RUHR
ECONOMIC PAPERS

Ulrich Haskamp

Improving the Forecasts of European Regional Banks' Profitability with Machine Learning Algorithms

#705

Imprint

Ruhr Economic Papers

Published by

RWI Leibniz-Institut für Wirtschaftsforschung
Hohenzollernstr. 1-3, 45128 Essen, Germany

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #705

Responsible Editor: Volker Clausen

ISSN 1864-4872 (online) – ISBN 978-3-86788-819-6

Ruhr Economic Papers #705

Ulrich Haskamp

Improving the Forecasts of European Regional Banks' Profitability with Machine Learning Algorithms

Ruhr Graduate School in Economics

UNIVERSITÄT DUISBURG ESSEN

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

Das RWI wird vom Bund und vom Land Nordrhein-Westfalen gefördert.

http://dx.doi.org/10.4419/86788819
ISSN 1864-4872 (online)
ISBN 978-3-86788-819-6

Ulrich Haskamp[1]

Improving the Forecasts of European Regional Banks' Profitability with Machine Learning Algorithms

Abstract

Regional banks as savings and cooperative banks are widespread in continental Europe. In the aftermath of the financial crisis, however, they had problems keeping their profitability which is an important quantitative indicator for the health of a bank and the banking sector overall. We use a large data set of bank-level balance sheet items and regional economic variables to forecast protability for about 2,000 regional banks. Machine learning algorithms are able to beat traditional estimators as ordinary least squares as well as autoregressive models in forecasting performance.

JEL Classification: C53, C55, G21

Keywords: Profitability; regional banking; forecasting; machine learning

July 2017

1 Ulrich Haskamp, RGS Econ and UDE. – All correspondence to: Ulrich Haskamp, University of Duisburg Essen, Universitätsstr. 12, 45117 Essen, Germany, e-mail: ulrich.haskamp@uni-due.de

1. Introduction

Banks in Europe suffer from low profitability. Further, more than five percent of all credits in Europe are considered as non-performing by the EBA (2016) and deteriorate banks' ability to build up equity. Profitability ratios are highly dispersed across European countries and regions. This also holds true for regional banks as cooperative and savings banks. Furthermore, the recent financial crisis has had a strong regional dimension. While some regions have recovered quickly from the crisis, output in other regions is still much below its pre-crisis level affecting banks' profitability and vice versa (Hasan et al., 2009; Bolt et al., 2012; Hakenes et al., 2015; Belke et al., 2016). Lastly, also the ECB's and the Swedish Sveriges Riksbank's zero interest rate policy diminishes banks opportunities to generate profits - especially for banks relying on traditional business models as regional banks (Demirgüc-Kunt & Huizinga, 1999; Genay, 2014; Alessandri & Nelson, 2015). Our paper, therefore, deals with regional bank profitability from a forecasting perspective. Getting sound forecasts for banks' profitability helps economic agents and regulators in assessing the future stance of the banking system. However, in the literature there are only a few studies devoted to banks which, moreover, look at listed banks only (Flannery et al., 2013; Anolli et al., 2014).

We base our forecasting exercise on the data of about 2,000 savings and cooperative banks from eight European countries. The data includes the 2000-2015 period which includes the interesting pre-crisis sample, the financial crisis and the "new normal of monetary policy". We include bank-level balance sheet items as well as regional variables in our estimations following Hasan et al. (2009) and Belke et al. (2016). Regional variables as GDP growth are likely to influence banks' profitability. Then we use the estimations to forecast profitability using classic methods as ordinary least squares (OLS), autoregressive models (AR) and state-of-the-art machine learning techniques. We examine whether the latter are able to provide better forecasts and, if so, during which times they outperform. In the literature such techniques have been shown to perform well in forecasting exercises (Khandani et al., 2010; Butaru et al., 2016; Fitzpatrick & Mues, 2016).

The outline of this work is as follows: Firstly, the used data set is introduced. Secondly, the empirical methodology and results are presented. Lastly, concluding remarks will be given.

2. Data

We have obtained unconsolidated financial data for 2,349 European savings and cooperative banks from Bankscope. The covered period is from 2000 until 2015 such that we can study the pre- and post-crisis period. For earlier years the coverage is slightly more limited (Englmaier & Stowasser, 2017). Furthermore, Bankscope includes information about whether a bank is a commercial, savings, cooperative or other bank. Nevertheless, we had to correct this information for some banks (Haskamp, 2016). Most of the banks in our sample are active for the complete sample which is to some extent due to numerous mergers during this time period (Koetter, 2008; Schmieder et al., 2010; Behr & Heid, 2011). Continental countries as Denmark, the Netherlands and Portugal have been neglected due to missing data.

