BANCO CENTRAL DE RESERVA DEL PERÚ

Bank Competition and Risk-Taking

Jorge Pozo* y Youel Rojas*

* Banco Central de Reserva del Perú.

DT. N°. 2021-007
Serie de Documentos de Trabajo
Working Paper series
Diciembre 2021

Bank Competition and Risk-Taking*

Jorge Pozo† & Youel Rojas‡

October, 2021

Abstract

This paper studies empirically the relationship between competition in the loan market and risk-taking 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 finds evidence that supports this nonlinear relationship in a developed economy as Spain. In contrast, we find empirical evidence of an inverted U-shaped relationship between competition and risk-taking in Peru, an emerging economy. We develop a theoretical model to rationalize our empirical findings. One possible explanation for our findings is the greater presence of borrowing constraints faced by entrepreneurs, which is a common feature in emerging economies.

Keywords: Competition, bank risk-taking, financial stability.

JEL Classification: E44, G21, L11

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 that 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.

*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.

†Researcher at Central Reserve Bank of Peru. Email: jorge.pozo@bcrp.gob.pe

‡Researcher at Central Reserve Bank of Peru. Email: youel.rojas@bcrp.gob.pe

To provide a rationality for our empirical findings, we build up a simple two-period closed economy model with banks and two types of entrepreneurs that invest in risky projects and face borrowing constraints. This latter feature, which is not modeled in Martinez-Miera and Repullo (2010), is to better characterize an emerging economy, where entrepreneurs seem to be more financially constrained than in an advanced economy (see, Blalock et al., 2008). Entrepreneurs differ in their strategic decisions. One type (entrepreneur Type 1) chooses the optimal loan size, where the probability distribution of the project's uncertainty source is fixed but there is a binding borrowing constraint if it takes too much lending; the other type (entrepreneur Type 2) chooses the failure probability of their projects, but loan size is fixed.

The model is able to describe three phases of the relationship between bank competition and bank risk-taking. In the first one (low competition), there is a negative relationship explained by the fact that the lower cost of loans (due to increasing competition) reduces entrepreneurs Type 2' incentives to take risk (*risk-shifting effect*), which in turn reduces bank default probability. In the second one (mid competition), there is a positive relationship. This is essentially explained due to that a lower cost of loans increases loan demand of entrepreneurs Type 1, and hence it increases entrepreneurs Type 1 default probability, which in turn increases bank default probability. In this phase, this effect dominates the *risk-shifting effect*. And finally, in the third phase (high competition), there is a negative relationship. This is because the borrowing constraint of entrepreneurs Type 1 starts binding and hence default probability of entrepreneurs Type 1 does not change. In this case the *risk-shifting effect* drives this negative relationship. Interestingly, the model is able to report a U-shaped relationship when moving through the first and second phases, while it reports an inverted U-shaped relationship when moving across the second and third phases.

With respect to the empirical part of the paper, 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 estimate 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 (SBS), while for the second specifications we use more granular data from the Credit Registry Data (RCC), which is restricted information. As in Jiménez et al. (2013), the main measure of competition is an indicator that reflects the number of relevant competitors that a financial institution faces. We also consider other proxies for competition measures as concentration indicators: loan share of four-largest financial institutions and the Herfindahl-Hirschman index. And our measure of risk-taking is the non-performing loan ratio, which is taken from the SBS or is built using the granular data from the RCC.

In the model with bank-time level data for the 2003-2019 period, results suggest that for Peruvian banks there is an inverted U-shaped relationship between competition and risk-taking, unlike Jiménez et al. (2013), when using the number of banks as our competition measure. Results are robust if we include non-bank financial institutions assuming non-competition across different groups of financial institutions.

Since the results under financial institution-time level specification cannot control for omitted variables that may affect the dynamics of the relationship between competition and risk-taking (e.g., changes in business opportunities and risk profiles at the departmental level and/or market strategies and diversification of financial institutions over time) 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 within-region and within-bank estimators. Furthermore, only loans to firms (commercial loans and loans to microenterprises) are considered and it is assumed that there is competition between the different groups of financial institutions.[1] The results indicate the significant existence of an inverted U-shaped relationship between competition and risk-taking when considering the number of banks as the competition measure. This is robust even when controlling for supply or demands shocks. Results are also robust when non-bank financial institutions are included; however, estimates are not statistically significant. We believe, as demonstrated in the theoretical model, 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. 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.

As a result, in contrast to Jiménez et al. (2013), we find empirical evidence in an emerging economy as Peru of an inverted U-shaped relationship between competition and risk-taking when considering the number of banks as our competition measure. This result holds when studying only banks or all financial institutions (assuming non-competition across groups) at the bank-time level. In addition, this result is robust when considering granular data for banks and hence being able to work at the region-bank-time level, and even when controlling for supply and demands shocks. And, according to our theoretical model, a possible explanation for this feature is the relatively stronger borrowing constraint faced by entrepreneurs in an emerging economy as Peru.

In Latin America, the Peruvian banking sector is more concentrated than in other countries. Only the banking system in Colombia is more concentrated than in Peru. Table 1 shows different measures of concentration and competition. For instance, in 2016 the share of the three largest banks assets in Peru was 71.9%, while in Chile it was 43.2%. The same is observed with the 5-bank asset concentration measure. Also, in

[1]The latter assumption is because we believe there is stronger competition across different groups of financial institutions at a more credit local market, i.e., we expect more competition at the provincial level rather than at the regional level.

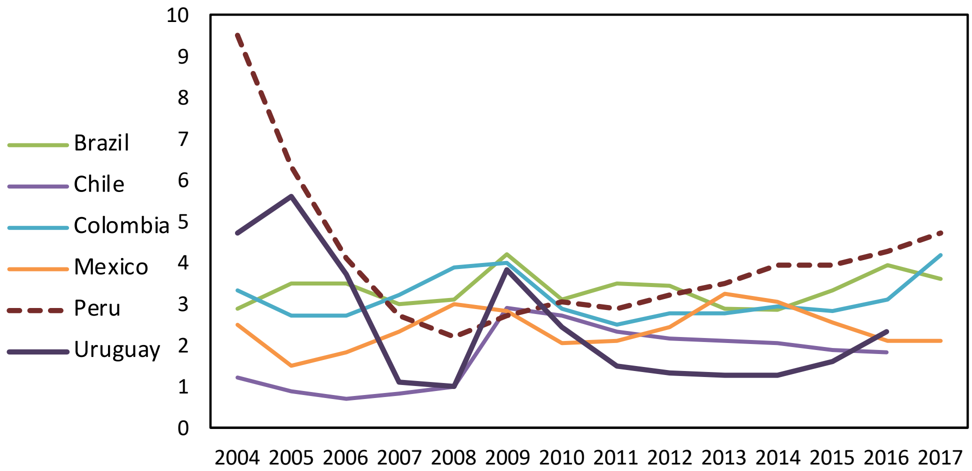
Table 1: *Bank competition and concentration in Latin America*

	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

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.

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. Relative to other emerging market economies or advanced economies, the Peruvian banking sector shows high levels of concentration and market power. In terms of financial stability, after the 2008 global financial crisis, in Peru, the bank nonperforming loans have been increasing steadily, while the tendency is not clear in other countries in Latin America (see figure 1).

Figure 1: *Bank non-performing loans to gross loans (%) in Latin America*

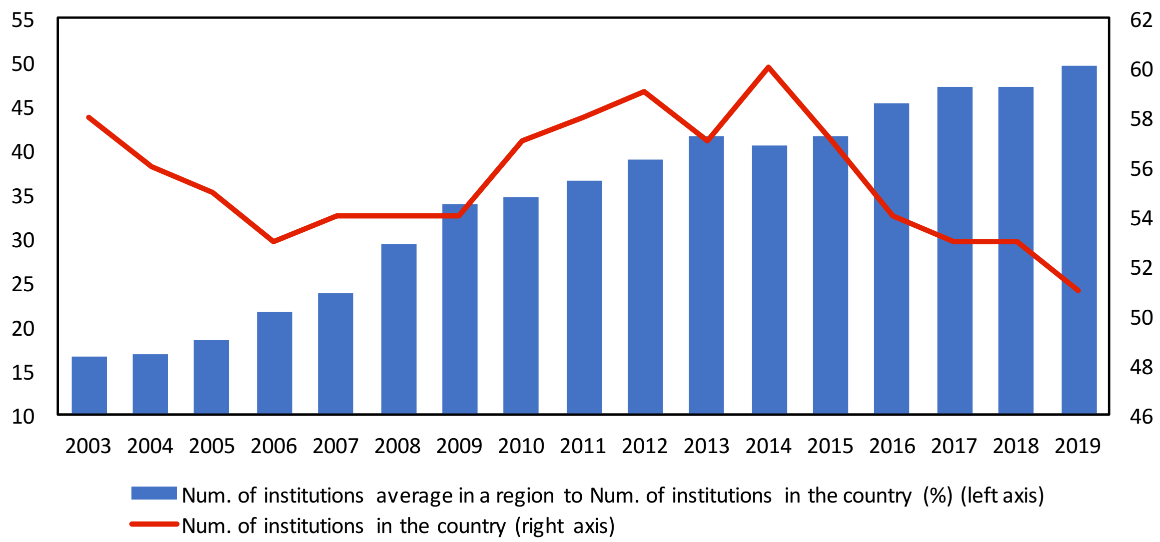


Source: Global Financial Development.

For the purpose of this paper, our definition of the Peruvian financial system includes 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). In 2018, 53 financial institutions that provide loans to the private sector constitute the Peruvian financial system. 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.[2] Even though from a country perspective banks are relatively more important, this is not necessarily true within a region.[3] There are regions where the lending role of nonbank financial institutions becomes relatively more important. Thus, in our analysis we focus not only on banks but also on nonbank financial institutions when studying the lending market.

Figure 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 a large number of 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.

Figure 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.

[2]The Peruvian financial system also includes other more specialized institutions, however, since the participation in the credit market is very small, we omit them.

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

Because of these features, the regional composition of the Peruvian financial system provides us with an interesting case to study the relationship between lending competition and bank risk-taking.

The remainder of this paper is partitioned as follows. Section 2 presents the literature review. Section 3 develop the theoretical model. Section 4 shows the data, the empirical model and the empirical results. Section 5 presents the granular assessment results. Finally, section 6 concludes.

2 Literature Review

This paper is 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 et al. (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.[4]

Boyd et al. (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 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 et al. (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 proba-

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

bility. They identify a *risk-shifting effect*, which is the one described in Boyd et al. (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 also using granular data. And, in contrast to Jiménez et al. (2013), we find an inverted U-shaped relationship between bank competition and risk-taking.

A key friction in our theoretical model, that departs from Martinez-Miera and Repullo (2010), is the financial constraint faced by entrepreneurs that aim to capture the idea that entrepreneurs in emerging economies are more financially constrained, which drives the presence of an inverted U-shaped relationship between bank competition and risk-taking in the model, as implied by the literature. Related to this, Blalock et al. (2008) suggest that domestic firms, in an emerging economy as Indonesia, are more financially constrained than foreign-owned firms. Beck (2007) state that the weakness in the financial system of many developing countries exacerbates its finding that small and medium enterprises are more financially constrained than large enterprises. Poschke (2018) finds that the mean of the firm size distribution is larger in rich countries. This together with the argument of Beck (2007) might imply that emerging economies have a large proportion of financially constrained firms.

3 Theoretical Model

The motivation of this theoretical part is to rationalize the relationship between bank competition and bank risk-taking in an emerging economy where the fraction of entrepreneurs financially constrained is larger than in advanced economies.

We develop a simple two-period closed economy model with two types of entrepreneurs, identical banks, households and government.[5] All agents are risk-neutral and households own firms, managed by entrepreneurs, and banks[6]. Entrepreneurs demand bank loans

[5]Notice that even though we refer to banks in the model, the results also hold for any financial institution that is going to be studied in the empirical section.

[6]Assuming agents are risk neutral avoids additional complications about asset pricing in this economy that are beyond the scope of the paper.

to invest in risky projects. Banks supply loans to entrepreneurs and capture deposits from households. We assume banks compete for loans à la Cournot. The novelty of this framework is that it includes two types of entrepreneurs that differ on the strategic investment decision that they have to take. The first type (Type 1) chooses privately the optimal investment level for a specific project, and hence they choose the privately optimal demand of bank loans as in Pozo (2019), where the probability distribution of the uncertainty source of the project is fixed.[7] The second type (Type 2) chooses the probability of failure of their projects as in Martinez-Miera and Repullo (2010) where the investment size (or bank loan size) of each individual entrepreneur is fixed. In contrast to Martinez-Miera and Repullo (2010), the returns of the projects faced by entrepreneurs Type 2 are perfectly correlated.

For simplicity, we assume entrepreneurs cannot borrow from another entrepreneur. We assume all entrepreneurs face a borrowing constraint on bank loans. In addition, entrepreneurs and banks have limited liability, and bank deposits are fully insured by the government. Households consume and save only through bank deposits and government collects sum lump taxes to fund her activity as a deposit insurer.

In general, the timing of the model is as follows, at $t = 0$ households make bank deposits, banks decide how much to lend to entrepreneurs, and entrepreneurs take investment decisions. At the beginning of $t = 1$ payoffs of entrepreneurs' investments are realized and bank loans are repaid.

