Irving Fisher Committee on
Central Bank Statistics

BANK FOR INTERNATIONAL SETTLEMENTS

IFC workshop on *"Combining micro and macro statistical data for financial stability analysis. Experiences, opportunities and challenges"*

Warsaw, Poland, 14-15 December 2015

A micro-powered model of mortgage default risk for full recourse economies, with an application to the case of Chile[1]

Diego Avanzini, Juan Francisco Martínez and Víctor Pérez,
Central Bank of Chile

[1] This paper was prepared for the meeting. The views expressed are those of the authors and do not necessarily reflect the views of the BIS or the central banks and other institutions represented at the meeting.

A micro-powered model of mortgage default risk for full recourse economies, with an application to the case of Chile*

Diego Avanzini† — Central Bank of Chile

Juan Francisco Martínez — Central Bank of Chile

Víctor Pérez — Central Bank of Chile

October 21, 2015

Abstract

This paper develops a customized model of mortgage loans default for a full recourse economy. We combine the more usual analysis for nonrecourse economies, adding a non-pecuniary cost for defaulting in order to account for possible loss of utility due to the full-recourse framework. This model applies to economies such as Spain, Australia, and Chile, where defaulters can be prosecuted until their debts are completely settled. Under the proposed model, we obtain an analytical expression involving default determinants for mortgage loans, both micro and macro. As a case study, we estimate

*This paper has been prepared for the Workshop on Housing, Santiago, April 25, 2014, at the Central Bank of Chile. The views and conclusions presented in this paper are exclusively those of the authors and do not necessarily reflect the position of the Central Bank of Chile or of the Board members. We are grateful to Natalia Gallardo for helping us and contributing in a preliminary stage of this work. All errors are our own.

†Corresponding author: davanzini@bcentral.cl, Statistical Division, Central Bank of Chile.

this relationship for the Chilean economy using information from the Chilean Survey of Household Finance (EFH). As stated by our micro-macro model, we find that household financial conditions and their interactions with systemic determinants account for an important part of the cross sectional probability of mortgage default.

Keywords: Default, Mortgage loan, Survey data, Full recourse economy, Rare events, Credit market

JEL Classification Number: C35, D53, E44, G21

1 Introduction

The recent financial crisis has tested the academics and regulators in their ability to anticipate possible sources of instability. Some of the lessons that can be learnt from the latest financial crisis is that policy models and their estimates should take into account factors that might have been at some point overlooked. The recent experience suggests the existence of a mortgage channel that acts as an accelerator or amplifier element of financial instability. This is why, for central banks and policy makers, it is crucial to understand the mechanisms at place, and to use richer sources of data to measure them. Hence, in this work we contribute to the existing literature by presenting a theoretical model that incorporates both micro and macro financial determinants and their interactions for a full recourse economy. We apply the framework to the Chilean economy as a case study, and test it empirically using using a rich source of micro-data coming from the Chilean Survey of Household Finance.[1]

It is widely accepted in the current economic literature, that –in aggregate terms– a mortgagor facing a negative equity is prone to default. In this case, when the value of debt is higher than the value of the collateral, there are incentives to optimally decide not to repay, since it is cheaper to give up the property and to stop repaying. This mechanism is in the heart of various recent crisis related papers such as those of Fostel and Geanakoplos (2008) and Goodhart et al. (2011). Other influential papers that model a collateral financial economy are Kiyotaky and Moore (1997) and Bernanke et al. (1999), but instead of households, these works model firms dynamics under the presence of a credit market affected by asymmetric information. In their settings, the risk premium depends on the firms' probabilities of default, which in turn are related to the value of net worth.

The model that inspires our treatment of the default phenomenon is developed in Geanakoplos and Zame (2013).[2] They abstract from the explicit modeling of asymmetric information,

[1] From now on "EFH", for its name in Spanish ("Encuesta Financiera de Hogares"). The survey is conducted by the Central Bank of Chile (BCCh) since 2007.

[2] Note that this is a published version of a working paper that was first issued in 1997.

but allow for an equilibrium with endogenous default. To do this, they base their model – as all the theoretical literature in this field– on a *non-recourse* credit regulation policy. Non-recourse regulatory frameworks state that the lender gets the collateral only after the occurrence of a default and it is not allowed to prosecute other payments or compensations from the debtor. In this way, modelers avoid the imposition of deadweight losses associated to other punishments in case of default.[3]

However, when modeling the determinants of mortgage default for economies such as Australia, Chile, and Spain, we need to introduce the case of *full recourse* regulation. Under this type of regulation, agents can be prosecuted until their debts are completely settled. This implies that the incentives to repay should be enhanced, since in a *full recourse* framework, the incentive to default due to negative equity is counteracted by the fact that if the collateral is not sufficient to cover the promises, the creditor has the right of going after other assets of the debtor. There is also a reputational effect given that the prosecuted debtors are publicly listed. All these issues motivate us to develop a modeling alternative that adapts the seminal framework of Geanakoplos and Zame (2013).

Economic intuition indicates that our model should incorporate a mechanism that additionally discourages mortgagors from defaulting. This is consistent with the idea that mortgagors do not always default when they face negative equity on their homes as in Harrison et al. (2004) or Ellul et al. (2010). In fact, Foster and Van Order (1984) and Bhutta et al. (2010) find that many borrowers with negative equity do not default; and, conversely, default is often associated with shocks, such as unemployment. Also, the cost of continuing to repay a mortgage also depends on the agent's idiosyncratic discount factor and thus on his liquidity position, as Elmer and Seelig (1999), Gerardi et al. (2007), and Bajari et al. (2010) explain.

