The impact of loan-loss-provision regulation on credit: Evidence from administrative data in Chile*

Mauricio Calani

Manuel Paillacar

First version: November 2017.

This version: August 2019

Abstract

In January 2016 the Chilean banking supervisor raised required loan-loss-provisions (LLP) for mortgage credit risk non-uniformly, arguing in favor of its prudential nature. How was the mortgage market affected by the introduction of this prudential policy tool? We conclude that the loan-to-value (LTV) ratio was 2.8% lower for the mean borrower, and 9.8% lower for the median borrower, because of the regulation. We reach this conclusion by developing a stylized imperfect information model that we use to guide our empirical analysis of administrative data. We argue that financial institutions responded by raising their acceptable borrowing standards to borrowers, i.e. lower loan-to-value ratios –contracting their supply of mortgage credit–, rather than raising interest rates. Our paper contributes to the literature on the evaluation of macro-prudential policies, which has mainly exploited cross-country macro data. In turn, our analysis narrows down to one particular policy in the mortgage market and dissects its effects by exploiting unique administrative tax data on the census of all real estate transactions in Chile, together with administrative data on mortgage credit operations.

Keywords: Loan loss provisioning, loan-to-value, screening, matching, bunching estimation, regression discontinuity, macroprudential policy evaluation

JEL Codes: G21, R31

---

*Both authors are affiliated with the Central Bank of Chile, 1180 Agustinas St. Santiago, Chile. Calani: `mmcalani@gmail.com`, Paillacar: `mpaillacar.almonacid@gmail.com`. We thank useful discussion with Rodrigo Alfaro, Solange Berstein, Ken Kuttner and participants at the 2018 BIS-CCA Research Conference in Rio de Janeiro. This research did not receive any specific grant from funding agencies in the public, commercial, or non-profit sectors.

1 Introduction

On December 30, 2014, the Chilean Banking Regulator ("Superintendencia de Bancos e Instituciones Financieras", and SBIF, henceforth) announced that starting January 2016, it would enforce a new regulation on provisioning against credit risk, stemming from mortgage loans portfolio.[1] Before this regulatory change, banks would use their models and decide on their provisions. However, the view of the regulator was that these provisions were insufficient. Starting in January 2016, the SBIF requires to effectively raise financial provisioning for *each* granted loan. But more importantly, this requirement varies over the maturity of a loan and is contingent on realized delinquency of the borrower, and borrowers' leverage at the moment entering said delinquency. The chosen measure of borrower's leverage is the loan-to-value of collateral (LTV) ratio. This new (or rather modified) regulation implies substantially higher financial cost for banks if compared to observed pre-regulation provisions.

Did the new regulation affect the mortgage credit market? and if it did, what aspects and through which mechanism exactly? In this paper, we attempt to address these questions by using a two-step analysis. First, we analyze the features of the regulation using an off-the-shelve screening-under-imperfect-information model and adapt it to the problem at hand. Equipped with a model, we can learn about the properties of equilibrium under the new regulation, and grasp a sense of the effects under a wide family of parameters. In particular, we argue that to reduce the expected financial cost of the new regulation, banks tried to grant loans only to borrowers who were less likely to enter into delinquency; and therefore would entail less provisioning *ex-post*. But cherry-picking these borrowers is hard from an *ex-ante* perspective, so banks had to do this using a noisy signal; the LTV ratio. This model can produce an endogenous threshold for the signal (LTV limit) which we later document in the data. The second step in our analysis is empirical. We use a unique administrative dataset from the Chilean Internal Revenue Service (Servicio de Impuestos Internos, or SII) that records all nation-wide real estate transactions from 2002 onwards. In this dataset, we can observe transactional variables such as the property price, downpayments, and the financial institution involved in the mortgage loan. We can also observe the characteristics of buyers and sellers, such as income, or if any party is a firm. Lastly, we can observe many features of real estates, such as size, type, and location. This data is unique, and to the best of our knowledge comparable data has only been gathered and used in IL, USA by Ben-David (2011) to analyze inflated house prices in the years before the International Financial Crisis. Besides, we complement the information on property

[1] We refer to regulation "Provisiones por Riesgo de Crédito para Bancos", in Chapter B-1 of Compendium of Accounting Standards, SBIF, Chile. A friendly explanation can be found here. Other related material can be found here.

transactions with administrative data related to loan contracts collected by the SBIF. This dataset includes information about contract-specific features of all commercial, consumer and mortgage loans granted in Chile from 2012 to date. There we can find information such as lending institutions, loan amount, term and interest rates. In this paper, we analyze the before / after of the new regulation by exploiting the above described administrative micro data through the coarsened exact matching method by Iacus, King and Porro (2012), the bunching estimation techniques by Saez (2010) and Chetty et al. (2011) and the RD design methodology originally introduced by Thistlethwaite and Campbell (1960). We can use such hungry-data methods because of the richness of our data.

Our main findings are: (i) the new regulation had an effect on loan-to-value ratios for new loans: fewer loans with lower LTV ratios were granted. We estimate that, because of the regulation, the LTV ratio is 2.8% lower on average. Furthermore, the median borrower is granted a 9.8% lower LTV. We also find that, because of the way the regulation differentiates provisioning below and above 80% of the loan-to-value ratio, a large fraction of loans are granted at exactly that LTV. In particular, we calculate that the fraction of loans granted at 80% LTV more than tripled and represented one-fourth of all loans in 2016-17. This agglomeration effect is predicted by our stylized model. Finally, we use our model to rationalize the reason why higher financial costs were not off-loaded onto costumers, via higher mortgage rates. We argue that such an outcome is an equilibrium outcome stemming from the combination of imperfect information and competition between banks.

The rest of the section is devoted to placing our contribution within the related literature, explaining in detail the exact change in regulation and the data. Section 2 presents the stylized model, and section 3 develops our empirical examinations of the data. Finally 4 concludes.

1.1 Related literature and our contribution

The new regulation on loan loss provisions for mortgage credit was not introduced explicitly as a macroprudential tool, though one of its explicit objectives was "to promote active credit risk management" by financial institutions (Pacheco, Pugar and Valdebenito, 2014). Thus, in practice, it relates to the myriad of macroprudential tools used to deal with excessive credit booms. In particular, under the definition of macroprudential tools by Cerutti, Claessens and Laeven (2017), provisions are similar to capital requirements, which are considered fully-fledged macro-prudential tools[2]. Then, this paper joins the literature evaluating the effect of macro-prudential tools on different aspects of the credit markets.

There is robust cross-country evidence on the effects of the introduction of macro-prudential policies on housing markets. For instance, Crowe et al. (2013), Hott (2015), Cerutti, Dagher and Dell'Ariccia (2017) and IMF (2011) discuss the policy options to cope with real estate booms and stress the importance of LTV limits for subduing increasing leverage of households, preventing negative home equity, as well as limiting the number of borrowers who access mortgages and fuel real estate booms. Cerutti, Claessens and Laeven (2017) also takes a cross country perspective to study the effectiveness of the macro-prudential policy menu. From their analysis, we learn that LTV limits are important for the dynamics of mortgage loans, house prices, and overall financial fragility. In turn, Kuttner and Shim (2016) raise the issue of complementarity, and find that LTV and debt-to-income measures, together, are more effective in taming house price booms, than each on their own. From Qi and Yang (2009) we learn that LTV limits are not only important to prevent default, but that LTV is the single most important determinant of loan loss, given default. Country-based cases have also been studied. We contribute to this literature by exploiting administrative data –instead of cross-country data– and argue that the richness of our data coupled with our identification strategies allow us to single out the causal effect of one particular macro-prudential policy, in a given country.

More broadly, we contribute to an extending group of papers that uses (micro) administrative data to address macro-financial questions. This avenue has proven to be very rewarding for many strands of the literature, and particularly for analyzing the housing and mortgage markets: For instance, Albanesi, De Giorgi and Nosal (2017) use administrative credit file data for the U.S., to examine the evolution of household debt and defaults between 1999 and 2013. They find a new narrative at odds with the role of sub-prime borrowers in the crisis and find instead, that credit growth between 2001 and 2007 –and later mortgage defaults– were concentrated in the prime segment, mostly among real estate investors. Beltratti, Benetton and Gavazza (2017), use Italian administrative data to evaluate the effects on mortgage credit of the elimination of pre-payment penalties of mortgage loans. Similarly, Ben-David (2011) uses transaction data from a county in Illinois to examine the possibility of inflated house prices, and their use by financially constrained households. More related to our work, papers that have evaluated the effectiveness of macroprudential policies within a specific country, are scarce; perhaps due to the evident difficulty in ac-

[2]Notably, in their paper, Cerutti, Claessens and Laeven (2017) define five groups of different macro-prudential tools: (a) quantitative restrictions on borrowers, (b) capital and provisioning requirements, (c) quantitative restrictions on banks' balance sheets, (d) taxation, (e) accounting and compensation rules on credit origination. Only the first one would be a *demand* side policy. In this paper, in particular, we will show that the line dividing (a) and (c) will become diffuse, and credit rationing on the supply side will look like a quantitative restriction on the borrowers through endogenous limits on loan-to-value ratios.

cessing the necessary data. A few notable exceptions are Kinghan, Lyons and McCarthy (2017) and Acharya et al. (2018), who study the transmission mechanism of macroprudential policies on the mortgage market, to bank lending, using loan-level microdata from the five largest Irish banks. They stress the re-allocation effect between different types of borrowers. Another notably related paper which; which like ours, focuses on the evaluation of macroprudential policies using administrative data is Epure et al. (2018). These authors use household credit register to study the effectiveness of macroprudential policies on household lending, and how these can mitigate spillovers from the global financial cycle in Romania. In turn, our paper focuses on one specific macro-prudential policy, and the time around the policy change; thus allowing us to argue that changes in lending are due to this specific policy change.

1.2 The new regulation on loan loss provision in the Chilean mortgage market

The change in regulation by the banking authority (SBIF) is a (non-explicit) macro-prudential measure toward making consistent the expected probability of loss due to delinquency (credit risk), with accounting provisions. It was formally announced in December 2014 and entered into force in January 2016. Before this change in regulation, banks decided on their level of provisioning following internal models. Banks were often supervised by the regulator when the latter assessed these provisions as being too low, or notably different to the rest of the system[3]. The change in regulation removes discretionary provision accumulation. The most important features being:

1. **Timing:** Provisions have to be calculated monthly, and not only at origination.
2. **Loans included:** All outstanding loans are included, not just new loans.
3. **Size:** The required loan loss provision for a delinquent borrower is hefty and can go as high as 30% of the outstanding loan. Furthermore, a borrower can be re-labeled as non-delinquent only after he has paid all debt in arrears on time for four consecutive months. During this period, the bank needs to keep provisions unchanged.
4. **Contingencies:** Provisions are explicit functions of (a) time in delinquency, and (b) the LTV ratio. No formal LTV limit exists, but LTV is important because it interacts with time in delinquency to determine the size of the required provisions. Figure (1) shows this complementarity.

[3]For a very detailed exposition of the evolution of provisioning due to credit risk in Chile in the last three decades, see Matus (2015).

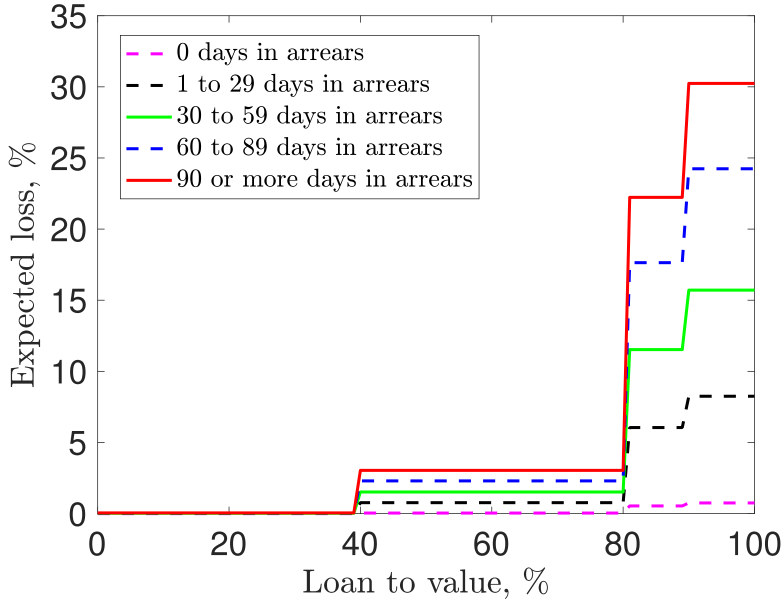


Figure 1: **Financial provisions under new regulation**: Expected loss (vertical axis, in percentage), according to Loan to Value ratio (horizontal axis), and days in arrears at the end of the month. Source: SBIF Chaper B-1 in "Compendio de Normas Contables"

1.3 Data

This paper exploits novel and unique administrative records from the Internal Revenue Service (IRS), for all real estate transactions in Chile, spanning 2002 to 2016[4]. Every real estate transaction in Chilean territory needs to be filed in the presence of a notary of faith ("Notario de Fe" in Spanish) who later submits all details of the transaction to a centralized archive of properties called "Conservador de Bienes Raices". Both, the notary and the archive, are obliged to inform the IRS using the "Declaration on Alienation and Registration of Real Estate" form (colloquially known as "Form F-2890")[5]. Currently, this dataset is used in the computation of the Housing Price Index by the Central Bank of Chile (Banco Central de Chile, 2017). The information contained in the F-2890 form includes the price of the property, mortgage loans, cash down-payments, name of the lender financial institution, and whether the buyer/seller is a person or a company. It also collects information on the identity of the buyer/seller, though this last piece of information is kept confidential. Combined with the Non-Farming Real Estate Property Cadastre ("Catastro de Propiedades no Agrícolas", also collected by the IRS) it is also possible to observe characteristics of the real estate in the transaction. In particular, whether it is residential or commercial property;

[4]Access to this data has been possible due to a Cooperation Agreement between the Central Bank of Chile and the IRS, signed in 2013

[5]This is in virtue of exempt resolution Nº8655 of December 27, 1999. More details to be found here.

a house, an apartment, a parking lot, or storage facility; its size, and age. We restrict our analysis to residential properties –houses and apartments– with some kind of mortgage financing.