The problem of the household is straightforward. Let say that $\bar{r}$ is the required gross return on bank deposits D from $t = 0$ to $t = 1$ and agreed at $t = 0$. Since deposits are fully protected by deposit insurance, it holds that $\bar{r}$=r, where r denotes the risk-free interest rate. Since households are risk-neutral, the equilibrium condition that avoids a corner solution yields r=$\frac{1}{\beta}$, where β is the household's exogenous discount factor. Hence, households are indifferent to the amount they deposit in banks. It follows that the deposit supply faced by banks is perfectly elastic at the interest rate of r.

Next, we describe the behavior of entrepreneurs Type 1 and Type 2 and banks, and present the equilibrium conditions.

3.1 Entrepreneurs Type 1

There is a continuum of measure one of identical entrepreneurs Type 1. They invest K in a project that gives them a return of ZK^α plus the leftover capital $(1-\delta)K$, where Z is the only source of aggregate uncertainty and hence is drawn from a distribution where $ln(Z) \sim \mathcal{N}(\mu, \sigma^2)$. Investments are funded by bank loans and entrepreneurs' exogenous

[7] In contrast to Pozo (2019) we do not model banks, but entrepreneurs, and we do not open the economy, for simplicity and since doing so in this framework does not affect our main results.

equity, N_f^1, i.e.,

$$K = L^1 + N_f^1, \tag{1}$$

where superscript 1 refers to Type 1. Entrepreneurs have the following borrowing constraint,

$$L^1 \leq \bar{L}^1. \tag{2}$$

This constraint captures the financial frictions that might exist between entrepreneurs and banks due to some informational frictions that as suggested in the literature (see, Blalock et al., 2008; Beck, 2007; Poschke, 2018) are relatively more important in emerging economies than in advanced economies. These informational frictions can be associated with monitoring cost, limited enforcement and asymmetric information. We assume, for simplicity, that $\bar{L}^1$ is exogenous.

Entrepreneurs borrow from banks at the market lending rate $r^{L,1}$.[8] An entrepreneur's expected discounted profits are,

$$\Pi^f = \beta \mathbb{E} \left\{ max \left[(1-\delta)K + ZK^\alpha - r^{L,1}L^1, 0 \right] \right\}. \tag{3}$$

An entrepreneur defaults if revenues $(1-\delta)K + ZK^\alpha$ are smaller than obligations $r^{L,1}L^1$. Indeed, we calibrate the model so in equilibrium entrepreneurs Type 1's default probability is positive. Hence, there exist a $Z^* > 0$ so that entrepreneurs' profits are zero, i.e.,

$$(1-\delta)K + Z^*K^\alpha - r^{L,1}L^1 = 0. \tag{4}$$

Hence, entrepreneur default probability is given by,

$$p_f^1 = F(Z^*),$$

where F is the cumulative distribution function of the random variable Z. Entrepreneurs optimally choose L^1 to maximize Π^f subject to entrepreneur balance sheet equation (1) and to the borrowing constraint (2) taking as given the lending rate $r^{L,1}$. For convenience, we rewrite (3) as,

$$\Pi^f = \beta \int_{Z^*}^{+\infty} ((1-\delta)K + ZK^\alpha - r^{L,1}L^1)dF(Z).$$

The first order condition with respect to L^1 yields,[9]

$$\beta \int_{Z^*}^{+\infty} (1-\delta + Z\alpha K^{\alpha-1} - r^{L,1})dF(Z) - ((1-\delta)K + Z^*K^\alpha - r^{L,1}L^1)f(Z^*)\frac{\partial Z^*}{\partial L^1} - \mu = 0.$$

[8] Note that the market lending rate of loans to entrepreneurs Type 1 and Type 2 are not necessarily identical since in the general equilibrium they depend on the risk level of the investment activities performed by entrepreneurs Type 1 and Type 2, which are not necessarily identical.

[9] Using Leibniz's rule for differentiation.

where μ is the Lagrange multiplier associated with the borrowing constraint (2). If the borrowing constraint binds, i.e., $L^1 = \bar{L}^1$, the bank loan demand is perfectly inelastic at $\bar{L}^1$.

As we will see, in our baseline calibration we start by assuming the borrowing constraint does not bind. However, it becomes binding if economic conditions change, e.g., after higher competition that pushes down the lending rate. In what follows, we focus on a non-binding borrowing constraint. In that case using (4) the first order condition with respect to L^1 becomes,

$$\beta \int_{Z^*}^{+\infty} (1 - \delta + Z\alpha K^{\alpha-1} - r^{L,1}) dF(Z) = 0, \tag{5}$$

which using (1) can be rewritten as,

$$r^{L,1}(L^1) = 1 - \delta + \alpha Z^+ (L^1 + N_f^1)^{\alpha-1}, \tag{6}$$

where $Z^+ = \mathbb{E}\{Z|Z > Z^*\}$ is the expected aggregate shock conditional on entrepreneurs not going bankrupt. Notice that the interaction of the limited liability and deposit insurance overestimate the return of entrepreneurs' investments (see, Collard et al., 2017; Pozo, 2019), which creates an excessive risk-taking. Let $r^{L,1}(L^1)$ denote the inverse loan demand function of entrepreneurs Type 1. Since it is not easy to analytically prove, numerically we find that[10]

$$\frac{\partial L^1}{\partial r^{L,1}} < 0 \qquad \textit{and} \qquad \frac{\partial p_f^1}{\partial r^{L,1}} \leq 0.$$

In other words, we claim that the slope of the loan aggregate demand curve of entrepreneurs Type 1 is negative and that the lower the lending rate the higher the default probability of entrepreneurs Type 1. The intuition is that according to (6), for a given Z^+, the lower lending rate, the higher the demand for bank loans. This is known as the *credit channel effect*. This in turn increases entrepreneur default probability p_f^1, which in turn increases the conditional expected return of investments, through the higher Z^+, and hence raises the demand of bank loans. This latter is called the *excessive risk-taking channel effect*. The higher bank loans raise entrepreneurs' default probability and so on.

If entrepreneurs do not default, bank receives the agreed repayment, $r^{L,1}L^1$, otherwise they receive, $(1-\delta)K + ZK^{\alpha}$. We define a recovery rate x^1, which measure the fraction of the promised repayment that is effectively repaid to banks. The recovery ratio is one

[10] The analytical expression is found in Pozo (2019). It shows that this result is qualitatively robust. Also, it shows that $\frac{\partial p_f^1}{\partial r^{L,1}} = 0$ only if entrepreneurs are fully leverage, i.e., $N_f = 0$.

if entrepreneurs do not default; otherwise, it becomes $\frac{(1-\delta)K+ZK^{\alpha}}{r^{L,1}L^{1}}$. In general,

$$x^{1}(L^{1},r^{L,1}(L^{1}),Z)=min\left\{1,\frac{(1-\delta)(L^{1}+N_{f}^{1})+Z(L^{1}+N_{f}^{1})^{\alpha}}{r^{L,1}L^{1}}\right\}. \tag{7}$$

3.2 Entrepreneurs Type 2

There is a continuous of penniless entrepreneurs Type 2 characterized by a continuous distribution of reservation utilities. Let $G(u)$ denotes the measure of entrepreneurs that have reservation utility less than or equal to u. Entrepreneurs are identical except for their reservation utility level. They demand loans to finance their investment activities. The entrepreneur j has a project that requires a unit of investment and yields the following payoff,

$$R(p_{fj}^{2})=\begin{cases}(1+\theta(p_{fj}^{2}), & \text{with probability} \quad 1-p_{fj}^{2}\\ 1-\lambda, & \text{with probability} \quad p_{fj}^{2}\end{cases}$$

where p_{fj}^{2} is the probability of failure of the project and is chosen by entrepreneur j. Recall, we assume payoffs across entrepreneurs Type 2 are no correlated and these are neither correlated with the aggregate shock face by entrepreneurs Type 1. We assume $\theta(p_{fj}^{2})$ is increasing on p_{fj}^{2}. Also, we assume $\theta(p_{fj}^{2})$ is concave on p_{fj}^{2} in order to get an interior solution. Investments are funded by bank loans and entrepreneur's exogenous equity, n_f, i.e.,

$$1=l^{2}+n_{f} \tag{8}$$

So, bank loans to an individual entrepreneur Type 2 and entrepreneur's leverage, defined as $\phi_{f}^{2}=1/(1-l^{2})$, are exogenous. Entrepreneurs face the following borrowing constraint,

$$l^{2}\leq\bar{l}^{2}. \tag{9}$$

For simplicity, we assume $\bar{l}^{2}$ is exogenous. We assume that the constraint binds. Actually, whether the constraint binds or not for entrepreneurs Type 2 is not qualitatively important when assessing the impact of bank competition on entrepreneurs Type 2's risk-taking decision and on aggregate demand of loans. This is because whether the constraint binds or not, the individual demand of loans is going to be fixed and hence independent of project failure probability decision.

Entrepreneur borrow from banks at the market lending interest rate $r^{L,2}$. Since entrepreneurs only differ in their reservation utilities, the solution $p_{f}^{2}(r^{L,2})$ and hence $u(r^{L,2})$ do not depend on j. Entrepreneur discounted profits are,

$$\pi^{f}=\beta\mathbb{E}\{max[R(p_{f}^{2})-r^{L,2}l^{2},0]\}, \tag{10}$$

where expectations are taken with respect to $R(p_{f}^{2})$. Entrepreneur defaults if revenues

$R(p_f)$ are smaller than obligations $r^{L,2}l$. Indeed, we parametrize the model so that if $R(p_f) = 1 - \lambda$, entrepreneur defaults, and if $R(p_f^2) = 1 + \theta(p_f^2)$, entrepreneur does not default. This is, the entrepreneur default probability is the same that the project failure probability, p_f^2, chosen by entrepreneur. Entrepreneur optimally chooses p_f^2 to maximize π^f taking as given the lending rate $r^{L,2}$. For convenience, we rewrite (10) as,

$$\pi^f = \beta(1 - p_f^2)(1 + \theta(p_f^2) - r^{L,2}l^2).$$

The first order condition with respect to p_f^2 yields,

$$-(1 + \theta(p_f^2) - r^{L,2}\frac{\phi_f^2 - 1}{\phi_f^2}) + (1 - p_f^2)\theta'(p_f^2) = 0. \tag{11}$$

Also since $\theta'(p_f^2) > 0$ and $\theta''(p_f^2) \leq 0$, the partial derivative of the first order condition with respect to $r^{L,2}$ yields,

$$\frac{\partial p_f^2}{\partial r^{L,2}} = \frac{\phi_f^2 - 1}{\phi_f^2})\frac{1}{2\theta'(p_f^2) - (1 - p_f^2)\theta''(p_f^2)} > 0. \tag{12}$$

Hence, the lower the lending rate, the lower the failure probability chosen by the entrepreneur Type 2. This positive effect of lending rate on entrepreneur default probability is known in Martinez-Miera and Repullo (2010) as the *risk-shifting effect.* A lower cost of funding increases profits when the entrepreneur does not default, which in turn reduces entrepreneurs' incentives to take risk.

Entrepreneur j undertake the project if reservation utility u_j is smaller or equal to expected discounted payoff $\pi^f(r^{L,2})$. Hence, the measure of entrepreneurs that want to borrow from banks at the rate $r^{L,2}$ is given by $G(\pi^f(r^{L,2}))$. Since each one requires $1 - n_f$ unit of loans, the loan demand function is,

$$L^2(r^{L,2}) = (1 - n_f)G(\pi^f(r^{L,2})). \tag{13}$$

Clearly, for $0 \leq r^{L,2} \leq 1 + \theta(p_f^2)$, the slope of the aggregate demand of bank loans of entrepreneurs Type 2 is negative, i.e., $L^{2'}(r^{L,2}) < 0$. Let $r^{L,2}(L^2)$ denote the corresponding inverse loan demand function of entrepreneurs Type 2.

