Journal of Financial Stability 66 (2023) 101119

Contents lists available at ScienceDirect

Journal of Financial Stability

journal homepage: www.elsevier.com/locate/jfstabil

Bank competition and credit risk: The case of Peru[☆]

Jorge Pozo[a,*,1], Youel Rojas[b,2]

[a] Monetary Statistics Department at the Central Reserve Bank of Peru, Jr. Santa Rosa 441-445, Lima-1, Peru
[b] Macroeconomic Modelling Department at the Central Reserve Bank of Peru, Jr. Santa Rosa 441-445, Lima-1, Peru

ARTICLE INFO

JEL classification:
E44
G21
L11

Keywords:
Competition
Credit risk
Risk-taking
Financial stability

ABSTRACT

This paper studies empirically the relationship between competition in the loan market and credit risk in the Peruvian financial system. Our finding challenges the theoretical work of Martinez-Miera and Repullo (2010) that finds a U-shaped relationship between competition and risk-taking, as well as the empirical work of Jiménez et al. (2013) that provides evidence that supports this nonlinear relationship in a developed economy as Spain. In contrast, we find that in Peru the shape of the relationship between competition and credit risk is more complex and it depends on the competition measure.

1. Introduction

This work is motivated by the theoretical work of Martinez-Miera and Repullo (2010) that finds a U-shaped relationship between competition and risk-taking and the empirical work of Jiménez et al. (2013) that finds support for a nonlinear relationship in Spain. However, we depart from Jiménez et al. (2013) since we aim to test the hypothesis of Martinez-Miera and Repullo (2010) in an emerging economy as Peru and also we make use of more granular data in addition to the standard bank-time level employed in Jiménez et al. (2013), to control for unobserved factors that can bias the results. Hence, this work aims to study empirically the relationship between bank competition in the loan market and credit risk in the Peruvian financial system.

To do so, we perform an empirical analysis that consists in two parts. In a first part, we estimate a model as in Jiménez et al. (2013) with bank-time level data; but in a second part, we go one step further and use micro level data to add a geography dimension, allowing us to specify a model in bank-time-region dimensions. For the first model, we use the public information from the webpage of Superintendency of Banking, Insurance and Private Pension Fund Administrators (which we refer henceforth SBS, by its spanish acronym), while for the second model specifications we use more granular data from the Credit Registry Data (which we refer henceforth RCC, by its spanish acronym), which is restricted information. As in Jiménez et al. (2013), we use three well-known measures of competition: the number of relevant competitors that a financial institution faces, loan share of four-largest financial institutions (C4) and the Herfindahl–Hirschman index (HHI). Our measure of credit risk is the non-performing loan ratio (referred in other papers as bank risk-taking measure), which is taken directly from official sources (SBS) or is built using the granular data from the RCC.

In the first model, with bank-time level data for the 2003–2019 period, when using the number of banks as our competition measure, results suggest that for Peruvian banks there is an inverted U-shaped relationship between competition and credit risk, unlike Jiménez et al. (2013). This results might hold when considering all institutions and assuming non-competition across groups.[3] However, if any we find a U-shaped relationship when using the C4 *y* HHI. Results are robust if we include non-bank financial institutions assuming non-competition across different groups of financial institutions. However, when using the C4 or HHI, we might find a U-shaped relationship, when studying only banks or all financial institutions.

☆ The views expressed in this paper do not necessarily represent those of the Central Reserve Bank of Peru. We thank Carola Müller for her valuable comments during the XXV Meeting of the Central Bank Researchers Network of CEMLA. We thank the Monetary Statistics Division and the Financial Analysis Division of the Central Reserve Bank of Peru for giving us access to the micro data. We also thank two anonymous referees for the very constructive comments.

* Corresponding author.
E-mail addresses: jorge.pozo@bcrp.gob.pe (J. Pozo), youel.rojas@bcrp.gob.pe (Y. Rojas).
1 Head.
2 Senior Analyst.

3 The latter assumption is because we believe there is a weaker competition across different groups of financial institutions at more aggregated market, i.e., we expect more competition at the provincial level rather than at the regional level.

https://doi.org/10.1016/j.jfs.2023.101119
Received 19 January 2022; Received in revised form 22 February 2023; Accepted 27 February 2023
Available online 6 March 2023
1572-3089/

Since in the specification at the institution-time level we cannot control for omitted variables that may affect the dynamics of the relationship between competition and credit risk,[4] we use micro data at the client-bank level to build a panel with region-bank-time dimensions for the 2004–2019 period. Our analysis starts from the assumption of segmented regional loan markets to achieve identification, and adopts a within-region and a within-bank estimators. Furthermore, only loans to firms (commercial loans and loans to microenterprises) are considered. In this case we find evidence of an inverted U-shaped relationship between competition and credit risk when considering the number of banks as the competition measure. This is robust even when controlling for supply or demands factors and when non-bank financial institutions (assuming competition across groups) are included, though results are no longer statistically significant in the latter specification.[5] However, when using the C4 or the HHI, we find a more statistically significant U-shaped relationship, when studying only banks or all financial institutions (assuming competition across groups).

All of the previous empirical findings show evidence, in an emerging economy like Peru, of an non-linear relationship between competition and credit risk, which is conditional to the competition measure used. Thus, With the number of banks, in contrast to Jiménez et al. (2013), we find evidence of an inverted U-shaped relationship between competition and credit risk. And with C4 and HHI we might find evidence of a U-shaped relationship.

Table 1 compares Peru to other Latin American countries on various measures of concentration and competition. The Peruvian banking sector is one of the most concentrated in the region. For instance, in 2017 the share of the three largest banks assets in Peru was 72.7%, while in Chile it was 42.6%. The same is observed with the 5-bank asset concentration measure. Also, in Peru in 2014 the elasticity of bank revenues to input prices (H-statistic) was one of the smallest, providing evidence of relatively low competition in the Peruvian banking system. Finally, the markup is largest in the Peruvian banking system suggesting relatively poor competition levels within Latin America. Thus, relative to other emerging market economies or advanced economies, Peru shows high levels of concentration and market power. In terms of financial stability, one important trend after the 2008 global financial crisis, in Peru, is that the bank nonperforming loans have been increasing steadily, while this tendency is not clear in other countries in Latin America (see Fig. 1).

For the purpose of this paper, our definition of the Peruvian financial system includes all financial intermediaries: banks and non banks.[6] By 2018, 53 financial institutions, that provide loans to the private sector, constitute the Peruvian financial system.[7] Even though from a country perspective banks are relatively more important, this

Table 1
Bank competition and concentration in Latin America.
Source: Global Financial Development. 3-bank asset concentration: Assets of three largest banks as a share of total banking assets. 5-bank asset concentration: Assets of three largest banks as a share of total banking assets. H-statistic: A measure of the degree of competition in the banking market. It measures the elasticity of banks' revenues relative to input prices. The closer to 1, the higher the competition. Lerner index: A measure of market power. It compares output pricing and marginal costs (that is, markup). A high value suggests less competition.

	3-bank asset concentration (%) 2017	5-bank asset concentration (%) 2017	H-statistic 2014	Lerner index 2014
Brazil	56.6	83.4	0.72	0.21
Chile	42.6	68.7	0.77	0.25
Colombia	78.7	89.4	0.51	0.48
Mexico	49.4	69.0	0.83	0.38
Peru	72.7	88.1	0.60	0.50
Uruguay	70.1	87.4	0.80	0.19

is not necessarily true within a region.[8] There are regions where the lending role of nonbank financial institutions becomes relatively more important. For this reason, in our analysis we focus not only on banks but also on nonbank financial institutions.

Fig. 2 shows the importance of a regional analysis. The red line shows that in recent years the total number of financial institutions in the country has been decreasing, but there is not a clear long-term trend. However, the ratio of the average number of institutions in a region to the total number of financial institutions (blue bars) has been increasing steadily from 16.7% since 2003 to 49.4% in 2019. In other words, in 2019 on average a region has the presence of 49.4% of all financial institutions in the country. It suggests that although the number of financial institutions does not consistently increase, the presence of these institutions in many regions has raised. The increment in the presence of financial institutions has been heterogeneous across regions, which allows us to gain variability in a measure of competition or concentration at a regional level. Crucially, regional level data allows us to control for regional demand trends that can influence bank entry, bank competition and risk-taking behavior.

It is precisely because of these features, that the regional composition of the Peruvian financial system provides us with an interesting case to study the relationship between lending competition and credit risk.

The remainder of this paper is organized as follows. Section 2 presents the literature review. Section 3 shows the data, the first model specification using bank-level data and its the empirical results. Section 4 presents the granular assessment results. Finally, Section 5 concludes.

