DEUTSCHE
BUNDESBANK
EUROSYSTEM

Discussion Paper

Deutsche Bundesbank
No 53/2013

Banks' concentration versus diversification in the loan portfolio: new evidence from Germany

Nadya Jahn
(University of Muenster)

Christoph Memmel
(Deutsche Bundesbank)

Andreas Pfingsten
(University of Muenster)

Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-0

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Internet http://www.bundesbank.de

ISBN 978–3–86558–994–1 (Printversion)
ISBN 978–3–86558–995–8 (Internetversion)

Non-technical summary

By carefully selecting and monitoring the borrower, a bank can influence the risk-return-profile of the loan. Based on the theory on financial intermediation, a bank can accumulate expertise if it concentrates lending on certain industries or sectors. This expertise is valuable to the bank when it comes to choosing and monitoring borrowers, as it allows the bank to keep write-offs and write-downs in its credit portfolio below average compared with loans to similar borrowers.

We investigate the validity of this theory for all universal banks in Germany using a unique dataset. This dataset includes each bank's lending activity to the German real economy, broken down into 27 industries or sectors and into three maturity brackets as well as the corresponding write-offs and write-downs. On the basis of the dataset, we can show the following for the period 2003-2011:

- The more concentrated the credit portfolio of a bank is (with respect to industries/sectors), the lower are the expected write-offs and write-downs in its credit portfolio. We control for the composition of a bank's credit portfolio.
- The rate of write-offs and write-downs is calculated for each bank and for each industry/sector. We reveal that the average rate over all banks and the observation period is lower for each industry/sector which represents a focus of lending for a bank. This result holds for 25 out of 27 industries/sectors.
- The more concentrated the credit portfolio of a bank is, the lower is the unexpected risk of the credit portfolio, where the unexpected loss is measured by the timely standard deviation of the loan loss rate. However, this relationship is less pronounced than in the case of the expected losses.

Nichttechnische Zusammenfassung

Eine Bank kann durch die sorgfältige Auswahl des Schuldners und dessen Überwachung das Ertrags- und Risikoprofil des Kredites beeinflussen. Gemäß der Theorie der Finanzintermediation kann eine Bank Fachkenntnisse aufbauen, wenn sie die Kreditvergabe auf wenige Wirtschaftszweige oder Sektoren konzentriert. Diese Fachkenntnisse helfen der Bank bei der Auswahl und Überwachung der Schuldner, so dass die Abschreibungen und Wertberichtigungen in ihrem Kreditportfolio unter dem Durchschnitt für Kredite an ähnliche Schuldner liegen.

Wir untersuchen die Stichhaltigkeit dieser Theorie für alle Universalbanken in Deutschland, wobei wir einen einzigartigen Datensatz verwenden. Dieser Datensatz enthält für jede Bank deren Kreditvergabe an die deutsche Realwirtschaft, aufgeteilt in 27 Wirtschaftszweige oder Sektoren und in drei Laufzeitbänder. Darüber hinaus sind in diesem Datensatz die entsprechenden Abschreibungen und Wertberichtigungen aufgeführt. Auf Basis dieses Datensatzes können wir für den Zeitraum 2003 bis 2011 Folgendes empirisch zeigen:

- Je konzentrierter (im Hinblick auf die Wirtschaftszweige/Sektoren) das Kreditportfolio einer Bank ist, desto geringer sind deren erwartete Abschreibungen und Wertberichtigungen im Kreditportfolio. Dabei wird der Zusammensetzung des Portfolios Rechnung getragen.
- Die Abschreibungs- und Wertberichtigungsrate wird für jede Bank und jeden der 27 Wirtschaftszweige/Sektoren berechnet. Es zeigt sich, dass diese Rate im Mittel über die Banken und den Beobachtungszeitraum geringer ist, wenn der Wirtschaftszweig/Sektor für eine Bank eine große Bedeutung hat. Dieses Ergebnis gilt für 25 der 27 Wirtschaftszweige/Sektoren.
- Je höher die Konzentration im Kreditportfolio, desto geringer das unerwartete Risiko des Kreditportfolios, gemessen als zeitliche Standardabweichung der Verlustrate im Kreditportfolio. Dieser Zusammenhang ist aber weniger deutlich ausgeprägt als bei den erwarteten Verlusten.

BUNDESBANK DISCUSSION PAPER NO 53/2013

Banks' concentration versus diversification in the loan portfolio: new evidence from Germany

Nadya Jahn[*]
(University of Muenster)

Christoph Memmel[†]
(Deutsche Bundesbank)

Andreas Pfingsten[‡]
(University of Muenster)

Abstract

Using a unique data set on German banks' sector specific loan exposures to the real economy and the corresponding write-offs and write-downs, we examine the impact of loan portfolio sector concentration on credit risk. By controlling for common risk factors, we separate the bank-specific selection and monitoring abilities from the composition of the loan portfolio. In our empirical study for the period 2003-2011, we find that (a) banks which are specialized in certain industries have, on average, lower loan losses, (b) the loss rate of a given industry in a bank's loan portfolio is lower if the bank has major exposures to this industry, and (c) the standard deviation of the loan losses is lower in the case of more focused banks.

Keywords: loan portfolio, credit risk, loan losses, concentration.

JEL Classification: G11, G21, C23, C43.

[*] Finance Center Muenster, University of Muenster, Universitaetsstraße 14-16, D-48143 Muenster, Germany, phone: +49 (0) 251 83 21881, e-mail: nadya.jahn@wiwi.uni-muenster.de.
[†] Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, D-60431 Frankfurt, Germany, phone: +49 (0) 69 9566 8531, e-mail: christoph.memmel@bundesbank.de.
[‡] Finance Center Muenster, University of Muenster, Universitaetsstraße 14-16, D-48143 Muenster, Germany, phone: +49 (0) 251 83 22881, e-mail: andreas.pfingsten@wiwi.uni-muenster.de.
We would like to thank Philipp Koziol and the participants of the 16th Conference of the Swiss Society for Financial Market Research 2013, the Bundesbank Research Seminar and the ECB Seminar. The views expressed in this paper are those of the authors and do not necessarily reflect the opinions of the Deutsche Bundesbank.

Banks' concentration versus diversification in the loan portfolio: new evidence from Germany

1 Introduction

Credit business is the core of traditional commercial banking. By carefully selecting and monitoring the borrower, the bank can, on average, reduce the credit risk of the loan portfolio. Selection and monitoring abilities depend on bank characteristics. One important feature is the sector concentration in the loan portfolio, whereas other characteristics are barely quantifiable – for instance, the management and process quality. The more a bank focuses on certain industries, the more it can acquire industry-specific knowledge and thereby realize specialization benefits, i.e. reduce, on average, the credit risk of the loan portfolio. However, a concentrated loan portfolio harbours increased concentration risks due to higher default correlations of borrowers within a given industry. The question of whether knowledge gains from loan portfolio concentration outweigh the foregone diversification benefits is rarely discussed in the theoretical and empirical banking literature.

We have an especially suitable data set for Germany for the period 2003-2011 at our disposal, comprising the banks' loan exposures to the non-financial private sector and the corresponding write-offs and write-downs, both of which can be broken down into different industries and maturity brackets. The banks' loan exposures on an industry-specific level are used to measure the concentration in the loan portfolio. Further, common risk factors are included to control for the composition of the loan portfolio. These are calculated from averages of the corresponding write-offs and write-downs, conditional on various characteristics, such as industry or maturity. This allows us to analyse the impact of loan portfolio concentration on credit risk, where credit risk is approximated using the historic loan losses.[1]

The contribution of our analysis is twofold. In general, the credit risk of a loan portfolio can be further separated into an ex ante expected portion as well as an unexpected portion of credit risk. The historic loan losses used below capture both the expected and unexpected parts of credit risk, whereas a narrow definition of the loan portfolio's credit risk may only address the unexpected portion. Therefore, we further use the standard deviation of historic loan losses to evaluate the impact of banks' loan portfolio concentration on this narrower definition of credit risk. As information to quantify the loan portfolios' (ex ante) credit risk is not easily observable, the literature commonly suggests alternative measures, such as the loan

[1] Historic loan losses are defined as the realized full or partial write-offs and write-downs of loans in the bank's loan portfolio as a percentage of their initial loan amount. In the following, we refer to write-downs comprising both write-offs and write-downs.

loss provision ratio, the non-performing loan ratio or their respective standard deviations.[2] However, these concepts are generally subject to the subjective judgement of the bank management within the framework of accounting standards or models suggested by regulators to estimate and quantify the loan portfolio's credit risk. We therefore regard these variables as biased proxies for the real underlying credit risk of the loan portfolio. In comparison, both the historic loan losses and their standard deviation used in this study are suggested as more comprehensive estimators of the credit risk in general as well as the unexpected portion of it. The historic loan losses are less exposed to the influence of accounting practices and assumptions of credit risk models. Thus, we argue that our database constitutes a considerable improvement in the measurement of the loan portfolio's credit risk and its common risk factors.

Second, by controlling for common risk factors, we separate the bank-specific selection and monitoring abilities from the composition of the loan portfolio. To be more precise, when measuring the performance of an investment fund, the return of the market portfolio is subtracted in order to adjust for the systematic risk. In our case, the loan losses from a hypothetical loan portfolio are subtracted from the bank's actual loan losses. The hypothetical loan portfolio losses are those of a reference portfolio with the same composition as that of the bank, but where the nationwide loss rates are applied.[3] The excess return can then be attributed to the bank's selection and monitoring abilities. As a result, we control for the impact of, for example, industrial concentration risks due to a concentrated loan portfolio and investigate the compensatory benefits through banks' monitoring qualities. By doing this, we not only investigate the impact of loan portfolio concentration on banks' performance measures but analyze the question of whether specialization gains outweigh the foregone diversification benefits.

The main results of our study are: Concentrated banks have, on average, lower loan loss rates after controlling for their portfolio composition. Examining in more detail each individual bank's loan exposure, we find that loss rates of the largest industry-specific loan exposures are, on average, significantly lower than the loss rates of the smallest industry-specific loan exposures. Further, concentrated banks have less unexpected credit risk, as the standard deviation of their loss rates is lower. These findings suggest that specialized German banks acquire considerable selection and monitoring abilities that reduce the loan portfolios' credit risk beyond associated concentration risks.

The paper proceeds as follows. Section 2 discusses the related literature. Section 3 presents the data used in this study. Concentration measures are introduced in Section 4, and Section 5 develops hypotheses as regards the relation between loan portfolio concentration and banks'

[2] See distinction by Liu and Ryan (1995).
[3] In a robustness check, we replace the nationwide loss rates for regional banks with the regional loss rates. This procedure does not alter our results in a significant way; see Section 8.

credit risk. Section 6 describes the empirical methodology, followed by a discussion of results in Section 7. Section 8 contains robustness checks, and Section 9 concludes.

2 Related Literature

Banks face a trade-off between monitoring benefits and concentration risk. Banks with a concentrated loan portfolio are expected to have better monitoring abilities[4], which might lower the loan portfolio's credit risk, while they are confronted with increased credit risk due to industrial concentrations. If the risk-return-profile of a loan were exogenous, i.e. outside the influence of a bank, the banks' credit portfolio risk would be higher for banks with lower diversification in the credit portfolio. However, the loan's risk-return-profile is to some extent endogenous, i.e. it can be influenced by a bank. Due to, for instance, its monitoring activities, it is not per se clear whether diversified banks are less risky than concentrated banks.