If entrepreneur does not default, bank receives the agreed repayment, $r^{L^2}l^2$, otherwise, it receives $1 - \lambda$. We define a recovery rate x^2, that measures the fraction of the promised repayment that is effectively repaid to banks. The recovery ratio is one if entrepreneur do not default; otherwise, it yields $\frac{1-\lambda}{r^{L,2}l^2}$. In general,

$$x^2(\phi_f^2, r^{L,2}(L^2)) = min\left\{1, \frac{1 - \lambda}{r^{L,2}}\frac{\phi_f^2}{\phi_f^2 - 1}\right\}. \tag{14}$$

3.3 Banks

Banks supply loans to both types of entrepreneurs and capture deposits from households. Bank deposits are fully insured. This is, when banks default, government ensures depositors are fully repaid. Banks face a supply of deposits that is perfectly elastic at the risk-free interest rate, r. Bank lending activities are financed by fully insured deposits and by exogenous equity. In the two loan markets, banks compete for loans à la Cournot, so the strategy variable of a bank $i = 1, ..., n$ is its supply of loans l_i^1 and l_i^2 to entrepreneurs Type 1 and Type 2, respectively. The aggregate supply of loans in each market is given, respectively, by,

$$L^1 = \sum_{i=1}^{n} l_i^1, \qquad L^2 = \sum_{i=1}^{n} l_i^2.$$

These determine the lending rates $r^{L^1}(L^1)$ and $r^{L^2}(L^2)$, respectively. The balance sheet of bank i is,

$$l_i^1 + l_i^2 = d_i + n_{bi}, \tag{15}$$

where d_i are government-insured deposits from households, and n_{bi} is the exogenous initial equity. If entrepreneurs Type 1 default, the recovery rate x^1 is given by (7). If entrepreneurs Type 2 default, the recovery rate x^2 is given by (14). Hence, $x^1 r^{L,1}$ and $x^2 r^{L,2}$ are the effective gross returns received by loans to entrepreneurs Type 1 and Type 2, respectively. Then, bank i aims to maximize its discounted expected value of future profits,

$$\pi_i^b = \beta \mathbb{E}\{max(x^1 r^{L,1} l_i^1 + x^2 r^{L,2} l_i^2 - r d_i, 0)\}. \tag{16}$$

If entrepreneurs Type 1 do not default (i.e., $Z_i \geq Z_i^*$) and entrepreneurs Type 2 do not default neither, $x^1 r^{L,1} l_i^1 + x^2 r^{L,2} l_i^2 - r d_i$ yields,

$$r^{L,1} l_i^1 + r^{L,2} l_i^2 - r d_i > 0,$$

since $r^{L,1} > r$ and $r^{L,2} > r$ in equilibrium. If entrepreneurs Type 1 default (i.e., $Z_i < Z_i^*$) and entrepreneurs Type 2 do not default, $x^1 r^{L,1} l_i^1 + x^2 r^{L,2} l_i^2 - r d_i$ yields,

$$\frac{(1-\delta)K + Z K^\alpha}{L^1} l_i^1 + r^{L,2} l_i^2 - r d_i.$$

We define Z_i^{**},

$$Z_i^{**} = max\{0, Z_i^{***}\}. \tag{17}$$

where,

$$\frac{(1-\delta)K + Z_i^{***} K^\alpha}{L^1} l_i^1 + r^{L,2} l_i^2 - r d_i = 0. \tag{18}$$

Hence, $F(Z^{**})$ is the bank default probability given that entrepreneurs Type 2 do not default. If entrepreneurs Type 2 default, $x^1 r^{L,1} l_i^1 + x^2 r^{L,2} l_i^2 - r d_i$ yields,

$$x^1 r^{L,1} l_i^1 + (1-\lambda)\frac{\phi_f^2}{\phi_f^2 - 1} l_i^2 - r d_i < 0,$$

since we assume a high enough value for λ so that banks always default whether or not entrepreneurs Type 1 default. With this, bank i default probability is,

$$p_{bi} = p_f^2 + (1 - p_f^2) F(Z_i^{**}). \tag{19}$$

For convenience, discounted bank i profits (16) are rewritten as,

$$\pi_i^b = \beta(1 - p_{fi}^2)\left[\int_{Z_i^{**}}^{Z_i^*} (x^1 r^{L,1} l_i^1 + r^{L,2} l_i^2 - r d_i) dF(Z) + \int_{Z_i^*}^{+\infty} (r^{L,1} l_i^1 + r^{L,2} l_i^2 - r d_i) dF(Z)\right].$$

When bank chooses l_i^1 and l_i^2, it takes into account the effects on inverse demand curves of bank loans, $r^{L,1}$ and $r^{L,2}$, on the recovery ratios, x^1 and x^2, and on Z_i^* and Z_1^{**}. The first order condition with respect to l_i^1 yields,[11]

$$\int_{Z_i^{**}}^{Z^*} \left(x^1 r^{L,1} + \frac{\partial (x^1 r^{L,1})}{\partial l_i^1} l_i^1 - r \right) dF(Z) + \int_{Z^*}^{+\infty} \left(r^{L,1} + \frac{\partial r^{L,1}}{\partial l_i^1} l_i^1 - r \right) dF(Z) = 0, \tag{20}$$

Proceeding in a similar way, the first order condition with respect to l_i^2 yields,

$$r^{L,2} + \frac{\partial (r^{L,2})}{\partial l_i^2} l_i^2 - r = 0. \tag{21}$$

Equations (20) and (21) represents the loan supply curve of bank i to entrepreneurs Type 1 and Type 2, respectively.

3.4 General Equilibrium

Here, we characterize the aggregate behavior of entrepreneurial and banking sectors. Indeed, equation (6) represents the aggregate demand of bank loans of entrepreneurs Type 1. In the case of entrepreneurs Type 2, since they are identical (except for their reservation utility level), equation (11) represents the aggregate demand curve of loans of entrepreneurs Type 2. Since banks are identical, it is straightforward to aggregate equations (20) and (21) in order to derive the aggregate supply of loans to entrepreneurs

[11] Proof in Appendix A.

Type 1 and entrepreneurs Type 2, respectively,

$$\int_{Z^{**}}^{Z^*}\left(x^1 r^{L,1}+\frac{\partial(x^1 r^{L,1})}{\partial l^1}L^1\frac{1}{n}-r\right)dF(Z)+\int_{Z^*}^{+\infty}\left(r^{L,1}+\frac{\partial r^{L,1}}{\partial l^1}L^1\frac{1}{n}-r\right)dF(Z)=0, \tag{22}$$

$$r^{L,2}+\frac{\partial(r^{L,2})}{\partial l^2}L^2\frac{1}{n}-r=0, \tag{23}$$

where $1/n = l^1/L^1 = l^2/L^2$. Also, the default probability of the representative bank is given by equation (19),

$$p_b = p_f^2 + (1-p_f^2)F(Z^{**}), \tag{24}$$

and it shows bank's default probability is driven by its lending exposure to both entrepreneurs Type 1, via Z^{**}, and entrepreneurs Type 2, via p_f^2. Z^{**} is obtained from the aggregate version of equation (17) as,

$$Z^{**} = max\{0, Z^{***}\},$$

where

$$(1-\delta)K + Z^{***}K^\alpha + r^{L,2}L^2 - rD = 0.$$

If bank default probability given that entrepreneurs Type 2 do not default, is positive, i.e. $Z^{***} > 0$, we find,

$$Z^{**}=\frac{1}{K^{\alpha-1}}\left(\frac{\phi_f^1-1}{\phi_f^1}\frac{L}{L^1}\left[r\frac{\phi_b-1}{\phi_b}-r^{L,2}\frac{L^2}{L}\right]-(1-\delta)\right). \tag{25}$$

where $L = L^1 + L^2$, $\phi_b = (L^1 + L^2)/N_b$ and $\phi_f^1 = K/N_f^1$. In this case, according to equations (24) and (25), in the general equilibrium bank default probability is ceteris paribus affected by:

- Entrepreneurs Type 2's default probability p_f^2: the higher the default probability of entrepreneurs Type 2, the higher the bank default probability.
- Entrepreneurs Type 1's leverage ϕ_f^1: given that entrepreneurs Type 1 default, the higher ϕ_f^1, the smaller is what banks can recover from defaulting entrepreneurs Type 1 per unit of these loans, and hence the higher the bank default probability. This is called the *credit and excessive risk-taking effect* , as described in subsection 3.1.
- Bank leverage ϕ_b: The higher the bank leverage, the lower bank capacity to absorb losses and hence the higher its default probability. This also increases bank incentives to take more risk and issue more loans. This is called the *leverage effect*.
- Entrepreneurs Type 1' loan share L^1/L: given that entrepreneurs Type 1 default,

the higher share of these loans the higher the share of loans that are not fully repaid and hence the higher the bank default probability. This is called the *composition effect*.

- Entrepreneurs Type 2' loan share L^2/L: given that entrepreneurs Type 2 do not default, the higher the share of Type 2 loans, the higher fraction of bank loans that are fully repaid, and hence the lower the bank default probability. This is also called the *composition effect*.

- The lending rate $r^{L,2}$: the higher the lending rate of the non-defaulting loans (L^2), the lower the bank default probability. This is known in Martinez-Miera and Repullo (2010) as the *margin effect*.

Notice that when bank default probability, given that entrepreneurs Type 2 do not default, is zero, i.e., $Z^{***} < 0$, only entrepreneur Type 2's default probability p_f^2 drives bank default probability. This is, $p_b = p_f^2$.

3.5 Implications of Bank Competition

In this framework the number of banks n is our measure of bank competition. In what follows, we focus on the effects of a higher competition, which is driven by a higher number of banks.

As it is suggested in the Cournot equilibrium, ceteris paribus the higher the number of banks, the smaller the market share (i.e., the smaller $1/n$) and hence the smaller the markups (i.e., the smaller $\frac{\partial(x^1 r^{L,1})}{\partial l_i^1} L^1 \frac{1}{n}$, $\frac{\partial r^{L,1}}{\partial l_i^1} L^1 \frac{1}{n}$ and $\frac{\partial(r^{L,2})}{\partial l_i^2} L^2 \frac{1}{n}$) on both loan markets. Consequently, the smaller the markups, the smaller the required returns (lending rates) of both loans, or equivalently the higher the supply of loans to both entrepreneurs Type 1 and Type 2.

After a lower lending rate $r^{L,2}$, entrepreneurs Type 2, as suggested by equation (12), are going to choose projects with a lower probability of failure and hence the default probability of entrepreneurs Type 2, p_f^2, decreases. According to (24) this pushes down bank default probability (*risk-shifting effect*). In addition, according to (13), the lower lending rate increases aggregate level of loans to entrepreneurs Type 2. Also, according to (25) this lower lending rate increases bank default probability (*margin effect*). As commented before, whether the borrowing constraint binds or not, it does not affect the impact of bank competition (n) on entrepreneurs Type 2' risk-taking decisions.

In the case of entrepreneurs Type 1, a lower lending rate $r^{L,1}$ is expected to increases bank loans, and with this there is going to be a higher default probability of Type 1. This turns to a higher leverage of entrepreneurs Type 1. And this higher leverage, as explained before, pushes up bank default probability (*credit and excessive risk-taking* *effects*). Importantly, the lower the lending rate, the higher the likelihood that the borrowing constraint binds. If the constraint binds, the demand of loans becomes perfectly inelastic at $\bar{L}^1$ and hence entrepreneurs Type 1 default probability does not change.[12] In this case the *credit and risk-taking channel effects* are turned off.

Also, the lower lending rates that increase lending activities, increase bank leverage as well. And as explained before, this pushes up bank default probability (*leverage effect*). Notice that this effect is on even if borrowing constraint on entrepreneurs Type 1 bind, since loans to entrepreneurs Type 2 are going to still increase. Also, if this borrowing constraint binds, the loan share of entrepreneurs Type 1 (Type 2) decreases (increases). According to (25), this pushes down bank default probability (*composition effect*).

To sum up, if the borrowing constraint for entrepreneurs Type 1 does not bind, a higher number of banks might produce two opposites effects on bank default probability. According to the *risk-shifting effect*, a high number of banks pushes down bank default probability; while according to the *margin effect*, the *credit and excessive risk-taking effects*, and the *leverage effect*, a higher number of banks pushes up bank default probability. And it is not clear the effect of the implications of the *composition effect*. While if the constraint binds, only the *risk-shifting effect*, the *margin effect*, the *leverage effect*, and the *composition effect* are on, and the other effects are turned off.

As we show in the quantitative analysis, this framework is able to capture three phases of the relationship between bank competition and bank risk-taking:

- In the first phase, the number of banks (i.e., low competition) is low enough so the leverage of entrepreneurs Type 1 is low enough so bank default probability is zero given the entrepreneurs Type 2 do not default. In this case, only the *risk-shifting effect* is on. Hence, we should observe a lower bank default probability after a higher number of banks.

- In the second phase, when the number of banks is mid (i.e., mid competition), the leverage of entrepreneurs Type 1 is high enough so bank default probability, given that entrepreneurs Type 2 do not default, is positive. This means that all studied *effects* are on. In this case, we claim that we should observe a higher bank default probability after a higher number of banks, and hence the higher bank default probability $F(Z^{**})$, given that entrepreneurs Type 2 do not default, dominates the *risk-shifting effect*. And the main force driving this higher $F(Z^{**})$ is the *credit and risk-taking effect*.

- In the third phase, when the number of banks is high enough, the borrowing con-

[12]Notice that the lending rate $r^{L,1}$ does not change, this is because banks do not have incentives to reduce $r^{L,1}$. And if banks try to increase $r^{L,1}$ it reduces entrepreneurs Type 1' to demand loans, moving away from a binding borrowing constraint, and hence $r^{L,1}$ need to go down again since it is not consistent to observe a higher lending rate with more competition when the borrowing constraint does not binds.

straint of entrepreneurs Type 1 binds. In this case, only the *risk-shifting effect*, the *margin effect*, the *leverage effect* and the *composition effect* are on. Here, we claim that we should observe a lower bank default probability after a higher number of banks, although the speed is not the same as in the first phase. This is because the *margin effect* and the *leverage effect* diminish the negative impact on bank default probability of the *risk-shifting effect*.

As a result, if we consider the first and second phase only, the model characterizes the U-shaped relationship between bank competition (i.e., number of banks) and bank default probability, also found in Martinez-Miera and Repullo (2010). More importantly, if we consider the second and third phases only, the model features the inverted U-shaped relationship between bank competition and bank default probability.

Next, we present the quantitative analysis and verifies the presence of these three phases. Instead of solving for the partial derivatives of bank default probability and other endogenous variables with respect to the number of banks (i.e., an analytical solution), which is not straightforward and we might not be able to easily conclude about the signs, we resort to numerical solutions for a simple parametrization of the model.

3.5.1 Quantitative Analysis

Here, we compute the equilibrium of the model and we quantify the impact of the number of banks of bank default probability. To do so we propose a simple parametrization and linear functions for the inverse demand of loans of entrepreneurs Type 2, $r^{L,2}(L^2)$, and non-failure payoff of the projects of entrepreneurs Type 2, $\theta(p_f^2)$. The parameters are chosen for illustrative purposes and not necessarily to reproduce realistic variable values.

