type of distances, in some cases, performs better compared with others.

Functional data

We assume that the data under investigation $\boldsymbol{x}^{(i)}(t) = \big(x_1^{(i)}(t), \ldots, x_d^{(i)}(t)\big)$, $t = 1, \ldots, T$, $i = 1, \ldots, N$, constitute observations of random curves:

$$\boldsymbol{X}^{(i)}(t) = (x_1^{(i)}(t), \ldots, x_d^{(i)}(t)), \quad t \in [0, T],\ i = 1, \ldots, N.$$

Moreover, we assume that the sampled curves are observed at discrete instants of time. Hence we have

$$\boldsymbol{x}_j^{(i)} = \boldsymbol{X}^{(i)}(j/T) + \varepsilon^i(j/T), \quad j = 1, \ldots, T.$$

We reconstruct the functions $\boldsymbol{x}^{(i)}(t), t \in [0, T]$ by smoothing techniques (see e.g. [132]), thus obtaining functional data

$$\widehat{\boldsymbol{x}}^{(i)}(t), \quad t \in [0, T],\ i = 1, \ldots, N,$$

which are a subject for the functional clustering analysis. It is worth mentioning that each function $\widehat{\boldsymbol{x}}^{(i)}$ is d-dimensional. In Fig. 2.2, we present an example of the 1-dimensional functional data under consideration.

Just like in the time series clustering, we apply the clustering methodology to 1-dimensional curves, i.e. we will cluster banks according to each ratio separately.

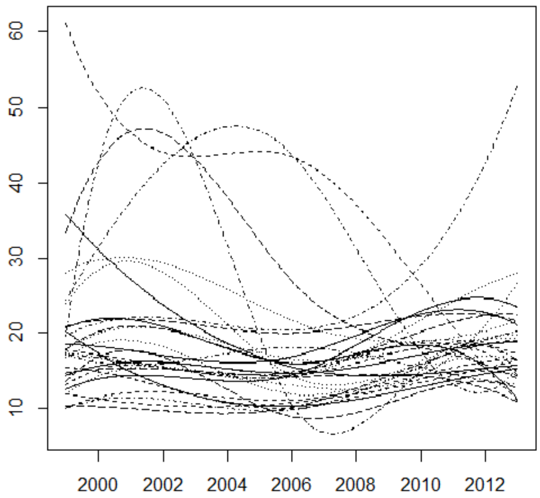

Figure 2.2: Capital adequacy ratio (smoothed using B-spline approximation)

Source: Bankscope data and authors' calculation.

Functional data dissimilarity measures

In this section, we review six dissimilarity measures, used in the functional data clustering.

Hausdorff distance between two curves
A distance between two curves can be measured by the Hausdorff distance. This distance measures the maximum distance from a point in one curve to the nearest point in the other curve. Suppose that $G(x) = \{(t, x(t)) : t \in [0, 1]\} \subset R^2$ and $G(y) = \{(t, y(t)) : t \in [0, 1]\} \subset R^2$ are graphs of the curves x and y, respectively. The Hausdorff distance $D_{\text{Hausdorff}}(x, y)$ is defined by:

$$D_{\text{Hausdorff}}(x, y) = \max\left\{\sup_{x\in G(x)} \inf_{y\in G(y)} D_{\mathcal{L}_2}(x, y), \sup_{y\in G(y)} \inf_{x\in G(x)} D_{\mathcal{L}_2}(x, y)\right\}, \tag{2.17}$$

where $D_{\mathcal{L}_2}$ is the Euclidean distance.

A formal definition of the Euclidean distance between functional data is:

$$D_{\mathcal{L}_2}(x,y) = \sqrt{\left(\frac{1}{\int_a^b w(t)\mathrm{d}t}\int_a^b |x(t)-y(t)|^2\cdot w(t)\mathrm{d}t\right)},$$

where $w(t)$ is a weighting function.

Distance, based on B-spline approximation
Ferraty and Vieu (2006) [56] proposed a two-stage approach for functional data clustering. The proximity between two curves x and y could be estimated using

$$D^q(x,y) = \sqrt{\frac{1}{T}\int_T (x^{(q)}(t)-y^{(q)}(t))^2\mathrm{d}t},$$

where $x^{(q)}(t)$ is the q-th derivative of x. In the first stage, [56] used B-spline to approximate functional data.

Consider a B-spline basis as a set of functions $B=\{b_1,\dots,b_N\}$. Then, the derivatives of the curves approximated by n elements of B-spline are expressed as: $\hat{x}^{(q)} = \sum_{n=1}^N c_n B_n^{(q)}$. The second stage is the proximity measure, expressed as:

$$D_{\mathrm{B}}(x,y) = \sqrt{\frac{1}{T}\int_T (\hat{x}^{(q)}(t)-\hat{y}^{(q)}(t))^2\mathrm{d}t}. \tag{2.18}$$

In our analysis, we have considered two cases, i.e. we estimated a distance with $q=0$ (D_{BASIS}) and $q=1$ (D_{DERIV}).

Distance, based on continuity properties of curves
In this work, we introduce two dissimilarity measures. One of them is based on the Hölderian property of a function. The dissimilarity measure is constructed from two parts. The first part shows how close the functions are to each other. In this part, we calculate supremum between two curves. The second part shows how similar curves are changing together. Hölder distance measure is defined by:

$$D_{\text{Hölder}}(x,y) = \sup_t |x(t)-y(t)| + \sup_{t\neq s}\frac{|(x(t)-y(t))-(x(s)-y(s))|}{|t-s|^\alpha}, \tag{2.19}$$

where the number $\alpha\in(0,1]$ is called the Hölder exponent.

The second dissimilarity measure, proposed by us is also constructed from two parts. The first part shows how B-spline approximations are close together. The second part uses the q-th derivative to capture how close is the change of curves. The dissimilarity measure is estimated as follows:

$$D_{\mathrm{SUP}}(x, y) = \sup_t |x(t) - y(t)| + \sup_t |x^{(q)}(t) - y^{(q)}(t)|. \tag{2.20}$$

This distance measure takes into account both the closeness and behaviour of the data.

Distance, based on functional principal components
Functional principal components give another tool to reduce a dimension of functional data. This distance measure is also considered as a two-stage approach. The functional data can be decomposed in a finite orthonormal basis: $\hat{x}_i(t) = \sum_{i=1}^{K} f_{ik}\xi_k(t)$, where f_{ik} is the score of the principal component $\xi_k(t)$. In this case, the distance between two curves is calculated as:

$$D_{\mathrm{FPCA}}(x, y) = \sqrt{\sum_{k=1}^{K} \left(\sum_{j=1}^{T} (f_x(t_j) - f_y(t_j)) \right)^2}, \tag{2.21}$$

where f_x and f_y are scores of the principal component of the curves x and y, respectively.

2.3.3 Multivariate clustering

In the previous two sections, we considered the data as univariate time series or 1-dimensional curves. In this section, we interpret the data as N observations of d-dimensional time series or d-dimensional curves.

Multivariate Euclidean distance
The Euclidean distance of univariate time series can be easily expended to the multivariate case. This disssimilarity measure is expressed as:

$$D_{\lambda,\mathrm{EUCL}}(\boldsymbol{x}, \boldsymbol{y}) = \Big(\sum_{j=1}^{d} \sum_{t=1}^{T} ((x_{jt} - \lambda x_{j,t-1}) - (y_{jt} - \lambda y_{j,t-1}))^2 \Big)^{1/2}. \tag{2.22}$$

The next two measures of dissimilarity of time series are obtained by introducing a certain correction of the Euclidean distance. We address to the adaptive dissimilarity index, introduced in [33], and to the complexity invariant distance measure introduced in [16], [17]. We define the analogues for the d-dimensional time series.

Multivariate adaptive dissimilarity index
An extended adaptive dissimilarity index for the d-dimensional time series $\boldsymbol{x}$ and $\boldsymbol{y}$, is defined as follows:

$$CORT(\boldsymbol{x}, \boldsymbol{y}) = Q_{\boldsymbol{x}}^{-1/2} Q_{\boldsymbol{x},\boldsymbol{y}} Q_{\boldsymbol{x}}^{-1/2},$$

where

$$Q_{\boldsymbol{x}} = \sum_{t=1}^{T-1} (\boldsymbol{x}_{t+1} - \boldsymbol{x}_t)^\tau (\boldsymbol{x}_{t+1} - \boldsymbol{x}_t), \quad Q_{\boldsymbol{x},\boldsymbol{y}} = \sum_{t=1}^{T-1} (\boldsymbol{x}_{t+1} - \boldsymbol{x}_t)^\tau (\boldsymbol{y}_{t+1} - \boldsymbol{y}_t).$$

Let $\lambda_{max}(CORT(\boldsymbol{x}, \boldsymbol{y}))$ denote the largest eigenvalue of the matrix $CORT(\boldsymbol{x}, \boldsymbol{y})$. Multivariate addaptive dissimilarity index is then expressed as:

$$D_{\text{CORT}}(\boldsymbol{x}, \boldsymbol{y}) = \phi_k[\lambda_{max}(CORT(\boldsymbol{x}, \boldsymbol{y}))] \cdot D_{\text{EUCL}}(\boldsymbol{x}, \boldsymbol{y}), \tag{2.23}$$

where $\phi(u)$ is an exponential adaptive tuning function.

Multivariate complexity-invariant distance measure
Let us consider that the complexity estimate $CE(\boldsymbol{x})$ for a multivariate case can be written as follows:

$$CE(\boldsymbol{x}) = \Big(\sum_{j=1}^{d} \sum_{t=1}^{T-1} (x_{jt+1} - x_{jt})^2 \Big)^{1/2}.$$

Then a complexity-invariant dissimilarity measure is expressed as:

$$D_{\text{CID}}(\boldsymbol{x}, \boldsymbol{y}) = CF(\boldsymbol{x}, \boldsymbol{y}) \cdot D_{\text{EUCL}}(\boldsymbol{x}, \boldsymbol{y}), \tag{2.24}$$

where $CF(\boldsymbol{x}, \boldsymbol{y})$ is a complexity correction factor.

Multivariate case of other dissimilarity measures

Other dissimilarity measures (D_{ACF}, D_{DTW}, $D_{\text{Hausdorff}}$, $D_{\text{Hölder}}$, D_{BASIS}, D_{DERIV}, D_{SUP}, D_{FPCA}) for the multivariate case are calculated using the following expression:

$$D_{\text{M}}(\boldsymbol{x}, \boldsymbol{y}) = \Big(\sum_{j=1}^{d} [D_j(x, y)]^2 \Big)^{1/2}, \tag{2.25}$$

where $D_j(x, y)$ is a coordinate-wise dissimilarity measure.

2.3.4 Clustering algorithm and validity assessment

In this section, the clustering algorithm and clustering validity indices, used in this work, are presented.

Clustering algorithm

In this study, we used the conventional agglomerative hierarchical clustering algorithm. This method works by clustering time series into a tree of clusters (dendrogram). In the beginning each observation is assigned to its own cluster. Afterwards, clustering algoritm works iteratively, at each step joining two most similar clusters into larger and larger ones. This process continues until a single cluster is formed or until certain termination conditions are satisfied. The complete linkage algorithm, which was applied in this study, measures the similarity between two clusters as the similarity between the farthest pair of data belonging to different clusters.

The iterative procedure of the complete linkage algorithm can be written as follows:

1. At the start each element is assigned to its own cluster. The level of dendrogram is set to $L(0) = 0$ and the sequence number $n = 0$.
2. One finds a pair of clusters, say C_i and C_j, with the lowest dissimilarity ($D(C_i, C_j)$). Set the sequence number to $n = n+1$. These two clusters are then joined at the level $L(n) = D(C_i, C_j)$.

3. The dissimilarity matrix is updated by reducing its order by one. In complete linkage clustering the distance between clusters is the distance between the farthest pair of points, i.e. $D(C_i \cup C_j, C_k) = \max_{i,j \in C_i \cup C_j, k \in C_k}(D(i,k), D(j,k))$.

4. Steps 2 and 3 are repeated until a single cluster is obtained ($N-1$ times).

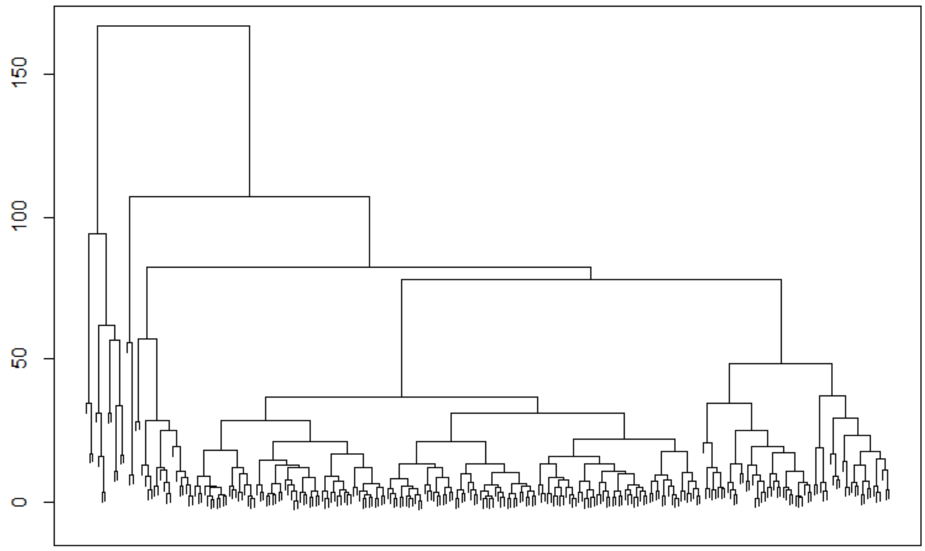

Figure 2.3: Dendrogram using D_{FPCA} distance measure: Capital adequacy ratio

Source: Bankscope data and authors' calculation.

Fig. 2.3 gives an example of dendrogram from the hierarchical clustering algorithm. We have chosen the hierarchical clustering algorithm, because it is more efficient in dealing with outliers than partitional algorithms.

Cluster validity assessment

The next step is to choose the optimal number of clusters. We calculated three different measures which are used for validating the results of clustering analysis in clustering literature. Dunn index and Caliński and Harabasz index are used for choosing the number of clusters. Meanwhile, average silhouette width helps us to choose the number of clusters and to compare different distance measures.

Average silhouette width

Rousseeuw (1986) [136], Kaufman and Rousseeuw (1990) [94] introduced the average silhouette width as a measure to evaluate clustering. Let $\mathcal{C} = \{C, , \ldots, C_K\}$ be a particular clustering partition of N observations into K disjoint clusters. The silhouette value measures the degree of confidence in the clustering of an observation. For the observation i, the value is defined by:

$$S(i) = \frac{b_i - a_i}{\max(b_i, a_i)}, \tag{2.26}$$

where a_i is the average dissimilarity of i to all other objects of C_i (cluster containing observation i) and b_i is a minimum of the average dissimilarity between i and the elements of the other cluster which is different from C_i. Thus:

$$a_i = \frac{1}{n(C_i)} \sum_{j \in C_i} D(i,j), \quad b_i = \min_{C_m \in \mathcal{C} \setminus C_X} \sum_{j \in C_m} \frac{D(i,j)}{n(C_m)}$$

where $D(i,j)$ is a dissimilarity measure and $n(C)$ is the cardinality of the cluster C. The average silhouette value is in the interval $[-1, 1]$. A value close to 1 means that the particular clustering partition is well classified, and the value close to -1 means that observations are misclassified.

Dunn index

Another cluster validity index was proposed by Dunn (1974) [52]. This index tries to identify compact and well separated clusters. The Dunn index is calculated as a ratio of the smallest dissimilarity between observations not in the same cluster to the largest intra-cluster dissimilarity:

$$Dunn(\mathcal{C}) = \frac{\min_{C_m, C_l \in \mathcal{C}, C_m \neq C_l} \left(\min_{i \in C_m, j \in C_l} D(i,j) \right)}{\max_{C_n \in \mathcal{C}} diam(C_n)} \tag{2.27}$$

where $diam(C_n)$ is the diameter of a cluster, i.e. maximum distance between observations in the cluster C_n.

