UKRAINIAN BANKS' BUSINESS MODELS CLUSTERING: APPLICATION OF KOHONEN NEURAL NETWORKS

Vladyslav Rashkovan
SD Capital

Dmytro Pokidin
National Bank of Ukraine

This paper clusters and identifies six distinct bank business models using Kohonen Self-Organising Maps. We show how these models transform over the crisis and conclude that some of them are more prone to default. We also analyze the risk profiles of the bank business models and differentiate between safest (valid) and riskiest ones. Specifically, six risk types (Profitability, Credit, Liquidity, Concentration, Related parties lending, and Money Laundering) are used to build risk maps of each business model. The method appears to be an efficient default prediction tool, since a back-testing exercise reveals that defaulted banks consistently find their place in a "risky" region of the map. Finally, we outline several potential fields of application of our model: development of an Early Warning System, Supervisory Review and Evaluation Process, mergers and acquisitions of banks.

JEL classification: G210, L100, C450

Keywords: neural networks, clustering, SOM, business model, banking

I. INTRODUCTION

As believed by many experts, the recent financial turmoil in Ukraine stemmed from more than a decade of reckless monetary and supervisory policy that allowed huge imbalances to be accumulated. Poor supervision gave rise to the unchecked growth of bad business practices at banks. Related party lending, large assets concentration, and money laundering are among the most pronounced risks of the Ukrainian banking system. Before the crisis, a number of banks had been growing fast, with the market having reached the peak of almost two hundred banks.

By collecting individuals' savings, most banks didn't provide funds to small- and medium-sized businesses. Instead, they mainly served business groups related to banks' owners, shoring up monopolization of the market at best. In the worst scenarios, the banks were used as intermediaries for illegal money laundering schemes.

According to many scholars, business model analysis must become a cornerstone in modern banking supervision.[1] Banking regulators also share this opinion, as the European Central Bank launched the Single Supervisory Review and Evaluation Process (SREP) in which business model analysis plays a key role. Indeed, business model analysis provides the regulator with valuable information on the structure of the financial sector. Knowing the dominant business models and their respective risks aids in the implementation of proper macroprudential policy. It also helps in ensuring proportionality in supervision, as stipulated by SREP.

This paper is fully devoted to identification and research of current Ukrainian banks' business models, how they changed over the crisis period, outlining risk areas, and finding out new possibilities for development. To the best of our knowledge, this is the first work of its kind in Ukraine. The ultimate goal of it is to develop a policy-oriented methodology that would aid in advancing the supervisory practices of the National Bank of Ukraine (NBU).

We did a clustering analysis of the Ukrainian banking industry with the aim of identifying business models. The clustering model we deployed in this paper was a Kohonen Self-Organizing Map (SOM). We identified six business models endemic to the Ukrainian banking system: Households-to-Corporates, Retail, Universal, Corporate, Investment/Wholesale, and Frozen/Undecided. Then, we demonstrated what kind of transformations Ukrainian banking underwent during the financial crisis.

1 See Ayadi R et al. (2015)

To supplement our findings, we built a risks map based on a set of risk indicators, one developed specifically for Ukrainian market. The map serves as a tool for assessment of each business model as well as for default prediction of an individual bank. We proved the efficiency of this tool by conducting backtesting, which showed that a majority of the defaulted banks lie in some particular (risky) region of the maps.

The vast majority of existing works on the clustering of banks' business models utilize k-means or hierarchical clustering methods.[2] Our paper proposes a SOM as a valid alternative to it. It is not only good in its primary function of dividing data into homogeneous groups, but it also has very nice features for data visualization as well as other functionality such as trajectories analysis, which we deployed in our work as well.

The paper is structured in the following way. In the second section, we propose an overview of some literature on the topic and compare it with our methodology. The third section presents the methodology, data, and software we deployed for the analysis in detail. In the fourth section, we present our most important findings. The fifth section sets out further work on the topic. Finally, the sixth section provides a summary of our study and concluding remarks.

II. LITERATURE OVERVIEW

In recent years, many works on identification and analysis of banking business models emerged in response to increased demand from regulators. In this section, we identify the main tendencies that appear in the literature and discuss their relative pros and cons.

Banks' business models

A business model is something that differentiates businesses within an industry from one another. The choice of a business model ultimately shapes all the essential characteristics of the firm: target clients, regions, products, marketing channels, suppliers, etc. All these features in one way or another find their quantitative representation in the data. Therefore, the problem of business model identification is purely a clustering one. All the works presented below use clustering algorithms to find out which banking business models prevail on the market. Nevertheless, the model, timing, variables, their granularity, and resulting number of clusters varied greatly.

The scholars generally try to keep their models parsimonious, i.e., using a modest number of variables. Ayadi et al. (2014/2015) and Ferstl, Seres (2014) used merely five variables, while overall a common range is from five to eight variables. Halaj, Ochowski (2009), however, stand out from the list using fifteen variables. For business model clustering that is aimed at solving policy related issues, the number of variables should indeed be limited. An increased number of variables commonly results either in an increased number of groups or less homogeneous ones. For macroprudential purposes, we want to see the general picture of the system and the main groups of banks comprising it. If the analysis is on a micro-level, e.g., for the purpose of mergers and acquisitions, we want to see as granular data as possible. In this case, the number of variables must be greater.

Scholars are divided in their approaches towards variable selection and construction. Ayadi R et al. (2014/2015), Roengpitya, Tarashev, Tsatsaronis (2014), and Tomkus (2014) use exclusively standardized balance sheet data so that banks' size does not matter. The advantage of this approach is that it is universal since financial statement data is always open sourced. The key assumption here is that all relevant information regarding a bank's business model shows up in its balance sheet ratios, which is not necessarily true. Other authors try to complement data with other characteristics. Halaj, Ochowski (2009) included some product-specific information such as the amount of housing loans and business-specific ones such as assets per employee. The European Central Bank (2016) used information on the proportion of domestic balance sheet exposure. Such information, of course, might be very helpful in achieving the goal of business model identification; however, it is not always openly available.

The European Central Bank (2016), among others, included a size variable in the form of Risk Weighted Assets. In such a way, the authors added another dimensionality to their analysis: they not only differentiate banks by business models, but by their size as well. However, in our study, we try to avoid inclusion of information that could in any way describe the banks' size. We believe that concentration only on the key business ratios could ensure clarity and consistency of results.

The methodology of Ferstl, Seres (2014) strikingly differs from previous ones. The authors made an amalgam of profitability, liquidity, and balance structure variables assuming they all reflect business models. For the reason discussed in the preceding paragraph, we believe that mixing the data that describes long-term business choices with volatile performance or risks indicators is not a good idea. Some business models might indeed correlate with risk level; others might occasionally outperform their peers in terms of profitability. However, this commonly has a temporary nature and depends on the financial cycle. In the long run, such indicators only contribute to noise in the data related to the business model.

2 See the literature review section for details.

Some of the authors do a post-clustering assessment of the resulting business models. Ayadi R et al. (2014/2015), and Roengpitya, Tarashev, Tsatsaronis (2014) analyzed the performance of business models in terms of their efficiency and riskiness by calculating some standard banking metrics. However, we think that more can be done about the issue. In this paper, we tried to extend the methodology for business model assessment. We made the methodology specifically targeted for Ukraine. However, it is also applicable to other post-soviet economies.

The abovementioned works reveal major strands in the literature on banking business model identification. All the authors agree that this is a clustering problem. To tackle the problem, they try to keep their clustering models parsimonious by keeping the number of variables moderately low. However, they are divided in the choice of the variables: some of them stick to financial statement data only, while others complement their analysis with more granular data. A few scholars went deeper in their assessments of the resulting models, while we believe that there is a lot of undiscovered space here. In our work, we tried to build a comprehensive methodology for both business model identification and their assessment. The methodology is tailored specifically for Ukraine, although it is also can be applied to many other emerging markets.

A SOM and its application in finance[3]

In the previous subsection, we revealed that scholars use clustering algorithms for business model identification. The algorithms they use are either hierarchical or k-means clustering. We propose a SOM as an alternative to them. A SOM is a clustering method based on neural computations. Kohonen (1982) first introduced it in the field of biology. Later on, it became popular in other areas, including economics.

We cannot claim that a SOM is any better than other clustering algorithms. Neither do other scholars researching the topic who often make controversial conclusions about the clustering efficiency of a SOM compared to other algorithms. Abbas (2008) did an experiment and showed that a SOM is better than its peers in almost all instances. Bação, Lobo, Painho (2005) found that a SOM is less prone to local minima than k-means. On the other hand, Mingoti, Lima (2006) showed that a SOM does not outperform hierarchical and k-means clustering, and often turns out inferior. However, we picked a SOM mainly due to its extensive functionality in data visualization. Additionally, it allows performance of a trajectories analysis (see the next paragraph), which we heavily deployed in our study.

There are not many works that apply a SOM to the business model identification problem. To the best of our knowledge, a paper by Vagizova, Luire, Ivasiv (2014) is the only in the field. The authors used a SOM to identify business models of interactions of the banking sector and the real economy of Russian banks. However, there are plenty of applications of SOM in broader economics and finance. Sarlin, Peltonen (2011) built a financial stability map of European banks in their paper aimed to predict financial crises. The authors of this work featured the attractive functionality of a SOM – trajectories analysis, showing how (by what trajectory) some countries moved across the map over time. Zarutska (2012) also used this feature in her analysis of the riskiness of Ukrainian banks.

Summarizing this subsection, one can assert that a SOM has its advantages over classical clustering methods. Although there is no strong evidence that a SOM is more efficient in the division of banks by homogeneous groups, it has an obvious data visualization advantage. Additionally, it allows for conducting a trajectories analysis, which we used in our study. Therefore, we propose it as a valid alternative to hierarchical and k-means algorithms commonly used in business models clustering.

III. METHODOLOGY AND DATA

Brief intro to Self-Organizing Maps

Kohonen SOMs is an algorithm from the Artificial Neural Networks (ANN) family. It is a two-layer neural network consisting of input and output layers. The following is a short theoretical summary of the method. It will be supported with examples specific to this paper, such that the reader can grasp the general idea of the method more easily.

Let $x=\{x_i : i=1,\dots,n\}$ be a set of size *n* of vectors of banks' variables on the input layer and $w=\{w_j : j=1,\dots,k\}$ be a set of size *k* of vectors of neurons' weights on the output layer, where $dim(x_i)=dim(w_j)$. In a Kohonen SOM, neurons are located on a two-dimensional grid.

In a SOM, algorithm weights *w* are typically initialized to have small random values. This, however, may result in the so-called dead-neurons problem – phenomena when some neurons do not ever take part in the learning process due to a high distance from each point from the input data (the essence of the problem will get clearer further). To avoid this problem, the weights are initialized along with two principal eigenvectors that correspond to the two highest eigenvalues of the input data. Such initialization ensures that all the data points are close enough to at least one output layer neuron.

3 Refer to Bullinaria (2016) for a comprehensive introductory guidance to SOM and neural computation. The text of Deboeck, Kohonen (1998) gives many additional examples of SOM applications in Finance.

After initialization, the vectors *x* on the input layer are matched with *w* to find the closest neuron by the formula $d(x_i, w_j)=(x_i-w_j)'(x_i-w_j)$, which is the squared Euclidian distance between the variables' vector of bank and the weights' vector of neuron *j*. The neuron with the minimum distance is declared the *winning* or the *Best Matching Unit (BMU)*.

Then, the input vectors start being supplied to the model iteratively. The SOM iterative process consists of two phases: *rough* and *fine-tuning*, which differ by learning parameters described below. In our case, the rough phase consisted of 10,000 iterations (*epochs*), and the fine-tuning of another 20,000 epochs. Each time, the neuron weights are updated by the formula $\Delta w_{i,j}=a(t)\, N_{i,j}(t)\,(x_i-w_j)$.

The term $a(t)$ is a time (epoch) dependent learning parameter, which determines by how much the weights would be updated. It starts with a moderately large value and then decays as the iteration process continues. In our application, the learning parameter decayed from 0.5 to 0.05 during the rough phase, and from 0.05 to zero in the fine-tuning phase.

The term $N(t)_{i,j}$ is a neighborhood parameter, which adjusts the weights' update according to the distance of the neuron to the BMU. It's defined as follows: $N(t)_{BMU,j} = \exp(-\frac{D^2_{BMU,j}}{2\sigma(t)^2})$, where $D^2_{BMU,j}$ is a map distance between neuron j and BMU. The term $\sigma(t)$ is a radius parameter. By analogy with the learning parameter, $\sigma(t)$ should decay over the epochs. In our case, it starts from 2.5 at the rough phase and decreases to 1, in the fine-tuning phases it is constant at 1.

To keep it simple, the formulas above ensure the following. Once the input vector of banks' variables is fed to the model, the weights of the output layer adjust in a way that the BMU's weights get closer to the input vector the most, while the neighbor neurons adjust by fewer values depending on their distance to the BMU. The farther the neuron from the BMU, the less its adjustment is. In such a way, over many iterations, our two-dimensional map takes on a topological structure, which corresponds to the original highly dimensional data.

Another parametric choice we faced was selection of the map's size. We opted for a 20x15 square, i.e., 300 neurons overall. The map size choice was stipulated by the data sample size.

There are a couple of clustering efficiency criteria we use to assess the quality of the resulting maps: quantization and topological errors. Quantization error is the average distance between each input data vector and its BMU. Topological error is the fraction portion of all input data vectors for which the first and second BMUs are not adjacent.