2. Literature review

This paper is related to the literature of competition and financial stability. One strand of this theoretical literature is the "competition-fragility" view (Marcus, 1984; Smith, 1984; Boot and Greenbaum, 1993; Demsetz et al., 1996; Hellman et al., 2000; Repullo, 2004) and argues that bank competition reduces the value of a bank's franchise and, as a result, encourages more risk-taking to increase returns (Keeley, 1990; Demsetz et al., 1996; Beck et al., 2006, 2007; Turk-Ariss, 2010; Fungacova and Weill, 2010). The "competition-stability" view (Boyd and De Nicoló, 2005; Allen et al., 2011) is the contrary strand of this theoretical literature, and contends that increased bank competition reduces market power and interest rates charged to borrowers, resulting in fewer incentives to take on more risky projects and create healthy banks (Boyd et al., 2006; De Nicoló and Loukoianova, 2007; Schaeck et al., 2009; Schaeck and Cihák, 2014; Clark et al., 2018).

[4] For instance, changes in business opportunities and risk profiles at the regional level and/or market strategies and diversification of financial institutions over time.

[5] We believe that the various conflicting effects of competition on credit risk may be visible in data if all financial institutions are mixed at a more granular level. In Peru, credit markets are segmented and with varying degrees of competition among financial groups serving each market. By combining all financial institutions at once within a specification, we may be mixing all of the various effects at once.

[6] The financial system in Peru is divided in five financial groups: Commercial banks (banks), *empresas financieras*, municipal credit and saving institutions (CMACs by its Spanish acronym), rural credit and savings institutions (CRACs by its Spanish acronym) and small business and microenterprises development institutions (EDPYMEs by its Spanish acronym).

[7] These were composed by 16 banks, that represent the 88% of the total loans, and other non-bank financial institutions as 10 *empresas financieras*, 12 CMACs, 6 CRACs and 9 EDPYMEs. The Peruvian financial system also includes other more specialized institutions, however, since the participation in the credit market is very small, we omit them.

[8] Peru is divided into 25 regions (24 *departamentos* and the constitutional province of Callao).

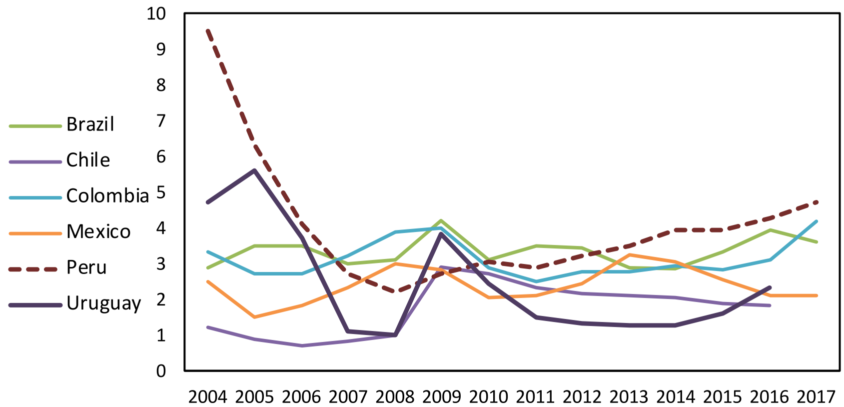


Fig. 1. Bank non-performing loans to gross loans (%) in Latin America. Source: Global Financial Development.

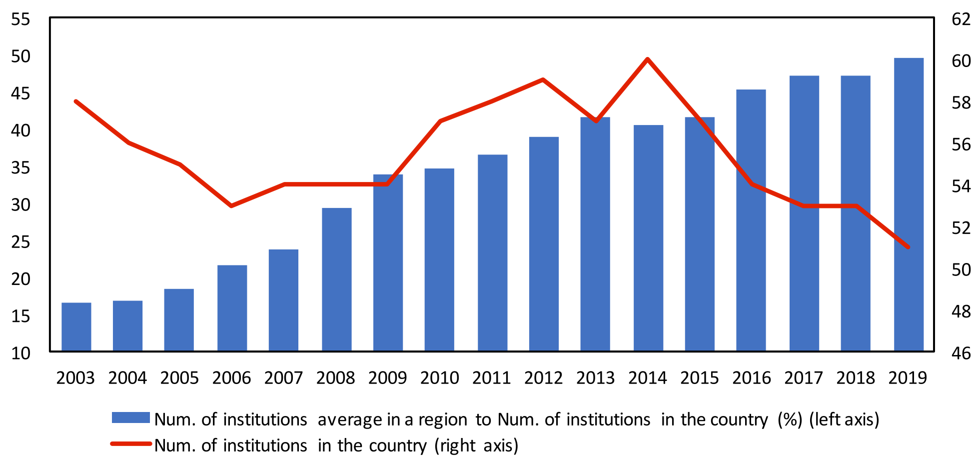


Fig. 2. Presence of financial institutions across regions in Peru.
Source: The Superintendency of Banking, Insurance and Private Pension Fund Administrators (SBS by its Spanish acronym). Own calculations. Here, we do not include the foreign market as another region.

Moreover, our paper shares the viewpoint of a third strand of literature that demonstrates that competition and financial stability are more complex and non linear (Caminal and Matutes, 2002; Allen and Gale, 2004; Martinez-Miera and Repullo, 2010). This is consistent with Zigraiova and Havranek (2015), which reports a large heterogeneity in the sign and magnitude of this relationship in the literature, which seems driven by differences in the definitions of financial stability and bank competition used. For instance, Liu and Wilson (2013) for Japan demonstrate that the relationship between competition and risk varies across bank types based on different initial levels of risk. Similarly, we contribute to the literature by taking advantage of bank heterogeneity and using granular data to control for demand and supply factors that may influence this relationship. Also, Schaeck et al. (2009) demonstrate that competition and concentration have different effects on bank risk: competition reduces the likelihood of a crisis, whereas concentration decreases the probability of a crisis. Similarly, in our paper, we distinguish between a measure of competition, the number of banks, and a measure of concentration, the market loan share of the four largest banks or the HHI.[9] We show robust evidence that both of these measures have an effect on the relationship between bank risk and competition. Finally, Fernández and Garza-García (2015), Berger Allen et al. (2017) show that more competition increases bank portfolio risks while decreasing overall bank risk. In our paper, we use non-performing loans as a measure of bank risk, and to distinguish it from risk portfolio effects, we include bank leverage to control for bank capacity to absorb losses.

In particular, our paper is closely related to the empirical and theoretical literature that aims to explore the relationship between bank competition and bank risk-taking. As commented in Martinez-Miera and Repullo (2010), the conventional wisdom is that increasing competition leads banks to take more risk. The key assumption, as commented by the authors, is the exogenous distribution of returns of bank assets. For instance, Bolt and Tieman (2004) conclude that higher competition leads to higher bank risk-taking. They develop a dynamic framework where banks compete for loans by establishing acceptance criteria. Their model suggests that competition reduces margins and thus bank's charter value declines. This provides higher incentives to take more risk raising the bank failure probability. In other words, less strictness to issue loans decreases loan quality. Similarly, in a dynamic model of imperfect competition with prudent and gambling asset, Repullo (2004) finds that in the absence of regulation if banks' margins are small, the equilibrium features banks investing only on risky assets.[10]

Boyd and De Nicoló (2005) show that there exists a risk-incentive mechanism that operates in the opposite direction of the one suggested by the previous literature. In their work the key assumption is that bank loan defaults are perfectly correlated. They also assume that bank

[9] We are aware that concentration measures may be a noisy proxy for competition and that concentration and competition describe different aspects of banking systems (Claessens and Laeven, 2004).

[10] The prudent asset has a higher expected return, and the gambling asset yields a higher payoff if the gamble succeeds.

borrowers optimally choose a higher project risk, the higher the loan interest rate set by banks. As competition increases banks have less market power to raise loan rates and hence with smaller loan interest rates borrowers choose lower risk projects. Due to the perfect correlation, loan default probability coincides with bank default probability. As a result, through that mechanism, as competition increases bank failure probability decreases.

Martinez-Miera and Repullo (2010) argue that the findings of Boyd and De Nicoló (2005) does not necessarily hold in the more realistic case of imperfect correlation of loan default. This is because bank competition reduces the interest rate pay from performing loans, which provides a buffer to cover loan losses, and hence increases the bank default probability. They identify a *risk-shifting effect*, which is the one described in Boyd and De Nicoló (2005) that suggests small loan rates after a higher competition reduces borrower incentive to take risk, which in turn pushes bank default probability down. And a margin effect that suggests that small loan rates also reduce bank capacity to avoid defaulting. They find that in a very competitive market the margin effect dominates, while in a less competitive market the *risk-shifting effect* dominates. As a result, they formulate a U-shaped relationship between the number of banks (bank competition measure) and the risk of bank failure.

The empirical work of Jiménez et al. (2013) using annual Spanish data and different measures of lending competition for the 1988–2003 period supports the nonlinear relationship found in Martinez-Miera and Repullo (2010). We depart from Jiménez et al. (2013), since we test the hypothesis of Martinez-Miera and Repullo (2010) in an emerging economy as Peru and use granular data. And, in contrast to Jiménez et al. (2013), the relationship between bank competition and risk-taking depends on the competition measure used.

Next, we move to the empirical section and provide the empirical evidence of the relationship of bank competition and credit risk in an emerging economy like Peru.