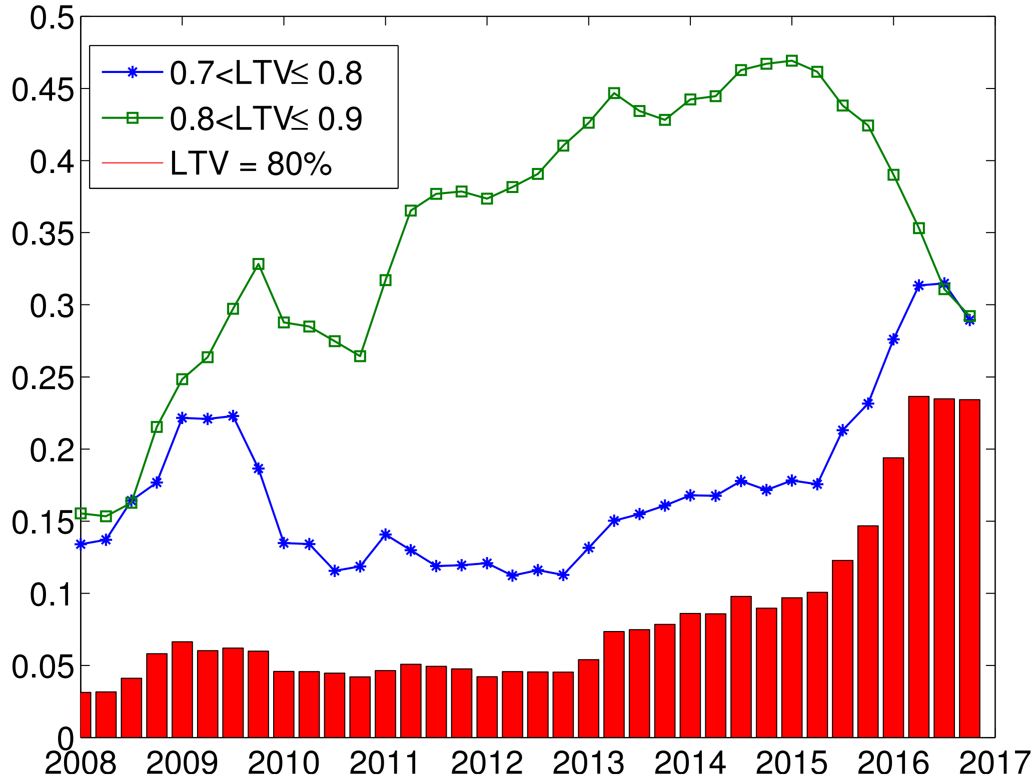


Figure 2: **Fraction of loans given at different LTV ratios.** The green line shows the fraction of loans given by all banks with loan-to-value ratios greater than 80 percent and lower or equal than 90 percent. The blue line does the same for loans with loan-to-value ratios higher than 70 percent and lower or equal than 80 percent. The red bars are the fraction of loans higher than 79.8 and lower than 80.2 percent. Quarterly averages. Source: Own calculation based on data from the IRS.

A quick examination of the data portraits one of the main arguments of this paper. There is a substantial difference in the distribution of LTV ratios before and after the introduction of the new regulation. The LTV ratio for the median borrower declined from 88% in 2014 to 80% in 2016. Of course, this decline cannot be directly attributed to the regulation without further examination of other covariates but provides a sense of relevance. The one other episode in recent memory where such a decline was observed coincides with the aftermath of the International Financial Crisis and the following recession. A different way to approach the same data is to consider the kinks of the regulation. In particular, from Figure (1) we can see the expected loss –and therefore the provisioning costs– considered under the new regulation depends positively on both, the LTV ratio, and on the number of days in delinquency. This relation is highly non-linear. In particular, the difference in provisioning between a non-delinquent credit and a delinquent one is negligible when the LTV ratio is below 80 percent but is very large when the LTV ratio is equal or above 80 percent. The 80 percent threshold represents a discontinuity which will prove key in the analysis. In Figure (2) we can see that after the announcement of the new regulation in December 2014, the fraction of loans with LTV lower or equal than 80% raised steadily (blue line) in detriment of the fraction of loans with LTV higher than 80% (green line). More importantly, the fraction of loans granted at exactly 80% grew very fast after December 2014, unlike any previous episode in the near past.

In the following section, we relate the cost of provisioning and all the contingencies specified in the regulation to an endogenous LTV limit in a model of financial screening. We parameterize the model to gain some insight into the quantitative effects of the new regulation, on the variable of interest. In the next section, we focus on the empirical counterpart.

2 The New Regulation under the Lens of a Simple Model of Financial Screening

2.1 Benchmark Model Setup

In the previous section, we elaborated on how the new regulation adds a non-negligible (expected) cost contingent on two conditions. First, after a mortgage is granted, the borrower goes into arrears. And second, that said borrower's mortgage debt represents a large fraction of the pledged collateral (high loan-to-value ratio). Furthermore, for this second condition, the regulation is highly non-linear around the 80% threshold (see figure 1). At this point, it is important to stress that the regulation does not legally impose a cap on LTV, but only disincentives granting new loans with high LTVs to low-quality borrowers. If banks were able to perfectly observe borrowers' quality, they would refrain from granting loans to those who will later become costly, or immediately offset this higher cost onto them. However, banks cannot separate high from low-quality borrowers *ex ante*. There is an incomplete information problem from the perspective of the lender.

In the rest of this section, we assess the problem of the financial intermediary using a benchmark model of imperfect information with screening. We do so because this model allows us to understand why it is that we care about LTV ratios; why the *ex-post* distribution of LTVs concentrates probability mass at exactly 80% of LTV; and why we should expect pass-through of higher financial costs onto mortgage rates, be very limited. Our small model below builds on the canonical models of imperfect information presented in

Stiglitz and Weiss (1981), Mas-Colell, Whinston and Green (1995), and some features of the application by Ates and Saffie (2013)..

2.1.1 Borrowers Heterogeneity

Every period a mass of size one of new borrowers shows up at the bank asking for a loan to purchase a house. These borrowers are indexed by $e \in [0, 1]$. Every one of them has an unobservable idiosyncratic probability $\theta(e)$ of being a (high) H-type borrower, and $1 - \theta(e)$ of being a (low) L-type borrower. H-type borrowers never enter delinquency, and therefore, never meet one of the two contingencies under which the provisioning cost is higher. L-type borrowers, on the other hand, have a positive and constant probability δ of entering delinquency at every given period. If $\theta(e)$ is non-decreasing in e, then the higher e, the higher the chances of the borrower of being H-type. In a way, then, e is the idiosyncratic quality ranking of borrowers. Note this is not a model of hidden action –which would raise moral hazard considerations–, or hidden information –which would bring along adverse selection–. This is a model of imperfect information. Borrowers know their quality ranking index e, but do not know their final type (H or L) for certain until after a mortgage is granted. Even more, they cannot credibly communicate their quality ranking e, and instead can do so only up to a noisy signal, $\tilde{e} \propto e$, which the financial intermediary can use to determine if it should grant the mortgage loan.

We will assume throughout that $\theta(e) = e^{\nu}$, with $\nu > 1$. Note that if $\nu < 1$, $\theta(e)$ is a concave function of e, which implies that H-type borrowers are relatively more abundant. On the other hand, if $\nu > 1$, H-type borrowers are relatively more scarce; meaning that high probabilities of being a good payer can only be achieved with values of e close to 1[6]. Put differently, ν governs the scarcity of H-type borrowers, and while it is a constant parameter in this model, nothing stops it from being countercyclical.

2.1.2 The value of lending to ex-post heterogeneous borrowers

Let us elaborate on the value of lending to an H(L) type borrower from the perspective of the lender. The financial intermediary is assumed to be exactly that; an intermediary who borrows funds at rate r_t from a deep-pocketed investor, and lends the proceeds to mortgage borrowers at rate $\hat{r}_t > r_t$. For simplicity let us assume that the financial intermediary only lends on perpetuity. We also assume that a full default is not a possible event. This assumption buys simplicity, but also allows us to put the emphasis on the effects of the new regulation, i.e. that the higher cost of lending to an L-type borrower comes from the

[6]It is possible to characterize the probability distribution $f(\theta)$ by $f(\theta) = \frac{1}{\nu}\left(\frac{1}{\theta}\right)^{1-\frac{1}{\nu}}$, with $\mathbb{E}(\theta) = \frac{1}{1+\nu}$.

financial burden of continuously provisioning a fraction of the loan in distress[7]. Recall then, that H-type borrowers are those who will not enter into arrears, and the value of lending to one of the said borrowers is given by the flow of period earnings derived from the lending/funding interest rate spread times the loan size, L_t,

$$V^H(L_t) = (\hat{r}_t - r_t)\, L_t + \frac{1}{1+r} V^H(L_{t+1}) \tag{1}$$

Also, note that under the assumption that the mortgage is to perpetuity, the loan amount L remains constant. Hence,

$$V^H(L) = \left(\frac{1+r}{r}\right)(\hat{r} - r)L$$

Analogously, the value of lending to an L-type borrower is similar to (1), except that there is a probability δ that borrower will enter into arrears, and trigger the cost of provisioning for a non-negligible period before they go back into good standing. Hence the value of lending to an L-type borrower includes this cost,

$$V^L(L_t) = (\hat{r}_t - r_t - r_t\delta\psi)L_t + \frac{1}{1+r} V^L(L_{t+1}) \tag{2}$$

with ψL_t the associated provision the bank has to make in such contingency. Again, because of the perpetuity assumption, we have that:

$$V^L(L) = \left(\frac{1+r}{r}\right)(\hat{r} - r - r\delta\psi)L$$

Note that *ex-ante* both types of borrowers are indistinguishable. It is only after the loan is granted that the borrower learns her type. Clearly, from the perspective of the lender, it is better to *ex-post* lend to an H-type borrower, and the difference in values is:

$$\Delta(L) = V^H(L) - V^L(L) = (1+r)\delta\psi L, \tag{3}$$

which, under the new regulation on provisions for mortgage loans, is positive ($\Delta > 0$). This implies that if the financial intermediary could observe a signal that points to a higher probability that the borrower will end up being H-type, then it should choose such borrower over another. In particular, under perfect information, the financial intermediary

[7] It can (correctly) be pointed out that banks could liquidate the house pledged as collateral in order to recover the capital lent to a defaulting borrower. While this is true, in practice, it is very uncommon. First, the Chilean case is one of full recourse. If a household defaults entirely on their debt, the bank can liquidate the house, other assets and could potentially go after earned income. This feature makes mortgage default an extremely rare event. Second, foreclosure is not only costly but takes a long time (more than 30 months until final liquidation); during which most borrowers go back into good shape.

would like to lend to costumers with higher quality ranking e, but it can only observe such statistic up to a noisy signal $\tilde{e}$. We elaborate next on this information friction.

2.1.3 The Signal

The bank knows that borrowers' ability to honor their commitments is related to many factors. Some of which are: financial education; household size; income volatility; the value of pledged collateral; total financial burden; to name a few. The two latter are efficiently summarized in two known statistics; the LTV ratio, and the debt-service-to-income (DSTI) ratio. For reasons elaborated above, the LTV ratio is by and large the most reliable signal in this regard. On top of that, while many developed economies extensively use credit scoring to separate high from low-quality borrowers; that is not the case in Chile. Information on debt in arrears is collected by the banking supervisor, but it is not publicly available to lending institutions in real time[8]. They must rely on own credit risk analysis from the information they request from the borrower. In practice, DSTI and LTV are used to allocate scarce credit funding, with LTV being the most frequently binding constraint[9].

Let us assume then, that borrowers' quality ranking can be imperfectly observed thorough the complement of the LTV ratio, $\tilde{e} = 1 - LTV$. That is, the down-payment; how much skin the borrower is willing to put in the game. A second interpretation is that (all else the same) higher savings at the moment of dwelling purchase point towards higher inter-temporal discount factor, and higher propensity to save. Thus, the higher the down-payment, the stronger the signal of the commitment of the borrower to honor their obligations. Let $\tilde{e} \in [0,1]$ stand for the noisy signal that is related to the true e quality ranking in the following way,

$$\tilde{e} = \begin{cases} e & \text{with probability} \quad \rho \\ \sim U[0,1] & \text{with probability} \quad 1-\rho \end{cases} \tag{4}$$

where ρ is the bank's screening technology accuracy; meaning that if screening works accurately (with probability $\rho = 1$), we have that lower LTV is signal of a borrower with higher e, and higher probability $\theta(e) = e^{\nu}$ of being H-type. On the other hand, with probability $1-\rho$ we have that the observed signal $\tilde{e}$ is simply noise. Even though the signal is

[8] Nonetheless, information of borrowers with recent default history is collected and sold by Equifax - Dicom, as long as the lender who was defaulted on reports such information. All in all, this information only gathers the very left of the distribution of borrowers' quality.

[9] Another reason why the LTV ratio is more widely used is that it is possible to extend the maturity of the mortgage contract and lower the debt service to income in any given period, it is not possible to do the same with the LTV.

imperfect, as long as $\rho > 0$, the signal is positively correlated to the true borrower's quality ranking; and therefore the optimal policy for the financial intermediary is then to set a cut-off threshold $\bar{e}$ on the realizations of $\tilde{e}$. This cut-off rule will have two effects on the rationing of credit. First, the extensive margin is affected as a more restrictive cut-off rule implies less acceptable borrowers. And second, the intensive margin is affected because on average borrowers (including H-type borrowers) are granted smaller loans, creating a trade-off.

