INSPER INSTITUTE OF EDUCATION AND RESEARCH
DOCTORAL PROGRAM IN BUSINESS ECONOMICS

CLAUDIA BRUSCHI MARTINS

ESSAYS ON BANKING

São Paulo
2021

CLAUDIA BRUSCHI MARTINS

ESSAYS ON BANKING

Thesis presented to Insper - Institute of Education and Research as partial fulfillment of the requirement for obtaining the title of Doctor in Business Economics.

Advisor: Marco Antonio Cesar Bonomo

São Paulo

2021

Resumo

Nos dois primeiros capítulos, esta tese avalia os impactos das mudanças na regulação bancária sobre o mercado de crédito ocorridas no Brasil. No último capítulo, expõe-se como o uso de dados bancários alternativos pode melhorar a compreensão da economia, especificamente na análise do mercado de trabalho.

O Capítulo 1 (com Marco Bonomo e Ricardo Schechtman) explora uma mudança na regulamentação que melhorou as informações das instituições financeiras sobre tomadores com bom histórico de pagamento de crédito como um experimento natural para investigar o efeito causal da redução das assimetrias informacionais no mercado de crédito. A mudança provocou uma redução no custo do crédito de 33,5 p.p., mas sem efeito no montante de novos empréstimos e no prazo médio contratado. Mostrou-se também que esses resultados foram impulsionados por bancos menores. Os resultados indicam que a assimetria de informação é um importante determinante de juros de empréstimo.

O Capítulo 2 analisa se seguro de depósito aumenta a concorrência no setor bancário e se afeta a estabilidade desse setor. O estudo explora a criação de um fundo de seguro de depósitos para cooperativas de crédito no Brasil que afetou apenas parte do sistema financeiro. Constata-se que depois que as cooperativas de crédito (CU) começaram a oferecer seguro de depósito, os bancos perderam 3% de seus depósitos, e que esse efeito foi mais forte em mercados onde as CU tinham anteriormente uma participação de mercado maior, indicando que o seguro de depósito pode afetar a concorrência no setor bancário. Observa-se também que o aumento da concorrência não afetou a estabilidade do setor ao estimular os bancos a assumirem mais riscos. Esses resultados mostram que o seguro de depósito pode ser importante para melhorar o ambiente competitivo para instituições financeiras menores.

Por fim, o capítulo 3 mostra que o uso de dados bancários das informações de processamento de folhas de pagamento permite a mensuração do emprego em tempo hábil, e tem potencial para melhorar a compreensão dos movimentos do mercado de trabalho no Brasil durante a pandemia de Covid-19, período em que as séries oficiais sobre emprego sofreram inconsistências metodológicas devido a mudanças no tamanho da amostra, métodos de coleta de dados e representatividade.

Palavras-chave: Assimetria de informação; Crédito; Seguro Depósito; Cooperativa de crédito; Mercado de trabalho; Big Data

Abstract

In its first two chapters, this thesis assesses the impacts of changes in banking regulation that have taken place in Brazil on the credit market. In the last chapter, it shows how the use of alternative bank data can improve the understanding of the economy, specifically in the labor market.

Chapter 1 (with Marco Bonomo and Ricardo Schechtman) explores a change in regulation that improved financial institutions' information about borrowers with good credit repayment history as a natural experiment for investigating the causal effect of reducing informational asymmetries on credit outcomes. The change caused a decrease in the cost of credit of 33.5 p.p., but the size of new loans and the average maturity were not affected. It is also documented that the results were driven by smaller banks. The results indicate that information asymmetry is an important determinant of lending rates.

Chapter 2 analyzes if deposit insurance increases competition in the banking industry and if it affects sector stability. The study explores the creation of a deposit insurance fund for credit unions in Brazil that affected only part of the financial system. It is found that after credit unions started offering deposit insurance, banks lost 3% of their deposits and that this effect was stronger in markets where CU had previously a larger market-share, indicating that deposit insurance can affect competition in the banking industry. It also observes that the increase in competition didn't affect sector stability by encouraging banks to take more risk. These findings show that deposit insurance can be important in leveling the playing field for smaller financial institutions.

Lastly, chapter 3 shows that the use of bank data from payroll processing information allows the measurement of employment at a timely frequency and has the potential to improve the understanding of the labor market in Brazil during the Covid-19 pandemic, a period in which government official series for employment underwent methodological inconsistencies due to changes in sample sizes, data collection methods, and representativeness.

Keywords: Information asymmetry; Loans; Deposit Insurance; Credit Union; Labor Market; Big Data

Summary

List of Figures

List of Tables

1. Effects of Sharing Positive Credit Information on Consumer Loans

1.1. Introduction

Loans interest rate spread[1] are relatively high across emerging markets (eg. 32.2% in Brazil, 12.3% in Argentina, 7.4% in Colombia[2]) and access to credit limited (Reed et al. (2015)). These affect the ability of households to smooth consumption overtime[3] (Antunes et al. (2013)), to invest in physical (e.g. capital to start a small business), and human capital (e.g. training) with implications for welfare and economic growth (De Gregorio (1996)). These facts could be explained by the lack of information sharing about a potential borrower's capability to repay its debt (Pagano and Jappelli (1993)). In this context, information sharing of credit records by credit registries appear as a potential remedy for information asymmetry (Demirguc-Kunt and Klapper (2012)). Yet, credit registries have a broad heterogeneity in borrower's information across countries and there is still a limited understanding of its effects on the credit market.

Credit registries store both positive (ie. outstanding loan amounts and a pattern of on-time repayments) and negative (ie. late-payments and the amount of defaults) borrowers' information. Despite empirical evidence indicating that sharing positive information improves the ability of banks to predict the probability of clients repaying their loans (Powell et al. (2003); Turner (2010)), 40% of countries' credit registries do not distribute positive credit information[4,5].

In this paper, we study the effect of sharing positive credit information on households' cost of credit, loan size, maturity, and risk assessment by banks. In July 2016, the Central Bank of Brazil (BCB) started collecting credit information on loans of borrowers that had maximum total liabilities per financial institution between R$ 200 and R$ 1,000, encompassing information from 41 million new users[6]. Until then, only loans from clients that had total liability per financial institution above R$ 1,000 were reported. With detailed information from the Brazilian public credit registry (SCR) and using this change in regulation as a source of exogenous variation, we can evaluate the causal effect of sharing positive credit

[1] Lending rate minus deposit rate

[2] Source: World Development Indicators (World Bank). Data from 2018.

[3] In 2018, American households spent 8% of their income in interest rates and amortization (Source: Bank for International Settlements), while Brazilian families spent 17% (Source: Central Bank of Brazil).

[4] Source: 2018 Doing Business (World Bank)

[5] 88% are developing countries and sharing positive credit information is part of the World Bank's directives in Doing Business.

[6] Source: 2016 Annual Report - Central Bank of Brazil

information on credit variables, once negative information was already shared by private bureaus. Additionally, we explore the different effects that information can have for small and mid-sized banks, and large banks.

Regardless of the information type, theory predicts that exchanging credit information eases adverse selection and helps financial institutions better screen good borrowers. Consequently, those institutions can offer larger loans with lower interest rates for on time payers (Pagano and Jappelli (1993)). However, regardless of theoretical prognosis, the few empirical papers that analyze reduction of information asymmetry on the credit market use limited databases or study aggregate effects, not finding significant consequences of sharing positive credit information on interest rates (eg. Foley et al. (2019); Behr and Sonnekalb (2012)). Thus, the implications of improving access to positive credit information of households on their loans outcomes needs to be better understood.

We use a difference-in-differences (DiD) framework to estimate the effects of sharing positive credit information. We define the treated group as individuals that had maximum total liabilities in any financial institution between R$ 500 and R$ 1,000, as of June 2016, and whose information was not disclosed in SCR previously the threshold change. We use as control the group of borrowers that had maximum total liabilities per bank between R$ 1,000 and R$ 1,500 [7] and, consequently, had information mandatorily reported to the SCR before June 2016. The identifying assumption is that absent the threshold change, the outcomes in treated and non-treated individuals would have followed parallel trajectories.

Our set of empirical results is divided in three parts. First, we focus on the effect of the threshold change on financial variables. After the change, we find a decrease in spreads of approximately 33.5 p.p. (of a sample average of 304%) and no significant change in new credit origination, maturity, and risk rating given to the new borrower by the bank.

Second, we investigate if sharing positive information can have heterogeneous effects depending on the size of the bank. Small and middle sized banks can have an advantage as they could develop better ways to screen specific types of borrowers (Nakamura (1991)). At the same time, getting information about borrowers could be more costly as they are smaller and not part of bank conglomerates. The results indicate that the impact of the change in the information environment matters just for smaller banks. With the new regulation, the loan spread rate decreased 42 p.p., and the contract size increased 24%. Lastly, we show that our results are robust to different specifications.

Our paper is connected to several strands of the academic literature. First, our work relates to the theoretical literature that assesses the effects of credit information sharing. Theory predicts that sharing information (not only positive) can decrease the cost of loans

[7]The equivalent in US dollars of 2016 is 143-287 USD and 287-430 USD, respectively.

and increase credit access through different channels: i) by mitigating adverse selection it increases lender's ability to screen good payers and, consequently, leads to better pricing (Pagano and Jappelli (1993))[8]; ii) by reducing moral hazard it increases the default costs of higher-risk borrowers (Padilla and Pagano (2000))[9]; iii) breaking the banks' information monopoly, which can limit banks from charging higher interest rates, decreasing the possibility of extracting rents from high-quality borrowers (Pagano and Jappelli (1993)). Our paper contributes to this literature by providing evidence consistent with those arguments, highlighting a potentially positive effect of the type of information shared on credit market.

Second, our work relates to the empirical literature that investigate the relation between information asymmetry and credit market outcomes. The empirical evidence on the effects of information sharing on the credit market is based on both macroeconomic and microeconomic data. Only a few papers use micro data to assess the impact on interest rates. Behr and Sonnekalb (2012) uses loan-level data from only one bank in Albania to analyze the effect of adopting a public credit registry (with both positive and negative information). Their results suggest that information sharing does not affect credit access or cost but improves loan performance. Foley et al. (2019) identify the causal effect of information sharing on ex-post competition and conclude that public credit information increases competition in credit markets for good borrowers but can hinder financial inclusion for clients with low creditworthiness. With a different and limited database (Chile has interest caps), they also conclude the main margin of adjustment is credit limits rather than interest rates, as they don't find any difference in the rate at the time of origination. Our work distinguishes itself from the empirical work done so far by focusing on borrowers with positive information and using a comprehensive database that comprises consumer credit loans of all banks in Brazil.

Third, our work relates to the literature that explores the distinct effects of positive or negative information sharing. Powell et al. (2003) tested the predictive power of debt default and showed that the added value of having positive and negative information about borrowers is significantly higher. In the same way, Turner (2010) finds that systems that cover positive credit information are more accurate than systems with only negative information, and this helps to identify borrowers with low and high credit risk, generating positive net effects such as broader access to credit.

We proceed as follows section 1.2 describes the institutional background. Section 1.3 describes the data, its characteristics and also discusses our empirical framework. Section 1.4 presents the empirical results. Section 1.5 concludes.

[8]de Janvry et al. (2010) note that interest rates can remain fixed over the short run.

[9]Yet, sharing positive information may not produce a disciplinary effect, since a high-quality borrower is less concerned about the consequences of default as long as they can remain a high-quality borrower (Hahm and Lee (2011)

1.2. *Institutional Background*

The Credit Information System (SCR) is managed by the Central Bank of Brazil. Credit suppliers (banks, credit unions, and non-banks) mandatorily report monthly detailed information on credit relationships with their clients[10], including overall debt exposure, maturities, and defaults. Also, banks can only have access to aggregate information of a client on a credit type basis and comprising only the past 12 months[11]. Thus, lenders don't have information about who lent to each client.

At the beginning of SCR, in 1997, financial institutions were required to submit individualized loan information of each borrower with total liability equal to or greater than R$ 5,000. This amount was later reduced to R$ 1,000 in 2012 and, as of June 2016, this limit was further reduced to R$ 200[12], which increased the number of individuals at SCR from 64 million to 105 million, encompassing 99% of all credit transactions in the financial system. This characteristic makes the SCR the most comprehensive credit registry operating in Brazil regarding financial institutions' credit transactions.

In Brazil, there are also private credit bureaus. Despite covering almost the same proportion of the adult population in Brazil as public credit registries[13], credit bureaus provide mainly socio-economic characteristics and negative information of borrowers, due to the lack of legal framework[14]). Thus, the basic difference between SCR and private bureaus in Brazil is that the former adds information about the total consumer's exposure with financial institutions and its characteristics (whether it is performing, type of loan, etc.).

1.3. *Empirical Strategy and Data*

1.3.1. *Empirical Strategy*

We want to identify the average effect that the decrease in information asymmetry of households with good payment history has on consumer credit cost, loan size, and credit maturity (i.e., the average impact of treatment on the treated group). In an ideal experimental setting, we would have randomly selected households to have their information

[10] All financial institutions must send the detailed information about their active loans to the BCB up to the 9th business day of each month. Usually, BCB process the information and make it available to be consulted by banks within a week.

[11] After May 2018, banks can access the information history of borrowers from the past 24 months.

[12] This change, despite being public, wasn't campaigned by the media and banks to consumers.

[13] Data from 2018 Doing Business shows that private credit bureaus covered 79% of adult population, while credit registry 76%.

[14] As of October 2019, credit bureaus can share positive information of borrowers. Until then, only 5% of borrowers had positive credit information shared by credit bureaus (Information based on a Administrative Council for Economic Defense vote).

revealed and would compare credit outcomes with clients that didn't have their information disclosed. Unfortunately, the information disclosure process was not random. Nevertheless, we treat the information threshold change of the credit registry by the BCB in July 2016 as a quasi-natural experiment and apply a difference-in-differences methodology to analyze the improvement in information sharing on credit market variables.

As the change in regulation affects all credit lines and we were operationally unable to work with all the database, we investigate all the originations of June 2016 (month of the threshold change) for clients with maximum total liability per bank ranging between R$ 500 and R$ 2,000. Only 0.3% were earmarked loans. Of the non-earmarked, the two main credit lines were revolving: 40% credit card purchases, in which clients don't incur in interest rate unless the payment is past due, and 20% overdraft. The third most important credit line was consumer credit without payroll-deduction, accounting for 12% of originations. All loans originations are reported in Table 1

As revolving credit is pre-approved and for very short term (less than a month), consumer credit without payroll-deduction is ideal to evaluate how information asymmetry impacts the credit market of longer-term loans equilibrium and without guarantee. Also, in a country with 38%[15] of employees working in the informal market, credit registries can greatly alleviate the lack of information about income and payment capability.

We divide all observations of individuals with personal credit into two groups i) treatment: individuals that had total liabilities between R$ 500 and R$ 1,000 in any financial institution as of June 2016 and had no information disclosure in SCR since January 2015.[16]; ii) control: individuals that had total liabilities between R$ 1,000 and R$ 1,500 per financial institution and, consequently, had information mandatorily shared through SCR before June 2016. We also restrict to individuals that didn't have new information disclosed with the threshold change [17]. The individuals from the first group are the households for which we expect a reduction in information asymmetry after the threshold change and, consequently, have the cost of credit relatively lowered and supply of credit relatively increased.