Our model formulation closely follows part of the Goodhart et al. (2011) work, which

[3] As Geanakoplos and Zame (2013) state, the only seizure of collateral "[...] avoids the moral and ethical issues of imposing penalties in the event of bad luck." Although, they recognize that in practice, the seizure of collateral implies deadweight losses on its own.

is a recent application based in the Geanakoplos and Zame (2013, and the earlier work dated 1997) model. Yet, in order to modify the mortgage default incentives, we add a non-pecuniary cost that arises as a reputational loss because of the burden of being enforced to repay the debt and other possible losses of utility (e.g. being blacklisted from the credit market or being banned for future credit opportunities). Natural references that we consult to address this type of default cost are those of Shubik and Wilson (1977) and Dubey et al. (2005). Together with Geanakoplos and Zame (2013), these works lie within the general equilibrium literature. Our contribution consists in combining both frameworks to generate a full recourse economy, such as that of Spain, Australia and Chile.

To test our framework, we apply it to the Chilean case. The literature that models default decisions in the Chilean economy is rather brief. One of the recent investigations appears in Alfaro and Gallardo (2012). These authors estimate an empirical characterization using the EFH 2007 survey. They find that income and income-related variables are the only significant and robust factors that explain default for both types of debt (consumer and mortgage). Additionally, demographic variables can help to further explain default probabilities.

Empirical evidence suggests that there is some rich information hidden away in aggregated (or macro) data. Thus, the use of micro-data seems a promising research avenue. That is why, instead of using financial aggregates of credit and macro-financial conditions to determine default dynamics through time, we follow Alfaro and Gallardo (2012) in the sense that we use micro-information coming from the EHF survey. As compared to the latter, beyond contributing with a theoretical modeling framework, in this work we are able to profit from the additional information that the EFH survey has gathered through the years since it was launched in 2007.

The main objective of this work is to improve the identification and estimation of default determinants in the mortgage market. Our results can contribute to support credit policy measures and also enhance the analysis and application of mortgage banking stress tests. In particular, we apply this methodology to the Chilean economy, but it may be relevant to other countries, such those already mentioned.

In recent years, theoretical advances and economic necessity have stimulated the study of the default decision specially on mortgage loans. Since Kau et al. (1994) and Capozza et al. (1997), theoretical models and empirical testing have being integrated in the search for a better understanding of the phenomenon. Following those efforts we elaborate on the theoretical modeling of mortgage default and estimate the resulting model using microdata from households in Chile.

The rest of the work is organized as follows. Section 2 introduces the theoretical model. Section 3 highlights the role of the determinants as they emerge from the model. In Section 4 we describe the empirical methodology as well as the dataset we use for the estimations. All the results are presented in Section 5. Some final remarks are in Section 6. The Appendix contains derivations and proofs from the model, a supporting glossary, and some data description.

2 A theoretical model of the determinants of default

To provide a general view of the mortgage default mechanism at play, in this section we introduce and describe a partial equilibrium model that motivates our empirical estimations. We describe the economy, agents, financial structure, transactions and its timing.

Our model considers a two-period setup. The beginning ($t = 0$) is deterministic, but the second period ($t = 1$) is stochastic. In the latter, there are s possible states of nature that occur with probability π_s.[4] To help notation, we also define $S = \{0, s\}$, a $t-$uple that groups both periods and consider the possible states at the second period.

The economy is composed by one representative household. There are two goods traded in this economy: a consumption commodity and housing. We also have two types of financial assets: one short-term (intra-period) unsecured loan and one long-term (inter-period)

[4] Later we will specify that state variables are in our case represented by commodity endowments and liquidity conditions (i.e. interest rates and monetary endowments).

collateralized loan.

The household is endowed with a perishable commodity that consumes and sells, and money. The endowment is deterministic in the first period, and depends on the realized state of nature in the second period. The agent also consumes housing in both periods that has to be purchased. There are two differences between commodities and housing. On the one hand, commodities are perishable (they only provide utility in the period they are available) while durable goods (house) purchased at the initial period ($t = 0$) also give utility at the final one ($t = 1$). On the other hand, housing also serves as collateral when contracting a long term loan.

2.1 Financial intermediation

The financial intermediation in this model is assumed to be exogenous and it is treated in a stylized manner. However, it can be thought of as a representative bank (or banking system) that lends short-term to the representative household,[5] as well as an inter-period collateralized mortgage loan. The latter is only available at $t = 0$. An exogenous amount of liquidity is pumped into the system and it is reflected in the short- and long-term interest rates (r_S and $\bar{r}$, respectively).[6] Nominal flows of the economy are depicted in Figure 1.

2.2 Timing

At the initial period ($t = 0$) short-term liquidity borrowing plus long-term mortgage borrowing from the representative household occur. Provided the financing, transactions in commodity and mortgage markets take place. The representative agent sells part of his endowment of the consumption commodity and uses the revenues to repay the short-term loan obligations. Under the current setup, short-term obligations arise as a consequence of

[5] To help the focus of the analysis and without loss of generality, this type of loan is assumed to be default-free.

[6] It can be thought of as the monetary authority that injects certain amount of resources in the money market, influencing the short-term rates (e.g. through M_S) and long-term interest rates (e.g. through $\bar{M}$).

the assumption of a cash-in-advance financial economy. The agent also purchases housing, which is financed with a collateralized long-term mortgage loan. Finally, consumption of housing and commodity takes place.

During the second period ($t = 1$) the nature realizes and reveals s, which mainly affects endowments –including monetary– and monetary supply, as we observed before. At this period, the intra-period transactions work in a similar manner than at the beginning. However, there are slight, but significant differences. First, it is easy to see that the long-term interest rate is already fixed (e.g. because the long-term money market does not open at every period). Second, the mortgage loan is settled by the representative household, given by his delivery (or default) decision. At this period, he uses the short-term money market to repay the long-term loans and finances his housing consumption.

2.3 Economy and budget set

The economy is composed by a positive **endowment of** the perishable **commodity**, e_{S1}; positive **monetary endowment** m_S; **preferences** U; **financial assets** consisting of long and short-term loans at rates $\bar{r}$ and r_S, respectively. We assume that in this economy preferences are monotonic and utility is quasi-concave.

