Bank Competition and Risk-Taking*

Jorge Pozo† & Youel Rojas‡

July, 2020

Abstract

This paper studies empirically the relationship between competition in the loan market and risk-taking in the Peruvian financial system. This work is motivated by the theoretical work of Martinez-Miera y Repullo (MMR, 2010) that finds a U-shaped relationship between competition and risk-taking and the empirical work of Jimenez, Lopez and Saurina (JLS, 2013) that finds support for that nonlinear relationship in Spain. In contrast to those studies, our findings suggest an inverted U-shaped relationship between competition and risk-raking for an emerging economy as Peru.

Keywords: Competition, bank risk-taking, financial stability.
JEL Classification: G21, E44, L11

1 Introduction

This work is motivated by the theoretical work of Martinez-Miera y Repullo (MMR, 2010) that find a U-shaped relationship between competition and risk-taking and the empirical work of Jimenez, Lopez and Saurina (JLS, 2013) that finds support for that nonlinear relationship in Spain. However, we depart from JLS 2013 since we aim to test the hypothesis of MMR 2010 in an emerging economy as Peru and also we make use of more granular data in addition to the standard bank-level employed in JLS 2013, to control for unobserved factors that can bias the results.

In a first part, we estimate a model similar as in JLS 2013 with bank-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 form the webpage of Superintendency of Banking, Insurance and Private Pension Fund Administrators (SBS), while for the second specifications we

*The views expressed in this paper do not necessarily represent those of the Central Reserve Bank of Peru.

†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

use, and a more granular data from the Credit Registry Data (RCC), which is a restricted information. As in JLS 2013, the main measure of competition is an indicator that reflects the number of relevant competitors that a financial institution faces. It is constructed as a weighted average of the number of financial institutions present in the regions where it operates, considering its loans to each regions as weights. And our measure of risk-taking is the non-performing loan ratio, which is taken from the SBS or it is built using the granular data from the RCC.

In the model with bank-level data, the results suggest that for Peruvian banks there is an inverted U-shaped relationship between competition and risk-taking, unlike MMR 2010 and JLS 2003, although it is still not significant. However, when we include also information for non-banks financial institutions the relationship becomes statistically significant even when considering fixed effects over time.

Since the previous results cannot control for omitted variables that may affect the dynamics of the relationship between competition and the underlying risk of the borrowers: such as changes in business opportunities and risk profiles at the departmental level and/or market strategies and diversification of financial institutions over time, micro data at the client-bank level is used to build a panel with region-bank-time dimensions. Our analysis starts from the assumption that segmented regional loan markets to achieve identification, and adopts within-region and within-bank estimators. Furthermore, only the category of commercial loans is considered and it is assumed that there is competition between the different groups of financial institutions. The results indicate the significant existence of an inverted U-shaped relationship between competition and risk-taking.

As a result, in contrast to MMR 2010 and JLS 2003, we find empirical evidence in an emerging economy as Peru of an inverted U-shaped relationship between competition and risk-taking.

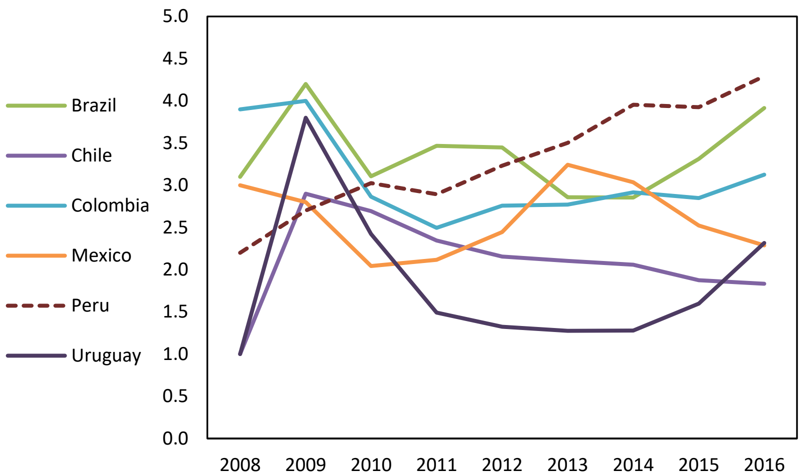
In Latin America, the Peruvian banking sector is more concentrated than in other countries. Only the banking system in Colombia is more concentrated than in Peru. Table 1 shows different measures of concentration and competition. For instance, in 2016 the share of the three largest banks assets in Peru was 71.9%, while in Chile it was 43.2%. The similar is observed with the 5-bank asset concentration measure. Also, in Peru in 2014 the elasticity of bank revenues to input prices (H-statistic) was one of the smallest, providing evidence of relatively low competition in the Peruvian banking system. Finally, the markup is largest in the Peruvian banking system suggesting relatively poor competition levels within Latin America. 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).

Table 1: Bank competition and concentration in Latin America

	3-bank asset concentration (%) 2016	5-bank asset concentration (%) 2016	H-statistic 2014	Lerner index 2014
Brazil	69.8	85.0	0.72	0.21
Chile	43.2	69.3	0.77	0.25
Colombia	78.7	89.4	0.51	0.48
Mexico	52.6	68.0	0.83	0.38
Peru	71.9	87.5	0.60	0.50
Uruguay	69.2	88.2	0.80	0.19
EME	63.2	75.9	0.57	0.35
AE	67.3	81.9	0.64	0.27

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. EME and AE correspond to simple averages across emerging market economies and advanced economies, respectively.

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



Source: Global Financial Development.

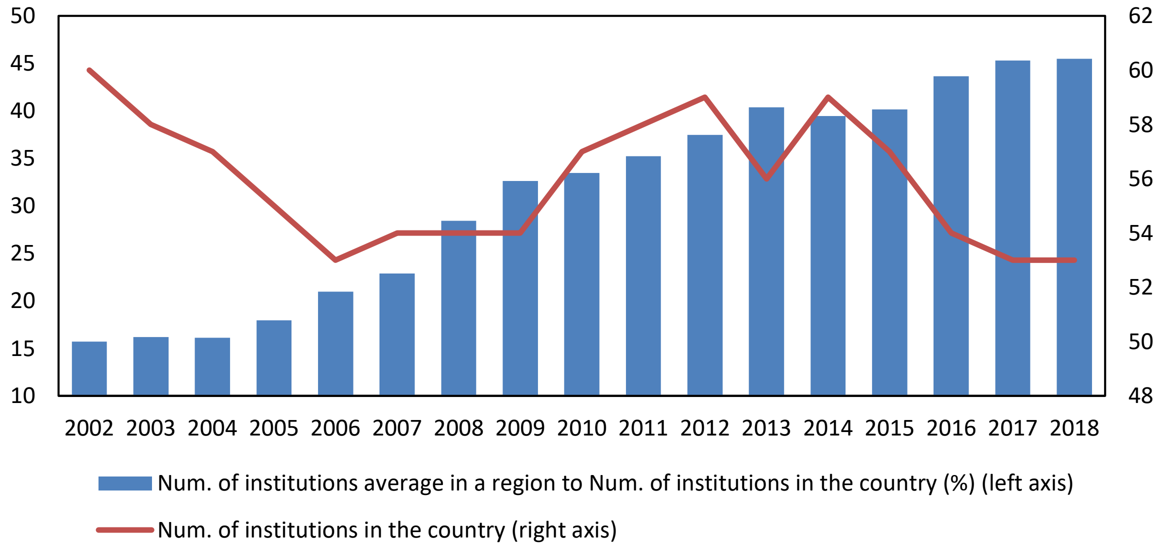
In 2018, 53 financial institutions that provide loans to the private sector constitute the Peruvian financial system. These were composed by 16 banks, that represent the 88% of the total loans, and other non-banks financial institutions as 12 municipal savings and credit funds (CMACs by its Spanish acronym), 6 rural savings and credit funds (CRACs by its Spanish acronym), 9 development entities for small and micro enterprise (EDPYMEs by its Spanish acronym) and 10 *empresas financieras* (other small financial institutions that take deposits and issue credits).[1] Even though from a country perspec-

[1]Note that the financial system also includes other more specialized institutions, however, since the

tive banks are relatively more important, this is not necessarily true within a region.[2] 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 there is not a clear trend in the number of total financial institutions in the country by the end of each year. 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 15% since 2002 to 45% in 2018. In other words, in 2018 on average a region has the presence of 45% 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 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.