The software we deployed in this paper is MATLAB and open sourced SOM Toolbox.

Clustering methodology

Our methodology consists of two main blocks. The first is business model clustering and the second is the assessment of the resulting business models. In both cases, we use a SOM: in the first application, we use it to cluster the data; in the second, we use it to build a risk map for the assessment of the riskiness of the business models in whole and of individual banks.

Classifying business models

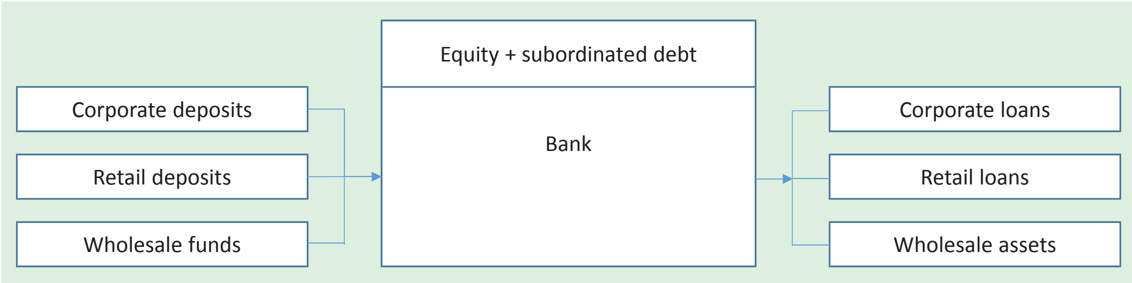
In broad strokes, a bank's business model can be described by answering four general questions:

- Who are a bank's target clients?
- Which products are offered to them?
- Which marketing channels does it deploy (chain of branches, alternative channels, etc.)?
- How does it generate profit (scale, low costs, high tariffs, etc.)?

Our goal for business model identification is to provide an unbiased quantitative view of the balance sheet structure of Ukrainian banks. We believe that the balances of a bank, coupled with some auxiliary indicators, can reveal the underlying business decisions that shape its business model.

The Figure 1 describes what the business model could be. How much equity does bank have, i.e., how leveraged it is? What sort of funds does it attract? What kind of revenue sources (i.e., assets) does it have? Are they classic loans only or some mixture of loans with wholesale assets? All these define the bank as a business.

Figure 1. Business model concept



The data we used was semi-annual, spanning a period of 3.5 years from January 2013 to July of 2016. Thus, a unit of measurement was a bank in a given period. Overall we had 169 banks as of 2014, of them only 93 were left as of mid-2016. This corresponds to 799 observations. The variables that we used to identify business models along with their descriptive statistics are displayed in Table 1.

Table 1. Business model variables descriptive statistics

Variable	*mean*	*sd*	*min*	*max*	*median*
Assets/Bracnhes (UAH)	602 951 772	1 344 209 368	2 212 142	6 499 324 617	102 137 583
Average loans maturity (years)	1.95	1.39	0.00	7.08	1.56
Average loans size (thousands UAH)	5 165.15	11 026.46	0.66	127 528.34	1 476.76
Equity and subordinated banks ratio	0.30	0.22	-0.20	1.00	0.23
Retail assets ratio	0.11	0.16	0.00	0.94	0.05
Retail deposits ratio	0.38	0.20	0.00	0.91	0.40
Loans ratio	0.74	0.21	0.00	1.00	0.79

The Assets/Branches variable shows how intensively the bank uses a chain of branches in its operations. Since we did not want this variable to implicitly represent banks' size, we standardized it by assets value. In the results, the variable shows assets' value per branch. A high value is supposed to indicate a relatively small number of branches.

The average loan maturity is calculated as the weighted average loan maturity in years applied to loan stocks as of a particular date. This indicator reflects the timing horizon in which a bank operates on its product side. The problem with this indicator is that it is calculated from stocks' values. Therefore, it reflects a decision made in the past (probably a distant one). What we instead would like to see is the flow information, i.e., the maturity of newly issued loans for a period. Unfortunately, data limitations did not allow us to construct such a variable.

The same problem applies to the average loan size variable, which was constructed as the overall loans portfolio divided by the number of loans. To tackle the possible problem of outlying values, we first dropped the top decile of each bank's loans. The difference between the mean and median of this variable indicates the presence of outliers from the top side. That means that some banks credit big businesses by issuing large loans.

The equity and subordinated debt ratio shows how leveraged is a bank. The distribution of the variable is centered around 0.23, while the mean is 0.3. As previously stated, this indicates the presence of some very deleveraged banks, which is very uncommon to the banking business.

The retail loans ratio is the proportion of retail loans to revenue generating assets.[4] It reveals the main target clients of the bank. A high value of this variable evidences that a bank serves individuals mainly. If the value is low, a bank orients more on the corporate or wholesale market. Descriptive statistics show that Ukrainian banking has more corporate or wholesale exposures, while there are banks that serve mostly individuals.

4 Revenue generating assets include loans, interbank exposures, and securities.

The retail deposits ratio is the proportion of retail funding to the sum of overall liabilities minus subordinated debt. It shows to what extent a bank relies on individuals to fund its operations. We can see from descriptive statistics that, despite assets exposure to individuals is on average very low, Ukrainian banks rely much more on them to fund their activities.

Finally, the loans ratio is a the share of loans (excluding interbank) to assets. It shows to what extent a bank is engaged in non-classical banking activities. If the value is low, then a bank has a high interbank or trading exposure. From descriptive statistics, we can see that Ukrainian banking is mostly traditional, having a median value of the variable equal to 0.8.

Note that no qualitative indicator is included in the list above, since we strived to give as objective a result as possible, without the use of subjective qualitative indicators. We also did not explicitly differentiate banks by size since all the ratios are standardized by assets value where applicable.

Also, to provide for equal weighting of all the variables in the SOM algorithm, they were normalized to have a mean of zero and variance of one. We did not want to see outliers in our training sample. Therefore, we replaced outlying values in the sample with the nearest value in a non-outlying range. Generally, we qualified a value as an outlier if it was more than 4 standard deviations away from the median. Appendix 1 contains boxplot graphs of the normalized variables.

After application of a SOM algorithm to the data, we additionally needed to join the output layer neurons into groups, such that we get the resulting clusters (i.e., business models). For this purpose, we applied a k-means algorithm to the neurons' weights.[5] The number of clusters (k's) was determined by an elbow method.[6]

Given the optimal number of clusters, the optimal division is achieved by a bootstrap procedure with 100 iterations. At each iteration, the criterion was constructed using the formula $Cr = \frac{BCSS}{WCSS}$, with BCSS (between clusters sum of squares) and WCSS (within clusters sum of squares). $BCSS=\sum_i (\bar{w}_i-\bar{w})(\bar{w}_i-\bar{w}), WCSS=\sum_i\sum_j(\bar{w}_i-\bar{w}_j)(\bar{w}_i-\bar{w}_j)$, where $\bar{w}$ is the overall sample mean, $\bar{w}_i$ is the cluster i mean, and $\bar{w}_j$ is the cluster j mean. Ultimately, the clustering with maximum *Cr* is selected.

Risk mapping

For the purpose of risk assessment, we propose concentrating on the six types of risks:

1) Profitability risk
2) Credit risk
3) Liquidity risk
4) Concentration risk
5) Related party lending risk
6) Money laundering risk

The first three types of risks come directly from the Basel framework. Profitability risk here partially quantifies the market risk from Basel, as will be explained below. Unfortunately, we could not include operational risk here as we could not find proper quantification of it. We admit that this type of risk might be material and contribute to the severity of the banking crisis.

The remaining three types of risks deal with those problematically specific to Ukraine and many other emerging market countries, namely high concentration, related party lending, and money laundering. The variables that we used to quantify the risks are presented in Table 2.

The time span is the same as business models clustering, so is the unit of measurement and outliers' treatment. However, the frequency this time is higher. We choose quarterly data since risk indicators are usually less stable over time than the ones for business model identification. As a result, the sample size for risk clustering is 1,475.

Our approach to normalization was slightly different too. We did separate normalization for each point of time. The reason for this was the fact that some variables we used for risk clustering experienced structural shifts in means.[7] Therefore, doing so ensured some sort of mean stationarity of the data.

5 A SOM algorithm is closely related to a k-means one. In fact, application of a k-means algorithm to the output weights of a SOM adds another layer to the neural network in the form of k-mean clusters. Therefore, the overall model may be considered as a three-layer network.
6 The method was first proposed by Thorndiket (1953).
7 For example, a real NPL that had been hidden by banks for a long time was revealed with a recent Assets Quality Review.

Table 2. Risk variables descriptive statistics

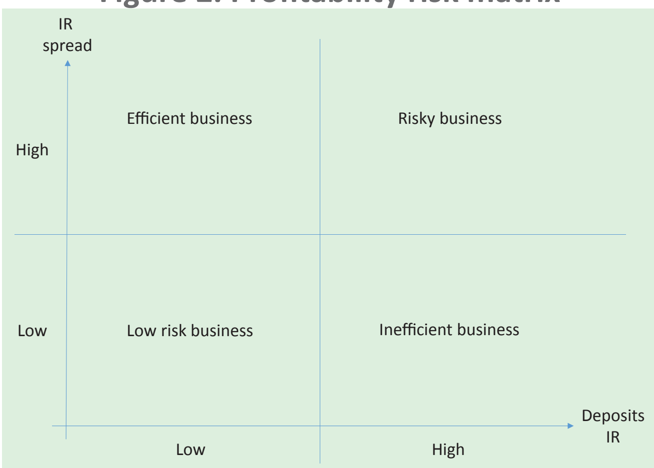
Variables	*mean*	*sd*	*min*	*median*	*max*
Deposits IR	15.36	5.49	0.00	16.34	33.64
IR spread	6.76	7.04	-8.34	5.97	28.33
NIM	0.03	0.03	-0.08	0.02	0.19
NPL ratio	0.13	0.21	0.00	0.04	1.00
NPL coverage	1.10	0.77	0.02	1.00	3.03
Liquid assets ratio	0.09	0.12	0.00	0.05	0.90
Assets concentration	0.49	0.25	0.00	0.49	1.00
Liabilities concentration	0.17	0.16	0.00	0.12	0.84
Unique borrowers ratio	0.36	0.23	0.00	0.34	1.00
Turnover	2.22	2.16	0.01	1.51	10.80

When analyzing profitability of banks, we address their ability to raise funds cheaply and allocate profitably. These imply the efficacy of a bank's target clients, market, regional, and other strategic choices. The variables that aid to quantify this are deposits' interest rate (**Deposits IR**) and interest rate spread (**IR spread**). These indicators deal with the interest rate and interest rate spread risk according to the Basel definition of market risk.[8]

The banking business is deemed efficient when it raises funds at a low interest rate and lends at a higher one (given a reasonable risk profile) and the other way around (see the Figure 2). Nevertheless, if a bank raises expensive funds and lends them with a high spread, it may suggest that the bank may undertake risky projects. From descriptive statistics, you can see that the average deposits interest rate and interest spread are very high, reflecting the high risk profile of the Ukrainian market.

Another useful indicator of profitability is Net Interest Margin (**NIM**). It is the relation of net interest, commission, and trade income to revenue generating assets. The average value of 0.03 is commensurate with the similar figure for developed markets. Therefore, higher risks are not on average compensated for by higher returns on assets.

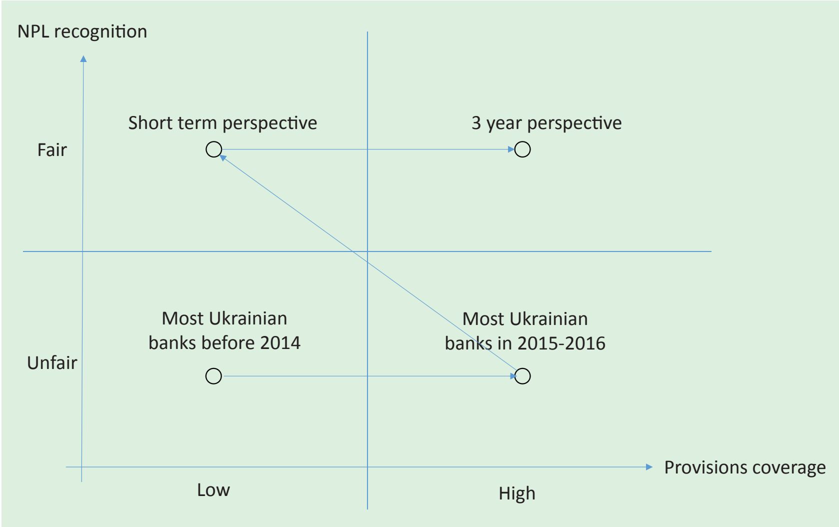
Figure 2. Profitability risk matrix



When dealing with credit risk, an obvious choice was to consider the non-performing loans level (**NPL ratio**) and to check if it is covered with provisions (**NPL coverage**). Dealing with the NPL level is a bit tricky since many banks hide the real level of NPLs by rolling over, restructuring, etc. Hence, we face an issue of fair recognition of NPLs. Before 2014, most banks had been hiding the real level of NPLs and kept too little provisions, as shown in the bottom-left box of Figure 3. An Assets Quality Review and stress test exercise conducted by the NBU from 2015-2016 forced banks to raise their provisioning levels, thereby moving them to the bottom-right box. In the short-term period, banks are expected to show the real NPL level, thus moving themselves to the top left box. Over a three-year period, banks are expected to fully cover these NPL with provisions, thereby appearing in the top-right box. Given the above information, we regarded having abnormally low NPL levels risky in our analysis. To the contrary, having high NPL and little provisions coverage should not always be taken myopically, because, in some occasions, it might signal the willingness of a bank to represent the real picture of its assets and to provision them shortly. The subjectivity of the issue allows us to make inferences only with some degree of confidence and subject them to professional judgment.