If observations are in the compact and well separated clusters, then the value of Dunn index should be large, because the dissimilarity between the clusters is expected to be large and the diameter of the clusters is expected to be small.

Caliński and Harabasz index
Caliński and Harabasz (1974) [27] introduced a criterion that can be used to determine the number of clusters in cluster analysis. Milligan and Cooper (1985) [109] showed that the $CH(k)$ index works in many cases.

Caliński and Harabasz index is defined by the following expression:

$$CH(k) = \frac{\mathbf{B}_k(N-k)}{\mathbf{W}_k(k-1)}, \tag{2.28}$$

where $\mathbf{W}_k$ is the overall within-cluster variance:

$$\mathbf{W}_k = \sum_{h=1}^{k} \frac{1}{|C_h|} \sum_{i,j \in C_h} D(i,j)^2$$

and $\mathbf{B}_k$ is the overall between-cluster variance:

$$\mathbf{B}_k = \frac{1}{N} \sum_{i,j=1}^{N} D(i,j)^2 - \mathbf{W}_k.$$

Well separated clusters have large $\mathbf{B}_k$ and small $\mathbf{W}_k$. Therefore, a larger value of $CH(k)$ indicates a better data partition.

2.3.5 Clustering results

In this section, we present the results from the banking data clustering exercise. We performed the clustering experiment on 6 bank performance ratios. We used 12 dissimilarity measures for each ratio to assess the closeness of banks. Then, an agglomerative hierarchical clustering algorithm was applied to group banks into clusters. Since the true number of clusters is unknown, we divided banks into 2, 4, 6, 8, 10 and 20 clusters. Finally, cluster validity indices were calculated, which are presented in Appendix.

In the univariate case, based on the average silhouette width, we can conclude that there is no dissimilarity measure that could be the best one for all ratios. For example, for some ratios (ROAA, ROAE) the dissimilarity measured between the first derivatives gives a high average silhouette width (see Appendix, Fig. C.1 and Fig. C.2). However, this distance measure performs poorly for the capital adequacy ratio, especially if we consider more than two clusters (Appendix, Fig. C.6). Similarly, a distance measure, based on supremum between two curves, gives good clustering results for loan loss provisions over the gross loan portfolio ratio (Appendix, Fig. C.5), but it is not suitable to cluster banks, if we use ROAE. Our proposed distance measures ($D_{\text{Hölder}}$, D_{SUP}) performed well in this study. A dissimilarity measure, based on L_∞ norm between B-spline approximations and their first derivatives, showed the highest average silhouette width for ROAA. Meanwhile, a distance measure, based on Hölder's exponent, provided the best results, if we take CAR ratio. One thing that could be noted is that dissimilarity measures, based on functional data properties ($D_{\text{Hölder}}$, D_{SUP}, D_{BASIS}, D_{DERIV}, D_{FPCA}), performed better than the measures which use time series properties. In our case, distance measures, based on autocorrelation and *CID* yields the lowest average silhouette width values (see e.g. Appendix C, Fig. C.3). It could be also noted that a simple Euclidean distance performed rather well for clustering banking data. This result is consistent with [49], who also found that the Euclidean distance provides a relatively good clustering outcome. Thus, we can conclude that it is useful to use dissimilarity measures which employ the functional data properties.

From these results we also see that the average silhouette width usually is the highest, if we take two clusters. Only in some cases a higher value is obtained with 4 clusters, for example, ROAA, if we measure the dissimilarity with D_{BASIS} distance or functional principal components and CIR if we use the Euclidean distance (see Appendix, Fig. C.4). In many cases, the average silhouette width drops significantly, if we consider more than four clusters. However, other clustering validity indices yield mixed results. In some cases, for instance, if we take CAR ratio or LLP, Dunn index mostly shows that we should consider two ar four bank clusters (see Appendix, Table C.1). Caliński and Harabasz index also mostly suggests to use 2 or

4 clusters. Similar results are with LLP ratio, where both indices indicate that the best option is to choose two or four clusters (see Appendix, Table C.2). An interesting case is with CIR ratio. If the Dunn index suggests that we should consider up to 20 clusters, Caliński and Harabasz index gives the opposite result, i.e. the index is highest, if we take 2 or 4 clusters. Such a discrepancy may arise due to the fact that CH index calculates average variance, meanwhile Dunn index takes the maximum distance between observations. Since our data are noisy and clusters are not well-separated, few oservations may have a strong impact on the Dunn index values.

The analysis of data, grouped into 20 clusters, revealed that, in many cases, there are few clusters formed by a larger number of banks and other clusters are formed only by few banks. For example, Fig. 2.4 shows 6 clusters that are formed from a larger number of banks, using D_{FPCA} distance measure. From the figure we can distinguish a few patterns in the development of the capital adequacy ratio: one group of banks kept their CAR ratio more or less at the same level, another group showed a decreasing trend, and the third group increased CAR significantly after the Global financial crisis in 2009. Other clusters included only 1 or 3 banks. The clustering results show that we can extract 6 larger clusters, if we take ROAA or ROAE. Taking capital adequacy ratio results in 6 clusters, whereas LLP in 8 clusters. Banks could be clustered into 10-11 groups, if we take NIM and into 12 larger clusters if we take CIR. The clusters formed by few banks could be considered as outliers. Therefore, only the larger clusters could be further examined in the development of macroprudential policy instruments.

In a multivariate case, we take three profitability measures and the efficiency ratio (CIR) to form d-dimensional time series. Furthermore, we normalized data of each ratio to take into account differences between the values of each ratio. We take these four ratios because most of the banks in our sample had data about them. Basing on the average silhouette width, we see that it is reasonable to cluster banks into groups, based on few ratios at the same time. Most of the values of the ASV index are comparable with the univariate cases. Another finding is that in the multivariate case it is important to take into account both clossenes and behaviour of time series, because $D_{\text{Hölder}}$, D_{SUP} and $D_{\Delta,\mathrm{EUCL}}$ give better clustering results.

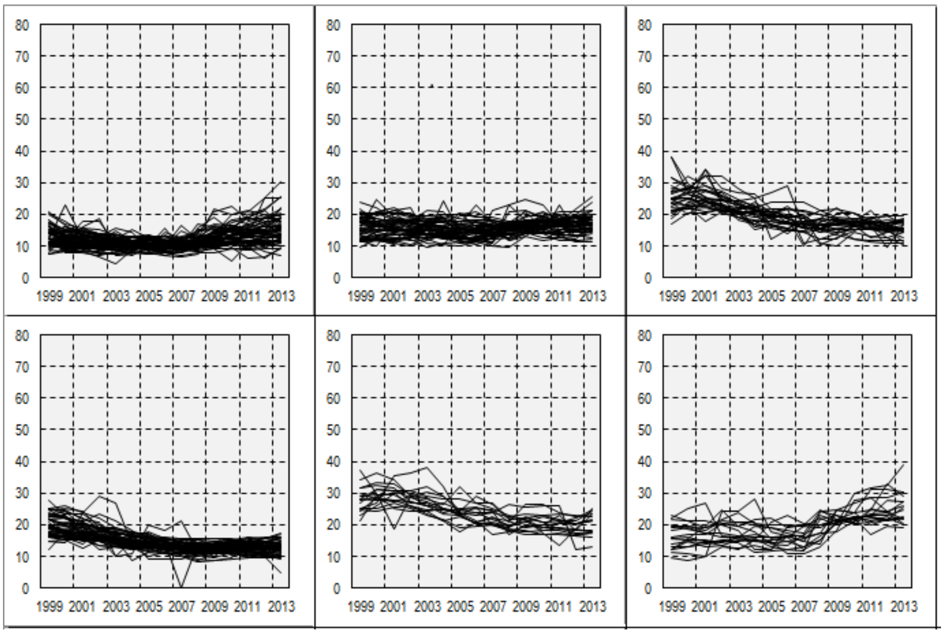

Figure 2.4: Clustering results using D_{FPCA} distance measure: Capital adequacy ratio

Source: Bankscope data and authors' calculation.

Other than in a univariate, case where $D_{\Delta,\text{EUCL}}$ does not improve the results of Euclidean distance (D_{EUCL}), in a multivariate case, a change in ratios improves the clustering results. In a multivariate case, ASV mostly suggests two clusters, but based on Dunn index and CH index we should take a larger number of clusters. As the multivariate clustering revealed it is possible to find homogeneous groups of banks, taking into account all the ratios. Of course, if we analyze separate ratios of the clustered banks, we see that some banks would not be grouped into the same cluster in a univariate case. For this result, there are also economic reasons as banks might reach a similar ROAA ratio having a different share of equity and/or performance efficiency. Nevertheless, multivariate dissimilarity measures, proposed in this paper, might be useful in other cases.

2.3.6 Conclusions of the cluster analysis

There are two main purposes of this part. The first one is to compare various dissmilarity measures that are used to cluster time series data. We considered dissimilarity measures based on the raw time series data and measures which take into account some properties of time series (e.g. autocorrelation). Another group of dissimilarity measures is based on the functional data properties. Furthermore, we analyzed clustering based on multivariate data. The second purpose is to consider clustering of banks according to their performance ratios and to find a proper number of clusters. We took 6 ratios that are commonly used to compare the performance of banks. Three ratios measure profitability: return on average assets, return on average equity, and net interest margin. Cost to income shows the efficiency, capital adequacy ratio shows how much risk a bank is taking and loan loss provisions show the quality of loan portfolio.

The results of the cluster analysis show that the choice of a dissimilarity measure may significantly change the way banks are grouped. The same could be addressed to the choice of the number of clusters, which depends on the clustering validation method. As pointed out Batista et al. (2014) [17] a dissimilarity measure is the key component in clustering. Therefore, it is a good option to take few distance measures and compare the results. Furthermore, basing on the average silhouette width, we may conclude that no dissimilarity measure worked best for all ratios. In some cases the dissimilarity measured between the first derivatives or that based on the Hölder condition gives a high average silhouette width, in other cases, the distance measure based on functional principal components gives better clustering results. However, clustering methods, based on the functional data properties mostly outperformed distance measures based on time series properties. In our study, D_{ACF} and D_{CID} provided relatively poor clustering results for many ratios. Another conclusion could be that a simple Euclidean distance is a relatively good distance measure for clustering banking data. The third conclusion, based on the average silhouette width, is that both proposed measures, $D_{\mathrm{Hölder}}$ and D_{SUP}, were among the best for clustering banking data.

The choice of the number of clusters is not that clear as well. For some banking ratio clustering validation indices suggest a low number of clusters. If we consider CAR, LLP or CIR, then the optimal number of clusters would be 2 or 4. But if we take profitability ratios, then the results are mixed and the number of clusters could be chosen by the expert judgement.

Division of banks into 20 clusters has revealed that there are few larger clusters and other clusters are formed by a small number of banks. According to different banking ratios, there are from 6 to 12 clusters. The larger clusters could be further analyzed and used to develop some new macroprudential tools.

Multivariate clustering has revealed that it is reasonable to group banks into clusters according to profitability and efficiency ratios. The average silhouette width is comparable with univariate cases. Indeed, if we analyze separate ratios of the clustered banks, we see that some banks would not be grouped in a univariate case. Nevertheless, in some data samples multivariate clustering might be useful as it divides time series, based on some features.

2.4 Forecasting with functional data: case study

In any specific practical application, usually it is difficult to argue on theoretical grounds which forecasting approach - top-down or bottom-up - should be correct. Therefore, this question is usually settled empirically by trying both approaches.

The top-down versus bottom-up forecasting problem appears in many fields of time series data forecasting, for example, manufacturing demand ([149], [148] or [139]), sales forecasts ([135]), tourism data ([12]), economic data ([51], [88]), energy-economy ([134]), or crime forecast ([81]). In general, it is not clear which of the methods is better. While Widiarta et al. (2009) ([148]) prove that, under some restrictions, both methods are equally efficient in terms of MSE, other authors ([149], [139]) argue that neither approach should be preferred a priori in any empirical application and that an appropriate aggregation level depends on the underlying data generation process. Meanwhile, Duarte and Rua (2007) [51] have found that bottom-up approach is better for a short-term forecasting, but the required disaggregation level decreases over the forecast horizon. In some applications time series data might be naturally organized in a hierarchical structure, using the attributes such as the product type, geographical location, etc. The hierarchical forecasting combines the top-down approach with bottom-up and ensures that bottop-up forecasts add up to the top-down forecast ([12], [84]). Hyndman et al. (2011) [84] have proposed a method that independently forecasts all time series at all levels of the hierarchy and then uses a regression model to optimally combine and reconcile these forecasts. The authors argue that their approach provides better forecasts than that produced either by a top-down or bottom-up approach.

Functional data analysis is one of the fields in statistics that has attracted great attention in recent years from both theoreticians and practitioners. Functional data are defined as discrete observations of curves. Forecasting by means of functional data is used in many fields, for example, energy market ([68], [101], [28]), finance ([76]), environmental data ([13]), mortality and fertility ([83], [140]). Most of the studies are dealing with the temporal dependencies among functional data, i.e. functional autoregressive models (FAR) are applied. Usually their purpose is to forecast new curves. However, in our case, the purpose is to forecast the future of the particular stochastic process, i.e. the future development of a curve.

In this section, we consider the estimation and forecasting problems of the banking data. We analyze the capital adequacy ratio which determines the capacity of the bank to meet potential losses arising from credit risk, market, operational rink, and others. The capital adequacy ratio helps to measure the riskiness of the banking sector and is closely monitored and regulated by micro-prudential and macro-prudential supervisors. Therefore, the aggregate forecast is important to macro-prudential supervision as it shows a possible future development of the whole sector. The forecast of individual institutions is important to microprudential supervision which is responsible for the stability of each bank individually.

2.4.1 Theoretical models

Let us consider for each $j = 1, \dots, N$, a random function $X_j = (X_j(t), 0 \leq t \leq T)$. By choosing the reference time points $0 = \tau_0 < \tau_1 < \cdots < \tau_n = T$, for each j assume a random sample $(X_j(\tau_k), k = 0, 1, \ldots, n)$ and assume the following model:

$$X_j(\tau_k) = \beta_j X_j(\tau_{k-1}) + \boldsymbol{\gamma}_j' \boldsymbol{Z}_k + \varepsilon_{jk}, \quad k = 1, \ldots, n, \tag{2.29}$$

where $\boldsymbol{Z_k}$ is a common explanatory p-dimensional vector independent with white noise process $(\varepsilon_{jk}, j, k \geq 0)$: $\mathsf{E}\,\varepsilon_{jk} = 0$ and $\mathsf{E}\,\varepsilon_{ik}\varepsilon_{j\ell} = \sigma^2 \delta_{ij}\delta_{k\ell}$, where $\delta_{st} = 0$ if $s \neq t$ and $\delta_{ss} = 1$. We also assume that $\{\beta_j\}$ are independent identically distributed (iid) random variables with a common distribution in the interval $(-a, a)$, $\{\boldsymbol{\gamma}_j\}$ are independent identically distributed p-dimensional random vectors with a finite mean. We assume a joint independence of $\{\beta_j\}$, $\{\boldsymbol{\gamma}_j\}$, and $\{\varepsilon_j\}$.

We find unknown coefficients by the least squares method, thus taking

$$(\widehat{\beta}_j, \widehat{\boldsymbol{\gamma}}_j) := \operatorname*{argmin}_{(b, \boldsymbol{c})} \sum_{k=1}^{n} \Big[X_j(\tau_k) - b X_j(\tau_{k-1}) - \boldsymbol{c}' \boldsymbol{Z}_k \Big]^2. \tag{2.30}$$

One can interpret then $\widehat{\beta}_1, \ldots, \widehat{\beta}_N$, and $\widehat{\boldsymbol{\gamma}}_1, \ldots, \widehat{\boldsymbol{\gamma}}_N$ as random samples and perform appropriate statistical inference on distributional properties of random coefficients of the model (2.29) (see Appendix D for further discussions).