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

The remainder of this chapter is partitioned as follows. Section 2 presents the lit-

participation in the credit market is still very small, we omit them.

[2]Peru is split into 24 regions, but as is explain later we consider 25 regions by considering the foreign market as an additional region served by the Peruvian financial system.

erature review. Section 3 shows and data y model description. Section 4 discusses 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 decline. 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.[3]

Boyd and De Nicolo (2005) show that there exist 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 increase banks have less market power to raise loan rates and hence with smaller loan interest rates borrowers choose lower risk projects. Due to the perfect correlation, loan default probability coincides with bank default probability. As a result, through that mechanism, as competition increases bank failure probability decreases.

Martinez-Miera and Repullo (2010) argue that the findings of Boyd and De Nicolo (2005) does not necessarily hold in the more realistic case of imperfect correlation of loan default. This is because bank competition reduces the interest rate pay from performing loans, which provides a buffer to cover loan losses, and hence increases the bank default probability. They identify a risk-shifting effect, which is the one described in Boyd and Nicolo (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 reduces bank capacity to avoid defaulting.

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

They find that in a very competitive market the margin effect dominates, while in a less competitive market the risk-shifting effects dominate. As a result, they formulate a U-shaped relationship between number of banks (bank competition measure) and the risk of bank failure.

The empirical work of JLS 2013 using annual Spanish data and different measures of lending competition for the 1988-2003 period supports the nonlinear relationship found in MMR 2010. We depart from JLS 2013 since we test the hypothesis of MMR 2010 in an emerging economy as Peru and also using granular data.

3 Data and Model Description

We use two levels of information. The first dataset is a bank-level data as used in the related literature and the second dataset is a bank-region level data, which is more granular that it allows us to control for demand and supply characteristics that can bias our results. In the following we describe the data in its first level and the model description and in section 5 we focus on the more granular data.

3.1 Data

Similar to JLS 2013 we use different bank-level weighted average measures of bank competition 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), and the Herfindahl index (HHI), which is the sum of banks' squared market shares in loans in each region. Peru is slit in 24 regions, however as part of our analysis we include the foreign marker as an additional region. Peruvian financial institutions serving foreign markets face higher competition as they encounter as competitors larger international financial institutions. 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. we use the four-largest financial institutions instead of the three or the five since in Peru the four-largest bank represents almost the 90% of the total loans. 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, the public information at the regional level is only about total credit, and there is no disaggregation by type of credit or status of credit to firms and households.

Since the Peruvian credit market is segmented geographically into 25 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 its credit competitors in the representative region (or representative market) where it is operating. This bank-competitors measure is calculated as a weighted average (by total loans) of bank-region-competitors across all of a bank's regional operations. 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 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 institution i grants loans at year t. The HHI in each region is computed as the sum of squared market shares in loans of financial institutions granted in the region.

Also, we include data to control for individual bank characteristics, as return on assets (ROA) and bank size (the market share of loans in a country level). We also control for aggregate trends, such as the Peruvian business cycle. Three control variables not included in JLS 2013 are bond issued by non-financial institutions to credit ratio, the risk weighted asset (RWA) to capital ratio and participation of foreign debt on credit funding. The first is to control for the preferences and/or opportunities for non-bank funding, while the second is to control for individual bank characteristics regarding bank capacity to handle a financial crisis. Since banks might hold buffers, the RWA to capital ratio does not necessarily only reflect the risk of bank loans portfolio, but also bank preferences or capacities for handling a crisis.[4] The third is to consider the capacity of bank from borrowing from foreign markets, which in turn might affect banks' incentives to take 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,

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

Information about nonperforming loan (NPL) ratio at an institution-time level is also provided by the SBS.

We also assess the relationship between competition and risk-taking considering the five main financial groups: banks, CAMCs, CRACs, EDPYMES and *empresas financieras*, that exist in the Peruvian financial system. In this case, the construction of the different competition measures for any financial institution that belongs to any of the financial

[4]See, the leverage channel explained in Agur et al. (2015) and Dell'Ariccia et al. (2014).

group follows the same procedure provided for banks. In this case, for instance, we can built the indicator "number of institutions". We use annual data and the period of study is 2004-2018. And we made the analysis in two steps: we start focusing on banks and then on all financial institutions.

Tabla 2 presents the descriptive statistics for the variables when considering only banks. There are in total 20 banks for the 2004-2018 period and 210 bank-year observations.[5] The average of the NPL ratio is 2.83% and it features a large degree of dispersion. Regarding one of our competition measures, "number of banks", the average number of banks that exists in the representative region where a bank competes is 13.67. This variable also exhibits a relatively high degree of dispersion. In general, the control variables report high degree of dispersion.

Tabla 3 presents the describe statistics for variables used for all financial institutions. In the period of study 2004-2018 we consider 72 financial institutions across the five financial groups. In this case, the competition measures can be computed under two different assumptions: no competition across groups, and competition across groups. In the former, I assume financial institutions only compete with those within its financial group, while in the latter case financial institutions can compete with institutions from any financial group. Regarding one competition measure, "number of institutions", the average number of financial institutions that exists in the representative region where a financial institution compete is 7.3 when no competition across groups is assumed, while this is 31.6 when we allow competition across groups.

Table 2: Description statistics for bank-year observations

Variables	Obs	Mean	S.D.	Minimum	Maximum
NPL_{it}	210	2.83	2.00	0.00	11.00
Number of banks_{it}	210	13.67	1.66	9.01	16.00
C4_{it}	210	0.83	0.02	0.74	0.86
HHI_{it}	210	0.21	0.02	0.17	0.31
Size_{it}	210	0.07	0.10	0.00	0.35
ROA_{it}	206	0.02	0.02	-0.13	0.08
$\text{For\_debt\_cred}_{it}$	210	0.10	0.09	0.00	0.74
RWA_Cap	210	7.09	1.49	0.94	10.03
GPD_rg_{t}	15	0.05	0.02	0.01	0.09
Bond_cred_{t}	15	0.12	0.06	0.06	0.24

Source: SBS. Own elaboration.

[5] This is the final number of bank-year observations after allowing for lags in the variables is 196.

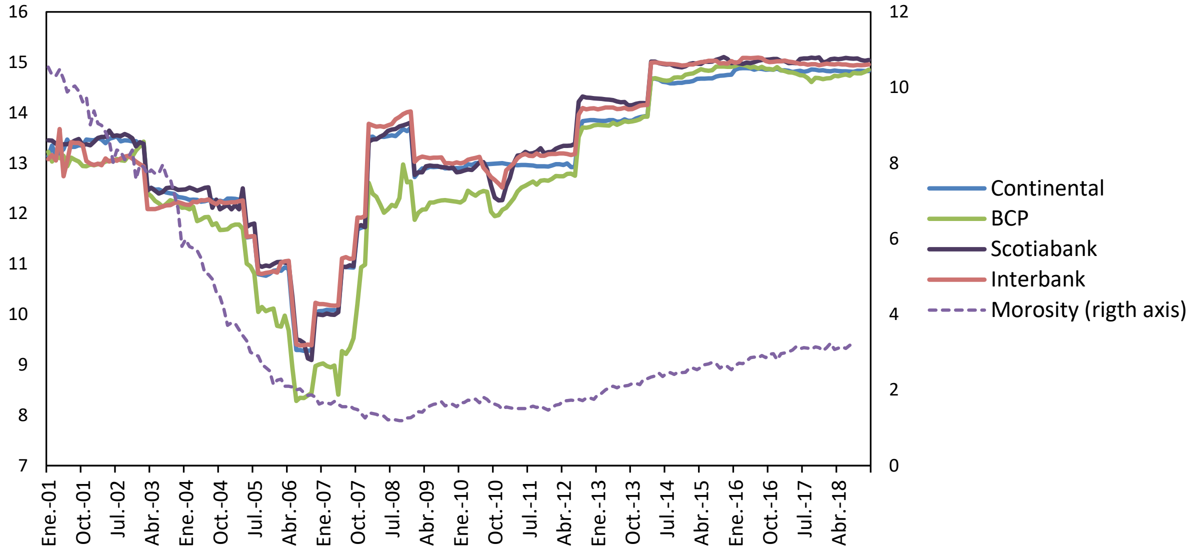
Table 3: Description statistics for financial institution-year observations

