Bank Complexity and Risk

Elizaveta Kamaraeva, International College of Economics and Finance, National Research University Higher School of Economics elizabeth.kamaraeva@yandex.ru

The consolidation of banks into banking groups and holdings has been a prominent recent trend in Russia’s banking sector. To evaluate the effect of consolidation on the risk of a banking group, one needs a specific metric that captures organisational, business, and geographic complexity. In this paper, I consider different complexity types and proxies, and examine how complexity affects the risk of a banking group. Using data for 76 banking groups in Russia for 2015–2019, I find that for most of the complexity indicators there is a positive relationship between the organisational and business complexity on the one hand, and the risk of a banking group on the other hand. I also show that, in combination, different types of complexity have a positive effect on risk.

Keywords: bank complexity, risk taking, regulation, banking group, agency problem, diversification, Z-score

JEL Codes: G21, G32

Citation: Kamaraeva, E. (2020). Bank Complexity and Risk. Russian Journal of Money and Finance, 79(3), pp. 75–104. doi: 10.31477/rjmf.202003.75

1. Introduction

The Global Financial Crisis, a change in the monetary policy regime by the Bank of Russia, developments in banking regulation, and economic sanctions are all factors that have had a significant influence on the business models and organisational structures of Russian banks. A prominent recent trend in Russia’s banking system is the consolidation of banking institutions into banking groups and banking holding companies. The impact of this trend on the financial stability and the riskiness of individual institutions is ambiguous: on the one hand, larger and more complex banks are better able to diversify their sources of income (Laeven and Levine, 2007; Buch et al., 2013; Berger et al., 2017), benefit from economies of scale, and take advantage of regulatory arbitrage (Argimón and Rodríguez-Moreno, 2019), but on the other hand, the consolidation process may cause moral hazard and agency problems (Morgan, 2002; Dam and Koetter, 2012; Beck et al., 2017) and banks might also take on too much risk because they believe that they are too big to fail (Acharya et al., 2016; Cetorelli and Traina, 2018) or because of strategic complementarities (Farhi and Tirole, 2012).
To evaluate the effect of consolidation on risk, one needs a specific metric that captures a bank’s organisational, business, and geographic complexity. One proxy for complexity has been the size of a bank as measured by total assets, for example. But size and complexity do not necessarily exactly capture the same feature of a banking firm, and I confirm this in my study. The complexity of banks still remains underexplored in literature. Key stylised facts are not firmly established, and there is no single generally accepted metric of complexity.

This paper addresses two questions. First, how do we measure complexity of different types? In this part, I examine various approaches to defining and measuring bank complexity. Second, is there a relationship between bank complexity and risk? I examine not only the individual effects of different types of complexity on risk, but also their joint effect on the riskiness of a banking group.

In this study, I employ data from supervisory reports of 76 banking groups in Russia for the period between 2015 and 2019. The choice of this specific time interval for the analysis is determined by certain changes in bank regulation (e.g. the definition of ‘banking group’ was changed in 2014) and by the availability of data. As a proxy for risk, I use two metrics, the natural logarithm of the Z-score and the standard deviation of the rate of return on assets.

Following the literature (e.g. Cetorelli and Goldberg, 2014), I study three types of bank complexity: organisational, business, and geographic. Organisational complexity is proxied by the total number of affiliates and the number of non-financial affiliates under a parent organisation. Business complexity is measured by the count of business activities in which affiliates are involved and by the Herfindahl-Hirschman Index that measures the dispersion of affiliates across different business activities. As proxies for geographic complexity, I use a dummy variable that equals 1 if a banking group has foreign affiliates and 0 otherwise, and a metric that equals the total number of foreign affiliates.

There are a few reasons why this study is relevant to policy. First of all, familiarity with empirical facts about complexity is important for policymakers, allowing them to better understand the potential consequences of banking regulation reforms and analyse the changing structure of global financial markets. Moreover, understanding the relationship between complexity and bank risk can presumably help determine new criteria for systemically important banking organisations. Complexity, in addition, is a natural subject of policy concern given the systemic implications of resolving failing institutions. Many regulatory initiatives aim at influencing the stability of the financial system through making organisations more resilient and resolvable. Central banks should be able to evaluate a failing credit institution, its organisational structure and business processes, to estimate the costs of resolution and to make the decision whether to bail out the organisation or not. Furthermore, the relationship between complexity and risk is likely to be informative to domestic thinking about evolving features of national and hosted foreign banks. This research may also provide new insights into the dynamics of banks and the banking system.
This paper is organised as follows. Section 2 describes main facts about banking groups in Russia. Section 3 presents an overview of recent studies on bank complexity, relevant theory, banking group statistics and trends. Section 4 describes the data and methodology used in this research and outlines the econometric specification employed in this study. Section 5 presents empirical results and hypothesis tests on the relationship between complexity and risk. Section 6 concludes.

2. Institutional background: banking groups in Russia

The object of this research is a banking group. Under Russian banking legislation, a banking group is a collection of legal entities which is not a legal entity itself and in which one or several legal entities (banking group members) are under the control or significant influence of one credit institution (banking group parent credit institution). It is important to mention that this definition has been applied only since 2014. Before then, a different definition was in use: a banking group was a collection of legal entities which was not a legal entity itself and in which the parent organisation had significant influence, either directly or through third parties, on the decisions of the credit institutions that belonged to the group.

The main distinction between these two definitions is that, according to the old definition, the affiliates of a banking group could only be credit organisations, whereas, according to the new one, it is permissible for a banking group to include also non-bank companies. Moreover, the old definition did not specify whether the parent organisation of a banking group had to be a credit organisation or not. In other words, this definition did not distinguish between banking groups and banking holding companies. This fact is crucial for my study because I am interested precisely in the effect of the complex structure of banking groups on their riskiness. That is why I have chosen to run this research using data for 2015–2019.

According to Russian banking legislation, banking groups must submit, on a quarterly basis, a consolidated balance sheet statement, a consolidated income statement, a report with a calculation of own funds (capital), and comply with the list of mandatory standards and requirements for banking groups.

Figure 1 shows some descriptive statistics of banking groups in Russia. By inspecting this graph, one can draw several conclusions:

1. The number of banking groups gradually declined between 2016 and 2018.
2. The number of credit institutions that are members of banking groups also gradually declined.
3. At the same time, the concentration of assets in the banking sector increased; in 2019, 89% of overall banking sector assets were under the control of banking groups.

According to the Annual Report of the Bank of Russia for 2018, the largest banking groups have emerged around state-owned banks, and the number of affiliates in these groups can reach several hundreds.
Preliminary analysis of this information suggests that the operations of banking groups may have a non-negligible effect both on the default risk of the respective parent banks and on overall financial stability. But to evaluate this effect, it is not sufficient to use common metrics, such as the size of a bank, exposure to retail lending, exposure to interbank loans, etc. One also needs to take into account the organisational and business structure of a banking group, or, in other words, one needs to evaluate its complexity.

It is important to emphasise that complexity can be expressed not only by the number of affiliates in a banking group, but also by metrics that can capture the geographic complexity, the number of business types in which the banking group is involved, the number of non-bank affiliates, the number of affiliates that are located abroad, etc.

In the next section, I overview the literature on bank complexity and discuss relevant metrics that can be used to measure the complexity of banking groups.

### 3. Related literature

#### 3.1. Complexity metrics

While both size and complexity are important features of financial organisations, specific types of complexity and their patterns across banks are not well examined. In this part of the paper I consider various concepts of complexity that have been studied in the literature.

Among the most influential papers that explore the complexity of financial institutions are Cetorelli and Goldberg (2014) and Goldberg and Meehl (2019). Cetorelli and Goldberg (2014) identify such potential dimensions of complexity as the breadth of business scope, the width of geographic scope, and organisational
complexity through the numerosness of affiliated legal entities. Organisational complexity is measured as the count of all affiliated entities that belong to a banking group: these are bank, nonbank, and foreign entities. The reason why it is essential to consider this type of complexity was shown in the study of Cetorelli and Goldberg (2016), which finds evidence that the organisational complexity of a bank’s family is a main driving force behind the choice of business model of a financial institution. Using data from global US banks, the authors demonstrate that branches of conglomerates in more complex families are less sensitive to fluctuations in the cost and availability of funding. Moreover, the paper shows that the scale of the lending channel of global banks was changed by more than 30% due to the complexity of the respective business family.

Business complexity is shaped by the concentration of subsidiaries across types of business sectors. In theory, this type of complexity has an ambiguous effect on financial stability: on the one hand, it can mitigate the impact of shocks originating in one business branch, but, on the other hand, there exists a risk of the propagation of these shocks across all business sectors and even of an amplification of these disturbances in interlinked systems.

Geographic complexity captures the allocation of subsidiaries across regions and countries. On the one hand, geographic diversification can help to diminish the risk of large fluctuations in a certain region or country (i.e. the regional and country risk), but on the other hand, a shock originating in a certain region or country can produce a ‘domino effect’.