We parametrize the model with standard values, whenever is feasible. In our baseline calibration, we set $\delta = 0.20$, $\alpha = 0.33$ and $\beta = 0.9606$.[13] We set $\mu_= - (\sigma^2)/2$, so the unconditional mean of aggregate shock Z equals one and hence is independent of σ.[14] To calibrate the loan composition, we add an exogenous multiplicative factor ω to the payoffs of the projects of entrepreneurs Type 1, i.e., $\omega Z K^{\alpha}$, and we set $\omega = 4.5$ so it yields $L^1/L^2 = 7.58$. We set the number of banks to 10.

The inverse aggregate demand of loans of entrepreneurs Type 2 is set to $r^{L,2} = 2 - 0.40L^2$ and the payoff of succeeding is set to $\theta(p_f^2) = 0.72 + 0.80p_f^2$. The loss given default parameter λ is set at 1. Finally, we set N_f^1, $\bar{l}^2$, N_b and σ, so that leverage of entrepreneurs Type 1, ϕ_f^1, leverage of entrepreneurs Type 2, ϕ_f^2, leverage of banks, ϕ_b,

[13]Notice that one period in the model represents one year. The household subjective discount factor is then the annualized of the one used in the literature $\beta = 0.99$. In addition, in order to have a positive default probability of banks, we need to set δ with a parameter value higher than what is used in the literature $\delta = 0.10$.

[14]This helps us to measure the impact of a higher volatility of the aggregate shock through a higher σ without affecting the mean.

and entrepreneurs Type 1' default probability, p_b, equate 7, 7, 15 and 25%.[15] This results in $N_f^1 = 2.75$, $\bar{l}^2 = 0.86$, $N_b = 1.25$ and $\sigma = 1.27$. In equilibrium, the privately optimal failure probability of projects of entrepreneurs Type 2 yields 0.03, bank default probability yields 0.04. In addition, the gross lending interest rates of loans to entrepreneurs Type 1 and Type 2, $r^{L,1}$ and $r^{L,2}$, are 1.07 and 1.13, respectively. And we set $\bar{L}^1$ enough high so it does not bind in our baseline calibration.[16]

Figure 3 shows the effects of the number of banks n on bank default probability. In particular, we reduce and increase n keeping the other parameters unchanged. The figure shows us that under this baseline calibration we are able to characterize the three phases previously described.

For a low number of banks, bank default probability decreases (solid red line), where this latter is driven only by the lower default probability of entrepreneurs Type 2 (solid blue line), since bank default probability given that entrepreneurs Type 2 do not default (dashed black line) is zero. This is the first phase. Then, when the number of banks becomes higher, bank default probability increases, which is driven by the increase of the bank default probability given that entrepreneurs Type 2 do not default (dashed black line). This is the second phase. Up to this point, we observe a U-shaped relationship between bank competition and bank risk-taking as in Martinez-Miera and Repullo (2010). If the number of banks continues increasing, the borrowing constraint (for entrepreneurs Type 1) turns out to be binding, we start observing a reduction of bank default probability although with a low speed. This lower bank default probability is driven by the lower default probability of entrepreneurs Type 2. Finally, this is the third case. As stated before, considering the second and third phases we might be able to observe an inverted U-shaped relationship between bank competition and bank risk-taking.

Figure 3 also reports that the lending rates decrease after a larger number of banks. However, the lending rate of loans to entrepreneurs Type 1 keeps unchanged after the borrowing limit for entrepreneurs Type 1 starts binding. Also, the loan portfolio composition suggests a higher participation of loans to entrepreneurs Type 1, which pushes up bank default probability (in the second phase); however, it starts decreasing after the borrowing constraint of entrepreneurs Type 1 starts binding. As suggested before, this pushes down bank default probability.

[15] We set relatively high leverage targets, especially for banks, in order to have a positive bank default probability.

[16] In particular, $\bar{L}^1$ is 11% higher than the baseline value of bank loans to entrepreneurs Type 1.

Figure 3: *Competition and risk of bank failure*

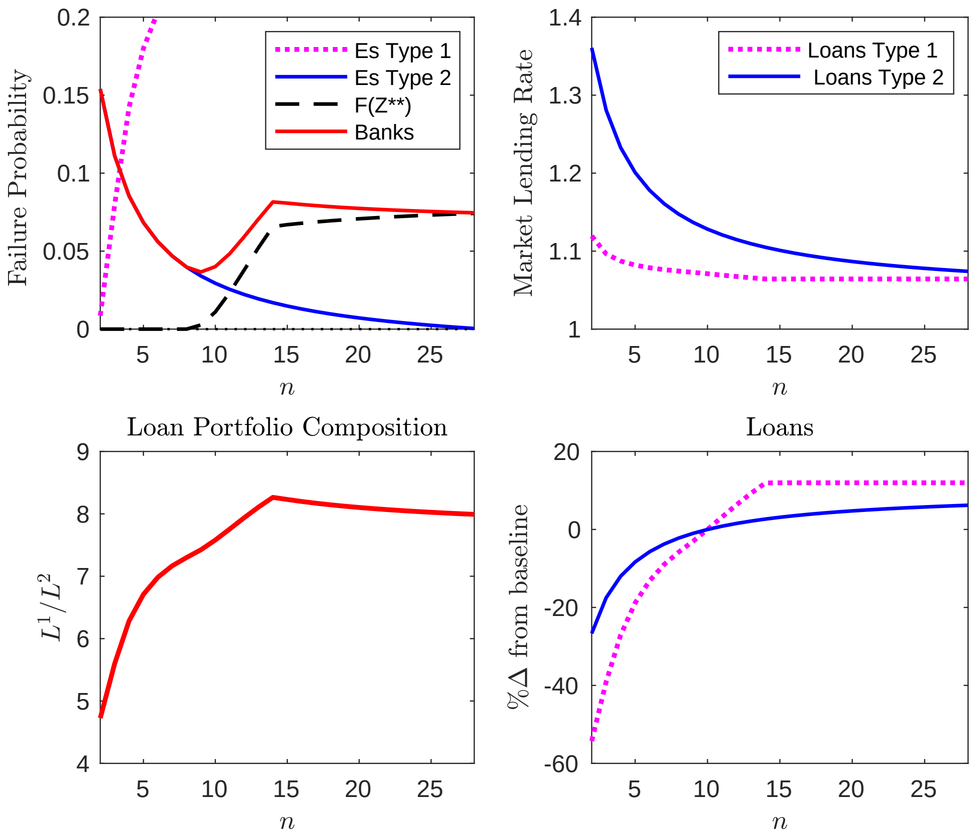


Figure shows the responses of some variables to changes of the number of banks while keeping the other parameters unchanged. $F(Z^{**})$ is the bank default probability given that entrepreneurs Type 2 do not default. Es Type 1(2): Entrepreneurs Type 1(2). Loans Type 1(2): Bank loans to entrepreneurs Type 1(2).

3.5.2 Robustness

If we assume other functional forms for the payoff in case of success for entrepreneurs Type 2, such as $\theta(p_f^2) = a + bq^\eta$ with $\eta \leq 1$ to ensure concavity, and for the inverse demand for loans $r^{L,2}$, such as $r^{L,2}(L^2) = c - dL^\gamma$ with $\gamma > 1$ to ensure concavity, the results might not be robust. In particular, for a low enough η, we do not find that a binding borrowing constraint for entrepreneurs Type 1 leads to a reduction of bank default probability after a higher number of banks (see figure 6 in Appendix B). This is because, according to (12), with $\eta < 1$ the response of default probability of entrepreneurs Type 2 after the reduction of the lending rate becomes smaller. Therefore, the *risk-shifting effect* is dominated by the *margin effect* and the *leverage effect*.

In figure 7 in Appendix B we find that the results are robust for small changes of ω, which defines the L^1/L^2 ratio. In particular the higher the ω (red dotted line). the higher the L^1/L^2 ratio while the borrowing constraint of entrepreneurs Type 1 does not bind. This also implies the borrowing constraint becomes binding for a lower number of banks or for a lower bank competition.

Next, we move to the empirical section of the paper. According to the model in order to find evidence of an inverted U-shaped relationship of bank competition and bank risk-taking, we might need to control by bank characteristics and by investment opportunities. Next, we turn to the empirical strategy and provide the empirical evidence of the relationship of bank competition and bank risk-taking in an emerging economy as Peru.

4 Data and Model Description

We use two levels of information. First, following the related literature, we use a bank-time (or financial institution-time) level data. Second, 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 (*Superintendencia de Banca, Seguros y Administradoras de Fondos de Pensiones*, SBS, website)[17]. The main limitation of this dataset is that the measure of risk taking 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). This latter is important since our model is based on the behavior of banks that issue loans to firms.

Firs, in this section we describe the bank-time level dataset, the empirical model, and our first regression results. And in section 5 we focus on the more granular data.

4.1 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 number of banks operating in each region, the share of loans of the four-largest financial institutions operating in each region (C4)[18], and the Herfindahl-Hirschman index (HHI), which is the sum of banks' squared market shares in loans in each region. The latter two measures are also typically used in the literature as concentration measures.

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[19].

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

[18] 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.

[19] This is to account for loans issued from branches abroad. We include these loans since from the

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 financial regulator from Peru, *Superintendencia de Banca, Seguros y Administradoras de Fondos de Pensiones* (SBS). 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-Hirschmann 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.[20] FD controls the bank's capacity to borrow from foreign markets, which in turn might affect banks' incentives to take risk.

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 as they encounter as competitors larger international financial institutions.

[20] See Agur et al. (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.

Our dependent variable is bank risk-taking. In this document, it is measured as the ratio of nonperforming loans to total loans under the same criterion defined by the Peruvian financial regulator, SBS,[21]

$$\frac{\text{loan arrears (Big firms(15d), small firms(30d) mortgage(30d), personal(90d))}}{\text{Total credits}}100. \tag{26}$$

The information about nonperforming loan (NPL) ratio at an institution-time level is also provided by the SBS.[22]

We not only assess the relationship between competition and risk-taking in the sample of banks in Peru, but also consider the five main financial groups, as a whole, that exist in the Peruvian financial system: banks, *empresas financieras*, CAMCs, CRACs and EDPYMEs. In this case, 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.

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 including the Covid-19 shock period. We begin our investigation by focusing solely on banks, then we expand it to include all financial institutions.

To ensure data consistency, we correct and merge the time series information for the same institution that for some reason changed its name or changed both its name and its financial group to which it belongs, causing to appear as two or more different time series.[23] In addition, we include dummies to control for these events and other financial events.[24]

Tabla 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.[25] 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.

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

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

[23]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.

[24]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.

[25]This is the number of bank-year observations before allowing for lags.

Tabla 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.[26] 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.[27] 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.

Table 2: *Descriptive statistics for bank-year observations*

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

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%).

[26]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*.

[27]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.

Table 3: *Descriptive statistics for financial institution-year observations*

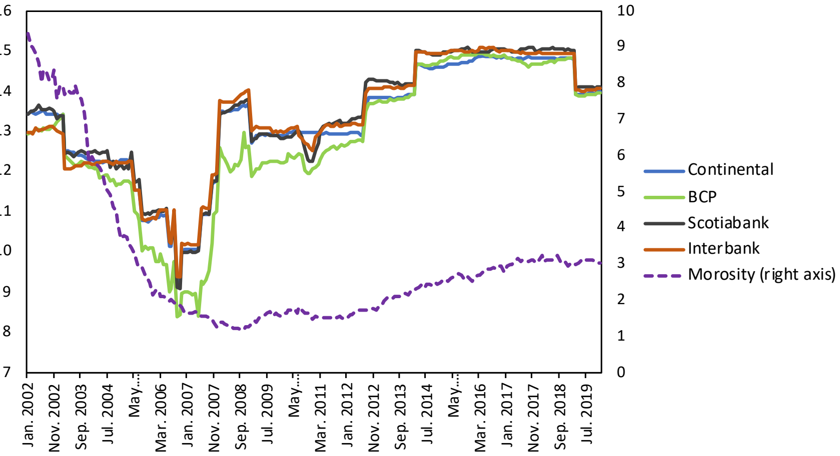
Variables	Obs	Mean	S.D.	Minimum	Maximum
NPL_{it} (%)	881	5.64	5.17	0.02	73.34
Non-competition across groups					
Number of $institutions_{it}$	881	7.25	4.39	1.00	16.00
$C4_{it}$	881	0.91	0.09	0.70	2.00
HHI_{it}	881	0.41	0.22	0.17	1.00
$SIZE^{*}_{it}$	881	0.09	0.10	0.00	0.88
Competition across groups					
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

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%).

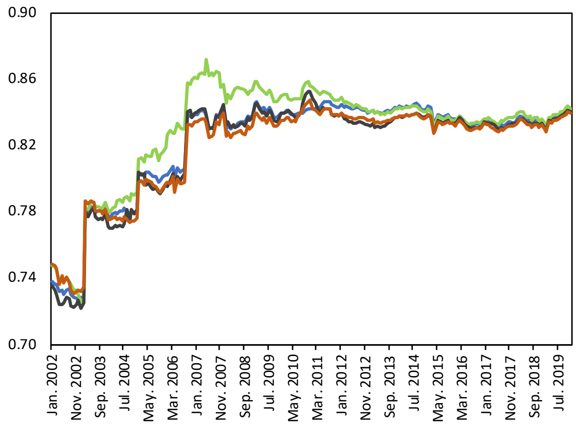
Figure 4.a reports the behavior of "number of banks" for the four largest banks from 2001 to 2018. In general, there is a common trend that governs the long-term dynamics in this competitive measure. Around the 2008 global financial crisis, these banks exhibit relatively more dispersion, compared to 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 this measure, for example, from 2004 to 2012 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, figure 4.b and figure 4.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.