Define the aggregated process by:

$$\overline{X}_N(t) = N^{-1} \sum_{j=1}^{N} X_j(t), \quad t \in [0, T].$$

The following result describes the asymptotic behaviour of $\overline{X}_N(\tau_k)$ as $N \to \infty$.

Theorem 1. *Assume that $X_j(0) = 0$. Then, for each $1 \leq k \leq n$, it holds*

$$\lim_{N \to \infty} \mathsf{E}\left[|\overline{X}_N(\tau_k) - Y_k|^2 \right] = 0,$$

where

$$Y_k = \sum_{i=0}^{k-1} \mathsf{E}\,(\beta_1^i) \mathsf{E}\,(\boldsymbol{\gamma}_1') \boldsymbol{Z}_{k-i}.$$

From Theorem 1, we see that asymptotically the behaviour of finite dimensional distributions of the aggregated random functions $(\overline{X}_N(t), t \in [0,T])$ controls the explanatory variables $Z_1, \ldots, Z_p$ and coefficients $a_1, \ldots, a_p$ in a sense that in probability as N is large

$$(\overline{X}_N(\tau_k))_{1\le k\le n} \approx \Big(\sum_{i=0}^{k-1} a_i' \boldsymbol{Z}_{k-i}\Big)_{1\le k\le n}.$$

In the case, where $(X_j(\tau_k), k = 0, \ldots, n)$, follows random coefficient regression model

$$X_j(\tau_k) = \boldsymbol{\gamma}_j' \boldsymbol{Z}_k + \varepsilon_{jk}, \quad k = 1, \ldots, n,$$

we have

$$\overline{X}_N(\tau_k) \xrightarrow[n\to\infty]{\mathsf{P}} \mathsf{E}\,(\boldsymbol{\gamma}_1')\boldsymbol{Z_k}$$

by the law of large numbers.

Since the sample under investigation constitutes curves, another possibility for statistical analysis is to exploit the Karhunen-Loèv expansion: every centered square integrable process, say, $X(t), t \in [0,T]$, can be written as:

$$X(t) = \sum_{\ell=1}^{\infty} \xi_\ell \psi_\ell(t), \quad t \in [0,T],$$

where (ψ_ℓ) are orthonormal functions (we refer to [132], [80] for the background on functional data analysis). To implement this approach, we consider each X_j as a random element in the Hilbert space $L_2[0,T]$, endowed with the inner product $\langle f, g\rangle = \int_0^T f(t)g(t)dt$ and the norm $||f|| = \sqrt{\langle f, f\rangle}$ for $f, g \in L_2[0,T]$. Fix an orthonormal basis $(\psi_k, k \ge 1)$ in $L_2[0,T]$, and assume that for each k,

$$\langle X_j, \psi_k\rangle = \beta_j \langle X_j, \psi_{k-1}\rangle + \boldsymbol{\gamma}_j' \boldsymbol{Z}_k + \varepsilon_{j,k}, \tag{2.31}$$

where the explanatory vectors $\boldsymbol{Z}_k$, random coefficients $\boldsymbol{\gamma}_j$, $\beta_j, j = 1, \ldots, N$, and a white noise process $(\varepsilon_{jk}, j, k \ge 0)$ are as written above. The explanatory variables $\boldsymbol{Z}_k' = (Z_{1k}, \ldots, Z_{pk})$ are obtained by projecting random processes $\{(Z_i(t), 0 \le t \le T), i = 1, \ldots, p\}$ on ψ_k: $Z_{ik} = \langle Z_i, \psi_k \rangle$.

Since, in our case, randomness of β_j and $\boldsymbol{\gamma}_j$ comes solely from j, we estimate each model (2.31) by the least squares estimator, thus minimizing the quantity

$$\sum_{k=1}^{n} \Big[\langle X_j, \psi_k \rangle - \beta \langle X_j, \psi_{k-1} \rangle - \boldsymbol{\gamma}_j' \boldsymbol{Z}_k \Big]^2 .$$

Just like in the case above, one can use $\widehat{\beta}_1, \ldots, \widehat{\beta}_N$, and $\widehat{\boldsymbol{\gamma}}_1, \ldots, \widehat{\boldsymbol{\gamma}}_N$ to perform the statistical inference on distributional properties of random coefficients of model (2.31).

Assuming $\psi_0 = 0$ we have the following result on the asymptotic behaviour of projections of the aggregated function $\overline{X}_N$.

Theorem 2. *For each fixed $k \ge 1$,*

$$\lim_{N \to \infty} \mathsf{E} \Big[N^{-1} \sum_{j=1}^{N} \langle X_j, \psi_k \rangle - \sum_{i=0}^{k-1} \mathsf{E}\,(\beta_1^i) \mathsf{E}\,(\boldsymbol{\gamma}_1') \langle \boldsymbol{Z}, \psi_{k-i} \rangle \Big]^2 = 0.$$

Hence, under model (2.31) we have for large N and M, in probability

$$\overline{X}_N(t) \approx \sum_{k=1}^{M} \sum_{i=0}^{k-1} \mathsf{E}\,(\beta_1^i) \mathsf{E}\,(\boldsymbol{\gamma}_1') \langle \boldsymbol{Z}, \psi_{k-i} \rangle \psi_k(t) = \sum_{k=1}^{M} \xi_k \psi_k(t), \quad t \in [0, T],$$

where the scores (ξ_k) constitute a linear process

$$\xi_k = \sum_{i=0}^{k-1} \boldsymbol{a}_i' \langle \boldsymbol{Z}, \psi_{k-i} \rangle = \sum_{i=1}^{k} \boldsymbol{a}_{k-i}' \langle \boldsymbol{Z}, \psi_i \rangle.$$

Note that model (2.31) ensures that each random function X_j has the same covariance. Indeed, noting that

$$\langle X_j, \psi_k \rangle = \sum_{i=0}^{k-1} \beta_j^i [\boldsymbol{\gamma}_j' \langle \boldsymbol{Z}, \psi_{k-i} \rangle + \varepsilon_{j,k-i}]$$

we find

$$\mathsf{E}\,\langle X_j,\psi_k\rangle\langle X_j,\psi_k\rangle=\sum_{i,v=0}^{k-1}\mathsf{E}\,\beta_1^{i+v}\langle\boldsymbol{\gamma}_j'\boldsymbol{Z},\psi_{k-i}\rangle\langle\boldsymbol{\gamma}_j'\boldsymbol{Z},\psi_{k-v}\rangle+\sum_{i=0}^{k-1}\mathsf{E}\,\beta_1^{2i}\sigma^2.$$

So one can use the functional principal components instead of (ψ_k).

For the aggregated process $X=P-\lim_{N\to\infty}\overline{X}_N$, Theorem 1 suggests to consider the functional regression model

$$X(t)=\sum_{j=1}^{p}\int_0^t Z_j(t-s)\beta_j(s)ds+\varepsilon(t),\quad t\in[0,T], \tag{2.32}$$

where the functions $\beta_1,\ldots,\beta_p$ are unknown parameters of the model and $Z_1,\ldots,Z_p$ are explanatory functions and can be either deterministic or random. The difference of this model from the classical functional regression is such that we do not have more information except one realization of the response function X and of each explanatory variable $Z_1,\ldots,Z_p$. From this one sample we have to estimate the parameters of the model and to make statistical inferences.

To estimate the parameters (β_j) we proceed as follows. We assume that

$$\beta_j(t)=\sum_{k=1}^{d}\beta_{jk}u_k(t),$$

for a given set of functions $u_1(t),\ldots,u_d(t)$, reducing the model (2.32) to

$$X(t)=\sum_{j=1}^{p}\sum_{k=1}^{d}\beta_{jk}y_{jk}(t)+\varepsilon(t),\quad t\in[0,1], \tag{2.33}$$

where

$$y_{jk}(t)=\int_0^t Z_j(t-s)u_k(s)ds.$$

Setting

$$Y(t)=(y_{11}(t),\ldots,y_{pd}(t))',\quad B=(\beta_{11},\ldots,\beta_{pd})'$$

we rewrite model (2.33) in a more compact form

$$X(t) = Y'(t)B + \varepsilon(t), \quad t \in [0, T]. \tag{2.34}$$

To estimate B we have several possibilities. By taking the reference points, say, $\tau_k, k = 1, \ldots, n$, we obtain

$$X(\tau_k) = Y'(\tau_k)B + \varepsilon(\tau_k), \quad k = 1, \ldots, n. \tag{2.35}$$

The ordinary least squares estimator of B is then

$$\widehat{B} = \Big(\sum_{k=1}^{n} Y(\tau_k)Y'(\tau_k)\Big)^{-1} \sum_{j=1}^{n} Y(\tau_k)X(\tau_k). \tag{2.36}$$

For asymptotic properties of the estimator $\widehat{B}$ we refer to [146].

Another possibility to estimate the parameter B from (2.34) is by using a normalized basis in $L_2(0, T)$, say (ψ_k). From (2.33) we derive for $\ell = 1, \ldots, n$,

$$\langle X, \psi_\ell\rangle = \langle Y', \psi_k\rangle B + \langle \varepsilon, \psi_\ell\rangle. \tag{2.37}$$

The least squares estimator of parameters B is given by

$$\widehat{B} = \Big(\sum_{k=1}^{n} \langle Y, \psi_k\rangle\langle Y', \psi_k\rangle\Big)^{-1} \sum_{j=1}^{n} \langle Y, \psi_j\rangle\langle X, \psi_j\rangle.$$

2.4.2 Case study

Data

In this study, we used the data of annual unconsolidated bank accounts covering the time period from 1999 to 2013. A dataset is obtained from Bureau van Dijk *Bankscope* database and includes all commercial, savings, and cooperative banks from the European Union countries. These institutional bank types are mainly focused on financial intermediation. The final dataset consisted of the capital adequacy ratio (CAR) from 260 banks. The capital adequacy ratio is defined as a bank's total capital expressed as a percentage of its risk-weighted assets. Exogenous variables, i.e. European Union GDP growth and Eonia overnigth rate are taken from Eurostat and European central bank databases, respectively.

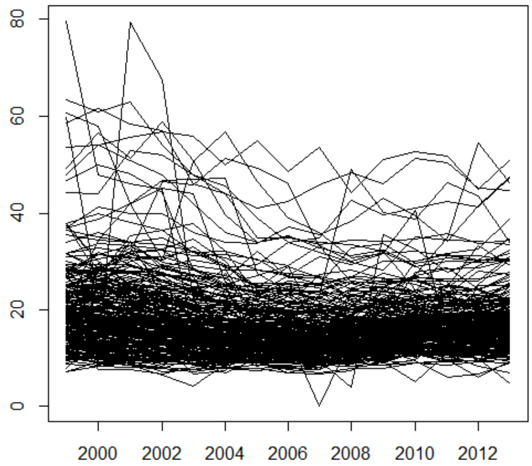

Figure 2.5: Capital adequacy ratio

Source: Bankscope data.

Usually, a higher CAR implies a more stable banking system. From Fig. 2.5, which shows our sample of the data, we can see that most of banks had CAR between 8 and 25, though some of the banks are clear exceptions with a much higher capital ratio.

Clustering

In this section, we aim to predict the general tendency of the CAR development, rather than of each value separately. Therefore, we assume that the data under investigation $x_j = \big(x_1, \ldots, x_T\big)$, $j = 1, \ldots, N$, constitute observations of random curves:

$$X_j = (X_j(t),\, 0 \leq t \leq T), \quad j = 1, \ldots, N.$$

Moreover, we assume that the sampled curves are observed at discrete instants of time. Hence we have:

$$x_j = X_j(i/T) + \varepsilon_j(i/T), \quad i = 1, \ldots, T.$$

We reconstruct the functions $x_j(t)$, $t \in [0,T]$ by smoothing techniques (see e.g. [132]) thus obtaining the functional data below

$$\widehat{x}_j(t), \quad t \in [0,T],\ j = 1,\dots,N,$$

In our case, we used B-spline smoothing to derive the functional data from the initial data (see Fig. 2.6). The smoothed data were used further to form clusters from the whole sample.

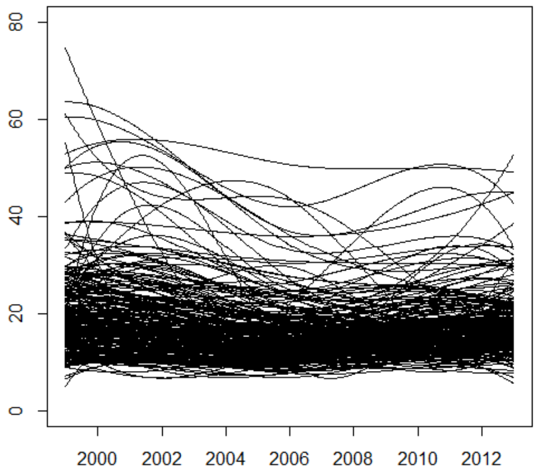

Figure 2.6: Smoothed data of the Capital adequacy ratio

Source: Bankscope data and authors' calculation.

Since the functional data have dynamic character, the concept of their similarity is complex, it is important to evaluate not only the closeness of the curves, but also their behaviour. The dissimilarity measure, which we used in this analysis, is constructed from two parts. The first part shows how B-spline approximations are close to one another. The second part uses q-th derivative to capture how close is the change of curves. The dissimilarity measure is estimated as follows:

$$D_{\mathrm{SUP}}(x,y) = \sup_t |x(t) - y(t)| + \sup_t |x^{(q)}(t) - y^{(q)}(t)|. \tag{2.38}$$

This measure takes into account both the closeness and behaviour of the data. In our case, we took $q = 1$, i.e. we took the first derivative of the function only.

In this study, we used the conventional agglomerative hierarchical clustering algorithm. This method is helpful in clustering functional data into a tree of clusters (dendrogram). The complete linkage algorithm, applied in this study, measures the similarity between two clusters as the similarity between the farthest pair of data, belonging to different clusters.

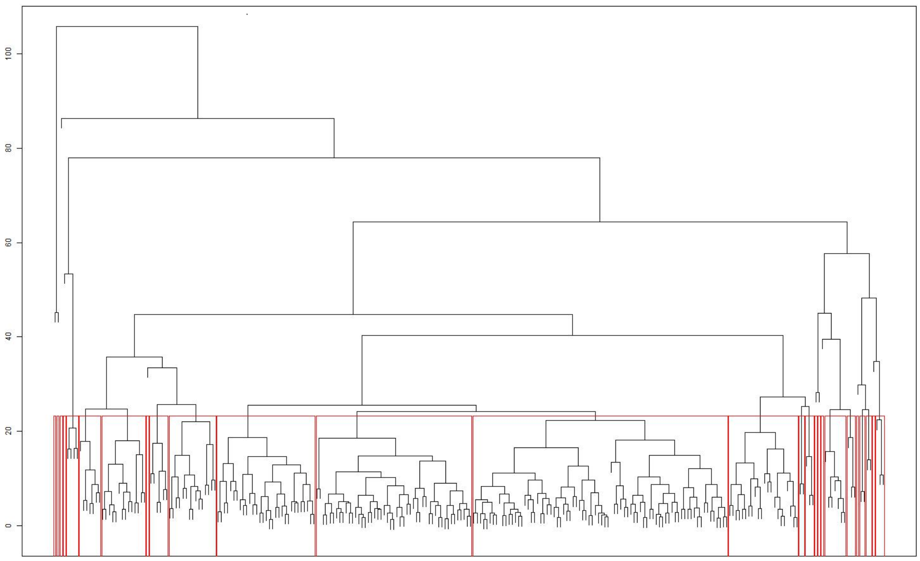

Figure 2.7: Dendrogram of the Capital adequacy ratio

Source: Bankscope data and authors' calculation.

Figure 2.7 shows a dendrogram that is formed from the banks in our sample. Clustering algorithms divides unlabelled data into significant groups and the precise number of clusters is not known a prior. It is important to evaluate the clustering results and find partitioning that fits data the best. In our analysis we wanted that a cluster would have a clear tendency, but at the same time, it would be made up of more than ten banks. In this way, we have six clusters that are made up from 14 to 80 banks. The clusters that are formed by a fewer number of banks could be considered as outliers.