Variables	Obs	Mean	S.D.	Minimum	Maximum
NPL_{it}	820	5.39	5.71	0.00	100.00
No competition across groups					
Number of institutions_{it}	823	7.33	4.40	1.00	16.00
C4_{it}	823	0.91	0.08	0.70	1.00
HHI_{it}	823	0.40	0.21	0.17	1.00
Competition across groups					
Number of institutions_{it}	823	31.66	11.31	6.00	50.00
Size^{*}_{it}	823	0.72	0.06	0.53	0.89
C4_{it}	823	0.72	0.06	0.53	0.89
HHI_{it}	823	0.16	0.02	0.10	0.29
Size^{**}_{it}	821	0.02	0.05	0.00	0.31
ROA_{it}	810	0.02	0.04	-0.39	0.17
$\text{For\_debt\_cred}_{it}$	820	0.12	0.19	0.00	1.08
Lev_{it}	824	6.08	1.82	0.47	10.93
$\text{Bond\_cred}^{**}_{t}$	15	0.11	0.06	0.05	0.21

Source: SBS and own elaborations. * Total credit: The credit of each group. **Total credit: The credit of all five groups.

Figure 3 reports the behavior of "number of banks" for the four largest banks from 2001 to 2018. In general, there is 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. Also, from 2002 to 2006, there was a morosity rate 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 morosiy rate. Since 2008 bank competition and morosity rate has 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 was 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.

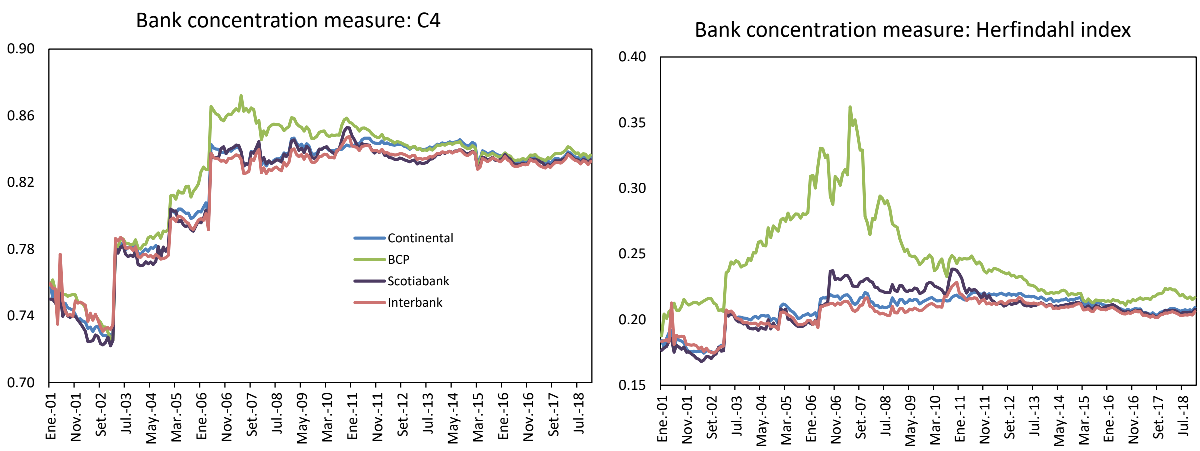
Figure 3: Bank competition measure "number of banks" and morosity rate



Source: SBS and own calculations. Morosity rate = Banking sector non-performing loans share (SBS criterion). Period January 2001 - August 2018.

Other two alternative measures for competition, typically used in the literature, are the C4 and the HHI. Although these are measures of concentration, literature typically use those as competition measures as well. As in the previous competition measure, figure 4 display C4 and HHI measures 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 were the concentration level (measure with C4 or Herfindahl index) was relatively high, than in those regions where the other banks operate.

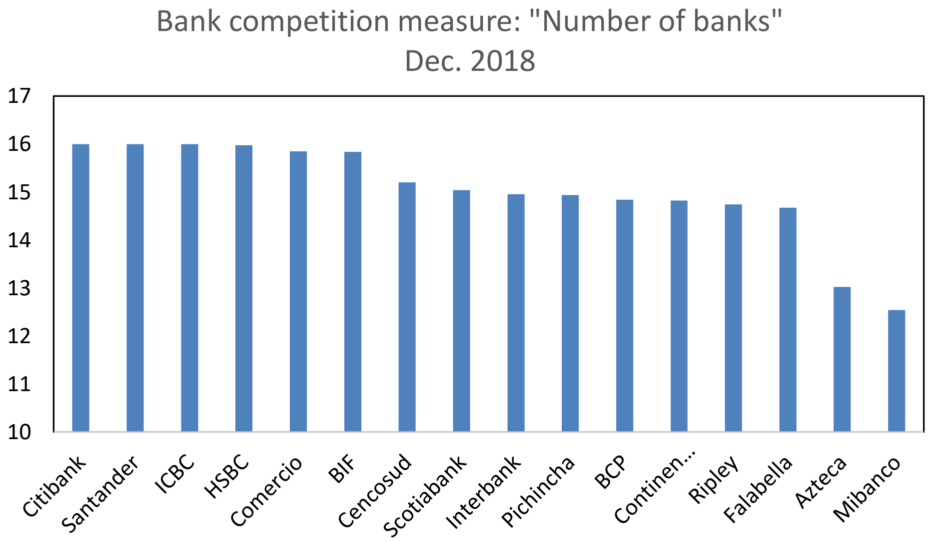
Figure 4: Bank concentration measures



Source: SBS. Own calculations.

In order to have an idea of the heterogeneity of bank competition, figure 5 reports the "number of banks" at December 2018 across banks. 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, Citibank and ICBC 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



Source: SBS. Own calculations. Number of banks: the number of banks that has the representative region for bank i at time t, calculated as the weighted average (by total loans) of the number of banks over all the regions where banks grant loans.

3.2 Model Description

Similar to JLS 2013 the model is as follows:

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

where the i subscript refers to a financial institution and 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.

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 (1). As in JLS 2013 we include the lagged dependent variable as an explanatory variable.

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 JLS 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 competitive measure and if β_1 is negative and β_2 is positive, the results would support the U-shaped pattern proposed in the MMR model, which was supported in JLS 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 ($cont_{it}$) we include business cycle conditions by introducing the current and lagged values of the annual real GDP growth rate, rg_gdp and $L.rg\_gdp$, respectively. 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 ($adeu\_cred$), bond issued by non-financial institutions to credit ratio ($bond\_cred$), and the RWA to capital ratio (lev). These five variables are introduced as lagged values.

4 Empirical results

4.1 Correlations

Table 4 reports the pairwise correlations between the variables considering only banks. We find a negative correlation with our measures of competition and our measure of bank risk-taking (NPL ratio). Specifically, the relationship between the number of banks and NPL ratio is negative and significant, while the correlation between the C4 or HHI with the NPL is positive but not statistically significant.

As expected, the NPL ratios are negatively and significantly correlated with the real GDP growth (the business cycle). The bank size is negatively and significantly correlated with NPL ratios. It suggests that relatively larger banks are less motivated to take risk. Also, NPL ratios are negatively and significantly correlated with the foreign debt share. This could be because the relatively larger banks have higher access to international credit market. Regarding the correlations when considering the five groups, figure 15 in the Appendix C shows the correlation coefficients when considering the five groups. The results are very similar.