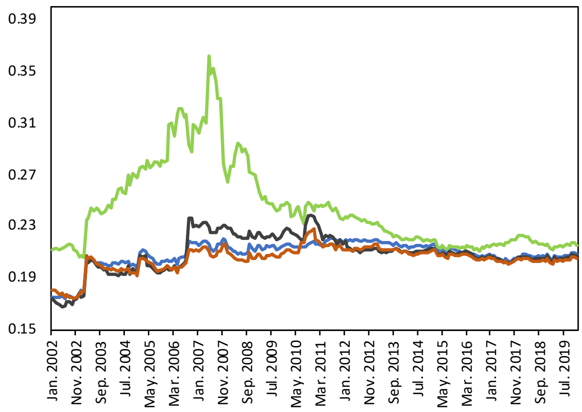
Figure 4: *Bank competition and concentration measures*



(a) Competition measure: "Number of banks" and morosity rate



(b) Concentration measure: C4

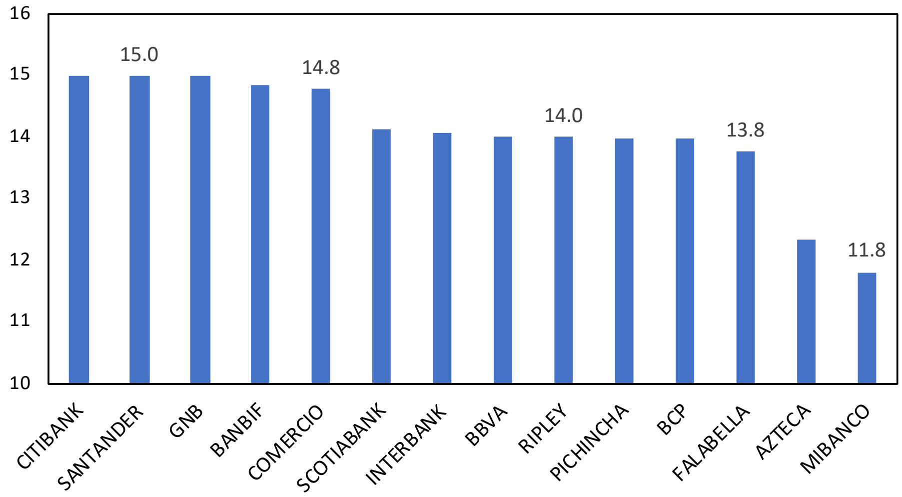


(c) Concentration measure: HHI

Source: SBS. Own calculations. Morosity rate = Banking sector non-performing loans share (SBS criterion) or NPL ratio. Monthly frequency. Period: January 2002 - December 2019.

In order to have an idea of the heterogeneity of bank competition, figure 5 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.

Figure 5: *Bank competition measure: "number of banks" - December 2019*



Source: SBS. Own calculations.

4.2 Model Description

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

$$endo\_var_{it} = \alpha + \beta_0 * endo\_var_{it-1} + \beta_1 * exo\_var_{it-1} + \beta_2 * exo\_var^2_{it-1} + \beta_3 * ctrl_{it} + \epsilon_{it}, \quad (27)$$

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 risk-taking 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.[28]

The dependent variable ($endo\_var_{it}$) is the log-odds transformation of the bank NPL ratio, which changes the variables's support from the unit interval to the real number line. In other words, $endo\_var_{it} = ln(NPL_{it}/(100 - NPL_{it})$, NPL_{it} is the non-performing loans ratio, defined in equation (26).[29] 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

[28]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.

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

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.

4.3 Regression Results

In this subsection we present the results when considering only banks and all financial institutions assuming non-competition between 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 risk-taking. 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 risk-taking only when considering banks. Also, bonds issued by non-financial institutions to credit ratio (BOND) is negatively and statistically significant associated with risk-taking. Finally, the risk weighted assets to capital ratio (RWA) is positively associated with risk-taking. 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 risk-taking when considering only banks, while positively associated but less statistically significant with risk-taking when considering all financial institutions.

Table 4 reports the estimation results for the model when considering only banks. It shows the results for nine different regressions. For each measure 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 risk-taking and loan market bank competition. The results are statistically significant when we do not include any fixed effects and even when time fixed effects are included. When considering bank fixed effects, the signs are the same but the relationship is not significant.[30]

When using the concentration measures (C4 or HHI), in general, results suggest a U-shaped relationship between bank risk-taking and bank competition as suggested by Martinez-Miera and Repullo (2010) and Jiménez et al. (2013). In the case of C4 and HHI are only significant with time fixed effects,

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

*** Statistically significant at 1%, ** statistically significant at 5%, * statistically significant at 10%. 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).

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 regional level two institutions from different groups compete in the same intensity as two financial 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 to 0.78, confirming the persistence in the NPL ratio.

When using the number of financial institutions, the estimation results show an inverted U-shaped relationship between bank risk-taking and loan market bank competition. Results are significant when omitting fixed effects and when considering time fixed

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

effects, while when considering financial institution fixed effects results are not significant, but keep the same direction.[31]

When using the concentration measures (C4 and HHI), in general, results suggest a U-shaped relationship between bank risk-taking and bank competition as suggested by the literature, but these results are not statistically significant. Table 10 in Appendix C reports the results when assuming competition across group. In that case, we do not find statistically significant estimates.[32]

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

*** Statistically significant at 1%, ** statistically significant at 5%, * statistically significant at 10%. 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.

5 Granular Assessment

Our previous analysis has some shortcomings, as it can not control for trends on demand and supply sides other than through the aggregate trends in the economy. In addition, our theoretical framework, as is also in Martinez-Miera and Repullo (2010), is based on the borrowing behavior of commercial firms; however, in our previous assessment we are not able to consider loans to firms only, as it is done in Jiménez et al. (2013).

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

[32] 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.

In order to overcome these problems and avoid a biased estimator we make use of more granular data. In particular, we add another dimension to the institution-time data: "region".[33] 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 microenterprises[34]. As a result, our baseline specification (27) becomes,

$$endo\_var_{icrt} = \alpha + \gamma_{rt} + \lambda_{it} + \mu_c + \beta_0 * endo\_var_{icrt-1} + \beta_1 * exo\_var_{icrt-1} + \beta_2 * exo\_var^2_{icrt-1} + \epsilon_{icrt}, \tag{28}$$

where the r subscript refers to a region, the c subscript refers to type of credit, 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.

5.1 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.[35] 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).[36] It contains data on firm Tax ID

[33]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).

[34]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.

[35]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.

[36]SUNAT information source http://www.sunat.gob.pe/descargaPRR/mrc137_padron_reducido.

(RUC) and Location codes (UBIGEO). (2) Datos Perú.[37] It contains information on businesses identified by RUC and their geographic location. And (3) Credit Report of Debtors (RCD by its Spanish acronym)[38]. It contains information on borrowers identified by RUC and/or DNI and their geographic location, UBIGEO. [39]

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.[40] As a result, we are able to build up a panel-data at loans type-financial institution-region-time level. Since our model is based on the behavior of firms, we focus only on loans to firms, i.e., we focus only on two types of credit: commercial loans and loans to microenterprises.[41]

We assume that loans go to the province-region registered as the location of the borrower.[42] 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. This is a potential borrower located in region A cannot move to region B and ask for loans. Similarly, a financial institution's branch located at region A cannot offer loans to clients located in region B.

As before the risk-taking measure is given by the non-performing loans ratio, which is built using the SBS criterion (see, equation 26) 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

`html` accessed on 20/06/2018.

[37]https://www.datosperu.org/. We resort to web scraping techniques to extract the required information from the website in June-July 2021.

[38]RCD is not publicly available. It is restricted information provided by the SBS.

[39]Table 11 in Appendix D shows the pairs of RUC-UBIGEO and DNI-UBIGEO from using these three datasets and the strategy followed in case of conflicts.

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

[41]According to figure 8 in Appendix D, 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.

[42]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.

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-Hirschmann 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 comes at no surprise, since we are measures 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.

Table 6: *Descriptive statistics for bank-type-region-year observations*

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

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%).

Table 7: *Descriptive statistics for financial institution-type-region-year observations*

Variables	Obs	Mean	S.D.	Minimum	Maximum
NPL_{icrt} (%)	17370	9.74	14.50	0.00	100.00
Non-competition across groups					
Number of $\text{institutions}_{icrt}$	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 $\text{institutions}_{icrt}$	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

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%).

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.

5.2 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 9 in Appendix E illustrate the representativity of our sample at the aggregate level. Figure 9.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 9.b suggests that our sample of loans to firms mimics fairly well the dynamics of the official data since 2005.[43] In addition, according to figure 10 in Appendix E 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 11 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 12 our sample mimics very well the credit growth at bank-time level. Similarly, figure 13 our sample mimics

[43] This is because the 2004 growth rate contains 2004 information which according to Figure 9.b does not represent well the official data.

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 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.[44] 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 14 in Appendix E 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 15 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 equation (28) considering banks and only the 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 always validate the inverted U-shaped relationship between bank risk-taking and bank competition, independently if we control by demand (column 2) or supply (column 1) or even if we do not (column 1).[45]. This is the opposite of Martinez-Miera and Repullo (2010) findings. 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 when using HHI as the competition measure.

Robustness provided in Table 12 Appendix F shows that our results are consistent across different specifications. Statistical significance does not change when we omit extreme value observations of the NPL ratio.[46] 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. And when considering only loans to microenterprises, none of the estimates are significant. These two latter robustness exercises come as no surprise, since the inverted U-shaped is only consistent when two types of firms in the credit market are considered:

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

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

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

constrained and unconstrained firms, as shown in the analysis with the theoretical model in Section 3.5.1.

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

*** statistically significant at 1%, ** statistically significant at 5%, * statistically significant at 10%. Robust standard errors in columns (1), (4) and (7). Clustered (at region-time level) standard errors in columns (2), (5) and (8). Clustered (at financial institution-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.

5.3 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. This is, 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 16 in Appendix E, 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 oficial data.

Table 9 reports the regression results of the empirical model, equation (28), when considering all financial institutions, two types of loans to firms (commercial credit and loans to microenterprises), (i) non-competition and (ii) competition across groups within a province. In the case of non-competition, when considering the C4 and HHI as our competition measures, results validate an inverted U-shaped relationship between bank competition and bank risk-taking. For the case of HHI coefficient estimates are significant either we control by demand or supply shocks, columns 8 and 9 respectively. 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.

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 risk-taking but the 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 inconclusive.

Notice that in contrast to our regressions at the financial institution-time level in subsection 4.3, if we assume competition across groups, this time results can validate (but without statistical significance) the inverted U-shaped relationship when considering the number of financial institutions as our competition measure. These findings could be a result of data construction. The likelihood that two financial institutions compete if they are located in the same province is higher than if they are located in the same region. This might suggest that assuming non-competition (competition) across groups when estimating the model at region-financial firm-time level makes more sense than assuming competition (non-competition) across groups. We believe also, as demonstrated in the theoretical model, 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. Robustness shown in Appendix G confirms these results.

In short, we have found evidence of an inverted U-shaped relationship between competition, measured as the number of financial institutions, and the bank risk-taking, measured as the NPL ratio, in an emerging economy as Peru, in contrast to what is found in Martinez-Miera and Repullo (2010) for an advanced economy as Spain. And our theoretical provides a possible explanation based on the stronger financial constraints faced by entrepreneurs in emerging economies.

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**
$\text{exo\_var}^2_{icrt-1}$	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
$\text{exo\_var}^2_{icrt-1}$	-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

*** statistically significant at 1%, ** statistically significant at 5%, * statistically significant at 10%. Robust standard errors in columns (1), (4) and (7). Clustered (at region-time level) standard errors in columns (2), (5) and (8). Clustered (at financial institution-time level) standard errors in columns (3), (6) and (9). In all regressions we control by $\text{SIZE}_{ict,ct}$, $\text{SIZE}_{icrt,ict}$ and $\text{SIZE}_{icrt,ct}$. $\text{SIZE}_{x,y}=\text{credit}_x/\text{credit}_y$. 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.

6 Conclusions

In this paper we can conclude that in the Peruvian financial system there is evidence of a nonlinear relationship between competition and risk-taking. In contrast to Martinez-Miera and Repullo (2010), when considering the number of banks as our competition measure, there is an inverted U-shaped relationship, which according to our model this can be explained by the stronger financial constraints faced by entrepreneurs in emerging markets. This result holds when studying only banks or all financial institutions (assuming non-competition across groups) at the bank-time level. In addition, this result is robust when considering granular data for banks and hence being able to work at the region-bank-time level, and even when controlling for supply and demands shocks.

As far as policy implications are concerned, our results point that, increasing the number of financial institutions does not necessarily increase the risk-taking behavior. However, our estimates that consider concentration measures show the opposite picture. Finding a common basis to explain this otherwise perplexing behavior of risk-taking, competitiveness, and market concentration is something that we leave as part of our research agenda.

References

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

Agur, Itai and Maria Demertzis. 2019. "Will Macroprudential Policy, Counteract Monetary Policy's Effects on Financial Stability?" *The North American Journal of Economics*, 48, 66-75.

Beck, T.H.L.. 2007. "Financing constraints of SMEs in developing countries : Evidence, determinants and solutions." *Other publications TiSEM*, Tilburg University, School of Economics and Management.

Blalock, Garrick, Paul J. Gertler and David l. Levine. 2008. "Financial constraints on investment in an emerging market crisis." *Journal of Monetary Economics*, 55(3), 568-591.