Furthermore, we test whether the difference between two clusters is statistically significant. To evaluate that, the absolute value of a t-statistic at each point is considered ([132]):

$$T(t) = \frac{|\bar{x}(t) - \bar{y}(t)|}{\sqrt{\frac{1}{N_x}Var(x(t)) + \frac{1}{N_y}Var(y(t))}}. \tag{2.39}$$

The test statistic used is the maximum value of the multivariate T-test, $T(t)$ ([132]). To find the critical value of this statistic, a permutation test was used, which performs the following procedure:

1. Randomly shuffle the labels of curves.

2. Recalculate the maximum of $T(t)$ with the new labels.

Repeating this procedure many times allows a null distribution to be constructed. This process provides a possibility for evaluating the maximum value of observed $T(t)$.

Table 2.2: Permutation test p-value

	Cluster 2	Cluster 4	Cluster 5	Cluster 6	Cluster 10
Cluster 1	0.00	0.00	0.00	0.00	0.00
Cluster 2	-	0.00	0.00	0.00	0.00
Cluster 4	-	-	0.00	0.00	0.005
Cluster 5	-	-	-	0.00	0.00
Cluster 6	-	-	-	-	0.00

Source: authors' calculation.

The results in Table 2.2 show that the p-value of the functional T-test is less than 0.05, indicating that the clusters are significantly different. Therefore, our cluster analysis partitioned banks into clusters that have characteristic patterns and those clusters are used to estimate the proposed functional data models.

Functional principal components analysis

The next step of our analysis is to investigate functional principal components (FPCA) of the data, which allows us to display cluster in a few components. In FPCA eigenfunctions of the variance-covariance function describe the main variability in the data.

Table 2.3: Variability explained by FPCA

	Number of FPC	Variability explained with the first PC	Variability explained with the second PC	Variability explained with the third PC	Total variability explained with all three PC
Cluster 1	3	52.6	24.6	14.9	92.3
Cluster 2	3	65.8	20.6	7.3	93.7
Cluster 4	3	73.9	14.8	8.4	97.1
Cluster 5	3	45.0	31.9	15.4	92.3
Cluster 6	3	68.5	17.8	9.1	95.4
Cluster 10	3	57.6	20.0	15.5	93.2

Source: authors' calculation.

Even though banks were grouped into clusters with a similar development, there is still some variability among them. Table 2.3 shows the results of the functional principal component analysis. We see that for all clusters three PC can explain more than 90 per cent of variation within the cluster. It is known that unrotated functional principal components display the same sequence of variation. For example, if we have three principal components, then the first is a constant shift, the second one is a linear contrast between the first and second half, and the third one is a quadratic pattern ([132]).

The VARIMAX rotation algorithm is used in order to have more meaningful explanation for the components of variation. The results of the first cluster are plotted in Figure 2.8 where a black line is the mean function, a red dashed line shows what happens when the principal component is added to the mean and a green dotted line shows the effect of subtracting the component. In this case, the first component reflects the variation in the middle of the sample. The second component explains the variation until 2006, and the last one captures the variation that is strongest at the end of the sample. The analysis of the functional pricpipal component anal-

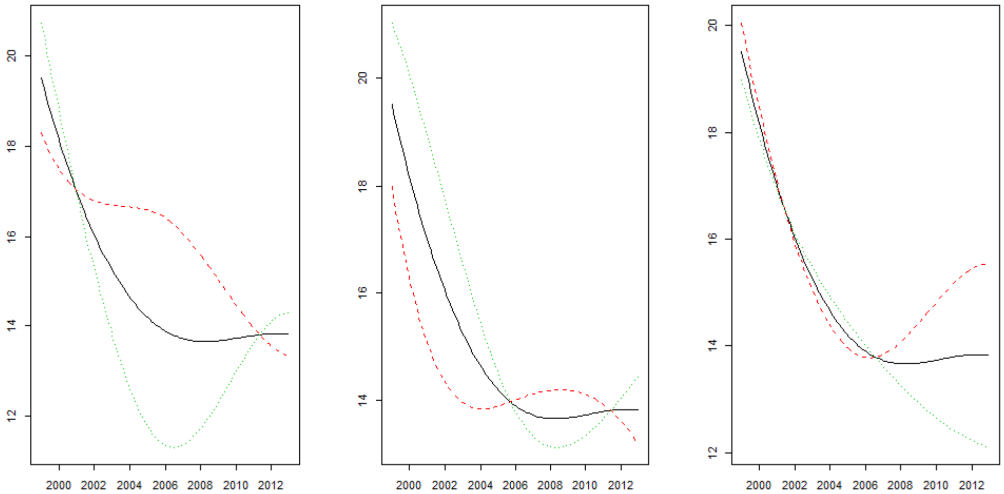

Figure 2.8: The three rotated principal component functions are shown as a perturbation of the first cluster mean expressed by a black line

Source: authors' calculation.

ysis has revealed that it is not enough to have only the mean function to characterize the cluster as the variation within the cluster is still present.

Model estimation

Bottom-up estimation

In the bottom-up case we estimated the following equation to each bank that belongs to a particular cluster:

$$X_j(\tau_k) = \alpha + \beta_j X_j(\tau_{k-1}) + \sum_{i=1}^{p} \gamma_i Z(\tau_k) + \varepsilon_j(\tau_k), \tag{2.40}$$

where $X_j(\tau_k)$ is the bank's capital adequacy ratio, and $Z(\tau_k)$ are exogenous explanatory variables. We assume that CAR ratio depends not only on the capital level that bank has or its performance, but also on the environment in which it operates. As explanaroty variables we consider the annual growth rate of the Europeans Union's GDP which shows the overall economic situation. During the period of economic growth banks can lend more, customers need more of their services, etc., which may increase the profitability and affect other performance indicators. The second one is the Eonia overnight rate that influences cost of lending, deposit interest rate, and cost of wholesale financing. Thus Eonia overnight rate has an impact on the main activities of the bank.

We used the two-stage least squares estimation method, which allowed us to have consistent estimates of the coefficients. As instruments we used $X_j(\tau_{k-2})$ and $Z(\tau_k)$. Thus, each bank was evaluated with the same set-up.

After the initial assessment, we calculated the estimated mean values of the cluster:

$$\hat{\bar{X}}_{BU} = \frac{1}{N}\sum_{j=1}^{N}\hat{X}_j, \tag{2.41}$$

where $\hat{X}_j$ is the fitted values from equation 2.40. In this way, we have reveived the bottom-up estimate of the mean process for each cluster.

Top-down estimation

Similarly to the bottom-up case, we used the following equation in the top-down estimation:

$$\bar{X}(\tau_k) = \alpha + \beta\bar{X}(\tau_{k-1}) + \sum_{i=1}^{p}\gamma_i Z(\tau_k) + \varepsilon(\tau_k), \tag{2.42}$$

where $\bar{X}(\tau_k)$ is the mean value of CAR in the cluster, i.e. we estimate the mean process directly. In this case we have obtained $\hat{\bar{X}}_{TD}$ which is the fitted values from the estimation of model 2.42. Furthermore, we compared the estimation accuracy in terms of RMSE and MAE.

Estimation results

During the estimation process we gradually increased the number of points from 15, the initial time span of the data, to 500. Figure 2.9 shows the results of the average β_j from the bottom-up estimation and the β coefficient from the top-down estimation, i.e. AR(1) coefficients, obtained from the estimation of the first cluster. The results gave us several findings. The first one, is that both values of the β estimates are close to each other. The average β_j shows the same pattern and the value is similar to the β

estimate from the top-down equation. The second one, as the number of points is increasing from 15 to 500, both values of the AR(1) coefficient are getting closer to 1. In the beginning, the average β_j is around 0.65 and TD estimate is close to 0.75. When we increased the number of points to 100, the estimates were equal to 0.97 and 0.96, respectively. Moving forward the difference between the estimates decreased and they both were getting closer to one. The estimation of another clusters provided similar results and findings.

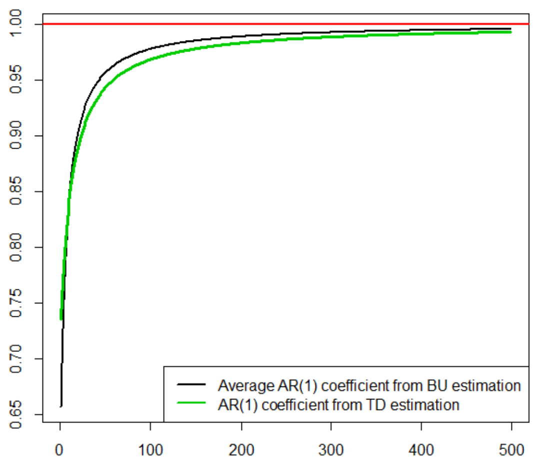

Figure 2.9: AR(1) coefficients from the estimated equations (Cluster 1)

Source: authors' calculation.

Meanwhile, Figure 2.10 shows that both the root mean square error (RMSE) and mean absolute error (MAE) are decreasing, i.e. the estimated values are closer to the actual mean process when we increase the number of points. From this figure, we can also notice that residuals from the top-down estimation are smaller in terms of RMSE and MAE. Therefore, we may conclude that the top-down estimated model yields better results for in sample estimation for the first cluster.

However, RMSE and MAE from other clusters are much closer to one another. Thus, the bottom-up estimation accuracy is similar to the top-down estimation accuracy. We cannot distinguish which of them is better in the case of in sample estimation. Nevertheless, in all cases, RMSE and MAE are smaller as the number of points is increasing.

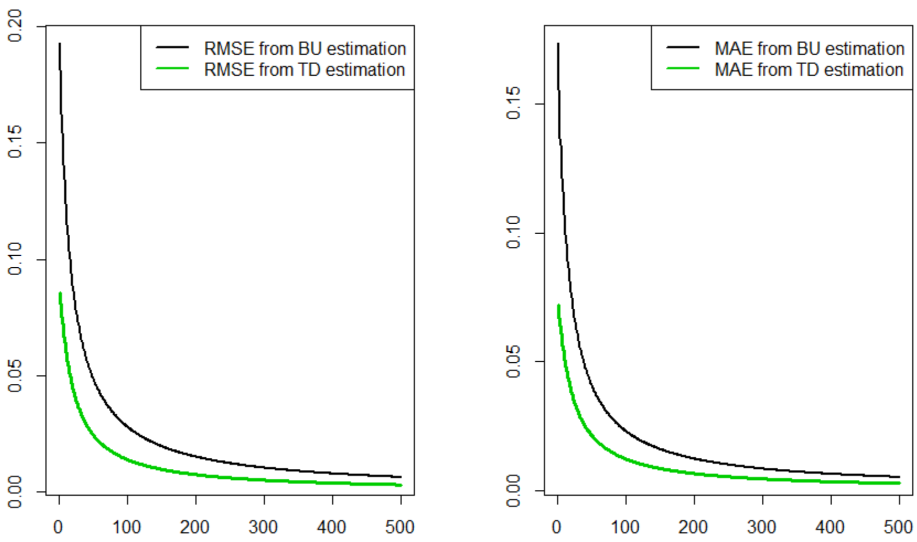

Figure 2.10: RMSE and MAE from in sample estimation (Cluster 1)

Source: authors' calculation.

Forecasting

We split our data in the clusters into a training sample (including data until the end of 2011) and a testing sample (including data from the first data point in 2012 to the end of 2013). Then we estimate the bottom-up model (eq. 2.40) and the top-down model (eq. 2.42), taking the number of data points (τ_k) which will be the same with that of annual, semi-annual, quarterly or monthly data. In this way we need to forecast 2, 4, 8, and 24 data points, respectively. In order to decide which the top-down or bottom-up model gives a better forecast of the mean process, we calculate forecast errors by comparing the forecast with the actual out-of-sample data.

To measure the overall point forecast accuracy, we use RMSE and MAE. Based on error statistics presented in Table D.1 and Table D.2 (see Appendix D) we can conclude that, in this way we also cannot distinguish which method is better. The top-down model provided a better out-of-sample forecast for three out of six clusters (Cluster 1, Cluster 2 and Cluster 10). For another, estimation of each bank separately and calculating the mean is a better strategy. If we forecast the annual data, then Cluster 5 and Cluster 6 yield a better forecast using the TD model.

We may conclude that it is useful to use both forecasts, i.e. top-down forecast and bottom-up forecast. The top-down model may provide a better out-of-sample forecast for the mean process, but at the same time, taking BU approach we can achieve similar results and, in addition, we have the forecasts for each bank individually.

Functional regression model

In this section, we propose a procedure to estimate the functional regression model:

$$X(t) = \sum_{j=1}^{p} \int_0^t Z_j(t-s)\beta_j(s)ds + \varepsilon(t), \quad t \in [0, T], \tag{2.43}$$

where the functions $\beta_1, \ldots, \beta_p$ are unknown parameters of the model and $Z_1, \ldots, Z_p$ are regressors. As in the previous case, we have two exogenous regressors: annual GDP growth and daily information of Eonia short-term interest rate. Thus, we need to estimate the functions β_1 and β_2. Here we explore one of the advantages of the functional data analysis, i.e. we can use data that have different frequencies. In this case the annual banking data and the GDP growth data are combined with the daily interest rate data. By smoothing the initial data with B-splines, we can later use them in the same regression.

After smoothing exogenous variables with B-spline functions, we need to construct variables $y_{jk}(t)$, $j = 1, 2;\ k = 1, \ldots, d$. To do that we calculate inner products between regressors and the Fourier functions (i.e. we used Fourier basis functions in model 2.33) at discretized points that correspond to the data points of $X(\tau_k)$. Next, we need to estimate the following model:

$$X(\tau_k) = Y'(\tau_k)B + \varepsilon(\tau_k), \quad k = 1, \ldots, n. \tag{2.44}$$

Setting $X(\tau_0) = 0$ and using the least squares estimator we got the estimates $\hat{\beta}_{1k}$ and $\hat{\beta}_{2k}$, $\ k = 1, \ldots, d$. In our case, we used three Fourier basis functions to approximate $\beta_j(t)$, therefore $d = 3$, and one needed to estimate 6 coefficients in total.

Figure 2.11 shows the actual (black solid line) and fitted (red dashed

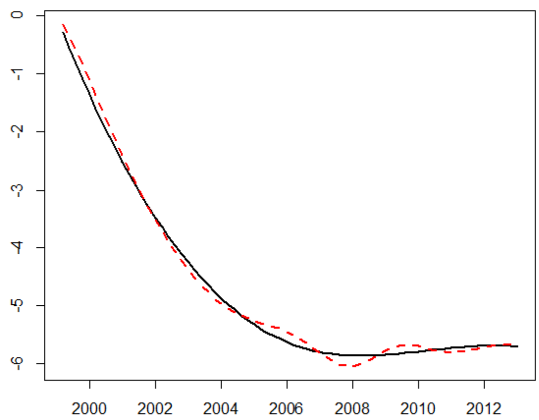

Figure 2.11: Actual mean function (black solid line) and the fitted (red dashed line) function (Cluster 1)

Source: authors' calculation.

line) values of the mean function from the first cluster. The results demonstrate that the model captures the values of the actual data quite well in the beginning of the sample period. Then there are some divergences in explaining the development of the data approximately between 2006 and 2010, and the last part is captured well again. We can conclude that the GDP growth and Eonia overnight interest rate has some influence on the development of the banks' capital adequacy ratio during the sample period, while in 2006-2010, other variables would also be useful.

The estimated model shows similar results for other clusters as well. The model captures the development of the mean function at the beginning of the sample, but then it is not able to explain the movement of the CAR ratio, though the fitted values are close to the actual data. A greater divergence is observed for Cluster 4 and Cluster 5. For those clusters, additional explanatory variables could be considered.

The empirical application shows that the functional regression model and its estimation method, proposed in thesis, are useful when modelling the main characteristic behaviour of the data. This model also allows us to combine data with the values observed at different frequencies. Another useful feature is the availability to discretize the function to as many points as necessary to estimate the model.