Acharya et al. (2006) empirically examine the impact of loan portfolio concentration versus diversification on performance indicators of Italian banks. The authors use the Herfindahl-Hirschman Index (HHI) as a measure of loan portfolio concentration across different industries and sectors.[5] They find that industrial or sectoral diversification implies unaffected or marginally increased return and increased credit risk for banks with a moderate downside risk in the loan portfolio, whereas banks with a high credit risk in their loan portfolio experience decreased bank performance through diversification. The authors conclude that "diversification per se is no guarantee of superior performance or greater bank safety and soundness".[6] Recent single country evidence is provided by Tabak et al. (2011) regarding the impact of loan portfolio concentration on Brazilian banks' return and credit risk, measured by return on assets and non-performing loans over total loans, respectively. The authors perform both static and dynamic regression analyses using traditional concentration as well as distance measures. They find a positive relationship between bank returns and loan portfolio concentration for both the HHI and the Shannon Entropy as well as for an absolute and a relative distance measure.[7] In addition, the HHI and Shannon Entropy have a negative influence on banks' loan portfolio credit risk; in sum, the authors find that concentration has an overall positive effect on banks' performance. The above-mentioned studies suggest a slight positive impact of banks' concentration strategy on performance measures, whereas Rossi et al. (2009) and Bebczuk and Galindo (2007) come to the opposite conclusion examining large commercial Austrian banks and banks from Argentina, respectively.

[4] These abilities have been analyzed in the literature, see, for instance, Boot (2000), Carey et al. (1998) and Sharpe (1990).
[5] Although loan exposures could theoretically be separated into 23 industries, Italian banks have to report the top five exposures only and combine the remaining industries into a sixth category. The HHI is usually defined as a concentration measure with higher values indicating a concentrated loan portfolio, Hirschman (1945), Herfindahl (1950). See also Section 4.
[6] Acharya et al. (2006), p. 1405.
[7] Shannon (1948), Pfingsten and Rudolph (2004). See also Section 4.

Two inherent shortcomings of most studies are addressed by Kamp et al. (2005) following the methodology suggested by Pfingsten and Rudolph (2004). The first deals with concentration measures mostly applied in empirical studies, such as the HHI, which assign perfect naïve diversification to the case that loan exposures to all segments – for example, industries or sectors – are equal. This concept does not account for the diverging importance of different industries or sectors in an economy. Hence, the authors suggest six benchmark indices that measure the distance between a bank's loan portfolio and the loan market portfolio. Behr et al. (2007) adopt this improved toolkit of concentration measures and analyse all German banks to find out whether a concentration strategy is superior to a diversification strategy in terms of the banks' risk-return characteristics. The authors conclude that a concentrated loan portfolio brings a slightly higher return on assets. Measuring the bank's loan portfolio credit risk with the loan-loss provision (LLP) ratio and the non-performing loan (NPL) ratio, respectively, reveals that concentrated banks tend to have lower credit risk. However, when the standard deviation of the loan loss provision ratio and respective non-performing loan ratio, as a more straightforward risk measure of a loan portfolio's unexpected credit risk, is used instead, concentrated banks are more risky than diversified banks.[8] In sum, a bank's concentration strategy seems to reflect the typical risk-return trade-off.

Closely related to our study, Boeve et al. (2010) analyse German cooperative banks and savings banks from 1995 until 2006. Much as in our approach, the authors separate the bank-specific selection and monitoring abilities defined by the actual over expected loss ratio and include average values of bank-specific and other controls as common determinants of the loan portfolio's credit risk. They observe that concentrated banks show, on average, a significantly higher monitoring quality. Comparing concentration benefits with associated concentration risks using a common credit risk model, the authors find strong evidence for cooperative banks that a higher concentration level is, on average, associated with a lower credit risk of the loan portfolio, whereas results for savings banks depend more on the applied diversification measure. In contrast to Boeve et al. (2010) and Behr et al. (2007), we possess a more detailed database comprising the loan exposures and corresponding write-downs on individual bank-, industry- and maturity-specific levels. This enables us to analyse historic loan losses and their standard deviation as unbiased proxies of the loan portfolio's credit risk and to control for common factors of credit risk in a very direct way.

To distinguish the bank- specific selection and monitoring abilities from the composition of the loan portfolio, we follow the basic idea of credit portfolio models (see, for example, Crouhy et al. (2000), Wilson (1998)) addressing systematic and idiosyncratic drivers of loan portfolios' default risk. More precisely, our approach is similar to intensity-based credit portfolio models that apply sector-specific average default rates as systematic factors of credit risk. These common factors are considered by the calculations of the loss rates of a hypothetical loan portfolio which has the same composition as that of the bank, but where the

[8] Behr et al. (2007), p. 14.

nationwide loss rates are applied.[9] The literature has identified several systematic drivers of credit risk. Examining a panel data set of Italian financial intermediaries, Quagliariello (2007) relates loan loss provisions and non-performing loans to GDP growth. Much like a macroeconomic systematic risk factor, a nationwide driver of credit risk reflecting the business cycle is included in this analysis. Beyond the general economic cycle, Aretz and Pope (2013) decompose firms' default risk into common factors such as global, country and industry effects by analysing firms from 24 countries and 30 industries. They find that around 61% of the systematic variance in changes in firms' default risk is due to global and industry effects. Industry-specific effects as well as regional differences are also included in our analysis. By examining the drivers of contract features in revolving bank credit agreements, Dennis et al. (2000) also give a valuable insight into the relationship between maturity and credit risk. The authors extract data on global loan transactions and private placements from TR Dealscan and find that factors which can be related to an increased loan portfolio credit risk, such as less secured loans or a higher earnings variance, imply a shorter revolving loans' maturity. As our data can be broken down into different maturity brackets, the impact of banks' maturity composition on credit risk is likewise considered.

3 Data

We use the borrowers statistics (*Kreditnehmerstatistik*) provided by the Deutsche Bundesbank as our main database. Short-, medium- and long-term loan exposures to the real economy and (since end-2002) the respective changes in the valuation of these loans are reported on quarterly basis by all German banks.[10] Following Deutsche Bundesbank (2009), write-downs and write-ups are defined as "valuation [...] changes caused by individual value adjustments and any write-downs/write-ups of nonperforming debt".[11] Any write-offs are similarly reported as valuation change. Thus, the database contains both write-offs and write-downs.[12] A considerable advantage of the database is that it allows loan exposures and valuation changes to be broken down into different industries (i.e. loans to enterprises, households and non-profit institutions) and maturity brackets. Note that lending to monetary financial institutions (MFIs) and all layers of government are excluded from this database. The borrowers statistics collects valuation changes as net write-downs; for example, a negative value is reported if write-downs exceed write-ups (in this case, the value is used as gross

[9] The corresponding methodology used in this study closely follows the approach applied by Memmel et al. (2012).
[10] Concerning the different maturity brackets, the Bundesbank differentiates between loans up to one year (short-term), between one year and up to five years (medium-term), and more than five years (long-term).
[11] Deutsche Bundesbank (2009), p. 148.
[12] Write-offs and write-downs also impact on the outstanding loan exposure. For example, if the outstanding loan amount is 110 at end-2012 and the exposure has to be written off by 100 in the first quarter of 2013, the corresponding loan exposure reduces to 10 and the valuation change reported amounts to -100. In comparison, any write-ups following previous write-downs will affect the valuation changes only.

write-downs in this study) and a positive value is reported if write-ups exceed write-downs (in this case, we set the value to zero). Thereby, we focus on the historic loan losses only.[13]

Table 1 lists the composition of the database as regards industries and groups of retail borrowers. Besides the group of corporate borrowers, which can be broken down into 23 industries, three subgroups of private borrowers as well as non-profit institutions are examined, which gives 27 "industries" that are relevant for the empirical study.[14]

[13] Similar to the procedure adopted by Memmel et al. (2012).
[14] A further analysis as regards the distribution of borrowers within a specific industry is not possible and left to anecdotal evidence.

Item	Borrowers	Dec. 2003	Dec. 2011
Enterprises		69.4%	71.9%
1	Agriculture, forestry, fishing and aquaculture	1.5%	2.1%
2	Electricity, gas and water supply; refuse disposal, mining and quarrying	3.2%	6.5%
	Manufacturing	10.7%	9.2%
3	Chemical Industry, manufacture of coke and refined petroleum products	0.9%	0.7%
4	Manufacture of rubber and plastic products	0.6%	0.5%
5	Manufacture of other non-metallic mineral products	0.5%	0.3%
6	Manufacture of basic metals and fabricated metal products	1.7%	1.7%
7	Manufacture of machinery and equipment; manufacture of transport equipment; repair and installation of machinery and equipment	2.1%	2.4%
8	Manufacture of computer, electronic and optical products	1.1%	0.9%
9	Manufacture of wood and wood products; manufacture of pulp, paper and paper products, printing; manufacture of furniture	1.9%	1.3%
10	Textiles, apparel and leather goods	0.4%	0.2%
11	Manufacture of food products and beverages; manufacture of tobacco products	1.4%	1.1%
12	Construction	2.9%	2.9%
13	Wholesale and retail trade; repair of motor vehicles and motorcycles	9.5%	7.2%
14	Transportation and storage; post and telecommunications	4.1%	4.0%
15	Financial intermediation (excluding MFIs) and insurance companies services	3.9%	14.6%
	Services (including self-employment)	33.6%	25.5%
16	Housing enterprises	4.8%	4.3%
17	Holding companies	3.3%	3.1%
18	Other real estate activities	9.0%	6.3%
19	Hotels and restaurants	1.2%	0.9%
20	Information and communication; research and development; membership organisations; publishing activities; other business activities	4.7%	4.1%
21	Health and social work (enterprises and self-employment)	3.9%	3.6%
22	Rental and leasing activities	2.0%	0.8%
23	Other service activities	4.5%	2.3%
Private households		29.8%	27.4%
24	Instalment loans (excluding housing loans)	10.7%	11.8%
25	Other loans (excluding housing loans)	6.6%	4.2%
26	Housing loans	12.6%	11.4%
Non-profit institutions			
27	Non-profit institutions	0.8%	0.7%
Total		100.0%	100.0%

Table 1: This table shows the shares of lending to all German borrowers (without MFI and government; long-term mortgage loans also not included), broken down into enterprises, private households and non-profit institutions at end-2003 and end-2011. Long-term mortgage loans are not included.

Whereas the general pattern between the groups marginally changed over time with about 71.9% accounting for loans to enterprises, 27.4% to private households and 0.7% to non-profit institutions at end-2011, some interesting observations within the groups can be made. The percentage share of electricity, gas and water supply, for example, doubled between end-2003 and end-2011 from 3.2% to 6.5%. Also, the share of financial intermediation (without MFIs) and insurance companies services increased from 3.9% at end-2003 to 14.6% at end-2011. Financial vehicle corporations, in particular, have been driving this development since 2009.[15] Please note that mortgage loans, which constitute a major portion of the loans to the real economy, are not included in our analysis. We refrain from including these loans because their loss rate is considerably smaller than that of the other loans.