Also, we segregate the database into two periods: originations made between January 2016 and June 2016 (pre-change) and between August 2016 and January 2017 (post-change). We exclude July 2016, the transition month.

As described in the previous section, prior to the change, banks could only know if an

[15]World Bank. Data from 2015.

[16]The use of the threshold cut-off is not enough to define the treatment group as one individual could have had in the past a loan above R$ 1,000 and had information available for other banks through SCR.

[17]There is a group of borrowers who have more than one banking relationship and who, despite being in the SCR for having loans with a specific bank above R$ 1,000, may have exposure with another bank below the threshold.

individual from the treatment group had a bad payment history[18]. Thus, the threshold change would only affect clients that were on-time with their loans. This setting in Brazil is ideal for assessing how positive information (good payment history) impacts credit market.

The impact of the shift in the threshold might affect differently banks' perception of the creditworthiness of the client if she is a current client or a prospective clients. A strand of literature supports that because of the opaqueness of small borrowers, lenders value the preexisting client relationship to reduce information asymmetries (eg. Boot (2000)). Consequently, for current clients, the analyzed event would not change the hard information that the bank had. For potential clients, however, the change alleviates the lack of information. Thus, another important consideration for our analysis is that we restrict our sample to first-time borrower with a financial institution, as the information revealed with the threshold change might not be important for old clients, as the bank could already assess the client's capability of paying.

The first identifying assumption in our approach is that the particular outcome variable develops in a parallel way for the two groups if information sharing is unchanged. We assume that while there may be a difference in the outcome between groups, this difference is time-constant and would have prevailed had the improvement in information sharing not occurred. Although we cannot observe what would have happened in the absence of the credit registry reporting change, observing a parallel trend before would make us confident about this assumption. A graphical analysis to support this conjecture is provided below. If this assumption holds, we can attribute any divergence of this parallel trend after the change to the effect of improving information sharing between lenders. Figure 1 shows that this assumption seems justified.

Another concern is that financial institutions could be supplying credit just below R$ 1,000 before June 2016 not to have their creditworthiness customers exposed to other banks. In Figure 2, we present the frequency of loans' size as a piece of evidence that banks were not manipulating the size of contracts.

1.3.2. Data and Descriptive Statistics

The change in the threshold required that financial institutions in July 2016 transmitted to BCB all active credit relationships of their clients that had total exposure to the bank above R$ 200 in June 2016. Before July 2016, this threshold was R$1,000. Thus, for our treatment group - new borrowers with loans between R$500 and R$1,000 that were not in

[18]In some States (eg. Law 15.659/2015 of São Paulo State), banks can send information of non-performing clients to credit bureaus after 15 days of payment delay and it is unexpensive (the cost is less than 1% of the minimum loan amount considered).

the SCR after January 2015[19] - we only observe a one month picture of their loans before the threshold change. However, with the information of the date of origination, we can generate our monthly database of new originations since January 2016.

We need to point that our database has a bias for longer term credit for our treatment group, as we only observe active loans, not having information of those that expired before June 2016. To deal with this, we use the same selection (active loans in June 2016) for the control group and show in robustness tests that our result still holds considering different maturities.

Our sample ranges from January 2016 until January 2017[20]. We winsorize the financial variables at the 2/98% level. Table 2 has the descriptive statistics for treatment and control groups, divided into pre and post threshold change. The final dataset has 1,475 observations. When we restrict our analysis to borrowers with only first time loan with a bank, we loose a considerable amount of observations, as borrowers tend to stay in the same bank.

The average spread (lending rate minus national level deposit rate) is around 326% across groups. As initially highlighted, Brazil has one of the highest credit spreads in the world, especially for lower income clients that are considered as more risky. Also, from the simple descriptive statistics, it is not straightforward that borrowers who were not in the SCR before the regulation had higher interest rates compared to those who already had information being shared. However, this fact could be due to the composition of the banks' profile (shareholding control and size) in which they borrowed in comparison to the control group: higher percentage in state-owned and larger banks. The average size of the contract and total financial debt have a considerable difference after the change, but this is because more financial institutions had to report about the same individual after the threshold change.

In terms of demographics (Panel B), the heterogeneity across group is not prominent. 64% of the originations were to female borrowers. The average age of the borrowers was 45 years old. The average monthly income was R$ 1,241, equivalent to 1.2 times the minimum wage of 2016. The usual income of Brazilian borrowers is above 5 minimum wages, indicating that the threshold change impacted mostly small borrowers.

Regards the financial characteristics (Panel C), around 90% of the originations came from private banks. The total indebtedness of individuals was, on average, 18% of their annual income. However, the monthly commitment of the monthly income paying credit amortization and interest rates were close to 47%. An important characteristic of our sample is that the majority of the borrowers in the control group got credit with small and mid-sized

[19]We use the 12-month search window imposed by the BCB to delimit our treatment

[20]Before arriving at the final data sample, we apply several data filters, excluding i) clients that are not Brazilian; ii) loans with variable interest rate rates; iii) clients' with age below 18 or above 65; iv) clients with contract risk above D rate (renegotiation).

banks (SMB) and the share for our treatment group increased after the change in the SCR, indicating a market segmentation.

1.4. Empirical Framework and Results

This section presents evidence on the effects of improving credit information using data described in the previous section.

1.4.1. Empirical Framework

In our baseline estimation, we investigate the effect of reducing information asymmetry for good borrowers on credit market outcomes using an exogenous change in regulation employing a DiD research design. We compare credit contract outcomes of treated individuals (exposed to SCR with the regulation change) with those in the control group (already exposed at the time of the registry expansion), before and after the threshold variation.

We estimate the following regression using OLS:

$$y_{ijt} = \alpha + \beta_1 Treatment_i Post + \theta_i + \theta_{jt} + X_{it} + u_{ijt} \tag{1}$$

Where y is our variable of interest - cost of credit (spread), log of the size of contract, and maturity -, for client i in bank j at time t. $Treatment_i$ is a binary variable equal to one for the individuals that had information revealed by the change in the threshold of SCR (change in the cut-off from R$ 1,000 to R$ 200) and zero for individuals that already had the information in the SCR before June 2016 and no new data were disclosed[21]. $Post$ is a binary variable equal to one for observations after the change of the credit registry in June 2016 and zero otherwise. The coefficient of interest is β_1, from the interaction of $Treatment_i$ and $Post$, that measures the average treatment effect of the improvement in information about the borrower through the credit registry. Additionally, we control for the general time-varying individual characteristics (X_{it}): income, total indebtedness in the financial system[22], debt-to-service ratio[23], number of bank relationships, number of banks in the municipality of the borrower. Also, we include month-bank fixed effects (θ_{jt}), which account for bank policy unobservables. Finally, we control for individual fixed effects (θ_i), which captures time-invariant unobserved heterogeneity across borrowers. As we control for the described fixed effects, the single variables $Treatment_i$ and $Post$ are not estimated. The

[21] A small group of borrowers who already had information in SCR but had loan with a different bank below R$1,000 was affected by the new regulation. In our baseline estimation, we exclude this group, but use for robustness check in the next section.

[22] Total debt in the financial system divided by the annual income.

[23] Ratio of debt service payments (principal + interest) to monthly income.

standard errors are clustered at the bank level to account for a possible dependence of setting contract levels[24]

We note that controlling for contract characteristics after the registry expansion in eq. 1 would bias the coefficients. Thus, our empirical setting is the reduced form effect of reduction in the information asymmetry.

1.4.2. Baseline Results

Table 3 presents the results estimates of equation 1 on borrower-bank level data using different dependent variables: loan's spread, natural log of contract size, and maturity. In all estimations, we control for client and bank-month fixed effects, and some client's time-varying characteristics. In Column (1), our outcome variable is the lending spread and we observe that spread has decreased by 33.5 p.p. (of a sample average of 304%) and is statistically significant at 5% after the positive information of a borrower is made available.

Column (2) shows the result for the size of the contract as dependent variable. Despite being positive, the coefficient is not statistically significant. We don't exclude the possibility that positive information decreases credit restriction, since we are analyzing only one credit line (personal loan) and the increase may occur in other lines of credit[25].

In Column (3) of Table 3, we evaluate change in the regulation on loan maturity. The effect is unexpected negative, but not statistically significant.

Lastly, we ran Equation 1 but with data on client level:

$$y_{it} = \alpha + \beta_1 Treatment_i Post + \theta_i + \theta_t + X_{it} + u_{it} \tag{2}$$

The outcome variable is number of different banks that the borrower has a loan. Despite not finding any result about the size of loan (intensive margin), it might be that the effect was on the number of loans with different banks (extensive margin). The result of the coefficient of interest is shown in column (4) of Table 3. No significant effect is found (economically and statically).

In sum, sharing positive information of borrowers has meaningful changes in the cost of credit for borrowers, but not in other loan characteristics. We recognize the limitations of our database in capturing the whole effect. However, the initial results suggest the importance of information sharing in the well functioning of the credit market.

[24]Our results are still robust if estimated without this clustering option.
[25]Data from other credit lines are not available for our analysis.

1.4.3. *Heterogeneity*

Different characteristics of management practices of financial institutions could drive heterogeneous outcomes even in the absence of differences in their credit information disclosure requirements. Although these differences can be understood as being at least partly endogenous to the credit information setting, in our analysis, any change in the lending outcomes can be causally assigned to the change in the informational environment.

The size of the bank can affect differently credit outcomes. Small and middle sized banks can have an advantage as they could develop better ways to screen specific types of borrowers (Nakamura (1991)). At the same time, getting information about borrowers could be more costly as they are smaller and not part of bank conglomerates. In the first case, the change of the information set would not largely affect their lending practice. In the last case, this could give them a greater advantage in better screening good borrowers. To test the heterogeneous response of different size of financial institutions, we split our sample into two based on the size of bank: i) small and middle size (the majority of our sample), and ii) large banks[26]. The results of Eq. 1 with the subsamples are in Table 4. Columns (1) and (2) in Panel A show that the impact of the change in the information environment matters just for smaller banks. With the new regulation, the loan spread rate decreased 42 p.p., and the contract size increased 24%. The sample size of our regressions for big banks is very small, so we interpret it with a pinch of salt. These findings indicate that the average result found in the previous section were driven by smaller banks that could have problems better screening borrowers.

1.4.4. *Robustness Tests*

In this section, we present a series of robustness tests and extension exercises reported in Table 5 for loan spreads. First, to stress our point that it is the change in the informational environment that affected banks' evaluation of borrowers' creditworthiness, we drop the sample restriction for first-time borrowers in a bank and run Eq. 1 with borrowers that got loans in financial institutions that he already had a previous loan - for both treatment and control group. The result is shown in Column (1) which we see no decrease in spread, indicating that information setting didn't change for relationship bank, underscoring our result. We highlight that this finding doesn't exclude the competitive response of banks to a contestable market as in Foley et al. (2019). Our analysis ends six months after the change in the regulation and the response can take more time.

[26] Classification based on the number of clients. The first group are those with less than 4 million clients and the second, above this number.

Another concern is that not only the positive information disclosed by SCR is the driver of the decrease in spreads. If the threshold change was to improve other type of information, we would expect an impact for clients with any delay history. As pointed, financial institutions could cheaply have this information from private bureaus. Column (2) presents the results of Eq. 1 with the sample of clients with delay history. We find no statistical and economic impact on the spread. Additionally, we test if the result would change with different levels of default. With our data, we only know if a client is in default within a range of days. We test the change in the information setting for interest rate spread including clients that delayed repaying their debts between 0 and 15 days. Column (3) report the result. Our sample almost doubles, but the previous result vanishes. By law, institutions can send the information to private credit bureaus just after a few days.

Further, to rule out that our result came from other financial information other than the positive, we run again Eq. 1 but with the treatment being borrowers that already had info in SCR at the time of regulation but new information was made available with the change in the threshold (ie. had loans below R$1,000 with another bank). Banks could already infer the creditworthiness of these clients with the information that were already visible at the SCR. So, we would expect no effect with the threshold change. Indeed, Table 6 shows the results for this new estimation and no variable significantly changed with more information of those borrowers.

In sum, we find evidences that the decrease in the credit cost for clients is due to the change in informational environment about good borrowers' creditworthiness.

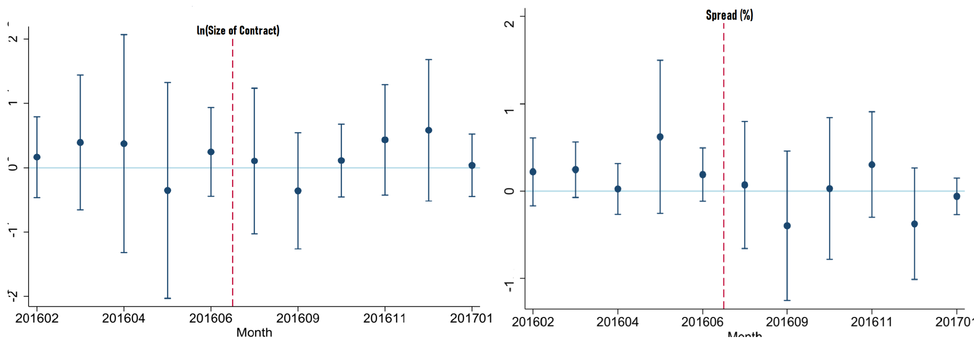
1.5. Conclusion

This paper presents evidence on the causal effect of reducing informational asymmetries on credit outcomes using a comprehensive dataset of Brazilian loans to households. We use a change in regulation, in June 2016, that enhanced the information set of financial institutions about borrowers with good credit repayment history to identify impacts on the cost of getting credit, loan size, and maturity. We employ a difference-in-differences approach to compare households whose financial institutions could already access positive information about them, with households whose banks could not have any positive information before the change in regulation.

Our main result shows a moderate decrease in the credit cost for borrowers. We estimate a reduction of 33.5 p.p. in the loan spreads. Also, we show evidence that our result is driven by smaller banks that could better and cheaply screen good borrowers.

Our result may have potential effects on the real economy. On the one hand, it can reduce financial inclusion for bad payers (Foley et al. (2019)). On the other, it can alleviate credit restrictions for good payers. All, information asymmetry is an important determinant of lending rates. Thus, measures to improve positive information access should increase the efficiency of credit markets, benefiting on-time borrowers.

Figure 1. Development of log of the size of credit and interest rate spread for first-time borrowers with no delays in loan repayment



Note: Coefficients β_1 from Eq. 1 for each month before June 2016. Bars show 99% confidence intervals. We normalize Jan/06 $= 0$.