Forecasting with functional regression model

Similarly to the previous models, where we used the bottom-up and top-down approach to compare forecasting performances, we employed the functional regression model to get another forecast of the top-down approach. In this case, the coefficients $\hat{\beta}_{1k}$ and $\hat{\beta}_{2k}$, $k = 1, \ldots, 3$ were estimated using the data until the end of 2011. Then we constructed fitted values to the end of the sample. Table D.1 (Appendix D) gives the forecasting accuracy in terms of RMSE and MAE. Forecast errors were calculated taking 2, 4, 8, and 24 data points, which is identical to the BU and TD models.

If we compare the forecasting error statistics of the functional regression model with the bottom-up and top-down statistics, we can see that the functional model was not able to capture the future development of the mean process for Cluster 4 and Cluster 5. The errors of those clusters are significantly larger. However, for Cluster 2, Cluster 6 and Cluster 10, the functional regression model was able to give better forecasting results.

We may conclude that proposed functional regression model and it's estimation method might be used to forecast mean process of the data. Forecasting errors are comperable with results obtained from more traditional estimation techniques.

2.4.3 Conclusions of the cluster estimation and forecasting

The top-down versus bottom-up forecasting problem is addressed in this work. We explore the advantages of the functional data estimating top-down and bottom-up models for the capital adequacy ratio of European banks. In the beginning the initial data were smoothed and banks were clustered into similar groups. Then each cluster was analyzed by two types of models: the first is an autoregressive model with exogenous variables and the second is a functional regression model.

Estimation by the first model included one of the advantages of the functional data, i.e. availability to discretize curves into many points. The top-down and bottom-up models provided similar results in terms of in sample accuracy. The average of the estimated bottom-up fitted values was close to the top-down estimation of the mean process. Moreover, in both cases the AR(1) coefficient was getting closer to 1, as the number of points in the estimation process was increasing. Thus, we may conclude that, in this case, there is no significant difference between estimating mean behaviour of the capital adequacy ratio or estimating each bank separately and then adding them up. The out-of-sample forecasting exercise gave similar results, for three out of six clusters the top-down model provided better forecasts of the mean process, but the mean of the other three clusters was better forecasted with bottom-up approach.

The functional regression model and the proposed estimation method are useful to capture the main characteristic behaviour of the banking data. In this case, we used another feature of the functional data, i.e. it allowed us to use data which have different frequencies. Though at the beginning of the sample functional regression model is able to fit the data well, later some divergence between the actual data and fitted values was observed. However, main characteristic pattern of the stochastic process was captured with this model. The proposed functional regression model and its estimation is novel in a way that it allows us to estimate the functional relationship having only one realization of the stochastic process. The out-of-sample forecasting performance is comparable with the more traditional estimation techniques. Nevertheless, a further analysis of the model and empirical applications is still needed.

To sum up, the estimation results, in this work, show that it might be useful to apply the functional data analysis methods and to take advantages of some features when analysing banking data. Such methods can give additional insights about the possible relationship in this sector.

Conclusions

After the Global financial crisis of 2007-2008 there has been a strong interest of research that is related with financial system or research that is dealing with the banking system. One of the main topics is how to estimate the stability of the banks and the whole banking system. In the first part of this thesis we were working with the Lithuanian banking data and tried to address the riskiness and profitability issues of the system. In the second part, we considered the data from European banks and proposed clustering and forecasting methods that could be applied to these data. The problems, considered in this thesis, help to get some new insights into this stream of research and into the field of functional data analysis.

In the thesis, we presented a macroeconomic top-down stress testing methodology, which is used to assess the resilience of the Lithuanian banking sector. In this work, we briefy introduced a stress testing methodology and described the main components of it. Then we reviewed the macroeconomic scenario design. The main focus, however, was on the development and estimation of the so-called satellite models, which help link the dynamics of macroeconomic variables to credit risk and profitability of a bank. It is worth noting that the focus was on testing bank solvency, i.e. assessing the potential insolvency of the bank due to regulatory capital shortages in the adverse macroeconomic scenario.

The presented methodology is used to run a top-down stress test. It means that the models and assumptions used are the same for all banks, which makes the results comparable across different banks. On the other hand, such an approach does not fully capture the specific aspects of individual banks. The proposed stress testing methodology is in line with the literature on stress testing and is capable of producing robust and meaningful results.

In the thesis we further analyzed the Lithuanian banking data and pay closer attention to the profitability of our banks. We examined the long-term and short-term relationship between bank profitability and bank-specific or macroeconomic variables. In the analysis, we applied a pooled mean group estimator, developed by Pesaran et al. (1997, 1999) [124, 125], which constrains long-term coefficient across cross-sectional units and at the same time allows intercept, short-term coefficients and adjustment to the equilibrium relationship to differ. The empirical results supported the idea that the Lithuanian banking sector is still developing and banks could be attributed to small and medium size banks and, therefore, they can exploit the economies of scale and scope. Further, in line with other studies, we found a pro-cyclical behaviour of banks revenues and expenses, i.e. GDP development is an important determinant which makes the impact on banks' profits. However, a larger data set for Lithuanian banks would help us to include more determinants in the model and have a better understanding of the long-term and short-term relationship.

In the second part of the thesis, we considered clustering of the banks according to their performance ratios. We approached the problem from two perspectives. One way was to look at the data as time series and apply dissimilarity measures that are based on raw data or some properties of time series. Another approach taken here was to consider that our data is continuous functions (i.e. functional data) and then we used dissimilarity measures, based on the functional data properties. In the thesis we have introduced two dissimilarity measures that are based on functions and take into account not only how close the two curves are, but also how similarly they change over time. In addition to the univariate clustering, where banks are grouped into clusters according to one bank performance ratio, a multivariate clustering is applied, where banks are clustered based on several ratios. For that reason, we extended several dissimilarity measures from the univariate case into a multivariate case.

The results of cluster analysis show that there is no clear answer which dissimilarity measure is best for banking data and how many clusters we get. However, the clustering, based on distance measures that used data as functions mostly outperformed distance measures, based on time series properties. Since in the cluster analysis data are unlabelled, the number on clusters is not known a priori. Division of banks into 20 clusters has revealed that there are from 6 to 12 clusters (depending on the banking ratio) that are formed by a larger number of banks and other clusters are formed by one or several banks. Those larger clusters could be taken as the basis for further research.

In the last part of the thesis, we considered the top-down versus bottom-up forecasting problem. We explored the advantages of the functional data and estimated top-down and bottom-up models to forecast a future development of the European banks' capital adequacy ratio. Furthermore, in this work, we proposed a novel functional regression model and introduce its estimation method. This model could be used to estimate the functional relationship between one realization of the stochastic process and other functional covariates.

The estimation of the models showed that, in this case, there are no significant difference between the top-down and bottom-up approaches while forecasting a future development of the mean process. The autoregressive models with exogenous variables and functional regression models provided similar results in terms of RMSE and MAE. Moreover, the proposed functional regression model was able to capture the main characteristic behaviour of the banking data and the result of out-of-sample forecasting performance was comparable with more traditional estimation techniques. Nevertheless, a further analysis of the model and more empirical applications is still necessary.

To sum up, in the thesis, various aspects of the banking data analysis were considered. The stress testing, profitability, clustering and forecasting problems were discussed and analyzed. The results that were received here help to add some new insights into the stream of research that is dealing with banking data, in particular concerting stability of the banking system, as well as into the field of functional data analysis. Of course, future work and improvements are still needed. For example, it would be useful to apply quantile regression in stress testing procedure, allowing the estimation of nonlinear relationships between macroeconomic variables and bank indicators. Since various banking ratios are overlapped, it would be also interesting to apply fuzzy clustering algorithms to the banking data. These and other issues could be addressed in the future studies.

Bibliography

[1] Abraham, C., P.A. Cornillon, E. Matzner-Løber, and N. Molinari (2003). *Unsupervised curve clustering using B-splines*. Scandinavian Journal of Statistics, Vol. 30(3), pp. 581-595.

[2] Abreu, M., and V. Mendes (2002). *Commercial Bank Interest Margins and Profitability: Evidence from E.U. Countries*. Working Paper Series, Porto.

[3] Acharya V., and T.Yorulmazer (2008). *Information Contagion and Bank Herding*. Journal of Money, Credit and Banking, Vol. 40(1), pp. 215-231.

[4] Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2013). *Systemic Risk and Stability in Financial Networks*. NBER Working Paper, No. 18727.

[5] Akaike, H. (1970). *Statistical predictor information*. Annals of the Institute of Statistical Mathematics, Vol. 22, pp. 203–217.

[6] Akaike, H. (1974). *A new look at the statistical model identification*. IEEE transactions on Automatic Control, Vol. 19(6), pp. 716–723.

[7] Albertazzi, U., and L. Gambacorta (2009). *Bank profitability and the business cycle*. Journal of Financial Stability, Vol. 5(4), pp. 393-409.

[8] Andersen, H., S.A. Berg, and E.S. Jansen (2008). *The dynamics of operating income in the Norwegian banking sector*. Norges Bank Working Paper 13.

[9] Apergis, N. (2009). *Bank profitability over different business cycles regimes: evidence from panel threshold models*. Banks and Bank Systems, Vol 4(3), pp. 59-70.

[10] Asberg, P., and H. Shahnazarian (2008). *Macroeconomic Impact on Expected Default Frequency*. Sveriges Riksbank Working Paper No. 21

[11] Athanasoglou, P.P., S.N. Brissimis, and M.D. Delis (2008). *Bank-specific, industry-specific and macroeconomic determinants of bank profitability*. Journal of International Financial Markets, Institutions and Money, Vol. 18(2), pp. 121–136.

[12] Athanasopoulos, G., R.A. Ahmed, and R.J. Hyndman (2009). *Hierarchical forecasts for Australian domestic tourism*. International Journal of Forecasting 25(1), pp. 146–166.

[13] Aue, A, D.D. Norinho, and S. Hörmann (2015) *On the prediction of stationary functional time series*. Journal of the American Statistical Association, Vol. 110(509), pp. 378-392.

[14] Bardsen, G., K.G. Lindquist and D.P. Tsomocos (2006). *Evaluation of Macroeconomic Models for Financial Stability Analysis*. Norwegian University of Science and Technology Working Paper, 4.

[15] Basel Commitee on Banking Supervision (BCBS) (2009). *Principles for Sound Stress Testing Practices and Supervision.*

[16] Batista, G., X. Wang, and E.J. Keogh (2011). *A complexity-invariant distance measure for time series*. In Proceedings of the Eleventh SIAM International Conference on Data Mining, pp. 699-710.

[17] Batista, G., E.J. Keogh, O.M. Tataw and V. de Souza (2014). *CID: an efficient complexity-invariant distance for time series*. Data Mining and Knowledge Discovery, Vol. 28(3), pp. 634-669.

[18] Berlemann, M, M. Oestmann, and M. Thum (2014). *Demographic change and bank profitability: empirical evidence from German savings banks*. Applied Economics, Vol. 46(1), pp. 79-94.

[19] Berndt, D.J., and J. Clifford (1994). *Using dynamic time warping to find patterns in time series*. In AAA1-94 Workshop on Knowledge Discovery in Databases, pp. 359-370.

[20] Bezdek, J.C., R. Ehrlich, and W. Full (1984). *FCM: Fuzzy c-means algorithm*. Computers and Geosciences, Vol. 10(2-3) pp. 191-203.

[21] Bohte, Z., D. Cepar, and K. Kosmelij (1980). *Clustering of time series*. In Proceedings of 1980 Computational Statistics Meeting COMPSTAT80, pp. 587-593.

[22] Borio, C., M. Drehmann, and K. Tsatsaronis (2012). *Stress testing macro stress testing: does itlive up to expectations?* BIS Working Paper No. 369.

[23] Boss, M., G. Krenn, C. Puhr and M. Summer (2006). *Systemic Risk Monitor: a Model for Systemic Risk Analysis and Stress Testing for Austria*. OeNB Financial Stability Report, 11, 83–95.

[24] Buncic, D. and M. Melecky (2013). *Macroprudential stress testing of credit risk: A practical approach for policy makers*. Journal of Financial stability, Vol. 9(3), pp. 347-370.

[25] Burrows O., D. Learmonth, and J. McKeown (2012). *RAMSI: A Top-Down Stress-Testing Model*. Bank of England, Financial Stability Paper No. 17.

[26] Caccioli, F., M. Shrestha, C. Moore, and J.D. Farmer(2012). *Stability analysis of financial contagion due to overlapping portfolios*. University of Trier Working Paper 1210.

[27] Caliński, T., and J. Harabasz (1974). *A dendrite method for cluster analysis*. Communications in Statistics, Vol. 3(1), pp. 1-27.

[28] Canale, A., and S. Vantini (2016). *Constrained functional time series: Application to the Italian gas market*. International Journal of Forecasting, Vol. 32, pp. 1340-1351.

[29] Castren, O., S. Dees and F. Zaher (2008). *Global Macro-financial Shocks and Corporate Sector Expected Default Frequencies in the Euro Area*. European Central Bank (ECB) Working Paper, 875.

[30] Celov, D. (forthcoming) *LEMPA - a Quarterly Lithuanian Economy Model for Projections and Analysis.*

[31] Chan, K.-P., and W.-C. Fu (1999). *Efficient time series matching by wavelets*. In Proceedings of 15th International Conference on Data Engineering (ICDE '99), pp. 126-133.

[32] Choi, I. (2001). *Unit root tests for panel data*. Journal of International Money and Finance, Vol. 20(2), pp. 249–272.

[33] Chouakria, A.D., and P.N. Nagabhushan (2007). *Adaptive dissimilarity index for measuring time series proximity*. Advances in Data Analysis and Classification, Vol. 1(1), pp. 5-21.

[34] Chronopoulos, D.K., H. Liu, F.J. McMillan, and J. Wilson (2015). *The dynamics of US bank profitability*. The European Journal of Finance, Vol. 21(5), 426-443.

[35] Claeys, S., and R. van der Vennet (2004). *Determinants of bank interest margins in Central and Eastern Europe: a comparison with the West.* Economic Systems, Vol. 32(2), pp. 197-216.

[36] Coffinet, J., and S. Lin (2010). *Stress testing banks' profitability: the case of French banks*. Banque de France, Working paper No. 306.

[37] Christiano, L., R. Motto, and M. Rostagno (2010). *Financial factors in economic fluctuations*. ECB Working Paper, No. 1192.

[38] Čihák, M. (2004). *Stress Testing: a Review of Key Concepts*. CNB International Research and Policy Note, 2.

[39] Cuevas, A., F. Febrero, and R. Fraiman (2002). *Linear functional regression: the case of fixed design and functional response*. Canadian Journal of Statistics, Vol. 30, pp. 285-300.

[40] Darracq Pariés, M., C. Kok, and D. Rodriguez-Palenzuela (2011). *Macroeconomic propagation under different regulatory regimes: evidence from an estimated DSGE model for the euro area.* International Journal of Central Banking, Vol. 7(4), pp. 49-113.

[41] Davies, D.L., and D.W. Bouldin (1979). *A cluster separation measure.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 1(2), pp. 224-227.

[42] de Boor, C. (2001). *A Practical Guide to Splines, Revised Edition.* New York: Springer.

[43] Demirguc-Kunt, A., and H. Huizinga (1999). *Determinants of commercial bank Interest margins and profitability: some international evidence.* World Bank Economic Review 13(2), pp. 379–408.

[44] Díaz, S.P., and J.A. Vilar (2010). *Comparing several parametric and nonparametric approaches to time series clustering: a simulation study.* Journal of Classification, Vol. 27, pp. 333-362.

[45] Dickey, D.A., and W.A. Fuller 1979. *Distribution of the estimators for autoregressive time series with a unit root.* Journal of the American Statistical Association, Vol. 74, pp. 417–431.

[46] Dieppe, A., A. González Pandiella, S. Hall, and A. Willman (2011). *The ECB's new multicountry model for the euro area: NMCM – with boundedly rational learning expectations.* ECB Working Paper, No. 1316.

[47] Dietrich, A., and G. Wanzenried (2011). *Determinants of bank profitability before and during the crisis: Evidence from Switzerland.* Journal of International Financial Markets, Institutions and Money, Vol. 21(3), pp. 307–327.