2.1.4 The problem of the financial intermediary

Given the definition of the signal $\tilde{e}$, we can express loans in terms of this signal; $L = (1 - \tilde{e})P$. In the same way the value of lending to an H-type borrower, $V^H(\tilde{e})$, and to an L-type borrower, $V^L(\tilde{e})$, can also be written in terms of $\tilde{e}$. The problem of the financial intermediary is then: Given prices $\{r_t, \hat{r}_t, P_t\}$, the constant probability of entering into arrears for L-type borrowers, δ, and the provision required by the regulator (in percentage), ψ; the problem of the financial intermediary is to choose threshold $\bar{e}$ to solve the following program,

$$\pi(\bar{e}) = \max_{\bar{e}_t} \int_0^1 \int_0^1 \mathbb{1}\{\tilde{e}_t \geq \bar{e}_t | e_t\} \left[\theta(e_t)V^H(\tilde{e}_t) + (1 - \theta(e_t))V^L(\tilde{e}_t)\right] d\tilde{e}_t de_t \tag{5}$$

where the indicator function captures the fact that only borrowers with a quality ranking of $\bar{e}$ or more, are granted loans. We can re-express equation (5) as:

$$\pi(\bar{e}) = \max_{\bar{e}_t} \frac{1}{2}(1 + r)P(1 - \bar{e})^2 \left[\frac{\hat{r}}{r} - 1 - \delta\psi + \delta\psi\frac{1 - \rho}{\nu + 1}\right] + \rho(1 + r)\delta\psi P \left[\frac{1 - \bar{e}^{\nu+1}}{\nu + 1} - \frac{1 - \bar{e}^{\nu+2}}{\nu + 2}\right]$$

Taking the first-order condition and working through the algebra it is possible to solve for the threshold $\bar{e}$ in closed form,

$$\rho\bar{e}^{\nu} = 1 - \frac{1 - \rho}{\nu + 1} - \frac{\hat{r} - r}{r\delta\psi} \tag{6}$$

Then, it can be verified that this threshold is increasing in the cost of lending to an ex-post bad borrower, $\delta\psi$. This implies that the endogenous LTV is lower the costlier it is to have lent to an ex-post L-type borrower; which is exactly the direction the new regulation took. This conclusion, along with two others are summarized in Proposition 1.

Proposition 1. *A Loan to Value limit ($\bar{\ell} = 1 - \bar{e}$) is endogenously determined by the introduction of a provisioning cost for the contingent L-type borrower. This limit is*

1. *Non-increasing in the expected cost of the provision, $\delta\psi$*

2. *Non-increasing in the scarcity of good borrowers, governed by parameter ν.*

3. *Non-decreasing in the net profitability of each granted loan, as captured by the spread $\hat{r} - r > 0$*

Proof. Direct evaluation suffices. □

The stylized model above has all the intuition necessary to guide our empirical examination. Notably, in such a model, the lending interest rate $\hat{r}$ has been kept constant. That need not be the case. Price discrimination is a well-known strategy of firms to raise profit, but in the following subsection we show that such strategies do not pay off, and we can abstract from them. The key to this conclusion is the interaction of the imperfect information set-up and competition; both of which are good characterizations in the Chilean mortgage credit market.

2.2 Alternative Setups

2.2.1 Screening and charging two different interest rates

In the benchmark –and simplest– model we analyzed the case in which the bank charges a unique interest to all costumers once it has decided they should be granted a mortgage loan. It could be argued instead, the ex-ante heterogeneity of applicants implies different probabilities of them turning into L-type borrowers. Then, at least a two-interest-rate strategy should be implemented. The extension to more than two rates follows naturally. Consider such small extension to the benchmark problem of the bank: it has the possibility to deny granting a loan to costumers with signal $\tilde{e}$ below the cut-off level $\bar{e}$, charges interest rate r^h to borrowers with signal $\bar{e} \leq \tilde{e} \leq z$, and charges $r^l < r^h$ to borrowers with signal $\tilde{e} \geq z$. The case of perfect screening technology is sketched in Figure 3.

The problem of the financial intermediary is then: given prices $\{r^h, r^l, r, P\}$; the constant probability of entering into arrears for L-type borrowers δ; and the provision required by the regulator (in percentage) ψ; the problem of the financial intermediary is to choose $\{\bar{e}, z\}$ to solve the following program,

$$\pi(\bar{e}) = \max_{\{\bar{e},z\}} \quad \int_0^1 \int_0^1 \mathbb{1}\{\tilde{e} \geq \bar{e}|e\} \Big[\mathbb{1}\{\tilde{e} < z|e\}[\theta(e)V^H(\tilde{e}, r^h) + (1-\theta(e))V^L(\tilde{e}, r^h)] \\ \mathbb{1}\{\tilde{e} \geq z|e\}[\theta(e)V^H(\tilde{e}, r^l) + (1-\theta(e))V^L(\tilde{e}, r^l)]\Big] d\tilde{e}_t de_t \qquad (7)$$

where the first order condition with respect to z boils down to

$$(r^h - r - r\delta\psi)(1-z) - (r^l - r - r\delta\psi)(z-1) = 0$$

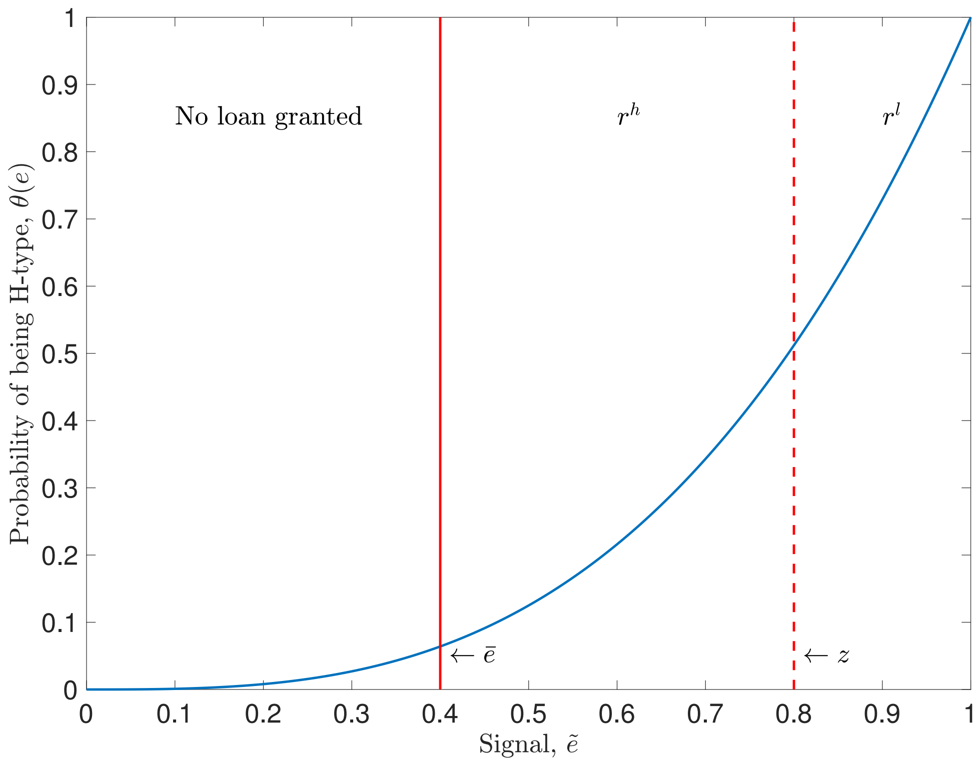


Figure 3: **Two interest-rate menu strategy:** The figure shows on the vertical axis the probability of turning H-type if the screening technology were perfect, namely $\rho = 1$, and $\tilde{e} = e$. For $e < \bar{e}$ the screening strategy dictates to not grant a loan; if $\bar{e} \leq e \leq z$ grant loan and charge a higher interest rate r^h; else, charge r^l to granted loans. Source: Own elaboration.

This last condition captures the fact that since loans are granted to the right of $\bar{e}$, and a financial cost is paid anyhow if the borrower becomes delinquent regardless of the charged interest rate, it is optimal for the bank to charge the highest possible interest rate. Hence its choices are,

$$\rho \bar{e}^{\nu} = 1 - \frac{1-\rho}{v+1} - \frac{r^h - r}{r\delta\psi} \tag{8}$$

$$z = 1 \tag{9}$$

that is, the bank charges effectively one interest rate, namely r^h, and the cut-off rule is the same as in the benchmark case if $\hat{r} = r^h$.

2.2.2 The case of two prices strategy and no screening

A third alternative model would be to simply separate the market and charge two different interest rates to costumers with signals below/above a threshold z. This is simply a special case of the previous extension with $\bar{e} = 0$, hence it delivers the same conclusions for the same reasons.

2.2.3 Including banking competition

In the benchmark model, we outlined the optimal cut-off rule for the screening problem of a bank that takes interest rates as given (see equation 5). We also outlined the case in which a given bank decides to charge two interest rates and argued that it is optimal for the said bank not to pursue such a strategy and charge the highest rate of the two offered. However, a third alternative equilibrium may be possible. It could be plausible to have an equilibrium in the mortgage credit market with two interest rates, charged by different banks; a leader bank that charges a lower interest rate and a follower bank that charges a higher interest rate. In the remainder of this section, we argue that this last outcome will, too, not be an equilibrium.

Consider the following set-up. There are (at least) two banks ($j = 1, 2$) who compete. For comparability with previous results, let the mass of costumers be normalized to two. Banks set interest rates first, and conditional on these decisions, choose $\bar{e}_j$. Suppose we start from equilibrium with positive profits and in which both banks, charge the same high-interest rate (r_j^h). Both banks are identical to the eyes of the potential borrowers, therefore they randomize which bank to go to first, and the result in equation (6) carries on for both banks. If bank j decides to deviate from this equilibrium and charge $r_j^l = r_{-j}^h - \epsilon$, its profits will differ for two reasons. First, the margin for each granted loan is lower as can be verified from (5). Second, and more importantly, costumers will no longer be randomly assigned between banks. They will go first to the cheapest bank (bank j), and if rejected, will go to the competitor bank which charges a higher interest rate for the same mortgage loan. This sequentiality is not only realistic but allows us to set the problem in a simple normal form game, where we can use the concept of dominant strategies.

Let us consider first the problem from the perspective of bank $j = 1$, who charges interest rate r^l while its competitor, bank $j = 2$ charges r^h. Let profits for this bank be denoted by $\pi_1(r^l, r^h)$ where the first argument in parenthesis denotes the action of the first bank and the second argument, the action chosen by its competitor. The optimal cut-off rule for bank 1 is given by equation (6) with $\hat{r} = r^l$, $V^k(\tilde{e}; \hat{r}) = V^k(\tilde{e}, r^l)$, $k = H, L$. Given this bank is the cheaper bank, costumers will go ask for a loan to bank 1 first, and if rejected, will turn to bank 2. We assume that it is costless for borrowers to apply for mortgage loans at any bank. The fact that bank 1 receives twice as many applications concerning the benchmark case, does not affect its choice of $\bar{e}_1$, and simply implies that profits will, too, be twice as those in the benchmark case. Hence,

$$\rho \bar{e}_1^{\nu} = 1 - \frac{1-\rho}{\nu+1} - \frac{r^l - r}{r\delta\psi} \tag{10}$$

Now consider the case of bank $j = 2$. Its problem is different because a fraction of borrowers (those with $\tilde{e}$ higher than $\bar{e}_1$, defined in equation (10)) already got their mortgage loans at bank 1. Then the problem of bank 2 is: Given prices $\{r^h, r, P\}$, and cut-off rule of the competitor bank, $\bar{e}_1$, choose $\bar{e}_2$ in order to solve the following program;

$$\pi_2(r^l, r^h) = \max_{\bar{e}_2} 2 \int_0^1 \int_0^{\bar{e}_1} \mathbb{1}\{\tilde{e} \geq \bar{e}_2 | e\} \left[\theta(e)\Delta(\tilde{e}) + V^L(\tilde{e}, r^h)\right] d\tilde{e}de \tag{11}$$

Working out the first order condition, we can obtain

$$\rho \bar{e}_2^{\nu} = 1 - \frac{1-\rho}{\nu+1} - \frac{r^h - r}{r\delta\psi}, \tag{12}$$

With this result at hand, we can compare the pay-offs to bank 2, for the two alternative interest rates it can charge: r^l, r^h. With r^l, both banks are charging a low-interest rate, hence we are back in the benchmark case. Alternatively, if the charged rated is r^h, then equilibrium profits $\pi_2(r^l, r^h)$ are given by plugging in (12) to (11). Both cases are depicted in figure (4) for different values of $r^h - r^l$.

From Figure (4) we can distil two insights. For small deviations of r^h from the competitor's charged interest rate, we have that (a) $\pi_2(r^l, r^h) < \pi_2(r^l, r^l)$. For large deviations of r^h from r^l the opposite is true; and we have that (b) $\pi_2(r^l, r^h) > \pi_2(r^l, r^l)$. For now, let us focus on the case (a). If bank 1 chooses r^l, then it is optimal for bank 2 to also choose r^l. If bank 1 chooses r^h bank 2 can choose an interest rate slightly lower than r^h and get all the market for itself, making (almost) twice as much profit as it would if it had chosen r^h. Hence, choosing r^l is a dominant strategy for bank 2. Next, consider the case (b). Suppose that bank 2 chose an interest rate r^h very much higher than r^l. It is clear that bank 1 has the incentive to raise r^l to $r^l = r^h - \epsilon$, with $\epsilon \to 0$. This way, bank 1 raises its profits and still keeps all the market to itself. But this implies that $r^h - r^l = \epsilon$ is very small, and we are back to the case (a). In summary, it is very hard for bank two to set an interest rate that is too high and expect the other bank not to set its interest rate a little below and steal all the market. Given this competitiveness, the optimal action for both banks to set its rate at a unique level of r^l. Thus, we go back to the benchmark case.