8 See BIS (2016).

Figure 3. NPL and fair recognition issue

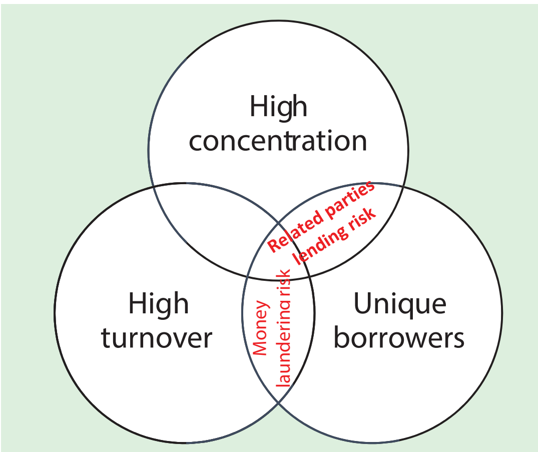


In analyzing Liquidity Risk, we constructed a **liquid assets ratio** indicator. It is essentially the portion of the bank's most liquid assets, which include cash, correspondent accounts with the NBU, deposits with NBU, and government securities that are refinanced by NBU. One can be sure that, in the case of a massive deposits outflow, a bank would certainly be able to survive the increased liquidity pressure of at least the value of this indicator. Unfortunately, there are not many ways to measure liquidity risk yet since such measures as Liquidity Coverage Ratio (LCR) are yet to be developed in the NBU.[9]

Last, but just as important as the previous risks, is the risk of a bank's malpractices. It includes related party lending, endemic to it concentration risk, and money laundering risk. Concentration risk is measured with **assets concentration** and **liabilities concentration** variables. These variables are constructed as the ratio of assets/liabilities that account for >2% of the total assets each.[10] Descriptive statistics show us that assets concentrations risk is more pronounced in Ukrainian banking.

Related lending detracts the banking system from the fulfillment of its primary function – provision of funds to the real sector. Instead, it causes market inequalities, inefficient resource allocation, monopolization, and many resultant issues. More on the destructive impact of related lending is laid out in La Porta, Lopez-de-Silanes, Zamarripa (2001). Identification of such practices is a tough and tricky task. Our approach to this issue is presented in Figure 4.

Figure 4. Banks' malpractice identification approach



9 In fact, there are three existing liquidity measures in accordance with NBU economic normatives – N4, N5, N6. However, they were proven inefficient in the current application. More details on this can be found in Figures A-C of the Appendix 1, where the signaling ability of some indicators is analyzed.
10 Liabilities excluded subordinated debt.

Here we incorporate the following logic. A variable of the **unique borrowers ratio** is the portion of large borrowers (>2mln UAH) within a particular bank that have not come across in other banks for a period of the four last years. We assume that if there is a high unique borrowers ratio and a high assets concentration, then the probability that a given bank is engaged in related party lending is greater, ceteris paribus. Here we bear in mind that banks practicing related party lending are most likely to serve some particular business group that is not interested in borrowing from someone else. In addition, many business groups have the practice of creating fictitious companies (so-called Special Purpose Vehicles) that manage financial flows of the business group and will most likely to be a client of only the bank also belonging to this group. Such companies usually do not create any value, have a few of employees, and do not have an office. Therefore, such companies naturally do not even have a chance to get a loan from a bank other than that owned by the business group.

In turn, we assume that a high **Turnover** on some balance sheet accounts[11] coupled with a high ratio of unique borrowers might indicate money laundering practices.

IV. RESULTS

Business models maps

Using the variables and optimal clustering solution from Section 3.1, we conducted a clustering analysis of the Ukrainian banking system. The purpose of this was to identify what types of business models are common for Ukrainian banks and how they transformed over the crisis. We identified six business models: Households-to-Corporates, Retail, Universal, Corporate, Investment/Wholesale, and Frozen/Undecided.

Figure 5 contains a SOM of the business models. It shows the location of each business model on it. From the figure, we can observe how 300 neurons are organized into a two-dimensional grid. Each neuron can contain one bank, several banks, or be empty. The coloring of the map represents different clusters. Neurons to be joined into one group were determined by a k-means clustering algorithm, as explained in Section 3.1.[12]

Figure 5. SOM of business models

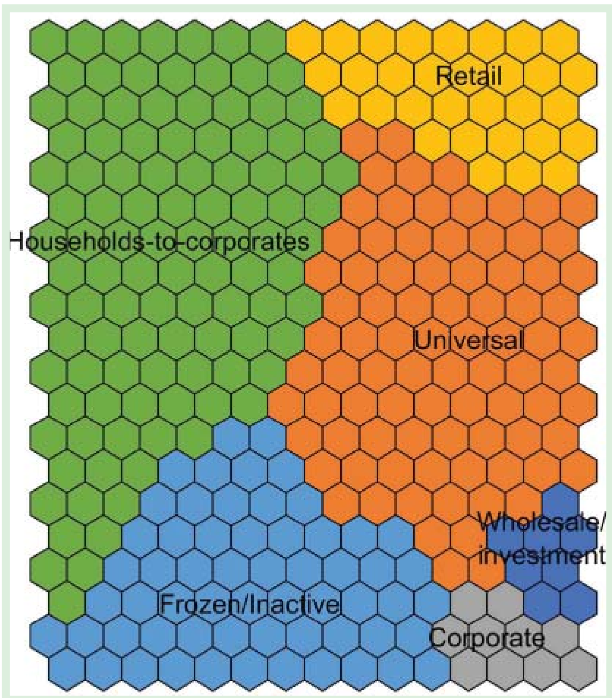
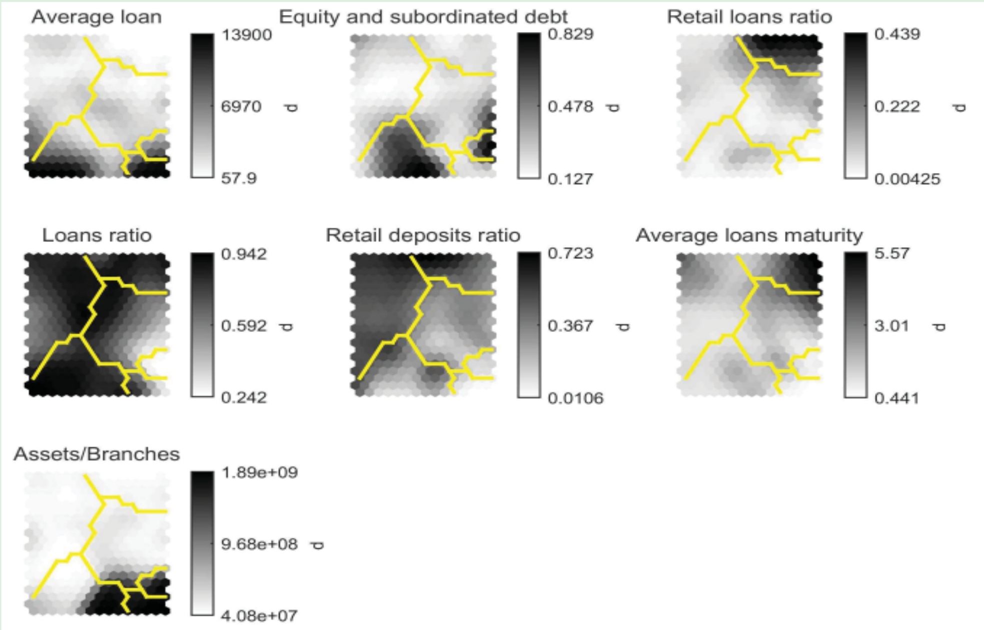


Figure 6 visualizes the variables used in a SOM algorithm. Each little map corresponds to some of the seven variables utilized for the business models clustering. These maps are colored according to the variables' values. The darker the region, the higher a variable's value it has, and hence, the less a variable's value the banks have in that area.

11 The list of accounts used is the result of an analysis conducted by the authors. This list includes accounts, increased turnover on which could be observed in banks, liquidated by Financial Monitoring Laws. Unfortunately, the authors cannot disclose the list of accounts.
12 We applied elbow criterion to the map's weights and found the optimal number of clusters - 14. However, this number was unreasonably high and the actual difference in the weights was not very material. Therefore, we expertly joined some clusters and came up with the six. You can find the map divided by these original 14 clusters in Figure N of the Appendix 1.

Figure 6. Components maps



On these maps, you can visually observe characteristics of the determined clusters.[13] For example, you can see that the share of retail loans is much higher in the retail cluster region; also, it has a very high retail deposits ratio, the lowest average loan size, the longest maturity of loans, and the largest number of branches.

The Households-to-Corporate (HTC) has a small portion of retail loans (high fraction of corporate loans) and a large portion of retail deposits. In other words, the banks from this cluster stream households' funds to corporates. It is not a bad business model per se. However, in Ukraine, it is highly over-represented and accounted for about half of the banking system before the crisis. In addition, due to its characteristics, this cluster bears the risk of related party lending, although this point will be disclosed in the risk clustering section.

A mix of loans and wholesale assets characterizes the universal cluster. Loans are issued to both retail and corporate clusters. The retail deposits ratio is high, but not much high as in the HTC and retail clusters.

The Frozen/Undecided segment is quite diverse in assets and liabilities structure. The feature that is common to this segment is very high equity and subordinated debt share, reaching up to 90%. It indicates that the banks from this group do not fulfill one of the main function of a banking institution (financial intermediation) since they do not attract deposits. This might happen for several reasons: the bank is young and not yet scaled up in its operations; the bank is inactive; the bank is undecided as to its business model; or the bank is engaged in activities not typical to traditional banking.

The Corporate segment does not have retail loans and deposits – it serves only corporates. In addition, it has the largest average loan and shortest loans' maturity: there is no surprise in it since enterprises naturally take larger loans than individuals do. Moreover, corporates in Ukraine take loans mainly to finance operational activities. Therefore, loans are mostly short-term. Finally, since the cluster does not serve individuals, it does not need branches, which is reflected on the Assets/Branches variable's map.

The Wholesale/Investment banks are extremely uncommon in Ukraine. There had been just five such banks before the crisis. The cluster is similar to Corporate; however, it has the lowest fraction of loans among all clusters. Therefore, the majority of its operations are wholesale.

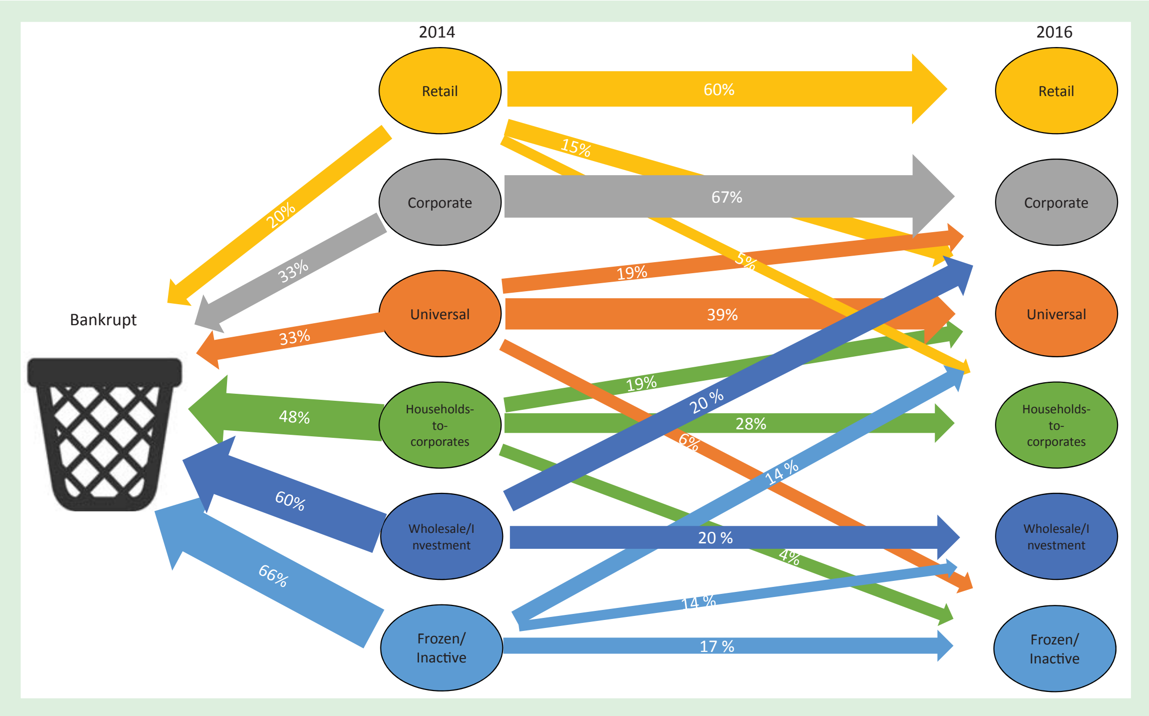
The quantization error for a business models SOM is 0.8, and the topological error is 1.75%, which is low enough for the map to be considered accurate. An optimal clustering solution has a Cr value of 0.37.