Table 4: Correlation coefficients

Variables										
NPL_{it}	1									
Number of $banks_{it}$	-0.1163	1								
$C4_{it}$	0.0134	0.021	1							
HHI_{it}	0.0269	-0.4250*	0.6132*	1						
$Size_{it}$	-0.2463*	-0.1974*	0.1708*	0.5516*	1					
ROA_{it}	0.0149	-0.108	0.2149*	0.2010*	0.1384*	1				
For_debt_$cred_{it}$	-0.3704*	0.072	0.0221	-0.0449	0.1601*	-0.3058*	1			
GPD_rg_t	-0.2208*	-0.4402*	0.04	0.1567*	0.0258	0.0449	0.1009	1		
Bond_$cred_t$	-0.1117	-0.7387*	-0.5566*	-0.0215	0.0651	-0.0548	-0.0059	0.4057*	1	
RWA_Cap_{it}	-0.1976*	-0.2458*	0.0453	0.1740*	0.2174*	0.1369*	0.1931*	0.2020*	0.2230*	1

* Significant at the 5% level.

4.2 Regression Results

In this subsection we present the results when considering only banks and all financial institutions assuming no 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 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 (foreign debt to credit ratio) has a positive and statistically significant association with risk-taking only when considering banks. Also, the bonds issued by non-financial institution to credit ratio is negatively and statistically significant associated with risk-taking. Finally, RWA to capital ratio 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 5 reports the estimation results for the model using the annual data for the 2004-2018 period considering only banks. The table shows the results for nine different regression. 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 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. The number of banks that maximizes the NPL ratio, as a measure of bank risk-taking, is 9.6 with no fixed effects (for column 1) and 9.8 with time fixed effect (column 3).

When using the concentration measures (C4 or HHI), in general, results suggest an U-shaped relationship between bank risk-taking and bank competition as suggested by MMR 2010. In the case of C4, this is only significant with bank fixed effects, while HHI is significant without fixed effects and with time fixed effects.

Table 5: Banks

exo_var	ln (# banks)			C4			Herfindahl index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
L.endo_var	0.759***	0.488***	0.770***	0.822***	0.477***	0.809***	0.792***	0.469***	0.807***
L.exo_var	10.27**	1.726	12.30*	-2.364	134.6**	-44.61	65.07***	12.53	66.03***
L.exo_var2	-2.271**	-0.238	-2.699*	4.430	-85.71**	32.91	-131.6***	-39.65	-133.0***
L.roa_1	-1.808	-3.331*	-2.022	-2.223	-3.440**	-2.386	-2.108	-3.347**	-2.349
L.size_1	-0.823***	3.911	-0.812***	-0.686***	6.361	-0.868***	-0.971***	4.297	-0.962***
L.adeu_cred_1	0.735**	0.494	0.719**	0.850**	0.624**	0.827**	0.861**	0.398	0.901***
L.bond_cred_nonf_1	-3.265***	-0.767		-0.276	-2.261***		-0.579	-1.361*	
L.lev	0.0331	0.0577**	0.0202	0.0223	0.0577**	0.0123	0.0176	0.0688***	0.00551
rg_gdp_1	-3.879***	-3.024*		-3.270**	-2.557**		-3.535**	-3.219**	
L.rg_gdp_1	0.0419	-0.914		0.427	-0.334		0.594	-1.003	
Observations	196	194	196	196	194	196	196	194	196
R-squared	0.824	0.904	0.838	0.820	0.909	0.834	0.825	0.909	0.837
F test (ρ-value)	0	1.50e-10	0	0	1.73e-10	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%

Table 6 presents the estimation results for the model when considering all institutions and no competition across groups using the annual data for the 2004-2018 period. We control for group and for several financial events (e.g. reallocation of institutions from one group to another, mergers, adquisitions, etc.)[6]

Recall, no competition across groups means that each institution competes only with those within its group. For example, we assume that a bank cannot compete with an institution from the CAMC group. By assuming no competition across groups, we assume that the market is segmented regarding the borrowers. We think this is not a very realistic assumption, but it is more realistic than assuming that in a regional level two institutions

[6]For example, ceteris paribus a financial institution that moves from one group to another faces a different competition.

from different groups compete in the same intensity as two financial institutions from the same group.

As in table 5, we show the results for nine different regression 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 banks, 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 effects, while when considering institution fixed effects results are not significant, but keep the same direction. In this case the number of financial institutions that maximizes bank risk-taking is 2.9 without fixed effects (column 1) and 3.3 with time fixed effects (column 3).

When using the concentration measures (C4 and HHI), in general, results suggest an U-shaped relationship between bank risk-taking and bank competition as suggested by the literature, but these results are not statistically significant.

Appendix B reports the results when assuming competition across group. Interestingly, in this case, we do not find statistically significant estimates. This is evidence that at the regional level (or within a region) there is not significant competition between financial institutions from different groups.

Table 6: All financial institutions: No competition across groups

exo_var	ln (# institutions)			C4			Herfindahl index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
L.endo_var	0.771***	0.571***	0.771***	0.776***	0.573***	0.778***	0.774***	0.568***	0.775***
L.exo_var	0.163*	-0.140	0.178*	0.0717	-6.202	2.111	0.392	-0.975	0.430
L.exo_var2	-0.0763*	0.0792	-0.0773*	0.389	3.630	-0.782	-0.395	0.676	-0.435
L.roa	-1.015*	-0.415	-1.139	-1.057*	-0.478	-1.182	-0.966	-0.375	-1.110
L.size_no_comp	0.0115	0.726**	0.0202	0.0200	0.414	0.0252	0.0811	0.719**	0.0861
L.adeu_cred	0.116	0.132	0.108	0.130	0.154	0.119	0.119	0.146	0.109
L.bond_cred	-1.278***	-1.079**		-1.240***	-1.566***		-0.987***	-1.293***	
L.lev	0.0193*	0.0329**	0.0204*	0.0208*	0.0295**	0.0219*	0.0208*	0.0318**	0.0218*
Observations	783	781	783	783	781	783	783	781	783
R-squared	0.786	0.847	0.792	0.786	0.846	0.792	0.785	0.847	0.791
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%

5 Granular Assessment

Our previous analysis have several shortcomings, as it can not control for trends on demand factors other than the aggregate trends in the economy and all construction of variables is based on total credit and not only on credit to firms. Thus, we were not testing the same hypothesis as in MMR 2010 and JLS 2013.

In order to overcome these problems and avoid a bias estimator we make use of more granular data. In particular, we add another dimension to the institution-time data: "region".[7] This additional dimension allows us to control for local lending opportunities & bank level strategies. These might include demand and supply credit shocks, respectively.

5.1 Granular Data

The source of the granular on credit data is the Credit Registry Data (RCC). This is a loan-level data, that contains debt classification at client-level and at loan-level originated in the financial system .[8] The data is available in quarterly frequency for the 2003Q1-2010Q3 period and in monthly frequency for the 2010M10-2018M08 period. Debtors are identified by an SBS code, tax ID (RUC) and national ID (DNI).

To match the credit data with geographic location, in a region, we use the information of the debtor (Tax ID) together with information from Peruvian tax administration (SUNAT) data on individual and firm Tax ID (RUC) and Location codes (UBIGEO). [9]. The goal is to obtain a panel-data at bank-region-time level on credit and non-performing loans ratio. We identify a sample of all formal loans from the financial institutions.

Specifically, for the construction of any bank-region-time level variable, we proceed as follows:

1. Identify a sample of clients with RUC in RCC.
2. Match clients with RUC in RCC with Locational data from SUNAT.
3. Select loans provided to private non-financial firms $\rightarrow$ Loans by RUC and Location
4. Construct credit information, risk-taking measures and competition/concentration measures at bank-region-time level.

We make two strong assumptions. Note that we assume that loans go to the region registered as the location of the borrower. It could be that the registered location is

[7]When working with granular data we omit considering the foreign market as another region and hence we consider the 24 regions within the Peruvian territory.