There are a lot of simplifications in the benchmark model. To start, we have assumed that the demand for mortgage loans is completely inelastic. Negatively sloped demand

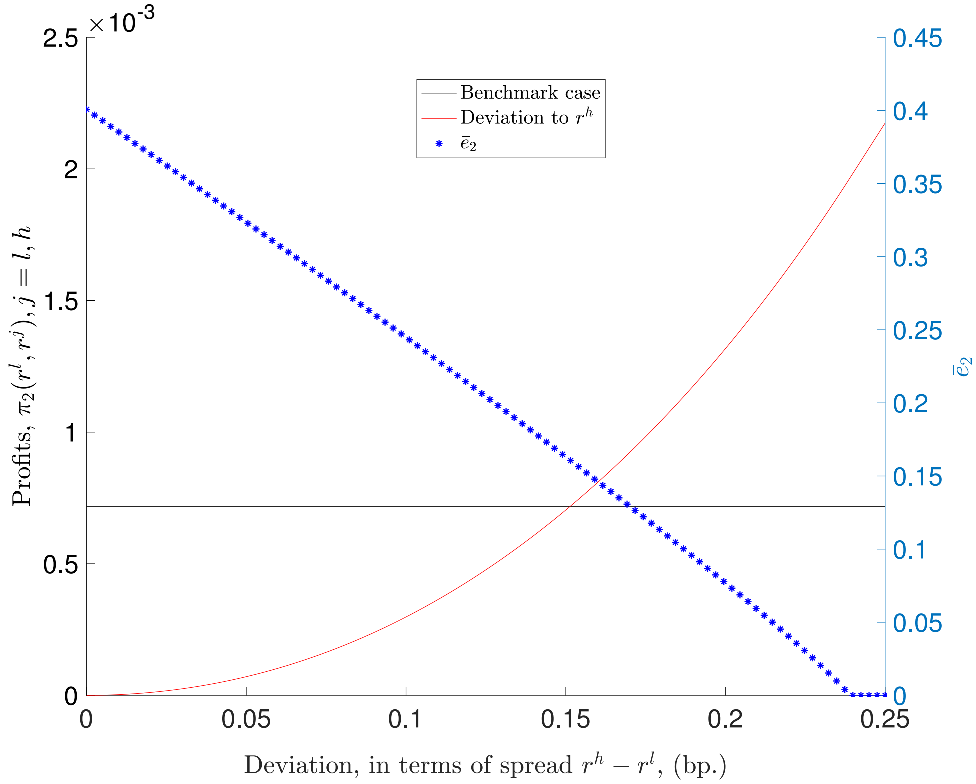


Figure 4: **Profits and cut-off strategy for bank 2:** Figure shows $\pi_2(r^l, r^l)$ in black, $\pi_2(r^l, r^h)$ in red, and the optimal cut-off rule as a function of $r^h - r^l$ in blue dots (right hand axis). Source: Own elaboration.

would further limit the ability of banks to set too-high interest rates. Second, we are assuming that there is no strategic interaction between borrowers and creditors. Instead, all bargaining power is assumed to belong to the financial institution. This means that borrowers do the best they can to provide the highest possible down-payment to value ratio, and if rejected they simply do not raise it again. A third major simplification of the model is to assume univariate signals. Instead of simply signaling good re-payment capacity with a high down-payment, borrowers could present proof of previous debt, add other properties as collateral, past behavior with the same creditor, etc. Highly important, banks could also use the debt-service-to-income to assess the probability of a borrower entering delinquency. We have abstracted from this as we discussed in the previous sections. That said, we still want to use the model to understand how scarce credit was allocated, after the coming into force of the new regulation on loan-loss provisions for mortgages. The LTV is particularly important to our analysis because of the signaling information it provides, and because of the regulation non-linearity in said ratio.

2.3 Calibration of the Model

In this subsection, we proceed to analyze a calibrated version of the benchmark model. We do so for two reasons. First, because it allows us to understand the ability of the proposed framework to generate effects on key variables that we can later examine empirically; and second, because it makes it easier to analyze the effects of the non-linearity of the loan-loss provision regulation around the 80% LTV threshold for a wide variety of plausible family parameters.

Table 1: **Baseline Calibration**

Parameter	Value	Target	Source/Target
ρ	0.90		Ates and Saffie (2013)
r	3.5		Banco Central de Chile (2017)
$\hat{r}$	3.7	2.73% markup (1)	Banco Central de Chile (2017)
δ	0.29	9% (2)	Pacheco, Pugar and Valdebenito (2014)
ν	0.69	90% LTV (3)	Median of LTV distribution, 2015

Notes: (1) mark-up is consistent with the CAR and ROE ratios reported in Chapter IV of Banco Central de Chile (2017); (2) Figure 2.1 in Pacheco, Pugar and Valdebenito (2014), share of borrowers who are delinquent, non-value weighted. To match this moment it is also necessary to calculate the probability of being L-type, conditional on being granted a loan. That is, $\mathbb{E}[\theta(e)|e > \bar{e}] = \frac{1}{\nu+1}(1 - \bar{e}^{1+\nu})$; (3) endogenous LTV limit of 90% at $\psi = 12.5\%$.

Our preferred calibration is summarized in table (1), for all parameters except for the provisioning cost ψ, as the new regulation implies substantial variation of this parameter with the leverage of the borrower in delinquency, and the time spent in said state. For instance; after the reform, the cost of provisioning was increased to more than 8% for highly leveraged borrowers who were delinquent for more than one day; but to more than 30% for the delinquency of more than 90 days –see Figure (1)–. Though we analyze large support for plausible values of ψ, our model does not distinguish one-day from 90-day delinquency; hence our quantitative conclusions should be understood only as an approximation to guide our empirical analysis.

We learn that the way the regulation was implemented, implies that the 80% threshold is very important for a wide set of plausible parameterizations. First, consider the benchmark calibration, with provisioning ψ not contingent on LTV. Panel (a) of Figure (5), depicts the optimal cut-off rule in equation (6) under different values of the scarcity of high-quality borrowers ν, and different values of provisioning ψ. If an applicant with a signal $\tilde{e}_j < \bar{e}$ (below any given curve) requests a loan, his application will be rejected. If the signal is $\tilde{e}_j \geq \bar{e}$ then they will be granted the loan. This threshold is non-decreasing in the provisioning cost, and in the scarcity of good borrowers, as we already stated in Proposition 1. Second, let us consider the non-linearity introduced in the regulation for parameter ψ. We learned from Figure 1 that below the LTV threshold of 80% the provisioning cost was negligible. Instead, above such threshold, and when a borrower enters delinquency, the provisioning cost can be as high as 30% of total asset value. If a signal $0.2 \leq \tilde{e} < \bar{e}$ we know that the associated provisioning cost in the horizontal axis, is not operative. This is depicted in the right-hand side panel in Figure (5), which shows the actual thresholds for mortgage loan approvals. For a wide set of provisioning costs, the 80% LTV limit is the one that matters. This is crucial to the understanding of the change in the distribution of LTV ratios documented in section 3. For (almost) all relevant values of the provisioning cost, ψ, it is an equilibrium outcome to observe that the distribution of LTV ratios of granted mortgage loans will gravitate towards 80%, as this will be the binding threshold above which banks will reject applicants.

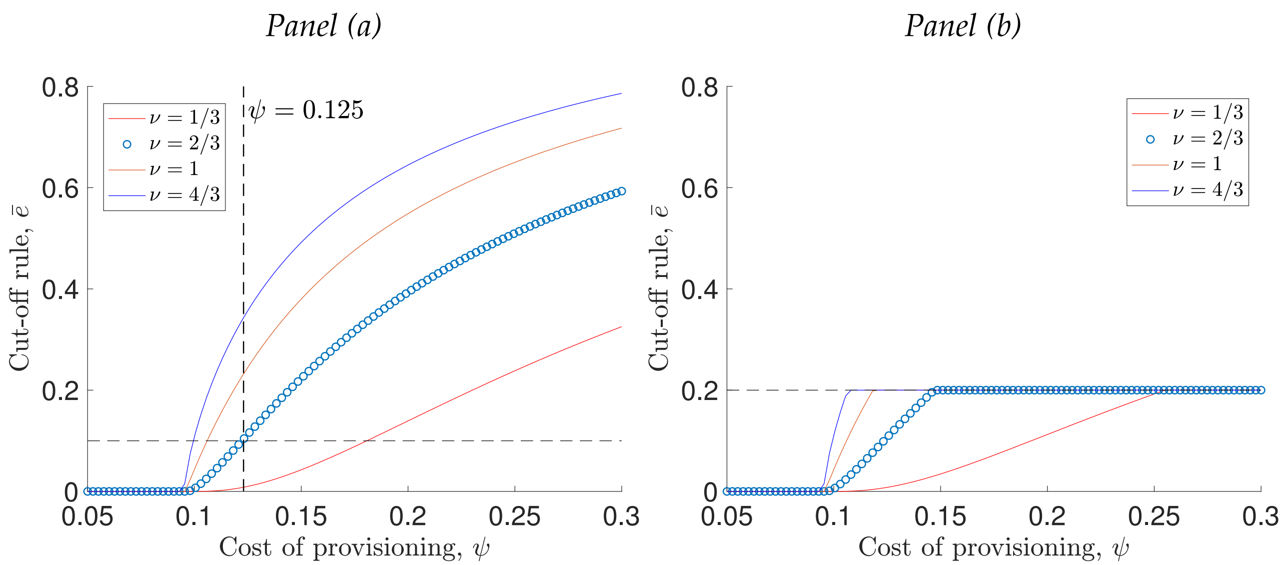


Figure 5: **Financial provision under new regulation**: Optimal threshold setting under of the simple model for different costs of financial provision, scarcity of good borrowers and non-linear regulation parameters.

3 Empirical examinations to loan-to-values and interest rates

We learned from the above model that the distribution of loan-to-value ratios of mortgages granted after the new regulation came into force, would necessarily gravitate toward lower values, for banks to accommodate the higher cost of ending up with a low-quality loan. Given a wide set of parameter values, and the specificities of the regulation, we concluded that the distribution of LTV would necessarily concentrate around 80%. This (predicted) swing in the LTV distribution, however, is unconditional on other relevant variables simultaneously changing around the date of implementation of the regulation under study. The most straightforward variables being: economic activity and house price growth, among others. In this subsection, we check for the main predictions of the model above: that controlling for a wide set of potentially relevant variables, the new regulation of financial provisioning in the mortgage market led unambiguously to lower LTV ratios, and that the threshold of 80% accumulates mass. We also check the prediction of the model about interest rates. Namely, that in equilibrium, financial institutions would not raise mortgage rates, but instead, do a better screening process. We also find support in the data for this claim.

3.1 A Matching Exercise

Our first conclusion is based on the results of a matching exercise that exploits the richness of our dataset; the coarsened exact matching (CEM) algorithm proposed in Iacus, King and Porro (2012). Matching is a widely-used method of evaluation of non-experimental treatments or programs. The principle behind this method is quite intuitive; it contrasts the outcomes of "program" participants (Y_1) with the outcomes of "comparable" non-participants (Y_0) (An extensive summary of the benefits of matching can be found in Heckman, Ichimura and Todd (1998)). The main idea is that differences in the outcomes between the two groups are attributed to the program or treatment, given that groups were indeed "comparable" in every other sense.

The method is powerful, so it is no surprise that applications can be found in the evaluation of an extensive list of policies. For instance, Heckman, Ichimura and Todd (1997), Lechner (2002), Jalan and Ravallion (2003) and Smith and Todd (2005) evaluate the impact of training programs on earnings; Galiani, Gertler and Schargrodsky (2005) evaluate the impact of privatization of water services on child mortality in Argentina in the 1990's; Encina (2013) studies the labor market effects of the 2008 pension reform in Chile; and Almus and Czarnitzki (2003) and Moser (2005) study the impact of subsidies and patent laws on research and development, patents, and innovation. The housing and credit markets are no exception. To name a couple among many others; Park (2016) studies mortgage performance for FHA and privately insured home purchases relative to uninsured mortgages; and Field and Torero (2006) study the impact on credit supply of obtaining a property title through a land titling program in Peru.

3.1.1 Some definitions

The introduction of the regulation of financial provisioning for mortgage loans is an exogenous event from the perspective of a given household's home buying decision, but it is not entirely experimental. The problem –as with any non-experimental data–, is that counterfactuals are unobserved. Ideally one would be interested in observing the outcome variable of an individual who received the treatment and the outcome for that same individual without the treatment. In our set-up, we would like a potential borrower to enter a bank and have a coin decide on whether the new regulation applies to him, and enter again and do the opposite; and compare the outcome. This experiment is not available, and we use a matching method to try to uncover two samples that mimic this sort of experiment.

Following Smith and Todd (2005), define a dummy variable D, which takes the value of one ($D = 1$) if the new regulation has come into force (starting 2016), and zero ($D = 0$) if not (before 2016). Our object of interest is the mean differential effect on the outcome variable (Y, LTV) on those households subject to the new regulation compared to their counterfactual under no-regulation: the mean effect of treatment on the treated for people with covariates X,

$$ATT = E(Y_1 - Y_0|D = 1, X) \tag{13}$$

where $E(Y_1|D = 1, X)$ represents the outcome (LTV) for agents who were affected by the new regulation, and $E(Y_0|D = 1, X)$ the outcome for agents had they not have been affected by the new regulation but the regulation was already active (the unobserved counterfactual). The first term can be directly identified from home buyers in 2016. The second term, however, is unobservable. As an approximation to the second term, $E(Y_0|D = 0, X)$ is used, meaning the no-treatment outcome of buyers when the regulation was not active. This approximation has a potential selection bias:

$$B(X) = E(Y_0|D = 1, X) - E(Y_0|D = 0, X) \tag{14}$$

Then, the fundamental identification condition for estimating (13) is *conditional mean independence* (see Heckman, Ichimura and Todd (1998)):

$$E(Y_0|D = 1, X) = E(Y_0|D = 0, X), \tag{15}$$

which amounts to saying that conditioning on X, eliminates the bias; or that conditional on X, studied agent samples are balanced. Exactly balanced data means that controlling further for X is unnecessary because it is unrelated to the treatment variable. It also means that model dependence is minimized and researcher's discretion along with it (Ho et al., 2007).