The prudential information system *(Bankaufsichtliches Informationssystem, BAKIS)*, which is a database provided by the Deutsche Bundesbank and the German Federal Financial Supervisory Authority *(BaFin)* for regular banking supervision, contains – along with other information – balance sheet data on a yearly basis and is used here as a source for bank-specific control variables.[16] To begin with, the natural logarithm of total assets (LN_TA) is included as a common control variable for bank size. Although we do not have strong expectations as regards the sign of the variable, one argument might refer to smaller banks having a higher number of cases relative to larger banks, which could be consistent with increased industry-specific experience and corresponding monitoring benefits. This might especially be the case for regional banks which predominantly serve small and medium-sized enterprises, suggesting a positive relation between bank size and credit risk. Following Rossi et al. (2009), we consider the overall quality of a bank's loan portfolio as, for example, a highly diversified loan portfolio of risky assets may have a different impact on the bank-wide loss rate than a concentrated loan portfolio which focuses on almost risk-free assets.[17] As risk-weighted assets are both a proxy for the size as well as for the risk position of a bank, the risk-weighted assets over total assets ratio (RWA_TA) is included to ensure that the variable captures only the banks' loan quality and is not multicollinear to the banks' total assets. According to banking regulation, as increased risk-taking is supposed to result in higher risk-weighted assets, we expect a positive relationship between risk-weighted assets over total assets and credit risk. Further, we anticipate increased risk to be accompanied by increased returns. We therefore include return-on-assets (ROA) and, alternatively, return-on-equity (ROE) in our model. As a further characteristic influencing the monitoring quality, Behr et al. (2007) suggest the employee ratio (ER) calculated by the average number of employees over total assets as an appropriate proxy, since the build-up of sector-specific knowledge is

[15] See "Special Statistical Publications" (*Statistische Sonderveroeffentlichungen*) with respect to the borrowers statistics available at www.bundesbank.de.
[16] Memmel and Stein (2008).
[17] Rossi et al. (2009), p. 2219.

assumed to be personnel-intensive. We likewise include the ER in our analysis and expect a negative relation with credit risk.[18]

To merge quarterly data taken from the borrowers statistics with annual data from the prudential information system, we calculate moving averages of the former, which is also why our regressions start only at end-2003. *BAKIS* also provides information on mergers and acquisitions. Despite ongoing consolidation trends in the German banking market, the numbers of regional banks, branches of foreign banks, and banks with special functions increased over the observation period.[19] A common procedure for handling mergers is applied. At the time of any type of acquisition or merger, a new (third) bank is constructed, which is why the number of banks in our sample artificially exceeds the number of existing banks.

We summarize savings banks, credit cooperatives as well as regional private commercial banks as regional banks which operate in areas around their headquarters. For these banks, regional effects are included with the help of the postcode system, as there are ten almost equally populated postcode areas in Germany.[20] All other banks, especially the big banks as well as the central institutions of savings banks and credit cooperatives, are expected to operate nationwide. To correct moderately for outliers, observations with bank-wide loss rates above the 99th percentile are deleted from the overall dataset.

4 Measures of Loan Portfolio Concentration

Classical portfolio theory according to Markowitz (1952) suggests optimal portfolio selection through diversification as regards the typical trade-off between portfolio risk and expected return. However, Kamp (2006) points out that data limitations concerning the calculation of risk, return and default correlations hinder the applicability to financial intermediation. More importantly, in contrast to the Markowitz model, banks can influence the expected return of a granted loan by monitoring it closely, which suggests that credit risk is endogenous. As a consequence, banks usually adopt simpler strategies to allocate their loan portfolios. The heuristic of naïve diversification is a possible starting point, assuming that increased diversification reduces credit risk. For the data used in our study, perfect naïve diversification corresponds to equal shares of the loan exposures to 27 industries. If loans are granted to one industry only, the loan portfolio is perfectly concentrated. The Herfindahl-Hirschman Index (HHI) comprises naïve diversification in a concentration measure and has been widely used in

[18] See Appendix I, II and III for a list and definitions of variables used in this study, summary statistics and correlation coefficients, respectively.

[19] At end-2011 (end-2003), there were 4 (5) big banks, 160 (144) regional banks, 85 (61) branches of foreign banks, 10 (14) state banks, 427 (490) savings banks, 2 (2) central institutions of the corporative sector, 1,121 (1385) credit cooperatives, 20 (23) real estate banks, 16 (14) special purpose banks, 23 (27) building societies, and 1 (42) other banks. In sum, there were 1,869 (2207) banks in Germany that are considered in our study.

[20] Germany is divided into ten postcode areas, each of includes roughly the same number of people. These postcode areas are much more similar than, for instance, the 16 Bundesländer. See Memmel et al. (2012), p. 10.

the loan portfolio concentration literature.[21] Another measure of naïve diversification is the Shannon entropy (S).[22] Among other things, it is used to indicate industrial concentration or income inequality.[23] Naïve diversification measures well illustrate the risk concentration on single industries or segments. However, neither measure accounts for possible differences in the importance of several industries and sectors in an economy. Pfingsten and Rudolph (2004) develop distance measures that describe the composition of a loan portfolio by comparing its distance to a benchmark portfolio. In doing so, the prevailing industry structure in an economy is considered. With distance measures, diversification is greatest if the loan portfolio composition equals the composition of the benchmark portfolio. Like naïve diversification measures, distance measures constitute a simple heuristic for banks to allocate the loan portfolio, assuming that increased diversification is accompanied by less credit risk. Regional banks operating around their location might be unable to reproduce a nationwide benchmark. Therefore, a regional distance measure is likewise included in this analysis.

We define

$$X_{i,t,j,k} \tag{1}$$

as the accounting value (in euro) of the loan exposure of bank i at time t to industry j in maturity bracket k. Correspondingly,

$$X_{i,t,j} := \sum_{k=1}^{3} X_{i,t,j,k} \tag{2}$$

denotes the accounting value (in euro) of loan exposure of bank i at time t to industry j where the loans are aggregated over the three maturity brackets. Further,

$$x_{i,t,j} := \frac{X_{i,t,j}}{\sum_{j=1}^{27} X_{i,t,j}} \tag{3}$$

defines the share of loans granted by bank i to industry j at time t. Since the borrowers statistics provides quarterly data, the specialization measures can be calculated on a quarterly basis.

As regards naïve diversification measures, the Herfindahl-Hirschman Index (HHI) is defined by

$$HHI_{i,t} := \sum_{j=1}^{27} x_{i,t,j}^2 \quad . \tag{4}$$

[21] Hirschman (1945), Herfindahl (1950).
[22] Shannon (1948).
[23] Frenken (2007).

The HHI is equal to 1 when all loan exposures are granted to a single industry, and it equals 1/27 when all 27 industries receive the same amount of loans.

The Shannon entropy (S) is given by

$$S_{i,t} := \sum_{j=1}^{27} x_{i,t,j} \cdot ln(\frac{1}{x_{i,t,j}}) \quad . \tag{5}$$

To ensure that all concentration measures point in the same direction, we consider $\hat{S}_{i,t} = -S_{i,t}$.[24] Thus, a loan portfolio exhibits perfect diversification if $\hat{S}_{i,t} = -\ln(27)$ and maximum concentration if $\hat{S}_{i,t}$ is equal to zero.

With respect to distance measures, the absolute distance measures are denoted by

$$D\_BENCHMARK_{i,t} := \max_j\{|x_{i,t,j} - y_{t,j}|\} \tag{6}$$

with $y_{t,j}$ as the share of the respective industry j in the benchmark portfolio at time t. D_NATION takes the Germany-wide market portfolio as the benchmark portfolio. Similarly, D_REGION is defined with $y_{j,t}$ reflecting the benchmark portfolio according to bank i's region, namely one of Germany's ten postcode areas.

5 Hypotheses on Loan Portfolio Concentration and Credit Risk

The following section establishes three hypotheses regarding the relation between loan portfolio concentration and both the expected and unexpected part of credit risk, approximated using the historic loan losses *(Hypothesis 1 and 2)* as well as loan portfolio concentration and the loan portfolio's unexpected part of credit risk *(Hypothesis 3)*.

We examine whether concentrated banks – in the sense of industry concentration – exhibit lower loan losses than banks with a diversified loan portfolio. While allocating their loan portfolio, banks face a trade-off between specialization benefits and associated concentration risks. Concentration risks arise due to the fact that a concentrated loan exposure entails increased default correlations of borrowers within a certain industry in comparison to a more diversified loan portfolio. Specialization benefits depend on various bank characteristics. As one important feature, loan portfolio concentration is expected to improve the selection and monitoring abilities through the build-up of sector-specific knowledge. Specialization in certain industries or sectors is accompanied by the acquisition of experience as regards industry- or sector specific characteristics over time. Alternatively, a bank might decide to expand into new business areas and therefore devote more resources to the build-up of sector-specific knowledge in the cross-sectional dimension. The gain in expert knowledge allows the loan officer to monitor the borrower more efficiently through a better assessment of, for

[24] Following Theil (1972), $x_{i,j,t} \cdot ln\left(\frac{1}{x_{i,j,t}}\right) = 0$ if $x_{i,j,t} = 0$.

example, a deteriorating borrower's business. As a result, the improved monitoring abilities allow the loan officer to realize specialization benefits, i.e. reduce the credit risk of the loan portfolio. Other characteristics associated with specialization benefits are specific to the individual bank management: For example, the ability to score the borrower's individual attributes and the propagation of information into appropriate terms and conditions before the loan agreement is concluded (e.g. collateral agreements) are essential aspects of banks' selection abilities but hardly measurable.[25] Similarly, general management and process quality during the loan period as well as the supervision of loan repayments are of major importance for the monitoring abilities, but can barely be disentangled. This is why there is scarce empirical evidence on the direct relation between loan portfolio concentration and monitoring abilities as well as monitoring abilities and credit risk in banking. Usually, empirical studies assume the above named relations to exist. The study by Boeve et al. (2010) is an initial attempt to separate monitoring abilities for German banks. Initial evidence for savings and cooperative banks confirms that banks with a concentrated loan portfolio have a higher monitoring quality. Further, the authors find that a higher level of specialization reduces credit risk after controlling for monitoring quality. However, results are robust only for the sample of cooperative banks and are more ambiguous for the sample of savings banks.

Benefitting from our unique database, we control for common risk factors and thereby separate the bank-specific selection and monitoring abilities from the composition of the loan portfolio in a very direct way. More precisely, when measuring the performance of an investment fund, the return of the market portfolio is subtracted to adjust for the systematic risk. In our case, the loan losses from a hypothetical loan portfolio are subtracted from the bank's actual loan losses. The hypothetical loan portfolio losses are those of the reference portfolio with the same composition as that of the bank, but where the nationwide loss rates are applied. The excess return can then be attributed to the bank's selection and monitoring abilities.[26] As a consequence, we control – among other common factors – for the impact of industrial concentration risks which arise due to a concentrated loan portfolio, and investigate the compensatory benefits through banks' monitoring qualities.

In sum, we expect that – after controlling for the portfolio composition – concentrated banks can build up industry-specific knowledge and thereby reduce their loan losses, i.e. both the expected and unexpected parts of credit risk. This leads to *Hypothesis 1*.

Hypothesis 1: Banks with a concentrated loan portfolio have, on average, lower loan losses.