Following Hasan et al. (2009) and Belke et al. (2016), the Bankscope database allows us to map European banks to NUTS 2 regions.[1] Bankscope contains information on the location of banks' headquarters. In almost all countries we were able to use zip codes to match banks to a specific region. In the other countries we matched banks by the headquarters' city name. In total we mapped the banks to 133 NUTS 2 regions which yields an average number of banks per region of about 17. Bank-based economies as Austria, Germany or Italy contain relatively more banks per regions due to their large savings and cooperative banking sectors. Then, we used this information to merge the bank-level data with the regional data which we obtained from Eurostat. Koetter & Wedow (2010), for example, examined whether German banks' operate across multiple NUTS 2 regions. Their approach was to check whether a bank's branches are located in the same *Raumordnungsregion*[2] which are even smaller than NUTS 2 regions. Using Bundesbank data they found that 93% of all cooperative banks' branches and 97% of the savings banks' branches lie into the same *Raumordnungsregion*. Therefore, we draw the conclusion that our mapping approach is eligible.

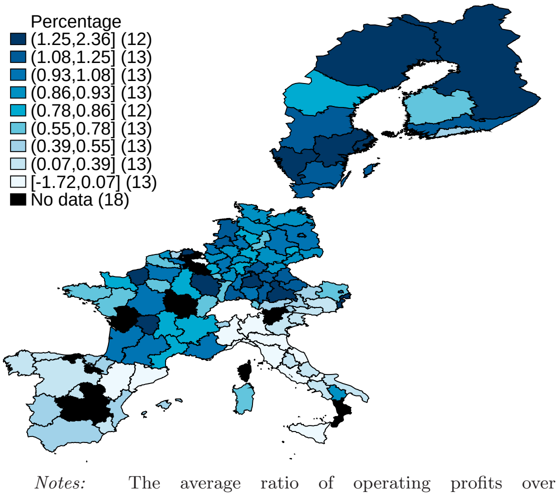
The obtained bank-level balance sheet data includes items as total assets, loans, equity ratio, interest income over average loans and others. The variable of interest, profitability, is defined as operating profits over total assets. The corresponding regional data incor-

[1] A NUTS 2 region has a population between 800,000 and 3 million persons. We use the NUTS version 2010 which is the latest one for the NUTS 2 regions we are considering.

[2] These are aggregations of NUTS 3 regions. They are created based on economic interdependencies between districts.

porates GDP and population growth. Savings and cooperative banks are similar in their characteristics and business models. Both are very active in lending as their loans to non-banks over total assets ratio is high. Further, savings banks seem to be slightly more risk averse as their interest charged over average loans and loan impairment ratio is lower while their ratio of mortgage loans over all loans is higher. The biggest difference is that savings banks are generally larger than their cooperative counterparts (Haskamp, 2016; Englmaier & Stowasser, 2017). To give also a graphical overview, we present profitability aggregated over NUTS 2 regions for the year 2013 in Figure 1. It can be seen that regional banks in Spain and Italy are less profitable than in other European countries. Regional differences do exists also within countries.

Figure 1: European Regional Banks' Operating Profits Ratio in European NUTS 2 Regions



Notes: The average ratio of operating profits over total assets of all the regional banks in our sample within the examined European NUTS 2 regions in 2013.

3. Empirical Evidence

3.1. Estimation Methodology

Our main empirical panel specification looks as follows:

$$y_{i,t+1} = \alpha + \beta y_{i,t} + \gamma X_{i,t} + \epsilon_{i,t} \tag{1}$$

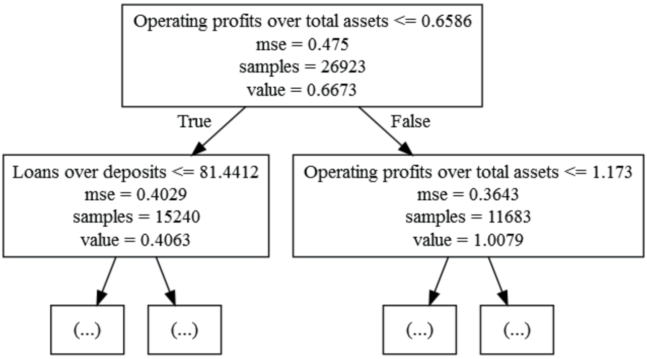
with $y_{i,t+1}$ being the dependent variable, profitability, for bank i in year t. Further, we include bank-specific and regional controls in vector $X_{i,t}$. The control variables used are the following: Firstly, we control for standard bank-level characteristics such as size measured by the log of total assets and its squared representations as well as current profitability before and after tax. Further, to include risk measures of the bank's business, we incorporate the equity over total assets ratio, interest income over loans ratio, the fees over loans ratio and the loan impairment ratio. To control for differing business models, we include the loans over assets and loans over deposits ratios. Furthermore, we add the GDP per capita ratio and regional growth rates of GDP and population on the NUTS 2 level as well as country dummies.