[8]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.

[9]Once we have a UBIGEO, we use the Peruvian Bureau of statistics' information on location of a UBIGEO in a region.

different to the one where the debtors' activities are performed. However, we assume this is an odd case. Furthermore, we also assume that the loans located in a certain region are issued by an agency from the same region.

As before the risk-taking measure is given by the non-performing loans ratio, that is built using the SBS criterion (see, equation 1) but this time at the bank-region-time level. To construct the competition measures at bank-region-time level, the approach is similar than when construction the measures at institution-time level. However, this time 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 where the institution i, located in the region r, operates, calculated as the weighed average of the number of competitors over all the provinces where institution i operates. The weights are given by the loans granted to each of these provinces divided by the loans granted to region r by the institution i. Also, C4 at the bank-region-time level denotes the share of the largest financial institutions in the representative province of institution i located at region r, calculated as the weighed average of the C4 over all the provinces where institution i operates. Similarly, HHI at the bank-region-time level denotes Herfindahl index of concentration for the representative province of institution i located at region r, calculated as the weighed average of the C4 over all the provinces where institution i operates

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

Representativeness of our sample: We assess the representativeness of our sample and hence how well it matches the characteristics of the official data from the SBS. The left plot in figure 7 reports our sample as a share of SBS total credit in the banking system. Our bank loans sample represents between the 42 to 52% of total loans since 2004. The right plot suggests that our sample mimics fairly well the dynamics of total bank credit since 2005. Figure 8 reports that at bank-time level credit shares in our sample mimics the behavior of official data specially in large banks. According to figure 9 in aggregate our sample also mimics fairly well the dynamics of non-performing loans (NPL) ratio. In particular, the correlation between our sample and official data is 0.92. Figure 10 reports the relationship between the bank credit growth at bank-time level in the official data and in our sample. It suggests a strong positive correlation between these two. As a result, our sample mimics very well the credit growth at bank-time level. Similarly, figure 11 shows the relationship between the NPL ratio in the official data and in our sample. In general, the dynamics of NPL ratio in our sample and at bank-time level from SBS are fairly equal.

Distribution of bank loans among regions: In numbers, the official data (SBS) has 2 164 bank-region-time observations of loans, while in our sample there are 3 255 bank-region-time observations. There are 2 139 cases where both sources report loans for the same time-location. Loans that are not located in the regions where the SBS reports represent only the 0.3% of our total credit RCC sample. Figure 6 in Appendix C reports at region-time level the ratio of our sample bank credit to SBS bank credit. In general, the ratios are between 0 and 60% and these seems to keep constant across time.

Estimation results

Table 7 shows the regression as the reported in Table 5 considering only banks in the 2004-2017 period at bank-region-time level. As before, the coefficient of the lagged endogenous variable is significant. However, for the competition measures the coefficients that describe the nonlinear relationship between NPL and competition are not statistically significant.

Table 8 is identical to table 7 but is considers only bank loans to firms, consistent with the hypothesis in MMR 2010. When considering the number of banks, as the competition measure, the results validate the inverted U-shaped relationship between bank competition and bank risk-taking.

The coefficient estimates are significant when we do not control by demand (region-time fixed effects) and supply (bank-time fixed effects) shocks, column1, and when we control by supply shocks (column 3). Even though if the results are not significant when we control for region-time fixed effects, the sign of the coefficients are in concordance with the other estimates. The number of banks that maximizes bank risk-taking is 2.9 and 3.3, respectively. Regarding the other measures of competition we find significant estimates for HHI (column 7), which suggests a U-shaped relationship, but when we do not control for demand and supply shocks. Notice the importance of using only the credit to firms margin, as there are no sign or size contradictions in the estimates across the different specifications.

Table 7: Regression results: Banks

exo_var	ln (# banks)			C4			Herfindahl index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
lag_endo_var	0.468***	0.473***	0.446***	0.468***	0.470***	0.447***	0.465***	0.465***	0.444***
lag_exo_var	0.293	-2.762	-0.190	-1.659	6.181	-0.686	2.955	10.07*	1.937
lag_exo_var2	-0.0548	0.626	0.0609	0.784	-1.643	-0.160	-2.215	-8.316	-1.232
lag_size_bt_t	12.90***	13.56***		12.87***	13.53***		12.74***	13.15***	
lag_size_brt_bt	-0.902	-1.039	-0.739	-0.901	-1.032	-0.777	-0.812	-0.996	-0.690
Observations	2,344	2,336	2,337	2,344	2,336	2,337	2,344	2,336	2,337
R-squared	0.395	0.494	0.461	0.395	0.494	0.461	0.396	0.497	0.462
F test (ρ-value)	0	7.98e-09	3.41e-08	0	5.84e-09	2.90e-08	0	5.27e-09	2.95e-08
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

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

Table 8: Regression results: Banks - credit to firms

exo_var	ln (# banks)			C4			Herfindahl index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
lag_endo_var	0.451***	0.448***	0.430***	0.449***	0.447***	0.430***	0.448***	0.444***	0.429***
lag_exo_var	1.131**	1.157	1.157*	-23.03	-35.65	-20.12	2.668**	4.072	2.023
lag_exo_var2	-0.349**	-0.377	-0.329*	13.25	20.22	11.43	-2.539**	-3.905	-2.064
lag_size_bt_t	9.833***	10.01**		10.33***	10.33***		10.20***	10.17***	
lag_size_brt_bt	-1.361	-1.496	-0.610	-1.292	-1.550	-0.613	-1.353	-1.546	-0.617
Observations	2,612	2,594	2,597	2,612	2,594	2,597	2,612	2,594	2,597
R-squared	0.349	0.445	0.404	0.349	0.445	0.403	0.349	0.446	0.403
F test (ρ-value)	0	1.62e-07	0	0	7.95e-09	1.69e-09	0	7.68e-08	1.24e-09
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

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

5.3 Financial system

Given that at the regional level the role of nonbank financial institutions in lending activities becomes relatively more important than the role of banks, in this section we consider all lenders in the lending market, both bank and nonbank financial institutions.

As before in our paper financial system is defined by the following five groups banks, CMACs, CRACs, EDPYMES and *empresas financieras*. Figure 12 in Appendix C shows the representativeness of credit data by type: commercial credit, small firms credit, mortgage, personal credit.[10] It reports that our sample essentially represents loans to commercial firms. Figure 13 in Appendix C shows the credit growth by type of credit in the SBS official data and in our sample. It matches fairly well the dynamics of the commercial credit.

Table 9 reports the regression results of the model in the 2004-2017 period at institution-region-time level when considering the financial system (all financial institutions), competition even between financial institutions from different groups, and only loans to firms (commercial credit and small firm credit). When considering the number of financial institutions as our competition measure, results validate an inverted U-shaped relationship between bank competition and bank risk-taking. All coefficient estimates are strongly significant either we control by supply shocks (Institution-time fixed effects, column 2) or demand shocks (region-time fixed effects, column 3). However, when observing the concentration measure C4, we conclude the opposite, which is in line with suggested by the theoretical work of MMR 2010 and the empirical work of JLS 2013.

Results presented in table 9 and in table 6 suggest that the competition between financial institutions from different groups within provinces is relatively more related with risk-taking than competition within regions. This might be an expected result since there is a higher likelihood that two financial institutions compete if they are located in the save province than in the same region. This is because it is more likely that within a province they potential clients are the same.

In general we have found an inverted U-shaped relationship between the competition, measured as the number of financial institutions, and the bank risk-taking measure as the NPL ratio. This contrast with the obtained by MMR 2010 for a developed economy as Spain. And we believe that an explanation of this different results is the particular characteristic of an emerging economy with a still undeveloped financial system.

A different research question: The data that we are exploring allow us to respond to a different but very related research question, or to investigate the relationship between bank risk-taking and competition following a different approach. In particular, the RCC

[10]Due to data availability reasons we follow the credit classification that took place before July 2010. This is commercial credit includes loans to small-size, medium-size, large-size and corporate firms. Small firms loans includes loans to micro-size firms.

data will allow us to identify how the lending competition that exists at a region affects the level of risk-taking in that region. As before we obtain an inverted U-shaped relationship. See Appendix A for the econometric model and results.