Banks migration from clusters over time is illustrated in Figure 7. It seems that the HTC and Frozen/Undecided segments were more prone to defaults over the crisis. Investment/Wholesale also has very high default rate, however, taking into account the very low number of its constituents, the absolute number of defaulted banks in this cluster is not material.[14] Universal, Corporate, and, especially, Retail segments have relatively low default rates. Therefore, they might be considered relatively safe from this perspective.

13 Appendix 2 contains descriptive statistics of the identified clusters.
14 Refer to the tabular representation of Figure 7 in Appendix 2.

Ayadi et al. (2015) also conducted a migration analysis among clusters in their regular *Banking business models monitor. Europe.* In Europe, the clusters behave quite stably. If we consider only the banks that survived the crisis in Ukraine, we would observe a similar picture in Ukraine.

Figure 7. Visualization of banks' migrations among models

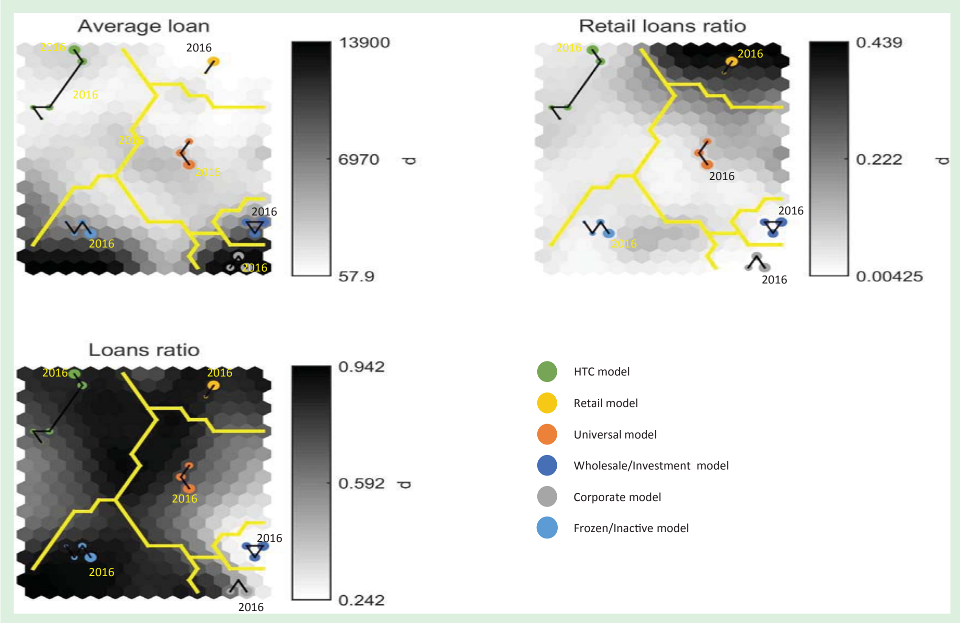


Another important feature of business models is the stability of their constituents, which is illustrated in Figure 8. The black dots there indicate the locations of banks that appeared at least once in the respective cluster over the studied period. It is seen that all groups, except Universal, are quite stable. Universal, though, has its constituents very scattered over the map. Therefore, this cluster can be deemed as a transition cluster. For example, if a bank decides to change its business model from a Retail to a Corporate one, it would certainly start changing its assets structure by reducing its retail loans ratio. But, it would not happen instantly. Hence, during the transition period, the bank would appear in the Universal segment, which is characterized by diversity of assets.

Figure 8. Banking constituents' stability

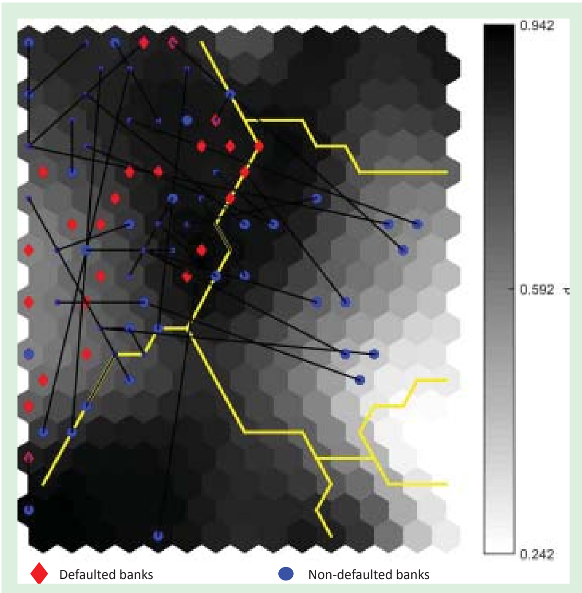
A SOM provides very useful functionality called trajectory analysis. Trajectory analysis tracks the movement of some units (for example, a bank or a cluster centroid) on the map over time. It allows visual observation of the changes undergone by that unit over time. Figure 9 shows the movement of cluster centroids starting from the beginning of 2014 on three variable's maps.[15] We can see that for nearly all clusters the average loan has increased, which is natural due to drastic depreciation and inflation that occurred in Ukraine during the crisis. However, the retail loans ratio increased exclusively in the Retail cluster, while in others, this indicator either decreased or remained unchanged. In a way, the Retail cluster reinforced its authenticity, which we believe is a good sign.

Figure 9. Trajectories of cluster centers on the business models map



Considering the Loans ratio, we can observe that the HTC segment somewhat got rid of its wholesale assets. There are a couple of reasons for that, namely the disappearance of the Ukrainian interbank market and high default rates of HTC banks that were engaged in wholesale operations. Figure 10 illustrates the trajectories of HTC banks on the Loans ratio map. Red diamonds are the banks that went bankrupt. We can see that a majority of banks located in the light region of the cluster either went bankrupt or moved out of the region.

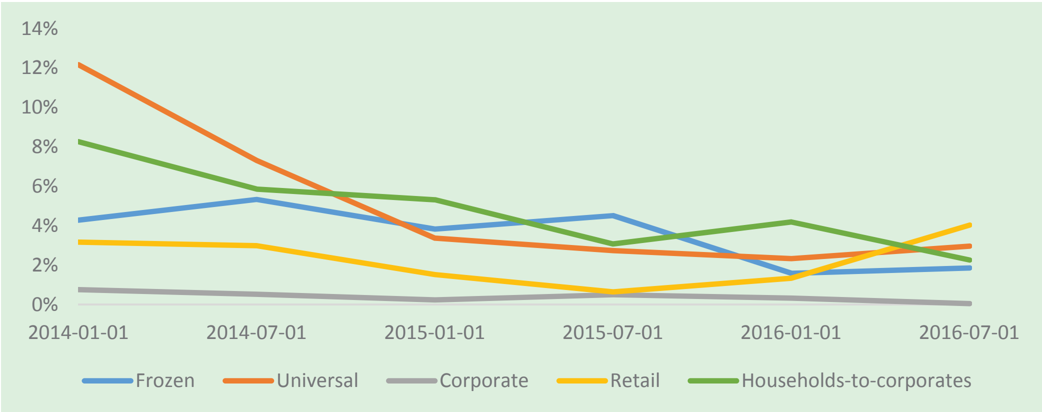
Figure 10. Trajectories of HTC banks on the Loans ratio map



15 Similar figure for all variables maps is in Appendix 1 (Figure O).

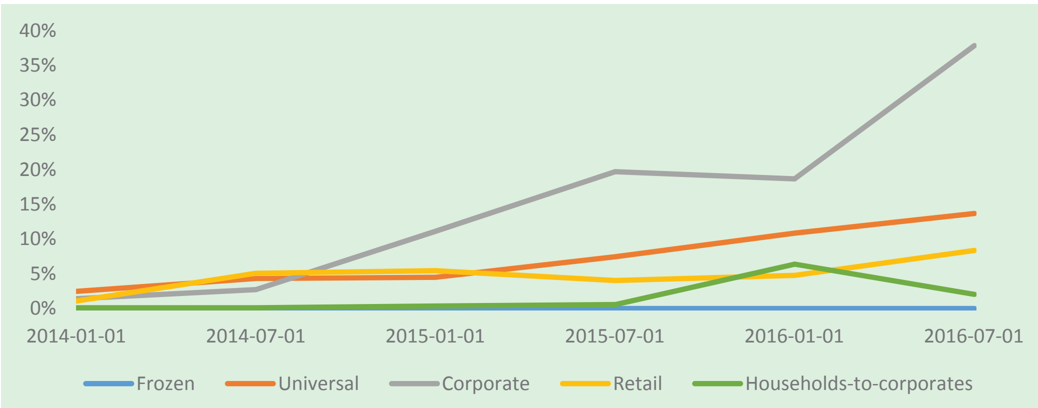
Before the crisis, banks had relied heavily on the interbank market to fund their short-term liquidity gaps; other banks had been ready to stream their free resources to them. However, with the onset of liquidity pressure in 2014, a decrease in overall lending, and a loss of trust within the banking system, this market naturally vanished. Therefore, the portion of the local interbank market in assets approached zero. Figure 11 shows that the HTC and Universal segments were active participants in the interbank lending market. Now the fraction of local interbank lending for each cluster is diminutive. This led the HTC cluster centroid to the top of the business model map in Figure 9, where the ratio of loans is relatively high.

Figure 11. The portion of local interbank lending by clusters



On the other hand, Figure 9 also tells us that the Universal and Corporate clusters reduced its loans share. The reason for this is illustrated in Figure 12. We see that the portion of government securities in assets soared dramatically with the onset of the crisis for the segments mentioned above. This phenomenon can be easily explained by the risk aversion of the clusters. Economic turbulence made the real sector very risky. As a result, these clusters seemed to prefer to invest in relatively safe government securities instead.

Figure 12. The portion of government securities by clusters



Business models clustering identified six distinct business models of Ukrainian banking system. Banks that went bankrupt were not equally spread among business models with HTC and Frozen/Undecided clusters accounting for more than 70% of all defaulted banks. We also showed that the clusters are relatively stable, except Universal one, which carries a bit of each cluster's characteristics, therefore can be regarded as a transfer point for banks switching between business models. Retail cluster reinforced its authenticity by accumulating retail loans fraction in its assets. The engagement in wholesale operations changed oppositely for HTC cluster and Corporate and Universal ones. The former reduced its wholesale assets fraction mainly due to the disappearance of the local interbank market, where it used to be the dominant player. Universal and Corporate clusters invested heavily in the government securities thus having accumulated wholesale assets fraction compared to the pre-crisis period. This was presumably dictated by risk aversion of the clusters.

Risks maps

Using variables from the section 3.2.2 we constructed a risks map, which you can find in Figure F of the Appendix 1. However, for the sake of visualization advantage, we transformed it somewhat by merging some of the map's weights (variables) using the logic of the section 3.2.2. In a way, we came up with six-dimensional weight map: each for a particular risk type. To be precise, we applied the following transformations[16]:

- Concentration risk=(Assets Concentration + Liabilities Concentration)/2;
- Related parties lending risk=(Assets Concentration + Unique borrowers Concentration)/2;
- Laundering risk=(Assets Concentration + Turnover)/2;
- Liquidity risk=(Deposits IR-Liquid Assets Fraction)/2;
- Profitability risk=(–IR spread – NIM)/2;
- Credit Risk=(|NPL level|–NPL coverage)/2;

As a result, we can visually illustrate six risk types with the maps in Figure 13. The darker the region on the maps, the more risk of particular type bear the bank located in that region. It can be easily seen that the top half of the mast overall is riskier than the bottom one: it is highly concentrated and has signs of related parties lending; it also has liquidity problems; the left flank of it bears the risk of money laundering; the top right corner has very high credit risk.