Given a set of securities (i.e. short and long-term loans) Λ; commodity price p_{S1}; housing price p_{S2}; rate of return of the short and long-term credit market r_S and $\bar{r}$, respectively; the mortgagor with commodity endowment e_{S1} makes consumption plans, credit demand, sales of commodity, and deliveries against promises.[7] Following Geanakoplos and Zame (2013), in the case of mortgages, optimal deliveries will be always the minimum of promises (i.e. credit demand) and the value of the collateral. Thus, we define the budget set for the representative household as $B(p_1, p_2, r_S, \bar{r}, e_{S1}, \Lambda)$ to be the set of plans $(\mu, \bar{\mu}, q_{S1}, b_{S2})$, where μ stands for the short term borrowing, $\bar{\mu}$ is the mortgage borrowing, q_{S1} represents the commodity sales

[7] Our notation for prices, quantities and monetary expenditures contains a subindex with two components. The first one is the period (or state at $t = 1$), while the second term indicates the type of good, 1 for commodities and 2 for housing.

and b_{S2} stands for housing expenditure.

We have that at the initial date ($t = 0$), the agent is subject to three budget constraints. First, the short term loans must not exceed the revenues from commodity sales. Second, the housing expenditure must be lower than or equal to its long and short term credits and monetary endowment. Finally, the third constraint considers an element of mortgage credit regulation. There is a fixed portion of collateral (i.e. ϕ) required for a mortgage loan.[8]

Formally, we have that, at $t = 0:$

$$\begin{aligned} \mu_0 &\leq p_{01}q_{01} \\ b_{02} &\leq \frac{\mu_0}{1+r_0} + \frac{\bar{\mu}}{1+\bar{r}} + m_0 \\ \frac{\bar{\mu}}{1+\bar{r}} &\leq \phi b_{02} \end{aligned}$$

Conversely, at the final date ($t = 1$), the agent is subject to two budget constraints. The first constraint reflects the fact that the short term loans must not exceed the revenues from commodity sales. The second budget constraint depends on the state of nature that is realized. Thus, under a good state of nature (i.e. $s \in S_G$, where no default is exerted), the constraint states that the repayment of the mortgage loans plus the new housing consumption of the household (i.e. a type of housing top-up) must not exceed the agent's short-term borrowing and monetary endowment. Under a bad state of nature (i.e. $s \in S_B$, there is default), the constraint accounts for the fact that there is no mortgage repayment.

Formally, at $t = 1$, the constraints are:

$$\begin{aligned} \mu_S &\leq p_{s1}q_{s1} \\ b_{s2} + \bar{\mu} &\leq \frac{\mu_s}{1+r_s} + m_s \qquad \forall s \in S_G \\ b_{s2} &\leq \frac{\mu_s}{1+r_s} + m_s \qquad \forall s \in S_B \end{aligned}$$

[8] Without loss of generality, we assume that this constraint is always binding in equilibrium.

2.4 Preferences

At the beginning of the time ($t = 0$), the representative household maximizes his expected utility subject to the budget constraints previously described. The agent's utility in our case has the particularity that default incentives are mitigated by seizing the collateral of the mortgagor.[9] However, to adequate the problem to a *full recourse* economy, we also introduce a non-pecuniary default penalty proportional to the defaulted amount, similar to Shubik and Wilson (1977) and Dubey et al. (2005). Thus –subject to his budget set–, the agent solves the following program:

$$\begin{aligned}\max_{\mu_s,\bar{\mu},b_{s2},q_{s1}} U \quad = \quad & u\left(e_{01}-q_{01}\right)+u\left(\frac{b_{02}}{p_{02}}\right)+\sum_{s\in S}\pi_s\left\{u\left(e_{s1}-q_{s1}\right)\right\}+ \\ & +\sum_{s\in S_G}\pi_s\left\{u\left(\frac{b_{02}}{p_{02}}+\frac{b_{s2}}{p_{s2}}\right)\right\}+\sum_{s\in S_B}\pi_s\left\{u\left(\frac{b_{s2}}{p_{s2}}\right)\right\}- \\ & -\lambda\sum_{s\in S}\pi_s\max\left\{\left(1-\frac{b_{02}p_{s2}}{p_{02}\bar{\mu}}\right),0\right\}\end{aligned}$$

That is, the agent optimizes his utility function, by solving his plan of short and long-term borrowing, housing expenditure (in monetary terms) and commodity sales (quantity), subject to his budget constraints. At the first period ($t = 0$), the agent obtains utility from consuming the commodity that remains from his sales and the housing units that he purchases. At the final period ($t = 1$), the representative household faces the two possible sets of states of nature, good or bad. If a good state realizes, there is no default. Thus, the agent enjoys from consuming his initial date housing consumption plus the additional housing he buys in the final period. Conversely, if the household faces a bad state of nature, he (optimally) defaults.[10] In such cases, the (non-modelled) banking sector seizes the collateral, so the agent cannot benefit from consuming the initial period housing and he is subject to

[9] We interpret this as an exogenous commercial banking sector that is compensated by receiving the collateral. In turn, it sells it back into the market to cash-in.

[10] As we previously mentioned, the optimal delivery is the minimum between the current value of collateral and the promises (i.e. $\min\{\bar{\mu}, (b_{02}p_{s2}/p_{02})\}$).

a default penalty (linearly) proportional to the defaulted amount. It has to be noted that in this model, the agent is not expelled from the market. He is still allowed to obtain short term credit.[11]

2.5 Equilibrium

This section provides a brief note on equilibrium. From the problem formulation, we have that this economy is composed only by the representative household. Here we assume that he behaves in a competitive manner, so takes prices as given.

The competitive collateral equilibrium (CCE) in this economy consists of a set of commodity prices (p_{S1}), security prices (interest rates r_S, $\bar{r}$) and consumption plans ($\mu, \bar{\mu}, q_{S1}, b_{S2}$), such that short-term commodity market clears, durable housing market clears, security markets clear, plans are budget feasible and consumers optimize.