Table 9: All institutions - credit to firms

exo_var	ln (# institutions)			C4			Herfindahl index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
lag_endo_var	0.222***	0.202***	0.127***	0.223***	0.203***	0.128***	0.222***	0.202***	0.126***
lag_exo_var	0.605	1.341***	0.950**	10.80***	16.77***	15.43***	0.537	1.568	1.200
lag_exo_var2	-0.158**	-0.298***	-0.213***	-6.876***	-10.73***	-9.804***	-1.435	-2.895**	-2.329**
Observations	5,889	5,881	5,753	5,889	5,881	5,753	5,889	5,881	5,753
R-squared	0.233	0.283	0.339	0.232	0.282	0.340	0.232	0.282	0.339
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

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

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 MMR 2010, there is an inverted U-shaped relationship. This results holds when studying only banks or all financial institutions. In addition, this result is robust when considering granular data and even when controlling for supply shocks. The competition between financial institutions from different financial groups within provinces is relatively more related with risk-taking than the competition within regions.

References

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

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

[3] 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.

[4] 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.

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

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

[7] Repullo, Rafael, 2002. *Capital requirements, market power, and risk-taking in banking*, Journal of Financial Intermediation 13, 156-1.

Appendix A A different research question

Our RCC sample also allows to responds a different question: How does lending competition in a region affects the risk of the loans? This is a different empirical approach to study the relationship between the risk-taking and competition.

To answer this question our indicators is only computed through two dimensions: region r and time t. Using our RCC sample, we compute the risk-taking of the representative financial institution that exists in a region at time t, and the number of competing financial institutions for each region and time. In particular, the risk-taking of the representative financial institution in the region r at time t is defined as the ratio of the sum of its all non-perming loans at region r and time t to the sum of all loans at region r and time t. The empirical model is,

$$endo\_var_{rt} = \alpha + \beta_0 * endo\_var_{rt-1} + \beta_1 * exo\_var_{rt-1} + \beta_2 * exo\_var^2_{rt-1} + \beta_3 * cont_{rt-1} + e_{rt}.$$

where $endo\_var_{rt} = ln\,(NPL_{rt}/(100 - NPL_{rt}))$ and NPL_{rt}: Non-performing loans ratio. We first run a regression only for banks. As before we use annual information and the period of study is 2004-2017. And as in previous regressions the explanatory variable and controls are lagged one period.

Table 10 reports a statistically significant inverted U-shaped relationship between number of competing banks in a region with the NPL ratio with time-fixe effects . The number of competing banks that maximizes the NPL in a region is 10.3. We then run a regression only for banks and considering only commercial loans. As before Table 11 reports a statistically significant inverted U-shaped relationship between number of competing banks in a region with the NPL ratio with time-fixe effects . This time, the number of competing banks that maximizes the NPL in a region is 6.4.

Finally, we do the same considering all financial institutions considering all loans (table 12) and only commercial loans (table 13). In both cases we find an inverted U-shaped relationship between number of competing banks in a region with the NPL ratio with and without time-fixe effects. And the number of competing financial institutions that maximizes the NPL in a region are 32 and 22.6, respectively, without time-fixed effects and 28.7 and 17.8 with time-fixed effects.

Table 10: RCC sample: Banks

exo_var	ln (# banks)			C4			Herfindahl index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
L.endo_var	0.609***	0.519***	0.605***	0.553***	0.448***	0.592***	0.520***	0.406***	0.531***
L.exo_var	3.504**	0.472	5.673***	76.13**	114.5**	21.68	8.736***	2.782	8.206***
L.exo_var2	-0.547	0.281	-1.216***	-42.77**	-66.54**	-12.22	-12.54***	-8.028	-11.01***
L.size_rt	-0.00639***	0.206**	-0.00364***	-0.00671***	0.187	-0.00469**	-0.00513***	0.162	-0.00422**
Observations	336	336	336	336	336	336	336	336	336
R-squared	0.530	0.621	0.602	0.484	0.579	0.583	0.558	0.645	0.631
F test (ρ-value)	0	0	8.46e-11	0	0	5.05e-09	0	1.52e-09	8.83e-09
Region 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%
We control for the credit market share (size_rt).

Table 11: RCC sample: Banks - commercial loans

exo_var	ln (# banks)			C4			Herfindahl index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
lag_endo_var	0.438***	0.333***	0.438***	0.465***	0.341***	0.469***	0.456***	0.341***	0.461***
lag_exo_var	2.696**	1.747	2.754*	51.06	51.19	20.27	5.120**	1.772	4.535*
lag_exo_var2	-0.619*	-0.207	-0.736*	-28.40	-29.81	-11.28	-7.087***	-4.459*	-6.236**
lag_size_crt_t	-0.596**	16.15**	-0.160	-0.804***	12.06*	-0.551***	-0.584***	14.01**	-0.416*
lag_size_ct_t	31.46**	39.51***		19.02	27.74**		27.45**	34.29***	
lag_size_crt_rt	-0.662	0.194	-1.297	0.703	1.680**	-0.224	-0.354	0.483	-1.180
Observations	334	334	334	334	334	334	334	334	334
R-squared	0.467	0.574	0.527	0.435	0.553	0.502	0.497	0.602	0.548
F test (ρ-value)	0	0	2.03e-08	0	1.29e-09	1.74e-07	0	0	3.60e-08
Region 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%
We control for the commercial credit market share (lag_size_crt_t), the share of commercial loans in the region (lag_size_crt_rt), and the proportion of commercial loans across time in the whole economy (lag_size_ct_t).

Table 12: RCC sample: All financial institutions - competition across groups

exo_var	ln (# institutions)			C4			Herfindahl index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
L.endo_var	0.620***	0.522***	0.639***	0.531***	0.422***	0.573***	0.558***	0.436***	0.578***
L.exo_var	4.114**	-0.207	3.417**	18.76**	25.38*	8.914	2.828***	1.218	2.401**
L.exo_var2	-0.574**	0.162	-0.509**	-12.68**	-17.33*	-6.514	-7.643***	-6.237***	-6.533***
L.size_rt	-0.0337	0.382	0.0156	-0.0978***	0.661	-0.0863***	-0.108***	0.697*	-0.101***
Observations	336	336	336	336	336	336	336	336	336
R-squared	0.555	0.631	0.620	0.558	0.621	0.628	0.589	0.655	0.649
F test (ρ-value)	0	5.51e-11	1.53e-06	0	2.97e-10	2.93e-07	0	0	1.50e-09
Region 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%

Table 13: RCC sample: All financial institutions - competition across groups - commercial loans

exo_var	ln (# institutions)			C4			Herfindahl index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
lag_endo_var	0.524***	0.384***	0.538***	0.486***	0.341***	0.503***	0.531***	0.367***	0.545***
lag_exo_var	2.307*	1.550	2.395***	25.45**	22.38*	19.85**	0.660	-1.182	0.154
lag_exo_var2	-0.370*	-0.223	-0.416***	-16.81***	-15.41**	-13.37**	-3.436	-1.665	-2.516
lag_loans_crt_t	33.71*	39.19***	23.30*	49.76***	37.07**	46.66**	37.24**	34.23***	38.43***
lag_loans_ct_t	14.70	21.86***		11.15	20.87***		9.703	19.31***	
lag_loans_crt_rt	-0.119	0.612	-0.508	0.229	0.657	-0.274	0.409	0.911	-0.121
lag_loans_rt_t	-32.95*	-33.59***	-22.45*	-48.96***	-31.70**	-45.83**	-36.81**	-30.13***	-37.85***
Observations	336	336	336	336	336	336	336	336	336
R-squared	0.495	0.586	0.544	0.515	0.600	0.559	0.513	0.608	0.556
F test (ρ-value)	0	0	1.64e-07	0	0	2.19e-08	0	0	1.04e-07
Region 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%
The control lag_loans_rt_t is identical to L.size_rt.