Figure 2. Frequency of loan contracts size before and after June 2016 (histogram)

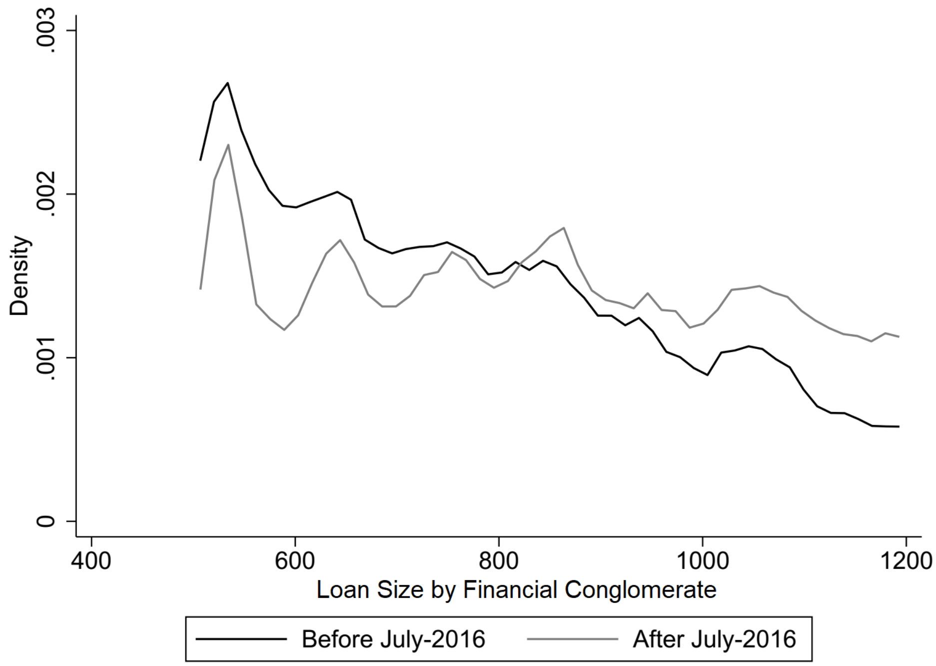


Table 1: Share of oiginations for clients with maximum total liability per bank ranging between R$500 and R$2,000 in June 2016

Credit Line	**Share**
Credit Card	40%
Overdraft	20%
Personal Loans	12%
Credit card revolving credit	12%
Payroll-deducted personal loans	5%
Other	12%

Table 2: Descriptive Statistics

	(1)	(2)	(3)	(4)
	Control		***Treatment***	
	Before	**After**	**Before**	**After**
Panel A: Loan Outcomes				
Origination per Bank (R$)	951	1,540	668	1,370
Spread	280%	350%	271%	402%
Maturity (Months)	10	14	9	13
Risk (AA = 1 to D = 5)	2.3	2.2	2.4	2.2
Panel B: Borrower Characteristics				
Monthly Income (R$)	1,327	1,342	1,076	1,219
Age (Years)	45	47	44	45
Gender (Female)	65%	63%	66%	65%
Panel C: Financial Characteristics				
Private Banks	90%	92%	84%	94%
Loans with Small & Medium Size Banks	69%	71%	57%	68%
Indebtedness (Total Credit/Annual Income)	14%	27%	11%	22%
Debt Service Ratio (Principal + Interest)	37%	68%	30%	54%
Personal Credit/Total Debt	74%	99%	81%	85%
# of banks in the Municipality	13	13	12	14
# of observations	108	311	441	615

Table 3: Baseline results for difference-in-differences estimation on the cost of credit, loan size, loan maturity, and number of banks that a borrow has a loan

	(1)	(2)	(3)	(4)
	Lending Spread	**ln(Loan Size)**	**Maturity**	**# of Bank Relationship**
Post * Info Disclosured	-0.335**	0.077	-3.127	-0.152
	(0.159)	(0.129)	(1.866)	(0.212)
Controls	Y	Y	Y	Y
Time-Bank FE	Y	Y	Y	N
Borrower FE	Y	Y	Y	Y
Time FE	N	N	N	Y
Avg Dep. Variable	304%	R$ 911	10.6 months	2.2
Obs		1,475		258
R-squared	0.954	0.851	0.842	0.893
# Banks		37		-

Note: OLS estimation of Eq.1 results for loans originations from January 2016 up to January 2017 for new clients in the conglomerate with no default history. Post is a binary variable equal to one for loan originations after June 2016, the date of the credit registry reduced the threshold. Treatment is a binary variable equal to one for clients with no history in the credit registry (SCR). Controls: income, total indebtedness in the financial system, debt-to-service ratio, number of bank relationships (except for column (4)), and number of banks in the municipality of the borrower. Standard errors are clustered at the conglomerate level and reported in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Results for difference-in-differences analyzes on with different banks classified by size

	(1)	(2)	(3)
	Lending Spread	**ln(Loan Size)**	**Maturity**
Panel A: Small & Middle Sized Banks			
Post * Info Disclosed	-0.424***	0.240***	0.157
	(0.107)	(0.082)	(1.499)
Obs		595	
R-squared	0.961	0.900	0.908
# Banks		30	
Panel B: Large Banks			
Post * Info Disclosed	-0.037	-0.473	-0.112
	(0.087)	(0.523)	(1.145)
Obs		142	
R-squared	0.957	0.934	0.913
# Banks		13	
Controls	Y	Y	Y
Time-Bank FE	Y	Y	Y
Household FE	Y	Y	Y
Time FE	N	N	N

Note: OLS estimation of Eq.1 results for loans originations from January 2016 up to January 2017 for new clients in the conglomerate with no default history. Post is a binary variable equal to one for loan originations after June 2016, the date of the credit registry reduced the threshold. Treatment is a binary variable equal to one for clients with no history in the credit registry (SCR). Controls: income, total indebtedness in the financial system, debt-to-service ratio, number of bank relationships, and number of banks in the municipality of the borrower. Small-Mid sized banks are those with less than 4 million clients and large banks above this number. Standard errors are clustered at the conglomerate level and reported in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Results for difference-in-differences analyzes - Robustness

Dependent Variable Spread	(1) **Clients with any default**	(2) **Clients w/ default below 15 days**	(3) **Old Clients**
Post * Info Disclosed	0.125	-0.264	-0.011
	(0.108)	(0.167)	(0.014)
Controls	Y	Y	Y
Time-Bank FE	Y	Y	Y
Household FE	Y	Y	Y
Obs	4,107	2,192	293,543
R-squared	0.953	0.936	0.959

Note: OLS estimation of Eq.1 results for loans spread originations from January 2016 up to January 2017. Post is a binary variable equal to one for loan originations after June 2016, the date of the credit registry reduced the threshold. Treatment is a binary variable equal to one for clients with no history in the credit registry (SCR). Controls: income, total indebtedness in the financial system, debt-to-service ratio, number of bank relationships, and number of banks in the municipality of the borrower. Column (1) uses a sample of borrowers that were on default. Column (2) uses a sample of borrowers that had arrears between 1 and 15 days. Last column uses a sample of borrowers that took a loan with a bank that she already had a credit relationship. Standard errors are clustered at the conglomerate level and reported in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6: Results for difference-in-differences analyzes for borrowers that already had info in SCR at the time of regulation but new information was made available with the change in the threshold

	(1)	(2)	(3)
	Lending Spread	**ln(Loan Size)**	**Maturity**
Post * Info Disclosured	-5.121	0.082	0.271
	(16.216)	(0.100)	(1.364)
Controls	Y	Y	Y
Time-Bank FE	Y	Y	Y
Household FE	Y	Y	Y
Time FE	N	N	N
Avg Dep. Variable	183%	R$ 879	8.3 months
Obs		1,261	
R-squared	0.941	0.847	0.853
# Banks		34	

Note: OLS estimation of Eq.1 results for loans originations from January 2016 up to January 2017 for new clients in the conglomerate with no default history. Post is a binary variable equal to one for loan originations after June 2016, the date of the credit registry reduced the threshold. Treatment is a borrowers that already had info in SCR at the time of regulation but new information was made available with the change in the threshold. Controls: income, total indebtedness in the financial system, debt-to-service ratio, number of bank relationships (except for column (4)), and number of banks in the municipality of the borrower. Standard errors are clustered at the conglomerate level and reported in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

2. Deposit Insurance Effects: Competition and Risk-Taking

2.1. Introduction

The debate on deposit insurance (DI) has focused on whether explicit DI benefits of attenuating financial instability by reducing bank runs (Diamond and Dybvig (1983)) outweigh moral-hazard costs, such as increased risk taking by insured institutions as a result of reduced depositors' incentives for monitoring institution's performance (see Calomiris and Jaremski (2016) and Anginer and Demirguc-Kunt (2018) for a review). However, the theoretical literature also points that DI can also change the degree of competition and, as a consequence, increase the chances of bank failure (Matutes and Vives (1996)) and financial instability (Baltensperger et al. (1987), Vives (1991)). These ambiguous theoretical predictions regarding DI, banking industry competition and sector stability have been sparsely explored by the literature with an empirical approach.

The effects of DI on competition are not clear. In one hand, DI can foster greater competition: in a scenario with a partially insured financial system or one with no insurance at all, depositors are incentivized to put their savings in larger banks[27] or public banks[28]. Thus, wider deposit insurance provision can level the playing field and allow smaller banks to attract deposits and thus foster greater competition (Camara and Montes-Negret (2006)). In addition, Matutes and Vives (1996) theoretically show that deposit insurance results in more deposits because it increases the expected payment to depositors, and that this market-expansion effect may be sufficiently important to transform the market structure from local monopolies into one of direct competition. On the other side, deposit insurance structure can increase the costs of compliance, which will weigh heavily on smaller financial institutions and discourage the continuation and entry of new competitors.

As DI can change the degree of competition for deposits, Matutes and Vives (1996) also point that this can increase the chances of bank failure by changing the cost of funding, compelling banks to make riskier investments to maintain profitability, and, as consequence, increase financial instability (Baltensperger et al. (1987), Vives (1991)).

Despite those conflicting evidences, some political institutions (such as the IMF) endorse DI adoption. As of 2019, 42% of countries still didn't have deposit insurers[29].

This paper uses a unique empirical setting, exploring if the introduction of an explicit policy on deposit insurance funds affected competition or stimulated more risk taking. In 2014, the Central Bank of Brazil (BCB) instituted that credit unions (CU) had to create a deposit insurance fund (FGCoop). The new regulatory rule affected only part of the financial system, without direct influence on the operations of banks[30] in municipalities where there were no credit unions. Using the creation of FGCoop as a quasi-natural experiment and the findings of Nguyen (2019) that banking markets are local, a difference-in-differences framework is used where the treatment group, in the baseline analyses, consists of bank branches in municipalities where there was at least one CU. The control group is bank branches in municipalities where no CUs were operating. The identification assumption is that without the creation of FGCoop, the outcomes in treated and non-treated bank branches in those municipalities would have followed equal trajectories.

[27]Larger banks tend to be more diversified and are candidates to benefit from too-big-to-fail measures.
[28]As the government is the main stakeholder, the chances of letting the bank fail are lower.
[29]Source: 2020 International Association of Deposit Insurers Survey with members and non-members.
[30]Banks in Brazil offer deposit insurance since 1995.

Results are divided in three parts. First, the study documents that bank deposits reduced 3% after CU's deposit insurance implementation and, at the same time, that this increased transaction deposit volume in CUs, indicating an effect on the capability of DI for attracting deposits to CU. Additionally, it observes that the effect was higher in municipalities where CU had more market share. Lastly, it demonstrates that deposit insurance didn't encourage an increase in riskier loans by banks.

Those findings make some noteworthy contributions to the literature on deposit insurance. The study highlights that most of the literature on DI is based on cross-country comparisons or on time series within a country, which does not allow a causal identification of mechanisms, or control for local specificities as this study's empirical setting. The closest to this work is the study of Calomiris and Jaremski (2019), that explores the context of early 20th century in US, when local states installed deposit insurance for unit state-chartered banks that would operate in parallel to the uninsured national banks. Similar to this study's results, they show that deposit insurance increased insured banks' deposits. Also, that the deposits flowed from relatively stable banks to risky banks and that this increased risk, by removing the market discipline in the deposit market. This work distinguishes itself by using a more recent framework where two types of insured financial institutions would be competing for deposits.

This work also relates to the literature that assesses the effect of competition on financial stability. The literature has found ambiguous evidence in this link. In one aspect, in an environment with greater competition, the pressure for profits will make banks choose riskier portfolios, leading to greater fragility (Hellmann et al. (2000)). In another, greater competition might lead to lower loan rates, reducing borrowers' riskier incentives and, thus, diminishing bank's exposure to risk of failure (Boyd and Nicoló (2005)). Matutes and Vives (1996) point that competition coming from DI makes banks more aggressive in bidding for deposits, thereby increasing failure probabilities. This paper differentiates itself from the empirical work done so far by linking the direct competition from DI. It is important to highlight that a large share of the literature focused on moral hazard problems as a result of reduced depositors' incentives for monitoring newly insured institutions. The focus of this work is to assess the moral-hazard problem arising from competition.

The study proceeds as follows: Section 2.2 describes the institutional background. Section 2.3 describes the data, its characteristics and also discusses our empirical framework. Section 2.4 presents the empirical results. Section 2.5 concludes.

2.2. Background

2.2.1. Credit Unions in Brazil

Credit Unions (CU) in Brazil, like banks, can accept deposits, make loans, and provide a wide array of other financial services, however, only for their members. CUs are also subject to Central Bank supervision with simplified regulatory rules, as their size and product shelf are small-scale compared to banks[31].

An important difference from banks is that credit unions are not-for-profit organizations and are also member-owned institutions. To become a member it is necessary to buy a common bond that can yield up to 100% of the Selic rate (Brazilian federal funds rate) and the CUs' operating result is redistributed to its members. A survey from 2017 has shown that the membership could be symbolic, with a R$1 price, but has a median value of R$30[32]. This bond gives the member only one vote in the general assembly - where

[31] For example, capital requirement rules are less strict and less accounting information is required to be sent to BCB.

[32] Equivalent to 3% of 2017's minimum wage. Source: Obervatório de Cooperativas - FEARP/USP (2017).

members can deliberate on business related to the CU -, regardless of the equity share. Nevertheless, the daily decision and management is left to the board of directors.

Members, households and firms, may be able to join based on specific characteristics (e.g. profession, place of work, business sector) or without any restriction or requirement (free admission). In 2014, free admission CUs represented 29% of all the segment.

In terms of risk, CUs are conservative in their risk management, as the average Basel ratio[33] of their segment was around 30% in 2019, much higher than the financial system as a whole, 18%, and the minimum legal requirement, 11%[34]. This translated in only few episodes of CU liquidation (bankruptcy). It is worth noting that, on average, CUs have lower interest rate compared to banks.

BCB data from 2014 shown that there were more than 1,100 CUs in Brazil against only 96 commercial banks (conglomerate). Despite the broad presence in Brazil, CUs attracted only 3.1% of deposits in the country[35]. Also, CUs' average size (assets) was R$86 thousand versus R$59,500 thousand of banks.

Figure 3 shows that 43% of municipalities in Brazil were served with at least one CU[36], but that those unions were concentrated in the South and Southeast regions of Brazil. CUs usually operate only in one municipality, while banks were present themselves, on average, on 90 different municipalities. In Panel B of Table 7 presents some descriptive statistics of municipalities with and without CU, pre and post FGCoop. On average, CUs operate in richer (higher GPD per-capita) and more populated cities.