There are important particularities to be highlighted in this type of economy. On the one hand –in this partial equilibrium model–, we abstract from a housing pricing amplification mechanism first described by Fisher (1933) and, more recently, by Goodhart et al. (2011). In these works, if a bad state is attained and there is mortgage default, the representative household will have to give-up the collateral and the housing goes back to the market. In such a situation, the housing market clearing condition incorporates that there is a decrease in prices that further increase default incentives, given that collateral losses value, and the debt face value remains. This spiral-type mechanism is called debt-deflation.

Another issue to be considerer concerns to rational expectations. In this work we also abstract from explicitly modelling the financial (banking) sector. Thus, we do not need to consider the rational expectations of the lending activity. Nevertheless, in cases where the banking sector appears (e.g. Goodhart et al., 2011), and to incorporate the default possibility under rational expectations, the effective return on the mortgage will be the ratio between: (i) the minimum between the collateral value at $t = 1$ and the mortgage promise, $\bar{\mu}$; and (ii)

[11] The agent's program first order conditions appear in the Appendix.

the initial mortgage extension (credit supply). That is, a bank should rationally anticipate the optimal household delivery and account for this into the risk premium.

Theorem 1 Existence: *Assuming rational expectations, a well-behaved (e.g. normally distributed) state probability space, monotonic preferences, quasi-concave utility (quasi-concavity is assumed to be maintained in the presence of linear penalties) and convex budget sets, and provided that all the markets clear, this economy admits a competitive equilibrium.*

Proof. It appears in Geanakoplos and Zame (2013). ■

3 Determinants of mortgage default

Proposition 2 Existence of full-recourse determinants of mortgage default: *From the household optimization problem solution, and assuming without loss of generality that there are only two possible states of nature (the good G, where there is full delivery, and the bad B, where there is default) we can obtain an equation that represents the mortgage default determinants. Furthermore, the default depends on idiosyncratic and systemic factors as well as their interactions. The explicit equation is given by:*

$$1-\frac{b_{02}p_{B2}}{\bar{\mu}p_{02}}=\omega_0+\omega_1\bar{\mu}u'\left(\frac{b_{02}}{p_{02}}\right)+\omega_2\bar{\mu}u'\left(\frac{b_{02}}{p_{02}}+\frac{b_{G2}}{p_{G2}}\right)+\omega_3\bar{\mu}u'\left(e_{01}-q_{01}\right) \tag{1}$$

where

$$\begin{aligned}
\omega_0 &= 1-\frac{\lambda^{\alpha}\pi_B p_{B2}}{p_{02}(1+\bar{r})\phi}\\
\omega_1 &= -\frac{1}{p_{02}\lambda\pi_B\left(1+\bar{r}\right)\phi}\\
\omega_2 &= \frac{\pi_G\left(\phi p_{02}\left(1+\bar{r}\right)-p_{G2}\right)}{p_{G2}p_{02}\lambda\pi_B\left(1+\bar{r}\right)\phi}\\
\omega_3 &= -\frac{\left(1+r_0\right)\left(1-\phi\right)}{p_{01}\lambda\pi_B\left(1+\bar{r}\right)\phi}
\end{aligned}$$

Proof. See the Appendix. ■

Notice that the left hand side of equation (1) corresponds to the expected default frequency on the mortgage loans (i.e. the complement of the expected repayment). The right hand side is composed by systemic and idiosyncratic factors, and their interactions. All of them constitute the determinants of mortgage default.

Remark 3 Systemic factors: *Notice that* ω_0*, the first term at the RHS of equation (1) accounts for the banking regulation and financing standards faced by the household. This is a* systemic *factor that affects households independently from their idiosyncracies (i.e. it is exogenous to them). It is easy to see that the default frequency decreases with: (i) a higher default penalty, (ii) a lower LTV* (ϕ)*, (iii) a higher interest rate, and (iv) a decrease in housing prices. Our interpretation is consistent with Jaffee and Stiglitz (1990) that the higher the screening, the lower the default incentives. On the other hand, a tighter regulation provides lower incentives to default, as in OECD (2011). Additionally, the terms* ω_1*,* ω_2 *and* ω_3 *depend only on exogenous variables or parameters. Thus, they also work as exogenous systemic factors that interact with agent's idiosyncracies.*

Remark 4 Systemic and idiosyncratic interaction – housing and mortgage demand in the first period: *The second term at the RHS of equation (1) shows a systemic factor interacted with credit demand and marginal utility of housing at* $t = 0$*. Given that* ω_1 *is always negative, we interpret that the higher the marginal utility of housing consumption and mortgage demand, the lower the incentives to default. Alternatively, given that we assume that preferences are quasi-concave, the higher the preference for housing consumption, the higher the expected mortgage default.*

Remark 5 Systemic and idiosyncratic interaction – housing and mortgage demand in the second period: *The third term at the RHS of equation (1) also contains the systemic factor interacted with the marginal utility, but in this case it is in the good state at* $t = 1$*. The systemic factor* ω_2 *in this case has an ambiguous sign. It is easy to see that if prices are in* *a growing phase*[12] *(i.e.* $p_{02} < p_{G2}$*) and credit supply standards are sufficiently relaxed (i.e.* $\phi(1+\bar{r}) > 1$*), then the dynamics of mortgage default and housing demand work in the same direction than in the second term of (1).*[13] *To sum up, under the specified conditions, higher mortgage or housing demands are associated to greater expected defaults.*

Remark 6 Systemic and idiosyncratic interaction - income: *The fourth term at the RHS of equation (1) works in a similar fashion with regard to the systemic component. There is a negative relationship between tighter (short-term) credit standards and mortgage default probabilities. The effect is proportional to the ratio between the housing and commodity current prices.*[14] *This result is conditioned on the marginal utility of consumption. Assuming risk aversion (e.g. logarithmic utilities), we have that –for a given mortgage credit standard and prices ratio– the higher the commodity endowment, the lower the mortgage default frequency. In this case, we interpret endowments as the availability of resources. Thus, it expresses the importance of housing income in this context. Our conjecture is that the combination of systemic credit standards and household income (or a related variable) is key to determine mortgage default frequency.*

From the findings that our theoretical model provides, we conjecture that mortgage default determinants are associated to credit cost and/or regulation standards and its interaction with idiosyncratic factors, such as the demand for housing and income. At this point, it is important to notice the interaction that we found. This will be key in our empirical specification which will be based on our theoretical model results.