Appendix B Competition across groups

Table 14 is as table 6 but assuming that there is competition across institutions from different groups (banks, CMACs, CRACs, EDPYMES and *empresas financieras*).

Results regarding the number of banks suggests an inverted U- shaped relationship but it is not significant. Regarding the other competition measures they suggest a U-shaped relationship, however they are not significant neither.

Table 14: All financial institutions: Competition across groups

exo_var	ln (# institutions)			C4			Herfindahl index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
L.endo_var	0.766***	0.565***	0.772***	0.771***	0.572***	0.773***	0.774***	0.574***	0.776***
L.exo_var	0.0167	-0.814	-0.113	4.754	1.533	4.481	6.456	7.332	6.127
L.exo_var2	-0.0105	0.131	0.0144	-3.211	-0.438	-3.024	-17.16	-16.04	-16.39
L.roa	-0.962	-0.383	-1.054	-0.865	-0.466	-0.999	-0.871	-0.456	-1.004
L.size_new	-0.831**	5.013	-0.800**	-0.836**	4.897	-0.818**	-0.822**	4.922	-0.807**
L.adeu_cred	0.103	0.170	0.0875	0.106	0.170	0.0895	0.105	0.174	0.0898
L.bond_cred	-1.297***	-1.481		-1.137***	-1.700***		-1.137***	-1.728***	
L.lev	0.0233**	0.0326**	0.0253**	0.0243**	0.0344**	0.0256**	0.0250**	0.0341**	0.0262**
Observations	783	781	783	783	781	783	783	781	783
R-squared	0.785	0.846	0.791	0.786	0.847	0.792	0.786	0.846	0.792
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

*** p<0.01, ** p<0.05, * p<0.1

Appendix C More Figures and Tables

Figure 6: Matching of loans distribution across regions

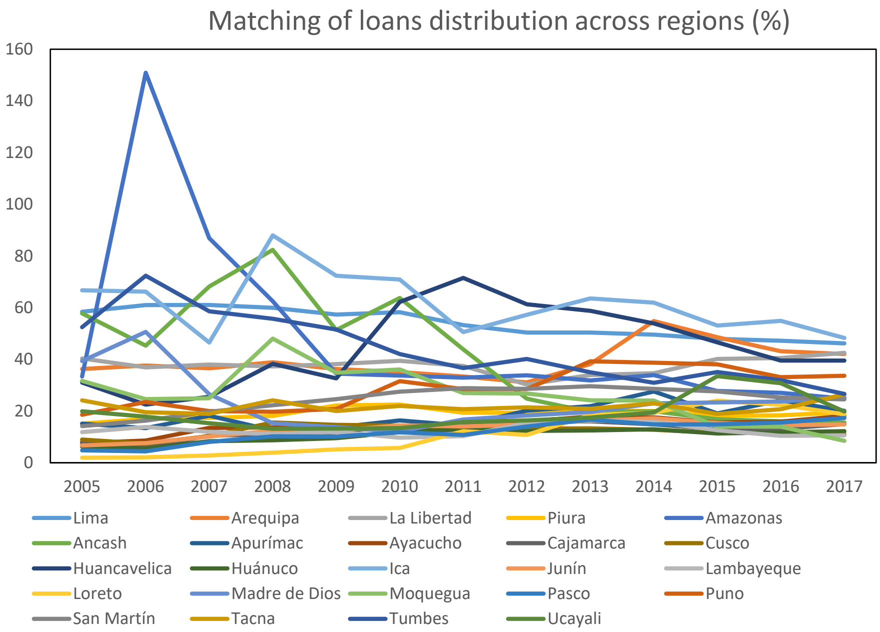
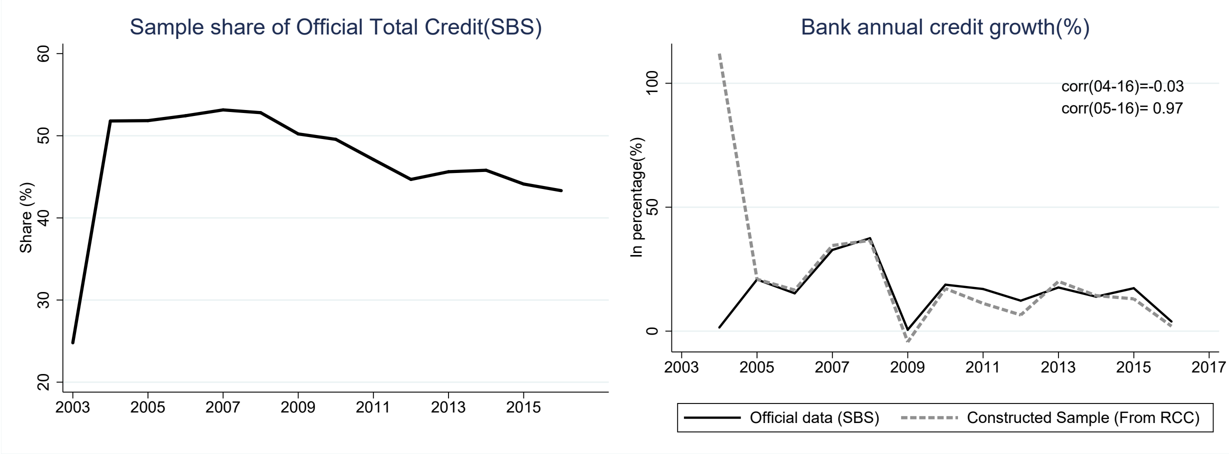


Table 15: Correlation coefficients - No competition across groups

Variables									
NPL_{it}	1								
Number of $banks_{it}$	-0.2391*	1							
$C4_{it}$	-0.1190*	0.4745*	1						
HHI_{it}	-0.0940*	0.3475*	0.9195*	1					
$Size_{it}$	-0.1322*	-0.1476*	-0.0327	0.0135	1				
ROA_{it}	-0.1454*	-0.0246	0.0865*	0.1183*	0.1836*	1			
$For\_debt\_cred_{it}$	-0.1215*	0.0635	0.0878*	0.0208	0.0828*	-0.0209	1		
RWA_Cap_{it}	-0.2275*	0.2198*	0.0125	0.0517	0.2082*	0.0695*	0.0014	1	
$Bond\_cred_{t}$	0.0512	-0.3436*	0.1973*	0.3006*	0.0116	0.2410*	-0.0900*	-0.1243*	1

* Significant at the 5% level.

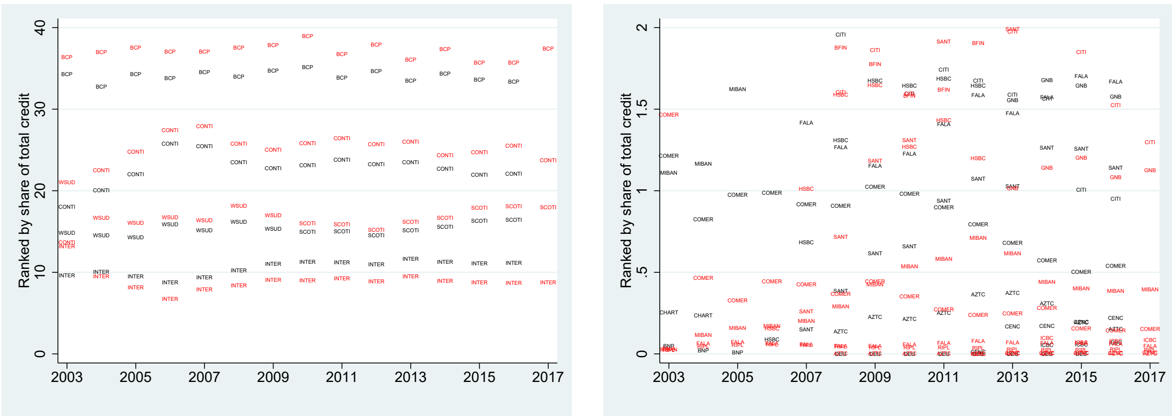
Figure 7: Representativeness of credit: Our sample vs official data