The next step in our framework is to use the estimation results to forecast profitability for the next year. We use all data which is available up to this year. In a robustness check we also use a rolling window procedure where we neglect data for the estimation which is more than five years older.

Instead of utilizing simple OLS or AR models, we employ the open-source library scikit-learn of the Python programming language to apply machine learning algorithms which were found to be fast and efficient also for big data sets and nonlinearities. Further, they are specifically useful for forecasting (Einav & Levin, 2013; Varian, 2014).[3] The models we employ in our forecasting exercise are decision tree, random forest and gradient boosting regressors. We present in Figure 2 an example decision tree for our task. Our algorithm chose that the current operating profits over total assets (OPoTA) is the best variable to start splitting the sample. If the variable is bigger than 0.6586% in the first node, we again look at operating profits in the second node which shows how decision trees deal with nonlinearities. In the other case the algorithm chose to split the sample again according to the loans over deposits ratio. Algorithms for constructing such decision trees work top-down. This means that it chooses a variable at each step that splits the sample best. We follow the default setting of using the mean squared error as the criterion to split the sample. The algorithm continues to split until no more useful information can be gained or the set maximum depth of the tree is reached.

[3] We continue to use the standard terminology in econometrics and speak about variables instead of features.

Figure 2: Example of a Decision Tree Estimation



A problem of decision trees is that they tend to overfit the data when it is allowed for lots of branches just as regression do if one has n observations and n variables. A random forest regressor uses multiple trees and tries to deal with overfitting by adding randomness. Firstly, the random forest regressor chooses a bootstrap sample of the observations and starts to grow a tree. Secondly, at each node of the tree a random subset of all variables is chosen for the next split which is therefore not necessarily the best split among all variables, but introduces randomness. These two steps are repeated until one has grown multiple (a forest of) trees. Then the trees' majority vote decides about the prediction value. Lastly, we also use a gradient boosting regressor which specifically puts more weight on misclassified observations during each repetition (Varian, 2014). We use the following specification for our machine learning algorithms: 15 variables which can be considered, a maximum of 15 nodes and a maximum depth of the tree of 10. Other specifications did not improve our results considerably.

3.2. Evaluating Forecasting Performance

In a last step the forecasting performance of the models is evaluated by calculating the mean absolute prediction error (MAPE). For a sample with T observations the MAPE of P forecasts is calculated as follows:

$$\hat{\sigma}_m = P^{-1} \sum_{t=T-P}^{T} |\hat{e}_{m,t+1}| \tag{2}$$

Where $\hat{e}_{m,t+1}$ is the one-year ahead forecast error of the respective model m. For the random walk the forecast is that profitability remains unchanged, $\hat{e}_{rw,t+1} = \Delta y_{i,t+1}$. With the Diebold-Mariano test (1995) it can be tested whether the MAPE of a model is different from a competing model. The null hypothesis is that both models' MAPE are equal which is tested against the alternative that model m has a higher MAPE:

$$H_0 : \sigma_m - \sigma_n = 0 \tag{3}$$

$$H_1 : \sigma_m - \sigma_n > 0 \tag{4}$$

For the construction of the Diebold-Mariano (DM) test statistic, the average difference between the forecast errors of the competing models is defined as:

$$\bar{f} = P^{-1} \sum_{t=T-P}^{T} (\hat{e}_{m,t+1} - \hat{e}_{n,t+1}) = \hat{\sigma}_m - \hat{\sigma}_n \tag{5}$$

Then the DM test statistic can be computed as follows:

$$DM = \frac{\bar{f}}{\sqrt{P^{-1}\hat{V}}} \quad \text{with} \quad \hat{V} = P^{-1} \sum_{t=T-P}^{T} ((\hat{e}_{m,t+1} - \hat{e}_{n,t+1}) - \bar{f})^2 \tag{6}$$

The denominator represents the standard error of the difference between the competing models' forecast errors. The test statistic of the Diebold-Mariano test can easily be gained by regressing the difference of the forecast errors on a constant (Elliot et al., 2006). Because the DM statistic is asymptotically normal distributed, the resulting t-statistic of the regression can be used to test the above mentioned hypothesis (Diebold & Mariano, 1995).