Table 1 provides a classification of banking organisations across domestic and foreign markets and by the type of entities within an organisation.

| Affiliate location | Affiliate type |
|--------------------|----------------|
|                    | Bank           | Non-bank financial | Non-financial |
| Home               | Counts of banks that are located in the home country | Counts of entities in financial industries that are located in the home country | Counts of entities in non-financial industries that are located in the home country |
| Foreign            | Counts of banks that are located in foreign countries | Counts of entities in financial industries that are located in foreign countries | Counts of entities in non-financial industries that are located in foreign countries |

Source: prepared by the author

To evaluate organisational complexity, Cetorelli and Goldberg (2014) use two main metrics, \( \text{Count} \) and \( \text{CountNBtoB} \). \( \text{Count} \) is the number of 50%+ owned affiliates under a parent organisation, and \( \text{CountNBtoB} \) is the ratio of the counts of non-bank and bank affiliates in which a parent has over 50%+ ownership. Non-bank subsidiaries can be used as a tool to conduct operations that are not permitted for banks. A business family may take advantage of relatively soft regulation of
non-bank affiliates, and, more precisely, it can incentivise banks to bear greater asset risk. Some authors also use the fraction of a parent bank’s foreign subsidiaries in all subsidiaries of the business family as a measure of organisational complexity. A high proportion of foreign affiliates may indicate greater opportunities for regulatory arbitrage. But there is the other side of the coin: where an affiliate is subject to the regulation of its host country, there exists a risk of coordination problems between the regulators of the home and host countries. There is much ongoing discussion about organisational complexity. Some economists suggest using size as a proxy for organisational complexity and, thus, equate these two concepts. Cetorelli and Goldberg (2014) show that this point of view is erroneous. As a metric for the size of a business, they use asset valuation. In their sample, all companies are divided into two groups: foreign banking organisations and the US global banks. The authors construct regressions between various measures of complexity and size. The main result is that the number of affiliates is positively correlated with the size of a bank holding organisation. In other words, larger organisations tend to have more subsidiaries. According to the data for foreign banks, an organisation that is twice as large as another one is likely to have 70% more affiliates. For other types of complexity, however, size is a poor predictor of the diversity of affiliate types. Similar results are obtained from the data for the US global banks.

The regression results of Cetorelli and Goldberg (2014) for the US banks and for foreign banks are reproduced in the Appendix (see Figures A1 and A2).

The metrics for business complexity and geographic complexity are constructed in a similar way. Both are based on the Herfindahl-Hirschman concentration index. For business complexity, the normalised Herfindahl-Hirschman Index (HHI) is determined by the distribution of subsidiaries across different business types relative to the total number of subsidiaries of a parent bank.

The HHI for business complexity is:

\[
BHHI_{i,t} = \frac{CountB_{i,t}}{CountB_{i,t} - 1} \left( 1 - \sum_{j=1}^{B} \left( \frac{count_{j,t}}{\sum_{k=1}^{B} count_{k,t}} \right)^2 \right),
\]

where \( CountB_{i,t} \) is the number of business types of the affiliates, \( count_{j,t} \) is the number of 50%+ owned subsidiaries, including the parent organisation, that are classified in accordance with each business type \( j \) for banking group \( i \) at time \( t \). To calculate \( BHHI_{i,t} \), the authors use the classification that is adopted in the Bankscope database, where all affiliate types are divided into five groups: banks, insurance companies, mutual and pension funds, other financial intermediaries, and non-financial subsidiaries (in this case \( B = 5 \)). As a possible extension, this system for the classification of banking group affiliates across different business types can be refined.

The output of this formula takes on values ranging between zero and one. A value of zero (one) indicates the lowest (highest) degree of complexity.
A more complex network of affiliates may yield economies of scale and, as a result, higher profitability. Moreover, a complex business family might create a buffer against the occurrence of losses in one or several sectors where its affiliates are active. However, companies with higher business complexity may incur increased transaction costs due to information exchange among subsidiaries in different business sectors. These increased costs can, in turn, diminish the performance of a bank: reduce its market power, decrease margins, etc. Higher business complexity can also cause difficulties for shareholders: the monitoring of business becomes more difficult and it provides incentives for increased risk-taking.

A similar approach is used for measuring geographic complexity. This metric may have different variations: for example, it can consider only companies which have foreign affiliates, and show the concentration of this business family across foreign countries, or it may be used within a certain country.

The \( HHI \) for geographic complexity is:

\[
GHHI_{t,t} = \frac{\text{Count}C_{i,t}}{\text{Count}C_{i,t} - 1} \left( 2 - \sum_{j=1}^{C} \left( \frac{\text{count}^j_{i,t}}{\sum_{k=1}^{C} \text{count}^k_{i,t}} \right) ^2 \right),
\]

where \( \text{Count}C_{i,t} \) is the set of countries (regions, geographic zones) and \( \text{count}^j_{i,t} \) is the count of subsidiaries in each country \( c \).

The index of geographic complexity varies from 0 to 1, where 0 means the lowest complexity and 1 means the highest complexity. The authors also used the classification adopted in the Bankscope database, where all regions are divided into 13 groups: 1) the euro area, 2) the United Kingdom, 3) Japan, 4) South Korea, 5) China, 6) Canada, 7) the United States, 8) Taiwan, 9) Middle East, 10) other Americas, 11) other Europe, 12) other Asia, 13) other.

A higher level of geographic complexity can help to alleviate local shocks and increase bank resilience, but at the same time it can also increase agency problems and the impact of the side effects of global shock.

Cetorelli and Goldberg (2014) made a great contribution to the research on the complexity of banking groups. Inspired by their study, many other scholars conducted their own research, some of them employing modified proxies for complexity.

Argimón and Rodríguez Moreno (2019) used an \( HHI \) based on the net worth of affiliates:

\[
HHI_{i,t} = \frac{R}{R - 1} \left( 1 - \sum_{j=1}^{R} \left( \frac{\text{net worth}^j_{i,t}}{\text{total net worth}_{i,t}} \right) ^2 \right),
\]

where \( R \) is the maximum number of affiliates observed in a banking group in the data sample and \( \text{net worth}^j_{i,t} \) is the net worth of affiliate \( j \) at time \( t \). Another proxy for complexity employed in their paper is
Deviation activity\textsuperscript{i,j}_{t,t}
\[= \left| \frac{\text{net worth abroad}_{i,t}^j}{\text{total net worth abroad}_{i,t}} - \frac{\text{net worth home}_{i,t}^j}{\text{total net worth home}_{i,t}} \right|,\]

where \text{net worth abroad}_{i,t}^j and \text{net worth home}_{i,t}^j are the net worths of affiliates of business type \(j\) located, respectively, abroad and at home; \text{total net worth abroad}_{i,t} and \text{total net worth home}_{i,t} are the total net worth of affiliates located, respectively, abroad and at home; \(|...|\) is the function returning the absolute value of its input.

The authors calculated these two metrics for five types of business activities: Banking, Other Financial Activities, Insurance, Real Estate, and Non-Financial Activities. The second metric indicates whether a banking group has more divergent activities carried out abroad in relation to those carried out at home. Argimón and Rodríguez Moreno (2019) are interested in the dimension of the difference, because their main hypothesis states that a greater degree of divergence leads to a lower degree of control, and, thus, to higher risk of a banking group. One of the reasons why these authors decided to use modified indices of complexity was the lack of information available for the research.

Following Cetorelli and Goldberg (2014), Kwan et al. (2019) constructed the two following measures of complexity, one for business complexity and the other for geographic complexity.

The index of business complexity is defined as:
\[HHI_{i,t} = \frac{T}{T-1} \left( 1 - \sum_{j=1}^{T} \left( \frac{\text{revenue}_{i,t}^j}{\text{total revenue}_{i,t}} \right)^2 \right),\]

where \(\text{revenue}_{i,t}^j\) is the revenue generated by affiliates of business type \(j\) and \(\text{total revenue}_{i,t}\) is the total revenue generated by a banking group \(i\) in period \(t\).

The index of geographic complexity is defined as:
\[G. \text{complexity}_{i,t} = 1 - \frac{\text{domestic revenue}_{i,t}}{\text{total revenue}_{i,t}},\]

where \(\text{domestic revenue}_{i,t}\) is the revenues generated by the global bank in its home country.

Moreover, based on the research of Bank for International Settlements Committee on the Global Financial System (2018) the authors constructed a metric for scope complexity:
\[\text{Scope Complexity}_{i,t} = \frac{\sum_{j=1}^{T} \text{distance}^j \times \text{revenue}_{i,t}^j}{\text{total revenue}_{i,t}},\]

where subscript \(j\) is associated with different business types from which a banking group generates revenues, and the variable \(\text{distance}^j\) takes on values of 1 to 5.
according to the North American Industry Classification System (NAICS). The scope complexity index measures the contributions of revenues that were generated by the non-bank affiliates of a banking group compared to its total revenues, weighted by the notional ‘distance’ of the respective non-bank business types from commercial banking.