Figure 13. Risks map

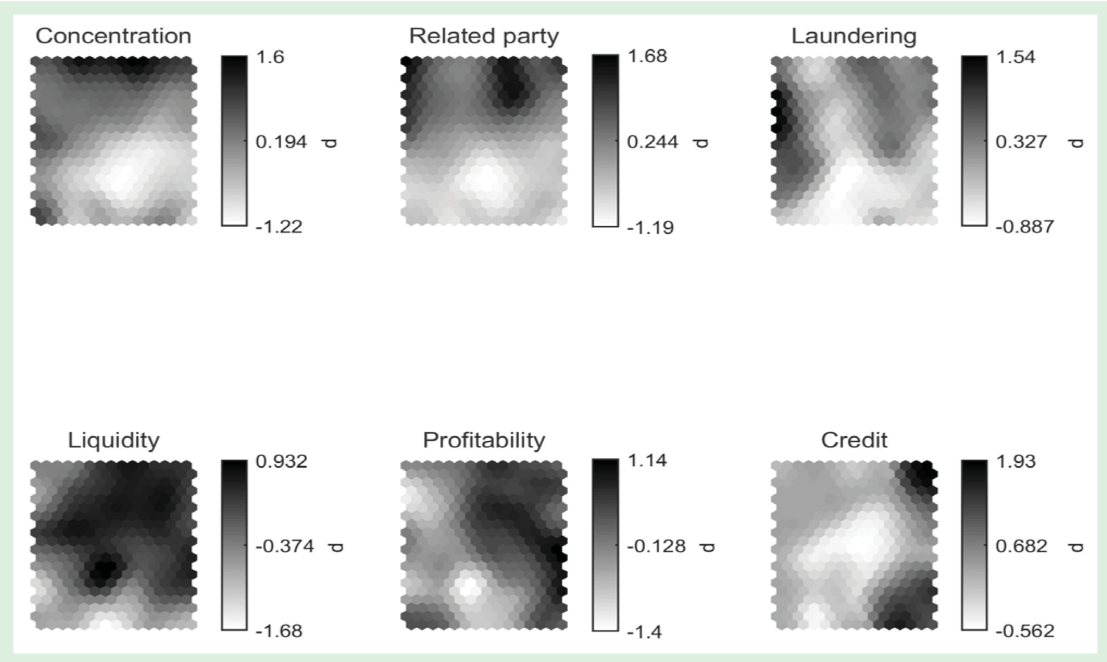
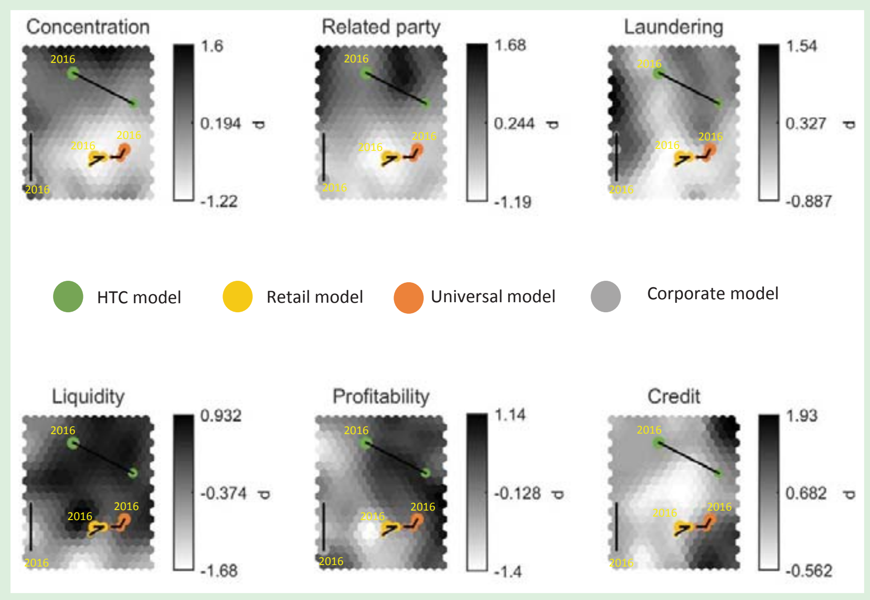


Figure 14 contains trajectories of business models' means on the risks map. It shows how business models looked at the beginning of 2014 and now in terms of riskiness. We disregarded Frozen/Undecided and Investment/Wholesale clusters due to their negligible size in the banking system and concentrated our attention on the HTC, Retail, Universal, and Corporate segments. There is no surprise that the HTC model is located at the top part of the map given its default rates. The model is highly concentrated, not very profitable, and has signs of related party lending. These risks, along with liquidity ones, are the major issues for this model. The Retail, Corporate, and Universal models are in the bottom part of the map. The Retail model seems to have the safest risk profile. The Universal model bears a bit of profitability and credit risk, while the Corporate model is somewhat concentrated.

Unfortunately, the risk profile of the HTC business model has not changed much over the crisis. It is seen that the model moved to the region with more concentration and credit risk. The Retail model remained in the safest area. The Universal cluster on average also has not changed its risk profile much, however its liquidity position slightly deteriorated. Regarding the Corporate model, we can see that it moved out of the region with high money laundering risk. It is natural because some of the banks from this segment were liquidated under the Financial Monitoring Law. Thus, we can reasonably assume that the remaining banks do not practice any illegal activity. The banks also enhanced their liquidity positions due to investments in liquid government securities. On the other hand, profitability deteriorated slightly.

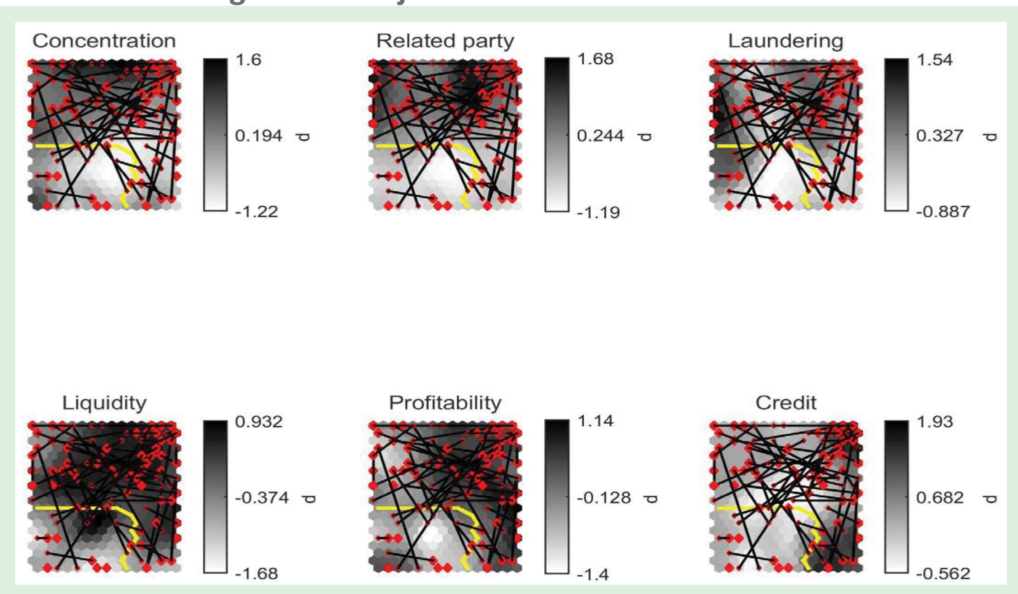
16 While reading the formulas, keep in mind that all the variable, and hence maps' weights, were normalized around zero. This makes the interpretation of the NPL level taken to the modulus in the sixth formula clearer: not only a high NPL level was considered risky by us, but an unusually low (below average, i.e., zero) as well.

Figure 14. Business models' movement on the risks map



Now, we are in a position to perform backtesting on our risks map with the use of actual data on banks' defaults. Figure 15 illustrates trajectories of the banks defaulted over the crisis from the beginning of 2014 until the last quarter of their existence. The circumscribed region in the bottom left region is considered the safest one, this is supported by the fact that only 8% of the defaulted banks were located there in the last quarter before their bankruptcy. There are also many instances when a bank had been located in the safe region and then moved out of it right before its bankruptcy. Based on the map, we can conclude that the region that is most densely populated with defaulted banks is the top right corner. The area holds each of the six types of risks to some extent.

Figure 15. Trajectories of defaulted banks[17]

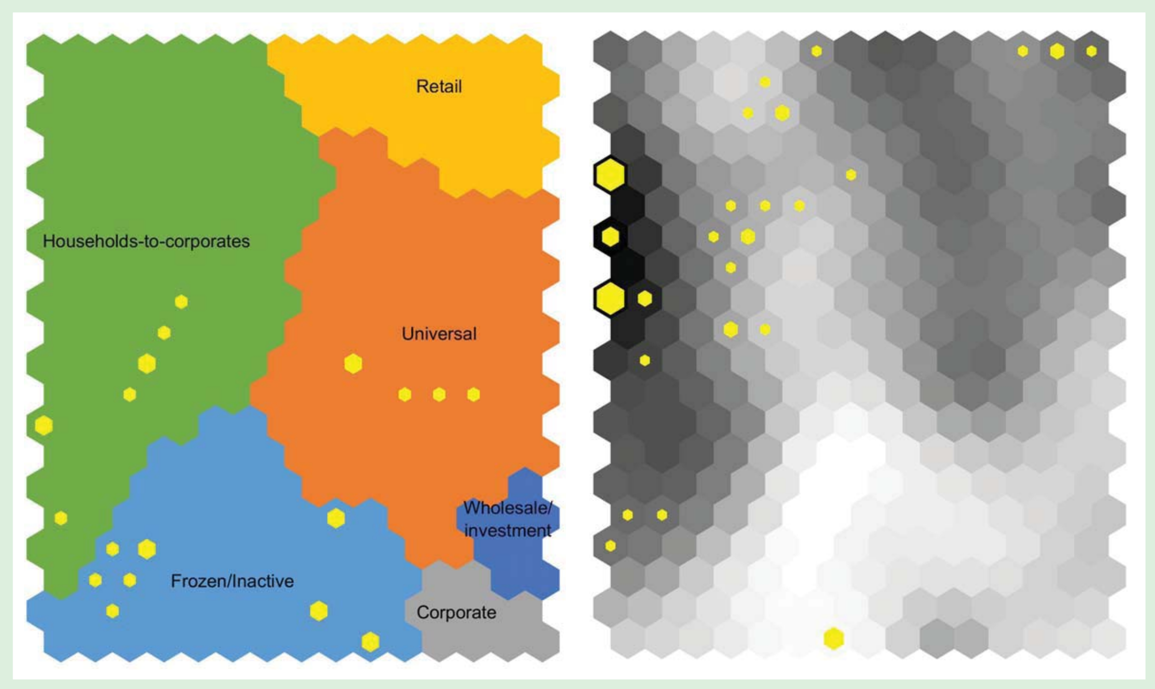


We can also test our hypothesis regarding identification of money laundering banks. Figure 16 shows the locations of money laundering banks on the business models map a) and on the laundering risk map b). On the business models map, the majority of these kinds of banks are expectantly located in the Frozen/Undecided business model or nearby it. On the laundering map, they lie mainly in the darkest region. This also confirms our hypothesis that high accounts turnover, coupled with a high unique borrowers concentration, may indicate illegal banking practices.

17 Figures H-M of the Appendix 1 contain the trajectories of individual bank by business models.

Figure 16. Money laundering banks location

a)BM map **b) Laundering map**



In this section, we proved that the risk indicators we constructed turned out to be quite informative in terms of default prediction. The risks map built based on them illustrated that 92% of the defaulted banks were located in a particular map's region that we consider risky. On the other hand, there is a relatively safe region, which contains only 8% of the defaulted banks. In addition, we confirmed our initial hypothesis that if the bank has high accounts turnover and a large unique borrowers concentration, then it is likely to be engaged in money laundering schemes. Such banks were mainly located in the Frozen/ Undecided cluster.

V. FURTHER RESEARCH

The previous section exhibited the great potential for SOM clustering in an analysis of the banking sector. However, some issues can be explored in more detail during further research. In particular, despite the risk maps showing a good signaling ability, this part of the analysis cannot yet be considered as fully comprehensive because of the complexity of the topic. For example, the analysis of such an important risk as liquidity risk and development of its measures requires much time and effort. This work is yet to be done within the NBU and well beyond the scope of this paper. In the meantime, we have just one measure of liquidity risk (Liquid Assets Ratio), which is proven to be informative retrospectively.

Risk mapping is not an attempt to create an Early Warning System since the time horizon of the risk mapping is much longer than an EWS must have. It does not preclude one from using the current methodology for the creation of such a system. Given the undeniable visualization advantages of the technique, it may appear as a very lucrative option. In this context, a trajectories analysis would be a very useful tool. If, for example, a regulator observes that it gradually approaches a map's "risky" region, this should become a clear warning signal.

Another fertile field for SOM usage might be banks' mergers and acquisitions. A bank seeking a partner for a merger or acquisition can outline key indicators that describe a desired profile of the target. Then, a SOM could be built based on these indicators. An analysis of the map can help identify the region that contains the most suitable targets.

In addition, the subject of our work might fit well in the SREP, which the NBU plans to develop. As prescribed by the European Banking Authority (2014), banks' categorization, business models, and strategic risks analysis are essential parts of the SREP. All the topics were disclosed in our methodology.

VI. CONCLUSIONS

In this paper, we developed a methodology and conducted a clustering analysis based on Kohonen neural networks to identify banking business models that prevail in Ukraine. We outlined six distinct business models: HTC, Retail, Universal, Corporate, Investment/Wholesale, and Frozen/Undecided.

Then, we showed how these models transformed as a result of the banking crisis. We showed that more than half of HTC and Frozen/Undecided models' constituents went bankrupt. This indicated that these models by default were riskier. We also revealed that some of the models had opposite changes in wholesale assets portfolios: while the HTC segment reduced its wholesale ratio due to the disappearance of the local interbank market, the Corporate and Universal segments accumulated the ratio due to increased investments in government securities. The latter happened presumably due to the risk aversion of the clusters' constituents. In addition, we showed that the Retail cluster is considered relatively safe due to its transparent and market-oriented business model: during the crisis it not only had the lowest default rate but additionally accumulated its retail loans ratio, thereby reinforcing its authenticity.

To complement our analysis, we constructed a risk map based on a set of risk indicators of six types: Profitability, Credit, Liquidity, Concentration, Related party lending, and Money laundering. It confirmed our previous findings regarding the riskiness of the HTC model and safeness of the Retail one. Then we conducted backtesting, which proved the efficiency of the proposed risk indicators: a majority of defaulted banks were located in some map's "risky" region before their bankruptcy. Hence, the presented SOM tool can be considered efficient in default prediction and other supervisory purposes.

Finally, we outlined a field for further research. In particular, in our plans is an improvement in the risk assessment methodology as new quantitative indicators of risks, such as LCR, come to life. Additionally, we gave some examples of where our methodology and method could also be applied. Particularly, in our opinion, such fields as EWS, SREP, or M&A are potentially good spheres to apply a SOM clustering approach.