3.1.2 Coarsened Exact Matching

The most straightforward (and ideal) matching would be exact matching. That is, emulating a fully blocked experiment in which two agents are matched with the same covariate variables (X), and then treatment is randomly applied to one of them. This type of matching not only balances unobserved covariates on average, but balances observed covariates exactly (Ho et al., 2007). Unfortunately, when using several covariates –and when at least one of them is a continuous variable–, this approach becomes impractical because finding exact matches becomes unlikely. Other methods of *approximate* matching rely on finding "close enough" covariates for the control and treated agents. Notably the Mahalanobis Distance Matching (MDM), or the popular Propensity Score Matching (PSM), or the Coarsened Exact Matching (CEM).

In this paper, we choose to use CEM over MDM and PSM. We choose not to use PSM as it is the least efficient of the three methods. PSM takes several X covariates, summarizes them into the "propensity score", and uses this one scalar as a measure of the distance between treated and control units –as opposed to using a distance which considers all k dimensions of X–. Then, it prunes any observations that do not get matched. But doing so results in loss of information because there is an inherently random component dictating which observations are dropped. Notably, it is not the pruning that makes the method less efficient. On the contrary, all matching methods rely on some form of pruning. PSM is less efficient than the alternatives because of the way such pruning is performed. The second method; the MDM emulates a fully blocked experiment defining a (euclidean) distance between covariates (X). Later, it prunes units that are not close enough and compares the outcome variable on those surviving matches. It deals more satisfactorily with continuous variables and seeks to compare treated and control covariates using a multidimensional notion of distance, therefore not incurring in random pruning. The shortcoming of the method is the not-so-obvious way to weight every covariate –with different units– in the euclidean distance. CEM addresses this point more directly while keeping all the advantages of the MDM.

The CEM is an approximation to exact matching. We have already made the point that while the exact matching provides perfect balance, it does so at the cost of producing very few matches, in particular when a covariate is a continuous variable. CEM attempts to address this weakness. The idea presented in Iacus, King and Porro (2012), is to temporarily coarsen each variable into substantively meaningful groups; exact match on these new data; sort observations in strata; prune any strata with no treated or control units, and pass on only original un-coarsened values after pruning. The method is more powerful if the coarsening is nourished by a meaningful grouping of covariates[10]. There are other attractive properties of the method. First CEM belongs to the group monotonic imbalance-reducing methods, which means that the balance between treated and control groups is chosen ex-ante (i.e. employing the coarsening), rather than post-estimation as in the propensity score matching. Also, CEM meets the congruence principle, which states that data and analysis spaces should be the same. This is achieved via pruning of observations whose strata (bins in the coarsening) fail to find a match in the complementary (treated/control) group. Finally, CEM restricts matched data to areas of common support by construction, which is a requirement to be checked post-estimation when using the PSM.

3.1.3 Results of the Matching Exercise

We explore our data in three complementary sets of experiments: a benchmark case (two alternative exercises), an anticipation case, and two placebo tests. In the benchmark case we compare individuals who were given credit before, and after the regulation came into force in January 2016. While we examined several periods as candidates for the control group, the results are very robust to this choice. Hence, we report the results of using loans granted during years 2012-14 (and 2013-14) as the control group, and 2016-17 as the treated group. The anticipation exercise uses loans granted in the year 2015 (after the regulation was announced but not yet enforced) as the treated group and those in the year 2014 as the control group. Finally, we present two placebo exercises, in which the year 2014 is considered the treated group against two alternative control groups: individuals who were given credit in 2013, or in years 2012-13.

Across all our experiments we have kept the coarsening of variables unchanged, to ensure comparability. In particular, the vector $X \in R^k$ includes the following seven dimensions in which we perform the matching: neighbourhood ("comuna"); property price in real terms; maturity of mortgage loan in years; lender institution; size of the property (square meters); income of the borrower (up to taxable income brackets); and type of property (apartment/house). Loan maturity is coarsened using the following cut-points (in years): $\{15, 20, 25, 30, 35\}$. The neighborhood, lender financial institution, income bracket, and property type are no further coarsened. All remaining variables, except loan maturity, are coarsened automatically using the CEM package by King et al. (2010) which uses Scott's method (Scott, 2015). Given the featuring role of pruning in the method, table (2) reports some summary statistics of the matching. Across all experiments, we can see that one of every four strata contains control and treated units and is therefore kept. All other

[10] For instance, if a covariate is years of schooling, we could group them into basic schooling, high school, college degree, post-graduate, etc. Or in our case below, the length of mortgage loan can be split into intervals centered around 15, 20, 25, and 30 years, which are typically the loan lengths used by the financial sector.

Table 2: **Descriptive statistics of matched samples**

	Benchmark (1)	Benchmark (2)	Anticipation (3)	Placebo (4)	Placebo (5)
Number of strata	110,635	94,738	71,063	62,495	81,279
Matched	27,648	24,827	20,692	15,768	20,237
Unmatched	82,987	69,911	50,371	46,727	61,042
Number of Control Units	385,223	251,950	127,683	124,267	257,540
Matched	287,198	190,365	102,954	92,195	190,227
Unmatched	98,025	61,585	24,729	32,072	67,313
Number of Treated Units	270,088	270,088	184,728	127,683	127,683
Matched	220,369	215,294	139,768	91,311	99,103
Unmatched	49,719	54,794	44,960	36,372	28,580
Overall Imbalance (L1)	0.393	0.390	0.439	0.423	0.422

Note: In this table we show the main results of sample and strata size after pruning, for five different exercises: First, two benchmark experiments in the first two columns. The third column tests anticipation effects given the regulation was announced a year before entering into effect. Lastly, statistics for two placebo tests. In particular, (1) Specification takes years 2012-14 as control and 2016-17 as treated. (2) Specification takes years 2013-14 as control and 2016-17 as treated. (3) Specification takes year 2014 as control and 2015 as treated. (4) Specification takes the year 2013 as control and 2014 as treated. (5) Specification takes years 2012-2013 as control and 2014 as treated.

strata contain no observations, or either only treated, or only control units. However, the method still uses three out of every four units in the control and treated groups, as can be seen from the ratio of matched units to total units in every group. That is, the matching method restricts to a small common support region, in which it uses intensively most of the observation units.

An overall imbalance metric $\mathcal{L}_1$ is also reported in table (2). This statistic is a distance notion between multidimensional histograms of treated and control group (Iacus, King and Porro, 2012). Intuitively, it provides information about how balanced the covariates in the two groups are. Technically, consider a total of s strata (multidimensional boxes) in which the covariates are coarsened and exactly matched, then record the k-dimensional relative frequencies for treated f and control g units. The measure of imbalance is the absolute difference overall s cell values: $\mathcal{L}_1(f,g) = 1/2\sum_{i=1}^{s}|f_i - g_i|$. If this statistic takes the value of zero, then we have achieved a perfect balance, if it takes the value of 1, then we have a total imbalance. As mentioned by Iacus, King and Porro (2012), this statistic is to matching as R^2 is to regression analysis. Next, we compare the results of the three sets of experiments in terms of the variable of interest, the loan to value ratio.

Benchmark Results. In columns (1) and (2) of table (3) and (4) we report some statistics of the distribution of loan to value ratios. Even though the control group in (2) is smaller by one third, the method proves very robust to this exclusion. We can see that borrowers in 2012-14 were granted loans that were on average, 81.5% of collateral value. During and after 2016, loans granted to a comparable group of borrowers were smaller; averaging 78.8% of collateral value. We attribute the -2.7% difference to the coming into force of the regulation on provisioning for credit risk in the mortgage market. The picture is clearer if we consider the percentiles of the distribution, as in table (4). From the first two benchmark experiments, we learn that the 25th and 75th percentiles were hardly changed. However, a large mass of borrowers did move. The median borrower pre-regulation borrowed 89.8% of collateral value. After the regulation was introduced that number dropped to 80% exactly, as the calibrated model in the previous section anticipated.

Table 3: **Loan to value ratio: means of treated and control groups**
(expressed in percentage)

	Benchmark (1)	Benchmark (2)	Anticipation (3)	Placebo (4)	Placebo (5)
Mean					
Control Units	81.75	81.50	81.62	81.35	81.82
Treated Units	78.80	78.82	81.11	81.61	81.64
Std. Err.					
Control Units	0.04	0.04	0.05	0.06	0.05
Treated Units	0.04	0.04	0.05	0.06	0.06
Difference					
Mean	2.95	2.68	0.51	-0.26	0.22
Std. Err.	0.06	0.06	0.08	0.09	0.08
Two-sample t test (t-stat)	53.21	44.69	6.63	-3.06	2.89

Note: In this table, we show the main results of the sample and strata size after pruning, for five different exercises: First, two benchmark experiments in the first two columns. The third column tests anticipation effects given the regulation was announced a year before entering into effect. Lastly, statistics for two placebo tests. In particular, (1) Specification takes years 2012-14 as control and 2016-17 as treated. (2) Specification takes years 2013-14 as control and 2016-17 as treated. (3) Specification takes the year 2014 as control and 2015 as treated. (4) Specification takes the year 2013 as control and 2014 as treated. (5) Specification takes years 2012-2013 as control and 2014 as treated.

Dealing with anticipation. In the previous baseline exercises, we assumed that a treated household, was one who got a mortgage loan after January 1st, 2016 –when the regulation was fully enforced– and that a control household was one who got a loan before December 2014, when the regulation was announced. Thus, dropping 2015 is a choice made to keep the exercise as clean as possible, but the downside is that we are missing a potentially important anticipation effect. In order to evaluate if this is the case, consider column (3) in tables (3) and (4). Note that while it is true that mean LTV is marginally (but statistically significant) lower post-December 2014, other moments of the distribution are unchanged. Percentiles 25 and 75 remain 80% and 90% respectively. Also, in contrast to the 9.8% drop in the baseline cases, the anticipation effect for the median borrower is only 0.8% (from 89.8% to 89%).

Table 4: **Loan to value ratio: quantiles for treated and control groups**
(expressed in percentage)

	Benchmark (1)	Benchmark (2)	Anticipation (3)	Placebo (4)	Placebo (5)
Percentile 25					
Control Units	80.0	80.0	80.0	80.0	79.9
Treated Units	75.1	76.2	79.7	80.0	79.8
Percentile 50					
Control Units	89.8	89.9	89.8	89.9	90.0
Treated Units	80.0	80.0	89.0	89.5	89.4
Percentile 75					
Control Units	90.0	90.0	90.0	90.0	90.0
Treated Units	90.0	90.0	90.0	90.0	90.0

Note: In this table we show the main results of sample and strata size after pruning, for five different exercises: First, two benchmark experiments in the first two columns. The third column tests anticipation effects given the regulation was announced a year before entering into effect. Lastly, statistics for three placebo tests. In particular, (1) Specification takes years 2012-14 as control and 2016-17 as treated. (2) Specification takes years 2013-14 as control and 2016-17 as treated. (3) Specification takes the year 2014 as control and 2015 as treated. (4) Specification takes the year 2013 as control and 2014 as treated. (5) Specification takes years 2012-2013 as control and 2014 as treated.

Placebo tests. In columns (4) and (5) of tables (3) and (4) we report the results for two placebo tests. In the first one our control group is households who were granted loans in 2013, and the treated group those who got theirs in 2014 (before the regulation was announced). The second placebo test extends the control period to 2012-13. As in the baseline case, the actual choice of control group (period) is immaterial as long as it is effective before the regulation came into force. The first placebo test presents evidence that the treatment resulted in a 0.26% hike, and the second a 0.22% drop in LTV ratios as is evident from table (3). In the same way, percentiles 25th and 75th are unchanged. Furthermore, our placebo tests imply that there was a drop in the median LTV of the treated group between 0.4% and 0.6%. These numbers are an order of magnitude lower than 9.8% reduction in the median of LTV in the baseline scenario. Taken together, all these experiments confirm the fact that the new regulation on provisions for credit risk in the mortgage market had a bite in the decision of banks to extend smaller loans, relative to the pledged collateral.