4. Results

We present our baseline results for Equation (1) in Table 1. In the table the MAPE of the random walk, an AR(1) statistical model, the OLS model and the three machine learning techniques are presented. We show the MAPE values for three different specifications. In contrast to the baseline model, the bank-level variables specification does not include regional variables. The last model uses a five year rolling window. All models are able to get considerably better forecasts than the random walk and a simple AR(1) model. Furthermore, we found that the forecasting performance shrinks if we use a rolling window for the estimation of the models. The best results are delivered by the gradient boosting algorithm. Furthermore, all machine learning models are according to the DM test statistic statistically significantly better in forecasting than the OLS model.[4]

Table 1: Mean Absolute Forecasting Error of Different Estimation Strategies

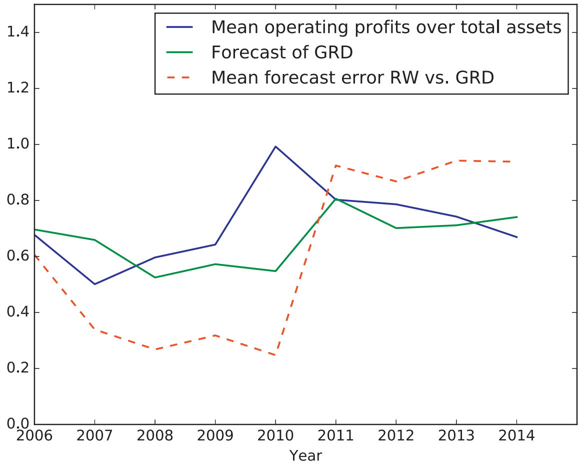
Specification	RW	AR(1)	OLS	RT	RF	GRD
Baseline	0.964	0.426	0.376	0.369	0.366	0.358
Bank-Level Variables	0.964	0.426	0.379	0.374	0.359	0.353
Rolling Window	0.964	0.429	0.389	0.402	0.386	0.374

Notes: The abbreviations stand for the random walk (RW), ordinary least squares (OLS), random tree (RT), random forest (RF) and gradient boosting (GRD). In contrast to the baseline specification, the second specification uses only bank-level variables and the third one used a five year rolling window for the estimation.

In Figure 3 the time series of profitability is depicted next to the forecast of our best-performing model, the gradient boosting algorithm. One can see that banks' profitability is going up again after the financial crisis with a peak in 2010 before the European debt crisis affected regional banks' balance sheets. The dashed red line shows the difference between the random walk's forecast error and the gradient boosting model. In line with Table 1, the random walk's error is in each year larger. Additionally, beginning with 2011, the relative forecasting performance of the gradient boosting model further improves.

[4] The t-statistics for the baseline model are -2.906, -5.853 and -10.132 corresponding to the order of Table 1.

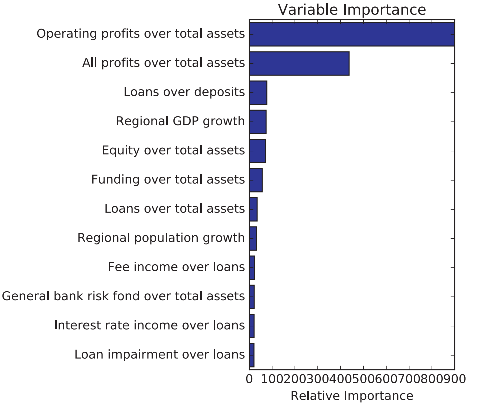
Figure 3: Mean Absolute Error per Year of Different Forecasting Techniques



Notes: The abbreviations RW and GRD denote the random walk and the gradient boosting model, respectively.

To inspect which variables are found to be most important in the estimations, we present Figure 4 for the gradient boosting model. It can be seen that current profitability measures are the most important variables for explaining future profitability. Variables which measure the bank's business riskiness as loans over deposits or the equity over assets ratio as well as regional variables as GDP and population growth are important to a smaller degree.

Figure 4: The Importance of the Included Variables for the Estimation



5. Conclusion

In the aftermath of the financial crisis regional banks had problems keeping up their profitability. Banks' profitability is an important indicator for the stability of the banking sector. We use a data set of bank-level balance sheet items and regional economic variables to forecast profitability. For the 2,000 savings and cooperative banks from eight European countries and the 2000-2015 time period, we found that machine learning algorithms are able to beat traditional estimators as ordinary least squares as well as autoregressive models in forecasting performance. Therefore, our paper is in line with the literature on machine learning models and their superior forecasting performance (Khandani et al., 2010; Butaru et al., 2016; Fitzpatrick & Mues, 2016).