3. Bank level evidence

We use two levels of information. First, in this section, following the related literature, we use a bank-time (or financial institution-time) level data. Later, we make use of a bank-region-time (or financial institution-region-time) level data, which is more granular and hence it allows us to control for demand and supply characteristics that can bias our results. The first dataset is publicly available at the Financial System Regulator of Peru's website (SBS, website).[11] The main limitation of this dataset is that the measure of credit risk is not available for firm loans but for all total loans. However, in the granular dataset, we can distinguish between the different type of loans (commercial loans, loans microenterprises, mortgage and personal loans) and hence we can focus on loans to firms, as in Jiménez et al. (2013).

First, in this section we describe the bank-time level dataset, the empirical model, and our first regression results. And in Section 4 we focus on the more granular data.

3.1. Bank level data

In this part, we follow Jiménez et al. (2013) and use bank-time level data to test the effects of competition on bank risk-taking behavior. We compute different measures of bank competition, calculated as a weighted average based on regional information. The first measure, it is the number of banks operating in each region. The other two measures are essentially concentration measures, that are used as proxies of competition measures: the share of loans of the four-largest financial institutions operating in each region (C4),[12] and the Herfindahl–Hirschman index (HHI), which is the sum of banks' squared market shares in loans in each region.

Peru is divided into 25 regions (24 regions and the Constitutional Province of Callao); however, as part of our analysis we include the foreign market as an additional region.[13] Consequently, in our analysis we consider 26 regional credit markets for Peru. In particular, the higher the number of banks the higher the competition, while the higher the C4 and HHI ratios, the higher the concentration and hence we might expect a lower competition. The information about loans and number of financial institutions in each region is provided by the SBS, the financial regulator in Peru, However, at the regional level, public data is only available for total credit, with no breakdown by type of credit or credit status (delay situation) for firms and households.

Since the Peruvian credit market is segmented geographically into 26 regions, the competition measures have to reflect the degree of competition that each bank faces in each of the regional market where it operates. Hence, we construct an aggregate competition measure faced by each bank using a weighted average, where the weights are the market loan share each bank holds in each region. For instance, the competition measure of "number of banks" for a bank i at year t is defined as the number of banks that has the representative region (or representative market) for bank i at year t. This competition measure is calculated as the weighted average (by total loans) of the number of banks over all regions where the bank grants loans. C4 denotes the share of the 4 largest banks in the representative market for bank i at time t, calculated as the weighted average (by total loans) of the C4 over all regions where the bank i grants loans at year t. Finally, HHI is the Herfindahl–Hirschman Index of concentration for the representative region of bank i at time t, calculated as the weighted average (by total loans) of the HHI over all regions where the bank i grants loans at year t.

We include data to control for individual bank characteristics, as return on assets (ROA) and bank size or loan market share (SIZE). We also control for aggregate trends, such as the Peruvian business cycle, using the real GDP growth rate (RGDPGR). In addition, we include three control variables not included in Jiménez et al. (2013): bonds issued by non-financial institutions to credit ratio (BOND), the risk weighted asset to capital ratio (RWA) and participation of foreign debt on credit funding (FD). BOND controls for the preferences and/or opportunities for non-bank funding, while RWA controls for individual bank characteristics regarding bank capacity to handle a financial crisis and also individual preferences on risk-taking.[14] FD controls the bank's capacity to borrow from foreign markets, which in turn might affect banks' incentives to take risk.

Our dependent variable is the bank credit risk. In this document, it is measured as the ratio of nonperforming loans to total loans, and

[11] Available at https://www.sbs.gob.pe/estadisticas-y-publicaciones/estadisticas-/sistema-financiero_

[12] We use the four-largest financial institutions instead of the three or the five since in Peru the four-largest bank represents around 90% of the total loans.

[13] This is to account for loans issued from branches abroad. We include these loans since from the SBS available data, at the bank-time level, we cannot build up a NPL ratio of only loans issued by domestic branches. Indeed, only a few banks have branches abroad and the associated loans have a small participation. It represents on average 2.1% of bank loans in the 2003–2019 period. Notice that Peruvian financial institutions serving foreign markets face higher competition since they encounter as competitors larger international financial institutions.

[14] See Agur and Demertzis (2012, 2019) and Dell'Ariccia et al. (2014) for a detailed explanation of the effect of bank leverage on bank risk-taking (i.e., the *leverage channel*). Intuitively, the higher the leverage the lower the participation of owners' wealth on funding bank investment activities, the smaller the losses of the owners if banks default (due to limited liability), and hence the stronger the preferences to take higher risk.

it follows under the same criterion defined by the Peruvian financial regulator, SBS,[15]

$$\frac{\text{loan arrears (Big firms(15d), small firms(30d) mortgage(30d), personal(90d))}}{\text{Total credits}}100. \tag{1}$$

The information about nonperforming loan (NPL) ratio at an institution-time level is also provided by the SBS.[16]

Notice that we refer to the NPL ratio as a measure bank credit risk instead of bank risk-taking, as previous literature does. This is because, the NPL ratio looks more like a "credit risk" measure, since this is about the ex-post risk of the loan portfolio, rather than the ex-ante "risk-taking" of the bank, which is an overall bank risk and responds to both sides of bank balance sheet. This is why in a model as Martinez-Miera and Repullo (2010), this "risk-taking" is well-captured by the bank default probability. To try to overcome this, in this empirical part we control for the risk weighted assets to capital ratio, as a measure of bank capacity to absorb losses.

We assess not only the relationship between competition and credit risk in a sample of banks in Peru, but also consider a sample that includes non-bank financial intermediaries, so we consider the five main financial groups, as a whole, that exist in the Peruvian financial system: banks, *empresas financieras*, CMACs, CRACs and EDPYMEs.[17] We start focusing solely on banks, then we include all financial institutions.

We use annual data and the time period analyzed spans from 2003 to 2019. We start from 2003 to consider only the inflation targeting period in Peru and stop at 2019 to avoid the Covid-19 shock period. We start focusing solely on banks, then we include all financial institutions.

To ensure data consistency, we merge the time series for the same institution that for some reason changed its name or changed both its name and its financial group to which it belongs.[18] In addition, we include dummies to control for these and other corporate events.[19]

Table 2 presents the descriptive statistics for the variables when considering only banks. There are 16 banks in the period analyzed and at least 221 bank–year observations.[20] The average NPL ratio is 3.17% and it features a large dispersion. The average "number of banks" that exists in the representative region where a bank competes is 13.62. This variable also exhibits a relatively high degree of dispersion. Similarly, the average C4 and HHI are 0.83 and 0.21, respectively, suggesting a relatively high concentration in the loan market.

Table 3 presents the descriptive statistics for variables used for all financial institutions. In our period of study we have 65 financial institutions across the five financial groups and at least 873 bank–year observations.[21] In this case, the competition measures can be computed under two different assumptions: non-competition across groups, and competition across groups. In the former, we assume a financial institution competes only with those within its financial group, while in the latter case financial the institution can compete with institutions from any financial group.[22] The average "number of institutions" that exists in the representative region where a financial institution competes is 7.25 when non-competition across groups is assumed, while this is 31.39 when we allow competition across groups. The average C4 and HHI are 0.91 and 0.41, respectively, assuming non-competition across groups, and 0.72 and 0.16, assuming competition across groups. In general competition, measures are smaller when assuming competition across groups.

Fig. 3.a reports the behavior of "number of banks" for the four largest banks from 2001 to 2018 (BCP, Interbank, Continental and Scotiabank). In general, there is a common trend that governs the long-term dynamics. Around the 2008 global financial crisis, these banks exhibit relatively more dispersion, compared to other periods, on the competition that they are exposed to. Also, from 2002 to 2006, there was a NPL ratio reduction that was accompanied by less competition faced by these banks. Just before the financial crisis, from 2006 to 2008, there was a considerable increase of bank competition accompanied by a slow reduction of the NPL ratio. Since 2008 bank competition and the NPL ratio have been increased slowly. According to

Table 2
Descriptive statistics for bank–year observations.
Source: SBS. Own elaboration. S.D.: Standard deviation. We omit financial institutions with less than three observations, and observations with extreme value of the NPL ratio (0% and 100%).

Variables	Obs	Mean	S.D.	Minimum	Maximum
NPL_{it} (%)	223	3.17	2.88	0.02	33.43
Number of $banks_{it}$	223	13.62	1.61	9.01	16.00
$C4_{it}$	223	0.83	0.02	0.74	0.86
HHI_{it}	223	0.21	0.02	0.17	0.31
$SIZE_{it}$	223	0.08	0.10	0.00	0.35
ROA_{it} (%)	221	1.79	1.88	−11.75	7.71
FD_{it}	223	0.09	0.09	0.00	0.74
RWA_{it}	223	7.18	1.31	2.47	10.03
$RGDPGR_t$	17	0.05	0.02	0.01	0.09
$BOND_t$	17	0.10	0.06	0.05	0.21

Table 3
Descriptive statistics for financial institution–year observations.
Source: SBS. Own elaboration. S.D.: Standard deviation. $SIZE^{*}_{it}$ = (credit of institution i at time t)/(total credit of institution i's group at time t.) $SIZE^{**}_{it}$ = (credit of institution i at time t)/(total credit of the five groups at time t). We omit financial institutions with less than three observations, and observations with extreme value of the NPL ratio (0% and 100%).