Bolt, Wilko and Alexander F. Tieman. 2004. "Bank Competition, Risk and Regulation." *The Scandinavian Journal of Economics*, 104(4), 783-804.

Boyd, John H. and Gianni De Nicoló. 2005. "The Theory of Bank Risk Taking and Competition Revisited." *The Journal of Finance*, 60(3), 1329-1343.

Collard, F., H. Dellas, B. Diba and O. Loisel, 2017, "Optimal Monetary and Prudential Policy," *American Economic Journal: Macroeconomics*, 9(1) 40-87.

Dell'Ariccia, G., L. Laeven and R. Marquez. 2014. "Real interest rates, leverage, and bank risk-taking." *Journal of Economic Theory*, 149(1), 65-99.

Jiménez, Gabriel, Jose A. Lopez and Jesús Saurina, 2013. "How does Competition affect bank risk-taking?" *Journal of Financial Stability*, 9, 185-195.

Martinez-Miera, David and Rafael Repullo, 2010. "Does Competition Reduce the Risk of Bank Failure?" *The Review of Financial Studies*, 23(10), 3638-3664.

Poschke, Markus. 2018. "The Firm Size Distribution across Countries and Skill-Biased Change in Entrepreneurial Technology." *American Economic Journal: Macroeconomics*, 10 (3): 1-41.

Pozo, Jorge. 2019. “Bank Risk-Taking in a Small Open Economy.” *BCRP Working Paper*, 16.

Repullo, Rafael, 2004. “Capital requirements, market power, and risk-taking in banking.” *Journal of Financial Intermediation*, 13, 156-1.

Appendices

A Equilibrium condition of banks

Recall the discounted bank i profits (16) are,

$$\pi_i^b = \beta(1-p_{fi}^2)\left[\int_{Z_i^{**}}^{Z^*}(x^1 r^{L,1} l_i^1 + r^{L,2} l_i^2 - r d_i)dF(Z) + \int_{Z^*}^{+\infty}(r^{L,1} l_i^1 + r^{L,2} l_i^2 - r d_i)dF(Z)\right].$$

The first order condition with respect to l_i^1 yields,[47]

$$\begin{aligned}
& (x^1(Z^*) r^{L,1} l_i^1 + r^{L,2} l_i^2 - r d_i) f(Z^*)\frac{\partial Z^*}{\partial l_i^1} - (x^1(Z_i^{**}) r^{L,1} l_i^1 + r^{L,2} l_i^2 - r d_i) f(Z_i^{**})\frac{\partial Z_i^{**}}{\partial l_i^1} \\
+ & \int_{Z_i^{**}}^{Z^*}(x^1 r^{L,1} + \frac{\partial(x^1 r^{L,1})}{\partial l_i^1} l_i^1 - r)dF(Z) + \int_{Z^*}^{+\infty}(r^{L,1} + \frac{\partial r^{L,1}}{\partial l_i^1} l_i^1 - r)dF(Z) \\
- & (r^{L,1} l_i^1 + r^{L,2} l_i^2 - r d_i) f(Z^*)\frac{\partial Z^*}{\partial l_i^1} = 0.
\end{aligned}$$

Since in equilibrium $x^1(Z_i^{**}) r^{L,1} l_i^1 + r^{L,2} l_i^2 - r d_i = 0$ and $x^1(Z^*) = 1$. The first order condition with respect to l_i^1 becomes,

$$\int_{Z_i^{**}}^{Z^*}\left(x^1 r^{L,1} + \frac{\partial(x^1 r^{L,1})}{\partial l_i^1} l_i^1 - r\right)dF(Z) + \int_{Z^*}^{+\infty}\left(r^{L,1} + \frac{\partial r^{L,1}}{\partial l_i^1} l_i^1 - r\right)dF(Z) = 0,$$

where,

$$\frac{\partial(x^1 r^{L,1})}{\partial l_i^1} = \frac{\partial\left[\frac{(1-\delta)K + ZK^\alpha}{L^1}\right]}{\partial l_i} = \frac{-(1-\delta)N_f + ZK^\alpha(\alpha L^1/K - 1)}{(L)^2} < 0,$$

and,

$$\frac{\partial r^{L,1}}{\partial l_i^1} = -\frac{\int_{Z^*}^{+\infty} Z\alpha(1-\alpha)K^{\alpha-2}dF(Z) - (1-\delta+Z^*\alpha K^{\alpha-1} - r^L)^2 f(Z^*)\frac{1}{K^\alpha}}{(1-F(Z^*)) + (1-\delta+Z^*\alpha K^{\alpha-1} - r^L) f(Z^*)\frac{L^1}{K^\alpha}}.$$

[47] Using Leibniz's rule for differentiation.

B Theoretical Model: Robustness

Figure 6: *Competition and risk of bank failure: Non-linear functional forms*

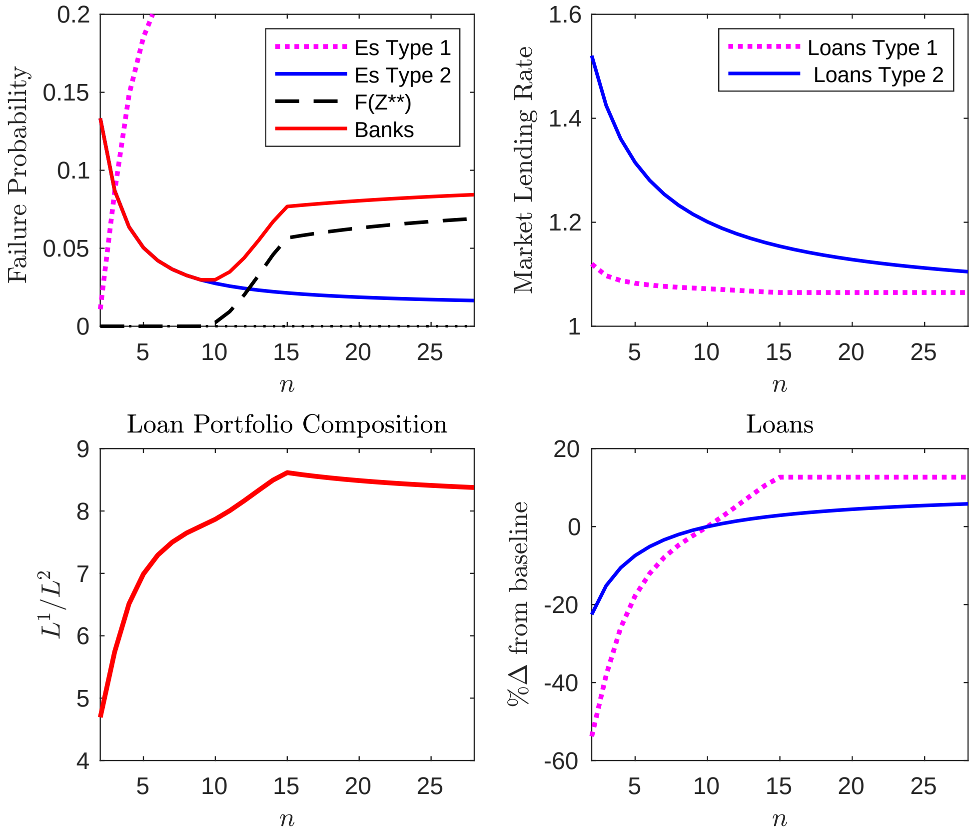


Figure shows the responses of some variables to changes of the number of banks while keeping the other parameters unchanged. $F(Z^{**})$ is the bank default probability given that entrepreneurs Type 2 do not default. We set $\theta(p_f^2) = 0.72 + 0.39q^{0.80}$ and $r^{L,2}(L^2) = 2 - 0.18(L^2)^2$, and we set $N_f^1 = 2.73$, $N_b = 1.23$ and $\sigma = 1.27$ in order to match the same targets. The rest of the parameters are keep unchanged.

Figure 7: *Competition and risk of bank failure: Loan composition*

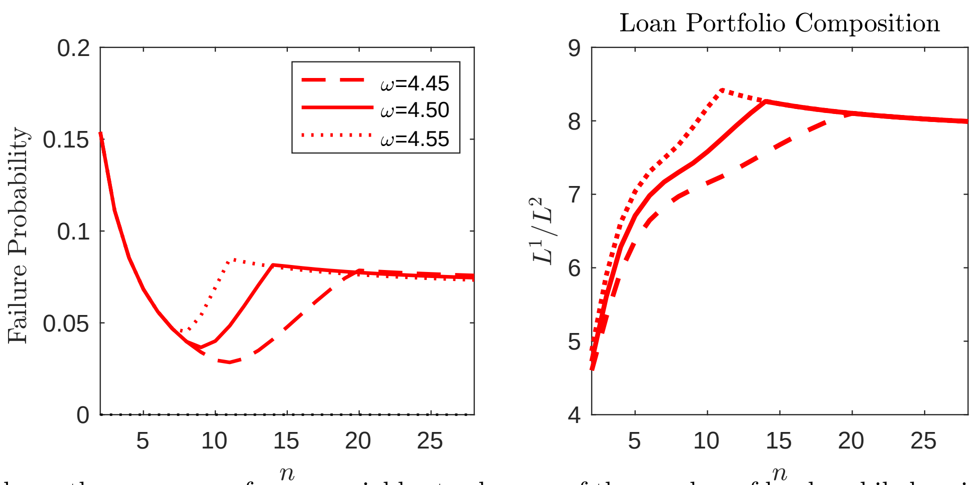


Figure shows the responses of some variables to changes of the number of banks while keeping the other parameters unchanged. $F(Z^{**})$ is the bank default probability given that entrepreneurs Type 2 do not default. There are three simulations that correspond to different values of ω, included the baseline value

C All Financial Institutions: Competition Across Groups

Table 10: *All financial institutions: Competition across groups*

exo_var	ln (# institutions)			C4			HHI		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
endo_var$_{it-1}$	0.769***	0.556***	0.777***	0.767***	0.560***	0.775***	0.770***	0.560***	0.777***
exo_var$_{it-1}$	0.00313	-1.010	-0.141	3.838	0.354	3.195	4.381	6.966	4.139
exo_var$^2_{it-1}$	-0.00305	0.188	0.0214	-2.732	0.214	-2.221	-12.73	-16.70	-11.70
ROA$_{it-1}$	-0.0117**	-0.00579	-0.0123	-0.0115**	-0.00670	-0.0122	-0.0115**	-0.00674	-0.0123
SIZE$_{it-1}$	-0.522*	4.107	-0.491**	-0.516*	4.027	-0.498**	-0.436	4.133	-0.434**
FD$_{it-1}$	0.104	0.122	0.0880	0.108	0.145	0.0922	0.110	0.146	0.0937
BOND$_{it-1}$	-1.796**	-1.320		-1.776**	-1.662**		-1.733**	-1.686**	
RWA$_{it-1}$	0.0294***	0.0434***	0.0316**	0.0295***	0.0440***	0.0314**	0.0301***	0.0447***	0.0318**
RGDPGR$_{it}$	-0.128	-1.262		-0.129	-1.427*		-0.139	-1.458*	
RGDPGR$_{it-1}$	-0.563	-1.993**		-0.580	-1.944***		-0.578	-1.969***	
Observations	802	802	802	802	802	802	802	802	802
R-squared	0.794	0.855	0.798	0.794	0.854	0.798	0.794	0.854	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

*** Statistically significant at 1%, ** statistically significant at 5%, * statistically significant at 10%. 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.