2.2.2. Credit Union Deposit Insurance Fund (Fundo Garantidor do Cooperativismo de Crédito - FGCoop)

In October 2012, the Central Bank of Brazil established a credit union deposit insurance fund for CUs and that all of them, compulsorily, would need to be associated[37]. The so called FGCoop started its operation in April 2014. Before that, CU depositors did not have any explicit insurance against default[38]. This regulation was part of a broader legal movement that aimed to regulate the operation of credit unions[39], not a response to higher risk-taking by CUs as illustrated in Figure 4.

CUs have to pay a monthly fee of 0.0125% on balances eligible to the insurance[40]. In the case of a CU's inability to pay back depositors, FGCoop will guarantee a maximum payment amount of R$250,000 per client and per institution to cover the deposits. The insurance covers a broad array of products, including transaction and time deposits[41]. The value and products insured were the same offered by FGC (banks deposit insurer) at the time of FGCoop's foundation.

[33] Risk weighted assets as the denominator of the capital ratio.

[34] Source: 2016's Central Bank of Brazil report on Credit Union System

[35] The Brazilian financial system is highly concentrated, with five institutions accounting for 87% of all deposits.

[36] Banks have branches in at least 66% of Brazilian municipalities.

[37] Resolution 4.150.

[38] In this paper we are excluding Credit Unions Banks and Federations, as they receive deposit only from CUs.

[39] Legislative Report of Complementary Bill No. 177, 2004

[40] Same rate as banks in 2014.

[41] 70% of their fund-raising.

2.3. Empirical Strategy and Data

2.3.1. Empirical Strategy

To assess the impact of a compulsory adoption of deposit insurance on competition and risk-taking, this study uses the creation of FGCoop as a quasi-natural experiment. The new regulatory rule affected only part of the financial system, without direct influence on the operations of banks. For this analysis, different identification strategies are used because of the availability of the data, but always within a difference-in-differences methodology.

Deposit Insurance, Competition and Risk-Taking

From a depositor's perspective, having two financial institutions to deposit their savings with the same risk and deposit rate, the one that offers an insurance will be preferable, granting an advantage for banks on attracting deposits. Thus, all else equal, FGCoop creation can level the playing field, facilitating CUs to attract deposits and thus fostering greater competition in the financial system. Despite this prediction, CUs may not be able to dilute the compliance costs with scale and, thus, would face a proportionately larger burden, discouraging both the shut down of CUs and the entry of new competitors.

To evaluate the impact of increasing competition, the study uses a geographical strategy, similar to Joaquim (2019): it is expected that in municipalities where CUs were operating when FGCoop was found, banks will find a more competitive environment for attracting deposits. Thus, bank branches in those municipalities are used as treatment. At the same time, in municipalities where there was no credit union, the competitive environment will not change for banks. Consequently, these are used as control. It is expected that if the competition is fostered, the study will find a decrease in bank deposits. It is also used a propensity score matching to approximate control and treatment at the municipality level, using GDP per capita and deposits per inhabitant pre-FGCoop as key characteristics (see Table 7).

The same empirical strategy applies for risk-taking by banks after the FGCoop foundation. In one aspect, in an environment with greater competition, the pressure on profits will make banks choose riskier portfolios, leading to greater fragility (Hellmann et al. (2000)). In another, greater competition might lead to lower loan rates, reducing borrowers' riskier incentives and, thus, diminishing bank's exposure to risk of failure (Boyd and Nicoló (2005)). Thus, it is expected that, if indeed competition is to be increased, the stimulus for risk-taking by banks can change. It is worth noting that risk management for national banks might not be on branch level, but profit seeking must be local, and may still change the risk profile coming from competition.

It is true that Credit Unions cannot attract deposits as easily as banks, due to the necessity of becoming a member or being part of a community/profession or fulfilling other requirements. Therefore, the study restricts the treatment sample only to municipalities where there are CUs of free admission. Since the information about the membership price of CUs isn't available, the study relies on the low prices evidenced before.

The empirical model is a difference-in-difference specification that takes the form of:

$$Y_{i,m,t} = \beta_1 CU_i^* Post_t + \beta_2 X_{i,m,t-1} + \beta_3 Z_{m,t-1} + \alpha_{i,t} + \alpha_m + \alpha_t + \epsilon_{i,m,t} \tag{3}$$

Where i stands for financial institutions during period t and in the municipality m. The dependent variables (Y) to evaluate competition will be the log of total deposits, transaction deposits or time deposits. To investigate the risk-taking, it will use as a dependent variable provision for credit losses (total provision to total credit balance) and the share of loans without guarantee in relation to total loans.

The dummy variables CU_i and $Post_t$ take on the value of one if, respectively, the observation is a bank in a municipality with a CU and are recorded after the introduction of the deposit insurance in April 2014. The coefficient β_1 represents the difference-in-differences estimates of the effect of deposit insurance. If competition is to be increased with deposit insurance it is expected $\beta_1 < 0$, and if banks become more prone to risk, as hypothesized, it is expect the coefficient to have the following sign $\beta_1 > 0$.

X represents a vector of bank-branch-level controls that varies over time across banks and municipalities. These include log of total assets, and loans-to-asset. The study also controls for time-varying indicator Z in the municipality level: share of government-owned banks. αs represents fixed-effects for financial institution-time, municipality, and time. Standard-errors are clustered in the municipality level.

In any of the analysis, the identification relies on the assumption that in the absence of the deposit insurance, depositors would not change their preferences between banks and credit unions. Despite this assumption not being testable, evidence of its validity is provided in Figure 5 by examining the outcomes of treatment and control before the institution of the FGCoop.

2.3.2. Data and Descriptive Statistics

All banks and credit unions are required to disclose their balance sheets to the Central Bank of Brazil on a monthly basis. The study's main database comes from banks' disclosure at branch level (ESTBAN). It uses, additionally, some controls and information from Credit Union aggregate level information (Cosif)[42]; the income statement and credit portfolio disclosed to the public in aggregate level and on a quarterly basis (IF.data). All information is made available to the public within three months.

Another important point for the study's identification is that the results are not driven by other changes in regulation, which may be correlated with unobservable, thereby biasing its estimates. Therefore, the sample is restricted to the period between June 2013 and March 2015 (one year before and after the creation of FGCoop). Quarterly data is used to ensure the same frequency as the aggregate data.

Descriptive Statistics are presented in Table 7. As CUs operate in larger municipalities, in Panel A is shown that financial institutions (CUs and banks) in municipalities with CUs are larger in assets and deposits than banks without CUs. While CUs are slightly larger than bank-branches in control group, banks in treatment group are almost twice the size of CUs. In terms of risk, banks, both in control and treatment group, are similar in the share of loans without guarantee and loan loss provision, while CUs provision more, but with smaller share of loans without guarantee.

2.4. Results

2.4.1. Baseline Results

Competition

Table 8 presents the result estimates of equation 3. In Column (1), the outcome variable is the log of total deposit balances of bank branches and results point that the deposit insurance for credit unions implemented in April 2014 might have induced a decrease of 3% in deposit balances. Columns (2) and (3) break total deposit into transaction and time deposit, respectively. Column (2) shows no statistical significant decrease in transaction deposit. However, column (3) indicates that the result is being driven by a decrease in time deposits of 3.2%.

[42]Despite not having data at branch level from CUs, data from BCB show that they operate (practically) in only one municipality.

To make sure that the decrease in bank deposits found in Table 8 is going to credit unions, equation 3 is run, where the treatment will be CUs, instead of banks[43]. Results are in Table 9. Column (1) shows that there's no increase in total deposit to credit unions. However, breaking into transaction and time deposits, it finds an increase of 7.9% in transaction deposit at 5% significance, but no outcome in time deposit. As previously highlighted, credit unions are member-owned financial institutions and the membership yields interest. Thus, part of decrease in time deposit in banks might have gone to membership payment, that are not guaranteed by FGCoop, and part left in demand deposits. I carefully interpret these findings, since increase in CUs' deposits might be a result from spillover from other municipalities.

2.4.2. Heterogeneity in Competition

The study also investigates if the intensity of credit unions' local market-share can have different impacts in attracting more deposits. For that, the following equation is run

$$Y_{i,m,t} = \beta_1 ShareCU_m^* Post_t + \beta_2 X_{i,m,t-1} + \beta_3 Z_{m,t-1} + \alpha_i + \alpha_m + \alpha_t + \alpha_{m,t} + \epsilon_{i,m,t} \tag{4}$$

Differently from Eq. 3, $ShareCU$ will be a continuous variable of the average market-share of deposits of credit unions in the municipality in March 2014 (control municipalities will be equal zero). In doing this analysis, the geographical control (m) will change for the microregion concept used in the Brazilian Census.

Table 10 presents results similar to Table 7: the higher the deposit market share of credit unions before the implementation of FGCoop, the more banks lost time deposit. Column (3) indicates that for 1 p.p. more of market share the credit unions had, after FGCoop it would decrease bank's deposit by 9.3%.

Combining results discussed above, the deposit insurance fund was able to attract deposits for credit unions, increasing competition of the financial market in Brazil. As the shift of deposits was relatively small, structural changes in the market structure might take time[44].

Loans

Despite being beyond the primary focus of the paper, it was also possible to assess if the change in deposit balances affected FIs' credit portfolio. Table 11 presents the result estimates of equation 3 with the outcome variable being the total outstanding loans. In Column (1), the treatment are banks and in Column (2) CUs. In both cases, results point to no effect in the credit portfolio. As noted, the immediate change in deposit balances seems to be small for any impact in FIs' credit decisions.

Risk-Taking

Table 12 presents results for deposit insurance in risk-taking by banks for different metrics. Ideally, information of new originations would be used. However, as this information is not publicly available at the financial institutions level, the balance of loans is used, which makes it unable to account for the term of contracts.

Column (1) provides evidence that after FGCoop, banks didn't increase their loan loss provision. Additionally, the study looks for more evidence that this result can be interpreted as no risk taking. Column (2) presents evidence that the share of credit balances without collateral didn't change after the creation of the deposit insurance fund. Altogether, banks seem to not have incurred in more risk taking after the increase in competition from DI.

[43]Regressions with CU as treatment also need to be controlled for specific time-trend for CUs

[44]With similar methodology, I ran a regression for the Herfindahl-Hirschman Index (omitted in this work) and find no statistical and economical significant change after FGCoop.

2.4.3. *Robustness*

In this section, some robustness tests are presented. The increase in competition finding can be a result of a pre-trend, thus, 3 is run, but restricting the sample to the period before FGCoop implementation (April 2014). $PostPlacebo_t$ takes the value of one for quarters after November 2013. Table 13 present the results and no statistical and economic effect is found in this exercise, thus excluding the possibility that the study's results are being driven by a pre-trend.

The increased competition found could also be due to the fact that CUs may have increased their deposit rates. To test this theory, I investigate if CU funding costs increased, as deposit rates are not available. The following equation is run:

$$Y_{i,t} = \beta_1 CU_i^* Post_t + \beta_2 X_{i,t-1} + \alpha_i + \alpha_t + CUtrend + \epsilon_{i,t} \tag{5}$$

Where the data will come from aggregate income statements at financial institution level. The treatment will be CUs and control will be banks not operating in municipalities where there is a CU. The sample is matched by log of assets and risk, to account for the fact that CUs are smaller. Results in 14 show that funding costs with deposits didn't increase if compared to banks.

2.5. *Conclusion*

This paper presents evidence on the causal effect of deposit insurance in stimulating competition, but not inducing more risk. A change in regulation in April 2014 is used, that required all credit unions to be part of a unique deposit insurance fund. The new regulatory rule affected only part of the financial system, without direct influence on the operations of banks. Employing a difference-in-differences methodology with different treatment groups, the study concludes that deposit insurance can be important in leveling the playing field for smaller financial institutions to attract deposits and thus foster greater competition in the financial system. Also, no persistent evidence that the deposit insurance increased the appetite for risk from banks is found.

The study highlights that the results can be important for the real economy, especially in Brazil where the credit market is concentrated, with high interest rates. On one hand, credit unions offer cheaper loans than banks - aggregate data from BCB show an average of 18 p.p. difference for lending for micro and small firms. On the other, imposing the costs of deposit insurance for credit unions that are member-owned, having lower risks of 'bank runs', can have a negative effect for consumers. The net effect of DI is difficult to estimate, but all consequences need to be considered. In a future research, it would be interesting to explore the real effects of deposit insurance on loans variables.

Figure 3. Municipalities with Credit Unions in Brazil

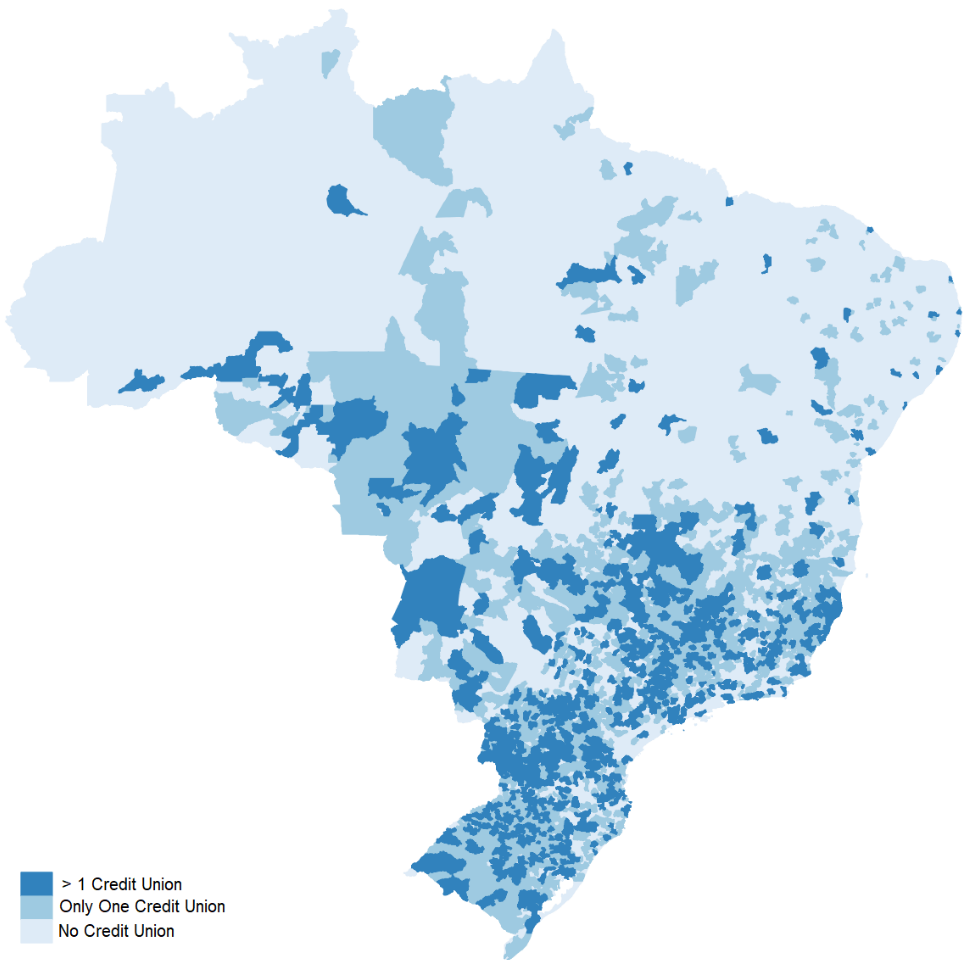


Figure 4. Credit Provision over Total Loans of Credit Unions before FGCoop regulation