Variables	Obs	Mean	S.D.	Minimum	Maximum
NPL_{it} (%)	881	5.64	5.17	0.02	73.34
Non-competition across groups					
Number of $institutions_{it}$	881	7.25	4.39	1.00	16.00
$C4_{it}$	881	0.91	0.09	0.70	2.00
HHI_{it}	881	0.41	0.22	0.17	1.00
$SIZE^{*}_{it}$	881	0.09	0.10	0.00	0.88
Competition across groups					
Number of $institutions_{it}$	881	31.39	11.58	5.71	65.03
$C4_{it}$	881	0.72	0.06	0.53	1.24
HHI_{it}	881	0.16	0.03	0.10	0.28
$SIZE^{**}_{it}$	881	0.02	0.05	0.00	0.31
ROA_{it} (%)	873	1.56	4.16	−39.09	17.20
FD_{it}	881	0.10	0.17	0.00	0.95
RWA_{it}	881	6.13	1.79	0.63	10.93

[15] Big firms correspond to corporate, big and mid firms. Small firms corresponds to small and micro business.

[16] According to IMF a loan is considered non-performing when more than 90 days pass without the borrower paying the agreed installments or interest.

[17] The construction of the different competition measures for any financial institution that belongs to any of the financial groups follows the same procedure provided for banks.

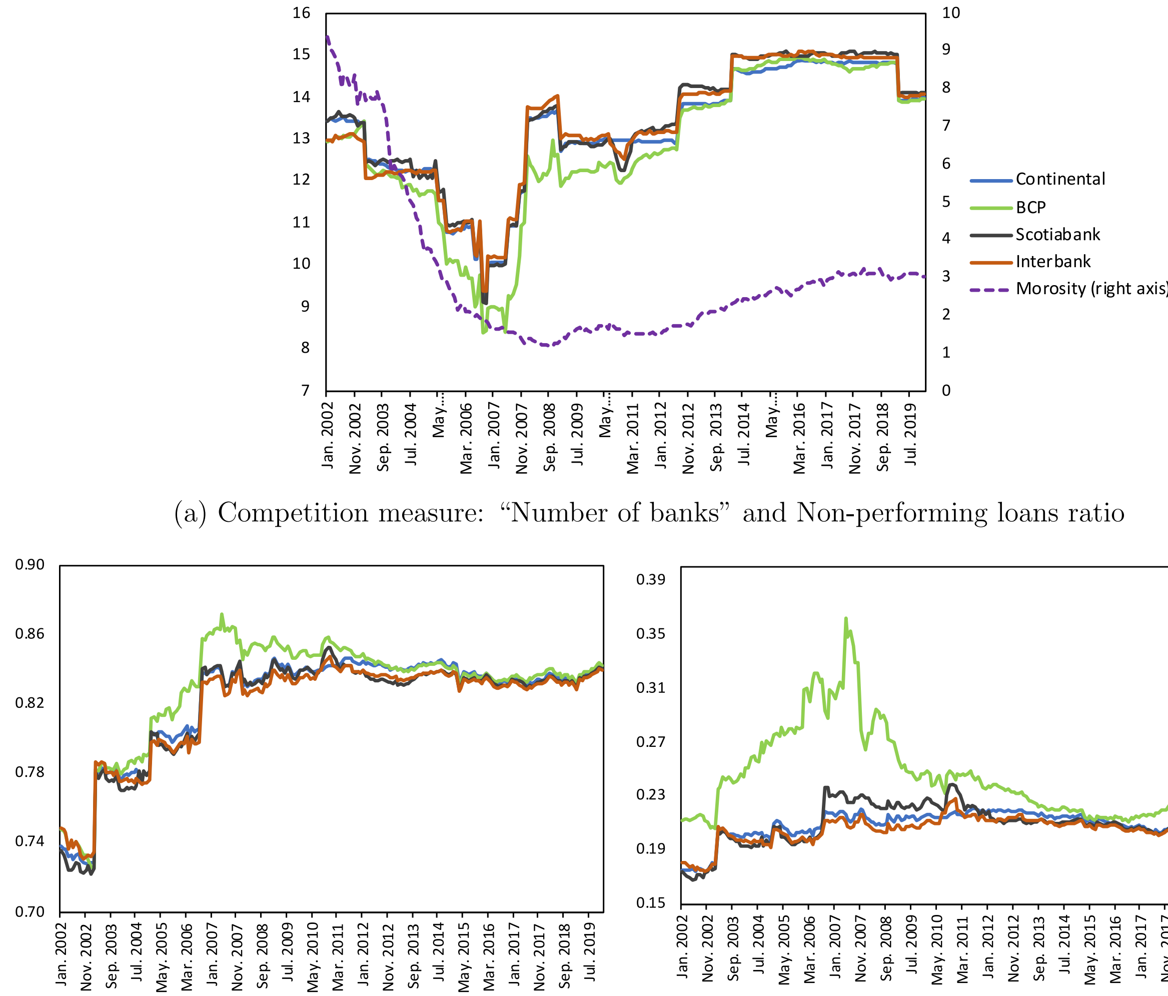
[18] For example, we might have information of institution A and institution B, which are actually the same institution, but they look different since it changes its name. Then, we collapse these two series into one.

[19] In all regressions we control for several financial events. In particular, we put dummies to control for five major economic events: only new institution names, mergers, new names and acquisitions, new names and mergers, new group affiliations (reallocation of institutions from one group to another), loan portfolio purchases.

[20] This is the number of bank–year observations before allowing for lags.

[21] The is the number of financial institution–year observations before allowing for lags. With respect to non-banks, we omit those with only one and two observations in our regressions: CMAC *Chincha*, EDPYME *Crear Cusco* and *Amerika Financiera*.

[22] For example, we assume that a bank cannot compete with an institution from the CAMC group. This also means that if ceteris paribus a financial institution moves from one group to another, it faces a different competition level.



(a) Competition measure: "Number of banks" and Non-performing loans ratio

(b) Concentration measure: C4

(c) Concentration measure: HHI

Fig. 3. Bank competition and concentration measures.
Source: SBS. Own calculations. This figure shows the measures of credit risk and competition for the four largest banks from 2001 to 2018: BCP, Interbank, Continental, Scotiabank. Number of banks = loan weighted average number of competitors; C4 = share of the 4 largest banks; HHI = Herfindahl–Hirschman Index of concentration; Morosity rate = Banking sector non-performing loans (NPL) ratio (SBS criterion) . Monthly frequency. Period: January 2002–December 2019.

this measure, for example, from 2004 to 2012 the largest bank in Peru, BCP, was operating in a less competitive representative market than the other three largest banks. This could be only explained by two reasons: BCP was operating in regions with a relatively small number of banks than in those regions where the other banks were operating. And/or BCP, compared to the other banks, increased its operation (or reallocate their loans) in regions where the presence of banks was relatively small.

As in the previous competition measure, Fig. 3.b and 3.c display C4 and HHI, respectively, for the four largest banks. Also, in this case there is a general trend and relatively more dispersion around the 2008 financial crisis. As with the number of banks measure of competition, in the case of BCP, from 2004 to 2012, it was operating in regions where the concentration level (measured with C4 or HHI) was relatively higher than in those regions where the other banks operate.

To have an idea of the heterogeneity of bank competition, Fig. 4 reports the "number of banks" with which each bank competes in December 2018. There are important differences in competition levels across banks. Note that in December 2018 the four largest banks operate in a market that features an intermediate level of competition. More specialized banks as Santander and Citibank operate in a more competitive market, while there are other also specialized banks as Mibanco and Azteca that operate in a less competitive market.

3.2. Model description

Similar to Jiménez et al. (2013) the empirical model is as follows:

$$endo\_var_{it} = \alpha + \beta_0 * endo\_var_{it-1} + \beta_1 * exo\_var_{it-1} + \beta_2 * exo\_var^2_{it-1} + \beta_3 * ctrl_{it} + \epsilon_{it}, \tag{2}$$

where the i subscript refers to a financial institution, the t subscript refers to a sample year and ϵ_{it} is a random error that has a normal distribution. The model describes the relationship between bank credit risk measure and bank competition measure, controlling for bank characteristics and the state of the business cycle. We might include bank fixed effects, to control for unobservable bank characteristics, or time fixed effects to control by real and business cycles. We put a dummy to control for the credit type reclassification.[23]

The dependent variable ($endo\_var_{it}$) is the log-odds transformation of the bank NPL ratio, which shifts the support of the variable from the unit interval to the real number line. In other words, $endo\_var_{it}$ =

[23] By the end of July 2010, credit types increases from four to seven. It leads to some reclassification from mortgage loans to loans to firms and loans to microenterprises to loans to commercial loans.