The main objective of this subsection is to provide an overview of various approaches to defining and measuring complexity, and that is why I cannot ignore the important concept that was recently put forward by the Basel Committee on Banking Supervision (BCBS). BCBS (2014) developed a number of criteria to designate globally systemically important financial institutions or, in short, G-SIFIs. These criteria include size, cross-jurisdictional activity, interconnectedness, and complexity. Each indicator related to a certain criterion is assigned a specific weight. Table A1 in the Appendix provides information about the indicators and their weights. Complexity, in this context, is defined as the opacity of balance sheets and off-balance sheet exposures of banking groups. It increases with holdings of assets that have a complex structure and are difficult to price. Because these designation criteria are used in day-to-day regulatory decision making, it is useful to examine their reliability and to consider additional perspectives on complexity. The reason why complexity was included in the methodology is straightforward: G-SIFIs characterised by greater complexity are more difficult to resolve and, therefore, their failure is likely to cause a significantly greater disruption to the wider financial system and economic activity.

According to the BCBS methodology, complexity is presented by three main parameters: (i) the notional amount of over-the-counter (OTC) derivatives; (ii) trading and available-for-sale (AFS) securities; and (iii) the so-called Level 3 assets. The notional amount of OTC derivatives parameter is used to measure the level of engagement of the company in OTC derivative transactions and should include all types of risk categories and instruments. The volume of trading and AFS securities captures the value of bonds and shares that, if sold quickly, would incur larger fire-sale discounts or haircuts to compensate for high market risk in times of severe market stress. Level 3 assets include the fair value of all assets that are priced on a recurring basis on the balance sheet using the approach adopted in the International Accounting Standards (IAS). This approach commonly applies a three-level fair value hierarchy that prioritises inputs used to measure fair value based on observability.

One more concept of complexity is under development by the International Banking Research Network (IBRN). This concept focuses on the extent to which liquidity is managed centrally in a banking organisation. One of the main objectives of this specific initiative is to construct a metric to capture the magnitude and variability of liquidity flows within a banking organisation. To the best of my knowledge, this metric is still under construction.
To summarise, there is no single accepted one-size-fits-all metric for complexity. Different authors construct and employ different indicators of complexity for banking organisations depending on the data available and the specific research questions.

3.2. Risk metrics

According to the literature, there are several indicators that can be used to measure banks’ riskiness. All these metrics can be divided into three groups: (i) proxies that are based on the balance sheet of a bank, such as leverage, asset risk, default risk, liquidity risk, etc.; (ii) market-based indicators, such as total risk, tail risk, residual risk, etc.; and (iii) systemic risk measures, such as marginal expected shortfall (MES), CoVaR, etc. All these metrics are summarised in the Appendix (Table A2).

In this subsection, I discuss the metrics that are more commonly used for measuring banks’ riskiness.

Perhaps the earliest method of measuring risk is the so-called Z-score, which was employed, to mention a few, by Boyd and Graham (1986), Hannan and Hanweck (1988), and Boyd et al. (1993). There are different approaches to the construction of a time-varying Z-score. In this paper, I apply the methodology employed Lepetit and Strobel (2013). The Z-score is defined as:

$$Z\text{-score}_{i,t} = \frac{\text{ROA}_{i,t-k+1,t} + (E/A)_{i,t-k+1,t}}{\sigma(\text{ROA})_{i,t-k+1,t}},$$

where $\text{ROA}_{i,t-k+1,t}$ is the average return on assets of bank $i$ between dates $t - k + 1$ and $t$, $\sigma(\text{ROA})_{i,t-k+1,t}$ is the standard deviation of return on assets of bank $i$ between dates $t - k + 1$ and $t$, $(E/A)_{i,t-k+1,t}$ is the average equity to total assets ratio of bank $i$ between dates $t - k + 1$ and $t$. All components of this formula are computed over $k$ periods between dates $t - k + 1$ and $t$, where $k$ can vary depending on data availability. Most authors use $k$ equal to 8 or 12 quarters (e.g. Berger et al., 2017).

The Z-score shows the number of standard deviations that a bank’s return on assets has to drop below its expected value before equity is depleted, but a more general definition states that the Z-score measures the distance from insolvency (Roy, 1952). A higher Z-score indicates that a bank or a banking group has a lower risk of default. The Z-score can be used both to evaluate the individual risk of a bank and to estimate the risk of default for the whole banking system. To calculate the Z-score based on market information, one can use the following modified formula:

$$\text{Market Z-score}_{i,t} = 1 - \frac{\text{Equity Returns}_{i,t} + 1}{\sigma(\text{Stock Returns})_{i,t}},$$
where $\sigma(\text{Stock Returns})_{i,t}$ is the standard deviation of the idiosyncratic stock returns of banking group $i$, and $\text{Equity Returns}_{i,t}$ is the return on equity of the banking group $i$ (Correa and Goldberg, 2020). This metric is based on the rolling daily returns over the 120 days preceding date $t$.

The Z-score has two main advantages. First, the calculation of the Z-score is based on the supervisory data which is usually available to researchers. Second, the Z-score can be applied for companies that are not listed on the stock exchange. But it also has two important shortcomings. First, the Z-score is highly skewed, and this is why most authors use the natural logarithm of the Z-score (Laeven and Levine, 2009). Second, the Z-score is consistent only under the assumption of normally distributed returns on assets, which is viewed by many as too strong (Lapteacru, 2016).

Correa and Goldberg (2020) also use, as alternative risk metrics, the standard deviations of the average returns on assets (ROA) and returns on equity (ROE) of a banking group. In these cases, a greater value of a metric indicates a higher level of risk.

A second type of proxy for bank risk was suggested by Krause et al. (2016). The authors use a binary variable that equals one if bank $i$ is exposed to the risk of default and is bailed out by the government on date $t$, e.g. via recapitalisation or liability guarantees, and zero otherwise. The main difference of this binary variable from the Z-score, which signals the distance to default, is that the state aid indicator is a more precise signal that a bank has a high risk of default.

A third concept of bank risk was developed by Adrian and Brunnermeier (2016) in a paper titled ‘CoVaR’. According to this approach, $\Delta \text{CoVaR}_{i,t}$ is an indicator that shows the dependence between a particular banking organisation $i$ and the whole financial sector. It is defined as

$$\Delta \text{CoVaR}_{i,t} = \text{Var}_{EU,t} - \text{Var}_{i,t},$$

where $\text{Var}_{EU,t}$ is the value at risk in the Euro Area financial system conditional on the particular bank $i$ being in financial distress, $\text{Var}_{i,t}$ is the value at risk in the financial system conditional on bank $i$ being in its median state. In contrast to the previous indicators, $\Delta \text{CoVaR}_{i,t}$ reflects not only the idiosyncratic risk associated with a certain company, but also the systemic risk or, in other words, the risk that is associated with the entire financial system.

Another way to measure the risk of a banking group is to track the evolution of non-performing loans and bad debt. To evaluate this type of risk, Bonfim and Felix (2020) use two metrics, $\text{New NPL}$ and $\text{NPL}$, where $\text{New NPL}$ is the ratio of impairments and provisions recorded in each year to total assets, and $\text{NPL}$ is the ratio of the stock of non-performing loans to total assets. The higher the value of any of these metrics, the higher the risk of a banking group.

Finally, Bonfim and Felix (2020) also use the probability of default as a measure of risk. To construct this metric, the authors calculate the average
default probability of loans granted to firms by each bank, using internal credit models (Antunes et al., 2016).

To summarise, there are multiple metrics that can be used to evaluate the risk of a bank or a banking group. Different metrics capture different aspects of risk and require a great variety of data as input. So, in accordance with my objectives and the available data, I employ the logarithm of the Z-score and the standard deviation of the return on assets as two alternative proxies for bank risk.

3.3. Complexity and risk

There are several papers in the literature that my study is based upon. Krause et al. (2016) evaluate how different types of complexity relate to banks’ idiosyncratic and systemic riskiness. According to the paper, bank risk is affected by complexity, although the estimated effects differ depending on the specific complexity measure used. The results of this research suggest that banking organisations with a higher level of geographic complexity and a higher percentage of foreign subsidiaries tend to be less stable. Furthermore, a higher share of non-bank affiliates in the total number of affiliates of a banking group significantly decreases the chances of a government bail-out in a state of crisis.

Goetz et al. (2016) obtain opposite results. The authors find that banks that are more geographically diversified across the US states have a lower risk. Moreover, according to their findings, geographic complexity does not affect loan quality. They use the US data on the expansion of banks across states, distinguishing geographic diversification from business activity diversification. International diversification is not explicitly studied; a robustness test excludes banks that are internationally diversified.

Faia et al. (2017) examine if expansion to foreign markets increases the riskiness of a financial institution and identify the channels through which this usually happens. They document a strong negative correlation between the risk of an individual bank, as proxied by credit default swap (CDS) spreads and loan loss provisions, on the one hand, and foreign expansion, on the other hand. The reason why a bank’s foreign expansion reduces the individual riskiness is that foreign expansion leads to a better asset portfolio diversification with no evidence of any relevant effect of possible regulatory arbitrage. Moreover, the authors show that expansions decrease riskiness only in the case that competition in the origin country is less intense than competition in the destination country.