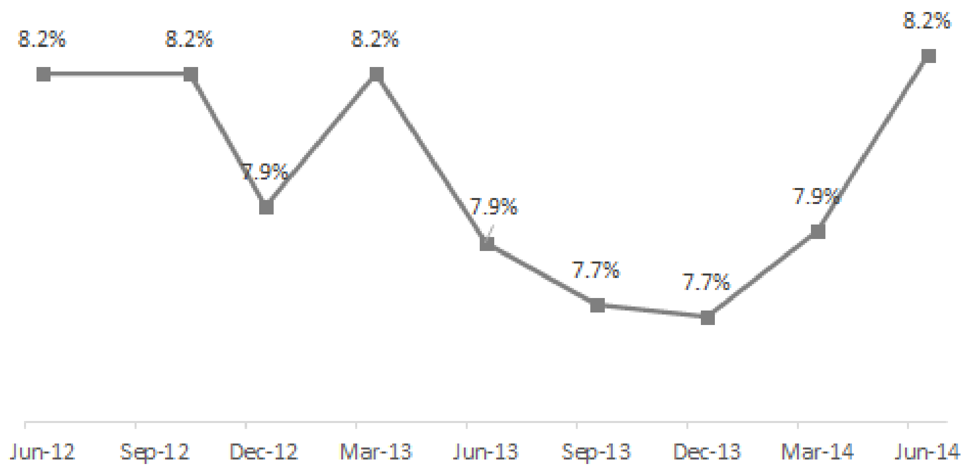
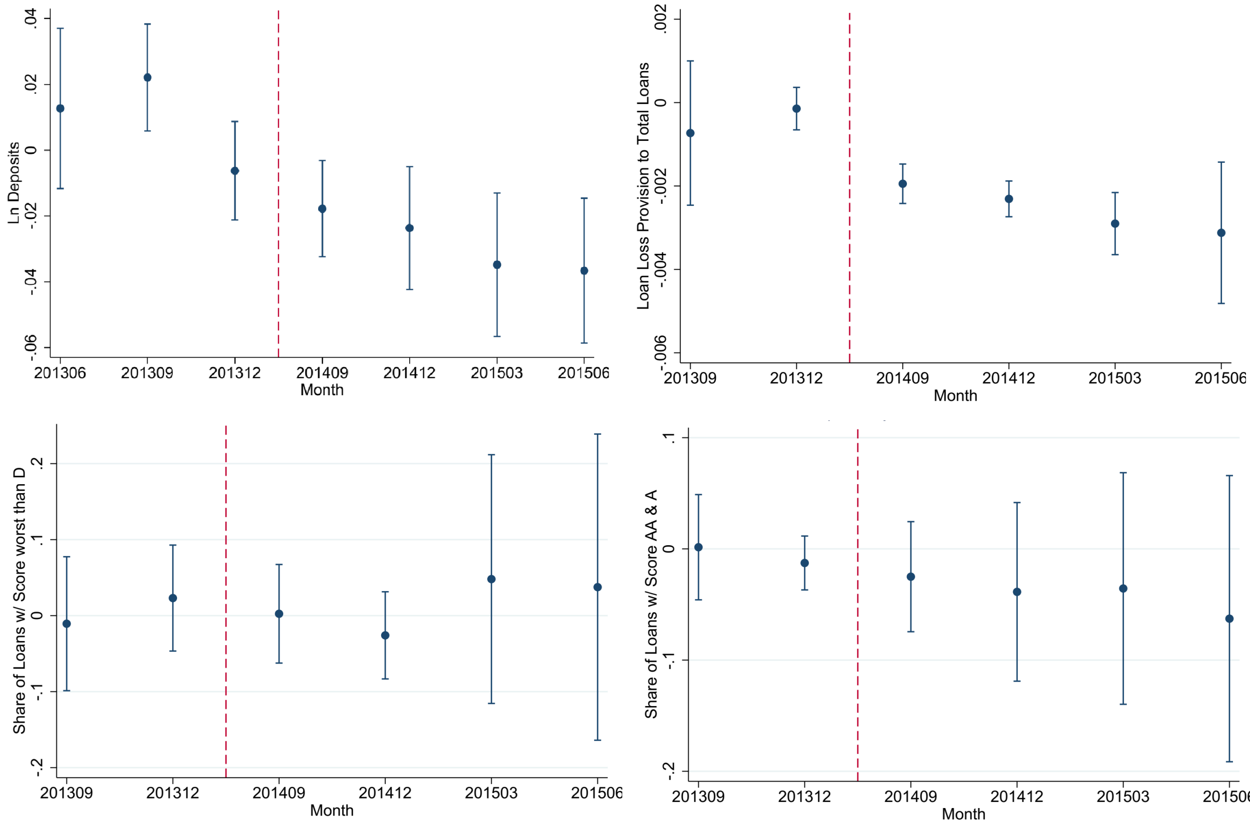


Figure 5. Parallel Trends: development of coefficient estimations



Note: Coefficients β_1 from Eq. 3 for each quarter. Bars show 95% confidence intervals. We normalize March 2014 = 0.

Table 7: Descriptive Statistics from the matched sample: means

	Pre			Post		
Panel A: FI-branch level	**Bank w/ CU**	**Bank wo CU**	**CU**	**Bank w/ CU**	**Bank wo CU**	**CU**
Total Asset (R$ k)	778,800	235,700	348,500	1,023,000	360,300	417,100
Loans to Asset	34%	34%	24%	33%	33%	23%
Time Deposit (R$ k)	114,400	41,600	43,820	126,500	46,620	52,190
Transaction Deposit (R$ k)	10,410	3,751	13,420	10,220	3,802	15,010
Total Deposit (R$ k)	124,900	45,350	57,240	136,700	50,420	67,200
Share of Loans wo Guarantee	86%	88%	67%	82%	86%	67%
Loan Loss Provision to Total Loans	3.7%	3.8%	4.5%	3.9%	4.3%	4.9%
Observations	1,297	4,710	1,326	1,297	4,710	1,326

	Pre		Post	
Panel B: Municipality level	**wo CU**	**w/ CU**	**wo CU**	**w/ CU**
GDP per-capita	18,520	22,933	18,011	22,229
Population	59,849	129,899	61,072	132,593
HHI (deposit)	9,039	6,844	9,029	6,829
Share CU	-	29%	-	31%
Obs	995	285	995	285

Note: All monetary variables are inflation adjusted by IGP-DI Index of March 2011

Table 8: Deposit insurance effect on competition: deposit balances of banks

	(1)	(2)	(3)
	Ln Total Deposit	**Ln Transaction Deposit**	**Ln Time Deposit**
CU * Post	-0.030***	-0.010	-0.032***
	(0.009)	(0.016)	(0.010)
Controls (t-1)	Y	Y	Y
Bank-Time FE	Y	Y	Y
Municipality FE	Y	Y	Y
Time FE	Y	Y	Y
Obs.		11,985	
R-squared	0.976	0.940	0.972
Dep. Var. Mean	R$ 65,600,000	R$ 5,177,971	R$ 60,400,000

Note: OLS estimation of Eq.3 results for deposit balances in bank branches from June 2013 up to March 2015. The dummy variables, *CU* and *Post*, take on the value of one if, respectively, the observation is a bank in a municipality with a CU of free admission and is recorded after the introduction of the deposit insurance in April 2014. Controls with one quarter lag: log of total assets, loans-to-asset ratio, liquid assets to deposit, share of government-owned banks in the municipality. Standard errors are clustered at the municipality level and reported in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 9: Deposit insurance effect on competition: deposit balances of credit unions

	(1) **Ln Total Deposit**	(2) **Ln Transaction Deposit**	(3) **Ln Time Deposit**
CU * Post	0.004	0.079**	0.001
	(0.016)	(0.036)	(0.016)
Controls (t-1)	Y	Y	Y
FI FE	Y	Y	Y
Municipality FE	Y	Y	Y
Time FE	Y	Y	Y
Obs.	14,633	14,567	14,630
R-squared	0.980	0.949	0.977
Dep. Var. Mean	R$ 64,900,000	R$ 7,000,264	R$ 57,900,000

Note: OLS estimation of Eq.4 results for deposit balances in financial institutions branches from June 2013 up to March 2015. The dummy variables, *CU* and *Post*, take on the value of one if, respectively, the observation is a credit union of free admission and is recorded after the introduction of the deposit insurance in April 2014. Controls with one quarter lag: log of total assets, loans-to-asset ratio, liquid assets to deposit, share of government-owned banks in the municipality, return-on-equity. Standard errors are clustered at the financial institution level and reported in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 10: Heterogeneous effect of deposit insurance on banks deposits: market-share of credit unions in the municipality

	(1)	(2)	(3)
	Ln Total Deposit	**Ln Transaction Deposit**	**Ln Time Deposit**
Post * Share CU	-0.089***	-0.039	-0.093***
	(0.034)	(0.040)	(0.036)
Controls	Y	Y	Y
Bank-Time FE	Y	Y	Y
Microregion FE	Y	Y	Y
Time FE	Y	Y	Y
Obs.		11,985	
R-squared	0.935	0.889	0.932

Note: OLS estimation of Eq.4 results for deposit balances in bank branches from June 2013 up to March 2015. The dummy *CU* take on the value of one if in column (1) if the observation is a bank in a municipality with a CU of free admission and in column (2) if it is a credit union of free admission. *Post* takes the value of one if is recorded after the introduction of the deposit insurance in April 2014. Controls with one quarter lag: log of total assets, loans-to-asset ratio, liquid assets to deposit, share of government-owned banks in the municipality, return-on-equity. Standard errors are clustered at the financial institution level and reported in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 11: Deposit insurance effect on loans

	(1)	(2)
	Ln Loans Bank	**Ln Loans CU**
CU * Post	-0.010	0.003
	(0.014)	(0.010)
Controls (t-1)	Y	Y
FI FE	Y	Y
Municipality FE	Y	Y
Time FE	Y	Y
Obs.	12,058	15,162
R-squared	0.920	0.952
Dep. Var. Mean	R$ 36,089,450	R$ 42,701,835

Note: OLS estimation of Eq.3 results for loans in FI branches from June 2013 up to March 2015. *ShareCU* is a continuous variable of the average market-share of deposits of credit unions in the municipality before the creation of the FGCoop. *Post* takes on the value of one if the observation is recorded after the introduction of the deposit insurance in April 2014. Controls with one quarter lag: log of total assets, loans-to-asset ratio, liquid assets to deposit, share of government-owned banks in the municipality. Standard errors are clustered at the microregion level and reported in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 12: Deposit insurance effect on risk-taking

	(1) **Loan Loss Provision to Total Loans**	(2) **Share of Loans wo Guarantee**
CU * Post	-0.080	-0.230
	(0.077)	(0.003)
Controls (t-1)	Y	Y
Bank-Time FE	Y	Y
Municipality FE	Y	Y
Time FE	Y	Y
Obs.	10,073	12,032
R-squared	0.752	0.761
Dep. Var. Mean	4.7	

Note: OLS estimation of Eq.3 results for risk-taking measures by banks from June 2013 up to March 2015. The dummy variables, *CU* and *Post*, take on the value of one if, respectively, the observation is a bank and is recorded after the introduction of the deposit insurance in April 2014. Controls with one quarter lag: log of total assets, loans-to-asset ratio, liquid assets to deposit, share of government-owned banks in the municipality, return-on-equity. Standard errors are clustered at the financial institution level and reported in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 13: Robusteness test for deposit insurance effect on competition on deposit balances of banks

	(1)	(2)	(3)
	Ln Total Deposit	**Ln Transaction Deposit**	**Ln Time Deposit**
CU * Post Placebo	-0.005	-0.003	-0.008
	(0.008)	(0.020)	(0.008)
Controls (t-1)	Y	Y	Y
Bank-Time FE	Y	Y	Y
Municipality FE	Y	Y	Y
Time FE	Y	Y	Y
Obs.		5,964	
R-squared	0.975	0.940	0.971

Note: OLS estimation of Eq.3 results for deposit balances in bank branches from June 2013 up to March 2014. *CU* is a dummy variable that takes on the value of one if the observation is a bank in a municipality with a CU of free admission. *PostPlacebo* takes the value of one for quarters after November 2013. Controls with one quarter lag: log of total assets, loans-to-asset ratio, liquid assets to deposit, share of government-owned banks in the municipality. Standard errors are clustered at the municipality level and reported in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 14: Robusteness test for deposit insurance effect: No increase in Funding Costs

	Funding Cost over Deposits
CU * Post	0.007
	(0.005)
Controls (t-1)	Y
Bank FE	Y
Time FE	Y
Obs.	6,004
R-squared	0.066

Note: OLS estimation of Eq.5 results for funding cost over total deposits in financial institutions from June 2013 up to March 2014. *CU* is a dummy variable that takes on the value of one if the observation is a credit union and zero if is a bank. *Post* takes the value of one for quarters after November 2013. Controls with one quarter lag: log of total assets, return on equity, credit over assets. Standard errors are clustered at the financial institution level and reported in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

3. Payroll Payments Microdata and Aggregate Labor Market in Brazil

3.1. Introduction

Monitoring the evolution of employment in Brazil became challenging with the onset of the Covid-19 pandemic, as the two main traditional sources of information - Continuous National Household Sample Survey (PNADC) and General Register of Employees and Unemployed (Caged) – suffered with methodological problems during this period.

The interviews for the PNADC had to change from in-person collection to telephone inquiries as of March 2020, a change that caused a significant drop in the number of interviewees - especially for those that would be responding for the first time and which are more likely to occupy formal positions (Corseuil and Russo (2021)). This could be one of the reasons for the significant drop in formal employment measured by PNADC as compared to Caged registries. At the same time, however, Caged itself underwent a methodological change in the beginning of 2020 which could have overestimated the number of new job openings, as the bankrupt and hibernating companies during the health crisis may have failed to report the dismissals of employees.

In this scenario, new data sources are becoming increasingly available through collaborations with private businesses engaged in commercial activities that record economic activity on a granular, frequent, and timely basis. New statistics, created using these nontraditional data sources, can complement existing government measurements and may enhance the ability of policymakers for measuring economic activity in real time. Thus, an alternative indicator for the formal labor market in Brazil is proposed in this study, using high frequency payroll payments from corporate clients of Itaú, the largest bank in Brazil, responsible for processing the payroll of roughly one hundred and forty thousand businesses representative of nearly eighteen percent[45] of Brazilian private-sector workers.

This paper proposes a monthly indicator for formal employment in Brazil based on payroll data from corporate clients of the largest bank in the country. The proposed measurement suggests that between 2018 and 2019 the evolution of this indicator is similar to official series, diverging, however, from official government indicators as of 2020. Itaú data indicator shows that formal employment did not fall as sharply as shown by PNADC data in the beginning of the Covid-19 pandemic, but the performance was slightly worse than Caged data. Also, it is shown that businesses of all sizes saw massive employment declines during the first few months of the current recession, especially those with less than 50 workers. This last analysis is not possible to assess using public data until RAIS is released, with nearly a one year delay.