Furthermore, we investigate in greater detail the composition of each individual bank's total loan exposure at time t. In particular, we develop the idea that the largest industry-specific loan exposures within the bank's loan portfolio are, on average, accompanied by increased monitoring experience as regards expert knowledge in comparison to the smallest industry-specific loan exposures. Similar to the argumentation concerning bank size and credit risk, the

[25] Hartmann-Wendels et al. (2010), p. 588, 589.
[26] Of course, this assumes that all relevant common factors of credit risk are included in our estimations.

increased experience arises, for example, through a higher number of cases corresponding to the largest industry-specific loan exposures, which might especially be the case for regional banks which predominantly serve the small- and medium- sized enterprises. Another reason why banks devote higher resources and increased monitoring intensity to the largest industry-specific loan exposures relative to the smallest industry-specific loan exposures could be related to corresponding concentration risks originating from substantial exposures. However, especially savings banks and credit cooperatives, in particular, might benefit from information sharing with their central institutions, which would imply that the bank does not increase the monitoring ability by itself. Thus, we expect that the corresponding monitoring benefits – arising either through increased experience or intensified supervision – lower, on average, the respective loan losses. We therefore compare the loss rates of the largest industry-specific loan exposures with the corresponding loss rates of the smallest industry-specific loan exposures. Similar to *Hypothesis 1*, we control for the composition of the loan portfolio and thereby separate the effects of selection and monitoring abilities.

Hypothesis 2: For a given industry, the loss rate in a bank's loan portfolio is lower if the bank has a major exposure to this industry.

The first hypothesis examines the average impact of overall loan portfolio concentration on the banks' credit risk in general, measured by the historic loan loss rates, whereas the second hypothesis addresses the loss rates within a bank's loan portfolio with respect to the largest and smallest industry-specific exposures.

Due to the trade-off between specialization gains and foregone diversification benefits, the relation between loan portfolio concentration and the loan portfolio's unexpected part of credit risk is of further interest. As noted by Acharya et al. (2006), risk is usually defined by unexpected losses and not by expected losses that are already reflected in commonly applied credit risk measures, such as the loan loss provision ratio or non-performing loan ratio.[27] Further, we argue that these concepts are generally subject to, for instance, discretionary decisions by the bank management and therefore constitute less suited proxies for the real underlying credit risk of the loan portfolio. Our database allows us to use the standard deviation of historic loan losses as a less biased proxy for the unexpected portion of credit risk in our study. If specialized banks acquire considerable selection and monitoring benefits – after controlling for the composition of the loan portfolio – such that specialization gains outweigh the associated concentration risks, a concentrated loan portfolio has a lower standard deviation of the loan loss rates, i.e. a lower unexpected part of credit risk. We address this issue in *Hypothesis 3*.

Hypothesis 3: For concentrated banks, the standard deviation of the loan loss rate is lower than for diversified banks.

[27] Another possible measure of unexpected credit risk is given by the concept of Value at Risk.

The empirical methodology and evidence for these three hypotheses is provided in the next two sections, whereas several robustness checks are conducted in section 8.

6 Empirical Methodology

The empirical analysis investigates the impact of loan portfolio concentration on banks' credit risk, approximated using the historic loan losses, after controlling for the composition of the loan portfolio and further bank-specific control variables. The bank-wide loss rate serves as a dependent variable in our study and is based on data from the borrowers statistics. The notations (7) - (14) and (20) - (28) follow Memmel et al. (2012).[28]

To start with, $C_{i,t,j,k}$ denotes the change in value (in euro) of bank i's loans in maturity bracket k and industry j from $t-1$ to t. This leads to the bank-wide change in value, defined as

$$C_{i,t} := \sum_{j=1}^{27} \sum_{k=1}^{3} C_{i,t,j,k} \quad . \tag{7}$$

Further, the bank-wide loan exposure is correspondingly denoted as

$$X_{i,t} := \sum_{j=1}^{27} \sum_{k=1}^{3} X_{i,t,j,k} \quad . \tag{8}$$

The bank-wide loan loss rate is then calculated on a quarterly basis as a moving average

$$q_{i,t} := \frac{4 \cdot \sum_{m=0}^{3} C_{i,t-m}}{0.5 \cdot X_{i,t} + X_{i,t-1} + X_{i,t-2} + X_{i,t-3} + 0.5 \cdot X_{i,t-4}} \quad .[29] \tag{9}$$

To be clear, we merge quarterly data taken from the borrowers statistics with annual data from the prudential information system in the empirical analysis by using the annual loss rates for the 2003-2011 period. The bank-wide loss rates are presented in Table 2 for the whole sample and both nationwide and regional banks. The whole sample of 13,605 observations shows a median loss rate at 1.01% over our observation period from 2003 to 2011. For the subsample of regional banks, the distribution is also broken down into size quintiles according to the loan portfolio exposure. Interestingly, regional banks in the 4th size quintile show the highest loss rates, whereas banks with the smallest loan portfolios, which are supposed to be less able to diversify credit risk, show rather lower loss rates.[30]

[28] Memmel et al. (2012), pp. 11, 12 and their Appendix II.

[29] Note that $q_{i,t}$ is the weighted average of the loss rates in the current and three previous quarters, where the weights are the portfolio size in a quarter. We calculate the average in this way to avoid extreme values that can otherwise arise when a bank's credit portfolio has nearly completely shrunk.

[30] As we do not further restrict our sample, the number of observations might differ between size quintiles. See Appendix IV for the average bank-wide loss rate over time.

To separate the bank-specific selection and monitoring abilities from the composition of the loan portfolio, common factors of credit risk, i.e. a nationwide, industry and maturity-specific factor, are included in the empirical analysis. These common factors are calculated by a hypothetical loan portfolio that has the same composition as that of the bank but where nationwide loss rates are applied. Further, a regional factor is likewise considered.

Banks	Number of observations	Bank-wide loss rate (in %) Percentile			
		99th	95th	90th	Median
Nationwide	387	4.82	2.77	2.11	0.55
Regional	13,218	5.05	3.54	2.75	1.03
1st size quintile	2,580	4.93	3.43	2.57	0.83
2nd	2,624	4.86	3.35	2.58	0.94
3rd	2,639	5.14	3.66	2.84	1.09
4th	2,689	5.17	3.74	3.00	1.15
5th	2,686	5.16	3.50	2.71	1.08
Whole sample	13,605	5.04	3.53	2.73	1.01

Table 2: This table shows the bank-wide loss rates (per annum, without MFI and government; long-term mortgage loans also not included) for the period 2003-2011. For regional banks, the bank-wide loss rates are broken down into loan portfolio size quintiles. The 1st quintile corresponds to the smallest and the 5th quintile matches the largest size quintile, respectively.

The loss rate of the hypothetical loan portfolio is calculated on a quarterly basis as

$$hq_{i,t}^{ind\times mat} = \sum_{j=1}^{27}\sum_{k=1}^{3} w_{i,t,j,k} \cdot Q_{t,j,k}^{ind\times mat} \tag{10}$$

with $w_{i,t,j,k}$ as the weight according to the industry- and maturity-specific share of loans made by bank i at time t with respect to the whole loan exposure and $Q_{t,j,k}^{ind\times mat}$ as time-, industry- and maturity-specific nationwide loss rate. See Appendix V for a detailed definition.

The loss rate of the hypothetical loan portfolio can be decomposed into the following common factors:

$$hq_{i,t}^{ind\times mat} = Q_t + \Delta hq_{i,t}^{ind} + \Delta hq_{i,t}^{mat} \tag{11}$$

with Q_t as nationwide loss rate of the entire loan portfolio at time t and

$$\Delta hq_{i,t}^{ind} \equiv hq_{i,t}^{ind} - Q_t \tag{12}$$

as well as

$$\Delta hq_{i,t}^{mat} \equiv hq_{i,t}^{ind \times mat} - hq_{i,t}^{ind}. \tag{13}$$

Q_t is a nationwide factor reflecting the phase of the overall credit cycle. The expression $\Delta hq_{i,t}^{ind}$ provides the industry factor, i.e. the differences in the loss rate of the hypothetical portfolio that are because of bank i's deviations in the industry composition. The maturity factor $\Delta hq_{i,t}^{mat}$ is similarly calculated to account for the maturity structure.[31] Thus, the hypothetical loan portfolio loss rates are those of a portfolio where the bank's portfolio weights and the nationwide loss rates are applied. As a result, beyond a nationwide factor, industry-specific effects and the impact of the loan portfolio's maturity composition can be considered.

For regional banks, differences in the loan portfolios' loss rates due to regional effects are similarly included. The region-specific loss rate is defined as

$$Q_{R,t}^{reg} := \frac{\sum_{bank\ i\ is\ in\ region\ R} C_{i,t}}{\sum_{bank\ i\ is\ in\ region\ R} X_{i,t}} \tag{14}$$

with R denoting the ten postcode areas. Thus, the variable $\Delta Q_{R(i),t}^{reg}$ denotes the difference between the loss rate in bank i's postal code area and the nationwide loss rate, and is set to zero for nationwide banks.

The variable $SM_{i,t}$ denotes the inclusion of a specialization measure, the HHI, $\hat{S}$ or D_BENCHMARK, respectively. In addition, banks' total assets (LN_TA), risk-weighted assets over total assets (RWA_TA), return-on-assets (ROA) or, alternatively, return-on-equity (ROE) as well as the employee ratio (ER) are included as bank-specific control variables, denoted by $X_{a,i,t}$ $(a = 6, \ldots, 9)$. These bank-specific controls taken from the prudential information system are available on a yearly basis only, which restricts our analysis to $t = 4, 8, 12, .., T$ with $T = 36$.

This leads us to the following regression for the empirical analysis of *Hypothesis 1*:

$$q_{i,t} = \beta_0 + \beta_1 Q_t + \beta_2 \Delta hq_{i,t}^{ind} + \beta_3 \Delta hq_{i,t}^{mat} + \beta_4 \Delta Q_{R(i),t}^{reg} + \beta_5 SM_{i,t} + \sum_{a=6}^{9} \beta_a X_{a,i,t} + \varepsilon_{i,t} \tag{15}$$

with the dependent variable $q_{i,t}$ as bank-wide yearly loss rate at time t regressed on loan portfolio concentration (β_5) while controlling for the composition of the loan portfolio (β_1 to β_4) and including further bank-specific control variables (β_6 to β_9). The coefficients β_0 to β_9 are the parameters to be estimated.

[31] See Appendix V for detailed definitions. Memmel et al. (2012) show that a change in the sequence of decomposition, i.e. at first the maturity, then the industry, makes only little difference.

The empirical investigation of *Hypothesis 2* proceeds as follows: To analyse the individual bank's loan exposure in more detail, we investigate the bank's largest and smallest industry-specific loan exposures. Therefore, we divide a bank's loan portfolio at time t into two subportfolios. The first subportfolio includes the bank's largest industry-specific loan exposures (L) with the in descending order aggregated share of $\leq 50\%$. The remaining industry-specific loan exposures (S) are included in the second loan portfolio.[32] As a result, within a bank's loan portfolio, industry-specific loan exposures contribute to (either) the upper (or lower) half of the portfolio size distribution. For simplicity, these two subportfolio groups are called the largest and smallest industry-specific loan exposures below.