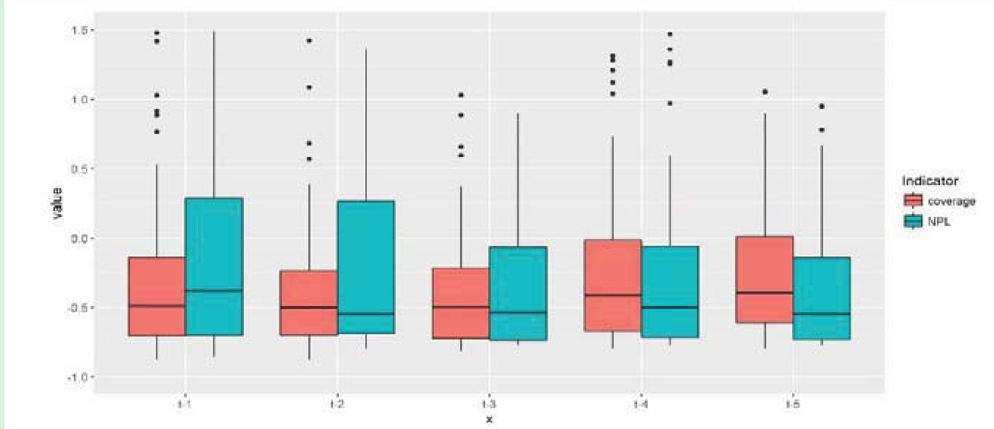
References

- Abbas O.A. (2008). Comparisons Between Data Clustering Algorithms, The International Arab Journal of Information Technology Vol.5.
- Ayadi R., Arbak E., de Groen D. (2011). Business models in European Banking: A pre – and post- crisis screening, Center for European Policy Studies, Brussels.
- Ayadi R., Arbak E., de Groen D. (2012). Regulation of European Banks and Business models: Towards a New Paradigm?, Center for European Policy Studies, Brussels.
- Ayadi R., de Groen D., Lapointe M., Michelet A., Rey H., Sassi I., Tita C. (2014). Banking Business Models Monitor 2014, Center for European Policy Studies, Brussels.
- Ayadi R., de Groen D., Rey H., Sassi I., Mathlouthi W., Aurby O. (2015). Banking Business Models Monitor 2015, Center for European Policy Studies, Brussels.
- Bação F., Lobo V., Painho M. (2005). Self-organizing Maps as Substitutes for K-Means Clustering, ICCS'05 Proceedings of the 5th international conference on Computational Science Volume Part III.
- BIS (2016). Minimum Capital Requirements for Market Risk.
- Bullinaria J. (2016). Self-Organizing Maps: Fundamentals, Introduction to Neural Computations. Available at http://www.cs.bham.ac.uk/~jxb/inc.html.
- Deboeck G., Kohonen T. (1998). Visual Explorations in Finance with Self-Organizing Maps.
- European Central Bank (2016). Financial Stability Review May, Recent trend in Euro area banks' business models.
- European Banking Authority (2014). Guidelines on common procedures for the supervisory review and evaluation process.
- Ferstl R., D. Seres (2014). Clustering Austrian banks' business models and peer groups in the European Banking Sector, Financial Stability Report, 24, pp. 79-95.

- Halaj G., Ochowski D. (2009). Strategic groups and banks' performance, Financial Theory and Practice, 33 (2), pp. 153-186.
- Kohonen T. (1982). Self-organized formation of topologically correct feature maps, Biological Cybernetics.
- La Porta R., Lopez-de-Silanes F., Zamarripa G. (2001). Related Lending, World Bank.
- Mingoti S.A., Lima J.O. (2006). Comparing SOM neural network with Fuzzy c-means, K-means and traditional hierarchical clustering algorithms, European Journal of Operational Research 174
- Roengpitya R., Tarashev N., Tsatsaronis K. (2014). Bank business models, BIS.
- Sarlin P., Peltonen T. (2011). Mapping the State of Financial Stability, European Central Bank Working Papers Series, No. 1382
- Thorndiket R. (1953). Who Belongs in the Family?, Psychometrika Vol. 18 No. 4.
- Tomkus M. (2014). Identifying Business Models of Banks: Analysis of Biggest Banks from Europe and United States of America, Aarhus University, Denmark.
- Vagizova V., Luire K., Ivasiv I. (2014). Clustering of Russian Banks: Business Models of Interaction of the Banking Sector and the Real Economy, Problems and Perspectives in Management, Vol. 12, Issue 1.
- Zarutska O. (2012). Obgruntuvannya pidhodu mashtabnogo rozpodilu bankiv Ukrayni na osnovi structurno-funktsyonalnih grup, Visnyk of the National Bank of Ukraine, (10), pp. 20-24.

Apendix 1. Figures

Figure A. Signalling ability of NPL level and provision coverage*



* Horizontal axis – quartes to default. Values are standartized such that non-defaulted banks have a value of zero. Therefore, the locations of bars are deviations of defaulted banks' indicators from non-defaulted ones.