[48] Dietrich, A., and G. Wanzenried (2014). *The determinants of commercial banking profitability in low-middle-, and high-income countries.* The Quarterly Review of Economics and Finance, Vol. 54(3), pp. 337-354.

[49] Ding, H., G. Trajcevski, X. Wang, and E.J. Keogh (2008). *Querying and mining of time series data: experimental comparison of representations and distance measures*. In Proceedings of the VLDB Endowment, Vol. 1(2), pp. 1542-1552.

[50] Drehmann, M. (2008). *Stress Tests: Objectives, Challenges and Modelling Choices*. Riksbank Economic Review 2, pp. 60-92.

[51] Duarte, C., and A. Rua (2007). *Forecasting inflation through a bottom-up approach: How bottom is bottom?* Economic Modelling, Vol. 24, pp. 941-953.

[52] Dunn, J.C. (1974). *Well separated clusters and fuzzy partitions*. Journal of Cybernetics, Vol. 1(24), pp. 95-104.

[53] Ester, M., H.-P. Kriegel, J. Sander, and X. Xu (1996). *A density-based algorithm for discovering clusters in large spatial databases with noise*. In Proceedings of 2nd International Conference On Knowledge Discovery and Data Mining, pp. 226-231.

[54] European Banking Authority (2014). *Methodological note EU-wide Stress Test 2014*.

[55] Faloutsos, C., M. Ranganathan, and Y. Manolopoulos (1994). *Fast subsequence matching in time-series databases*. In Proceedings of the 1994 ACM SIGMOD International Conference on Management of Data, pp. 419-429.

[56] Ferraty, F., and P. Vieu (2006). *Nonparametric Functional Data Analysis. Theory and Practice*. Springer Series in Statistics, Springer, New York, pp. 125-147.

[57] Foglia, A. (2009) *Stress Testing Credit Risk: A Survey of Authorities' Approaches*. International Journal of Central Banking, Vol. 5(3), pp. 9-45.

[58] Fu, T.-C. (2011). *A review on time series data mining*. Engineering Applications of Artificial Intelligence, Vol. 24, pp. 164-181.

[59] Gai, P., and S. Kapadia (2010). *Contagion in Financial Networks*. Bank of England Working Paper, No. 383.

[60] Gambacorta, L., M. Scatigna, and J. Yang (2014). *Diversification and bank profitability: a nonlinear approach*. Applied Economics Letters, Vol. 21(6), pp. 438-441.

[61] García-Herrero, A., S. Gavilá, and D. Santabárbara (2009). *What explains the low profitability of Chinese banks*? Journal of Banking and Finance, Vol. 33(11), pp. 2080–2092.

[62] Geleano, G., and D. Peña (2000). *Multivariate analysis in vector time series*. Statistics and Econometrics Series 15 WP, Vol. 1(24).

[63] Geršl, A., P. Jakubik, T. Konečný and J.Seidler (2012) *Dynamic Stress Testing: The Framework for Testing Banking Sector Resilience Used by the Czech National Bank*. CNB Working Paper, No. 11.

[64] Georg, C.P. (2013). *The effect of the interbank network structure on contagion and common shocks*. Journal of Banking and Finance, Vol. 37(7), pp. 2216-2228.

[65] Giannone, D., M. Lenza, and L. Reichlin (2012). *Money, credit, monetary policy and the business cycle in the euro area*. CEPR Discussion Paper, No. 8944.

[66] Goddard, J., H. Liu, P. Molyneux, and J. Wilson (2010). *Do bank profits converge*? European Financial Management, Vol. 19(2), pp. 346-365.

[67] Goddard, J., P. Molyneux, and J. Wilson (2004). *The profitability of European Banks: a cross-sectional and dynamic panel analysis*. Manchester School, Vol. 72(3), pp. 363–381.

[68] Goia, A., C. May, and G. Fusai (2010). *Functional clustering and linear regression for peak load forecasting*. International Journal of Forecasting, Vol. 26, pp. 700-711.

[69] Golay, X., S. Kollias, G. Stoll, D. Meier, A. Valavanis, and P. Boesiger (2004). *A new correlation-based fuzzy logic clustering algorithm for fMRI*. Magnetic Resonance in Medicine, Vol. 40(2), pp. 249-260.

[70] Golin, J. (2001). *The bank credit analysis handbook: a guide for analysts, bankers and investors*. Wiley Finance Series, Wiley, New York, pp. 261-336.

[71] Gray, D.F., M. Gross, J. Paredes, and M. Sydow (2013). *Modelling the joint dynamics of banking, sovereign, macro, and corporate risk using contingent claims analysis (CCA) in a multi-country global VAR*. IMF Working Paper, No. 218

[72] Halkidi, M., Y. Batistakis, and M. Vizirgiannis (2001). *On clustering validation techniques*. Journal of Intelligent Information Systems, Vol. 17(2/3), pp. 107-145.

[73] Haque, N.U., M.H. Pesaran, and S. Sharma (1999). *Neglected Heterogeneity and Dynamics in cross-country Savings Regressions*. International Monetary Fund Working paper No. 99/128.

[74] Härdle, W. (1994). *Applied Nonparametric Regression*. Cambridge University Press.

[75] Hausman, J.A. (1978). *Specification tests in econometrics*. Econometrica, Vol. 46(6), pp. 1251-1271.

[76] Hays, S., H. Shen, and J.Z. Huang (2012). *Functional dynamic factor models with application to yield curve forecasting*. The Annals of Applied Statistics, Vol. 6(3), pp. 870-894.

[77] Heise, S., and R. Kühn (2012). *Derivatives and Credit Contagion in Interconnected Networks*. European Physical Journal B, Vol. 85(4), id.115.

[78] Hemanth, D.J., J. Anitha, and V.E. Balas (2016). *Performance improved modified fuzzy c-means algorithm for image segmentation applications*. Informatica, Vol. 26(4), pp. 635-648.

[79] Henry, J., and Ch. Kok (2013). *A macro stress testing framework for assessing systemic risks in the bankink sector*. European Central Bank, Occasional paper series, No. 152.

[80] Horváth, L., and P. Kokoszka (2012). *Inference for Functional Data with Applications*. Springer, New York.

[81] Huddleston, S.H., J.H. Porter, and D.E. Brown (2015). *Improving forecasts for noisy geographic time series*. Journal of Business Research, Vol. 68, pp. 1810-1818.

[82] Hurd, T.R., and J. Gleeson (2013). *On Watts' Cascade Model with Random Link Weights*. Journal of Complex Networks, Vol.1(1), pp. 25-43.

[83] Hyndman, R.J., and M.S. Ullah (2007). *Robust forecasting of mortality and fertility rates: a functional data approach*. Computational Statistics and Data Analysis, Vol. 51(10), 4942-4956.

[84] Hyndman, R.J., R.A. Ahmed, G. Athanasopoulos and H.L. Shang (2011). *Optimal combination forecasts for hierarchical time series*. Computational Statistics and Data Analysis, Vol. 55(9), pp. 2579–2589.

[85] Ieva, F., A.M. Paganoni, D. Pigoli, and V. Viteli (2013). *Multivariate functional clustering for the morphological analysis of electrocardiograph curves*. Journal of the Royal Statistical Society: Series C (Applied Statistics), Vol. 62(3), pp. 401-418.

[86] Im, K.S., M.H. Pesaran, and Y. Shin (2003). *Testing for unit roots in heterogeneous panels*. Journal of Econometrics, Vol. 115(1), pp. 53–74.

[87] Jacques, J., and C. Preda (2014). *Functional data clustering: a survey*. Advances in Data Analysis and Classification, Vol. 8(3), pp. 231-255.

[88] Jakaitienė, A., and S. Dees (2012) *Forecasting the World Economy in the Short Term*. The World Economy, Vol. 35(2), pp. 331-350.

[89] James, G.M., and C.A. Sugar (2003). *Clustering for sparsely sampled functional data*. Journal of the American Statistical Association, Vol. 98(462), pp. 397-408.

[90] Jiménez, G., and J. Mencía (2007). *Modelling the Distribution of Credit Losses with Observable and Latent Factors*. Banco de España Working Paper, No. 0709.

[91] Jobst, A.A., L.L. Ong, and C. Schmieder (2013). *A Framework for Macroprudential Bank Solvency Stress Testing: Application to S-25 and Other G-20 Country FSAPs*. International Monetary Fund, Working Paper No. 13/68.

[92] Jones, M., P. Hilbers, and G. Slack (2004). *Stress Testing Financial Systems: What to Do when the Governor Calls*. IMF Working Paper, 04/127.

[93] Kanas, A. D. Vasiliou, and N. Eriotis (2012). *Revisiting bank profitability: A semi-parametric approach*. Journal of International Financial Markets, Institutions and Money. Vol. 22(4), 990-1005.

[94] Kaufman, L., and P.J. Rousseeuw (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley Series in Probability and Statistics, Wiley, New York, pp. 68-125.

[95] Keogh, E.J, K. Chakrabarti, M. Pazzani, and S. Mehrotra (2001). *Dimensionality reduction for fast similarity search in large time series databases*. Knowledge and Information Systems, Vol. 3(3), pp. 263-286.

[96] Krishnapuram, R., H. Frigui, and O. Nasraoui (1993). *Quadratic shell clustering algorithms and the detection of second-degree curves*. Pattern Recognition Letters, Vol. 14(7), pp. 545-552.

[97] Lee, C., and M. Hsieh (2013). *The impact of bank capital on profitability and risk in Asian banking*. Journal of International Money and Finance, Vol. 32, pp. 251-281.

[98] Lehmann, H., and M. Manz (2006). *The Exposure of Swiss Banks to Macroeconomic Shocks — An Empirical Investigation*. Swiss National Bank Working Paper, No. 4.

[99] Li, M., X. Chen, X. Li, B. Ma, and P.M.B. Vitányi (2004). *The similarity metric*. IEEE Transactions on Information Theory, Vol. 50(12), pp. 3250-3264.

[100] Liao, W.T. (2005). *Clustering of time series data - a survey*. Pattern Recognition, Vol. 38, pp. 1857-1874.

[101] Liebl. B. (2013). *Modeling and forecasting electricity spot prices: a functional data perspective*. The Annals of Applied Statistics, Vol. 7(3), pp. 1562-1592.

[102] Lin, J., E.J. Keogh, S. Lonardi, and B. Chiu (2003). *A symbolic representation of time series, with implications for streaming algorithms*. In Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, DMKD'03, pp. 2-11.

[103] Lukač, N., B. Žalik, and K.R. Žalik (2015). *Sweep-hyperplane clustering algorithm using dynamic model*. Informatica, Vol. 25(4), pp. 563-580.

[104] MacKinnon, J.G. (1996). *Numerical distribution functions for unit root and cointegration tests*. Journal of Applied Econometrics. Vol 11, pp. 601–618.

[105] MacQueen, J. (1967). *Some methods for classification and analysis of multivariate observations*. In Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, pp. 281-297.

[106] Maddala, G.S., and S. Wu (1999). *A comparative study unit root tests with panel data and a new simple test*. Oxford Bulletin of Economics and Statistics, Vol. 61(S1), pp. 631-652.

[107] Maharaj, E.A. (1996). *A significance test for classifying ARMA models*. Journal of Statistical Computation and Simulation, Vol. 54(4), pp. 305-331.

[108] Marcucci, J., and M. Quagliariello (2008). *Is Bank Portfolio Riskiness Procyclical? Evidence from Italy using a Vector Autoregression*.

Journal of International Financial Markets, Institutions and Money, Vol. 18(1), pp. 46–63.

[109] Milligan, G.W., and M.C. Cooper (1985). *An examination of procedures for determining the number of clusters in a data set*. Psychometrika, Vol. 50(2), pp. 159-179.

[110] Mirzaei, A., T. Moore, and G. Liu (2013). *Does market structure matter on banks' profitability and stability? Emerging vs. advanced economies*. Journal of Banking and Finance. Vol. 37(8), pp. 2920-2937.

[111] Molyneux, P., and J. Thornton (1992). *Determinants of European Bank Profitability: A Note*. Journal of Banking and Finance, Vol. 16(6), pp. 1173–1178.

[112] Montgomery D.C., CH.L. Jennings, and M. Kulahci (2007). *Introduction to Time Series Analysis and Forecasting*, A JOHN WILEY & SONS, INC., publication, USA.

[113] Newey, W.K., K.D. West (1987). *A Simple, Positive-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix*. Econometrica, Vol. 55, pp. 703–708.

[114] Newey, W.K., K.D. West (1994). *Automatic Lag Selection in Covariance Matrix Estimation*. Review of Economic Studies, Vol. 61, pp. 631–653.

[115] Ng, R.T., and J. Han (1994). *Effecient and effictive clustering methods for spatial data mining*. In Proceedings of the 20th VLDB Conference, pp. 144-155.

[116] Ouro, H., and L.B. Schumacher (2012). *Macrofinancial Stress Testing - Principles and Practices*. International Monetary Fund Policy Paper.

[117] Pasiouras, F., and K. Kosmidou (2007). *Factors influencing the profitability of domestic and foreign commercial banks in the European Union*. Research in International Business and Finance, Vol. 21(2), pp. 222–237.

[118] Pedroni, P. (1999). *Critical values for cointegration tests in heterogeneous panels with multiple regressors*. Oxford Bulletin of Economics and Statistics, Vol. 61(S1), pp. 653–70.

[119] Pedroni, P. (2004). *Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis*. Econometric Theory, Vol. 20(3), pp. 597–625.

[120] Peng, B.J., and H.-G. Müller (2008). *Distance-based clustering of sparsely observed stochastic process, with application to online auctions*. The Annals of Applied Statistics, Vol. 2(3), pp. 1056-1077.

[121] Perman, M., I. Pelivan, and J. Arnerić (2015). *Profit persistence and determinants of bank profitability in Croatia*. Economic Research, Vol. 28(1), pp. 284-298.

[122] Pesaran, M.H., and Y. Shin (1999). *An autoregressive distributed lag modelling approach to cointegration analysis*. In (ed.) S.Strom, Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium, chapter 11, Cambridge University Press.

[123] Pesaran, M.H., Y. Shin, and R.J. Smith (1999). *Bounds testing approaches to the analysis of long run relationships*. Cambridge working papers in Economics, No. 9907.

[124] Pesaran, M.H., Y. Shin, and R.P. Smith (1997). *Pooled estimation of long-run relationships in dynamic heterogeneous panels*. Cambridge working papers in Economics, No. 9721.

[125] Pesaran, M.H., Y. Shin, and R.P. Smith (1999). *Pooled mean group estimation of dynamic heterogeneous panels*, Journal of the American Statistical Association, Vol. 94(446), pp. 621-634.

[126] Pesaran, M.H., and R.P. Smith (1995). *Estimating long-run relationships from dynamic heterogeneous panels*. Journal of Econometrics, Vol. 68(1), pp. 79–113.

[127] Pesaran, M.H., T. Schuerman, B.J. Treutler and S.M. Weiner (2006). *Macroeconomic Dynamics and Credit Risk: a Global Perspective*. Journal of Money Credit and Banking, Vol. 38(5), pp. 1211–1261.

[128] Piccolo, D. (1990). *A distance measure for classifying ARIMA models*. Journal of Time Series Analysis, Vol. 11(2), pp. 153-164.

[129] Polson, N.G., and J.G. Scott (2011). *Explosive volatility: A model of financial contagion*. University of Chicago, Manuscript.

[130] Puhr, C., and S. Schmitz (2014). *A View From The Top – The Interaction Between Solvency And Liquidity Stress*. Journal of Risk Management in Financial Institutions Vol. 7(4), pp. 38–51.

[131] Quagliariello, M. (2009). *Stress-testing the Banking System: Methodologies and Applications*. Cambridge University Press, New York.

[132] Ramsey, J., and B.W. Silverman (2005). *Functional Data Analysis*. Springer Series in Statistics, second edition, Springer, New York.

[133] Réka, A., and A.L. Barabási (2002). *Statistical mechanics of complex networks*. Review of Modern Physics, Vol. 74(1), pp. 47–97.