Panel (a): Benchmark model *Panel (b): Placebo test*

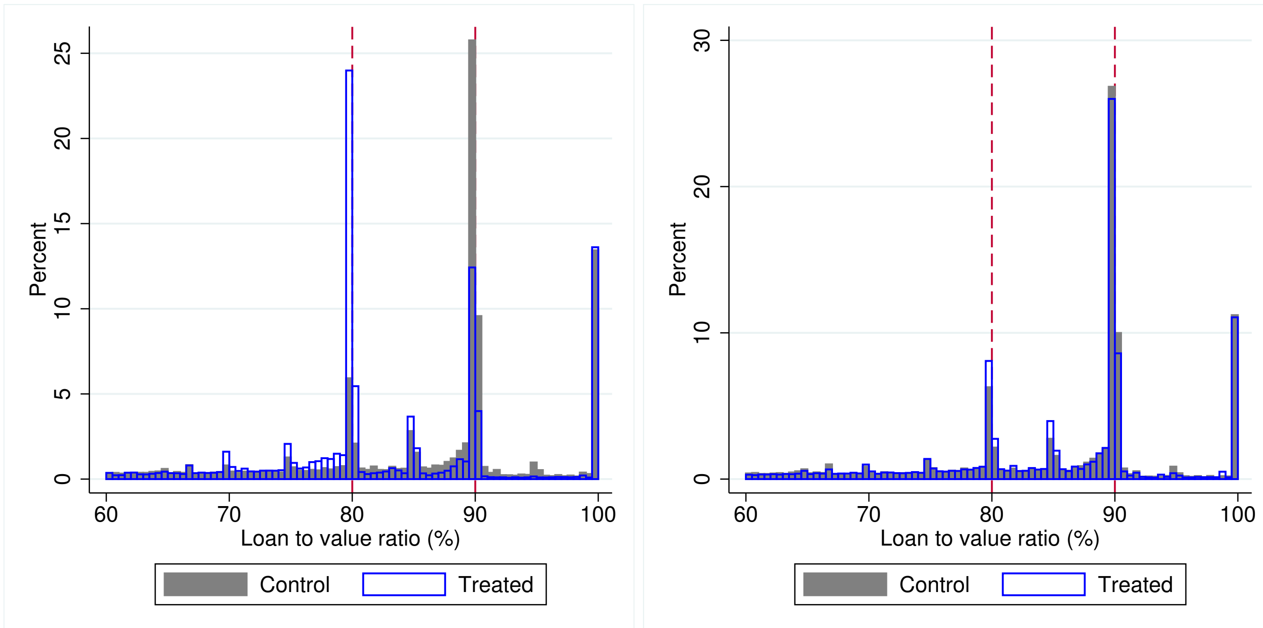


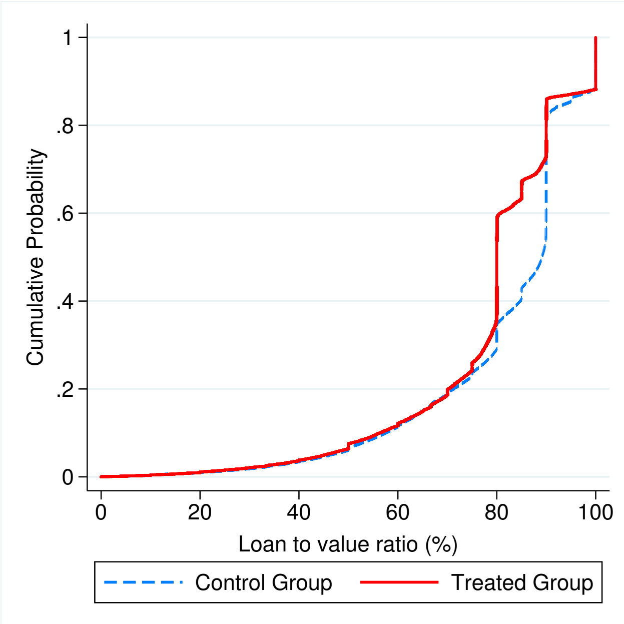
Figure 6: **Histograms of LTVs**: Bin widths are 0.5% to highlight that LTV ratios were concentrated around the 80% threshold. However, bar heights represent the fraction of the sample en each bin; e.g. 24% of loans had LTV ratios between 79.5% and 80% in the Benchmark matching. The figure shows the Benchmark specification (1) and a placebo test (4). (1) takes years 2012-14 as control and 2016-17 as treated. (4) Specification takes the year 2013 as control and 2014 as treated. Source: Own calculations based on IRS data.

To make our point more explicit let us present the histograms corresponding to specifications (1) and (4) in figure (6)[11]. The red pointed lines mark 80% and 90%. Panel (a) shows the baseline exercise. It is clear that after the regulation an important probability mass transited from just below 90% to just below 80%. In particular, the number of loans granted at exactly 80% more than tripled with the new regulation. On the contrary, Panel (b) shows a placebo test (2013 vs. 2014). We see that treatment indeed raises LTV at the 80% threshold level, but does so an order of magnitude relative to the baseline case. This is the same intuition conveyed from table (4). Also, in figure (7) we present the Cumulative Distribution Function for the same two experiments. On Panel (a) it is clear that treatment-LTV-CDF is different, both statistically and economically, from the control-group-LTV-CDF. On the contrary, in Panel (b) we show how similar the CDFs of treatment and control groups are in the placebo test. These figures are only another way to interpret the same information

[11] An earlier version of this paper used kernel density estimates for this evidence. We choose to use histograms with bins of 0.5% width to highlight the fact that there is a substantial concentration in the neighborhood below the 80% LTV.

Panel (a): Benchmark model

Panel (b): Placebo test



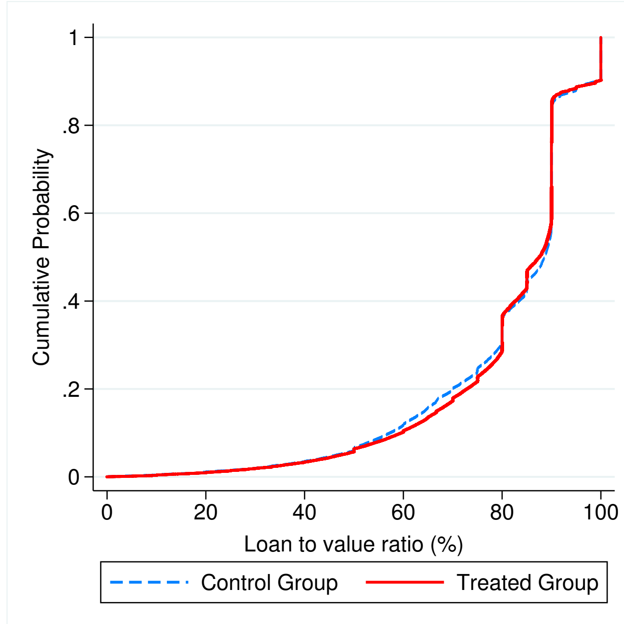


Figure 7: **Cumulative Distribution Functions**: The figure shows the Benchmark specification (1) and placebo test (4). (1) takes years 2012-14 as control and 2016-17 as treated. (4) Specification takes the year 2013 as control and 2014 as treated. Source: Own calculations based on IRS data.

as in figure (6), but highlight the effect on terms and conditions on granted loans due to the coming into force of the regulation of provisions for credit risk of mortgage loans.

3.1.4 Testing an alternative non-random control group through Diff-in-Diff

Through the CEM matching approach, we have constructed balanced CEM samples, which will prove to be very useful in our upcoming experiments. Besides, using those samples we have shown a swing in the LTV distribution towards 80% after the new regulation came into action, confirming one of the predictions from our theoretical model. Now, we exploit the fact that a small and non-random fraction of mortgage loans was not affected by the policy change: those associated with government housing subsidy programs. Since government housing programs impose a cap on the value of subsidized houses (Pacheco, Pugar and Valdebenito, 2014), we can roughly approximate those loans as mortgages of less or equal to UF 500, and set them as a control group to assess the impact of the change in the LLP regulation using the balanced CEM samples obtained previously. This exercise can help us validate not only the robustness of our previous empirical findings but also the use of this small fraction of mortgages as a control group for future analysis. If mortgages related to government subsidy programs were not affected by the regulation, then one must expect them to reduce their associated loan to value ratios in a lower magnitude than the other fraction under treatment, or even increase them.

A Difference in Differences Evaluation. We know that the regulation on loan-loss-provisions entered into force in 2016. Also, we know that a small fraction of mortgages was not affected by this policy. Therefore, a natural approach to evaluate its associated effects is a diff-in-diff estimation, making full use of the already constructed balanced CEM samples. The method relies on the existence of before and after periods as well as on two groups, namely control and treated. Among those groups, only the treated are delivered the treatment and both has no intervention in the before period. Formally, define before and after periods, $T = 0$ and $T = 1$, respectively and from the total sample of i individuals, establish $Z_i = 0$ and $Z_i = 1$, to be control and treated respectively. Therefore, for a particular outcome variable $Y_{i,T}$ the diff-in-diff procedure is given by:

$$\begin{aligned} DID = & [E(Y_{i,T=1}|D_{i,T=1} = 1, Z_i = 1) - E(Y_{i,T=1}|D_{i,T=1} = 0, Z_i = 0)] \\ & -[E(Y_{i,T=0}|D_{i,T=0} = 1, Z_i = 1) - E(Y_{i,T=0}|D_{i,T=0} = 0, Z_i = 0)] \end{aligned} \tag{16}$$

where, the above expected values can be obtained through the following linear regression:

$$Y_i = \beta_0 + \beta_1[T_i] + \beta_2[Z_i] + \beta_3[T_i Z_i] + \mu_i \tag{17}$$

and, the DID estimand is $\hat{\beta_3} \sim \beta_3$. Further details can be found in Angrist and Pischke (2008).

Results of the Diff-in-Diff Estimation. We now evaluate the effects of the loan-loss-provision regulation in the LTV by using a non-random control group: the mortgages related to government subsidy housing policies ($Z_i = 0$). All other loans are set as the treated group ($Z_i = 1$). We restrict our analysis to the periods 2014 ($T = 0$) and 2016 ($T = 1$).

Besides, to overcome the possible effects of other confounding covariates we rely on the balanced CEM samples. Table (5) presents the means and standard errors of the loan to value ratio, for both control and treated groups in the before and after periods. The negative *diff-in-diff* estimate implies a reduction in the LTV for the loans under treatment because of the new regulation. Specifically, one can see that in 2014, when there was no regulation, the mean LTV for the treated mortgages was 82.92%, approximately. In contrast, when the regulation entered into action, their mean LTV reduced to 80.11%. Surprisingly, loans related to government housing subsidy programs, did not decreased their mean LTV but increased

Table 5: **Diff-in-Diff: Estimates**

	LTV	S. Error	$\lvert t \rvert$	$P > \lvert t \rvert$
Before				
Control	46.338			
Treated	82.919			
Diff (T-C)	36.581	0.415	88.10	0.000***
After				
Control	51.757			
Treated	80.113			
Diff (T-C)	28.335	0.464	61.16	0.000***
Diff-in-Diff	-8.226	0.622	13.22	0.000***

Note: In this table we show the main results of the difference in differences experiment using the balanced CEM sample. The estimation process was carried out using the utility program in Villa (2016). Means and Standard Errors are estimated by linear regression. Inference: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

it from 2014 to 2016. In fact, their mean LTV increased from 46.34% in 2014 to 51.76% in 2016. This change suggests not only that the new regulation did not negatively affected loans related to government housing programs but also that credit conditions became less tight for them when the policy entered into force.

3.2 A Bunching Estimation

3.2.1 The Saez-Chetty Bunching Procedure

Having stablished that there exists an important probability mass swing towards a LTV of 80%, as predicted by our model in section 2, it seems natural to empirically analyze the degree of "bunching" or excess mass around this point. In a way we have already done so with the empirical distribution of the previous subsection, but we go beyond and follow the method proposed in Saez (2010) and Chetty et al. (2011), to estimate the excess probability mass around a particular LTV for 2016, and compare it to the period before the regulation was effective. Our interest relies not only in finding evidence of bunching but rather on the *change* of in probability mass around the bunching point.

We start by computing the counterfactual distribution of LTV ratios in the absence any kink, and compare it with its empirical counterpart. The computation process fits a q degree polynomial to the empirical distribution, "excluding" the points around threshold kink using the following regression:

$$C_j = \sum_{i=0}^{q} \beta_i^0 (Z_j)^i + \sum_{i=-R}^{R} \gamma_j^0 \times \mathbb{1}[Z_j = i] + \epsilon_j^0 \tag{18}$$

where C_j is the number of granted loans in LTV-bin j, Z_j is LTV relative to the 80% kink in 1% intervals, and $[-R, R]$ represents the region which is excluded around the threshold. In our case, this range is 1%. Since the above computation does not ensure that the area under both, counterfactual and empirical distributions, must be equal, the procedure defines the counterfactual as $\hat{C}_j = \hat{\beta}_i (Z_j)^i$, which is equal to the fitted values from:

$$C_j (1 + \mathbb{1}[j > R]) \frac{\hat{B}_N}{\sum_{j=R+1}^{\infty} C_j} = \sum_{i=0}^{q} \beta_i (Z_j)^i + \sum_{i=-R}^{R} \gamma_i \times \mathbb{1}[Z_j = i] + \epsilon_j \tag{19}$$

where $\hat{B}_N$ is the excess of loans granted around the implied kink in equation (18)[12]. Finally, the empirical estimate of excess mass around the threshold kink in proportion to the average density of the counterfactual LTV distribution in the range $[-R, R]$ is:

$$\hat{b} = \frac{\hat{B}_N}{\sum_{j=-R}^{R} \hat{C}_j / (2R + 1)} \tag{20}$$

3.2.2 Results of the Bunching Estimation

Our analysis around the discontinuity kink of 80% in LTV can be divided into two stages. First, we use the original data where we are not controlling for other variables and then, in a second stage we use our balanced sample from the CEM procedure, described in the previous section. In both cases, the periods under analysis are 2014, for the control group and 2016 for the treated group. We use a first-order polynomial ($q = 1$) to predict the counterfactual distributions. The samples are normalized exactly at the threshold kink with frequency expressed in percentages[13] and standard errors are estimated through a parametric bootstrap method with 5000 replications. As in Chetty et al. (2011), these standard errors indicate error owing to inappropriate degree polynomial specification instead of purely sampling error. Notice that, to ensure comparability across the experiments we kept these specifications fixed.

Figure (8) shows the empirical and counterfactual distribution from our bunching procedure using the original data of loan-to-value ratios for 2014 and 2016. We estimate the excess mass $\hat{b}$ to be equal to 57.23 in 2014 and 80.66 in 2016. For both years, we reject

[12] For further details, see Chetty et al. (2011).

[13] We count of the number of times that an LTV occurs and then transform it into a percentage.

Panel (a): Control Group *Panel (b): Treated Group*

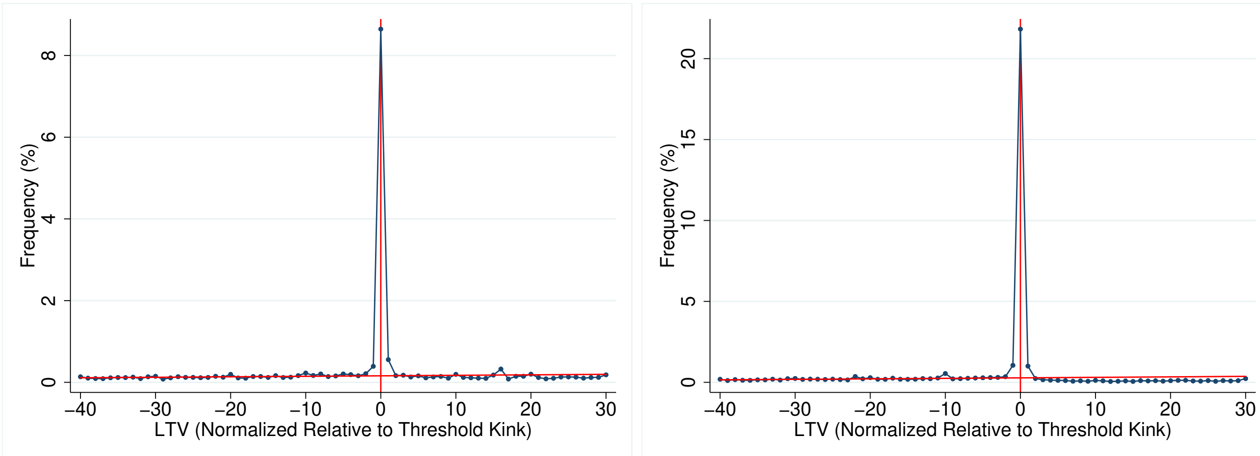


Figure 8: **Bunching in LTVs, Raw Samples**: The figure illustrates the counterfactual (red line) and empirical distributions of loan to value ratios in percentages for 2014 (*Control Group*) and 2016 (*Treated Group*) using the original data. The solid red vertical lines show the threshold kink at LTV 80%. Source: Own calculations based on IRS data.