Figure B. Signaling ability of the NBU's major economic normatives

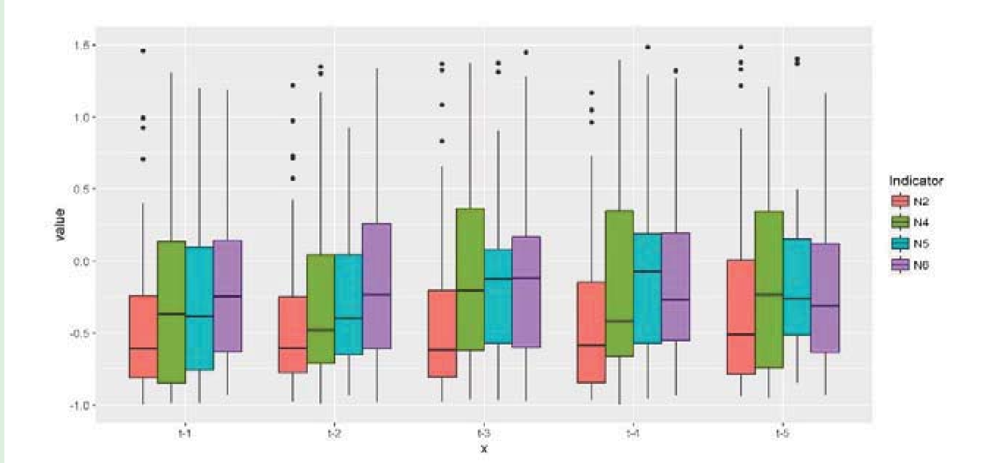


Figure C. Signaling ability of interest rates and spreads

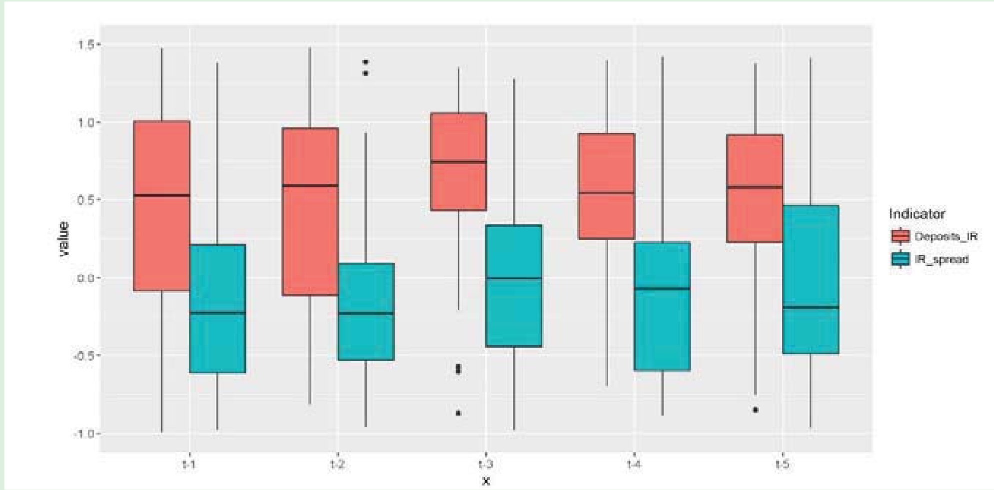


Figure D. Boxplots of normalized business models' variables

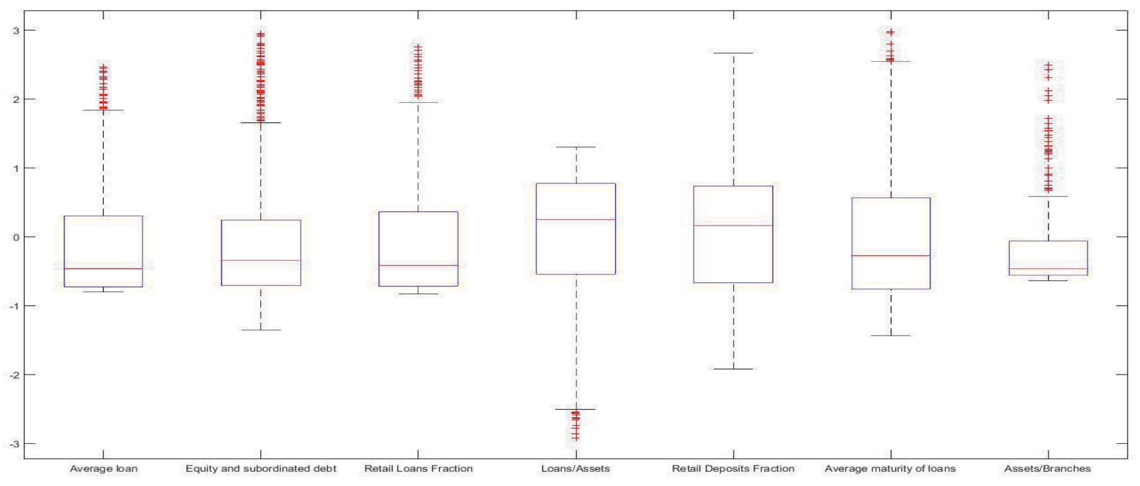


Figure E. Boxplots of normalized risk variables

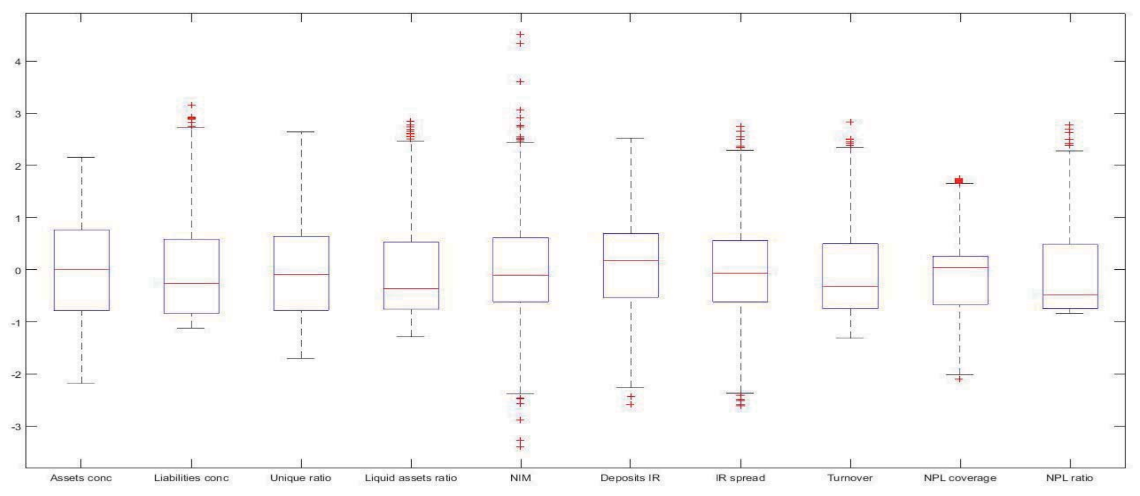


Figure F. Original risks map

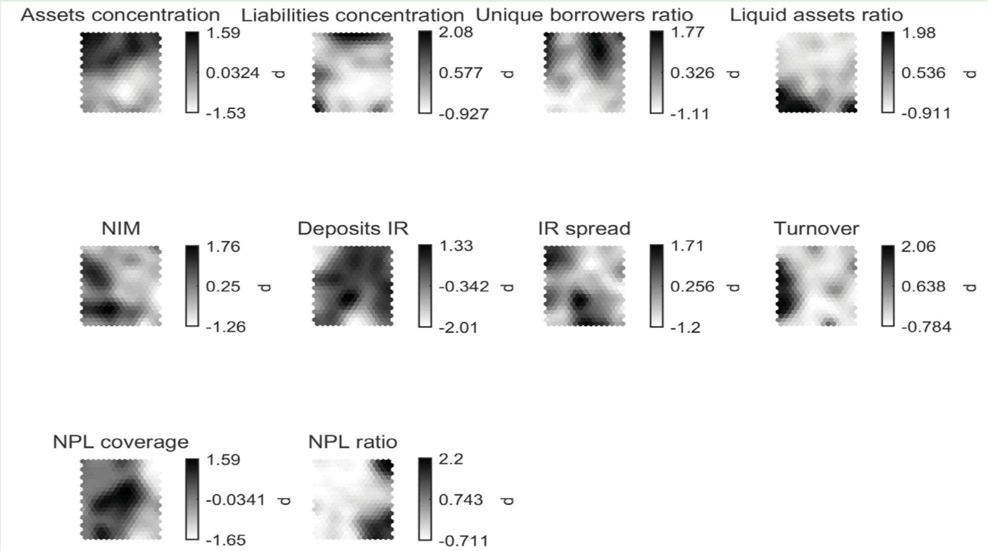


Figure G. Ownership structure by clusters as of mid-2016

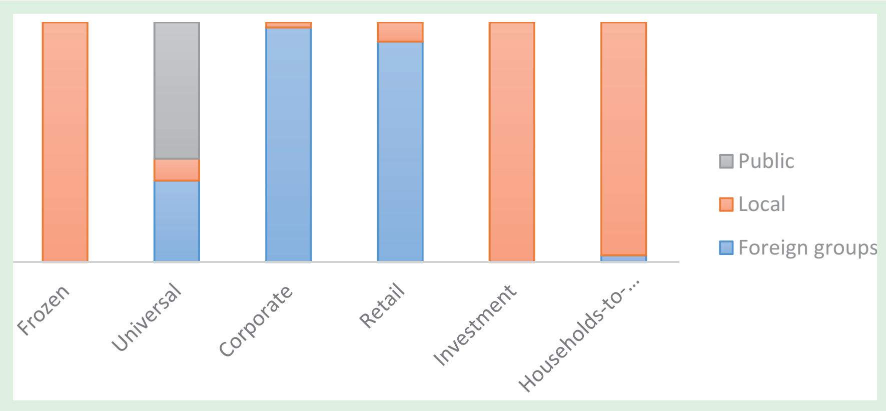


Figure H. HTC banks' trajectories on risks map

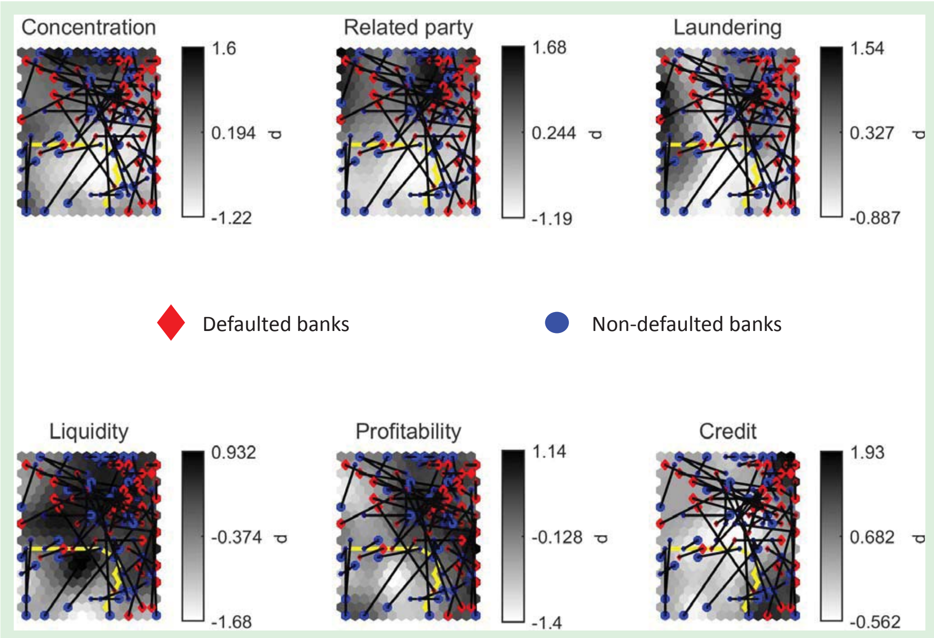


Figure I. Frozen/Undecided banks' trajectories on risks map

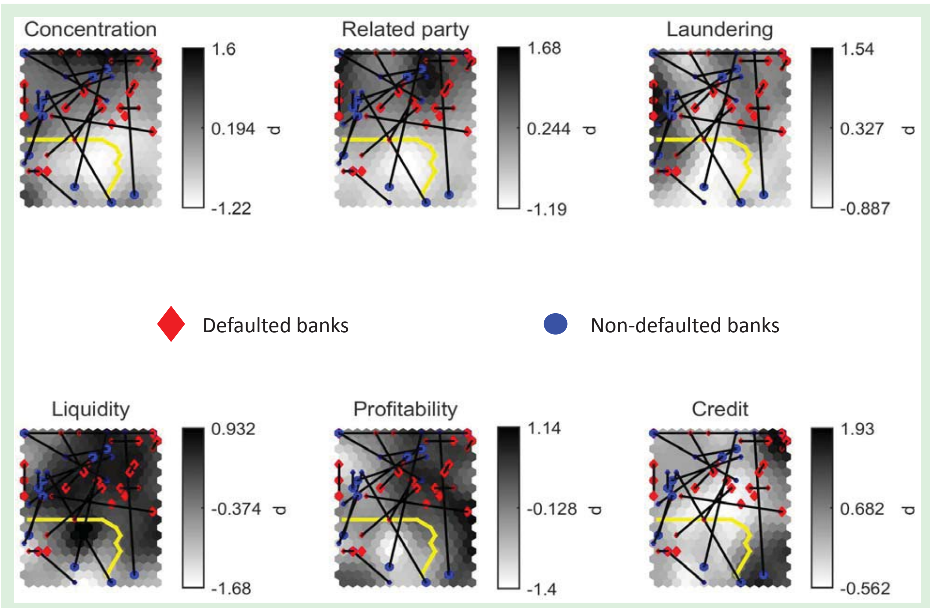


Figure J. Universal banks' trajectories on risks map

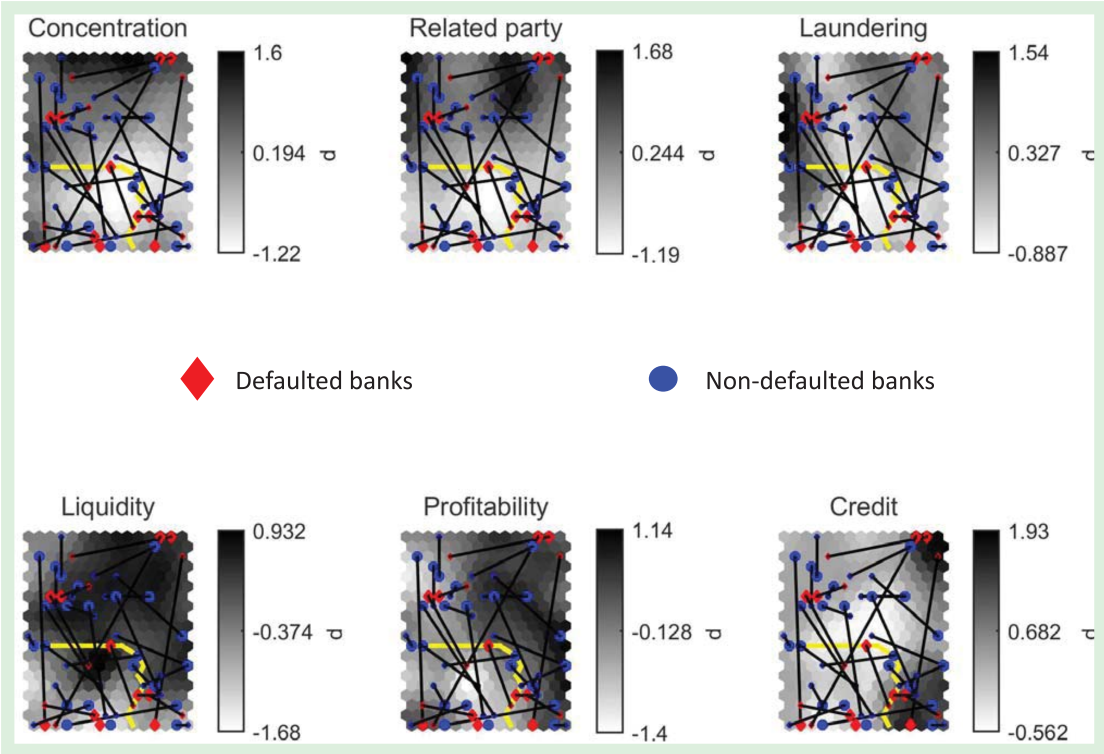


Figure K. Corporate banks' trajectories on risks map

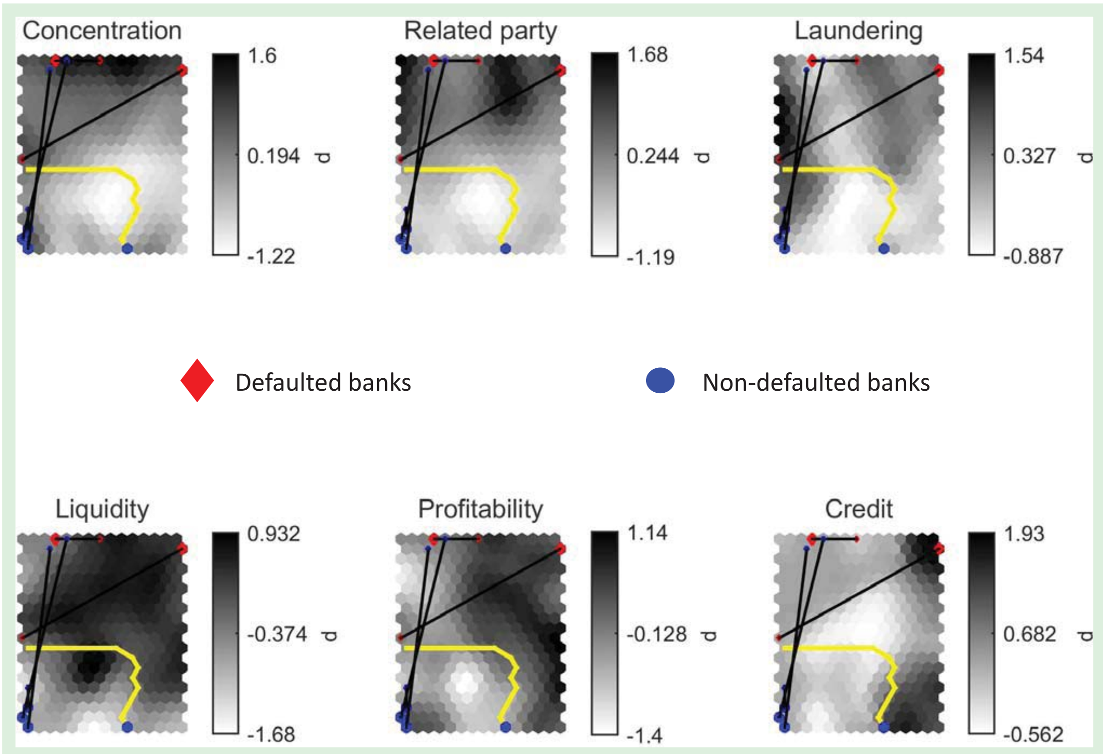


Figure L. Investments/Wholesale banks' trajectories on risks map

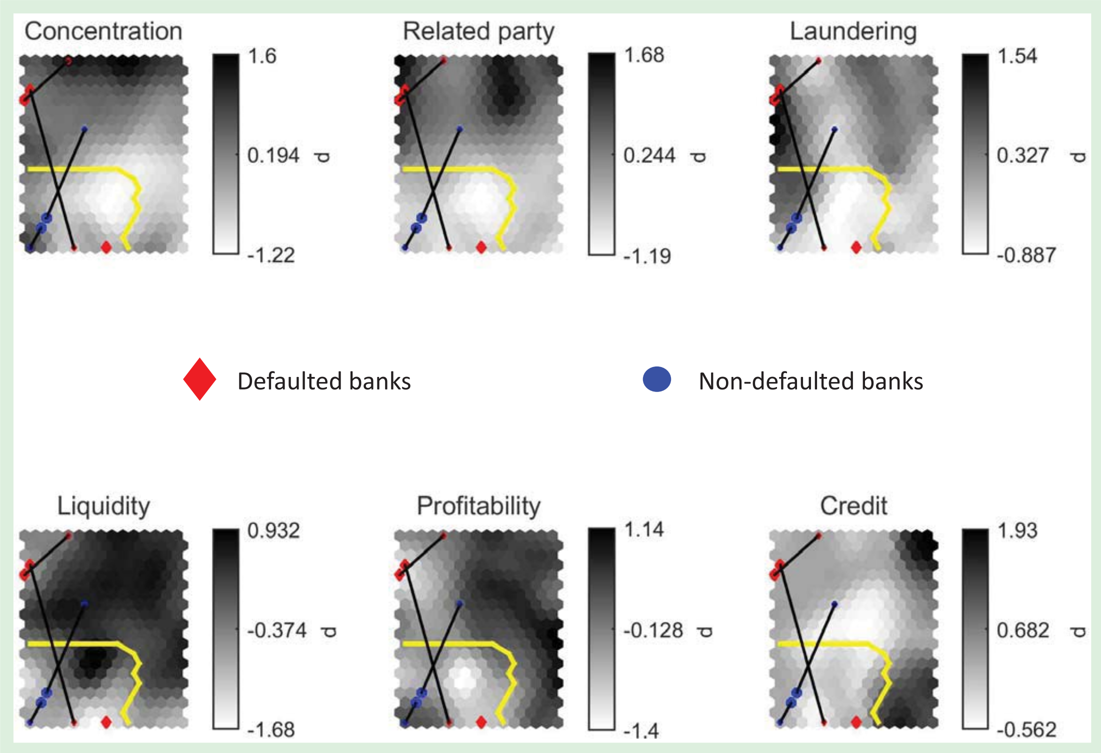


Figure M. Retail banks' trajectories on risks map

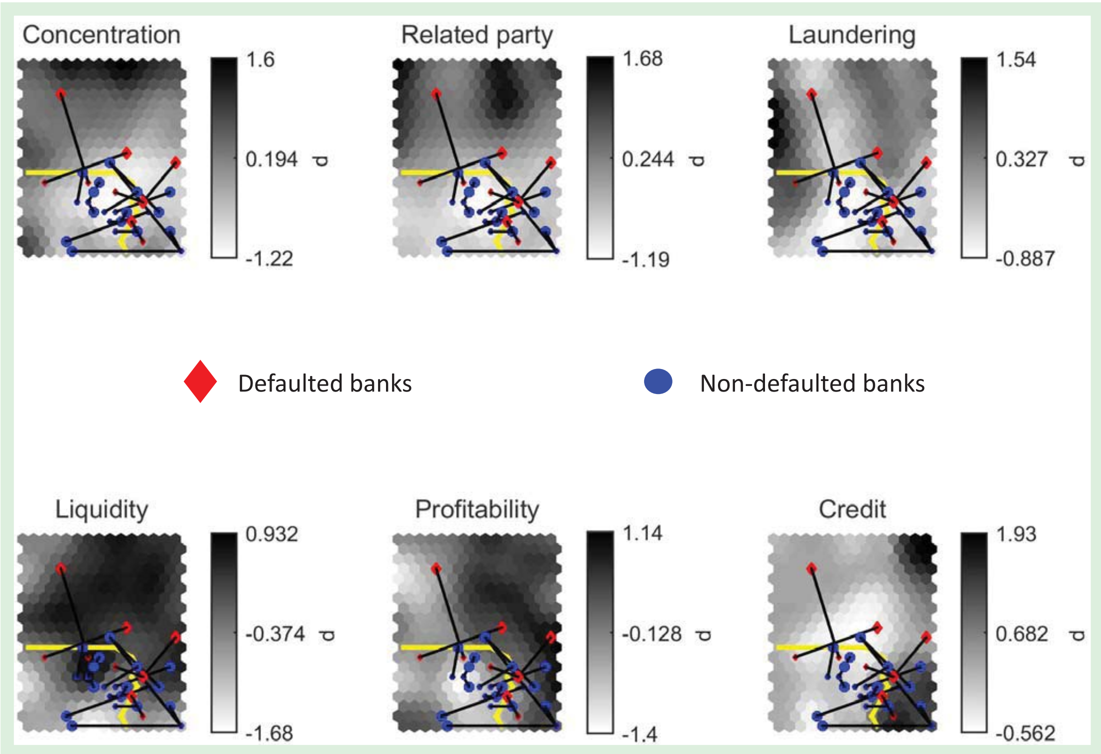
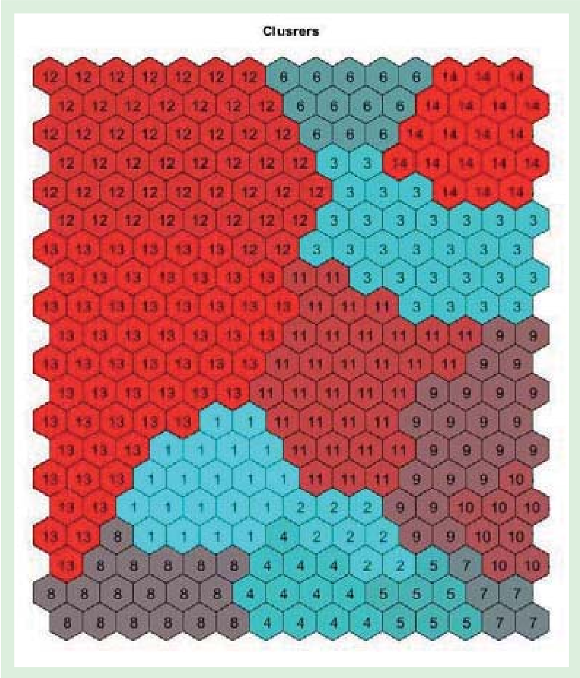
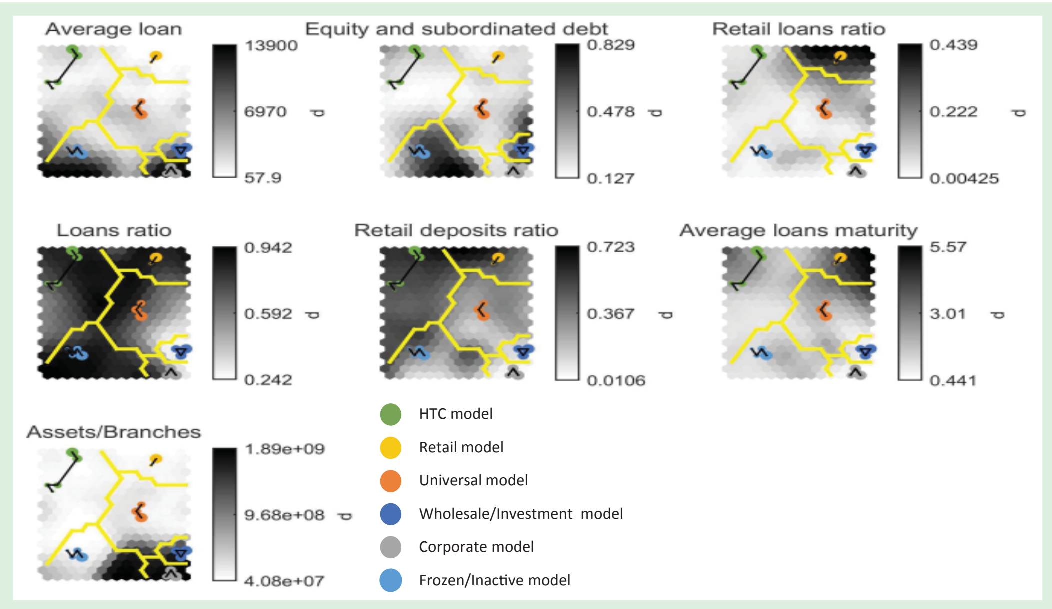


Figure N. Original clusters of the business models map*



*As you can see, we expertly joined some of the clusters: clusters 6 and 14 are the Retail model; cluster 3, 9, and 11are the Universal model; clusters 12 and 13 are the HTC model; clusters 1, 2, 4, and 8 are the Frozen/Undecided model; clusters 5 and 7 are the Corporate model; and cluster 10 is the Investment/Wholesale model.

Figure O. Business models' centroids movements on the business models map



Appendix 2. Tables

Table A. Descriptive statistics by business models over time

Clusters	*Variables*	*1.1.2014*					*7.1.2016*				
		mean	sd	max	median	min	mean	sd	max	median	min
Corporate	Retail deposits ratio	0.02	0.03	0.07	0.00	0.00	0.01	0.01	0.02	0.00	0.00
	Retail assets ratio	0.01	0.01	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00
	Loans ratio	0.63	0.30	0.90	0.72	0.16	0.45	0.29	0.82	0.29	0.16
	Equity and subordinated banks ratio	0.16	0.07	0.30	0.13	0.12	0.22	0.07	0.32	0.20	0.15
	Average loans size (thousands UAH)	13 397.7	1 539.0	14 556.0	14 149.2	11 038.8	14 556.0	0.0	14 556.0	14 556.0	14 556.0
	Average loans maturity (years)	0.59	0.32	1.06	0.61	0.07	0.42	0.12	0.55	0.45	0.25
	Assets/Bracnhes (UAH)	1 849 826 287	110 921 749	1 895 109 902	1 895 109 902	1 623 408 216	1 895 109 902	0	1 895 109 902	1 895 109 902	1 895 109 902
Frozen/Undecided	Retail deposits ratio	0.41	0.19	0.87	0.45	0.00	0.34	0.24	0.84	0.32	0.00
	Retail assets ratio	0.07	0.09	0.39	0.03	0.00	0.05	0.10	0.40	0.02	0.00
	Loans ratio	0.86	0.12	1.00	0.90	0.57	0.88	0.14	1.00	0.94	0.54
	Equity and subordinated banks ratio	0.41	0.25	0.87	0.35	0.10	0.62	0.23	0.92	0.66	0.20
	Average loans size (thousands UAH)	6 004.90	5 476.50	14 556.01	5 219.54	48.80	7 692.84	5 518.38	14 556.01	6 907.83	539.60
	Average loans maturity (years)	1.25	0.73	3.02	1.04	0.35	1.33	0.67	2.60	1.21	0.48
	Assets/Bracnhes (UAH)	877 840 137	817 760 092	1 895 109 902	543 254 739	2 272 634	584 917 205	742 758 024	1 895 109 902	202 027 212	9 032 576