For each of the two subportfolios, the actual and the hypothetical loss rates are considered, i.e. the loss rate based on the bank's portfolio weights and the nationwide loss rate of the subportfolio. The corresponding variable of interest is calculated both for each year separately ($\widehat{\Delta_t}$) and the overall sample ($\widehat{\Delta}$). Accordingly, we define

$$\widehat{\Delta_t} = \sum_{i=1}^{N} \left(r_{i,t}^{L} - r_{i,t}^{S} \right) \tag{16}$$

for each year $t = 1, .., T$ with $T = 9$. For the overall sample, $\widehat{\Delta}$ for N banks and $T = 9$ years is denoted by

$$\widehat{\Delta} = \sum_{i=1}^{N} \sum_{t=1}^{9} \left(r_{i,t}^{L} - r_{i,t}^{S} \right) \tag{17}$$

with

$$r_{i,t}^{L} = \frac{\sum_{j \epsilon L_{i,t}} C_{i,t,j}}{\sum_{j \epsilon L_{i,t}} X_{i,t,j}} - \frac{\sum_{j \epsilon L_{i,t}} Q_{t,j}^{ind} \cdot X_{i,t,j}}{\sum_{j \epsilon L_{i,t}} X_{i,t,j}} \tag{18}$$

and corresponding definition for $r_{i,t}^{S}$. See (23) and (27) in Appendix V for detailed definitions. By subtracting the hypothetical from the actual loss rates in (18) we control for the composition of the loan portfolio. Thus, $\widehat{\Delta_t}$ and $\widehat{\Delta}$ account for adjustments in different industries. According to our expectations formulated in *Hypothesis 2*, we expect $\widehat{\Delta_t}$ and $\widehat{\Delta}$ to be statistically significant and negative.

Finally, to empirically examine *Hypothesis 3*, we use the standard deviation of loan losses as proxy for the unexpected part of credit risk and apply the following cross-sectional regression:

[32] The division into two parts might seem arbitrary, but an inclusion of, say, a third group would imply loss of information. We thus use all information of the bank's loan portfolio size distribution for the analysis.

$$\hat{\sigma}_{lr}^{i} = \beta_0 + \beta_1 \hat{\sigma}_{cf}^{i} + \beta_2 \overline{SM_i} + \sum_{a=3}^{6} \beta_a \bar{X}_{a,i} + \varepsilon_i \qquad (19)$$

with $\hat{\sigma}_{lr}^{i}$ as the standard deviation of the bank-wide loss rates calculated for each bank over time. To control for the volatility of the reference loan portfolio with the same composition as that of the bank, we conduct a separate robust fixed effects regression including the loss rates as dependent and the common risk factors as independent variables. From this regression, we calculate the standard deviation of the predicted dependent variable for each bank which, in turn, serves as explanatory variable $\hat{\sigma}_{cf}^{i}$ in (19). The estimated coefficient of β_1 should therefore be close to one. The $\overline{SM_i}$ denote the serial average of the respective concentration measure for bank i and $\sum_{a=3}^{6} \bar{X}_{a,i}$ reflect the serial average of the bank-specific control variables.

7 Results

Evidence for the hypotheses derived above are presented for all banks as well as separately for nationwide and regional banks. Table 3 shows results for *Hypothesis 1*, focusing on the impact of the concentration measure HHI on the banks' credit risk in general, i.e. the historic loan losses.

For all banks, the model explains about 6% of the serial variation and about 10% of the cross-sectional variation in the bank-wide loss rate. The explanatory power varies approximately in the same range for regional banks as these constitute the majority in the dataset, whereas the R-squared within (R-squared between) is estimated at about 16% (20%) for nationwide banks, respectively. The common factors of the hypothetical portfolio, namely the nationwide (Q_T), industry (Q_TI), maturity (Q_TM) as well as the regional factor (Q_TR) are all positive and significant at least at the 5% level. We can, for example, interpret a 1 bp increase in the nationwide factor *ceteris paribus* to imply a 0.7 bp increase in the bank-wide loss rate for all banks, whereas a 1 bp increase in the regional factor *ceteris paribus* leads on average to a 0.1 bp increase in the bank-wide loss rate. The economic significance of the common factors is smaller for regional banks than it is for nationwide banks. In turn, one might expect a higher economic influence of the regional factor which is, however, rather low. One explanation for this is that regional banks are well diversified and closely map the nationwide benchmark portfolio; for example, regional banks have a low HHI (D_REGION) of 0.1422 (0.1683) at end-2011.[33]

Our first hypothesis claims that banks with a concentrated loan portfolio have, on average, lower loan losses than more diversified banks through the build-up of sector-specific knowledge. By including the common risk factors we separate the sources of credit risk (e.g. concentration risks due to a concentrated loan portfolio) from the bank-specific selection and

[33] For diversification trends in the German banking market over time, see Kamp et al. (2005).

monitoring abilities and thereby attribute the excess return to the latter. The corresponding variables of central interest are the concentration measures which show, in the case of the HHI, significance at the 1% level with negative sign for all banks as well as for nationwide and regional banks, respectively. This suggests that higher loan portfolio concentration - after controlling for the portfolio composition - implies, on average, a lower loss rate, which confirms our above-stated hypothesis and earlier results by Behr et al. (2007). The authors find that a concentrated loan portfolio implies, on average, a lower loan loss provision ratio and non-performing loan ratio. However, these proxies of the loan portfolio's credit risk are, in our view, biased measures for the real underlying credit risk of the loan portfolio. The effect of the concentration is also not negligible from an economic point of view: If the HHI in the credit portfolio increases by one standard deviation (0.136), then the portfolio loss rate goes down, on average, by 15 basis points. This reduction seems considerable, given the median loss rate of 1.01% p.a. However, due to the pronounced skewness of the distribution, the reduction of 15 basis points is relatively small when we look at the higher percentiles (See Table 2).

	Bank-wide loss rate		
Variables	All banks	Nationwide	Regional
Q_T	0.709***	0.838***	0.702***
	(0.043)	(0.240)	(0.044)
Q_TI	0.621***	0.802***	0.610***
	(0.082)	(0.2559)	(0.087)
Q_TM	0.619***	0.900**	0.600***
	(0.110)	(0.384)	(0.118)
Q_TR	0.104**		0.103**
	(0.042)		(0.042)
HHI	-0.011***	-0.009*	-0.011***
	(0.003)	(0.005)	(0.004)
LN_TA	0.002*	0.002	0.002
	(0.001)	(0.002)	(0.001)
RWA_TA	-0.007***	-0.007*	-0.008***
	(0.002)	(0.004)	(0.002)
ROA	-0.068**	0.053*	-0.109***
	(0.030)	(0.031)	(0.026)
ER	0.010***	0.011	0.010***
	(0.004)	(0.007)	(0.004)
Constant	-0.005	-0.020	-0.003
	(0.007)	(0.023)	(0.008)
R-squared (within)	6.1%	15.9%	6.1%
R-squared (between)	9.6%	19.6%	10.6%
Number of Obs.	13,605	387	13,218
Number of Groups	2,077	91	1,986

Table 3: This table shows regression results from a standard fixed effects estimation equation with robust standard errors. The dependent variable is the bank-wide loss rate. The right-hand side of the regression equation is based on a specialization measure (HHI), common risk factors and various bank-specific control variables, see Appendix I. Yearly data is used. ***,**,* denote statistical significance at the 1, 5 and 10 percent levels, respectively. Standard errors in parenthesis.

For the remaining bank-specific control variables, we can draw the following conclusions: The natural logarithm of total assets (LN_TA) as a control variable for bank size is significant with a positive sign for all banks, supporting our anticipation of a positive relationship. Further, the banks' loan portfolio constitutes, on average, a considerably higher share in the balance sheet total for smaller banks relative to larger banks.[34] This is why monitoring qualities are highly important for smaller banks, whereas larger banks might devote their resources to other parts of the balance sheet as well. However, the coefficient becomes insignificant for the subsamples, which is not surprising as the subsamples are rather homogeneous groups and imply little variation. This result can also be an indication that the effect of size is overlaid by the effect of the difference between nationwide and regional banks' lending behavior. In contrast to an expected positive relationship between the ratio of risk-weighted assets over total assets (RWA_TA) and loan losses, the regression results show a different picture as the estimated beta coefficient is negative. This result may be related to the argument that risk-weighted assets systematically overestimate credit risk. In that case, the level of the risk-weighted assets steadily suggests a higher risk than the actual credit risk in the loan portfolio. Interestingly, the return-on-assets variable (ROA) shows a positive sign for nationwide banks and a negative sign for regional banks. Thus, for nationwide banks increased return seems to be accompanied by increased credit risk, whereas for regional banks a reverse relationship turns out to be the case. These indications of inefficiencies in the German banking market, i.e. a negative relation between a bank's return and risk, have also been empirically revealed by Hayden et al. (2007), for example. In contrast, we do not find a statistical significant influence of the return-on-equity (ROE) ratio on loan losses, replacing the ROA in a regression not reported here. As regards the employee ratio (ER), we find a positive relationship with the bank-wide loss rate for regional banks and no impact as regards nationwide banks. This implies that, for regional banks, a higher number of employees relative to the bank's total assets is accompanied by a higher loss rate on average. In comparison to the results by Behr et al. (2007), who include a similar set of bank-specific control variables and find a negative relationship between the ER and loan loss provision- and non-performing loan ratios, this result is somewhat surprising. The considerable improvement as regards our database in the measurement of the loan portfolio's credit risk might be a reason for the different results. This reveals that a higher ER *per se* does not reflect an increased build-up of sector-specific knowledge with corresponding specialization benefits. Further, Appendix III shows that the ER and LN_TA are negatively correlated. Thus, especially for regional banks, a higher level of LN_TA does not proportionally induce a higher ER but, for example, a small number of persons in charge who have a leading role in making decisions. This observation likewise argues against the ER as an appropriate proxy for monitoring quality. In the robustness section, we consider similar variables related to the banks' employees situation.

[34] Deutsche Bundesbank (2013).

Similar to Table 3, regression results reported in Table 4 show evidence as regards *Hypothesis 1*, which analyses the impact of loan portfolio concentration on credit risk in general, but where different concentration measures are applied. We find that all considered concentration measures negatively affect the bank-wide loss rate on average. This confirms *Hypothesis 1* not only for naïve diversification measures, but also for distance measures. This outcome differs from earlier findings by Behr et al. (2007), who find changing signs for specialization measures affecting the loan loss provision ratio, and confirms the robustness of our results. The economic impact of the naïve concentration measure HHI and absolute distance measure D_NATION with the nationwide loan portfolio as a benchmark are approximately equal in magnitude, whereas the economic impact of the Shannon entropy $\hat{S}$ and the regional distance measure D_REGION is lower. For the latter, the estimated beta coefficient is, in fact, not significant.

	Bank-wide loss rate			
Variables	All banks			Regional Banks
Q_T	0.709***	0.716***	0.691***	0.694***
	(0.043)	(0.043)	(0.044)	(0.045)
Q_TI	0.621***	0.619***	0.620***	0.642***
	(0.082)	(0.082)	(0.082)	(0.087)
Q_TM	0.619***	0.607***	0.618***	0.614***
	(0.110)	(0.111)	(0.110)	(0.118)
Q_TR	0.104**	0.102**	0.104**	0.104**
	(0.042)	(0.042)	(0.042)	(0.042)
LN_TA	0.002*	0.002*	0.002*	0.002
	(0.001)	(0.001)	(0.001)	(0.001)
RWA_TA	-0.007***	-0.007***	-0.008***	-0.008***
	(0.002)	(0.002)	(0.002)	(0.002)
ROA	-0.068**	-0.068**	-0.068**	-0.112***
	(0.030)	(0.030)	(0.030)	(0.026)
ER	0.010***	0.010***	0.010***	0.011***
	(0.004)	(0.004)	(0.004)	(0.004)
HHI	-0.011***			
	(0.003)			
$\hat{S}$		-0.003***		
		(0.001)		
D_NATION			-0.009***	
			(0.002)	
D_REGION				-0.004
				(0.003)
Constant	-0.005	-0.014*	-0.005	-0.004
	(0.007)	(0.008)	(0.008)	(0.008)
R-squared (within)	6.1%	6.1%	6.2%	6.0%
R-squared (between)	9.6%	10.4%	9.3%	10.6%
Number of Obs	13,605	13,605	13,605	13,218
Number of Groups	2,077	2,077	2,077	1,986

Table 4: This table shows regression results from a standard fixed effects estimation equation with robust standard errors. The dependent variable is the bank-wide loss rate. The right-hand side of the regression equation is based on specialization measures (HHI, $\hat{S}$, D_NATION, D_REGION), common risk factors and various bank-specific control variables, see Appendix I. Yearly data is used. ***,**,* denote statistical significance at the 1, 5 and 10 percent levels, respectively. Standard errors in parenthesis.