Note: Annual Sample 2003-2016. SBS Official Data and Our sample

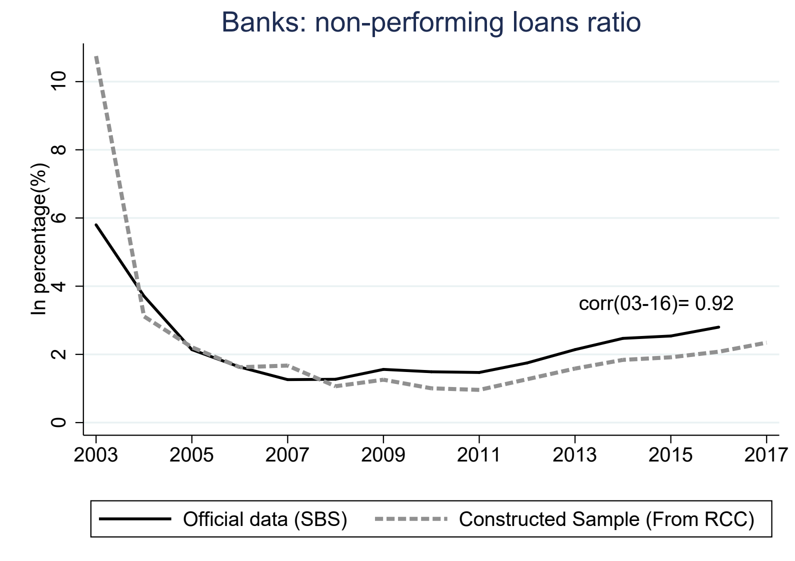
Figure 8: Ranking of credit shares: Our sample vs official data

Large banks with credit shares $> 6\%$ Small banks with credit shares $\leq 6\%$



Note: RCC sample data in red, SBS data in black. The graphs show ranking of banks in SBS Data vs ranking in the RCC sample. Banks ranked by credit shares.

Figure 9: Bank Non-Performing loans (NPL) ratio: Our sample vs official data

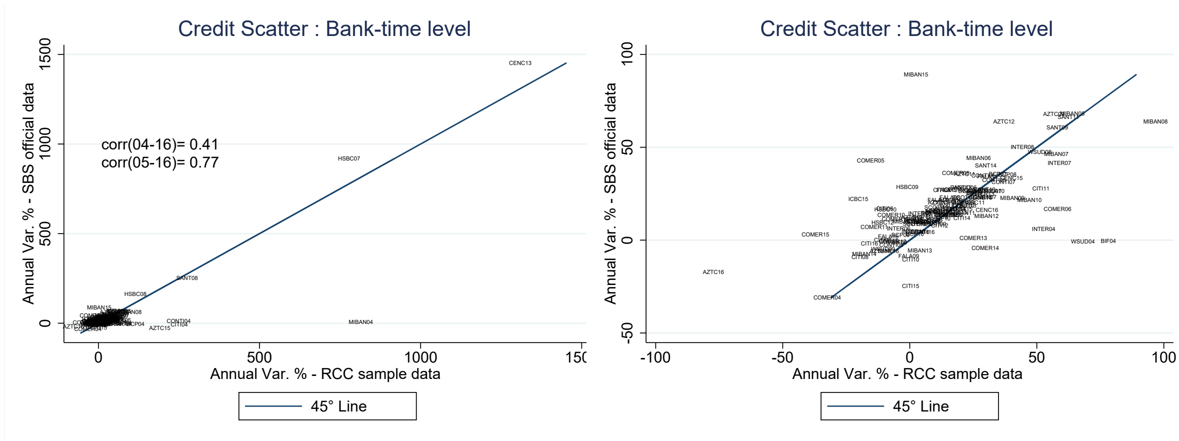


Note: Annual Sample 2004-2016. SBS Official Data.

Figure 10: Bank credit growth: Our sample vs official data

(a) Full data

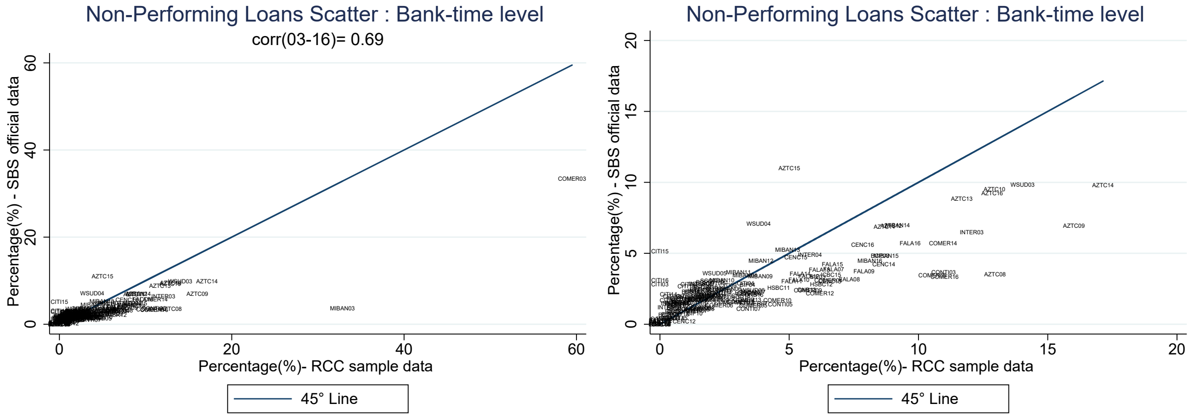
(b) Restricted var% rates < 100



Note: Annual Sample 2003-2016. Correlation computed as a cross-section average of bank correlation coefficients, i.e., $\frac{1}{\#banks}\sum_{b=1}^{\#banks} corr(\Delta Credit^{b}_{SBS,t}\,, \Delta Credit^{b}_{RCC\ sample,t})$. Each correlation computed over the period 2004-2016 or 2005-2016.

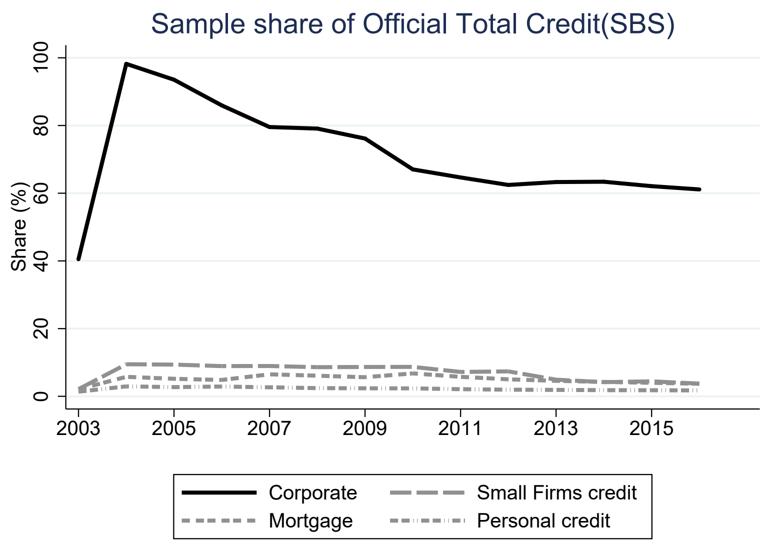
Figure 11: NPL ratio: Our sample vs official data

(a) Full data (b) Restricted % rates < 20



Note: Annual Sample 2003-2016. Correlation computed as a cross-section average of bank correlation coefficients, i.e., $\frac{1}{\#banks}\sum_{b=1}^{\#banks} corr(\Delta Credit^b_{SBS,t}\ ,\Delta Credit^b_{RCC\ sample,t})$. Each correlation computed over the period 2003-2016.

Figure 12: Share of official data by type of credit (%) - financial system



Corporate refers to comercial loans.

Figure 13: Representativeness credit growth - financial system