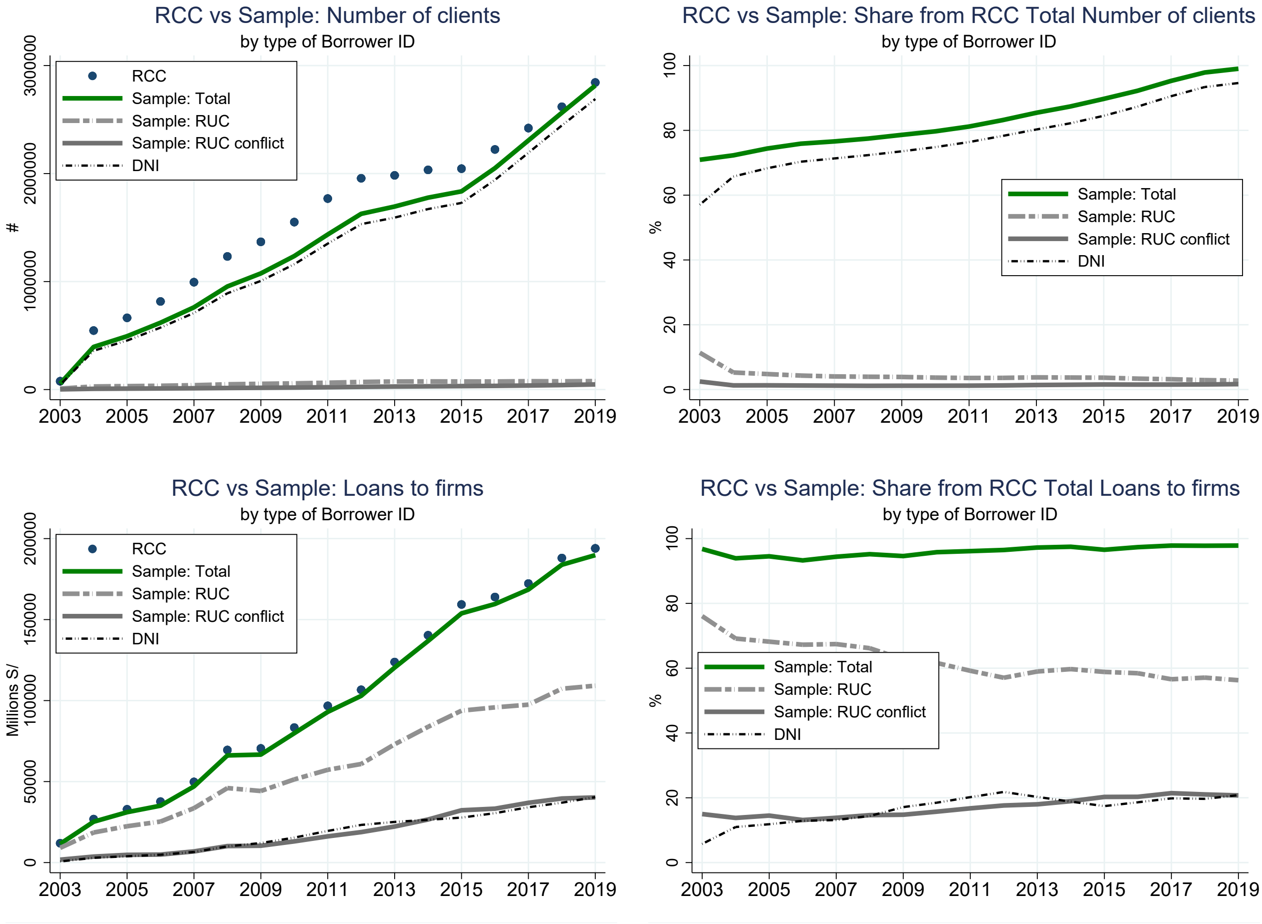
D Matching Process

Table 11: *Sources of information of RUC-Ubigeo and DNI-Ubigeo*

Source	# ID-Ubigeo	Share (%)	Cum. Share (%)
DNI: RCD	12,484,258	51.688	51.688
RUC: Sunat	5,897,657	24.418	76.106
RUC: Datos Peru	1,231,548	5.099	81.205
RUC: RCD	39,828	.165	81.37
RUC: Sunat & Datos Peru (No conflict)	2,496,240	10.335	91.705
RUC: Sunat & RCD (No conflict)	11,485	.048	91.752
RUC: Datos Peru & RCD (No conflict)	266,298	1.103	92.855
RUC: Sunat & Datos Peru & RCD (No conflict)	253,218	1.048	93.903
RUC: Sunat & Datos Peru (Conflict: Keep Sunat)	212,064	.878	94.781
RUC: Sunat & RCD (Conflict: keep Sunat)	68,517	.284	95.065
RUC: Datos Peru & RCD (Conflict: keep DP)	1,191,170	4.932	99.997
RUC: Sunat & DP & RCD(Conflict: Keep Sunat)	825	.003	100
Total	24,153,108	100	

Note: ID = RUC, DNI. RUC information from Datos Peru, RCD, Sunat. DNI information from RCD.

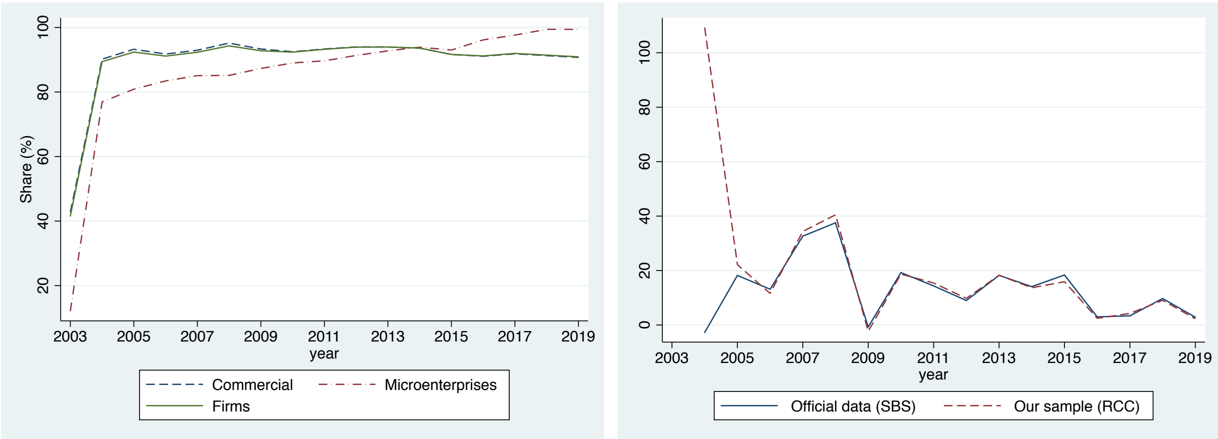
Figure 8: *Matching sample Representation: Formal Credit to Firms in RCC with Ubigeo, By type of Borrower ID*



Note: Annual Sample: 2003-2019.

E Granular Assessment: Figures and Tables

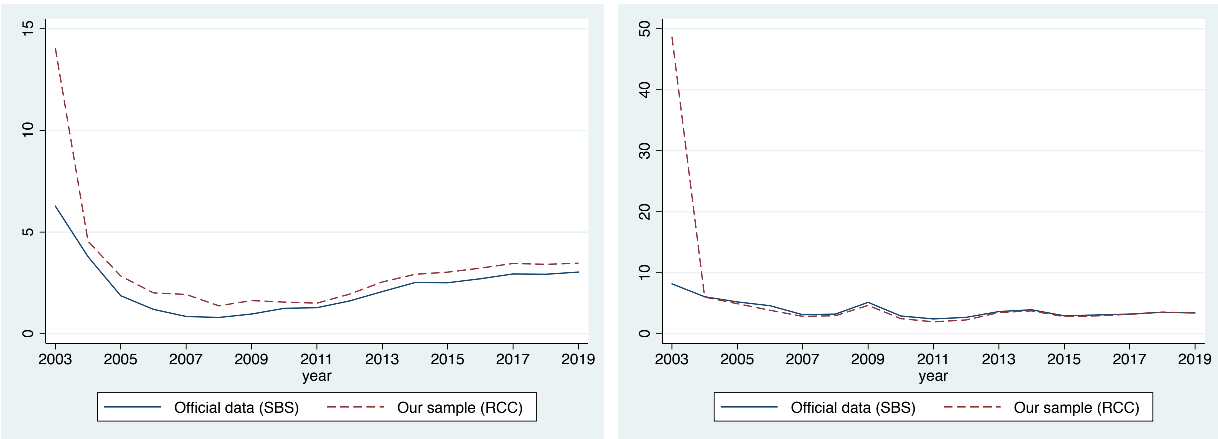
Figure 9: *Our sample vs official data: Aggregate level of bank loans*



(a) Sample share of SBS official data (b) Growth of loans to firms

Annual sample 2003-2019. SBS official data and our sample. Correlation of the growth of loans to firms of the SBS official data and our sample is 0.99 in the 2005-2019 period.

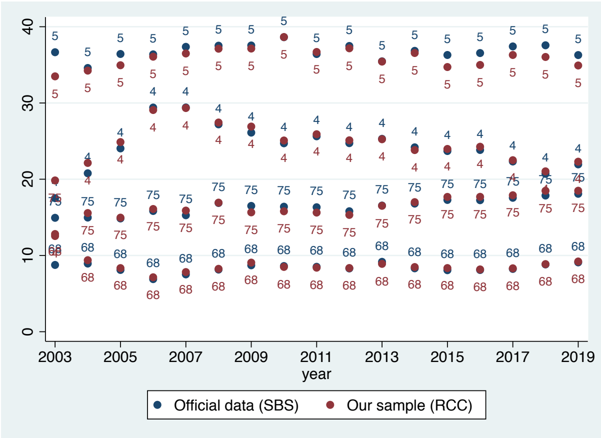
Figure 10: *Our sample vs official data: Aggregate level of bank non-performing loans (NPL) ratio*



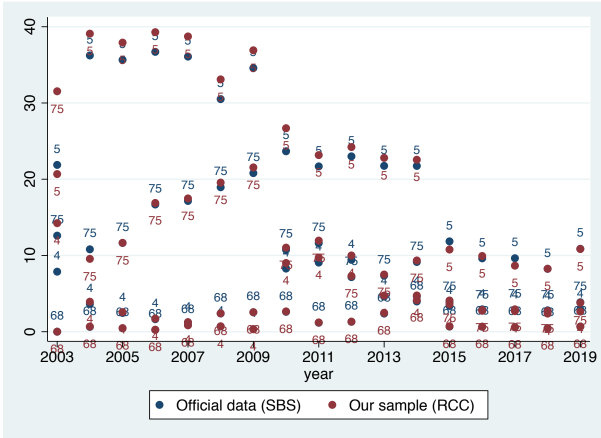
(a) Commercial loans (b) Loans to microenterprises

Annual sample 2003-2019. SBS official data and our sample. Correlation of the NPL ratio of the SBS official data and our sample is 0.97 for commercial loans and 0.98 for loans to microenterprises in the 2004-2019 period.

Figure 11: *Our sample vs official data: Bank credit shares (> 6%)*



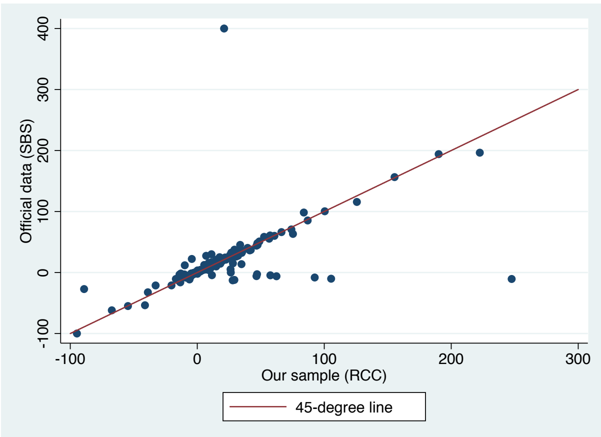
(a) Commercial loans



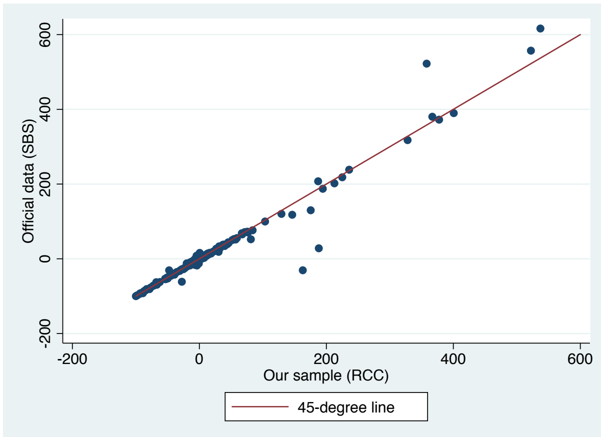
(b) Loans to microenterprises

Annual sample 2003-2019. The graphs show the credit shares of banks in SBS Data vs credit shares in the RCC sample. Name of banks: 5= BCP, 4=BBVA, 76=SCOTIABANK and 69=INTERBANK.

Figure 12: *Our sample vs official data: Bank credit growth*



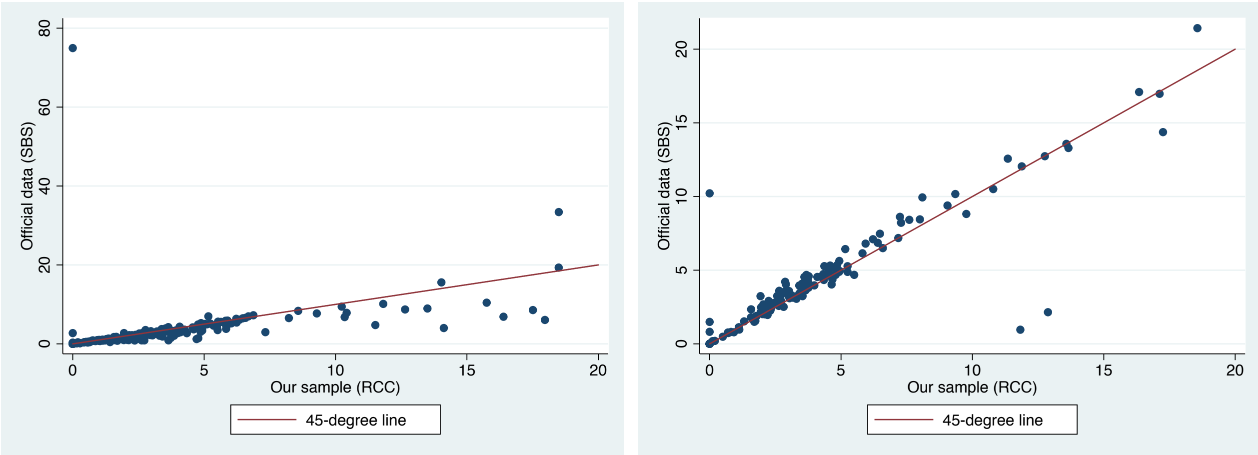
(a) Commercial loans



(b) Loans to microenterprises

Annual Sample 2004-2019. For illustrative purposes, figures omit some observations that are very outside the ranges showed. In total we omit 12 and 16 observations for commercial loans and loans to microenterprises, respectively. The final observation numbers at the bank-region-time level are 187 and 138 for commercial loans and loans to microenterprises.