Hypothesis 2 focuses on the composition of each individual bank's total loan exposure. We raise the idea that the largest industry-specific loan exposures are accompanied by increased monitoring experience or monitoring intensity by the loan officer and therefore show lower loss rates on average. The empirical procedure described in the previous section leads to a comparison of the loss rates of the largest industry-specific loan exposures and the smallest industry-specific loan exposures. This can be examined by a simple t-test of $\widehat{\Delta}$ for the overall sample and $\widehat{\Delta_t}$ for a t-test in each year. See Table 5 for the results of the t-test statistics.

One sample t-test	Variable: $\widehat{\Delta}$			
	Nobs	Mean	SE	t value
Overall Sample	13,605	-0.0085	0.000364	-23.4818
By year	Variable: $\widehat{\Delta_t}$			
2003	1,529	-0.0044	0.001495	-2.9160
2004	1,625	-0.0076	0.000982	-7.7141
2005	1,582	-0.0101	0.001108	-9.0731
2006	1,578	-0.0109	0.001106	-9.8708
2007	1,600	-0.0083	0.000906	-9.1167
2008	1,571	-0.0084	0.000882	-9.4857
2009	1,587	-0.0083	0.000909	-9.1769
2010	1,561	-0.0102	0.001103	-9.2730
2011	972	-0.0087	0.001248	-6.9794

Table 5: This table shows results of one sample t-tests for the variable delta hat according to *Hypothesis 2* (see section 6). Results are displayed for the overall sample and by year, period 2003-2011.

The test statistic for the overall sample reveals that the delta hat value is statistically significantly smaller than zero. Moreover, the result holds in the cross-section as the corresponding mean is statistically significantly smaller than zero for every year considered. These results confirm that banks gain increased knowledge and experience as regards their largest industry-specific loan exposures and thereby reduce the associated loan losses through corresponding monitoring benefits.

Table 6 presents the average loss rates for the industry-specific loan exposures in the case that these belong to the largest (smallest) subportfolio group. The corresponding results reveal an important finding that further confirms *Hypothesis 2*: For example, in the case that construction belongs to the banks' largest subportfolio group, the loan loss rate is, on average, 2.14%, whereas in the case that construction belongs to the banks' smallest subportfolio group, the loan loss rate is, on average, 2.39%. In general, the average loss rates of the largest industry-specific loan exposures are lower than the average loss rates of the smallest industry-specific loan exposures (with the sole exception of wholesale/retail trade and textiles). Our analysis cannot quantify the optimal trade-off between a concentrated vs. diversified loan portfolio, as we investigate the relationships around the mean of the distribution. However, as regards the considered 27 industries/sectors and within our observation period, banks could, on average, achieve the greatest specialization benefits by focusing on the chemical industry, c.f. the average loss rate if the sector is important for a bank, and otherwise.[35]

Results displayed in Table 5 confirm *Hypothesis 2* both for the overall sample as well as for the cross-section, whereas results provided in Table 6 endorse *Hypothesis 2* for the overall sample period, but broken down into every industry considered in this study.

[35] The five industries with the greatest specialization benefits (in descending order) are: 1. Chemical industry, 2. Holding companies, 3. Manufacture of rubber and plastic products, 4. Rental and leasing activities and 5. Financial intermediation (without MFIs).

Item	Borrower	Average loss rate (in %) If the sector is important for a bank	otherwise
	Enterprises		
1	Agriculture, forestry, fishing and aquaculture	0.53	1.18
2	Electricity, gas and water supply; refuse disposal, mining and quarrying	0.34	1.31
	Manufacturing		
3	Chemical Industry, manufacture of coke and refined petroleum products	0.00	3.72
4	Manufacture of rubber and plastic products	1.31	3.75
5	Manufacture of other non-metallic mineral products	2.14	2.56
6	Manufacture of basic metals and fabricated metal products	1.98	2.50
7	Manufacture of machinery and equipment; manufacture of transport equipment; repair and installation of machinery and equipment	1.62	2.75
8	Manufacture of computer, electronic and optical products	0.87	2.38
9	Manufacture of wood and wood products; manufacture of pulp, paper and paper products, printing; manufacture of furniture	2.16	2.42
10	Textiles, apparel and leather goods	3.62	3.31
11	Manufacture of food products and beverages; manufacture of tobacco products	1.70	2.34
12	Construction	2.14	2.39
13	Wholesale and retail trade; repair of motor vehicles and motorcycles	2.06	1.99
14	Transportation and storage; post and telecommunications	0.91	1.98
15	Financial intermediation (excluding MFIs) and insurance companies services	0.28	1.85
	Services (including self-employment)		
16	Housing enterprises	1.09	2.39
17	Holding companies	0.74	3.25
18	Other real estate activities	1.35	2.29
19	Hotels and restaurants	2.58	2.77
20	Information and communication; research and development; membership organisations; publishing activities; other business activities	1.37	1.77
21	Health and social work (enterprises and self-employment)	0.77	0.99
22	Rental and leasing activities	0.29	2.17
23	Other service activities	1.34	2.12
	Private households		
24	Instalment loans (excluding housing loans)	0.68	1.18
25	Other loans (excluding housing loans)	1.75	2.05
26	Housing loans	0.54	0.99
	Non-profit institutions		
27	Non-profit institutions	0.17	0.92

Table 6: This table shows the average loss rate (in %) if the sector is important for a bank and otherwise, according to above named procedure, period 2003-2011.

	SD (Bank-wide loss rate)		
Variables	All banks	Nationwide	Regional
SD_YHAT	1.1999***	2.3782***	1.1143***
	(0.147)	(0.474)	(0.152)
$\overline{\text{HHI}}$	-0.0034***	0.0004	-0.0015
	(0.001)	(0.002)	(0.001)
$\overline{\text{LN\_TA}}$	-0.0003***	0.0012**	-0.0003***
	(0.000)	(0.001)	(0.000)
$\overline{\text{RWA\_TA}}$	-0.0032***	0.0067	-0.0040***
	(0.001)	(0.005)	(0.001)
$\overline{\text{ROA}}$	-0.2641***	-0.1531	-0.2636***
	(0.057)	(0.302)	(0.058)
$\overline{\text{ER}}$	0.0020	0.0009	0.0021
	(0.002)	(0.009)	(0.002)
$\overline{\text{CONSTANT}}$	0.0093***	-0.0121*	0.0099***
	(0.001)	(0.006)	(0.001)
R-squared (between)	13.0%	74.4%	10.6%
Number of Groups	933	31	902

Table 7: This table shows regression results from a cross-sectional regression. The dependent variable is the standard deviation of the bank-wide loss rate. The right-hand side of the regression equation is based on the standard deviation of the hypothetical portfolios loss rates (SD_YHAT), the average values of the HHI and bank-specific control variables, see Appendix I. ***,**,* denote statistical significance at the 1, 5 and 10 percent levels, respectively. Standard errors in parenthesis.

Results displayed in Table 7 reveal evidence for *Hypothesis 3*, namely the impact of loan portfolio concentration on the unexpected part of credit risk, measured by the standard deviation of the bank-wide loss rates. Therefore, we consider only banks which have nine observations per bank and then conduct cross-sectional regressions. The variable SD_YHAT denotes the standard deviation of the common systematic risk factors and thereby controls for the volatility of the reference loan portfolio with the same composition as that of the bank.[36] For the samples of all banks and the regional banks, this coefficient is not statistically different from one, the theoretically expected value. A one standard deviation increase in the systematic risk factors, for example, implies an increase of about 1.2 in the standard deviation of the loan losses. According to *Hypothesis 3*, we find evidence for the overall sample that banks with a concentrated loan portfolio – after controlling for the portfolio composition – have a lower unexpected part of credit risk than diversified banks, as their standard deviation of the loan losses is lower. Results displayed in Appendix VI confirm this evidence also for other specialization measures. However, as Table 7 shows, the beta coefficient becomes insignificant for the subsamples of regional and nationwide banks. This result is different to earlier findings, for instance, Behr et al. (2007) who observe increased loan portfolio's credit risk for concentrated banks measured by the standard deviation of LLP ratio and the standard deviation of the NPL ratio. Again, we argue that the considerable improvement as regards our

[36] See Section 6.

database and the inclusion of common risk factors to control for the composition of the loan portfolio might constitute two important reasons for the different results.

8 Robustness Checks

As a first robustness check, we investigate in more detail the economic impact of the common risk factors on the bank-wide loss rates for regional banks. We therefore separate the corresponding subgroup into size quintiles. Regression results are reported in Table 8. We observe that the bank-wide loss rate is much more sensitive to all common risk factors (with the sole exception of the maturity factor) for the largest 20 percent regional banks (e.g. a nationwide factor of 0.882) than for the 20 percent smallest regional banks (e.g. a nationwide factor of 0.245). The economic impact of the estimated beta coefficients increases with the size quintile and shows an abrupt rise from the 2nd size quintile on. The regional factor is statistically relevant only for the 20 percent largest banks and small in economic magnitude. This result is in line with results reported in Table 3, i.e. that the regional factor is only of minor importance in driving the bank-wide loss rate, and according to Table 8, statistically significant only for the 20 percent largest regional banks.

	Bank-wide loss rate					
	All	Size quintile				
Variables	regional banks	1st	2nd	3rd	4th	5th
Q_T	0.683***	0.245***	0.732***	0.747***	0.682***	0.882***
	(0.042)	(0.094)	(0.092)	(0.106)	(0.090)	(0.078)
Q_TI	0.605***	0.171	0.583***	0.806***	0.819***	0.577***
	(0.087)	(0.164)	(0.175)	(0.214)	(0.182)	(0.219)
Q_TM	0.592***	0.853***	0.222	0.534*	0.566**	0.789***
	(0.118)	(0.223)	(0.256)	(0.278)	(0.252)	(0.270)
Q_TR	0.104**	-0.061	0.006	0.107	0.142	0.344***
	(0.042)	(0.074)	(0.086)	(0.102)	(0.103)	(0.091)
HHI	-0.012***	-0.009	-0.002	-0.008	-0.015***	-0.025***
	(0.004)	(0.006)	(0.008)	(0.010)	(0.006)	(0.009)
RWA_TA	-0.009***	-0.012***	-0.012***	-0.014***	-0.004	-0.005
	(0.002)	(0.004)	(0.005)	(0.004)	(0.004)	(0.003)
ROA	-0.107***	-0.031	-0.124**	-0.121***	-0.351***	-0.104*
	(0.026)	(0.025)	(0.052)	(0.034)	(0.094)	(0.059)
ER	0.009***	0.017***	0.009	0.002	0.039***	0.008
	(0.003)	(0.004)	(0.008)	(0.002)	(0.011)	(0.006)
Constant	0.008***	0.011***	0.007***	0.012***	0.000	0.006***
	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
R-squared (within)	6.1%	3.4%	5.5%	6.1%	9.6%	11.5%
R-squared (between)	8.7%	2.5%	8.3%	1.1%	13.7%	8.1%
Number of Obs	13,218	2,580	2,624	2,639	2,689	2,686
Number of Groups	1,986	415	386	396	408	381

Table 8: This table shows regression results from a standard fixed effects estimation with robust standard errors. The dependent variable is the bank-wide loss rate. The right-hand side of the regression equation is based on a specialization measure (HHI) and various bank-specific control variables, see Appendix I. Yearly data is used. ***,**,* denote statistical significance at the 1, 5 and 10 percent levels, respectively. Standard errors in parenthesis. Note that results for all regional banks differ from results reported in Table 3 since LN_TA is excluded here. The 1st quintile corresponds to the smallest and the 5th quintile matches the largest size quintile, respectively.