Panel (a): Control Group *Panel (b): Treated Group*

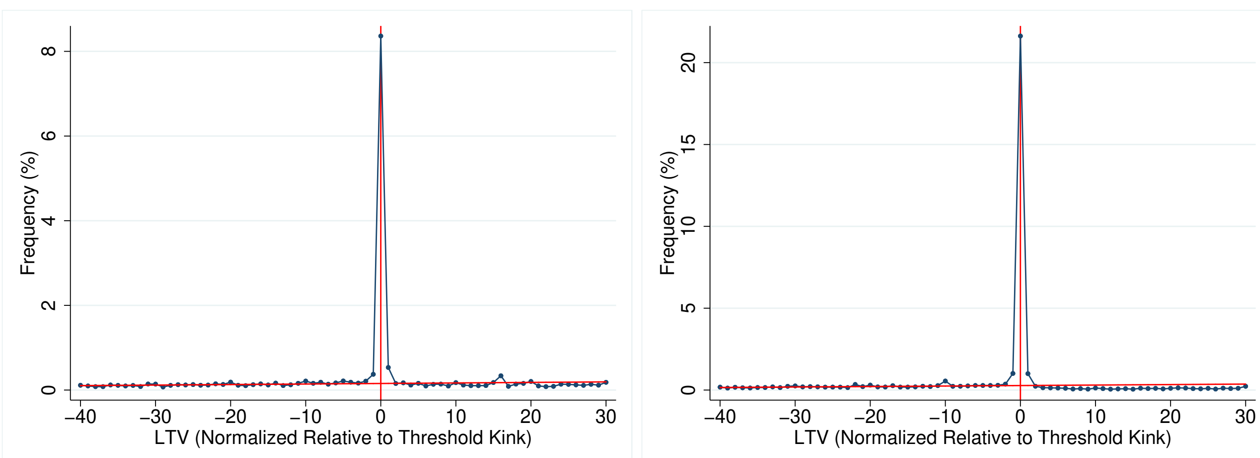


Figure 9: **Bunching in LTVs, CEM Balanced Samples**: The figure illustrates the counterfactual (red line) and empirical distributions of loan to value ratios in percentages for 2014 (*Control Group*) and 2016 (*Treated Group*) using the matched CEM samples. The solid red vertical lines show the threshold kink at LTV 80%. Source: Own calculations based on IRS data.

the null hypothesis of no-bunching or no-excess-mass relative, to the counterfactual distribution. To control for confounding factors different from the change in regulation, we use the balanced samples from the CEM procedure. The results of using these balanced samples for the bunching estimation are in Figure (9). We can confidently still reject the null hypothesis, yet our estimates of the degree of bunching are different. In particular we estimate $\hat{b}$ = 56.95 in 2014 and $\hat{b}$ = 79.85 in 2016. In both cases (see figures) the spikes at 80% LTV threshold, are higher in 2016 than in 2014. The associated standard errors and the changes in the excess mass from one period to the other are presented in Table (6).

Table 6: **Bunching in LTVs: Estimates**

Year	*LTV, Raw Samples*			*LTV, Balanced CEM Samples*		
	Excess Mass	Standard Error	Change	Excess Mass	Standard Error	Change
2014	57.23	2.80		56.95	2.85	
2016	80.66	3.47	40.94%	79.85	3.42	40.21%

Note: In this table we show the main results of the bunching exercise using two different samples. The estimation process was carried out using the utility program in Chetty et al. (2011).

As we can see from the results, the excess mass is highly concentrated in 2016 when the new regulation for loan-loss-provisions for mortgages came into force. Specifically, we estimate this change from 2014 to 2016 (Δ%) in the excess mass around the threshold kink equal to 23.4% and 22.9%, when using the raw samples and the balanced CEM samples, respectively.

3.3 Looking at interest rates

Our objective in this section is to explore how interest rates on granted mortgages were affected. It also serves us to a broader objective: to help us distinguish and validate our stylized model. In a different model where imperfect information were not important, financial institutions could simply offload the higher expected cost of granting a loan to the borrower with the highest loan-to-value ratios. Our model, predicts that this cost offloading is not an equilibrium outcome, a fact we proceed to verify empirically.

We use administrative data about loan contracts collected by the banking regulator. This exceptional dataset (informally known as archive ”D32”) includes all transactions on commercial, consumer and mortgage loans granted by banks in the period spanning from 2012 to 2019. Some of the variables included are identifiers for both banks and borrowers for all contracts, coupled with contract-specific features including daily date of transaction, term, loan amount and its corresponding interest rate. We merge the data on loan contracts (those related to mortgage credit) with our other administrative data on real state transactions (”Form F-2890”) from the IRS, and check for evidence of interest rate offloading to borrowers more likely affected by the new regulation –those with LTV ratios higher than 80%. We rely on evidence based on regression discontinuity (RD) to argue that there is no evidence of interest rate offloading for more leveraged borrowers.

3.3.1 The Regression Discontinuity Design Approach

In regression discontinuity design, we can quantify the magnitude of the effect of a particular treatment as the size of the vertical discontinuity in fitted regressions at a given cut-off. The method was originally presented by Thistlethwaite and Campbell (1960) and currently is one of the most popular among areas of social and natural sciences. Recent advances in RD design have been introduced by Cattaneo, Frandsen and Titiunik (2015) and Calonico, Cattaneo and Titiunik (2015). In the context of financial markets, this approach has been used before. For instance, Chava and Roberts (2008) used RD design to show that capital investment levels plummeted after a financial covenant violation in the U.S., and Albertazzi, Bottero and Sene (2017) explored the impact of the number of past rejections in loans applications on the loans approval rate in Italy, through the same approach.

Formally, define the random sample $[Y_i(0), Y_i(1), X_i]'$ from $\{Y(0), Y(1), X\}'$, where $Y(0)$ and $Y(1)$ are control and treatment groups of the outcome variable Y, given the covariate X (also known as *"forcing"* or *"running"* variable). The units of Y_i for which $X_i \geq \bar{x}$ (with $\bar{x}$ being a known cut-off), are set as treatment group $T_i = 1$, and the rest is known as the control group, $T_i = 0$. More specifically, following Imbens and Wooldridge (2009) the observed outcome of Y_i is given by:

$$Y_i = Y_i(0)(1 - T_i) + Y_i(1)T_i = \begin{cases} Y_i(0) & \text{if } X_i < \bar{x} \approx T_i = 0 \\ Y_i(1) & \text{if } X_i \geq \bar{x} \approx T_i = 1 \end{cases} \tag{21}$$

Where $Y_i(0)$ and $Y_i(1)$ are the potential outcomes of Y_i. Notice that we are interested in the sharp average treatment effect at the threshold, which is given by:

$$\tau = \mathbb{E}[Y_i(1) - Y_i(0)|X_i = \bar{x}] \tag{22}$$

Finally, we define an estimator of τ through kernel-based local polynomials on either side of the cut-off, as in Hahn, Todd and Van der Klaauw (2001) and Porter (2003). Specifically, the local polynomial based p-degree τ estimator is:

$$\hat{\tau}_p(h_n) = \hat{c}_{+,p}(h_n) - \hat{c}_{-,p}(h_n) \tag{23}$$

where $\hat{c}_{+,p}(h_n)$ and $\hat{c}_{-,p}(h_n)$ are the intercepts of a weighted polynomial regression at the cut-off $\bar{x}$ for control $(-)$ and treated $(+)$ groups with bandwidth h_n. Across all RD experiments, we build the polynomial estimator by using a uniform kernel function. This function assigns equal weighting to all $X_i \in [\bar{x} - h, \bar{x} + h]$.

3.3.2 Results from Regression Discontinuity Design

We estimate the effect of the policy change on interest rates at the 80% LTV cut-off ($\bar{x}$ = 80%). This result of this exercise is key for distinguishing between a theory in which imperfect information plays a role –like the model we presented in section (2)–, and one that does not, and in which banks can price (interest-rate)-discriminate leveraged households. Besides, as we have established previously, mortgages related to government housing subsidy policies (which represent a small and non-random fraction of all granted loans) defined as those less or equal to UF 500 were not affected by the policy change. Then, we can exploit this cut-off and use the loan amount (house price) as a running variable with $\bar{x}$ = UF 500 in a second RD experiment.

As in the previous subsection we use both the original data, as well as the balanced sample to account for confounding factors. Notably, in these RD exercises, the definition of the control group differs from that in the previous exercises. In particular, we use only data of loans granted after the change in regulation came into force. Then the reason to use the balanced sample aims to address a potential selection bias problem: if banks react on the extensive margin and cherry-pick borrowers, it could be the case that average interest rates are not representative for the sample of borrowers before the regulation came into force. By using the balanced sample, we are only considering individuals who were better representatives of the pre-regulation-change sample; thus alleviating potential selection bias.

Next, we present our results. We fit a $p = 4$ degree polynomial and a uniform kernel function, for the RD estimation. Based on our exercise that uses the loan-to-value ration as a *running* variable, we estimate a small but statistically significant sizes of the sharp average treatment effect $\hat{\tau}$. Specifically, our calculations show $\hat{\tau} = 0.1681$ when using the original data, and $\hat{\tau} = 0.1564$ when employing the balanced CEM sample. Notably, the estimate $\hat{\tau}$ becomes smaller once we can account for confounding variables through the use of the balanced sample defined in section 3.1. Graphically, we illustrate these results in Figure (10), where we can observe a small vertical discontinuity at the LTV 80% threshold. In contrast, when we set house price as *running* variable, our findings show the estimates of the sharp average treatment effect $\hat{\tau}$ to be statistically non-significant for both, original data and balanced CEM samples. Figure (11) illustrates this result, where there is no evidence of discontinuity in mortgage interest rates at house price threshold $\bar{x}$ = UF 500.

While there is evidence of a discontinuity in the LTV ratio at the threshold of $\bar{x}$ = 80%, the implied interest rate is only $\sim 0.16\%$ higher for leveraged households which contrasts

Panel (a): Raw Samples *Panel (b): Balanced CEM Samples*

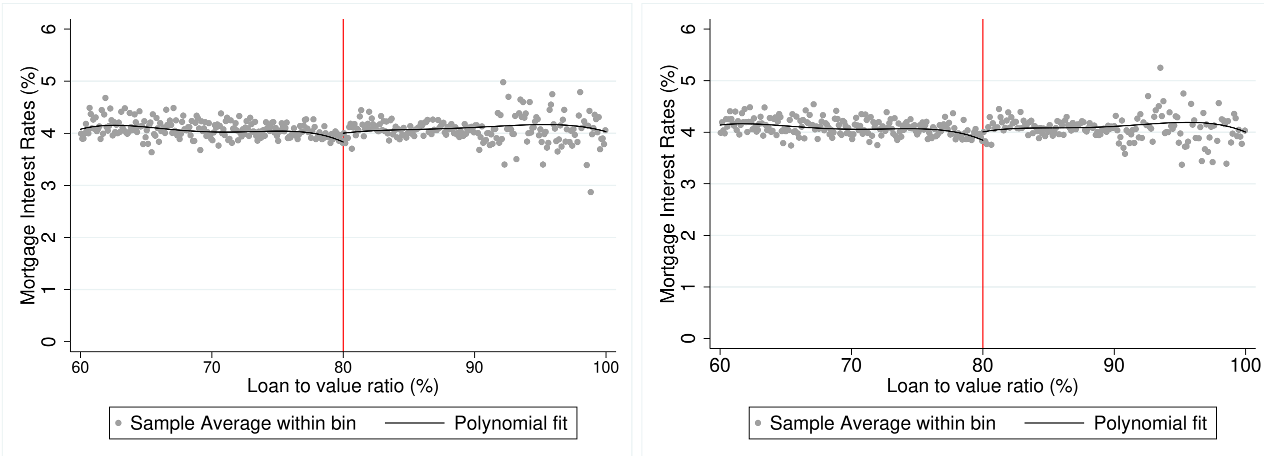


Figure 10: **RD Plots, Mortgage Interest Rates**: The figure illustrates the local polynomial approximation (solid black line) at the LTV 80% cut-off (solid red vertical line). Panel (a) uses the original data and Panel (b) the matched CEM samples. Source: Own calculations based on IRS data.