[134] Rivers, N., and M. Jaccard (2005). *Combining Top-down and Bottom-up Approaches to Energy-Economy Modeling Using Discrete Choise Methods*. The Energy Journal, Vol. 26(1), pp. 83-106.

[135] Rostami-Tabar, B., M.Z. Babai, Y. Dunq, and A. Syntetos (2015). *Non-stationary demand forecasting by cross-sectional aggregation*. International Journal of Production Economics, Vol. 170, pp. 297-309.

[136] Rousseeuw, P.J. (1986). *Silhouettes: a graphical aid to interpretation and validation of cluster analysis*. Journal of Computational and Applied Mathematics, Vol. 20, pp. 53-65.

[137] Rumler, F., and W. Waschiczek (2012). *Have Changes in the financial structure affected bank profitability? Evidence for Austria*. Oesterreichische Nationalbank, Workink paper series, No. 180.

[138] Samé, A., F. Chamroukhi, G. Govaert, and P. Aknin (2011). *Model-based clustering and segmentation of time series with changes in regime*. Advances in Data Analysis and Classification, Vol. 5(4), pp. 301-321.

[139] Sbrana, G., and A. Silvestrini (2013) *Forecasting aggregate demand: Analytical comparison of top-down and bottom-up approaches in a multivariate exponential smoothing framework*. International Journal of Production Economics, Vol. 146, pp. 185-198.

[140] Shang, H.L. (2013) *Functional time series approach for forecasting very short-term electricity demand*. Journal of Applied Statistics, Vol. 40(1), pp. 152-168.

[141] Shehzad, C.T., J. De Haan, and B. Scholtens (2013). *The relationship between size, growth and profitability of commercial banks*. Applied Economics, Vol. 45(13), pp. 1751-1765.

[142] Sufian, F. (2011). *Profitability of the Korean banking sector: panel evidence on bank-specific and macroeconomic determinants*. Journal of Economics and Management, Vol. 7(1), pp. 43–72.

[143] Trujillo-Ponce, A. (2013). *What determines the profitability of banks? Evidence from Spain*. Accounting and Finance, Vol. 53(2), pp. 561-586.

[144] Wang, W., J. Yang, and R. Muntz (1997). *STING: A statistical information grid approach to spatial data mining*. In Proceedings of 23rd VLDB Conference, pp. 186-195.

[145] Wasserman, L. (2006). *All of Nonparametric Statistics*. Springer Texts in Statistics, New York.

[146] White, H. (2001). *Asymptotic Theory for Econometricians*. Academic Press, USA.

[147] White, H., T.H. Kim, and S. Manganelli (2010). *VAR for VaR: Measuring systemic risk using multivariate regression quantiles*. MPRA working paper, No. 35372.

[148] Widiarta, H., S. Viswanathan, and R. Piplani (2009). *Forecasting aggregate demand: An analytical evaluation of top-down versus bottom-up forecasting in a production planning framework*. International Journal of Production Economics, Vol. 118, pp. 87-94.

[149] Zotteri, G., M. Kalchschmidt, and F. Caniato (2005). *The impact of aggregation level on forecasting performance*. International Journal of Production Economics, Vol. 93-94, pp. 479-491.

Appendix A

Table A.1: Credit risk modelling results

Variables	**Coefficient**	**Standard error**	**t-statistic**	**p-value**
Loans to industry sector (CLS_1)				
Constant	0.560	0.186	3.011	0.004
$\Delta\log(real\,GDP)_{t-1}$	-0.098	0.048	-2,031	0.048
$\Delta(Unemployment\,rate)_{t-2}$	0.300	0.139	2,154	0.036
$\Delta\log(Inflation)_{t-1}$	-0.238	0.132	-1.806	0.077
Loans to trade sector (CLS_2)				
Constant	0.315	0.105	2.998	0.004
$\Delta\log(nominal\,GDP)_{t-1}$	-0.071	0.025	-2.827	0.007
$\Delta(Unemployment\,rate)_{t-1}$	0.201	0.050	4.068	0.000
Loans to other non-financial corporations (CLS_5)				
Constant	0.770	0.211	3.655	0.001
Annual nominal GDP growth	-0.079	0.020	-4.037	0.000
Annual inflation growth$_{t-2}$	0.059	0.031	1.805	0.077
Loans to households - housing loans (CLS_6)				
Constant	-0.141	0.049	-2.913	0.005
Unemployment rate	0.016	0.004	4.247	0.000
Short term interest rate$_{t-2}$	0.021	0.007	3.155	0.003
Loans to households - consumer and other loans (CLS_7)				
Constant	0.422	0.079	5.359	0.000
Annual private consumption growth	-0.019	0.007	-2.818	0.007
Annual wages growth$_{t-2}$	-0.016	0.006	-2.765	0.008
Short term interest rate$_{t-2}$	0.056	0.015	3.789	0.000

Note: the sample period is 2000q1-2012q4. Δ - quarterly difference.
Source: Statistics Lithuania, bank data and author's calculation.

Table A.2: Description of the macroeconomic variables

Variables	Description
Real GDP	Lithuanian gross domestic product (constant prices; adjusted for seasonal and workday effects)
Nominal GDP	Lithuanian gross domestic product (current prices; adjusted for seasonal and workday effects)
Private consumption	Lithuanian private consumption (constant prices; adjusted for seasonal and workday effects)
Export	Lithuanian export (constant prices; adjusted for seasonal and workday effects)
Unemployment rate	Unemployment rate as a share of labour force (adjusted for seasonal effects)
Wages	Compensation of employees (current prices)
Inflation	Harmonised Index of Consumer Prices (HICP)
Short term interest rate	3 month Euribor rate

Source: Statistics Lithuania, Bloomberg.

Appendix B

Table B.1: Descriptive statistics

Dependent variables	Mean	Std. dev.	Min	Max
log(NII)	9.79	1.23	6.97	12.03
log(NCI)	8.79	1.30	6.41	11.04
log(OE)	9.62	1.04	7.53	11.37
Independent variables				
log(GDP)	9.89	0.09	9.71	10.02
log(REX)	9.48	0.23	9.04	9.88
$HICP$	3.84	3.14	-1.10	12.26
STI	3.15	2.20	0.40	8.02
log(CPE)	8.90	0.21	8.42	9.14
UNR	10.75	4.47	4.17	18.15
log(A)	15.15	1.31	11.98	17.07
log(NLS)	14.83	1.37	11.51	16.87
CL	0.27	0.65	-1.45	4.72
log(PRO)	10.79	2.28	1.39	14.43
HHI	0.19	0.01	0.16	0.21

Source: Statistics Lithuania, bank data and author's calculation.

Table B.2: Pair wise correlations

	log(NII)	log(NCI)	log(OE)
log(GDP)	0.311	0.224	0.268
log(REX)	0.265	0.230	0.257
$HICP$	0.182	0.070	0.112
STI	0.061	0.002	0.023
log(CPE)	0.335	0.260	0.307
UNR	0.026	0.055	0.058
log(A)	0.955	0.880	0.934
log(NLS)	0.946	0.853	0.918
CL	-0.094	-0.045	-0.072
log(PRO)	0.694	0.702	0.720
HHI	-0.143	-0.078	-0.108

Source: Statistics Lithuania, bank data and author's calculation.

Table B.3: Panel unit root tests (bank-specific variables)

Variable	Im, Pesaran and Shin		Fisher ADF		Fisher PP	
	Level	First differences	Level	First differences	Level	First differences
Loan stock (net)	2.644	-5.548	3.837	68.41	8.522	79.43
	(0.996)	*(0.000)*	*(0.999)*	*(0.000)*	*(0.932)*	*(0.000)*
Total assets	0.261	-6.849	12.98	84.58	8.344	131.5
	(0.603)	*(0.000)*	*(0.674)*	*(0.000)*	*(0.938)*	*(0.000)*
Provisions	1.210	-5.727	12.05	68.77	3.025	106.6
	(0.887)	*(0.000)*	*(0.741)*	*(0.000)*	*(0.999)*	*(0.000)*
Credit loss	-2.536	-20.976	31.79	440.5	58.19	955.8
	(0.006)	*(0.000)*	*(0.011)*	*(0.000)*	*(0.000)*	*(0.000)*

Notes: p-values are reported in the parenthesis. For Im, Pesaran and Shin, Fisher ADF panel unit root tests number of lags was selected using the AIC criterion. Panel unit root tests include intercept and trend.
Source: Statistics Lithuania, bank data and author's calculation.

Table B.4: Unit root test (macroeconomic variables)

Variable	ADF	
	Level	First differences
log(GDP)	-2.520 *(0.317)*	-4.463 *(0.005)*
log(REX)	-2.515 *(0.320)*	-5.233 *(0.001)*
$HICP$	-3.098 *(0.121)*	-3.924 *(0.020)*
STI	-1.799 *(0.687)*	-4.426 *(0.006)*
log(CPE)	-1.666 *(0.748)*	-5.257 *(0.001)*
UNR	-2.813 *(0.201)*	-2.796 *(0.207)*
HHI	-1.436 *(0.834)*	-5.616 *(0.000)*

Notes: p-values are reported in the parenthesis. Number of lags was selected using the AIC criterion. Unit root test includes intercept and trend.
Source: Statistics Lithuania, bank data and author's calculation.

Table B.5: Panel co-integration tests (net interest income)

	Panel ν-Statistic	Panel ρ-Statistic	Panel PP-Statistic	Panel ADF-Statistic	Group ρ-Statistic	Group PP-Statistic	Group ADF-Statistic
log(GDP)	0.731	0.013	0.000	0.000	0.702	0.054	0.691
log(REX)	0.774	0.005	0.000	0.000	0.845	0.523	0.511
$HICP$	0.526	0.000	0.000	0.000	0.179	0.020	0.016
STI	0.335	0.000	0.000	0.000	0.463	0.147	0.019
log(CPE)	0.002	0.000	0.000	0.000	0.020	0.001	0.000
UNR	0.903	0.001	0.000	0.000	0.773	0.670	0.318
log(A)	0.560	0.000	0.000	0.000	0.005	0.000	0.000
log(NLS)	0.716	0.000	0.000	0.000	0.032	0.002	0.000
CL	0.000	0.312	0.898	0.831	0.941	0.992	0.336
log(PRO)	0.599	0.000	0.000	0.000	0.504	0.141	0.019
HHI	0.718	0.001	0.000	0.000	0.704	0.313	0.301

Notes: p-values are reported. Number of lags was selected using the AIC criterion. Co-integration tests include intercept and trend.
Source: Statistics Lithuania, bank data and author's calculation.

Table B.6: Panel co-integration tests (net fee and commission income)

	Panel ν-Statistic	Panel ρ-Statistic	Panel PP-Statistic	Panel ADF-Statistic	Group ρ-Statistic	Group PP-Statistic	Group ADF-Statistic
log(GDP)	0.217	0.000	0.000	0.780	0.000	0.000	0.000
log(REX)	0.577	0.000	0.000	0.760	0.260	0.009	0.439
$HICP$	0.000	0.000	0.000	0.544	0.005	0.000	0.030
STI	0.090	0.000	0.000	0.000	0.014	0.000	0.000
log(CPE)	0.056	0.000	0.000	0.000	0.000	0.000	0.000
UNR	0.024	0.000	0.000	0.954	0.001	0.000	0.160
log(A)	0.054	0.000	0.000	0.000	0.000	0.000	0.000
log(NLS)	0.067	0.000	0.000	0.000	0.000	0.000	0.000
CL	0.333	0.000	0.000	0.000	0.453	0.477	0.373
log(PRO)	0.002	0.000	0.000	0.000	0.016	0.000	0.002
HHI	0.068	0.000	0.000	0.000	0.245	0.008	0.007

Notes: p-values are reported. Number of lags was selected using the AIC criterion. Co-integration tests include intercept and trend.
Source: Statistics Lithuania, bank data and author's calculation.

Table B.7: Panel co-integration tests (operating expenses)

	Panel *nu*-Statistic	Panel ρ-Statistic	Panel PP-Statistic	Panel ADF-Statistic	Group ρ-Statistic	Group PP-Statistic	Group ADF-Statistic
log(GDP)	0.733	0.080	0.020	0.557	0.058	0.008	0.540
log(REX)	0.958	0.882	0.520	0.336	0.944	0.665	0.144
$HICP$	0.079	0.053	0.014	0.015	0.155	0.024	0.081
STI	0.030	0.009	0.001	0.002	0.095	0.003	0.010
log(CPE)	0.065	0.000	0.000	0.073	0.000	0.000	0.006
UNR	0.553	0.514	0.227	0.602	0.428	0.132	0.483
log(A)	0.159	0.000	0.000	0.000	0.000	0.000	0.000
log(NLS)	0.212	0.000	0.000	0.000	0.000	0.000	0.000
CL	0.903	0.943	0.929	0.980	0.986	0.981	0.983
log(PRO)	0.379	0.321	0.049	0.290	0.591	0.121	0.531
HHI	0.680	0.568	0.213	0.226	0.696	0.240	0.263

Notes: p-values are reported. Number of lags was selected using the AIC criterion. Co-integration tests include intercept and trend.
Source: Statistics Lithuania, bank data and author's calculation.

Table B.8: Panel co-integration tests (long term variables of NII equation)

Panel ν-Statistic	Panel ρ-Statistic	Panel PP-Statistic	Panel ADF-Statistic	Group ρ-Statistic	Group PP-Statistic	Group ADF-Statistic
-0.967	-3.457	-10.667	-10.155	-1.306	-5.212	-5.043
(0.833)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.096)*	*(0.000)*	*(0.000)*

Notes: p-values are reported in parenthesis. Number of lags was selected using the AIC criterion. Co-integration tests include intercept and trend.
Source: Statistics Lithuania, bank data and author's calculation

Table B.9: Panel co-integration tests (long term variables of NCI equation)

Panel ν-Statistic	Panel ρ-Statistic	Panel PP-Statistic	Panel ADF-Statistic	Group ρ-Statistic	Group PP-Statistic	Group ADF-Statistic
1.039	-10.083	-16.715	-17.527	-3.527	-7.694	-6.630
(0.149)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*

Notes: p-values are reported in parenthesis. Number of lags was selected using the AIC criterion. Co-integration tests include intercept and trend.
Source: Statistics Lithuania, bank data and author's calculation

Table B.10: Panel co-integration tests (long term variables of OE equation)

Panel ν-Statistic	Panel ρ-Statistic	Panel PP-Statistic	Panel ADF-Statistic	Group ρ-Statistic	Group PP-Statistic	Group ADF-Statistic
0.976	-5.862	-9.143	-8.637	-5.446	-13.468	-8.216
(0.165)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*

Notes: p-values are reported in parenthesis. Number of lags was selected using the AIC criterion. Co-integration tests include intercept and trend.
Source: Statistics Lithuania, bank data and author's calculation

Appendix C

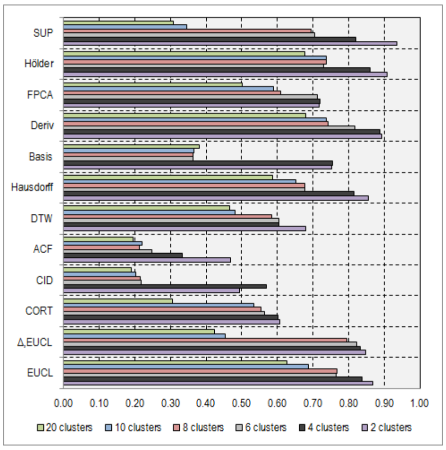

Figure C.1: Average silhouette width (ROAA)

Source: Bankscope data and authors' calculation.

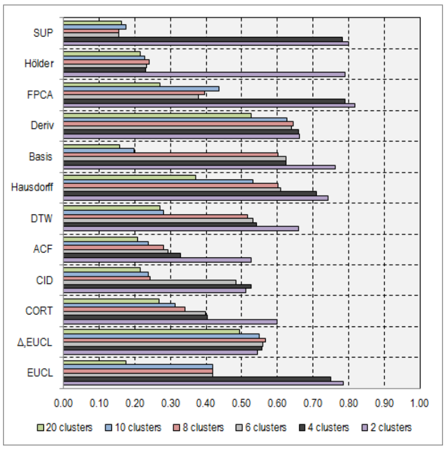

Figure C.2: Average silhouette width (ROAE)

Source: Bankscope data and authors' calculation.