The performance of the machine learning algorithms was particularly well during the European debt crisis which points out the importance of our forecasting exercise as during this time policy makers' interest in banks' profitability was enhanced as further potential rescue packages for banks could deteriorate fiscal stability. Policy makers and, especially, regulators should therefore use these algorithms instead of traditional estimators in combination with their even larger regulatory data sets in regard to size and frequency to forecast banks' profitability or other balance sheet items of interest.

References

Alessandri, P., & Nelson, B. D. (2015). Simple banking: Profitability and the yield curve. *Journal of Money, Credit and Banking*, *Vol. 47*, pp. 143–175.

Anolli, M., Beccalli, E., & Molyneux, P. (2014). Bank earnings forecasts, risk and the crisis. *Journal of International Financial Markets, Institutions and Money*, *Vol. 29*, pp. 309–335.

Behr, A., & Heid, F. (2011). The success of bank mergers revisited. An assessment based on a matching strategy. *Journal of Empirical Finance*, *Vol. 18*, pp. 117–135.

Belke, A., Haskamp, U., & Setzer, R. (2016). Regional bank efficiency and its effect on regional growth in normal and bad times. *Economic Modelling*, *Vol. 58*, pp. 413–426.

Bolt, W., de Haan, L., Hoeberichts, M., van Oordt, M. R., & Swank, J. (2012). Bank profitability during recessions. *Journal of Banking and Finance*, *Vol. 36*, pp. 2552–2564.

Butaru, F., Chen, Q., Clark, B., Das, S., Lo, A. W., & Siddique, A. (2016). Risk and risk management in the credit card industry. *Journal of Banking and Finance*, *Vol. 72*, pp. 218–239.

Demirgüc-Kunt, A., & Huizinga, H. (1999). Determinants of commercial bank interest margins and profitability: Some international evidence. *The World Bank Economic Review*, *Vol. 13*, pp. 379–408.

Diebold, F., & Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *Vol. 13*, pp. 253–265.

EBA (2016). *EBA Risk Assessment Report December 2016*. Available online.

Einav, L., & Levin, J. (2013). The Data Revolution and Economic Analysis. In *Innovation Policy and the Economy, Volume 14* NBER Chapters (pp. 1–24). National Bureau of Economic Research.

Elliot, G., Granger, C., & Timmermann, A. (2006). *Handbook of Economic Forecasting*. Amsterdam: Elsevier.

Englmaier, F., & Stowasser, T. (2017). Electoral cycles in savings bank lending. *Journal of the European Economic Association*, *Vol. 15*, pp. 296–354.

Fitzpatrick, T., & Mues, C. (2016). An empirical comparison of classification algorithms for mortgage default prediction: Evidence from a distressed mortgage market. *European Journal of Operational Research*, *Vol. 249*, pp. 427 – 439.

Flannery, M. J., Kwan, S. H., & Nimalendran, M. (2013). The 2007-2009 financial crisis and bank opaqueness. *Journal of Financial Intermediation*, *Vol. 22*, pp. 55 – 84.

Genay, H. (2014). What is the impact of a low interest rate environment on bank profitability? *Chicago Fed Letter*, *Vol. 324*.

Hakenes, H., Hasan, I., Molyneux, P., & Xie, R. (2015). Small banks and local economic development. *Review of Finance*, *Vol. 19*, pp. 653–683.

Hasan, I., Koetter, M., & Wedow, M. (2009). Regional growth and finance in Europe: Is there a quality effect of bank efficiency? *Journal of Banking and Finance*, *Vol. 33*, pp. 1446–1453.

Haskamp, U. (2016). Spillovers of banking regulation: The effect of the German bank levy on the lending rates of regional banks and their local competitors. *Ruhr Economic Papers*, *Vol. 664*.

Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking and Finance*, *Vol. 34*, pp. 2767–2787.

Koetter, M. (2008). An assessment of bank merger success in Germany. *German Economic Review*, *Vol. 9*, pp. 232–264.

Koetter, M., & Wedow, M. (2010). Finance and growth in a bank-based economy: Is it quantity or quality that matters? *Journal of International Money and Finance*, *Vol. 29*, pp. 1529–1545.

Schmieder, C., Marsch, K., & Forster-van Aerssen, K. (2010). Does banking consolidation worsen firms' access to credit? Evidence from the German economy. *Small Business Economics*, *Vol. 35*, pp. 449–465.

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, *Vol. 28*, pp. 3–28.