Figure 13: *Our sample vs official data: Bank NPL ratio*



(a) Commercial loans

(b) Loans to microenterprises

Annual Sample 2004-2019. For illustrative purposes, figures omit some observations that are very outside the ranges showed. In total we omit 5 and 20 observations for commercial loans and loans to microenterprises. The final observation numbers at the bank-region-time level are 207 and 146 for commercial loans and loans to microenterprises.

Figure 14: *Sample share of SBS data across regions*

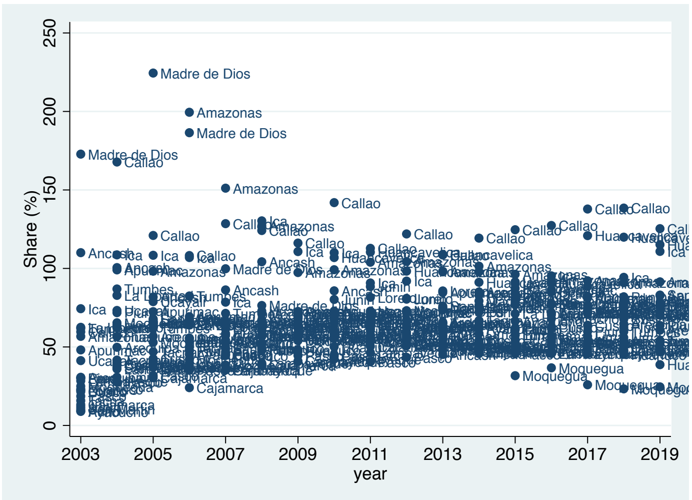
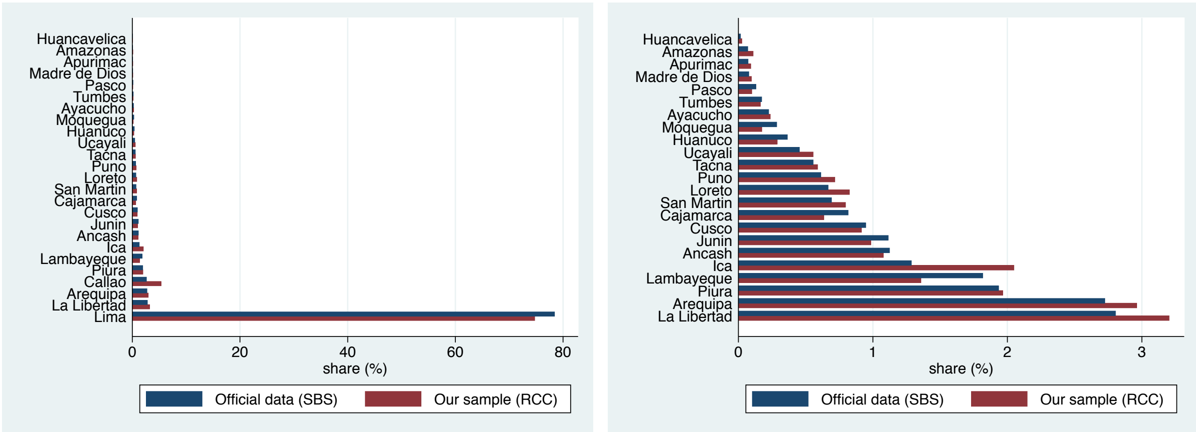


Figure shows the ratio of loans to firms (commercial loans and loans to microenterprises) of our sample (RCC) and total (all types) official loans (SBS) across regions and time. This is since in the SBS website there is not available credit by type at regional level. We limit of illustration to shares smaller than 300%, so we omit two observations.

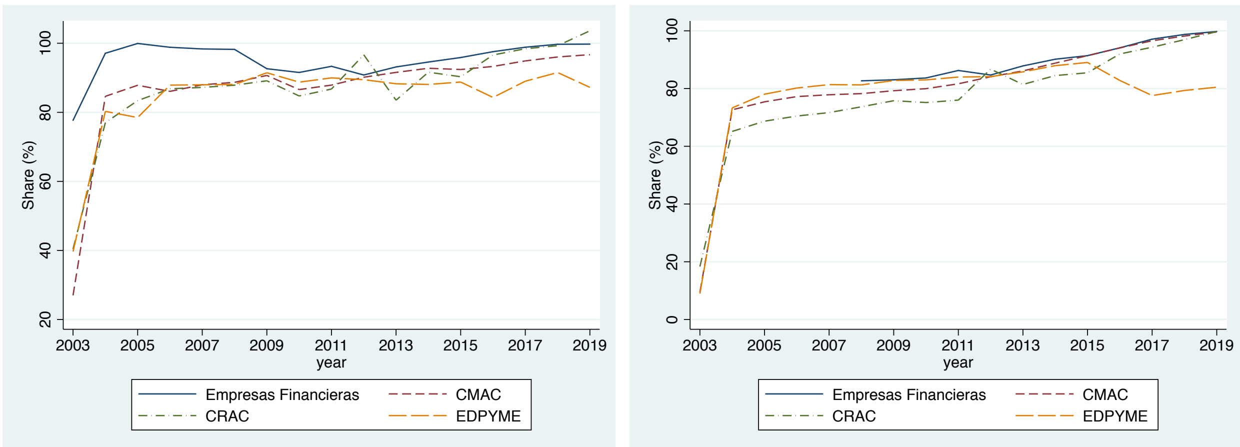
Figure 15: *Total loans distribution across regions: 2004-2019*



(a) All regions (b) Without Lima and Callao

Figure shows the ratio of average loans to firms (commercial loans and loans to microenterprises) in a region to average total loans in all regions for our sample (RCC) and total (all types) official loans (SBS) for the 2004-2019 period. This is since in the SBS website there is not available credit by type at regional level.

Figure 16: *Our sample vs official data: sample share of SBS data - non-bank groups*



(a) Commercial loans (b) Loans to microenterprises

F Robustness: Granular Assessment - Banks

Table 12: *Banks - Robustness*

exo_var	ln (# banks)			C4			HHI		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(i) 0.05%<NPL<94%									
endo_var$_{icrt-1}$	0.368***	0.373***	0.345***	0.369***	0.373***	0.344***	0.368***	0.371***	0.344***
exo_var$_{icrt-1}$	2.618***	2.588**	2.358**	65.05***	80.85***	70.34***	0.773	0.813	0.576
exo_var$^2_{icrt-1}$	-0.643***	-0.611**	-0.569**	-35.23***	-44.02***	-38.18***	-1.468*	-1.541	-1.285
Observations	4,171	4,171	4,160	4,171	4,171	4,160	4,171	4,171	4,160
R-squared	0.330	0.395	0.397	0.330	0.397	0.398	0.330	0.396	0.398
F test (ρ-value)	0	0	0	0	0	0	0	0	0
(ii) Without Metropolitan Area									
endo_var$_{icrt-1}$	0.441***	0.443***	0.401***	0.444***	0.447***	0.403***	0.439***	0.442***	0.399***
exo_var$_{icrt-1}$	2.222**	2.662**	1.847	41.04**	46.83**	53.38***	0.0507	0.270	-0.298
exo_var$^2_{icrt-1}$	-0.459	-0.522	-0.340	-22.26**	-25.66**	-29.24***	-1.039	-1.154	-0.659
Observations	3,726	3,726	3,723	3,726	3,726	3,723	3,726	3,726	3,723
R-squared	0.296	0.372	0.376	0.293	0.369	0.375	0.297	0.371	0.377
F test (ρ-value)	0	0	0	0	0	0	0	0	0
(iii) Only commercial loans									
endo_var$_{irt-1}$	0.411***	0.432***	0.400***	0.411***	0.428***	0.399***	0.412***	0.432***	0.401***
exo_var$_{irt-1}$	1.017	0.910	0.941	51.56**	83.00**	67.85***	-0.00362	3.835	0.145
exo_var$^2_{irt-1}$	-0.153	0.0653	-0.133	-29.15**	-45.74**	-38.25***	-2.246	-4.862	-2.762
Observations	2,282	2,273	2,271	2,282	2,273	2,271	2,282	2,273	2,271
R-squared	0.374	0.514	0.429	0.374	0.514	0.430	0.377	0.514	0.434
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

*** statistically significant at 1%, ** statistically significant at 5%, * statistically significant at 10%. Robust standard errors in columns (1), (4) and (7). Clustered (at region-time level) standard errors in columns (2), (5) and (8). Clustered (at financial institution-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. Metropolitan Area: Lima and Callao

G Robustness: Granular Assessment - All institutions

Table 13: *All financial institutions - Non-competition across groups - Robustness*

exo_var	ln (# institutions)			C4			HHI		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(i) 0.05%<NPL<94%									
endo_var$_{icrt-1}$	0.407***	0.409***	0.396***	0.407***	0.409***	0.396***	0.407***	0.409***	0.396***
exo_var$_{icrt-1}$	-0.0571	-0.0715	0.0936	-5.460***	-5.326**	-2.730	-0.315	-0.482*	-0.364
exo_var$^2_{icrt-1}$	0.0207	0.0250	-0.00502	3.102***	2.986**	1.477	0.309	0.457**	0.249
Observations	14,174	14,174	14,156	14,174	14,174	14,156	14,174	14,174	14,156
R-squared	0.367	0.389	0.449	0.368	0.389	0.449	0.367	0.389	0.449
F test (ρ-value)	0	0	0	0	0	0	0	0	0
(ii) Without Metropolitan Area									
endo_var$_{icrt-1}$	0.445***	0.447***	0.423***	0.445***	0.446***	0.424***	0.445***	0.447***	0.424***
exo_var$_{icrt-1}$	-0.0537	-0.0597	0.0791	0.823	7.740	13.86	-0.353	-0.398	-0.553
exo_var$^2_{icrt-1}$	0.0218	0.0248	0.00324	-0.145	-3.894	-7.439	0.313	0.356	0.333
Observations	12,203	12,203	12,191	12,203	12,203	12,191	12,203	12,203	12,191
R-squared	0.347	0.371	0.435	0.347	0.372	0.435	0.347	0.371	0.435
F test (ρ-value)	0	0	0	0	0	0	0	0	0
(iii) Only commercial loans									
endo_var$_{irt-1}$	0.440***	0.452***	0.429***	0.439***	0.451***	0.428***	0.439***	0.451***	0.429***
exo_var$_{irt-1}$	-0.0718	-0.142	0.175	-6.481*	-5.525	-4.087	-0.750*	-0.709	-0.618
exo_var$^2_{irt-1}$	0.0208	0.0346	-0.0361	3.596*	2.991	2.233	0.701*	0.759**	0.383
Observations	6,169	6,164	6,101	6,169	6,164	6,101	6,169	6,164	6,101
R-squared	0.423	0.478	0.509	0.423	0.479	0.509	0.423	0.479	0.510
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

*** statistically significant at 1%, ** statistically significant at 5%, * statistically significant at 10%. Robust standard errors in columns (1), (4) and (7). Clustered (at region-time level) standard errors in columns (2), (5) and (8). Clustered (at financial institution-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. Metropolitan Area: Lima and Callao

Table 14: *All financial institutions - Competition across groups - Robustness*

exo_var	ln (# institutions)			C4			HHI		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(i) 0.05%<NPL<94%									
endo_var$_{icrt-1}$	0.403***	0.405***	0.392***	0.406***	0.407***	0.395***	0.407***	0.408***	0.396***
exo_var$_{icrt-1}$	0.204	0.346	0.274	2.468**	3.777***	2.932***	-1.107	-1.312*	-0.951
exo_var$^2_{icrt-1}$	0.0147	-0.00655	0.00523	-1.972***	-2.948***	-2.302***	1.087	1.449	0.684
Observations	14,174	14,174	14,156	14,174	14,174	14,156	14,174	14,174	14,156
R-squared	0.370	0.391	0.451	0.368	0.390	0.450	0.368	0.390	0.449
F test (ρ-value)	0	0	0	0	0	0	0	0	0
(ii) Without Metropolitan Area									
endo_var$_{icrt-1}$	0.440***	0.441***	0.419***	0.444***	0.445***	0.423***	0.444***	0.446***	0.423***
exo_var$_{icrt-1}$	0.648	1.029	0.843	3.493**	5.515***	3.751**	-0.771	-0.850	-0.381
exo_var$^2_{icrt-1}$	-0.0462	-0.105	-0.0828	-2.701**	-4.101***	-2.872**	-0.134	0.139	-0.958
Observations	12,203	12,203	12,191	12,203	12,203	12,191	12,203	12,203	12,191
R-squared	0.349	0.374	0.437	0.347	0.372	0.436	0.347	0.372	0.436
F test (ρ-value)	0	0	0	0	0	0	0	0	0
(iii) Only commercial loans									
endo_var$_{irt-1}$	0.437***	0.448***	0.428***	0.437***	0.448***	0.428***	0.437***	0.449***	0.428***
exo_var$_{irt-1}$	0.331	2.888	0.472	2.760	9.995	2.559	-1.515	-1.900	-1.446
exo_var$^2_{irt-1}$	-0.0156	-0.392	-0.0418	-2.155	-6.927	-2.171	0.365	1.421	-0.0866
Observations	6,169	6,164	6,101	6,169	6,164	6,101	6,169	6,164	6,101
R-squared	0.423	0.480	0.511	0.423	0.479	0.510	0.424	0.479	0.512
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

*** statistically significant at 1%, ** statistically significant at 5%, * statistically significant at 10%. Robust standard errors in columns (1), (4) and (7). Clustered (at region-time level) standard errors in columns (2), (5) and (8). Clustered (at financial institution-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. Metropolitan Area: Lima and Callao