Chernobai et al. (2020) identify that business complexity has a strong positive effect on the occurrence of operational risk events. The authors claim that this result is more likely explained by managerial failure rather than by strategic risk taking. The explanation for this evidence is that business complexity generates greater operational risk in both the banking and
non-banking sectors of a banking group. Moreover, complexity does not appear to reward shareholders with a large or permanent improvement in a bank holding company’s (BHC) performance.

Correa and Goldberg (2020) show that complexity may have both positive and negative effects on risk, and the sign of this effect highly depends on banking governance and regulatory changes. Using data on large US BHCs for 2018, the authors find that the standard deviation of the return on equity (ROE) and the market Z-score of banking groups are decreasing in organisational and business complexity, while the geographic scope provides both diversification gains and increases market risk. Moreover, the authors show that banking groups with better governance are able to manage complex structures, and this benefits their income diversification and reduces their risk.

Gajewski and Kurowski (2019) investigate how the complexity of a parent organisation can affect risk taking by Polish subsidiaries of foreign banks. According to their results, more complex parent groups are associated with lower risk of their subsidiaries. It is worth mentioning that, initially, the authors conjectured the opposite relationship between the complexity of a parent organisation and the risk of its subsidiaries, due to agency problems within multinational bank companies. One possible interpretation of their finding is that the relationship between parent’s complexity and risk may not be driven by agency problems, but rather by some other mechanism; higher complexity does not necessarily exacerbate agency problems, and the documented empirical pattern may be a peculiarity of Poland’s banking system only.

Fritz et al. (2019) employ a sample of German banks for the period from 2005 to 2018 and show that banks that are subject to stricter capital regulation tend to increase their complexity in order to hide their risks and, thus, to reduce the regulatory burden.

One also needs to define the specific channels through which complexity may affect the risk of banking groups. The allocation of subsidiaries across different countries and/or across different business types within a banking group can make its management and monitoring more difficult, thus resulting in a higher risk. It might also lead banks to assume higher risks due to the intensification of agency problems (Liang and Rhoades, 1988). According to the study of Argimón and Rodríguez Moreno (2019) undertaken on Spanish data, the concentration of activity in business areas where a banking group has expertise should moderate agency problems if diversification is not carried out into higher risk activities, even if the returns are not highly correlated.

One of the channels where banking group complexity can have a significant impact on riskiness is regulatory arbitrage. Fritz et al. (2019) show that regulatory arbitrage can incentivise banking groups that are subject to strict capital requirements in their home country to increase geographic complexity in order to take higher risks. They also find that more stringent regulation at home forces
banks to reduce their riskiness, but overall, the increase in risk taking due to regulatory arbitrage offsets this risk reduction.

According to Argimón and Rodríguez Moreno (2019), complexity by itself is not a source of risk or diversification benefits. The authors conjecture that the differences in control or in the power that parent companies have over different subsidiaries may have a significant impact on the nature of these relationships.

4. Methodology and data

4.1. Hypotheses of interest

This section states the research question that I focus on and the main hypotheses that I test. My research question is: How does the complexity of a banking group affect its riskiness?

To answer this question, first, I explore the bilateral link between risk and different types of complexity for Russian banking groups. Second, I consider various combinations of complexity types and study their joint impact on risk.

My empirical analysis is focused on the effect of different types of complexity on a bank's risk. My first hypothesis is:

**H1: More complex banking groups engage in more risk taking.**

I conjecture that banking groups with high levels of complexity take more risk due to agency problems and moral hazard. As proxies for complexity, I use metrics that were suggested by Cetorelli and Goldberg (2014).

Next, I focus on estimating the impact of various combinations of different types of complexity on the risk of a banking group. I examine if the joint effect of different types of complexity on risk is also positive. So, my second hypothesis is:

**H2: In combination, different types of complexity have a positive effect on the overall risk of the banking group.**

4.2. Econometric specifications

To test my first hypothesis, I estimate the following model:

\[
Risk_{i,t} = \beta_0 + \beta_1 Complexity_{i,t} + \beta_2 X_{i,t} + \gamma_t + \alpha_i + \varepsilon_{i,t},
\]

where \( Complexity_{i,t} \) is a complexity indicator for banking group \( i \) on date \( t \), \( Risk_{i,t} \) is an indicator of the riskiness of banking group \( i \) on date \( t \), \( X_{i,t} \) is a set of bank-specific time-varying characteristics as controls, and \( \gamma_t \) is a set of time-fixed effects. One needs to control for potentially relevant banking group characteristics, because they can have an influence on both the banking group risk and complexity. For statistical inference, I use clustered standard errors to account possibility that the error term is serially correlated within entities.
The specification also contains entity-fixed effects \( \alpha_t \) to control for unobserved time-invariant heterogeneity across different groups that might be correlated with the regressor of interest. Time-fixed effects \( \gamma_t \) are included to capture the time-varying effect of macroeconomic developments such as changes in the Bank of Russia policy rate, the rate of inflation, the stage of the economic cycle, and other economy-wide events.

I am also interested in examining how three types of complexity in combination influence the riskiness of a banking group. To test my second hypothesis, I use the following specification:

\[
Risk_{i,t} = \beta_0 + \beta_1 \text{Combination Complexity}_{i,t} + \beta_2 X_{i,t} + \gamma_t + \alpha_t + \epsilon_{i,t},
\]

where \( \text{Combination Complexity}_{i,t} \) is the first principal component of proxies for three different types of uncertainty. I decided to use principal component analysis to construct an aggregate index that captures three different types of complexity simultaneously and avoids potential multicollinearity problems. Principal components are new variables that are constructed as linear combinations or mixtures of the input variables. Technically, this method employs the eigenvectors and eigenvalues of the variance-covariance matrix for the input variables, i.e. individual metrics of complexity of three different types. The first principal component is determined by the direction in the space of data observations such that the orthogonal projections of the input variables on it have the largest variance. In order to avoid a potential bias due to differences in the range of values for different input variables, I follow the conventional procedure and standardise the input variables by subtracting the mean and dividing by the standard deviation. As a robustness check, I also estimate the following regression:

\[
Risk_{i,t} = \beta_0 + \beta_1 \text{Mean Complexity}_{i,t} + \beta_2 X_{i,t} + \gamma_t + \alpha_t + \epsilon_{i,t},
\]

where \( \text{Mean Complexity}_{i,t} \) is calculated as the simple average of standardised indicators for three different types of complexity for banking group \( i \) at time \( t \).

4.3. Description of variables

In this subsection, I describe how complexity metrics, which are the regressors of interest in the analysis that follows, risk metrics, which are the dependent variables, and controls are constructed. Following Cetorelli and Goldberg (2014), I divide all complexity metrics into three groups: organisational, business, and geographic complexity. Table 2 defines the complexity metrics that are employed in this study.

Organisational complexity is measured by the number of non-financial affiliates of a banking group and by the total number of subsidiaries under the
parent credit organisation. To evaluate business complexity, I assign each affiliate of a banking group to one of 15 categories depending on its type of business according to Russia’s Classifier of Types of Economic Activity (OKVED). Then I calculate the number of different business types that are represented by the members of each banking group. Using this information, I also calculate the HHI according to formula (1) above.

**Table 2. Definition of complexity metrics**

| Type of complexity | Variable     | Definition                                                                 |
|--------------------|--------------|---------------------------------------------------------------------------|
| Organisational     | total_count  | The number of various entities within a banking group                     |
|                    | count_nonfin | Number of non-financial affiliates under the parent organisation of a banking group |
| Business           | HHI          | Normalised Herfindahl-Hirschman Index (HHI); determines the concentration of subsidiaries across different business types |
|                    | span_bus_type| The number of different business types in which the members of a banking group are engaged |
| Geographic         | has_foreign  | A dummy variable that equals one if a banking group has foreign affiliates and zero otherwise |
|                    | count_foreign| The number of affiliates of a banking group that are registered under foreign jurisdictions |

Geographic complexity is captured by the number of affiliates of a banking group that are located abroad and by a dummy variable that takes one if a banking group has foreign affiliates and zero otherwise.

The dependent variable in the regressions is the risk of a banking group. I proxy risk by two alternative metrics: (i) the standard deviation of the return on assets, and (ii) the natural logarithm of the Z-score. I abstract from any market-based measures of risk, since banking groups in Russia are not listed on the exchange (the parent organisation can be listed, but I am interested in consolidated data), and use proxies based on balance sheet and income statement data instead.

To calculate the Z-score, I use formula (2) above. Following the literature on bank complexity and risk, I employ a backward-looking version of the Z-score (Correa and Goldberg, 2020). Given the skewness of the distribution for the Z-score, I focus from here onwards on the natural logarithm of the Z-score, \( \ln(\text{Z-score}) \). The log Z-score measures the average profitability of a bank relative to the volatility of its returns (profitability) for a given period. A higher value of the log Z-score indicates a lower level of risk. I also use the standard deviation of the return on assets as an alternative proxy for the risk of a banking group. A lower value of \( \sigma(\text{ROA})_{i,t-k+1,t} \) means less volatile bank returns, and as a result, a lower level of risk.