Despite having been designed late in the the Covid-19 pandemic with a retrospective look on the beginning of this period, it is necessary to highlight that the proposed use of administrative data in this study may provide figures that are approximate to real-time indicators on the state of employment, which may represent an advantage in enabling faster policy responses. Private administrative data can also provide novel insights into employment by measuring variables more accurately than traditional survey data methods, or by capturing variables that may escape traditional data sources. In addition, the vast sample sizes available enable very detailed data segmentation and statistical precision, and can sometimes allow for new sources of variation and identification strategies. However, the use of administrative data may raise addi-

[45]Based on 2019's Relação Anual de Informações Sociais (RAIS), a database which consists of annual employment records that employers are required to report to the Ministry of Labor.

tional challenges and concerns of representativeness and external validity of such data, as those are often big concerns with private data. The greater statistical precision and quicker measurements made possible by this method and data do not necessarily equate with unbiased estimates, because large sample sizes do not solve, by themselves, identification problems.

This paper adds to the literature that uses administrative datasets to track real-time economic activity (see Vavra (2021) for comprehensive review in the U.S.). Several economists in the U.S. are turning to private-sector micro data to try to understand economic shocks while they are still unfolding, including data on employment based on payroll processing. For example, Cajner et al. (2018) were the first to use data from a payroll processor firm called ADP to track the U.S. workforce along many labor market dimensions. Using the same data during the early stages of the Covid-19 pandemic, Cajner et al. (2020) were able to quickly register a number of labor market facts that would not have been possible, at that moment, through official data. To date, this study represents the first initiative in Brazil that proposes the use of administrative data for this purpose[46].

The paper also contributes to the literature that tries to understand the evolution of the Brazilian labor market during the pandemic. Differing from the work of Corseuil et al. (2021), which tries to look for similarities with 2015's recession using the somehow biased PNADC, this work provides a perspective on the developments and status of employment by firm size in the country, with very recent data. Using official datasets, this analysis would only have been possible with almost one year lag, due to the periodicity of the publication of the data.

The paper is organized in three sections. First, the problems with PNADC and Caged series during the pandemics are described. In the second section, the aggregate employment indicator constructed using Itaú microdata is introduced, and the development and methodology used to create it are explained. In the third section, the usefulness of this indicator is explored to analyze the economic impacts of COVID-19 on employment in firms with perspective on its size. The last part concludes.

3.2. Formal Employment Measurements in Brazil

Figure 6 shows that, during the pandemic, a very significant difference was found between the two official labor market data sources commonly used to assess the evolution of formal employment in Brazil. Even considering methodological differences in the collection of data, represented population and periodicity (see Almeida et al. (2018); Corseuil et al. (2019) for more details), the evolution of formal employment computed by these two sources of information showed similar trends in the period between 2017 to 2019. The adherence between the two surveys, however, ceased from the beginning of 2020, when Caged started to report results considerably higher than PNADC for formal employment.

3.2.1. PNADC

PNADC is a quarterly household survey carried out by the Brazilian Institute of Geography and Statistics (IBGE) and its objective is to follow the evolution of the occupied and unemployed population (aggregate

[46]It is worth mentioning that confederations belonging to certain economic sectors regularly publish data on employment in their respective areas. However, those reports provide a limited view of their own sector, and are not representative on the state of employment in Brazil. As an example, the Brazilian National Confederation of Industry releases a monthly indicator of the sector, but with more than 30 days lag.

number of workers). IBGE estimates are based on surveying households (face-to-face) once a quarter for five consecutive quarters, with an interval of two months between interviews.

The pandemic brought challenges to IBGE regarding the interviews for the survey, forcing the institute to change its in-person collection of data to telephone inquiries from March 2020 until partially resuming in-person interviews in July 2021. As shown in Figure 7, the change caused a significant drop in the number of interviews carried out, reaching the lowest sample size since its beginning and, as first noted by Corseuil and Russo (2021), the drop was unevenly distributed across groups of individuals in regard to the number of times they had been interviewed. Figure 8 shows a significant drop in the number of respondents in all five groups from the second quarter of 2020 onwards compared to the respective quarter from 2019. The drop is more concentrated in the group of individuals that would be interviewed for the first time, reaching only 36% of the sample size in the second quarter of 2020 while the group that was being interviewed for the 5th time was 79%. Despite the total sample size differences, which remained if compared to before the pandemic, the distribution across those groups seems to have returned to a more balanced status in the last quarter of the available data (2Q21).

Corseuil and Russo (2021) show that the drop might have interfered in the evolution of formal employment indicators because of these changes in the composition of the sample in 2020. They note that the reduction in the sample of individuals who would be interviewed for the first time in the second quarter of 2020 may not have been neutral with regard to the composition of this group of individuals. In particular, that the differentiation may be related to characteristics associated with the probability of formal employment of the interviewee. Indeed, only in November 2021 did IBGE reweight the surveyed sample in order to correct the under representation of single-person male households, that are more prone to occupy a formal job.

Another evidence of the problem with the survey sample is shown in Figure 9 where PNAD data indicates the acceleration in growth of the working age population (population with more than 14 years old) since the beginning of the pandemic, which does not seem to be consistent with population growth dynamics in Brazil. The yearly pre-Covid growth of working age population trend was approximately 1%, having reached, however, 3% by the end of 2020.

Taking these evidences into consideration, it is of concern whether PNADC's series, due to its changes in collection, sample sizes and representativeness during the pandemic, was able to provide the correct evidence of the evolution of formal employment in Brazil for the period.

3.2.2. Caged

Caged is an administrative record of formal employment with monthly periodicity derived from compliance with labor obligations. The mandatory reporting of employment data is provided to government officials by establishments and lists the movements of employees throughout the month. Therefore, it represents the flow of workers entering and leaving formal employment in a given month.

In January 2020, there was a transition in the way companies provide registration information relating to their workers. This transition started in March 2018 for large private companies and began to reach micro and small companies in 2020. Adding to the change in data collection methodology, the emergence of the new coronavirus pandemic may have increased the difficulty in the transition by making it more complicated for companies to provide information correctly or in a timely manner.

Duque (2020) first raised suspicion that Caged was overestimating workers due to the bankruptcy and hibernation of companies during the health crisis, as those enterprises may have failed to report the dismissals of employees. Due to this concern, Caged has been revising its numbers and making their microdata available for the past 13 months, which has made it possible to calculate the difference between the first release versus current numbers. Figure 10 shows that 88% of companies that dismissed employees in March and in April of 2020 actually reported it on time compared to 90% that were hiring. In figure 11, the difference of net balance of workers flow is presented. Since the methodology change in January 2020 until August 2020 (the last available month with revisions released), the accumulated difference represented about 147 thousand dismissals underreported.

3.3. Itaú Data

Itaú's employment series is constructed based on payroll data measuring business level outcomes. Business records include both the number of paychecks issued in a given pay period and the total amount paid (regular paychecks, bonus checks or any other payment). The data encompasses more than 140,000 businesses across Brazil with nearly 18% of Brazilian private-sector workers. Since this study seeks to measure private sector employment, workers employed in public administration are excluded. Also, despite having daily data with a maximum delay of one day for all payments, the majority of companies pay workers on monthly basis in Brazil[47].] Total employment information is grouped for the main industries and by 2017's firm size bin defined by Eurostat[48] (e.g., 1-49 employees, 50-249 employees, >250 employees). It is not possible to break the data regionally, as most payments are made by the businesses headquarters accounts and do not necessarily match the place of work. Figure 12 shows that Itaú's payroll data appear to be quite representative across industries, but in Figure 13 data seems to modestly overrepresent large businesses. One hypothesis is that micro business might pay workers in cash or transfer from owners accounts and not corporate accounts[49]. Representativeness is ensured in terms of observables by reweighting the aggregate number of workers data to match RAIS employment shares by firm size.

The indicator constructed is reported as a yearly variation of total employment. The reason for this choice is that some steps are needed to prevent the change in the number of companies from artificially creating breaks in the employment series as a result of changes in market share. For each month, the total number of employees are calculated in t and recalculated in $t - 12$ with same companies, the newly created ones or the closed ones.

For new companies entering the database, the date of registration of the business available from Receita Federal (Brazilian IRS) is used to assess if it is a new job opening or not. If the company is more than one year old, it is considered as a gain on market share. Otherwise, it is considered a new company and its newly created job vacancy is included in the indicator. The company that is more than one year old will be incorporated in the aggregate employment after one year, when it is possible to compare the growth of employment for the same company.

The challenge is to assess the loss of market share, as companies take time to cancel their registration. In 2019, data from RAIS show that 15% of companies had no employees but had an active registry in Receita Federal. Since the study observes companies at t and $t - 12$, a few assumptions were made to separate

[47] Almost all employment series constructed in the country based on payroll data are generally available only with a lag, because people are paid monthly, after completing work over multiple prior weeks.

[48] Source: Eurostat Glossary on Enterprise size.

[49] A 2018 survey from the Central Bank of Brazil found that 37% of the sample employees received their wages in cash rather than bank transfer. However, the survey did not show the percentage of employees in the sample that were in the formal labor market. Link to the survey presentation.

enterprise accounts closed due to loss of bank market share and those that were closed as a result of a business terminating its activities. The study considers that a company was closed if in the past 12 months their payroll was decreasing, and the account was then terminated, disappearing from the dataset in the following months. If the company, however, reported a larger payroll or had no change in payroll in the past 12 months, indicating either growth or stability, and that was followed by a terminated account, that was considered as an indication of the loss of a client that took its payroll processing elsewhere, and thus a loss of market share. In this case, companies are not included in $t-12$ aggregate employment for the current month t.

To summarize, only the same companies are compared, but establishments are added if a company is new (less than 12 months old) or subtracted if it was closed (dismissing in the past 12 months). The indicator is shown in Figure 14 compared to national benchmarks.

The correlations between Itaú's employment growth rate indicator and official data indicators is 87% for PNADC's series and 90% with Caged, from 2018 until July 2021. The cyclical behavior between Itaú's series and the benchmark data are very much aligned until the Covid 19's pandemic, when both official indicators presented methodological problems. However, as of May of 2021, Caged's and Itaú's growth rates started to realign. It is important to highlight that PNADC's monthly data is a 3-month moving average. The higher correlation of Itaú's indicator with Caged might be due to the fact that both data sets monitor employment relationships, and a worker can have more than one employment relationship at a given time. Also, the slightly higher contrast with PNADC data may be attributed to a higher chance of measurement errors from self-reported data, since it consists of survey-based data.

The biggest advantage of the proposed indicator is the nearly real-time availability: within less than 5 days after the end of the previous month, compared to 3 to 6 weeks of official indicators. In periods of crisis, as the Covid-19 pandemic, it can play a crucial, early role in shaping our understanding of the consequences of the health crisis on employment and opens up the possibility for faster policy responses. Another important advantage is that employment can be measured with precision without the measurement error from series revision, as is currently happening in Brazil, or from self-reported data. In addition, the administrative data has advantages relative to traditional data sources, as it provides different lenses of the data with larger sample size and a panel element.

On the other side, a potentially significant concern is that the data poorly captures small employers, and does not encompass public-sector employment. It is also worth noting that it only measures employment declines once we observe a business's regularly scheduled payroll. This can mean that there is some lag in our measurement. For example, suppose a business pays all of its workers at the end of the month. If a worker is terminated in t, a payment will still be registered in t, and the loss of the employment link will only be registered in $t+1$. All of this is to say that the measurement may, at times, be shifted a month from when a dismissal actually took place.

3.3.1. Employment Changes in the Pandemic Recession

Much attention has been given to the preservation of small businesses and employment in the current recession. Small enterprises represent a large challenge for economic recovery, as they employ almost 40% of formal labor and are more financially constrained than its larger counterparts. In this sense, Figure 15 plots the change in employment by initial business size as of February 2020, an analysis that would not be possible to assess using public data until RAIS is released, with nearly a one year delay. The indicator plots estimated aggregate employment levels relative to February 2020. It is important to indicate that this panel shows employment changes that encompass continuing businesses, businesses that have shut down definitively and new businesses.

Businesses of all sizes saw massive employment declines during the first few months of the current recession. Businesses with fewer than 50 employees saw paid employment declines of almost 10 percent through July 2020, while those with more than 50 employees saw declines of less than 7 percent during that same time period. Notably, the employment rebounded first for companies with less than 250 workers in August 2020, while for larger companies it occurred in September 2020. Since then, larger and medium companies are recovering at similar pace and have already reached the aggregate level of employment as of February 2020, while small companies haven't returned to pre-pandemic levels. This result is different compared to U.S. dynamics. Cajner et al. (2020) shows that smaller business employment decreased as much as 15% in the U.S., but surpassed medium and larger business within 3 months of the beginning of the pandemic. Partially, this could be attributed to the greater flexibility of the country's labor legislation and government support.

The Brazilian government's stimulus packages that made special provisions to support medium and small businesses through a large expansion in credit was only signed in June 2020, with significant delay[50]. This late response might help explain why businesses with fewer than 50 employees reduced employment at a faster rate than their counterparts between March and July 2020. However, as early as April 2020 companies of all sizes were eligible for another government program that directly targeted the preservation of employment called Programa Emergencial de Manutenção do Emprego e da Renda (BEm). In this program, any registered company could reduce working hours up to 70% with a corresponding salary reduction in equal proportion. The government, in turn, would grant workers a financial compensation, observing limits established by the policy. Companies could also suspend employment contracts, maintaining a salary payment of 30% for the halted contracts.

While 25% of the formal private sector employment was enrolled in BEm, Guerrero et al. (2021) conducted a survey with small companies from different Latin American countries and found that, of all the programs made available to support small companies in Brazil, BEm received the smallest number of applications. The study points that programs based on credit were the most sought out by businesses. The authors also highlight that, in general, smaller firms were less aware of existing programs.

BEm BEm came into effect on April 1, 2020 and, with several successive renewals, it extended the maximum time workers could be enrolled on the program. Thus, it stayed in effect throughout 2020. In 2021, it was only valid from end of April to end of August.

The change in workload would last up to 240 days and the businesses were required to keep the worker employed for a period equal to that of the suspension or pay reduction after the end of the program's benefits. If the worker was fired before that, companies would be fined and had to refund the full amount of the compensation to the government. Companies could enroll any worker at any time during this period. That is, worker A could be enrolled in April and worker B in August. After the end of the period of first enrollment, the worker could then be re-enrolled.

The agreements could be signed individually or collectively, by association of workers. Thus, in an individual agreement, the company could dismiss part of the workforce, as long as these workers were not in the program, giving the company flexibility to dismiss part of the workforce and reduce/suspend the workload of those that remained.

[50] Measures for the (temporary) suspension of loan repayments were established in March 2020 by the Central Bank of Brazil.