In addition, as regards the 2008/2009 financial crisis period, the average bank-wide yearly loss rate over time did not show a similar increase compared to the 2003 recession (see Appendix IV). Whereas the average bank-wide loss rate for regional banks stood at about 1.2% in 2009 (compared to 1.6% in 2003), nationwide banks featured an average bank-wide loss rate at about 0.7% in 2009 (compared to 1.1% in 2003). This is mainly due to the fact that the German real economy exhibited only a short and moderate slump in growth and the majority of banks included in our data set were adversely affected only through second-round effects. According to the empirical analysis of our hypotheses, results hold throughout the 2008/2009 financial crisis period and also for a 2008-2010 subsample.[37]

In another robustness check, we investigate evidence as regards other monitoring measures beyond specialization indices. First, although Behr et al. (2007) suggest the employee ratio to be an appropriate proxy for monitoring quality as the build-up of sector-specific knowledge is

[37]A financial crisis dummy included in our regressions turns out to be insignificant. Corresponding regression results are available on request.

expected to be personnel-intensive, we do not find a corresponding negative relationship between the employee ratio and the bank-wide loss rate. This might be due to the fact that a high employee ratio *per se* does not reflect improved selection and monitoring abilities, say, if a bank employs a high number of low-skilled officers for day-to-day business. Instead, highly qualified employees that are able to efficiently select and monitor a bank's loan portfolio are expected to be cost-intensive. Thus, both labour costs and the ratio of labour costs over the total number of employees are included separately as further proxies for monitoring quality in our analysis. Further, we consider the idea that a relatively large portion of variable compensation in comparison to fixed compensation reflect both an incentive-oriented compensation scheme and high-skilled employees, which might be associated with increased monitoring abilities. Therefore, we examine the impact of the standard deviation of labour costs on the bank-wide loss rate. We do not find, however, noteworthy economic or statistical significance for the included variables. Our methodological procedure to separate the effects owing to the loan portfolio composition from a bank's selection and monitoring abilities leaves us with individual fixed effects resulting from the regression described in (15). These individual fixed effects can be attributed to the effectiveness of the bank. Relating them to a commonly applied efficiency measure, i.e. the cost-to-income ratio, reveals no statistically significant correlation and calls the usefulness of the proxy into question.[38] In contrast, we find a statistically significant and negative correlation of labour costs over the total number of employees ratio with the individual fixed effects and a statistically significant and positive correlation with the employee ratio. This might further support our finding that the employee ratio *per se* is a less useful proxy for the build-up of industry-specific knowledge and that more efficient banks tend to feature higher labour expenses per employee.

Further, we look at maturity diversification within a specific industry. In particular, the question of whether the maturity structure of loans changes according to bank concentration is investigated. Assuming that higher bank concentration is accompanied by increased experience or intensified monitoring, a concentration strategy might change a bank's monitoring activity. Initial evidence suggests that a higher loan portfolio concentration decreases the portion of short- and medium term loans, whereas it increases the portion of long-term loans. This might indicate that increased experience as well as close monitoring and supervision fosters greater confidence in the borrower not to default on its loan and thus favours longer-term loan contracts.

Aretz and Pope (2013) find that global, country and industry effects are common factors driving firms' default risk. Similar to our regional factor based on the postal code areas, regional GDP growth on district level (Landkreise) might be an important variable reflecting regional differences in the business cycle. Memmel et al. (2012) conduct an analysis which is closely related to our approach in terms of data and methodology. The authors focus on the

[38] Correlation of: cost-to-income ratio and fixed effects (pvalue) at -0.0075 (0.3802); labour costs over total number of employees and fixed effects at -0.0387 (0.0000); employee ratio and fixed effects at 0.1086 (0.0000).

common drivers of a loan portfolio's default risk, which are also used in our study, and find no significant explanatory power for the inclusion of regional GDP growth.

The hypothetical loan portfolio loss rates are those of a portfolio where the bank's portfolio weights and the nationwide loss rates are applied. This might be problematic for the majority of banks in our data set as regional banks generally operate around their location. Replacing the nationwide loss rates with the regional loss rates in the calculation of the hypothetical loss rates does not lead to significantly different results.

The explanatory power of the overall model is higher for a balanced panel in comparison to an unbalanced panel; but regression results remain qualitatively the same. We report results for an unbalanced panel in this paper to avoid any potential bias that might arise due to defaulted banks dropping out of the sample.

Our research question could be subject to a reverse causality issue. A bank may decide to concentrate its loan portfolio especially on those industries that are less exposed to risk, i.e. show comparatively low loan loss rates. In this case, less risky industries would affect the bank's concentration level. However, the decision of the bank's loan portfolio management to focus – whatever reasons – on specific industries or sectors is already included on the right-hand side of our estimation equation. The inclusion of common risk factors explicitly controls for the composition of the loan portfolio. Thereby, we examine the excess return only which can be attributed to the bank's selection and monitoring abilities. This eliminates above raised concern in our view. Finally, a lower loan loss rate could, on average, result from the borrowers' good quality and not from the banks' selection and monitoring abilities. Similar to above reasoning, the common risk factors included in our study take the average loan loss rate in a specific industry/sector (and thereby the performance of the corresponding borrowers) into account.

9 Conclusion

Concentration versus diversification in a bank's loan portfolio – this is an important topic in the banking literature. For stocks and bonds, the benefits of portfolio diversification are not questioned, but at the heart of modern finance. For banks' loan portfolios, however, our study shows that specialized banks have, on average, both a lower credit risk in general and a less unexpected portion of it.

Our data set is especially suitable for addressing this question because we not only have the loan exposures, but also the corresponding write-offs and write-downs at our disposal. In addition, these loan exposures and matching write-downs are broken down into different industries and maturity brackets. This makes it possible to verify the empirical findings from different angles.

First, it allows us to analyse the impact of loan portfolio concentration on historic loan losses and their standard deviation. The former captures both the expected and unexpected parts of credit risk, whereas the latter addresses its unexpected portion in a narrow definition of credit risk. In comparison to the measures commonly suggested by the literature, e.g. loan loss provisions or non-performing loans, the historic loan losses provide a more comprehensive estimator of the loan portfolio's real underlying credit risk. Our database therefore allows us to provide a considerable contribution to the literature.

Second, we can separate the bank-specific selection and monitoring abilities from the composition of the loan portfolio. By the inclusion of common risk factors, we control for the impact of, say, industrial concentration risks arising from a concentrated loan portfolio and investigate the compensatory benefits through banks' selection and monitoring qualities.

We find that concentrated banks reduce their credit risk in general through the build-up of sector-specific knowledge, i.e. they have lower loan loss rates on average. Examining in more detail the loss rates within a bank's loan portfolio reveals that the loss rates of the largest industry-specific loan exposures are, on average, significantly lower than the corresponding loss rates of the smallest industry-specific loan exposures. This result holds both for the overall sample period and in the cross-sectional dimension. Finally, our database enables us to examine the relation between loan portfolio concentration and the unexpected part of credit risk, measured by the standard deviation of the loan loss rates. We find that banks with a concentrated loan portfolio exhibit a lower unexpected part of credit risk in comparison to more diversified banks. These findings suggest that specialized German banks acquire considerable selection and monitoring abilities that reduce their loan portfolio's credit risk beyond associated concentration risks.

Regarding implications concerning the trade-off between concentration and diversification in the banks' credit portfolio, our analysis suggests the following: Allowing banks to realize specialization benefits from loan portfolio concentration enhances financial stability given that this specialization results in a more efficient allocation of credit risk. However, when assessing the effects on financial stability, one has to bear in mind that, in our study, we analyse the relationships around the mean of the distribution, not the tail, and that diversification in the credit portfolio becomes especially important in extreme events.

Appendices

Variable	Definition
Q_T	Nationwide loss rate of the entire loan portfolio at time t, i.e. nationwide write-downs / nationwide loan exposure
Q_TI	Industry factor, i.e. deviation in the loss rate of a hypothetical portfolio that are due to bank i's deviation in the industry composition
Q_TM	Maturity factor, i.e. deviation in the loss rate of a hypothetical portfolio that are due to bank i's deviation in the maturity composition
Q_TR	Regional factor, i.e. deviation in the loss rate of a hypothetical portfolio that are due to bank i's deviation in the regional composition
HHI	Herfindahl-Hirschman Index, i.e. $[1/27; 1]$
S	Shannon-Entropy, i.e. $[-\ln(27)\,; 0]$
D_NATION	Absolute distance measure with the German wide market portfolio as benchmark portfolio, i.e. [0; 1]
D_REGION	Absolute distance measure with a regional market portfolio as benchmark portfolio, i.e. [0; 1]
LN_TA	Natural logarithm of total assets (total assets in million euros)
RWA_TA	Risk-weighted assets (in million euros) / total assets (in million euros)
ROA	Return-on-Assets, i.e. net income (in million euros) / total assets (in million euros)
ER	Average number of employees / total assets (total assets in million euros)

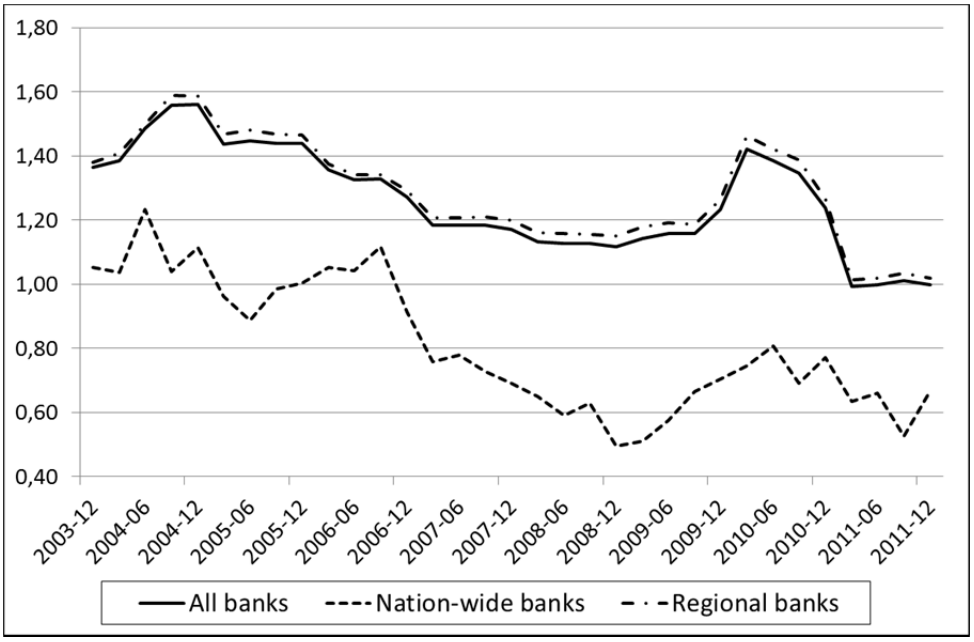
Appendix I: Definitions of variables.