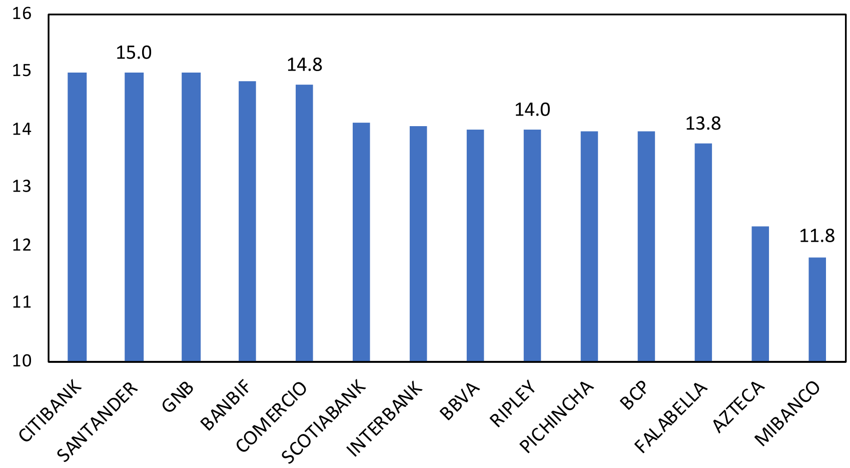


Fig. 4. Bank competition measure: "number of banks" — December 2018.
Source: SBS. Own calculations. This figures shows that heterogeneity of "number of banks" measure of competition. Number of banks = the number of banks that has the representative region for bank i at time t, calculated as the weighted average (by total loans) of the number of banks over all the regions where banks grant loans.

$ln(NPL_{it}/(100 - NPL_{it}))$, NPL_{it} is the non-performing loans ratio, defined in Eq. (1).[24] As in Jiménez et al. (2013) we include the lagged dependent variable as an explanatory variable; however, in contrast to Jiménez et al. (2013) our explanatory variables are not contemporaneous but lagged to help address reverse causality.

Our main explanatory variable (exo_var_{it-1}) is related to competition measures faced by a financial institution. To minimize simultaneity concerns, we include lagged values of the number of banks, C4 and HHI. We include, as in Jiménez et al. (2013), also the squared exo_var_{it-1}. In the model a statistically significant value of β_2 supports a nonlinear pattern. When using the number of banks as the competitive measure and if β_1 is negative and β_2 is positive, the results would support the U-shaped pattern proposed in the (Martinez-Miera and Repullo, 2010) model, which was supported in Jiménez et al. (2013). While when using C4 or HHI, the U-shaped pattern is associated with finding β_1 positive and β_2 negative.

Among the control variables ($ctrl_{it}$) we include business cycle conditions by introducing the current and lagged values of the annual real GDP growth rate (RGDPGR). We also control for the profitability of financial institutions measured by the return on assets (ROA), the size of the institution or the market share (SIZE), the foreign debt to credit ratio (FD), bonds issued by non-financial institutions to credit ratio (BOND), and the RWA to capital ratio (RWA). These latter five variables are introduced as lagged values. We solve the model with OLS and robust or clustered standard errors.

3.3. Regression results

In this subsection we present the results of the estimation of specification (2) when considering (i) only the sample of banks and (ii) a sample of all financial institutions assuming non-competition among financial institutions from different groups.

In general, the lagged endogenous variable is statistically significant, and the control variables have the expected sign. The ROA, a profitability measure is associated with low credit risk. The contemporaneous real GDP growth rate is negative and significant, while its lagged is not significant. The participation of foreign debt on loans funding (FD) has a positive and statistically significant association with credit risk only when considering banks. Also, bonds issued by non-financial institutions to credit ratio (BOND) is negatively and statistically significant associated with credit risk. Finally, the risk weighted assets to capital ratio (RWA) is positively associated with credit risk. This could be because the smaller the equity or owners' money is put in the table the higher the banks' incentives to take more risk. However, the market share of the financial institution (SIZE) is negatively associated with credit risk when considering only banks, while positively associated but less statistically significant with credit risk when considering all financial institutions.

Table 4 reports the estimation results for the model when considering only the sample banks. It shows the results of nine different regressions. For each of the three measures of bank competition, we estimate the model with no fixed effects, bank fixed effects and time fixed effects. In all cases, the lagged endogenous variable (NPL ratio) is statistically significant at 1% level with a parameter value between 0.46 to 0.81, confirming the persistence in the NPL ratio.

When using the number of banks, as the competition measure, the estimation results show an inverted U-shaped relationship between bank credit risk and loan market bank competition. This is statistically significant when we do not include any fixed effects and even when time fixed effects are included. With bank fixed effects, signs are the identical but the relationship is not longer statistically significant.[25] Interestingly, when C4 and HHI are used as competition measures, the results suggest a U-shaped relationship as suggested by Martinez-Miera and Repullo (2010) and Jiménez et al. (2013). In the case of C4 and HHI the estimates are only significant with time fixed effects.

Table 5 presents the estimation results for the model when considering all financial institutions and assuming non-competition across groups. We think this is not necessarily a very realistic assumption, but it is more realistic than assuming that in a region two financial institutions from different groups compete in the same intensity as two institutions from the same group. As in Table 4, we show the results for nine different regressions and in all cases, the lagged endogenous variable (NPL ratio) is significant at the 1% level with a parameter between 0.56 and 0.78, confirming the persistence in the NPL ratio.

When using the number of financial institutions as the competition measure, the results show an inverted U-shaped relationship between credit risk and competition. Results are significant when omitting fixed effects and when considering time fixed effects, while when considering financial institution fixed effects results suggest a U-shaped relationship

[24] Due the transformation extreme values of the NPL ratio (0 and 100) are dropped.

[25] The number of banks that maximizes the NPL ratio, as a measure of bank credit risk, is 3.5 with no fixed effects (for column 1) and 3.8 with time fixed effect (column 3).

Table 4
Banks.

exo_var	ln (# banks)			C4			HHI		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
endo_var$_{it-1}$	0.777***	0.472***	0.778***	0.846***	0.454***	0.842***	0.824***	0.451***	0.850***
exo_var$_{it-1}$	1.586*	4.091	2.432***	17.55	6.085	38.83**	6.859**	−5.266***	7.457**
exo_var$^2_{it-1}$	−0.639**	−0.732	−0.914***	−7.295	−5.191	−18.06**	−5.222	−1.413	−5.873*
ROA$_{it-1}$	−0.0386	−0.0142	−0.0433	−0.0383	−0.00866	−0.0432	−0.0342	−0.0179	−0.0376
SIZE$_{it-1}$	−0.576**	4.137*	−0.632**	−0.444**	1.423	−0.601**	−0.712***	4.896***	−0.666**
FD$_{it-1}$	1.023***	0.0716	1.003***	1.087***	0.133	1.144***	1.008***	−0.0540	1.090***
BOND$_{it-1}$	−6.303***	−3.073**		−1.299	−4.348**		−3.789***	−3.941***	
RWA$_{it-1}$	0.0368*	0.0470*	0.0316*	0.0302	0.0523**	0.0210	0.0385*	0.0634***	0.0292
RGDPGR$_{it}$	−2.256	−1.891		−2.250	−1.565		−0.355	−1.362	
RGDPGR$_{it-1}$	−0.941	−1.775		−0.185	−1.521		0.781	−1.440	
Observations	207	207	207	207	207	207	207	207	207
R-squared	0.841	0.911	0.857	0.835	0.911	0.849	0.828	0.915	0.840
F test (ρ-value)	0	3.78e−07	0	0	1.16e−10	0	0	3.61e−10	0
Bank FE	No	Yes	No	No	Yes	No	No	Yes	No
Time FE	No	No	Yes	No	No	Yes	No	No	Yes

Robust standard errors in columns (1), (4) and (7). Clustered (at bank level) standard errors in columns (2), (5) and (8). Clustered (at time level) standard errors in columns (3), (6) and (9).
***Statistically significant at 1%.
**Statistically significant at 5%.
*Statistically significant at 10%.