What would have happened to employment without government support through BEm? This sort of analysis is complex, as all registered companies that applied were granted access to the program, and the self enrollment process creates a selection bias. It is still of value to inspect how employment in those companies that applied for the program is evolving. Figure 16 shows employment changes relative to February 2020 for firms that did not apply and for those that applied for the program, dividing these companies in different quartiles of intensity of working force enrollment. As the figure shows, companies that did not enroll in the program saw, comparatively, a lesser decline in employment, and have already surpassed pre-pandemic levels of employment as of September 2020. As shown in Table 16 this could be justified by a large share of companies in the construction sector, which was less impacted by the pandemic because of low real interest rates that heated the housing market.

Interestingly, considering the companies that applied to the program, the ones in the bottom quartile of enrollment of working force distribution (which enrolled less than 14% of its working force) and the ones in the top distribution (which enrolled more than 64% of its working force) have both reached, as of October 2020, the same level of employment registered in February 2020, despite having different trends. In the first case (bottom quartile), the difference might arise from a higher concentration of larger companies. In the second case (top quartile), because most of the workforce was enrolled, the obligation of maintaining those contracts after the end of the enrollment explains the results. The last round of applications to BEm in 2021 was made available from April through August, and thus there are still employment dynamics to be monitored for those companies. However, it seems that, for companies that first applied in April 2020, workers were not laid off after the binding period, despite the uneven recovery of employment.

3.4. Remarks

This paper introduces a set of aggregate employment changes derived from Itaú payroll microdata that provide high-frequency information on the labor market. The series takes payroll data at the monthly frequency and aggregates it up with a RAIS-based weighting scheme to arrive at nationally representative series.

The underlying data that are used to construct the employment indicator covers almost 20 percent of the Brazilian private workforce. Moreover, analyzing industry composition indicates that the Itaú data are reasonably representative of the overall economy. It is necessary, however, to highlight a possible bias in the data, as it underrepresents small business. The biggest advantage is that the proposed indicator is more timely than traditional data sources, allowing for an estimate of payroll employment within only a couple of days after the end of the month.

Methodological problems that official labor market indicators had during the pandemics are discussed and, as the proposed indicator tracked movements in CAGED and PNADC employment quite well, the proposed indicator may be viewed as a complementary, more readily available source of information for the labor market. It is stressed that administrative micro data should be viewed as a complement rather than substitute for traditional data sources.

Without representative data sources for benchmarking and validation, it is very difficult to interpret results from particular administrative data sources. Thus, significant uncertainty remains about the direct comparability between private payroll data from Itaú and the officially produced statistics. First, significant questions remain about the selection issues for firms that chose the bank as their payroll processing. Second, the indicator has to approximate changes in market-share.

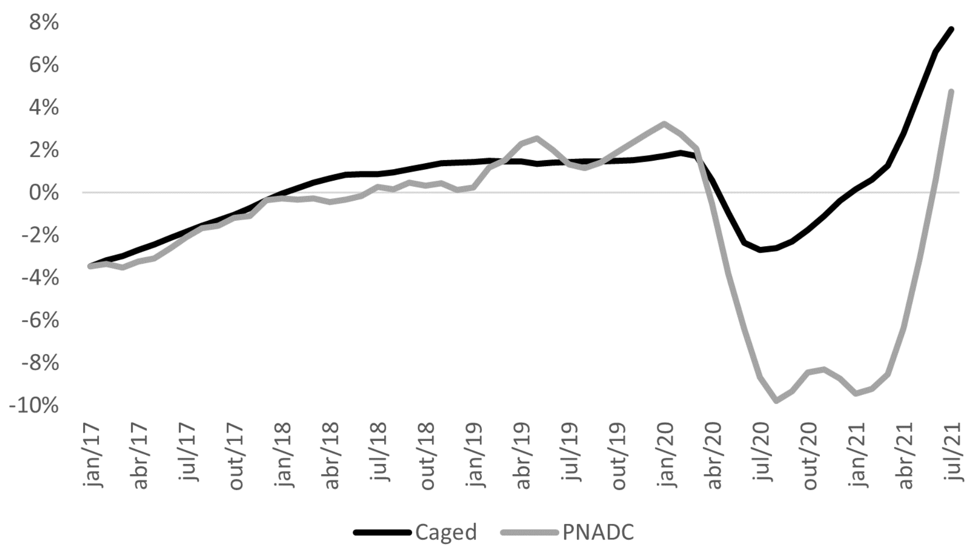
Hopefully, these concerns will be further addressed in future research to adequately represent the power of the administrative data. Moreover, there are broader questions that can be answered with the payroll microdata.

Table 15: Distribution in sector and size of companies that did not apply and the ones that applied for the Programa Emergencial de Manutenção do Emprego e da Renda (BEm), divided by quartiles of intensity of working force enrollment

	No Bem	Q1 (enrolled less)	Q2	Q3	Q4 (enrolled more)
Sector					
Agri	1%	1%	0%	0%	0%
Retail	24%	24%	29%	29%	29%
Construction	21%	7%	6%	4%	3%
Mining	0%	0%	0%	0%	0%
Manufacturing	15%	17%	15%	20%	23%
Utilities	2%	1%	1%	0%	1%
Service	37%	49%	48%	45%	43%
Size					
Small	87%	52%	80%	78%	85%
Medium	11%	32%	16%	17%	12%
Large	2%	16%	4%	5%	4%
Bin of working force enrollment	0%	]0%-14%[	[14%-36%[	[36%-64%[	[64%-100%]
Sample Size	41.586	7.004	7.004	7.004	7.004

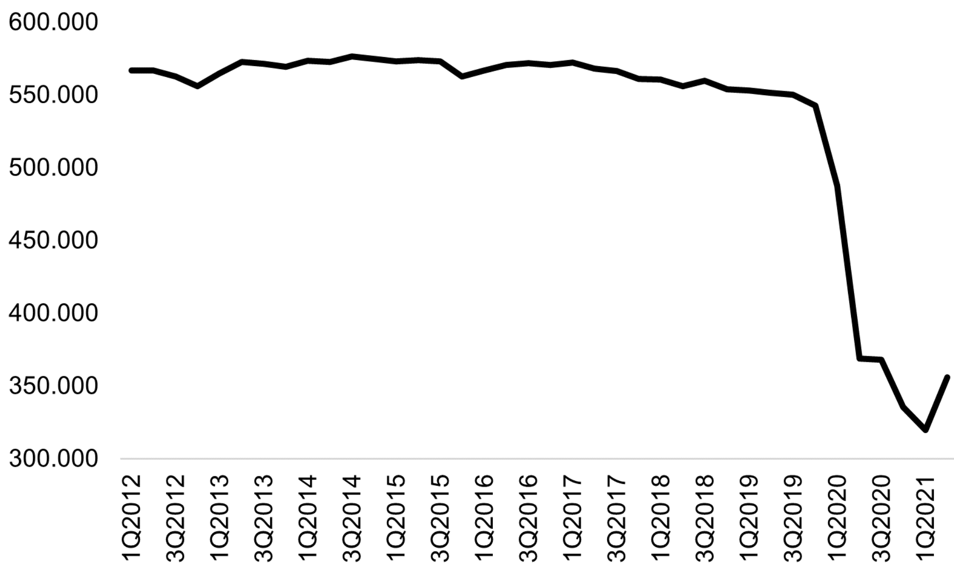
Note: Small: 1-49 workers; Medium: 50-249; Large: >250

Figure 6. Growth rate of employment from PNADC and Caged



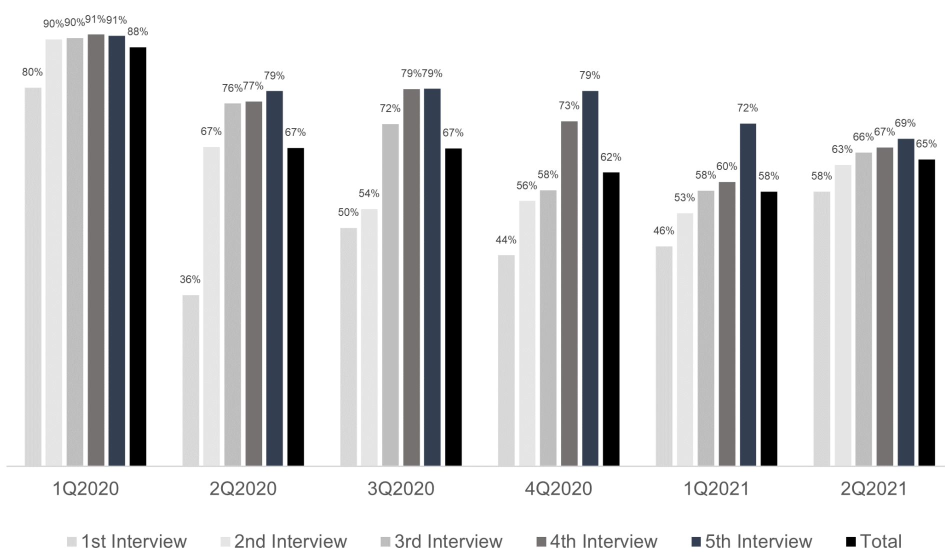
Note: This excludes Public Administration workers from Caged and the growth rate is adjusted based on total employees, as released on September 2021. PNADC contains only formal private workers and employers with registration in IRS. Both are 3-month moving averages.

Figure 7. Sample Size of PNADC survey



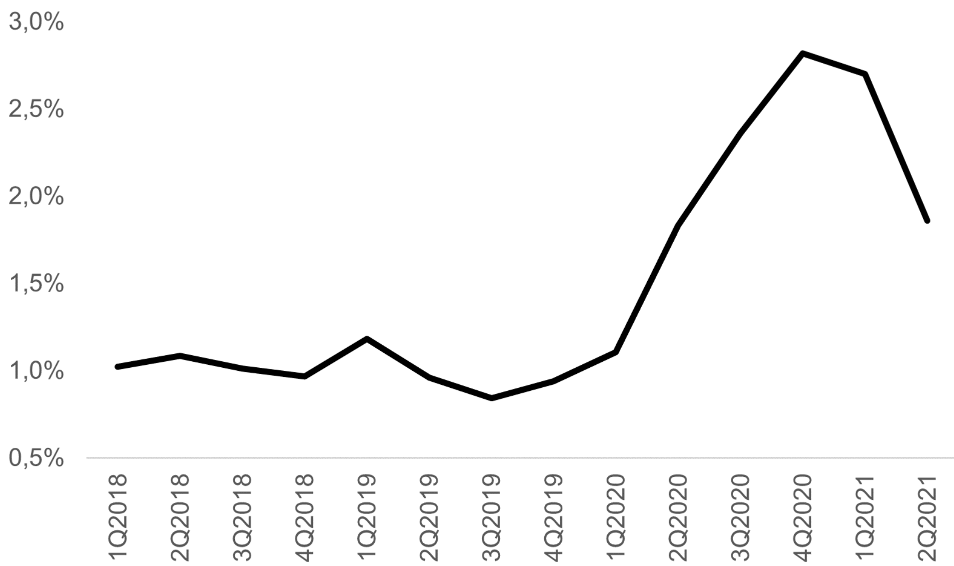
.

Figure 8. PNADC sample size relative to the same 2019 quarter



.

Figure 9. PNADC Working Age Population Growth

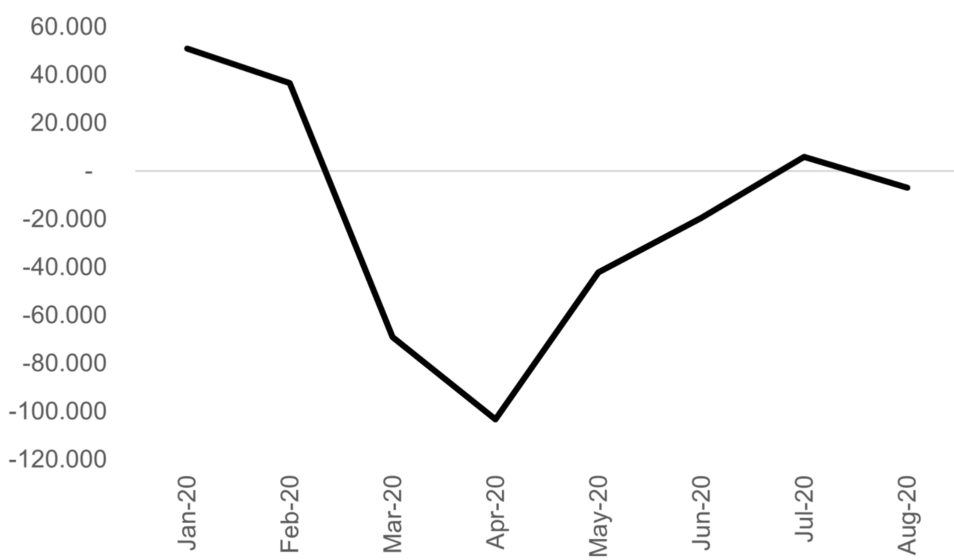


Note: Working age population is population with more than 14 years old

Figure 10. Number of companies reporting dismissal and hiring - First released over current

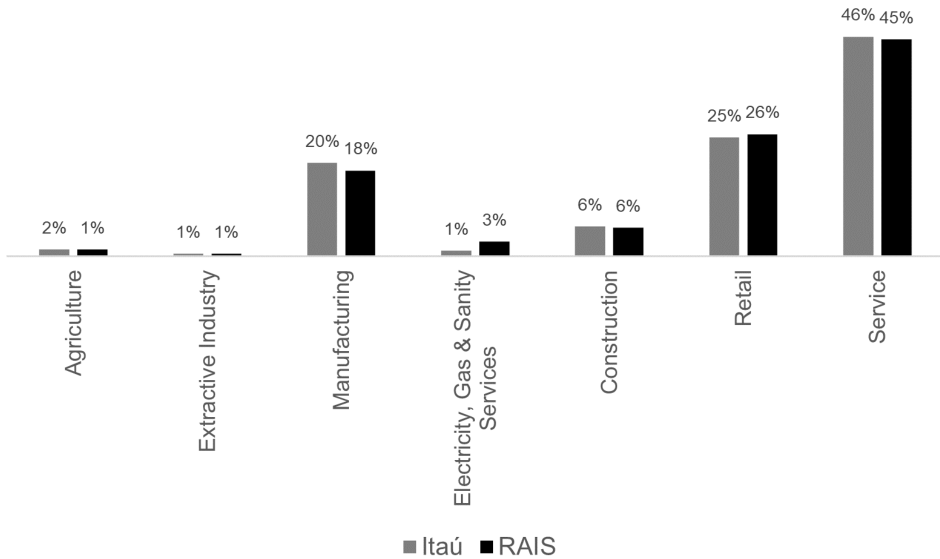
.