Table C.1: Number of clusters suggested by Dunn index

	EUCL	**Δ, EUCL**	**CORT**	**CID**	**ACF**	**DTW**	**Haus-dorff**	**Basis**	**Deriv**	**FPCA**	**Höl-der**	**SUP**
ROAA	2	8	4	20	10	10	4	4	4	6	2	2
ROAE	4	8	8	6	20	20	10	2	20	4	2	4
NIM	2	20	4	20	10	20	2	2	4	4	2	2
CIR	20	20	2	20	20	2	20	20	6	8	2	20
LLP	2	2	2	20	20	2	4	2	10	2	8	8
CAR	2	2	20	10	20	4	2	4	2	4	2	2
Multivariate	2	2	10	8	20	2	2	10	2	2	2	4

Source: Bankscope data and authors' calculation.

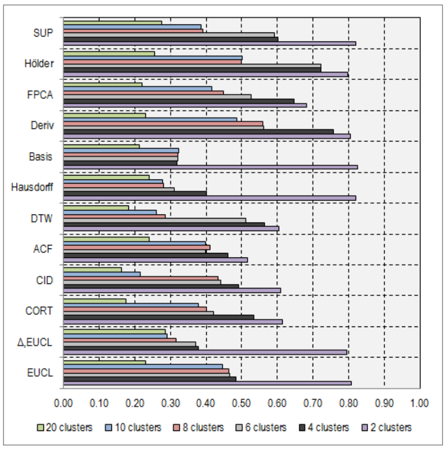

Figure C.3: Average silhouette width (NIM)

Source: Bankscope data and authors' calculation.

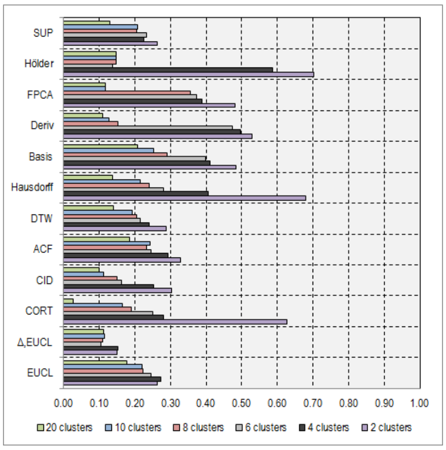

Figure C.4: Average silhouette width (CIR)

Source: Bankscope data and authors' calculation.

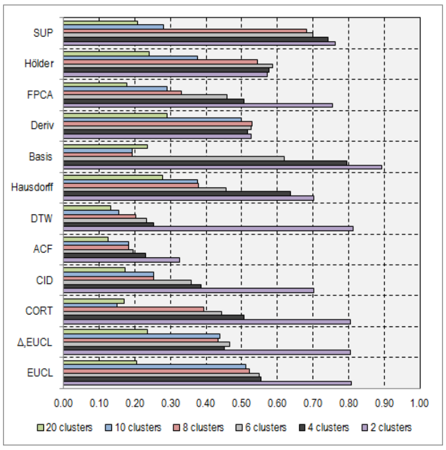

Figure C.5: Average silhouette width (LLP)

Source: Bankscope data and authors' calculation.

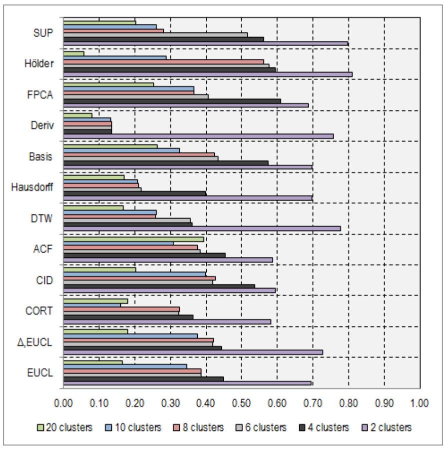

Figure C.6: Average silhouette width (CAR)

Source: Bankscope data and authors' calculation.

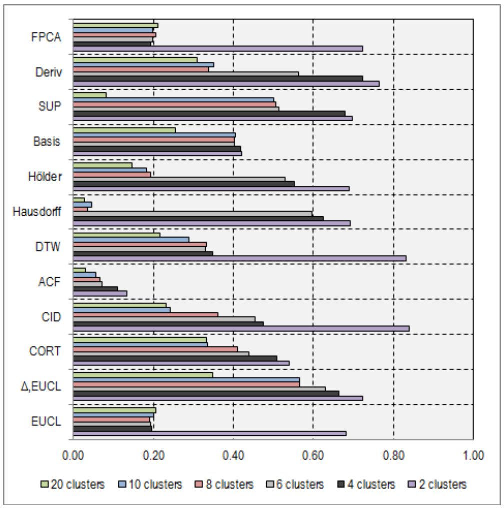

Figure C.7: Average silhouette width (multivariate case)

Source: Bankscope data and authors' calculation.

Table C.2: Number of clusters suggested by Caliński and Harabasz index

	EUCL	Δ, EUCL	CORT	CID	ACF	DTW	Haus-dorff	Basis	Deriv	FPCA	Höl-der	SUP
ROAA	10	2	20	6	2	2	2	6	6	20	2	2
ROAE	6	2	6	6	2	10	2	4	10	10	2	4
NIM	6	4	2	2	8	20	4	10	2	2	2	4
CIR	4	2	2	2	4	20	2	2	4	2	8	4
LLP	4	2	20	2	2	20	2	2	2	4	20	2
CAR	2	4	2	2	20	20	2	2	2	2	4	20
Multivariate	4	8	2	8	4	6	2	2	4	4	2	2

Source: Bankscope data and authors' calculation.

Appendix D

Proof of Theorem 1. Noting that

$$X_j(\tau_k) = \sum_{i=0}^{k-1} \beta_j^i [\boldsymbol{\gamma}_j' \boldsymbol{Z}_{k-i} + \varepsilon_{j,k-i}]$$

for any $k = 1, \ldots, n$, and summing along $j = 1, \ldots, N$, we obtain

$$N^{-1} \sum_{j=1}^{N} X_j(\tau_k) - Y_k = \sum_{i=0}^{k-1} U_{Ni} + \sum_{i=0}^{k-1} V_{Ni}$$

where

$$U_{Ni} = N^{-1} \sum_{j=1}^{N} \left[\beta_j^i \boldsymbol{\gamma}_j' - \mathsf{E}\,(\beta_1^i) \mathsf{E}\,(\boldsymbol{\gamma}_1')\right] \boldsymbol{Z}_{k-i}, \quad V_{Ni} = N^{-1} \sum_{j=1}^{N} \beta_j^i \varepsilon_{j,k-i}.$$

Since k is fixed one needs to prove that for any $i = 0, \ldots, k-1$

$$\mathsf{E}\,(U_{Ni}^2) \to 0, \quad \mathsf{E}\,(V_{Ni}^2) \to 0 \quad \text{as} \quad N \to \infty.$$

We have

$$\mathsf{E}\,(U_{Ni}^2) \leq \mathsf{E}\,||\boldsymbol{Z}_{k-i}||^2 \mathsf{E} \left\| N^{-1} \sum_{j=1}^{N} \left[\beta_j^i \boldsymbol{\gamma}_j' - \mathsf{E}\,(\beta_1^i) \mathsf{E}\,(\boldsymbol{\gamma}_1')\right] \right\|^2$$

where $||x||$ denotes the Euclidean norm in $\mathbb{R}^p$. Since the random vectors

$\beta_1^i \boldsymbol{\gamma}_1, \ldots, \beta_N^i \boldsymbol{\gamma}_N$ are independent and identically distributed, we have

$$\mathsf{E} \left\| N^{-1} \sum_{j=1}^{N} \left[\beta_j^i \boldsymbol{\gamma}_j' - \mathsf{E}\,(\beta_1^i) \mathsf{E}\,(\boldsymbol{\gamma}_1') \right] \right\|^2 \leq N^{-1} \mathsf{E}\,(\beta_1^{2i}) \mathsf{E}\,||\gamma_1||^2.$$

Hence, $\mathsf{E}\,(U_{Ni}^2 \to 0$ as $N \to \infty$. Next observing that

$$\mathsf{E}\,(V_{Ni}^2) = N^{-1} \sigma^2 \mathsf{E}\,(\beta_1^{2i})$$

we see that $\mathsf{E}\,(V_{Ni}^2) \to 0$ as $N \to \infty$. This completes the proof. □

Proof of Theorem 2. The proof goes along the lines of the proof of Theorem 1. □

Next we consider the estimates $\widehat{\beta}_j, j = 1, \ldots, N$ and $\widehat{\boldsymbol{\gamma}}_j, j = 1, \ldots, N$, obtained from (2.30). To this aim assume that for each fixed $j = 1, \ldots, N$, the observation Y_{jk} is associated with a vector of known d-dimensional covariates $\boldsymbol{X}_{jk}$ through the following equation:

$$Y_{jk} = \boldsymbol{X}_{jk}' \boldsymbol{b}_j + \varepsilon_{jk}, \quad k = 1, \ldots, n, \tag{2.45}$$

where $\boldsymbol{b}_j := (b_{j1}, \ldots, b_{jd})'$ is a d-dimensional random vector of regression coefficients and ε_{jk} represents an error. It is assumed that the errors are iid with mean 0 and constant variance $\sigma^2 > 0$. Moreover we assume independence of $\boldsymbol{b}_j, (\boldsymbol{X}_{jk}), (\varepsilon_{jk})$.

Let $Y_j = (Y_{jk})_{1 \leq k \leq n}$ be the vector of observations, $\mathcal{X}_j = (\boldsymbol{X}_{jk}')_{1 \leq k \leq n}$ be the matrix of covariates, and $\boldsymbol{\varepsilon}_j$ be the vector of errors. Then the linear regression (2.45) can be expressed as

$$Y_j = \mathcal{X}_j \boldsymbol{b}_j + \boldsymbol{\varepsilon}_j.$$

The least square method finds the regression coefficients $\boldsymbol{b}_j$ that minimize

$$||Y_j - \mathcal{X}_j \boldsymbol{b}_j||^2 = \sum_{k=1}^{n} (Y_{jk} - \boldsymbol{X}_{jk}' \boldsymbol{b}_j)^2.$$

If $\mathcal{X}_j$ is of full rank d, the solution, which is called the least square estimator of $\boldsymbol{b}_j$, can be expressed as

$$\widehat{\boldsymbol{b}}_j = (\mathcal{X}_j'\mathcal{X}_j)^{-1}\mathcal{X}_j'Y_j = \left(\sum_{k=1}^{n} \boldsymbol{X}_{jk}\boldsymbol{X}_{jk}'\right)^{-1} \sum_{i=1}^{n} \boldsymbol{X}_{ji}Y_{ji}.$$

Notting that

$$\widehat{\boldsymbol{b}}_j = \boldsymbol{b}_j + (\mathcal{X}_j'\mathcal{X}_j)^{-1}\mathcal{X}_j'\boldsymbol{\varepsilon}_j$$

and using independence of $\boldsymbol{\varepsilon}$ and $\mathcal{X}$ we have

$$\mathsf{E}\,[\widehat{\boldsymbol{b}}_j - \boldsymbol{b}_j] = 0 \quad \text{and} \quad \operatorname{cov}(\widehat{\boldsymbol{b}}_j - \boldsymbol{b}_j) = \mathsf{E}\,(\mathcal{X}_j'\mathcal{X}_j)^{-1}),$$

provided that $\mathsf{E}\,(\mathcal{X}_j'\mathcal{X}_j)^{-1})$ is well defined.

The estimators $\widehat{\boldsymbol{b}}_1, \ldots, \widehat{\boldsymbol{b}}_N$ can be used to perform statistical inference on distributional properties of random sample $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_N$.

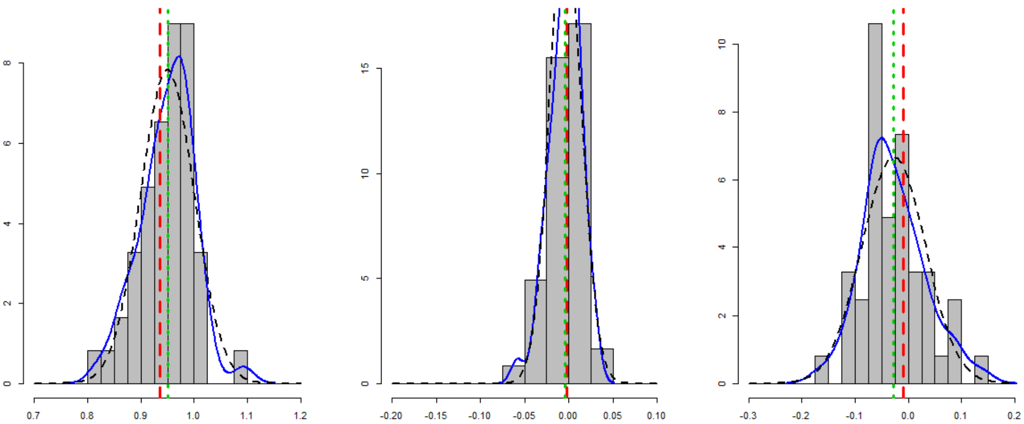

Figure C.8: Histogram of the estimated coefficients (β_j, γ_{j1} and γ_{j2}, respectivelly) from the bottom-up equation 2.40 with $\tau_k = 57$

Notes: Blue solid line - empirical density of the coefficients; black dashed line - density from the normal distribution; green dotted line - mean of estimated BU coefficients; red dashed line - corresponding estimated coefficient from the top-down equation (2.42).

Source: authors' calculation.

Table D.1: Forecasting error statistics (top-down)

	Cluster 1		Cluster 2		Cluster 4		Cluster 5		Cluster 6		Cluster 10	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
2 points forecasted	0.099	0.098	0.184	0.165	0.455	0.442	0.388	0.300	0.848	0.818	0.334	0.330
4 points forecasted	0.114	0.109	0.388	0.296	0.671	0.628	0.252	0.180	1.148	1.051	0.523	0.486
8 points forecasted	0.106	0.100	0.554	0.440	0.810	0.743	0.291	0.220	1.120	1.008	0.512	0.469
24 points forecasted	0.099	0.092	0.656	0.519	0.876	0.789	0.421	0.324	1.060	0.939	0.481	0.435

Source: authors' calculation.

Table D.2: Forecasting error statistics (bottom-up)

	Cluster 1		Cluster 2		Cluster 4		Cluster 5		Cluster 6		Cluster 10	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
2 points forecasted	0.243	0.237	0.387	0.294	0.307	0.302	0.883	0.882	0.867	0.844	0.556	0.547
4 points forecasted	0.228	0.214	0.633	0.500	0.277	0.246	0.192	0.170	1.024	0.957	0.628	0.592
8 points forecasted	0.238	0.218	0.600	0.462	0.521	0.487	0.187	0.145	1.080	0.980	0.664	0.609
24 points forecasted	0.248	0.222	0.578	0.439	0.667	0.615	0.217	0.154	1.018	0.905	0.641	0.574

Source: authors' calculation.

Table D.3: Forecasting error statistics (Functional regression model)

	Cluster 1		Cluster 2		Cluster 4		Cluster 5		Cluster 6		Cluster 10	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
2 points forecasted	0.819	0.788	0.155	0.135	6.556	6.358	3.156	3.045	0.646	0.610	0.200	0.176
4 points forecasted	0.723	0.683	0.161	0.150	5.852	5.504	2.790	2.642	0.557	0.525	0.162	0.143
8 points forecasted	0.674	0.631	0.162	0.153	5.488	5.068	2.605	2.443	0.514	0.486	0.143	0.129
24 points forecasted	0.642	0.597	0.162	0.154	5.239	4.775	2.480	2.312	0.486	0.462	0.132	0.121

Source: authors' calculation.