In regressions of banking group risk on complexity, one needs to control for those banking group characteristics that are determinants of risk and are
correlated with complexity. The omission of such variables would lead to the omitted variable bias (OVB) problem. The control variables that I add to my regressions are listed in Table 3.

**Table 3. Control variables**

| Variable          | Formula                               | Definition                                                                 |
|-------------------|---------------------------------------|-----------------------------------------------------------------------------|
| ln_assets         | ln(total assets)                      | Proxy for the size of a banking group                                       |
| ROA               | Net income/Total assets               | Banking group profitability                                                  |
| cost_to_income    | Operational costs/Operational income  | Proxy for the efficiency of banking group management                       |
| loan_to_assets    | Loans/Total assets                    | Proxy for the business model of a banking group                            |

I control for bank size, which is proxied by the natural logarithm of total assets \((ln\_assets)\), since larger banks are likely to be perceived as too big to fail and the belief of bank management in implicit guarantees of a bailout by the government can lead to excessive risk taking (Bonfim and Felix, 2020). Moreover, larger banks may enjoy economies of scale in foreign exchange management. An additional reason why size may affect risk taking is that larger banks have greater capacity to absorb risks (Berger et al., 2017).

I also control for the share of loans in total assets \((loan\_to\_assets)\), since this metric characterises the business model of a banking group. A higher proportion of loans in total assets would indicate a traditional intermediation profile of a banking group, which is often associated with less risk. I also need to control for management efficiency and for the profitability of a banking group, which can be captured by the cost-to-equity ratio and by ROA, respectively.

### 4.4. Data

This research is based on data that is collected from the quarterly supervisory reports submitted by banking groups in Russia for the time period between 2015Q1 and 2019Q4.\(^1\) This specific choice of time span for analysis is determined by changes in banking legislation, namely, by the fact that the legal definition of a banking group was changed in 2014, and also by data availability.

The data sample consists of 259 observations for 76 banking groups. For each banking group, there is information about its affiliates. It is worth mentioning that this panel of data is unbalanced, since data for some banking groups are not available for the entire sample period. There are a few reasons for this:

- Several new banking groups were organised in the middle of the sample period due to the trend towards consolidation in the banking sector during this time interval.

---

\(^1\) Form 0409802, Form 0409803, Form 0409805.
- The licences of five parent credit organisations of banking groups were revoked during the sample period.
- Some banking groups classified the supervisory reports, which serve as the main source of data for this study, as confidential information, so the data for those banks are missing for respective dates.

Table 4. Availability of banking groups data by years

| Percentage of banking groups | 5% | 25% | 50% | 75% | 95% |
|-----------------------------|----|-----|-----|-----|-----|
| Number of years for which data are available | 1  | 4   | 5   | 5   | 5   |

As shown in Table 4, at least 5% of the sample has information for only 1 year, at least 25% of the sample has data for 4 years, and at least 50% of the sample has information for the entire sample period.

Figure 2. Banking groups statistics

Figure 2 depicts the number of banking groups and affiliates in our sample over the research period. Until 2017, there was an increase both in the number of banking groups, from 60 groups in 2015 to 67 in 2017, and the number of affiliates under parent organisations, from 475 affiliates in 2015 to 582 in 2017. After 2017, this trend reversed and the number of banking groups decreased to 63 in 2019, whereas the number of affiliates in banking groups moderately declined and reached 543 by 2019.

Below, I briefly describe the data collection process for two key variables in my study, complexity metrics and the Z-score, and report descriptive statistics for them and also for the control variables.

For complexity metrics, I started by collecting data on the number of affiliates of banking groups for the time period between 2016 and 2019 using the consolidated supervisory reports. It is worth mentioning that I am interested
only in affiliates that are under significant control of the parent organisation, and that is why I took into account only 50%+ owned affiliates under a parent organisation.

Next, for each affiliate, I determined its core business type using the OKVED and the country of residence. All affiliates were divided into 20 groups according to the OKVED: agriculture, mining, manufacturing, power generation, water supply, construction, wholesale trade, transportation and storage, hotels and catering, telecommunications, financial services, real estate, professional services and research, administrative services, public administration, education, healthcare, entertainment, other services, and activities of extraterritorial organisations. Five types of these business activities are not represented by members of any banking group in the sample, and they were dropped.

Based on this information, I calculated the span of business types that is covered by banking groups, the number of non-financial affiliates, and the Herfindahl-Hirschman Index that characterises the concentration of subsidiaries across different business types.

Table 5. Complexity metrics: descriptive statistics

|                | span_bus_type | total_count | count_nonfin | count_foreign | has_foreign | HHI  |
|----------------|---------------|-------------|--------------|---------------|-------------|------|
| **Mean**       | 2.13          | 8.29        | 2.83         | 0.83          | 0.32        | 0.30 |
| **St. dev.**   | 1.46          | 11.06       | 5.70         | 1.92          | 0.47        | 0.28 |
| **Min**        | 1.00          | 2.00        | 0.00         | 0.00          | 0.00        | 0.00 |
| **Max**        | 8.00          | 83.00       | 35.00        | 11.00         | 1.00        | 0.90 |

The main descriptive statistics for different complexity indicators are shown in Table 5. On average, banking groups in my sample have eight affiliates, of which two affiliates are non-financial companies. The maximum number of affiliates of a single banking group is 83. On average, the affiliates of a single banking group do business in two different industries. The mean value of the HHI equals 0.3, and the maximum value is 0.9. Approximately 30% of the banking groups in the sample have foreign subsidiaries. To summarise, a typical banking group has affiliates of several business types, some of which are located in foreign countries.

I also checked the correlation between different pairs of complexity metrics. Table 6 shows the correlation matrix for different complexity indicators. There is a high correlation between total_count and count_nonfin, which is not surprising given that these two complexity metrics proxy the same type of complexity. There is also a high correlation between the total number of subsidiaries and the number of subsidiaries that are located abroad, while the span of business types covered by affiliates is strongly correlated with the number of non-financial subsidiaries. For other pairs of complexity metrics,
the correlation is less pronounced, with the correlation coefficient below 0.5. Surprisingly, the correlation between span_bus_type and HHI is negative and quantitatively small.

Table 6. Correlation matrix for different complexity metrics

| Correlation Matrix | Organisational | Business | Geographic |
|--------------------|----------------|----------|------------|
|                    | total_count    | count_nonfin | span_bus_type | HHI | has_foreign | count_foreign |
| total_count        | 1              |           |             |     |             |               |
| count_nonfin       | 0.9063         | 1         |             |     |             |               |
| span_bus_type      | 0.5678         | 0.6330    | 1           |     |             |               |
| HHI                | 0.0298         | 0.0756    | -0.1178     | 1   |             |               |
| has_foreign        | 0.4210         | 0.3056    | 0.2742      | -0.0276 | 1           |               |
| count_foreign      | 0.7060         | 0.4992    | 0.4569      | -0.0771 | 0.5548      | 1             |

To construct the Z-score, I collected data for the following variables: the average return on assets, the standard deviation of the return on assets, and the equity-to-total-assets ratio. All these variables are calculated on a $k$-quarter rolling window, i.e. for overlapping time periods between date $t - k + 1$ and date $t$, where $k$ is the width of the rolling window. In accordance with the literature and the availability of data on the web-site of the Bank of Russia (www.cbr.ru), I chose to set $k = 8$ quarters, and that is why I collected balance sheet data for the time period between 2014Q1 and 2019Q4.

It is worth mentioning that there are five banking groups in my sample whose parent credit organisations lost their banking licence and therefore were out of business afterwards. In such instances, I decided not to use the balance sheet data for the two years preceding the licence revocation.

Table 7. Risk indicators: descriptive statistics

|                          | Z-score, 8 | ln(Z-score) | σROA, 8 |
|--------------------------|------------|-------------|---------|
| Mean                     | 19.39      | 2.47        | 1.5%    |
| St. dev.                 | 16.34      | 1.28        | 2.1%    |
| Min                      | 0.007      | -4.83       | 0.1%    |
| Max                      | 91.70      | 4.94        | 19%     |

Table 8. Banking group characteristics: descriptive statistics

|                          | ln_assets | loan_to_assets | ROA  | Operational costs/Income |
|--------------------------|-----------|----------------|------|--------------------------|
| Mean                     | 18.77     | 57%            | 0.6% | 52%                      |
| St. dev.                 | 1.95      | 19%            | 4.35%| 17%                      |
| Min                      | 14.18     | 8%             | -16% | 16%                      |
| Max                      | 24.17     | 78%            | 13.6%| 103%                     |
Table 7 depicts the main descriptive statistics for risk indicators. The mean values of $Z\text{-score}_8$ and $\ln(Z\text{-score})$ are 18.77 and 2.45, respectively. The $Z$-scores range from 0.007 to 91.7, which means that banks in my sample are characterised by different degrees of riskiness. The descriptive statistics of $\sigma \text{ROA}_8$, which is the standard deviation of the return on assets for the past eight quarters, confirms this conjecture.