Panel (a): Raw Samples *Panel (b): Balanced CEM Samples*

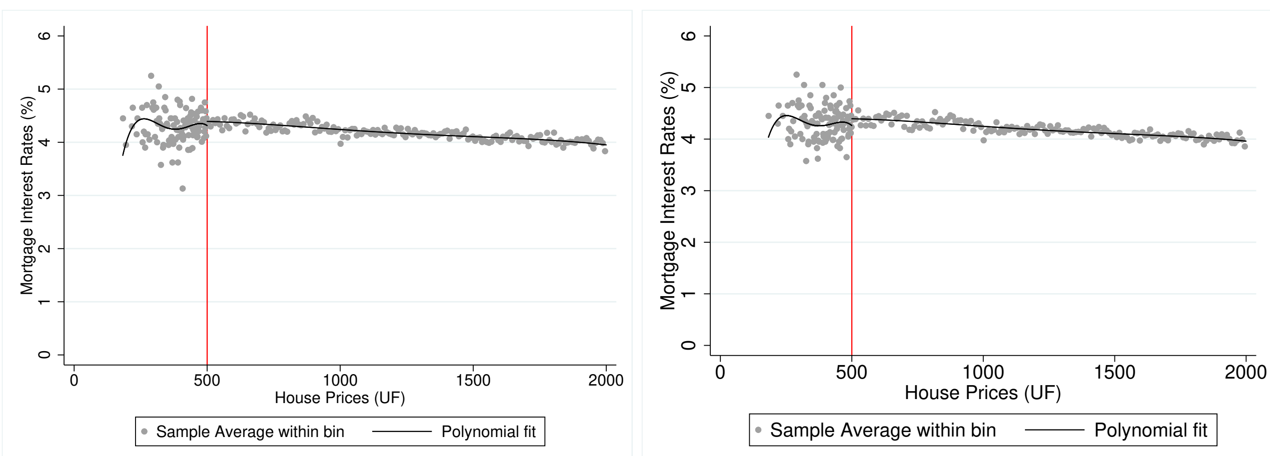


Figure 11: **RD Plots, Mortgage Interest Rates**: The figure illustrates the local polynomial approximation (solid black line) at the loan amount UF500 cut-off (solid red vertical line). Panel (a) uses the original data and Panel (b) the matched CEM samples. Source: Own calculations based on IRS data.

the $\sim$ 20% extra provision required by the new regulation. Even more, Figure (10) shows that the RD difference is not because charged interest rates for leveraved households is actually higher than for those with lower than 80% LTV ratios. In particular, borrowers with LTV ratios of less than 78% face the same interest rate as those with higher than 80% LTV ratios.

Table 7: **RD Design: Estimates**

X_i	Raw Samples			Balanced CEM Samples		
	τ	Standard Error	$P\|z\|$	τ	Standard Error	$P\|z\|$
LTV	0.1681	0.0168	0.0000***	0.1564	0.0182	0.0000***
Loan Amount	-0.0107	0.1361	0.9370	-0.0126	0.1412	0.7550

Note: In this table we show the main results of the regression discontinuity design experiments using two different samples. The estimation process was carried out using the utility program in Calonico (2014). Inference: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

4 Conclusion and road ahead

In December 2014 the Chilean Banking Supervisor announced that within a year, supervised institutions would need to effectively raise provisions for credit risk of mortgage loans, to match expected loss according to unified criteria. In this paper, we have analyzed the effect of such a change in regulation on the mortgage credit market. Notably, this new regulation raises required provisions contingent on leverage at the moment of *ex post* realized delinquency. Also, compared to previous regulation, the higher financial cost for banks is substantial.

We offer evidence that, as a consequence of the regulation, granted loans were on average, lower as a fraction of the value of pledged collateral. We do so by developing a small screening-under-imperfect-information model about borrowers' quality. In the said model, the introduction of higher provisioning cost, contingent on *ex-post* borrower payment behavior, and borrowers' leverage at the moment of delinquency affects the *ex-ante* screening of loan applicants by financial institutions. The LTV ratio is an informative but imperfect signal of borrowers' quality, hence financial institutions can use it to screen borrowers. By incorporating the features of the regulation into the model, we can generate an endogenous LTV limit, which helps us rationalize a clear bunching of loans in the data; which we otherwise could not. Equipped with the model we can more carefully examine the data.

We use novel and unique data from administrative records, collected by the Internal Revenue Service. Our data spans all transactions of real estate in Chilean territory from the year 2002 to present, though we focus our analysis in years 2012-17. We have access to buyers, sellers, and real estate characteristics. Using a matching algorithm and a bunching procedure we seek to evaluate the effect of the regulation on realized LTV ratios. We conclude that quantitatively the regulation had an effect: banks accommodated it by granting smaller loans as a fraction of pledged collateral. We estimate that, after the regulation came into force, the average granted LTV ratio is 2.8% lower. Also, for the median borrower, it is 9.8% lower. We also document that because of the calibration of the regulation, a large fraction of loans is granted at exactly 80% LTV. In particular, the fraction of loans granted at exactly 80% has more than tripled and represents now one-fourth of all loans. This is precisely the sort of bunching our theoretical model predicts and the one we confirm through a bunching estimation[14]. Besides, we merge our data on real estate transactions with administrative data on all loans contracts collected by the SBIF, and show by using regression discontinuity design techniques that higher mortgage interest rates because of the regulation was not an equilibrium outcome.

This paper left out other potential information sources of borrower quality (e.g. credit scores, alternative collateral, past behavior on loans with the same banking institution). We have assumed in our model that, besides all the observable characteristics we detailed in the text, the only other signal a borrower can provide is the size of the down-payment relative to the value of the property. This is clearly an abstraction. A prediction of the model, though, is that if there is a higher cost on having a low-quality borrower, on average the quality of the portfolio should be better after the introduction of the regulation. Unfortunately, evidence of ex-post delinquency rates is not observable just yet. This prediction could be tested in a few more years when enough time has passed to allow low-quality borrowers to enter into arrears.

[14] An interesting question is how many people are ousted of the market with this regulation? We know from other administrative data from the Superintendency of Banks and Financial Institutions, that the number of granted loans was lower in 2016 than before. Is this all to be attributed to the new regulation? The answer, is most likely, no. There is a myriad of potential explanations (from demographics to house prices) that contaminate such a simple answer. Our model cannot answer this question, as it would require a richer structure with housing demand in which a household could decide whether to raise the down-payment to value ratio to obtain a loan for buying a home instead of renting, or simply wait for another period to save more. This inter-temporal decision is completely abstracted from our simple model.

References

Acharya, Viral V, Katharina Bergant, Matteo Crosignani, Tim Eisert and Fergal McCann. 2018. "The Anatomy of the Transmission of Macroprudential Policies.". 5

Albanesi, Stefania, Giacomo De Giorgi and Jaromir Nosal. 2017. Credit growth and the financial crisis: A new narrative. Technical report National Bureau of Economic Research. 4

Albertazzi, Ugo, Margherita Bottero and Gabriele Sene. 2017. "Information externalities in the credit market and the spell of credit rationing." *Journal of Financial Intermediation* 30:61–70. 34

Almus, Matthias and Dirk Czarnitzki. 2003. "The effects of public R&D Subsidies on Firms' Innovation Activities: the Case of Eastern Germany." *Journal of Business & Economic Statistics* 21(2):226–236. 20

Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press. 29

Ates, Sina T and Felipe Eduardo Saffie. 2013. "Project Heterogeneity and Growth: The Impact of Selection." Unpublished, University of Pennsylvania . 9, 18

Banco Central de Chile. 2017. "Chapter 2: Local Financial Markets." *Informe de Estabilidad Financiera* 1. 6, 18

Beltratti, Andrea, Matteo Benetton and Alessandro Gavazza. 2017. "The role of prepayment penalties in mortgage loans." *Journal of Banking & Finance* . 4

Ben-David, Itzhak. 2011. "Financial constraints and inflated home prices during the real estate boom." *American Economic Journal: Applied Economics* 3(3):55–87. 2, 4

Calonico, Sebastian, Matias D Cattaneo and Rocio Titiunik. 2015. "Optimal data-driven regression discontinuity plots." *Journal of the American Statistical Association* 110(512):1753–1769. 34

Cattaneo, Matias D, Brigham R Frandsen and Rocio Titiunik. 2015. "Randomization inference in the regression discontinuity design: An application to party advantages in the US Senate." *Journal of Causal Inference* 3(1):1–24. 34

Cerutti, Eugenio, Jihad Dagher and Giovanni Dell'Ariccia. 2017. "Housing finance and real-estate booms: a cross-country perspective." *Journal of Housing Economics* . 4

Cerutti, Eugenio, Stijn Claessens and Luc Laeven. 2017. "The use and effectiveness of macroprudential policies: new evidence." *Journal of Financial Stability* . 3, 4

Chava, Sudheer and Michael R Roberts. 2008. "How does financing impact investment? The role of debt covenants." *The journal of finance* 63(5):2085–2121. 34

Chetty, Raj, John N Friedman, Tore Olsen and Luigi Pistaferri. 2011. "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records." *The quarterly journal of economics* 126(2):749–804. 3, 30, 31, 33

Crowe, Christopher, Giovanni Dell'Ariccia, Deniz Igan and Pau Rabanal. 2013. "How to deal with real estate booms: Lessons from country experiences." *Journal of Financial Stability* 9(3):300–319. 4

Encina, Jenny. 2013. "Pension Reform in Chile: A Difference in Difference Matching Estimation." *Estudios de Economía* 40(1). 20

Epure, Mircea, Irina Mihai, Camelia Minoiu and José-Luis Peydró. 2018. "Household credit, global financial cycle, and macroprudential policies: credit register evidence from an emerging country." *Mimeo* . 5

Field, Erica and Maximo Torero. 2006. Do property titles increase credit access among the urban poor? Evidence from a nationwide titling program. Technical report Working paper, Department of Economics, Harvard University, Cambridge, MA. 20

Galiani, Sebastian, Paul Gertler and Ernesto Schargrodsky. 2005. "Water for life: The Impact of the Privatization of Water Services on Child Mortality." *Journal of Political Economy* 113(1):83–120. 20

Hahn, Jinyong, Petra Todd and Wilbert Van der Klaauw. 2001. "Identification and estimation of treatment effects with a regression-discontinuity design." *Econometrica* 69(1):201–209. 34

Heckman, James J, Hidehiko Ichimura and Petra E Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *The Review of Economic Studies* 64(4):605–654. 20

Heckman, James J, Hidehiko Ichimura and Petra Todd. 1998. "Matching as an Econometric Evaluation Estimator." *The Review of Economic Studies* 65(2):261–294. 20, 21

Ho, Daniel E, Kosuke Imai, Gary King and Elizabeth A Stuart. 2007. "Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference." *Political analysis* 15(3):199–236. 21, 22

Hott, Christian. 2015. "A model of mortgage losses and its applications for macroprudential instruments." *Journal of Financial Stability* 16:183 – 194.
URL: *http://www.sciencedirect.com/science/article/pii/S1572308914000576* 4

Iacus, Stefano M, Gary King and Giuseppe Porro. 2012. "Causal Inference without Balance Checking: Coarsened Exact Matching." *Political Analysis* 20(1):1–24. 3, 20, 22, 24

Imbens, Guido W and Jeffrey M Wooldridge. 2009. "Recent developments in the econometrics of program evaluation." *Journal of economic literature* 47(1):5–86. 34

IMF. 2011. "Housing Finance and Financial Stability – Back to Basics. Global Financial Stability Report, IMF.". 4

Jalan, Jyotsna and Martin Ravallion. 2003. "Estimating the Benefit Incidence of an Antipoverty Program by Propensity-Score Matching." *Journal of Business & Economic Statistics* 21(1):19–30. 20

King, Gary, Matthew Blackwell, Stefano Iacus and Giuseppe Porro. 2010. "cem: Coarsened exact matching in Stata." *Stata Journal* . 23

Kinghan, Christina, Paul Lyons and Yvonne McCarthy. 2017. Macroprudential Measures and Irish Mortgage Lending: Insights from H1 2017. Economic letters Central Bank of Ireland. 5

Kuttner, Kenneth N and Ilhyock Shim. 2016. "Can Non-Interest Rate Policies Stabilize Housing Markets? Evidence from a Panel of 57 Economies." *Journal of Financial Stability* 26:31–44. 4

Lechner, Michael. 2002. "Some Practical Issues in the Evaluation of Heterogeneous Labour Market Programmes by Matching Methods." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 165(1):59–82. 20

Mas-Colell, Andreu, Michael Whinston and Jerry Green. 1995. *Microeconomic Theory*. Oxford University Press.
URL: *http://EconPapers.repec.org/RePEc:oxp:obooks:9780195102680* 9

Matus, José Miguel. 2015. "Provisiones por Riesgo de Crédito de la Banca Nacional: Análisis de los Cambios Normativos, Período 1975-2014." Studies in Economic Statistics, Central Bank of Chile . 5

Moser, Petra. 2005. "How do Patent Laws Influence Innovation? Evidence from Nineteenth-Century World's Fairs." *The American Economic Review* 95(4):1214–1236. 20

Pacheco, David, Carlos Pugar and Eduardo Valdebenito. 2014. Modelo Estándar de Provisiones para Cartera Hipotecaria Residencial. Technical report. ”Serie de Estudios Normativos”. 3, 18, 28

Park, Kevin A. 2016. “FHA loan performance and adverse selection in mortgage insurance.” *Journal of Housing Economics* 34:82–97. 20

Porter, Jack. 2003. “Estimation in the regression discontinuity model.” *Unpublished Manuscript, Department of Economics, University of Wisconsin at Madison* 2003:5–19. 34

Qi, Min and Xiaolong Yang. 2009. “Loss given default of high loan-to-value residential mortgages.” *Journal of Banking & Finance* 33(5):788–799. 4

Saez, Emmanuel. 2010. “Do taxpayers bunch at kink points?” *American economic Journal: economic policy* 2(3):180–212. 3, 30

Scott, David W. 2015. *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons. 23

Smith, Jeffrey A and Petra E Todd. 2005. “Does Matching Overcome LaLonde’s Critique of Nonexperimental Estimators?” *Journal of Econometrics* 125(1):305–353. 20, 21

Stiglitz, Joseph E and Andrew Weiss. 1981. “Credit rationing in markets with imperfect information.” *The American economic review* 71(3):393–410. 9

Thistlethwaite, Donald L and Donald T Campbell. 1960. “Regression-discontinuity analysis: An alternative to the ex post facto experiment.” *Journal of Educational psychology* 51(6):309. 3, 34

Villa, Juan M. 2016. “diff: Simplifying the estimation of difference-in-differences treatment effects.” *The Stata Journal* 16(1):52–71. 30