HTC	Retail deposits ratio	0.50	0.09	0.76	0.48	0.25	0.52	0.09	0.71	0.55	0.30
	Retail assets ratio	0.07	0.06	0.23	0.06	0.00	0.05	0.06	0.25	0.04	0.00
	Loans ratio	0.78	0.13	1.00	0.80	0.43	0.79	0.10	0.95	0.78	0.56
	Equity and subordinated banks ratio	0.18	0.07	0.41	0.16	0.08	0.26	0.14	0.71	0.26	0.08
	Average loans size (thousands UAH)	2 256.63	2 326.73	9 606.47	1 416.43	24.06	2 435.27	2 095.32	6 832.79	1 536.91	27.61
	Average loans maturity (years)	1.52	0.71	3.38	1.42	0.53	1.70	0.92	4.36	1.44	0.67
	Assets/Bracnhes (UAH)	143 364 968	197 503 788	932 301 122	70 491 584	13 862 860	85 951 467	75 960 799	330 881 069	60 984 865	10 286 533
Investment/Wholesale	Retail deposits ratio	0.15	0.18	0.43	0.16	0.00	0.05	0.03	0.08	0.05	0.03
	Retail assets ratio	0.01	0.02	0.05	0.01	0.00	0.00	0.00	0.00	0.00	0.00
	Loans ratio	0.21	0.09	0.37	0.16	0.16	0.33	0.10	0.40	0.33	0.25
	Equity and subordinated banks ratio	0.87	0.07	0.92	0.89	0.76	0.79	0.19	0.92	0.79	0.66
	Average loans size (thousands UAH)	10 595.22	5 749.26	14 556.01	14 556.01	1 956.17	1 074.45	755.46	1 608.64	1 074.45	540.26
	Average loans maturity (years)	1.34	0.67	1.97	1.42	0.29	0.15	0.18	0.27	0.15	0.02
	Assets/Bracnhes (UAH)	1 895 109 902	0	1 895 109 902	1 895 109 902	1 895 109 902	1 895 109 902	0	1 895 109 902	1 895 109 902	1 895 109 902
Retail	Retail deposits ratio	0.52	0.13	0.78	0.55	0.23	0.54	0.20	0.85	0.49	0.18
	Retail assets ratio	0.37	0.08	0.45	0.37	0.23	0.40	0.07	0.45	0.45	0.30
	Loans ratio	0.82	0.10	0.96	0.85	0.61	0.78	0.13	0.98	0.81	0.58
	Equity and subordinated banks ratio	0.20	0.06	0.28	0.20	0.10	0.21	0.15	0.59	0.16	0.02
	Average loans size (thousands UAH)	228.69	435.78	1 824.09	80.16	1.26	659.59	1 751.60	6 412.57	46.89	6.02
	Average loans maturity (years)	3.83	1.37	5.95	4.30	0.86	3.77	1.71	5.95	3.41	1.01
	Assets/Bracnhes (UAH)	95 693 262	142 406 621	656 114 208	55 712 310	10 905 967	95 265 551	71 837 636	260 219 890	81 476 473	19 114 667
Universal	Retail deposits ratio	0.27	0.13	0.59	0.27	0.02	0.25	0.12	0.54	0.25	0.01
	Retail assets ratio	0.11	0.09	0.30	0.10	0.00	0.08	0.08	0.26	0.04	0.00
	Loans ratio	0.58	0.21	0.93	0.63	0.16	0.63	0.21	0.97	0.64	0.23
	Equity and subordinated banks ratio	0.29	0.17	0.85	0.24	0.11	0.28	0.19	0.77	0.24	0.06
	Average loans size (thousands UAH)	2 354.81	2 487.22	9 210.24	1 691.69	9.07	1 768.14	2 022.20	6 223.23	686.55	27.11
	Average loans maturity (years)	2.08	1.14	4.45	2.00	0.39	2.26	1.40	5.95	2.09	0.42
	Assets/Bracnhes (UAH)	292 484 931	439 611 352	1 895 109 902	105 461 727	5 172 013	247 916 068	348 235 662	1 895 109 902	154 129 279	6 406 940

Table B. Migration of banks across business models*

As of 1h of 2016 / As of 2014	*Frozen/ Inactive*	*Universal*	*Corporate*	*Retail*	*Invest-ment/ Wholesale*	*House-holds-to-corporates*	*Went bankrupt*	*Total*
Frozen/ Inactive	6	5	0	0	1	0	23	35
Universal	2	14	0	1	0	7	12	36
Corporate	0	0	4	0	0	0	2	6
Retail	0	3	0	12	0	1	4	20
Invest-ment/ Wholesale	0	0	1	0	1	0	3	5
House-holds-to-corporates	3	13	0	0	0	19	32	67
Total	11	35	5	13	2	27	76	169

*The number of banks that participated in the analysis is slightly less than official number of the factual number of banks. We consider some banking institutions outliers, therefore their inclusion might have distorted the output of the model.