Table 8 shows the main descriptive statistics for the banking group characteristics that serve as control variables in my regressions.

The reported descriptive statistics suggest that there is a substantial variation across different banking groups in my sample in terms of business model, which is captured by the loan-to-asset ratio. Furthermore, profitability and management efficiency also vary much across banks, with ROA ranging from -16% to 13.6%.

5. Empirical results

In order to test the first hypothesis, I estimate fixed effects panel regressions with different pairwise combinations of the indices of organisational, business, and geographic complexity as the regressor of interest and two risk metrics, namely, the log of $Z$-score and the standard deviation of the return on assets, as the dependent variables. Tables 9 and 10 report estimation results for these alternative models. Notations *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. The regression with $\ln(Z\text{-score})$ as the dependent variable is a log-level model, whereas the regression with the standard deviation of ROA as the dependent variable is a level-level model. The estimation results for these regressions suggest that there is a positive statistically significant association between the risk of a banking group and the following complexity metrics: the total number of subsidiaries, the number of non-financial subsidiaries, and the span of different types of business activities within a banking group.

The negative sign of the coefficient of complexity in the regression with the log of the $Z$-score as a metric of risk can be understood if one recalls that a lower value of the log $Z$-score indicates higher idiosyncratic risk of a banking group. An increase in the number of affiliates by one leads to a decrease in the $Z$-score by 5%, or by 0.9695 in levels, which is obtained as the sample mean $Z$-score multiplied by 0.05. This result is statistically significant at the 1% level. An increase of affiliates by one legal entity is quite considerable, since the sample mean of the total number of legal entities in a banking group is 8.29. An increase in the total count of affiliates by one standard deviation causes a decrease in the $Z$-score by 54% ($= -0.04941 \times 100 \times 11.06$), which suggests that the estimated effect is economically significant.

The second regression is the estimated relationship between the number of non-financial affiliates of a banking group and the $Z$-score of a banking group. The coefficient can be interpreted as follows: an increase in the number of non-
financial subsidiaries by one is associated with a change in the Z-score by -7.4%. According to the descriptive statistics, the sample mean of the total number of non-financial subsidiaries is approximately 3, whereas the standard deviation of this complexity metric is 5.7. An increase in the number of non-financial affiliates by one standard deviation is associated with a decrease in the Z-score by 42.18%. This effect is statistically significant at the 1% level of significance.

I also find evidence that there is a marginally significant, at the 10% level, positive relationship between the idiosyncratic risk of a banking group and the span of different business types covered by its affiliates. If the number of different business types within a banking group increases by one, the Z-score decreases by 16%. In general, the results are quantitatively stronger for business complexity than for organisational complexity. The more dispersed a banking group is with regard to the activities it engages in, the higher its risk is.

Geographic complexity, which is proxied by the foreign dummy variable and by the number of subsidiaries located abroad, is also found to have a positive relationship with risk, but these results are not statistically significant.

### Table 9. Risk and complexity. Dependent variable: logarithm of the Z-score

| Type of complexity   | Organisational | Business | Geographic |
|---------------------|----------------|----------|------------|
| Complexity metric   | total_count    | count_nonfin | span_bus_type | HHI | count_foreign | has_foreign |
| Complexity          | -0.04941***    | -0.07411*** | -0.16114**  | -0.25515 | -0.18761 | -0.17559   | (0.00918) | (0.01912) | (0.08074) | (0.25452) | (0.16685) | (0.22466) |
| Opex/operational income | -0.13314      | -0.15247  | -0.10070  | 0.01385  | -0.03045 | -0.00306  | (0.62501) | (0.62600) | (0.64983) | (0.65115) | (0.64199) | (0.66673) |
| ROA                 | -1.30703       | -1.85718  | -0.87621  | -0.51382 | -0.47782 | -0.31273  | (0.94883) | (1.2215)  | (0.96521) | (0.92664) | (0.81039) | (0.96180) |
| Ln(assets)          | 0.03682        | 0.00967  | 0.04799  | 0.08132  | 0.11378  | 0.10985   | (0.35216) | (0.35474) | (0.36262) | (0.35864) | (0.37457) | (0.39748) |
| Credit ratio        | 0.16835***     | 0.16730*** | 0.18843*** | 0.19730** | 0.18856** | 0.19734** | (0.03671) | (0.03664) | (0.05336) | (0.06981) | (0.05537) | (0.06653) |
| Time fixed effects  | yes            | yes      | yes      | yes      | yes      | yes       |
| R²                  | 10.12%         | 10.28%   | 8%       | 7.01%    | 7%       | 8%        |

The results of regressions with the standard deviation of the return on assets as a risk metric are shown in Table 10. These results are consistent with the results obtained for regressions with the Z-score. The positive sign of the coefficient on complexity in the former set of regressions can be understood if one recalls that higher variability of the return on assets signals higher risk.

The estimated regressions suggest that an increase in the total number of affiliates under the parent organisation by one legal entity leads to an increase in the standard deviation of ROA by 0.093%. The estimated effect is statistically...
significant at the 5% level. Note that this is a level-level model, but due to the fact that ROA is measured as a percentage, I present my results as percentages. To demonstrate economic significance, I reformulate the result as follows: a change in the total number of affiliates by one standard deviation results in an increase in the standard deviation of ROA by 1.028% (= 0.093 × 11.06), which is significant in comparison with the sample mean (1.5%) and the standard deviation (2.1%) of this risk metric.

Another conclusion that can be drawn from the estimation results is that there is a significant positive relationship between the number of non-financial affiliates and the risk of a banking group. An increase by one standard deviation in the number of non-financial affiliates results in an increase in the standard deviation of ROA by 1.09% (= 0.192 × 5.7). In general, the results are stronger for count_nonfin than for the total_count metric, and suggest that a banking group with a higher number of non-financial affiliates tends to be more risky. This fact aligns with the findings based on regressions with business complexity as the regressor of interest. According to those results, an increase by one standard deviation in the number of business types represented by the members of a banking group (1.46) is associated with an increase in the volatility of ROA by 0.865% (= 1.46 × 0.593). The estimated effect is statistically significant at the 5% significance level.

Similarly to regressions with the Z-score, the results of regressions with the HHI, the number of foreign affiliates, and the has_foreign dummy establish a positive relationship with the standard deviation of ROA, but the estimated effects are not statistically significant. According to the results, I fail to reject my first hypothesis.

Table 10. Risk and complexity. Dependent variable: standard deviation of ROA

| Type of complexity | Organisational | Business | Geographic |
|-------------------|----------------|----------|------------|
| Complexity metric | total_count    | count_nonfin | span_bus_type | HHI | count_foreign | has_foreign |
| Complexity        | 0.00093***     | 0.00192*** | 0.00593**  | 0.00150 | 0.00165 | 0.00169 |
|                   | (0.00036)      | (0.00074)  | (0.00241)  | (0.00369) | (0.00127) | (0.00271) |
| Opex/operational income | -0.01303    | -0.01135  | -0.01115  | 0.01617 | -0.01570 | -0.01591* |
|                   | (0.00927)      | (0.00903)  | (0.00876)  | (0.01042) | (0.01016) | (0.01040) |
| ROA               | -0.12654       | -0.10600* | -0.12590* | -0.14336 | -0.14328 | -0.14479 |
|                   | (0.08440)      | (0.06346)  | (0.07584)  | (0.10145) | (0.10110) | (0.10175) |
| Ln(assets)        | 0.00256        | 0.00343   | 0.00258   | 0.00199  | 0.00164  | 0.00164   |
|                   | (0.00807)      | (0.00814)  | (0.00832)  | (0.00798) | (0.0082)  | (0.00845) |
| Credit ratio      | 0.00084        | 0.00060   | 0.00103   | 0.00143  | 0.00134  | 0.00142   |
|                   | (0.00071)      | (0.00064)  | (0.0007)   | (0.00117) | (0.00109) | (0.00114) |
| Time fixed effects| yes            | yes       | yes       | yes      | yes      | yes       |
| R²                | 22%            | 26.31%    | 23.3%     | 16%      | 16.26%   | 16%       |
While it is not straightforward to interpret the estimated relationship between risk and complexity as causal, one can view them through the lens of recent studies. One explanation is that more complex banks are more opaque and therefore are more difficult to manage and monitor, and this gives rise to moral hazard and agency problems (Morgan, 2002; Dam and Koetter, 2012; Duchin and Sosyura, 2014; Beck et al., 2017). Moreover, the positive relationship between the complexity of a banking group and risk can be explained by the conjecture that more complex banking groups, which are typically large, are likely to take on too much risk, since they believe that they are too big to fail (Acharya et al., 2016; Cetorelli and Traina, 2018). This explanation is in line with the fact that the largest banking groups in Russia have emerged around state-owned banks and the number of subsidiaries in these groups can reach several hundred.