Table 5
All financial institutions: Non-competition across groups.

exo_var	ln (# institutions)			C4			HHI		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
endo_var$_{it-1}$	0.766***	0.564***	0.773***	0.776***	0.570***	0.782***	0.769***	0.561***	0.776***
exo_var$_{it-1}$	0.209**	−0.0671	0.167*	1.522	1.610	1.574	0.398	−0.765	0.259
exo_var$^2_{it-1}$	−0.0917**	0.0497	−0.0743*	−0.418	−0.643	−0.483	−0.433	0.496	−0.314
ROA$_{it-1}$	−0.0122**	−0.00688	−0.0129*	−0.0124**	−0.00775	−0.0132*	−0.0117**	−0.00650	−0.0125
SIZE$_{it-1}$	−0.0729	0.708**	−0.0478	−0.0755	0.459	−0.0567	−0.00118	0.718**	0.0239
FD$_{it-1}$	0.116	0.108	0.101	0.126	0.126	0.110	0.117	0.142	0.104
BOND$_{it-1}$	−1.860**	−1.384		−1.927**	−1.740**		−1.653*	−1.472*	
RWA$_{it-1}$	0.0256**	0.0421**	0.0276**	0.0279**	0.0409**	0.0293**	0.0275**	0.0413**	0.0293**
RGDPGR$_{it}$	−0.393	−1.346		−0.119	−1.393		−0.214	−1.380*	
RGDPGR$_{it-1}$	−0.846	−1.856**		−0.712	−2.064***		−0.717	−1.848**	
Observations	802	802	802	802	802	802	802	802	802
R-squared	0.795	0.855	0.799	0.795	0.855	0.799	0.794	0.855	0.798
F test (ρ-value)	0	0	0	0	0	0	0	0	0
Bank FE	No	Yes	No	No	Yes	No	No	Yes	No
Time FE	No	No	Yes	No	No	Yes	No	No	Yes

Robust standard errors in columns (1), (4) and (7). Clustered (at bank level) standard errors in columns (2), (5) and (8). Clustered (at time level) standard errors in columns (3), (6) and (9). We control for unobservable characteristics at group level.
***Statistically significant at 1%.
**Statistically significant at 5%.
*Statistically significant at 10%.

but it is not statistically significant.[26] And when using C4 and HHI, in general, results suggest a U-shaped relationship between bank credit risk and bank competition, but these results are not statistically significant. Table 10 in Appendix A reports the results when assuming competition across group. In that case, we do not find statistically significant estimates.[27]

4. Granular evidence

Our previous analysis has some shortcomings, as it cannot control for trends on demand and supply sides other than through the aggregate trends in the economy. In addition, we are not able to consider loans to firms only, as it is done in Jiménez et al. (2013).

In order to overcome these problems and avoid a biased estimate we make use of more granular data. In particular, we add another dimension to the institution-time data: "region".[28] This additional dimension allows us to control for local lending opportunities and bank level strategies. In other words, we may control for demand and supply credit shocks, respectively. Also, with this granular data we can build up loans by type (commercial loans, loans to microenterprises, mortgage and personal loans), and then we can add the dimension credit type to our regression. In particular, we focus on commercial loans and loans to

[26] The number of financial institutions that maximizes bank credit risk is 3.13 without fixed effects (column 1) and 3.07 with time fixed effects (column 3).

[27] The non-significant estimates might be evidence that at the regional level (or within a region) there is not significant competition between financial institutions from different groups.

[28] When working with granular data we omit considering the foreign market as another region and hence we consider the 25 regions (24 regions and the Constitutional Province of Callao).

microenterprises.[29] As a result, our baseline specification (2) becomes,

$$endo\_var_{icrt} = \alpha + \gamma_{rt} + \lambda_{it} + \mu_c + \beta_0 * endo\_var_{icrt-1} + \beta_1 * exo\_var_{icrt-1} + \beta_2 * exo\_var^2_{icrt-1} + \epsilon_{icrt}, \tag{3}$$

where the r subscript refers to a region, the c subscript refers to type of credit, i subscript refers to a financial institution, the t subscript refers to a sample year, γ_{rt} are the region-time fixed effects, λ_{it} are the financial institution-time fixed effects, and μ_c are the type of credit fixed effects. Next, we describe how we build up our variables at the institution-type-region-time level using the granular data.

4.1. Granular data

The source of the more granular credit data is the Credit Registry Data (RCC by its Spanish acronym), and which contains loan-level data originated in the financial system and debt classification at client-level.[30] The data is available in quarterly frequency for the 2003Q1–2010Q3 period and in monthly frequency for the 2010M10–2019M12 period. Debtors are identified by an SBS code, tax ID (RUC by its Spanish acronym) and national ID (DNI by its Spanish acronym). However, RCC does not contain information about the location of the borrower.

We use a combination of information sources about the location of firms or individuals to match credit data with geographic location, at province-region level, of borrowers. These information sources provide a location code (UBIGEO by its Spanish acronym) for each tax ID (RUC for firms) and national ID (DNI for individuals). Hence, we use information of the RUC and/or DNI of the debtor and search for the UBIGEO in the following three datasets and in the following priority order: (1) Peruvian tax administration (SUNAT by its Spanish acronym).[31] It contains data on firm Tax ID (RUC) and Location codes (UBIGEO). (2) Datos Perú.[32] It contains information on businesses identified by RUC and their geographic location. And (3) Credit Report of Debtors (RCD by its Spanish acronym).[33] It contains information on borrowers identified by RUC and/or DNI and their geographic location, UBIGEO.[34]

Once we have a UBIGEO, we use the Peruvian Bureau of Statistics' information on location of a UBIGEO in a region. We identify a RCC sample of loans and non-performing loans with their geographical location of all formal loans from financial institutions in an annual frequency.[35] As a result, we can build up a panel-data at loans type-financial institution-region-time level. We focus only on loans to firms, i.e., we focus only on two types of credit: commercial loans and loans to microenterprises.[36]

Table 6
Descriptive statistics for bank-type-region-year observations.
Source: RCC. Own elaboration. We omit financial institutions with less than ten observations, and observations with extreme value of the NPL ratio (0% and 100%).

Variables	Obs	Mean	S.D.	Minimum	Maximum
NPL_{icrt} (%)	4906	8.24	15.07	0.00	100.00
Number of $banks_{icrt}$	4906	7.71	1.96	1.00	14.00
$C4_{icrt}$	4906	0.94	0.05	0.68	1.00
HHI_{icrt}	4906	0.39	0.16	0.14	1.00

We assume that loans go to the province-region registered as the location of the borrower.[37] This helps us to correctly control by credit demand shocks with region-time fixed effects. In addition, we also assume that the loans located in a certain region are issued by an agency from the same region. This ensures that we have fair competition measures. Furthermore, we assume regional credit markets are segmented. In other words, we assume that a potential borrower located in region A cannot move to region B to ask loans. Similarly, a financial institution's branch located at region A cannot offer loans to clients located in region B.

As before the credit risk measure is given by the non-performing loans ratio, which is built using the SBS criterion (see, Eq. (1)) but this time at the institution-type-region-time level. To construct the competition measures at institution-region-time level, the approach is similar than when constructing the measures at institution-time level. However, this time instead of working with total credit, we have two types of credit, and instead of having a "representative region", there is going to be a "representative province", where regions are built up of many provinces.

For instance, the competition measure "number of institutions" for an institution i at region r and at time t is defined as the number of institutions that has the representative province for institution i, located in the region r at time t. This is calculated as the weighted average of the number of financial institutions over all the provinces in region r where institution i grants loans. The weights are given by the loans granted to each of these provinces divided by the loans granted by the institution i to region r. Also, C4 at the bank-region-time level denotes the share of the largest four financial institutions in the representative province for institution i at region r, calculated as the weighted average of the C4 over all the provinces in region r where institution i operates. Similarly, HHI at the bank-region-time level denotes Herfindahl–Hirschman index of concentration for the representative province for institution i at region r, calculated as the weighted average of the C4 over all the provinces in region r where institution i operates.

Table 6 shows the descriptive statistics of our variables at the bank-type-region-year level for the 2004–2019 period. The average NPL ratio is higher than the average of the official data at bank-time level (see Table 2). This might suggest that are small credit market regions with high NPL ratio. Also the average number of banks is smaller compared to Table 2, while the average C4 and HHI are higher. This is not surprising, since we are measuring at a smaller geographical location, and consistently the smaller the number of competitors and larger concentration.

Accordingly, Table 7 shows the same but for all financial institutions, assuming non-competition and competition across groups. Compared to Table 3 the average NPL ratio is higher, the number of banks, C4 and HHI are slightly higher.

In the following we assess the representativeness of our RCC sample and present the regression results. We first focus on banks and then on all financial institutions.

29 Due to data availability reasons we follow this shorter credit classification that was in place before July 2010 in the Credit Registry data. In consequence, commercial credit includes loans to small-size, medium-size, large-size and corporate firms. Loans to microenterprises include loans to micro-size firms.

30 This information is restricted. We thank to *Dpto. de Estadísticas monetarias* and *Dpto. de Análisis Financiero*, at the Central Bank of Peru, BCRP, for giving us access to the datasets to construct credit type-regional aggregates.

31 SUNAT information source http://www.sunat.gob.pe/descargaPRR/mrc137_padron_reducido.html accessed on 20/06/2018.

32 https://www.datosperu.org/. We resort to web scraping techniques to extract the required information from the website in June–July 2021.

33 RCD is not publicly available. It is restricted information provided by the SBS.

34 Table 11 in Appendix B shows the pairs of RUC-UBIGEO and DNI-UBIGEO from using these three datasets and the strategy followed in case of conflicts.

35 Notice that since we use national identifiers (DNI and RUC), our RCC sample does not contain loans issued to foreigners.

36 According to Figure 5 in Appendix B, the fraction of debtors (loans) and of the loans that we were able to match with a location is greater than 70% (90%) in the 2003–2019 period. It also that the larger number of clients are identified by the DNI, but the larger share of loans corresponds to clients identified by RUC.