4 Estimation Methodology and Data

In this section we briefly describe the problems that we face to obtain unbiased and efficient estimators, and the characteristics of the data used in the estimations.

[12] This ensures that the good state is attained.

[13] This is straightforward if we notice that $\omega_2 = \omega_1 p_{G2} \pi_G$.

[14] This is easy to see if we notice that $\omega_3 = \omega_1 (p_{02}/p_{01})(1+r_0)(1+\phi)$.

4.1 Estimation Methodology

In this work we explain the default decision of households on their mortgage loans based on idiosyncratic and systemic variables. For doing this, we use logistic regressions to estimate the behavior of a discrete variable representing the fact that some households in our sample have defaulted their mortgages.

Our data on mortgage delinquency present two important features we need to account for in the estimation. First, defaulting a loan (specially mortgages) is not an usual event. Futhermore, it can be considered as a rare event when compared with the occurrence of the contrary. Thus, if we take into account the whole population of outstanding mortgage loans in an economy, delinquency is highly underrepresented in any sample. However, our sample captures an important mass of delinquent loans in order to better appreciate its behavior. This means that the sample is intentionally biased to account for the phenomenon. This is known in econometrics as *choice-based* or *endogenous stratified* sampling. This sampling method is often supplemented with known or estimated prior knowledge of the population fractions of the rare event (as in our case).

Second, popular statistical procedures including the logistic regression, can sharply underestimate the probability of rare events. As King and Zeng (2002) point out, that logit coefficients are biased in small samples is well documented in the literature, but that in rare events data the biases in probabilities can be substantively meaningful is less known and understood. In particular, finite sample, rare event data have at least three features we should account for: (i) increasing the size of the sample does not alleviate the bias, because the rare events do not increase their share of the sample more quickly than the growth rate of the sample; (ii) the bias of the coefficients estimated using logistic regression when rare events are present is known to underestimate the probability of the rare event; (iii) the computation of probabilities of events in logit analysis is suboptimal in finite samples of rare events data, leading to errors in the same direction as biases in the coefficients.

To deal with these features together, we follow the developments of King et al. (2000) and King and Zeng (2001, 2002). Assume that we have a dataset of $i = 1, ..., n$ observations for which we have gathered information on the variable Y_i and a set of $k-1$ variables recorded in $\mathbf{x}_i$. The variable Y_i is a dicothomous variable that represents an underlying continuous variable Y_i^* following $Y_i = 1$ if $Y_i^* > 0$ and $Yi = 0$ if $Y_i^* \leq 0$. Y_i follows a Bernoulli distribution, while Y_i^* follows a Logistic distribution. The behavior of Y_i^* can be adequately inferred from the maximum likelihood estimation of the behavior of Y_i.

We assume that the *ones* ($Y_i = 1$) are the rare outcome, representing those households that hold a delinquent mortgage loan. The *prior population information* refers to the fraction of ones in the population, τ, while the *observed fraction of ones in the sample* refers to the sampling probability, $\overline{y}$.

To compensate for differences in the sample ($\overline{y}$) and population (τ) fractions of ones induced by choice-based sampling we use a weighting procedure. The procedure maximizes the weighted log-likelihood:

$$\begin{aligned} \ln L_w(\boldsymbol{\beta}|\mathbf{y}) &= w_1 \sum_{\{Y_i=1\}} \ln(\pi_i) + w_0 \sum_{\{Y_i=0\}} \ln(1-\pi_i) \\ &= -\sum_{i=1}^{n} w_i \ln\left(1+\exp\left[(1-2Y_i)\,\mathbf{x}_i\boldsymbol{\beta}\right]\right) \end{aligned} \tag{2}$$

where the weights are $w_1 = \tau/\overline{y}$ and $w_0 = (1-\tau)/(1-\overline{y})$, and where

$$w_i = w_1 Y_i + w_0(1 - Y_i)$$

The resulting $\hat{\boldsymbol{\beta}}$ is the *weighted exogenous sampling maximum-likelihood estimator,* due to Manski and Lerman (1977). One of the advantages of this method is that weighting can outperform other approaches (e.g. prior correction) when both a large sample is available and the functional form is misspecified (see Xie and Manski, 1988). However, weighting is asymptotically less efficient in small samples (see Scott and Wild, 1986; Amemiya and

Vuong, 1987).

When information about τ is available, we can remedy these problems using an analytical approximation to the bias based on McCullagh and Nelder (1989). The bias in $\hat{\boldsymbol{\beta}}$ can be estimated by the following weighted least-squares expression:

$$bias\left(\hat{\boldsymbol{\beta}}\right) = \left(\mathbf{X}'\mathbf{W}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{W}\boldsymbol{\xi}$$

where $\boldsymbol{\xi}_i = 0.5Q_{ii}[(1+w_1)\hat{\pi}_i - w_1]$, Q_{ii} are the diagonal elements of $\mathbf{Q} = \mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'$, and $W = diag\{\hat{\pi}_i(1-\hat{\pi}_i)w_i\}$. The bias-corrected estimate is then

$$\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}} - bias\left(\hat{\boldsymbol{\beta}}\right).$$

Notice that this correction affects the constant term directly, but also the other coefficients primarily as a consequence (King and Zeng, 2001).