I am also interested in estimation of the effect of a combination of different complexity metrics on the risk of a banking group. I want to examine whether the estimated combined effect on risk is similar to the effect of individual complexity proxies. To evaluate this effect, I need a specific metric that will capture all three types of complexity jointly. Moreover, due to the high correlation between different complexity metrics, I need to eliminate the potential multicollinearity problem. One available solution is to employ the principal component analysis.

Table 11 represents the first principal components that I constructed, based on the indicators of three different types of complexity. The first column in Table 11 lists different combinations of the proxies for business, organisational, and geographic complexity. Columns 2, 3, and 4 show the factor loadings of individual complexity metrics, and the sum of squares of these coefficients equals one for each combination of complexity indicators. Column 5 shows the fraction of variation that is explained by the first principal component. In accordance with the Kaiser criterion, I use only one principal component as the aggregate complexity metric in my regressions since, for each combination of individual complexity metrics of three different types, only one eigenvalue of the variance-covariance matrix of inputs is above one.

Table 11. First principal component for individual metrics of complexity of three different types: factor loadings and fraction of variation explained

| Combination                                                   | Business | Geographic | Organisational | Variation |
|---------------------------------------------------------------|----------|------------|----------------|-----------|
| span_bus_type & has_foreign & total_count                    | 0.5850   | 0.5021     | 0.6369         | 72%       |
| span_bus_type & has_foreign & count_nonfin                   | 0.6300   | 0.4411     | 0.6393         | 71%       |
| span_bus_type & count_foreign & total_count                  | 0.5308   | 0.5833     | 0.2651         | 72%       |
| span_bus_type & count_foreign & count_nonfin                 | 0.5899   | 0.5368     | 0.6033         | 78%       |
| HHI & has_foreign & total_count                              | 0.0039   | 0.7070     | 0.7072         | 57%       |
| HHI & has_foreign & count_nonfin                             | 0.1122   | 0.6932     | 0.7120         | 49%       |
| HHI & count_foreign & total_count                            | -0.0475  | 0.7081     | 0.7045         | 57%       |
| HHI & count_foreign & count_nonfin                           | -0.0022  | 0.7073     | 0.7076         | 52%       |
Table 12 below reports the descriptive statistics of the first principal component for each combination of individual metrics of complexity of three different types. Even though the individual complexity metrics were standardised before extracting the first principal component, the aggregate metric has a standard deviation different from one for all cases.

**Table 12. First principal component for individual metrics of complexity of three different types: descriptive statistics**

|                          | Mean | St. dev. | Min   | Max   |
|--------------------------|------|----------|-------|-------|
| `span_bus_type & has_foreign & total_count` | 0    | 1.361    | -1.166 | 7.034 |
| `span_bus_type & has_foreign & count_nonfin` | 0    | 1.354    | -1.114 | 6.394 |
| `span_bus_type & count_foreign & total_count` | 0    | 1.469    | -1.020 | 8.687 |
| `span_bus_type & count_foreign & count_nonfin` | 0    | 1.436    | -0.995 | 7.911 |
| `HHI & has_foreign & total_count` | 0    | 1.192    | -0.893 | 5.856 |
| `HHI & has_foreign & count_nonfin` | 0    | 1.144    | -0.949 | 5.219 |
| `HHI & count_foreign & total_count` | 0    | 1.305    | -0.802 | 8.047 |
| `HHI & count_foreign & count_nonfin` | 0    | 1.224    | -0.664 | 7.326 |

Using these first principal components as the regressor of interest, I estimate two sets of regressions, one with \(\ln(\text{Z-score})\) and the other with standard deviation of ROA as the dependent variables. As before, the specification of the model also includes control variables, as well as banking group and time-fixed effects. Tables 13 and 14 report the estimation results for these regressions.

**Table 13. The effect of aggregate complexity on the riskiness of a banking group. Dependent variable: Z-score**

| Business & Geographical & Organisational                      | First principal component beta |
|---------------------------------------------------------------|--------------------------------|
| `span_bus_type & has_foreign & total_count`                  | -0.2927*** (0.1217)          |
| `span_bus_type & has_foreign & count_nonfin`                 | -0.2658*** (0.1076)          |
| `span_bus_type & count_foreign & total_count`                | -0.3363*** (0.1122)          |
| `span_bus_type & count_foreign & count_nonfin`               | -0.2923* (0.1028)            |
| `HHI & has_foreign & total_count`                            | -0.2664* (0.1473)            |
| `HHI & has_foreign & count_nonfin`                           | -0.2778* (0.1354)            |
| `HHI & count_foreign & total_count`                          | -0.4273* (0.1389)            |
| `HHI & count_foreign & count_nonfin`                         | -0.4124* (0.1086)            |
| Banking group controls                                       | yes                           |
| Time-fixed effects                                           | yes                           |
The estimation results for the regressions shown in Tables 13 and 14 suggest that there is a significant negative relationship between the aggregate complexity and the Z-score, and a significant positive relationship between the aggregate complexity and the standard deviation of ROA. Both sets of regressions provide evidence that aggregate complexity has a positive relationship with the risk of a banking group.

### Table 14. The effect of aggregate complexity on the riskiness of a banking group.

| Dependent variable standard deviation of ROA | First principal component beta |
|---------------------------------------------|-------------------------------|
| span_bus_type & has_foreign & total_count  | 0.006*** (0.002)              |
| span_bus_type & has_foreign & count_nonfin | 0.007*** (0.002)              |
| span_bus_type & count_foreign & total_count| 0.006*** (0.002)              |
| span_bus_type & count_foreign & count_nonfin| 0.007*** (0.003)              |
| HHI & has_foreign & total_count            | 0.004* (0.005)                |
| HHI & has_foreign & count_nonfin           | 0.005* (0.001)                |
| HHI & count_foreign & total_count          | 0.007* (0.009)                |
| HHI & count_foreign & count_nonfin         | 0.009* (0.020)                |

It is worth mentioning that principal component analysis is often used as a tool to overcome the multicollinearity problem by aggregating individual complexity metrics of three different types into a single index that captures the maximum variation of the data. It is quite difficult to interpret the results in terms of input variables, but it is possible to do so using the standard deviation of the first component, also taking into account the factor loadings of individual metrics in the first principal component.

The estimation results of the regression with the Z-score as a dependent variable suggest that an increase by one standard deviation in the aggregate complexity index constructed from the span of business types, the has_foreign dummy, and the total number of subsidiaries (span_bus_type & has_foreign & total_count) results in a decrease in the Z-score by 39.83% (= 0.2927 × 1.361) or, in levels, by 7.723 (= mean Z-score × 39.83%). I conclude that, in comparison with the standard deviation of the Z-score (16.34), the size of the effect (7.723) is economically significant. The same aggregate complexity index, with the exception that the total number of subsidiaries is replaced by the number of non-financial affiliates as a metric of organisational complexity (span_bus_type & has_foreign & count_nonfin), features an estimated effect on the Z-score equal to 35%.
An increase by one standard deviation in the combination of the span of business types, the number of foreign affiliates, and the total number of subsidiaries decreases the Z-score by 49% \((= 0.3363 \times 1.469)\). For all three of these specifications shown on the top of Table 13, the estimated coefficient on the aggregate complexity index is statistically significant at the 1% level, and the estimated effect is also economically significant.

To summarise, an increase by one in the aggregate index of complexity is found to have a negative impact on the Z-score. The size of this impact ranges from \(-42\%\) to \(-26\%\) depending on which specific inputs are used to construct the first principal component.

Table 14 shows the estimation results for regressions with the standard deviation of ROA as the dependent variable. The results obtained suggest that an increase by one in the aggregate complexity index induces an increase in the standard deviation of ROA in the range from 0.4% to 0.9% depending on which specific inputs are used to construct the first principal component. Since the standard deviation of ROA has a sample mean of 1.5% and a standard deviation of 2.1%, I conclude that the estimated effect of interest is economically significant. For all specifications, the estimated coefficient on the aggregate complexity index is statistically significant either on the 1% or 10% levels.

Based on the estimation results for the regressions shown in Tables 13 and 14, I fail to reject my second hypothesis. I conclude that the three different types of complexity in combination have a positive effect on risk. In other words, a banking group with a higher geographic, business, and organisational complexity tends to be riskier.

6. Conclusion

Banking group complexity may have an ambiguous effect on the risk of an individual credit organisation and the stability of the entire banking system. On the one hand, banking group complexity is often perceived by policymakers and regulators as a potential source of risk due to moral hazard and agency problems. But on the other hand, complexity may give rise to diversification and scale economies, thus, reducing risks. In this paper, I provide empirical evidence of the relationship between bank complexity and the riskiness of banking groups in Russia. Following the literature, I consider three different complexity types and a set of alternative proxies for each of them. As in Cetorelli and Goldberg (2014), I divide all complexity types into three groups: organisational, business, and geographic. For each type of complexity, I construct two metrics: the total number of affiliates and the number of non-financial affiliates for organisational complexity, the Herfindahl-Hirschman Index and the span of business types for business complexity, and the number of affiliates located abroad and the has_foreign dummy for geographic complexity. Then I run two sets of regressions
with the Z-score and the standard deviation of the return on assets, which are proxies for the risk of a banking group, as alternative dependent variables.