Figure 11. Caged difference between first released reports and actual figures, net balance of flow of workers



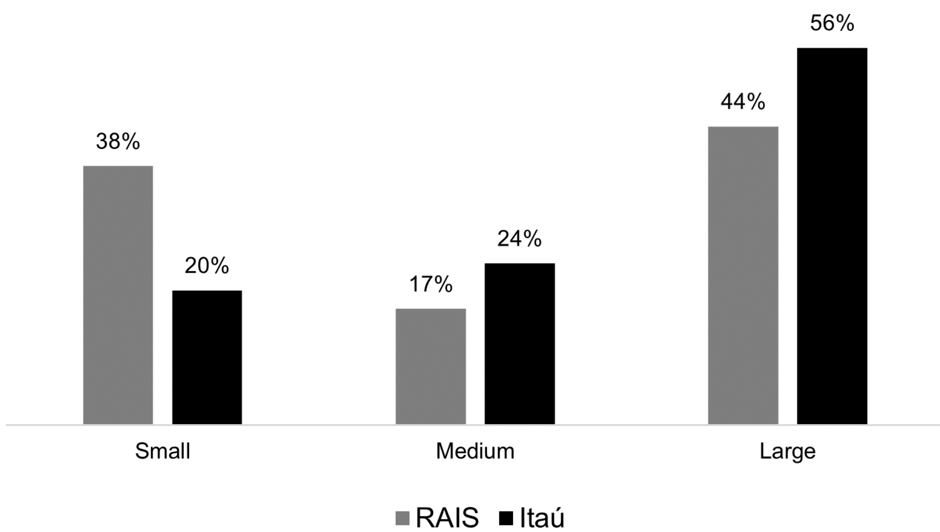
.

Figure 12. Share of Employment by Sector in 2017



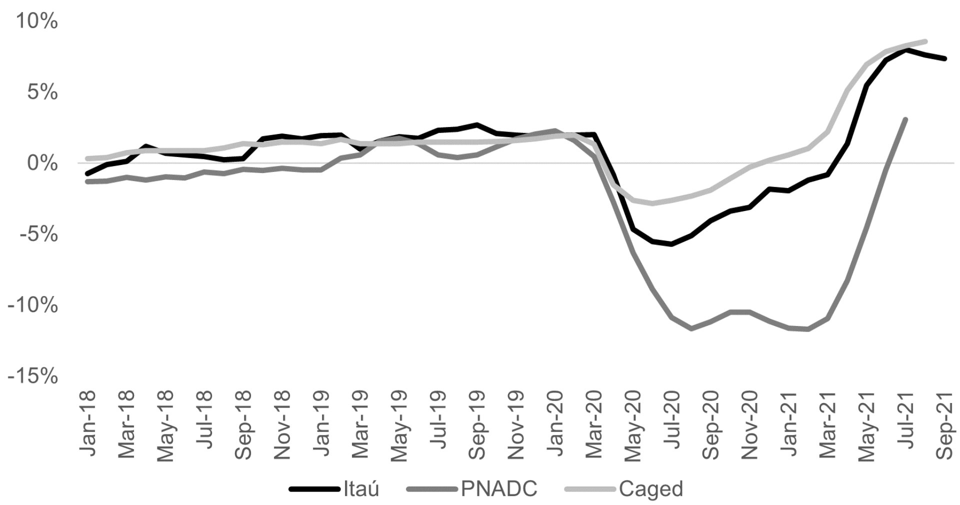
.

Figure 13. Share of Employment by Firm Size in 2017



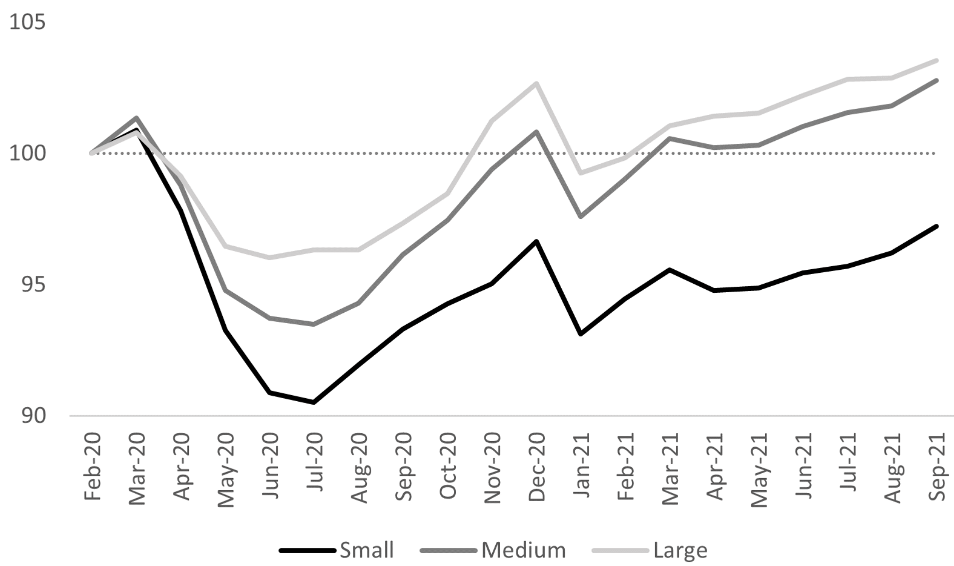
Note: Small: 1-49 workers; Medium: 50-249; Large: > 250

Figure 14. Growth rate of employment



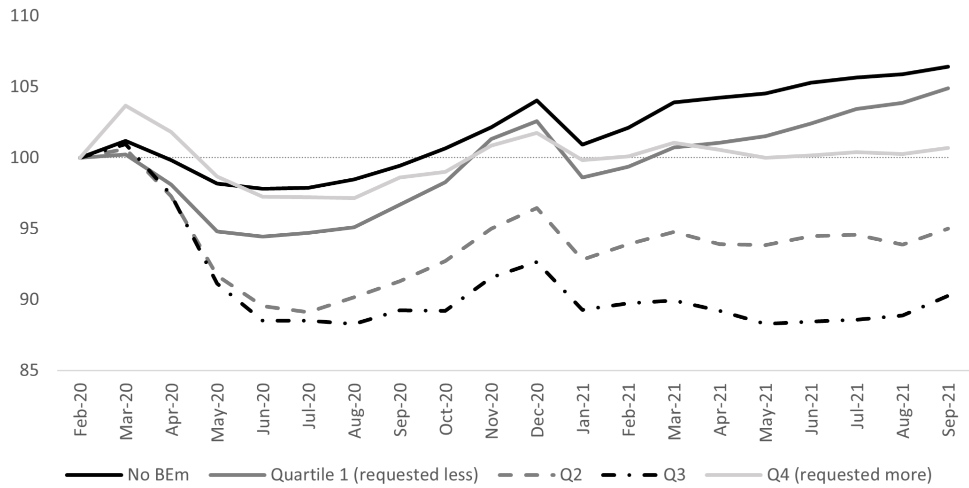
Note: This excludes Public Administration workers from Caged and the growth rate is adjusted based on total employees, as released on September 2021. PNADC contains only formal private workers and employers with registration at the IRS. Caged and Itaú's series are 3-month moving averages.

Figure 15. Aggregate Employment, by company size



Note: Figure shows employment changes relative to February-2020 within the Itaú business-level sample size. Small: 1-49 workers; Medium: 50-249; Large: > 250

Figure 16. Aggregate Employment, by quartile of intensity of working force enrollment at Emergency Job and Income Preservation Benefit (BEm)



Note: No BEm: companies that haven't enrolled in the program; Q1: companies that enrolled less than 14% of their working force; Q2: companies that enrolled between 14% and 36% of their working force; Q3: companies that enrolled between 36% and 64% of their working force; Q4: companies that enrolled more than 64% of their working force.

References

M. E. Almeida, M. d. P. Sousa, F. Vella, and A. V. S. M. Albuquerque. Comparando bases de dados: o caso do cadastro geral de empregados e desempregados (caged) e da pesquisa nacional por amostra de domicílios contínua (pnad contínua). *Boletim Mercado de Trabalho: conjuntura e análise*, 2018.

D. Anginer and A. Demirguc-Kunt. Bank Runs and Moral Hazard: A Review of Deposit Insurance. *Bank Runs and Moral Hazard: A Review of Deposit Insurance*, (September), 2018. doi: 10.1596/1813-9450-8589.

A. Antunes, T. Cavalcanti, and A. Villamil. Costly intermediation and consumption smoothing. *Economic Inquiry*, 51(1):459–472, 2013. ISSN 00952583. doi: 10.1111/j.1465-7295.2012.00471.x.

E. Baltensperger, J. Dermine, C. Goodhart, and J. Kay. Banking Deregulation in Europe. *Economic Policy*, 2(4):63, 1987. ISSN 02664658. doi: 10.2307/1344553.

P. Behr and S. Sonnekalb. The effect of information sharing between lenders on access to credit, cost of credit, and loan performance - Evidence from a credit registry introduction. *Journal of Banking and Finance*, 36(11):3017–3032, 2012. ISSN 03784266. doi: 10.1016/j.jbankfin.2012.07.007. URL `http://dx.doi.org/10.1016/j.jbankfin.2012.07.007`.

A. W. Boot. Relationship Banking: What Do We Know? *Journal of Financial Intermediation*, 9(1):7–25, 2000. ISSN 10429573. doi: 10.1006/jfin.2000.0282.

J. H. Boyd and G. D. E. Nicoló, O. The Theory of Bank Risk Taking. *Journal of Finance*, LX(3):1329–1343, 2005. ISSN 1540-6261.

T. Cajner, L. Crane, R. Decker, A. Hamins-Puertolas, C. Kurz, and T. Radler. Using payroll processor microdata to measure aggregate labor market activity. *Finance and Economics Discussion Series*, 2018 (005), 2018.

T. Cajner, L. D. Crane, R. A. Decker, J. Grigsby, A. Hamins-Puertolas, E. Hurst, C. Kurz, and A. Yildirmaz. The u.s. labor market during the beginning of the pandemic recession. Working Paper 27159, National Bureau of Economic Research, May 2020. URL `http://www.nber.org/papers/w27159`.

C. W. Calomiris and M. Jaremski. Deposit insurance: Theories and facts. *Annual Review of Financial Economics*, 8(1):97–120, 2016. doi: 10.1146/annurev-financial-111914-041923.

C. W. Calomiris and M. Jaremski. Stealing deposits: Deposit insurance, risk-taking, and the removal of market discipline in early 20th-century banks. *The Journal of Finance*, 74(2):711–754, 2019. doi: https://doi.org/10.1111/jofi.12753. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12753`.

M. K. Camara and F. Montes-Negret. Deposit insurance and banking reform in russia. Policy Research Working Paper Series 4056, The World Bank, 2006.

C. H. Corseuil, M. Franca, G. Padilha, L. Ramos, and F. Russo. Comportamento do mercado de trabalho brasileiro em duas recessões: análise do período 2015-2017 e da pandemia de covid-19. *Nota Técnica. Instituto de Pesquisa Econômica Aplicada (Ipea)*, 2021.

C. H. L. Corseuil and F. M. Russo. A redução no número de entrevistas na pnad contínua durante a pandemia e sua influência para a evolução do emprego formal. *Carta de Conjuntura. Instituto de Pesquisa Econômica Aplicada (Ipea)*, 2021.

C. H. L. Corseuil, L. Ramos, B. d. S. Azevedo, and F. M. Russo. Mais sobre as diferenças na evolução do emprego formal na pnad contínua e no caged. *Nota Técnica. Instituto de Pesquisa Econômica Aplicada (Ipea)*, 2019.

J. De Gregorio. Borrowing constraints, human capital accumulation, and growth. *Journal of Monetary Economics*, 37(1):49–71, 1996. ISSN 03043932. doi: 10.1016/0304-3932(95)01234-6.

A. de Janvry, C. McIntosh, and E. Sadoulet. The supply- and demand-side impacts of credit market information. *Journal of Development Economics*, 93(2):173–188, 2010. ISSN 03043878. doi: 10.1016/j.jdeveco.2009.09.008. URL `http://dx.doi.org/10.1016/j.jdeveco.2009.09.008`.

A. Demirguc-Kunt and L. Klapper. *Measuring financial inclusion: The global findex database.* The World Bank, 2012.

D. W. Diamond and P. H. Dybvig. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, 91(3):401–419, 1983. ISSN 1537534X. doi: 10.1086/261155.

D. Duque. Evidências da subnotificação de desligamentos do caged, Oct 2020. URL `https://blogdoibre.fgv.br/posts/evidencias-da-subnotificacao-de-desligamentos-do-caged#:~:text=Porqueaquedado,semreportá-losaogoverno`.

F. Foley, A. Hurtado, A. Liberman, A. Sepulveda, G. Maturana, A. Micco, S. Nelson, J. Tessada, and L. Zingales. The Effects of Information on Credit Market Competition: Evidence from Credit Cards. (October 2018):1–52, 2019. URL `http://pages.stern.nyu.edu/{~}aliberma/FHLS{_}Jan2019.pdf`.

M. Guerrero, J. Humphries, C. Neilson, N. Shimberg, and G. Ulyssea. Small firms and the pandemic: Evidence from latin america. *Journal of Development Economics*, Forthcoming, 2021.

J. H. Hahm and S. Lee. Economic effects of positive credit information sharing: The case of Korea. *Applied Economics*, 43(30):4879–4890, 2011. ISSN 00036846. doi: 10.1080/00036846.2010.498364.

T. F. Hellmann, K. C. Murdock, and J. E. Stiglitz. Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough? *American Economic Review*, 90(1):147–165, 2000. ISSN 00028282. doi: 10.1257/aer.90.1.147.

G. Joaquim. Bank Competition , Cost of Credit and Economic Activity : Evidence from Brazil Job Market Paper. 2019.

C. Matutes and X. Vives. Competition for deposits, fragility, and insurance. *Journal of Financial Intermediation*, 5(2):184–216, 1996. ISSN 10429573. doi: 10.1006/jfin.1996.0010.

L. I. Nakamura. Commercial bank information: implications for the structure of banking. Technical report, 1991.

H. L. Q. Nguyen. Are credit markets still local? Evidence from bank branch closings. *American Economic Journal: Applied Economics*, 11(1):1–32, 2019. ISSN 19457790. doi: 10.1257/app.20170543.

A. J. Padilla and M. Pagano. Sharing default information as a borrower discipline device. *European Economic Review*, 44(10):1951–1980, 2000. ISSN 00142921. doi: 10.1016/S0014-2921(00)00055-6.

M. Pagano and T. Jappelli. Information Sharing in Credit Markets. *The Journal of Finance*, 48(5):1693–1718, 1993. ISSN 15406261. doi: 10.1111/j.1540-6261.1993.tb05125.x.

A. Powell, N. Mylenko, M. Miller, and G. Majnoni. WPS3443. (December), 2003.

L. R. Reed, D. S. K. Rao, S. Rogers, C. Rivera, F. Diaz, S. Gailly, J. Marsden, and X. Sanchez. *Mapping Pathways out of Poverty*. 2015. ISBN 978-0-9965728-0-4. URL `https://commons.wikimedia.org/wiki/File:Snow-`.

M. A. Turner. The Consequences of Prohibiting Credit Inquiry Data in Chilean Credit Files. (March):1–25, 2010. URL `http://perc.net/files/ChileWhitePaper{_}web.pdf`.

J. Vavra. Tracking the pandemic in real time: Administrative micro data in business cycles enters the spotlight. *Journal of Economic Perspectives*, 35(3):47–66, 2021. doi: 10.1257/jep.35.3.47.

X. Vives. Regulatory reform in European banking. *European Economic Review*, 35(2-3):505–515, 1991. ISSN 00142921. doi: 10.1016/0014-2921(91)90152-9.