The variance of the approximately unbiased estimate of $\boldsymbol{\beta}$, $\tilde{\boldsymbol{\beta}}$, is approximated by a multiple of the usual variance matrix,

$$V\left(\tilde{\boldsymbol{\beta}}\right) = \left(\frac{n}{n+k}\right)^2 V\left(\hat{\boldsymbol{\beta}}\right).$$

A key point is that since $(n/(n+k))^2 < 1$, $V\left(\tilde{\boldsymbol{\beta}}\right) < V\left(\hat{\boldsymbol{\beta}}\right)$. Thus, reducing the bias also reduces the variance.

4.2 Data Description

The primary source of information for our estimations is the Chilean Survey of Household Finance (EFH). The EFH is an initiative of the Central Bank of Chile to collect information about the household's finance including socioeconomic characteristics, information about assets holdings, and detailed information on the debts of each household member. We use the five available waves of the EFH: the first wave (2007) collected information from 3828 urban households at the national level; the second (2008), third (2009), and fourth (2010) waves accounted for 1154, 1190, and 2037 urban households in the Metropolitan Region of Chile, respectively; finally, the fifth wave (2011-12) collected information from 4059 urban households at the national level.

For each wave, higher income households have been oversampled in order to obtain a better insight of the indebtedness. In Table 1 we show the percentage of households per income stratum.[15]

Table 2 shows information about mortgage holdings and delinquency as it comes from the EFH. However, to correct the bias induced by rare events, we also incorporate some prior information coming from the delinquency reported by the banks to the Superintendency of Banks and Financial Institutions (SBIF, for its name in Spanish) every month. Ideally we would use the number of delinquent loans versus the total number of outstanding loans. However, that information is not available. Instead, we use the amounts of loans. Thus, the prior information for us is given by the average ratio of the amount of nonperforming mortgage loans to the total amount of outstanding mortgage loans in the Chilean financial system, during the months each wave of the EFH was collected. Although imperfect, this measure is a good approximation given the almost even distribution of the amount of credits in delinquency in each wave of the EFH (see the Appendix for charts showing the distribution for each year). Notice that while the oversampling of defaulted loans is similar in the EFH, the delinquency rate (as reported by the SBIF) changes over time following the impact of the global crisis in Chile.

The EFH collects information on various aspects of the credits, but specific variables for our estimations must be constructed. For each mortgage loan, we constructed the loan-to-value ratio (LTV) for both initial loan amount and current outstanding debt. While most of the analysis uses the initial LTV (eg. Kau et al., 1994; Harrison et al., 2004; Bajari et al., 2010, 2012; Bhutta et al., 2010; Campbell and Cocco, 2010; Elul et al., 2010; Paniza Bontas,

[15] Stratum 1: percentiles 1-50; Stratum 2: percentiles 51-80; Stratum 3: percentiles 81-100.

2010; Hatchondo et al., 2012), we are more prone to use the current LTV (as in Capozza et al., 1997, 1998; Ambrose and Capone, 1998; Wong et al., 2004). Our perspective is that the current LTV better reflects the financial burden for the household and is the appropiate measure for the household to include in his optimization problem to evaluate the default decision. The initial LTV is useful when the loan is in its first stages of payment, or when we want to evaluate the risk involved at the time of granting the loan. However, for long-term credits such as mortgages, the changing nature of the economic context and the household financial balancesheet hinder the usefulness of the initial LTV. Nevertheless, in our analysis we tested the role of both measures, as shown in the next section.

In Table 3 we present information on the quartiles of the current and initial LTV, and the monthly installment. The median mortgage loan in our sample was granted to cover the 85% of the home value, and the current due amount represents the 45% of the current home value. The median monthly installment is about $185,000 (around US$ 350). Also notice that the median term of the credit is 20 years, while the median age of the credit (the time period elapsed from the granting of the mortgage) is 6 years. As it can be seen, most of the credits are granted for a term of around 20 years, while in more advanced economies the usual term is around 30 years.

We make a final note on the situation of renegotiated mortgages to have some insight on the behavior of those mortgagors that might be observed as defaulting but have renegotiated the loans to avoid default. In our complete sample (12,268 households) we have that 15.9% of them have an outstanding mortgage loan. Of them, the 19.5% have renegotiated the conditions of their loans at some point. Table 4 shows default situation of mortgagors that have renegotiated their loans. It is important to notice that there is a difference between households that are defaulting and those that are paying their debts after renegotiating a mortgage: the main motivation to renegotiate for defaulters was "to decrease the mortgage payment by increasing the term"; for those who kept paying their loan after renegotiation, the main motivation was "a decrease in the interest rate".[16]

5 Results

In this section we present estimation results using the method described in the previous section.[17] There are two sets of estimations in the following tables. In Table 5, each idiosyncratic and systemic variable enters directly in the regressions to explain the decision of defaulting. In Table 6, there are interactions between these variables included as regressors in the estimation, as suggested by our theoretical model.

In Table 5, we show four different models that differ in the inclusion of financial idiosyncratic variables and systemic variables. The set of demographic variables remain the same for all estimated models. Notice that the sign, magnitude, and significance of the coefficients accompanying the latter set of variables is stable through model specifications. Income is an important determinant of the probability of default, and it enters the regressions with the expected negative sign. Contrary to popular intuition, the number of persons in the household impacts positively in the probability of default. This result may be associated with the fact that in most cases at hand, a more densely populated household is not directly associated with higher household total income (for example, because the increased number of members of the household is due to the presence of children).

Regarding the financial variables, we find that suffering a negative shock in the recent past significantly increases the probability of defaulting a mortgage loan.[18] Households that renegotiated the conditions of their loans also are more likely to default them. Having rejected credit applications appears not to be significant to explain default behavior.

[16] The distinction of motivations is based on an specific question for mortgagors that have renegotiated their loans conditions, as collected by the EFH.