I find evidence that there is a significant positive relationship between the riskiness of a banking group and the total number of subsidiaries, the number of non-financial subsidiaries, and the span of business types. The estimated effect is generally quantitatively stronger for business complexity than for organisational complexity.

I also study the effect of aggregate complexity on the risk of a banking group. To estimate this effect, I construct the aggregate index of complexity using principal component analysis. I find evidence that different types of complexity in combination have a positive effect on risk.

The results obtained may be important for policymakers, as they allow a better understanding of the perceived effects of regulatory reforms and analyse the changing structure of global financial markets. Moreover, a better understanding of the relationship between complexity and bank risk may help determine new criteria for systemically important banking organisations.

Conducting this research, I encountered several problems that may affect the validity of my results. The first problem is the limited availability of data on banking groups in Russia. Some banking groups choose not to disclose data on their affiliates, or the country of their residence, etc., and this might lead to a sample selection bias. The second problem is that my dataset covers a relatively short period of time, namely, from 2015 to 2019. This choice of sample period is determined by some important developments in banking legislation, in particular, a change in the definition of a banking group in 2014, and also by data availability. Expansion of the sample period outside 2015–2019 may influence my results. The third problem is that, due to the relative novelty of this topic, there is still a shortage of publications with convincing interpretations of results.

This paper does not address many issues that are likely to be fruitful directions for further research. One of these is the determination of the channels through which complexity affects the riskiness of a banking group. It is also worth investigating how specific changes in regulation influence the complexity of banking groups and their riskiness. Another interesting extension of this study would be to examine whether the effect of complexity on risk depends on other banking group characteristics.

Appendix is available at
http://rjmf.econs.online/en;
dx.doi.org/10.31477/rjmf.202003.75

References

Acharya, V., Anginer, D. and Warburton, J. (2016). The End of Market Discipline? Investor Expectations of Implicit Government Guarantees. Working Paper, NYU Stern.
Adrian, T. and Brunnermeier, M. K. (2016). CoVaR. American Economic Review, 106(7), pp. 1705–1741. doi: 10.1257/aer.2012055

Antunes, A., Gonçalves, H. and Prego, P. (2016). Firm Default Probabilities Revisited. Economic Bulletin and Financial Stability Report Articles, Banco de Portugal.

Argimón, I. and Rodríguez Moreno, M. (2019). Business Complexity and Geographic Expansion in Banking. Working Paper, Bank of Spain.

Bank for International Settlements Committee on the Global Financial System (2018). Structural Changes in Banking after the Crisis. CGFS Papers, N 60.

Brunnermeier, M. K. (2014). The G-SIB Assessment Methodology – Score Calculation. Available at: https://www.bis.org/bcbs/publ/d296.htm [accessed on 26 July 2020].

Beck, T., Da-Rocha-Lopes, S. and Silva, A. (2017). Sharing the Pain? Credit Supply and Real Effects of Bank Bail-in’. CEPR Discussion Paper, N 12058. Available at: https://cepr.org/active/publications/discussion_papers/dp.php?dpno=12058 [accessed on 26 July 2020].

Berger, A. N., El Ghoul, S., Guedhami, O. and Roman, R. A. (2017). Internationalization and Bank Risk. Management Science, 63(7), pp. 2283–2301.

Bonfim, D. and Felix, S. (2020). Banks’ Complexity and Risk: Agency Problems and Diversification Benefits. Banco de Portugal Working Papers, N 10.

Boyd, J. H. and Graham, S. L. (1986). Risk, Regulation, and Bank Holding Company Expansion into Nonbanking. Federal Reserve Bank of Minneapolis Quarterly Review, 10(2), pp. 2–17.

Boyd, J. H., Graham, S. L. and Hewitt, R. S. (1993). Bank Holding Company Mergers with Nonbank Financial Firms: Effects on the Risk of Failure. Journal of Banking and Finance, 17(1), pp. 43–63. doi: 10.1016/0378-4266(93)90079-S

Buch, C. M., Koch C. T. and Koetter, M. (2013). Do Banks Benefit from Internationalization? Revisiting the Market Power – Risk Nexus. Review of Finance, 17(4), pp. 1401–1435. doi: 10.1093/rof/rof033

Cetorelli, N. and Goldberg, L. S. (2014). Measures of Global Bank Complexity. Federal Reserve Bank of New York Economic Policy Review, 20(2), pp. 107–126.

Cetorelli, N. and Goldberg, L. S. (2016). Organizational Complexity and Balance Sheet Management in Global Banks. NBER Working Papers, N 22169.

Cetorelli, N. and Traina, J. (2018). Resolving “Too Big to Fail”. Federal Reserve Bank of New York Staff Reports, N 859.

Chernobai, A., Ozdagli, A. and Wang, J. (2020). Business Complexity and Risk Management: Evidence from Operational Risk Events in U.S. Bank Holding Companies. Journal of Monetary Economics. Article in press. doi: 10.1016/j.jmoneco.2020.02.004.

Correa, R. and Goldberg, L. S. (2020). Bank Complexity, Governance, and Risk. NBER Working Papers, N 27547.

Dam, L. and Koetter, M. (2012). Bank Bailouts and Moral Hazard: Evidence from Germany. Review of Financial Studies, 25(8), pp. 2343–2380. doi: 10.1093/rfs/hhs056
Duchin, R. and Sosyura, D. (2014). Safer Ratios, Riskier Portfolios: Banks’ Response to Government Aid. *Journal of Financial Economics*, 113(1), pp. 1–28.

Faia, E., Ottaviano, G. and Sanchez Arjona, I. (2017). *International Expansion and Riskiness of Banks*. CEPR Discussion Papers, N 11951.

Farhi, E. and Tirole, J. (2012). Collective Moral Hazard, Maturity Mismatch, and Systemic Bailouts. *American Economic Review*, 102(1), pp. 60–93. doi: 10.1257/aer.102.1.60

Fritz, B., Martynova, N. and Vogel, U. (2019). *Banks’ Complexity-Risk Nexus and the Role of Regulation*. Working Paper, Deutsche Bundesbank.

Gajewski, K. and Kurowski, Ł. (2019). *Agency Problem in Multinational Banks? How Parent Complexity Affects Risk-Taking Behavior of Subsidiaries?* Working Paper, National Bank of Poland.

Goetz, M. R., Laeven, L. and Levine, R. (2016). Does the Geographic Expansion of Banks Reduce Risk? *Journal of Financial Economics*, 120(2), pp. 346–362. doi: 10.1016/j.jfineco.2016.01.020

Goldberg, L. S. and Meehl, A. (2019). *Complexity in Large U.S. Banks*. Federal Reserve Bank of New York Staff Reports, N 880.

Hannan, T. H. and Hanweck, G. A. (1988) Bank Insolvency Risk and the Market for Large Certificates of Deposit. *Journal of Money, Credit and Banking*, 20, pp. 203–211.

Krause, T., Sondershaus, T. and Tonzer, L. (2016). *The Role of Complexity for Bank Risk during the Financial Crisis: Evidence from a Novel Dataset*. IWH Discussion Papers, N 17.

Kwan, S. H., Ho, K. and Tan, E. (2019). *Complexity of Global Banks and Their Foreign Operation in Hong Kong*. HKIMR Working Paper, N 14. Available at: https://ssrn.com/abstract=3490461 [accessed on 26 July 2020].

Laeven, L. and Levine, R. (2007). Is There a Diversification Discount in Financial Conglomerates? *Journal of Financial Economics*, 85(2), pp. 331–367.

Laeven, L. and Levine, R. (2009). Bank Governance, Regulation and Risk Taking. *Journal of Financial Economics*, 93(2), pp. 259–275. doi: 1016/j.jfineco.2008.09.003

Lapteacru, I. (2016). On the Consistency of the Z-score to Measure the Bank Risk. *Larefi Working Papers*, N 1604.

Lepetit, L. and Strobel, F. (2013). Bank Insolvency Risk and Time-Varying Z-score Measures. *Journal of International Financial Markets, Institutions and Money*, 25, pp. 73–87. doi: 10.1016/j.intfin.2013.01.004

Liang, N. and Rhoades, S. A. (1988). Geographic Diversification and Risk in Banking. *Journal of Economics and Business*, 40(4), pp. 271–284.

Morgan, D. P. (2002). Rating Banks: Risk and Uncertainty in an Opaque Industry. *American Economic Review*, 92(4), pp. 874–888. doi: 10.1257/00028280260345006

Roy, A. D. (1952). Safety First and the Holding of Assets. *Econometrica*, 20(3), pp. 431–449. doi: 10.2307/1907413