Variable	Mean	SD	P5	P95
Ln (total assets) (LN_TA)				
2003-2011	6.3003	1.5028	4.0921	8.7402
2003	6.1164	1.4668	3.8813	8.5183
2011	6.9334	1.4906	4.6579	9.3978
Return-on-assets (ROA)				
2003-2011	0.26%	0.59%	0.02%	0.67%
2003	0.22%	0.38%	0.02%	0.54%
2011	0.28%	0.55%	0.00%	0.71%
Employee ratio (ER)				
2003-2011	0.2430	0.0980	0.0908	0.3609
2003	0.2697	0.0982	0.1321	0.3879
2011	0.2134	0.1002	0.0447	0.3089
Risk weighted assets/total assets (RWA_TA)				
2003-2011	0.5859	0.1345	0.3613	0.7828
2003	0.6245	0.1206	0.4193	0.7999
2011	0.5351	0.1461	0.3036	0.7569

Appendix II: Descriptive statistics of bank-specific control variables, averages for 2003, 2011, and 2003-2011 period.

	LN_TA	RWA_TA	ROA	ER
LN_TA	1			
RWA_TA	-0.2665	1		
ROA	-0.0886	0.0811	1	
ER	-0.4852	0.2644	0.0236	1

Appendix III: Pearson's correlation coefficients of bank-specific control variables.



Appendix IV: Average bank-wide yearly loss rate for all banks, nationwide banks and regional banks, respectively. Period December 2003 until December 2011.

Industry-wide exposure of maturity bracket k:

$$X_{t,j,k} := \sum_{i=1}^{N} X_{i,t,j,k} \tag{20}$$

Industry-wide exposure:

$$X_{t,j} := \sum_{k=1}^{3} X_{t,j,k} \tag{21}$$

Industry-wide, maturity-specific change in value:

$$C_{t,j,k} := \sum_{i=1}^{N} C_{i,t,j,k} \tag{22}$$

Change in value (in euro) of loan exposure of bank i at time t to industry j:

$$C_{i,t,j} := \sum_{k=1}^{3} C_{i,t,j,k} \tag{23}$$

Industry-wide change in value (in euro) at time t of industry j:

$$C_{t,j} := \sum_{k=1}^{3} C_{t,j,k} \tag{24}$$

Weight:

$$w_{i,t,j,k} = \frac{0.5 \cdot X_{i,t,j,k} + X_{i,t-1,j,k} + X_{i,t-2,j,k} + X_{i,t-3,j,k} + 0.5 \cdot X_{i,t-4,j,k}}{0.5 \cdot X_{i,t} + X_{i,t-1} + X_{i,t-2} + X_{i,t-3} + 0.5 \cdot X_{i,t-4}} \tag{25}$$

Industry-wide, maturity-specific yearly loss rate:

$$Q_{t,j,k}^{ind \times mat} := \frac{4 \cdot \sum_{m=0}^{3} C_{t-m,j,k}}{0.5 \cdot X_{t,j,k} + X_{t-1,j,k} + X_{t-2,j,k} + X_{t-3,j,k} + 0.5 \cdot X_{t-4,j,k}} \tag{26}$$

Industry-wide yearly loss rate:

$$Q_{t,j}^{ind} := \frac{4 \cdot \sum_{m=0}^{3} C_{t-m,j}}{0.5 \cdot X_{t,j} + X_{t-1,j} + X_{t-2,j} + X_{t-3,j} + 0.5 \cdot X_{t-4,j}} \tag{27}$$

Hypothetical industry-, but not maturity-specific nationwide average loss rate of bank i:

$$hq_{i,t}^{ind} = \sum_{j=1}^{27} Q_{t,j}^{ind} \cdot \left(\sum_{k=1}^{3} w_{i,t,j,k} \right) \tag{28}$$

Appendix V: Some additional notation. Note that time t is suppressed in the wording to increase readability.

	SD (Bank-wide loss rate)		
Variables	All banks	All banks	Regional banks
SD_YHAT	1.2004***	1.2098***	1.1322***
	(0.146)	(0.146)	(0.151)
$\overline{\text{LN\_TA}}$	-0.0003***	-0.0003***	-0.0003***
	(0.000)	(0.0009	(0.000)
$\overline{\text{RWA\_TA}}$	-0.0034***	-0.0031***	-0.0040***
	(0.001)	(0.001)	(0.001)
$\overline{\text{ROA}}$	-0.2555***	-0.2601***	-0.2639***
	(0.057)	(0.057)	(0.058)
$\overline{\text{ER}}$	0.0016	0.0018	0.0022
	(0.002)	(0.002)	(0.002)
$\bar{\hat{\text{S}}}$	-0.0011***		
	(0.000)		
$\overline{\text{D\_NATION}}$		-0.0034***	
		(0.001)	
$\overline{\text{D\_REGION}}$			-0.0010
			(0.001)
$\overline{\text{CONSTANT}}$	0.0066***	0.0094***	0.0098***
	(0.001)	(0.001)	(0.001)
R-squared (between)	13.4%	13.1%	10.6%
Number of Groups	933	933	902

Appendix VI: This table shows regression results from a cross-sectional regression. The dependent variable is the standard deviation of the bank-wide loss rate. The right-hand side of the regression equation is based on the standard deviation of the hypothetical portfolios loss rates (SD_YHAT), the average values of the respective concentration measures ($\hat{\text{S}}$, D_NATION, D_REGION) and bank-specific control variables, see Appendix I, II. ***,**,* denote statistical significance at the 1, 5 and 10 percent levels, respectively. Standard errors in parenthesis.

References

Acharya, Viral V., Iftekhar Hasan and Anthony Saunders (2006): Should banks be diversified? Evidence from individual bank loan portfolios, The Journal of Business, Vol. 79 (3), 1355-1412.

Aretz, Kevin and Peter F. Pope (2013): Common factors in default risk across countries and industries, European Financial Management, Vol. 19(1), 108-152.

Bebczuk, Ricardo and Arturo Galindo (2007): Financial crisis and sectoral diversification of Argentine banks, 1999-2004, Applied Financial Economics, Vol. 18 (3), 199-211.

Behr, Andreas, Andreas Kamp, Christoph Memmel and Andreas Pfingsten (2007): Diversification and the banks' risk-return characteristics – evidence from loan portfolios of German banks, Deutsche Bundesbank, Series 2, No. 05/2007.

Boot, Arnould W. A. (2000): Relationship Banking: What Do We Know?, Journal of Financial Intermediation, Vol. 9(1), 7-25.

Boeve, Rolf, Klaus Duellmann and Andreas Pfingsten (2010): Do specialization benefits outweigh concentration risks in credit portfolios of German banks?, Deutsche Bundesbank, Series 2, No. 10/2010.

Carey, Mark, Mitch Post and Steven A. Sharpe (1998): Does Corporate Lending by Banks and Finance Companies Differ? Evidence on Specialization in Private Debt Contracting, The Journal of Finance, Vol. 53, 845-878.

Crouhy, Michel, Dan Galai and Robert Mark (2000): A comparative analysis of current credit risk models, Journal of Banking and Finance, Vol. 24 (1), 59-117.

Dennis, Steven, Debarshi Nandy and Ian G. Sharpe (2000): The determinants of contract terms in bank revolving credit agreements, Journal of Financial and Quantitative Analysis, Vol. 35 (1), 87-110.

Deutsche Bundesbank (2009): Guidelines on the borrowers statistics, Banking statistics guidelines and customer classification, January.

Deutsche Bundesbank (2013): Banking statistics, Statistical Supplement 1 to the Monthly Report, April.

Frenken, Koen (2007): Entropy statistics and information theory, in: Horst Hanusch and Andreas Pyka (eds), The Elgar Companion to Neo-Schumpeterian Economic, Cheltenham, 544-555.

Hartmann-Wendels, Thomas, Andreas Pfingsten and Martin Weber (2010): Bankbetriebslehre Springer Verlag, 5th edition.

Hayden, Evelyn, Daniel Porath and Natalja von Westernhagen (2007): Does diversification improve the performance of German banks? Evidence from individual bank loan portfolios, Journal of Financial Services Research, Vol. 32 (3), 123-140.

Herfindahl, Orris Clement (1950): Concentration in the steel industry, Columbia University Press, New York.

Hirschman, Albert (1945): National power and the structure of foreign trade, University of California Press, Berkeley.

Kamp, Andreas (2006): Diversifikation versus Spezialisierung von Kreditportfolios - Eine empirische Analyse, Fritz Knapp Verlag, Frankfurt am Main.

Kamp, Andreas, Andreas Pfingsten and Daniel Porath (2005): Do banks diversify loan portfolios? A tentative answer based on individual bank loan portfolios, Deutsche Bundesbank, Series 2, No. 03/2005.

Liu, Chi-Chun and Stephen G. Ryan (1995): The effect of bank loan portfolio composition on the market reaction to and anticipation of loan loss provisions, Journal of Accounting Research, Vol. 33 (1), 77-94.

Markowitz, Harry (1952): Portfolio selection, Journal of Finance, Vol. 7 (1), 77-91.

Memmel, Christoph, Yalin Gündüz and Peter Raupach (2012): The common drivers of default risk, Deutsche Bundesbank, Discussion Paper, No. 36/2012.

Memmel, Christoph and Ingrid Stein (2008): The Deutsche Bundesbank's Prudential Database, Schmollers Jahrbuch, Vol. 128, 321-328.

Pfingsten, Andreas and Kai Rudolph (2004): Eine empirische Analyse der Kreditportfoliozusammensetzung deutscher Bankengruppen, Zeitschrift für Betriebswirtschaft, Ergänzungsheft, Vol. 2, 1-24.

Quagliariello, Mario (2007): Banks' riskiness over the business cycle: A panel analysis on Italian intermediaries, Applied Financial Economics, Vol. 17 (2), 119-138.

Rossi, Stefania P.S., Markus S. Schwaiger and Gerhard Winkler (2009): How loan portfolio diversification affects risk, efficiency and capitalization: A managerial behavior model for Austrian banks, Journal of Banking and Finance, Vol. 33 (12), 2218-2226.

Shannon, Claude Elwood (1948): A mathematical theory of communication, Bell System Technical Journal, Vol. 27 (10), 623-656.

Sharpe, Steven A. (1990): Asymmetric Information, Bank Lending, and Implicit Contracts: A Stylized Model of Customer relationships, The Journal of Finance, Vol. 45, 1069-1087.

Tabak, Benjamin M., Dimas M. Fazio and Daniel O. Cajueiro (2011): The effects of loan portfolio concentration on Brazilian banks' return and risk, Journal of Banking and Finance, Vol. 35 (11), 3065-3076.

Theil, Henri (1972): Statistical decomposition analysis, North-Holland Pub., Amsterdam.

Wilson, Thomas C. (1998), Portfolio credit risk, Economic Policy Review, Vol. 4 (10), 71-82.