37 It could be that the registered location is different to the one where the debtors' activities are performed. However, we assume this is an odd case.

Table 7
Descriptive statistics for financial institution-type-region-year observations.
Source: RCC. Own elaboration. We omit financial institutions with less than ten observations, and observations with extreme value of the NPL ratio (0% and 100%).

Variables	Obs	Mean	S.D.	Minimum	Maximum
NPL_{icrt} (%)	17370	9.74	14.50	0.00	100.00
Non-competition across groups					
Number of $institutions_{icrt}$	17370	7.89	2.67	1.00	14.00
$C4_{icrt}$	17370	0.95	0.07	0.59	1.00
HHI_{icrt}	17370	0.46	0.21	0.12	1.00
Competition across groups					
Number of $institutions_{icrt}$	17370	32.47	9.55	2.00	57.00
$C4_{icrt}$	17370	0.74	0.10	0.45	1.00
HHI_{icrt}	17370	0.19	0.06	0.07	0.91

4.2. Banks

Before we turn to the estimation results, we assess the representativeness of our sample of the banking system loans and hence how well it matches the characteristics of the official data from the SBS.

Figure 6 in Appendix C illustrate the representativity of our sample at the aggregate level. Figure 6.a reports our sample of commercial loans, loans to microenterprises and loans to firms (commercial and micro loans) as a share of their corresponding SBS official data. This plot reports that since 2004 our sample for both types of loans represents fairly more than 80% of the SBS official data. Figure 6.b suggests that our sample of loans to firms mimics fairly well the dynamics of the official data since 2005.[38] In addition, according to Figure 7 in Appendix C in aggregate since 2004 our samples of commercial loans and loans to microenterprises mimic fairly well the dynamics of their corresponding official non-performing loans (NPL) ratios. The poor representativeness of our 2003 sample, when working this granular data, justifies that time period analyzed spans from 2004 to 2019. This also holds when working with all financial institutions.

At the micro level, Figure 8 shows that our sample of commercial loans and loans to microenterprises mimics fairly well the credit shares that at bank-time level for shares larger than 6% of the official data. In addition, according to Figure 9 our sample resembles very well the credit growth at bank-time level. Similarly, Figure 10 our sample mimics very well the NPL ratio at bank-time level.

In Peru, there are 196 provinces. According to official SBS data and our RCC sample, 128 and 195 provinces, respectively, registered any type of credit activity, from 2004 to 2019. In addition, the official data (SBS) has 3 035 bank-region-time observations of total loans, while in our RCC sample there are 3 587 bank-region-time observations of loans to firms.[39] There are 2 497 cases where both sources report loans for the same bank-region-time. Loans that are not located in the regions where the SBS reports loans represent only the 0.55% of our RCC sample. Figure 11 in Appendix C reports at region-time level the ratio of loans to firms of our sample (RCC) and total official loans (SBS). In general, ratios are below 1 and seem to be fairly constant across time. Furthermore, according to Figure 12 on average the distribution of our sample of loans to firms across regions mimics fairly well the official distribution of total loans in the period 2004–2019. As in the official data, in our sample the larger proportion of loans are issued from branches located in Lima. The other two important credit markets are the regions of La Libertad and Arequipa.

Table 8 shows the regression results of the empirical model in Eq. (3) considering banks and only two types of loans to firms (commercial loans and loans to microenterprises) as in Martinez-Miera and Repullo (2010) but with the additional regional dimension. As usual, the coefficient of the lagged endogenous variable is significant. When considering the number of banks, as the competition measure, the results validate the inverted U-shaped relationship between bank credit risk and bank competition, independently if we control by demand (column 2) or supply (column 1) or even if we do not (column 1).[40] However, when considering C4 as the competition measure (columns 4–6), interestingly, results validate the U-shaped relationship as in Martinez-Miera and Repullo (2010). Results are inconclusive and not significant with the HHI as the competition measure.

Robustness provided in Table 12 Appendix D shows that our results are consistent across different specifications. Statistical significance does not change when we omit extreme value observations of the NPL ratio.[41] However, if we exclude the metropolitan area (i.e., the regions Lima and Callao), results become less significant when considering the number of banks as the competition measure. When considering only commercial loans, estimates are no longer significant when using the number of banks as the competition measure.

4.3. Financial system

Given that at the regional level the role of nonbank financial institutions in lending activities becomes more important, in this section we consider both bank and nonbank financial institutions. Thus, we consider the five financial groups that operate in the peruvian credit market: banks, *empresas financieras*, CMACs, CRACs and EDPYMEs.

We know already how bank loans sample matches the official data, so we focus on the rest of the groups. According to Figure 13 in Appendix C, since 2004 our RCC sample for the four non-bank groups and for commercial loans (loans to microenterprises) have represented no less than 80% (70%) of the SBS official data. We find also that the correlation between our RCC sample and official data of the growth of loans to firms, commercial loans and loans to microenterprises for any financial group in the 2005–2019 period is higher than 0.99. Similarly, the correlation between our RCC sample and official data of the NPL ratio of commercial loans and loans to microenterprises for any financial group in the 2004–2019 is higher than 0.85. So, our RCC sample matches fairly well the dynamics of the credit growth and NPL ratio of the SBS official data.

Table 9 reports the regression results of the empirical model, Eq. (3), when considering all financial institutions, two types of loans to firms (commercial credit and loans to microenterprises), and consider (i) non-competition and (ii) competition across groups within a province.

In the case of non-competition across groups, when considering the C4 and HHI as our competition measures, results validate an inverted U-shaped relationship between bank competition and bank credit risk. For the case of HHI coefficient estimates are significant even if we control by demand or supply shocks. However, for the case of C4 estimates are still significant only when controlling by demand shocks. When considering the number of banks results are not clear and significant. When we omit extreme values of the NPL ratio, results hold; but does not hold if we exclude the Metropolitan Area (see Table 13 in Appendix E).

Interestingly, in the case of competition across groups, according to Table 9 when considering the number of financial institutions as our competition measure results suggest an inverted U-shaped relationship between bank competition and bank credit risk but estimates are not statistically significant. And when considering C4 results validates the U-shaped relationship found in Jiménez et al. (2013). In this case results regarding HHI are neither significant nor conclusive. Results hold

[38] This is because the 2004 growth rate contains 2004 information which according to Figure 6.b does not represent well the official data.

[39] We use total loans since from the SBS official data there is not available credit by type at region level.

[40] The number of banks that maximizes bank credit risk is 8.98 (column 1), 9.88 (column 2) and 9.72 (column 3)

[41] We consider only 0.05%<NPL<94% and then we drop 86 observations.

Table 8
Granular estimation: Banks.

exo_var	ln (# banks)			C4			HHI		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$endo\_var_{icrt-1}$	0.446***	0.449***	0.407***	0.448***	0.452***	0.408***	0.444***	0.449***	0.405***
exo_var_{icrt-1}	2.736**	3.180***	2.515**	49.68***	60.29***	59.38***	0.0981	0.533	−0.110
$exo\_var^2_{icrt-1}$	−0.623**	−0.694**	−0.553*	−27.09***	−32.92***	−32.48***	−1.062	−1.395	−0.824
Observations	4,257	4,257	4,245	4,257	4,257	4,245	4,257	4,257	4,245
R-squared	0.312	0.385	0.381	0.310	0.383	0.380	0.313	0.384	0.382
F test (ρ-value)	0	0	0	0	0	0	0	0	0
Region Time FE	No	Yes	No	No	Yes	No	No	Yes	No
Bank Time FE	No	No	Yes	No	No	Yes	No	No	Yes
Bank FE	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Region FE	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
Time FE	Yes	No	No	Yes	No	No	Yes	No	No
Type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in columns (1), (4) and (7). Clustered (at bank level) standard errors in columns (2), (5) and (8). Clustered (at time level) standard errors in columns (3), (6) and (9). In all regressions we control by $SIZE_{ict,ct}$, $SIZE_{icrt,ict}$ and $SIZE_{icrt,ct}$. $SIZE_{x,y} = credit_x/credit_y$. For example, $credit_{icrt}$ is all credit of institution i of the type c in the region r at the year t. We control for unobservable characteristics at group level.
***Statistically significant at 1%.
**Statistically significant at 5%.
*Statistically significant at 10%.