[17] We have performed a complete set of estimations including a wide range of variables that were included in previous empirical analysis of mortgage default. We also tried single step and two step estimations to correct for possible endogeneity in the decision of getting a loan and then defaulting, as in Alfaro and Gallardo (2012). We did not find any technical or empirical reason to keep the track of the two step estimation procedure. However, all the additional results are available from the authors upon request.

[18] The EFH collects information on a specific question that inquires about the occurrence of an event that significantly lowered the income o increased the expenditure and that was not planned.

We included three variables capturing systemic impact on household default behavior. The current loan to value (CLTV) implies that a bigger CLTV increases the probability of default all other things equal. Our results indicate that a higher amount due (higher long-term financial burden) may impact positively the decision of defaulting (as in Elul et al., 2010). On the other hand, higher price houses (usually associated with higher amounts of loans) lessen the probability of default. The LTV of origination does not have major impact on the decision of default. This result supports our argument that CLTV is the relevant information for the household, specially if the relatively more defaulted loans are the older ones. The systemic variables as a whole tell us that bigger mortgage loans are less likely to default. These credits have being predominantly granted since 2005, accompanying important aggregate phenomena of the Chilean economy: a sustained high growth period (even during the global crisis), and a real estate boom. These two aspects require more detailed analysis in terms of the risk-taking decision of the financial system and the aggregate push-pull effect of commodity prices and other macro determinants, all of them far beyond the scope of this paper.[19]

The second set of results is presented in Table 6. Idiosyncratic demographic and financial variables, as well as systemic variables, play a similar role to that of the results in Table 5. The novel results come from the interaction between systemic and idiosyncratic variables, as our theoretical model suggested. We find that the interaction between income and current loan to value positively impact the chances of making default, as well as the interaction between the price of the home at the origination moment and the fact that the household suffered an unexpected negative shock. In the first case, even when the income has a negative effect on the default probability, the CLTV dominates in the overall. In the second case, the effect of the negative shock dominates the scene. These results emphasize that shocks affecting income (expenditure) or the remaining (long term) financial burden are key aspects when analyzing default probabilities. For example, sustained high unemployment rates may

[19] Our general equilibrium framework incorporates these two levels of analysis together with the partial equilibrium we are presenting in this paper.

impact income level configuring at the same time an unexpected shock that aggravates the financial burden.

As we have analyzed above –and in contrast to Alfaro Gallardo (2012), which also analyze the Chilean mortgage market–, the current LTV (CLTV) as a systemic factor is statistically and economically significant, as well as the house price at origination. This suggests that both idiosyncratic and systemic factors (and its interactions) are simultaneously determining the mortgage defaults.

6 Final Remarks

In this paper we propose a theoretical framework to model mortgage default in full recourse economies. In such framework, it is possible to derive an analytical expression for the household's default decision that involves idiosyncratic and systemic factors, and their combinations.

To empirically test this framework, we take into account two important issues of the type of phenomenon we are studying: first, its occurrence is rare; and second, the sampling framework favored the collection of data on default cases but contributes to worsen the statistical properties of estimators. Overall, traditional logistic regression provides biased and inefficient estimators. We overcome this problem adopting appropiate adjustments suggested by the literature.

Following our theoretical default description and the empirical strategy, we estimate the default decision determinants for the representative Chilean household, using information from the EFH survey.

According to our theoretical model, we find empirically that systemic and idiosyncratic factors are statistically important to determine the default decision. Moreover, the systemic and idiosyncratic interactions play a key role in determining default probability.

Our results suggest that aggregated macrofinancial variables –such as prices and loan to value ratios– and its interaction with microeconomic information, can contribute to explain an important portion of mortgage risks. This in turn implies that, when assessing the quality of a mortgage, or any other credit portfolio, the use of microeconomic or demographic information is not enough. Accordingly, financial regulators and supervisors should invest in developing aggregated measures that can act as early warning indicators, as well as incorporating market trends information into the analysis of risks. Furthermore, this constitutes another rationale that reinforces the idea that monetary and banking authorities should inform the market about this type of financial developments in a consistent and integrated report, that uses macro and micro financial information. Financial stability reports appear as natural channel to issue this information.[20].

7 References

Alfaro, R., and N. Gallardo (2012). "The Determinants of Household Debt Default," Revista de Análisis Económico – Economic Analysis Review, Ilades-Georgetown University, Universidad Alberto Hurtado/School of Economics and Bussines, 27(1)55-70.

Ambrose, B. W., and C. A. Capone (1998). "Modeling the conditional probability of foreclosure in the context of single-family mortgage default resolutions," Real Estate Economics 26(3):391-429.

Amemiya, T., and Q. H. Vuong (1987). "A Comparison of Two Consistent Estimators in the Choice-Based Sampling Qualitative Response Model." Econometrica 55(3):699–702.

Bajari, P., S. Chu, and M. Park (2010). "An Empirical Model of Subprime Mortgage Default from 2000 to 2007," mimeo.

[20] Although the financial stability reports were developed and published first in the mid 2000's, they increased their importance after the global financial crisis in 2008. Various central banks have developed their own periodic reports, such as UK, US, Spain, Australia, and Chile.

Bajari, P., S. Chu, D. Nekipelov, and M. Park (2012). "A Dynamic Model of Subprime Mortgage Default: Estimation and Policy Implications," mimeo.

Bernanke, B. S., M. Gertler, and S. Gilchrist (1999). "The financial accelerator in a quantitative business cycle framework," Handbook of Macroeconomics, chapter 21, pp. 1341-1393.

Bhutta, N., J. Dokko, and H. Shan (2010). "The depth of negative equity and mortgage default decisions," Finance and Economics Discussion Series (FEDS) 2010-35, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board (USA).

Campbell, J. Y., and J. F. Cocco (2010). "A Model of Mortgage Default," mimeo.

Capozza, D. R., D. Kazarian, and T. A. Thomson (1997). "Mortgage default in local markets," Real Estate Economics, 25(4):631-655.