Table 9
All financial institutions.

exo_var	ln (# institutions)			C4			HHI		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(i) Non-competition across groups									
$endo\_var_{icrt-1}$	0.454***	0.456***	0.433***	0.453***	0.455***	0.433***	0.454***	0.456***	0.433***
exo_var_{icrt-1}	−0.0427	−0.0476	0.104	−5.705***	−5.620**	−2.114	−0.514*	−0.573**	−0.623**
$exo\_var^2_{icrt-1}$	0.00758	0.00854	−0.0172	3.309***	3.244**	1.162	0.462**	0.519**	0.436*
Observations	14,298	14,298	14,279	14,298	14,298	14,279	14,298	14,298	14,279
R-squared	0.358	0.382	0.439	0.358	0.382	0.439	0.358	0.382	0.439
F test (ρ-value)	0	0	0	0	0	0	0	0	0
(ii) Competition across groups									
$endo\_var_{icrt-1}$	0.449***	0.451***	0.429***	0.453***	0.453***	0.432***	0.453***	0.454***	0.432***
exo_var_{icrt-1}	0.582	0.855	0.839	3.328***	4.880***	3.811***	−0.666	−0.812	−0.175
$exo\_var^2_{icrt-1}$	−0.0395	−0.0809	−0.0836	−2.584***	−3.705***	−2.908***	−0.231	0.171	−1.240
Observations	14,298	14,298	14,279	14,298	14,298	14,279	14,298	14,298	14,279
R-squared	0.361	0.384	0.441	0.359	0.383	0.440	0.359	0.382	0.440
F test (ρ-value)	0	0	0	0	0	0	0	0	0
Region Time FE	No	Yes	No	No	Yes	No	No	Yes	No
Bank Time FE	No	No	Yes	No	No	Yes	No	No	Yes
Bank FE	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Region FE	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
Time FE	Yes	No	No	Yes	No	No	Yes	No	No
Type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in columns (1), (4) and (7). Clustered (at bank level) standard errors in columns (2), (5) and (8). Clustered (at time level) standard errors in columns (3), (6) and (9). In all regressions we control by $SIZE_{ict,ct}$, $SIZE_{icrt,ict}$ and $SIZE_{icrt,ct}$. $SIZE_{x,y} = credit_x/credit_y$. For example, $credit_{icrt}$ is all credit of institution i of the type c in the region r at the year t. We control for unobservable characteristics at group level.
***Statistically significant at 1%.
**Statistically significant at 5%.
*Statistically significant at 10%.

when omitting extreme values of the NPL ratio and when excluding the Metropolitan Area (see Table 14 in Appendix E).

Notice that in contrast to our regressions at the institution-time level in Section 3.3, if we assume competition across groups, this time results can validate the inverted U-shaped relationship when considering the number of financial institutions as our competition measure and can validate the U-shaped relationship when considering the C4 and HHI, although the former estimates are not statistically significant.

These findings could be a result of the granular analysis. In other words, two financial institutions are more likely to compete if they are situated in the same province rather than the same region. This might suggest that assuming competition across groups when estimating the model at region-financial firm-time level makes relatively more sense than assuming non-competition across groups.[42]

In short, when using the number of financial institutions, we found evidence of an inverted U-shaped relationship between competition and credit risk in an emerging economy as Peru, in contrast to what is

[42] We believe also that the various conflicting effects of competition on risk-taking may be visible in data if all financial institutions are mixed at a more granular level and the assumption of no competition across groups is imposed. In Peru, credit markets are segmented and at varying levels of competition, and each financial group serves a different one. By combining all financial institutions at once within a specification, we may be mixing all of the various effects at once and being less successful to uncover the true relationship.

found in Martinez-Miera and Repullo (2010) for an advanced economy as Spain. And when using the C4 or the HHI we find evidence of a U-shaped relationship between competition and credit risk as in Martinez-Miera and Repullo (2010).

5. Conclusions

In this paper we can conclude that in the Peruvian financial system there is evidence of a nonlinear relationship between competition and credit risk. In particular, when considering the number of banks as our competition measure, in contrast to Martinez-Miera and Repullo (2010), we find an inverted U-shaped relationship. This result holds true whether only banks or all financial institutions are studied at the bank-time level (assuming non-competition across groups). In addition, the result is robust to granular data analysis for banks, where we can control for supply and demand factors.

However, when using the C4 or HHI, we might find a U-shaped relationship, when studying only banks or all financial institutions. This result is statistically more important at the granular level for banks and for all institutions (assuming competition across groups).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jfs.2023.101119.

References

Agur, Itai, Demertzis, Maria, 2012. Excessive Bank Risk Taking and Monetary Policy. ECB WP, p. 1457.

Agur, Itai, Demertzis, Maria, 2019. Will macroprudential policy, counteract monetary policy's effects on financial stability? North Am. J. Econ. 48, 66–75.

Allen, F., Carletti, E., Marquez, R., 2011. Credit market competition and capital regulation. Rev. Financ. Stud. 24 (4), 983–1018.

Allen, F., Gale, D., 2004. Competition and financial stability. J. Money Credit Bank. 36, 453–480.

Beck, Thorsten, Demirguc-Kunt, Asli, Levine, Ross, 2006. Bank concentration, competition, and crises: First results. J. Bank. Financ. 30, 1581–1603.

Beck, Thorsten, Demirguc-Kunt, Asli, Levine, Ross, 2007. Bank concentration and fragility: Impacts and mechanics. In: Carey, M., Stulz, R. (Eds.), The Risks of Financial Institutions. National Bureau of Economic Research, Cambridge, MA, pp. 193–231.

Berger Allen, N., Klapper, Leora F., Turk-Ariss, Rima, 2017. Bank competition and financial stability. In: Bikker, Jacob A., Spierdijk, Laura (Eds.), Handbook of Competition in Banking and Finance. Edward Elgar Publishing, pp. 185–204, chapter 10.

Bolt, Wilko, Tieman, Alexander F., 2004. Bank competition, risk and regulation. Scand. J. Econ. 104 (4), 783–804.

Boot, A., Greenbaum, S., 1993. Bank regulation, reputation and rents: Theory and policy implications. In: Mayer, Colin, Vives, Xavier (Eds.), Capital Markets and Financial Intermediation. University Press, Cambridge, pp. 262–285.

Boyd, John H., De Nicoló, Gianni, 2005. The theory of bank risk taking and competition revisited. J. Finance 60 (3), 1329–1343.

Boyd, J., De Nicolo, G., Jalal, A.M., 2006. Bank risk taking and competition revisited: new theory and evidence. IMF Working paper, WP/06/297.

Caminal, R., Matutes, C., 2002. Market power and banking failures. Int. J. Ind. Organ. 20 (9), 1341–1361.

Claessens, S., Laeven, L., 2004. What drives bank competition? Some international evidence. J. Money Credit Bank. 36, 563–583.

Clark, E., Radić, N., Sharipova, A., 2018. Bank competition and stability in the CIS markets. J. Int. Financial Mark. Inst. Money 54 (C), 190–203.

De Nicoló, G., Loukoianova, Elena, 2007. Bank ownership, market structure, and risk. IMF Working paper, WP/07/215.

Dell'Ariccia, G., Laeven, L., Marquez, R., 2014. Real interest rates, leverage, and bank risk-taking. J. Econom. Theory 149 (1), 65–99.

Demsetz, R., Saidenberg, M.R., Strahan, P.E., 1996. Banks with something to lose: The disciplinary role of franchise value. In: Federal Reserve Bank of New York Economic Policy Review. Vol. 2, (2), pp. 1–14.

Fernández, O.R., Garza-García, J.G., 2015. The relationship between bank competition and financial stability: a case study of the Mexican banking industry. In: Ensayos Revista de Economia. (1), Universidad Autonoma de Nuevo Leon, Facultad de Economia, pp. 103–120.

Fungacova, Z., Weill, L., 2010. How market power influences bank failures evidence from Russia. Working Papers of LaRGE Research Center 2010-08, Laboratoire de Recherche en Gestion et Economie (LaRGE), Université de Strasbourg.

Hellman, T., Murdoch, K., Stiglitz, J.E., 2000. Liberalization, moral hazard in banking and prudential regulation: Are capital requirements enough? Amer. Econ. Rev. 90, 147–165.

Jiménez, Gabriel, Lopez, Jose A., Saurina, Jesús, 2013. How does competition affect bank risk-taking? J. Financial Stab. 9, 185–195.

Keeley, M., 1990. Deposit insurance, risk and market power in banking. Amer. Econ. Rev. 80 (5), 1183–1200.

Liu, H., Wilson, J.O.S., 2013. Competition and risk in Japanese banking. Eur. J. Finance 19 (1), 1–18.

Marcus, A.J., 1984. Deregulation and bank financial policy. J. Bank. Financ. 8, 557–565.

Martinez-Miera, David, Repullo, Rafael, 2010. Does competition reduce the risk of bank failure? Rev. Financ. Stud. 23 (10), 3638–3664.

Repullo, Rafael, 2004. Capital requirements, market power, and risk-taking in banking. J. Financial Intermed. 13, 156–182.

Schaeck, K., Cihák, M., 2014. Competition, efficiency, and stability in banking. Financial Manag. 43 (1), 215–241.

Schaeck, K., Cihák, M., Wolfe, S., 2009. Are competitive banking systems more stable? J. Money Credit Bank. 41, 711–734.

Smith, B., 1984. Private information, deposit interest rates, and the 'stability' of the banking system. J. Monetary Econ. 14, 293–317.

Turk-Ariss, R., 2010. On the implications of market power in banking: evidence from developing countries. J. Bank. Financ. 34 (2010), 765–775.

Zigraiova, D., Havranek, T., 2015. Bank competition and financial stability: Much ado about nothing? J. Econ. Surv. 30 (5), 944–981.