Capozza, D. R., D. Kazarian, and T. A. Thomson (1998). "The conditional probability of mortgage default," Real Estate Economics, 26(3):359-389.

Dubey, P., J. Geanakoplos, and M. Shubik (2005). "Default and Punishment in General Equilibrium," Econometrica, 73(1):1-37.

Elmer, P. J., and S. A. Seelig (1999). "Insolvency, trigger events, and consumer risk posture in the theory of single-family mortgage default." Journal of Housing Research 10:1-25.

Elul, R., N. S. Souleles, S. Chomsisengphet, D. Glennon, and R. Hunt (2010). "What 'Triggers' Mortgage Default?" American Economic Review: Papers & Proceedings 100(May 2010):490–494.

Fisher, I. (1933). "The Debt-Deflation Theory of Great Depressions," Econometrica, 1(4):337-357.

Foster, C., and R. Van Order (1984). "An Option-based Model of Mortgage Default," Housing Finance Review, 3(4): 351–72.

Geanakoplos J. and A. Fostel, (2008). "Leverage Cycles and the Anxious Economy," American Economic Review, American Economic Association, vol. 98(4), pages 1211-44, September.

Geanakoplos, J., and W. Zame (2013). "Collateral Equilibrium: A Basic Framework," Cowles Foundation Discussion Paper No. 1906 (forthcoming, Economic Theory).

Gerardi, K., A. H. Shapiro, and P. S. Willen (2007). "Subprime outcomes: risky mortgages, homeownership experiences, and foreclosures," Working Papers 07-15, Federal Reserve Bank of Boston.

Goodhart, C. A. E., D. P. Tsomocos, and A. P. Vardoulakis (2011). "Modeling a Housing and Mortgage Crisis," in Rodrigo Alfaro (ed.), Financial Stability, Monetary Policy, and Central Banking, Vol. 15, chapter 8, pp. 215-253, Central Banking, Analysis, and Economic Policies Book Series, Central Bank of Chile.

Goodhart, C. A. E., P. Sunirand, and D. Tsomocos (2006). "A model to analyse financial fragility," Economic Theory, 27:107-142.

Harrison, D. M., T. G. Noordewier, and A. Yavas (2004). "Do Riskier Borrowers Borrow More?" Real Estate Economics 32(2):385-411.

Hatchondo, J. C., L. Martinez, J. M. Sanchez (2012). "Mortgage Defaults," IMF Working Paper WP/12/26.

IMF (2005). "The treatment of nonperforming loans," Eighteenth Meeting of the IMF Committee on Balance of Payments Statistics, BOPCOM-05/29.

Jaffee, D. and Stiglitz, J. (1990). "Credit Rationing," Handbook of Monetary Economics, Vol. 2:837-888.

Kau, J. B. , D. C. Keenan, and T. Kim (1994). "Default probabilities for mortgages," Journal of Urban Economics 35:278-296.

King, G., and L. Zeng (2001). "Logistic regression in rare events data." Political Analysis Seminar, 9(2):137–163.

King, G., and L. Zeng (2002). "Estimating risk and rate levels, ratios and differences in case-control studies." Statistics in Medicine, 21:1409-1427.

King, G., M. Tomz, and J. Wittenberg (2000). "Making the most of statistical analyses: improving interpretation and presentation," American Journal of Political Science, 44(2):341–355.

Kiyotaki, N., and J. Moore (1997). "Credit Cycles," Journal of Political Economy, 105(2):211-248.

Manski, C. F., and S. R. Lerman (1977). "The Estimation of Choice Probabilities from Choice Based Samples." Econometrica 45(8):1977–1988.

McCullagh, P., and J. A. Nelder (1989). Generalized Linear Models, 2nd ed. New York: Chapman and Hall.

OECD (2011). "Housing and the Economy: Policies for Renovation," Economic Policy Reforms 2011: Going for Growth, chapter 4.

Paniza Bontas, J. (2010). "Combined LTV at Origination and Household Default Probabilty," mimeo, University of Chicago.

Scott, A. J., and C. J. Wild (1986). "Fitting Logistic Models Under Case-Control or Choice Based Sampling." Journal of the Royal Statistical Society, B 48(2):170–182.

Shubik M., and C. Wilson (1977). "The Optimal Bankruptcy Rule in a Trading Economy Using Fiat Money," Zeitschrift fur Nationalokonomie, 37(3-4):337-354.

Tsomocos, D. (2003). "Equilibrium Analysis, Banking, and Financial Instability," Journal of Mathematical Economics, 39(5-6):619-655.

Wong, J., L. Fung, T. Fong and A. Sze (2004). “Residential mortgage default risk and the loan-tovalue Ratio,” Hong Kong Monetary Authority Quarterly Bulletin, December, pp. 35-45.

Xie, Y., and C. F. Manski (1989). “The Logit Model and Response-Based Samples.” Sociological Methods and Research 17(3):283–302.

8 Appendix

8.1 Glossary

Non-performing loans: A loan is nonperforming when payments of interest and/or principal are past due by 90 days or more, or interest payments equal to 90 days or more have been capitalized, refinanced, or delayed by agreement, or payments are less than 90 days overdue, but there are other good reasons —such as a debtor filing for bankruptcy— to doubt that payments will be made in full. After a loan is classified as nonperforming, it (and, possibly, replacement loan(s)) should remain classified as such until written off or payments of interest and/or principal are received (IMF, 2005).

Default: it essentially means a debtor has not paid a debt which he or she is required to have paid. Debt service default occurs when the borrower has not made a scheduled payment of interest or principal.

Reputational Effects of Default: A loan becomes *delinquent* the first day after the borrower misses a payment. The *delinquency* will continue until all payments are made to *bring the loan current*. Loan servicers report all delinquencies of at least 90 days to the SBIF and other credit bureaus (e.g. SINACOFI, DICOM). This information is used by the financial market agents to evaluate risk and grant credit. In particular, being in listed in DICOM implies that the debtor also may have trouble getting a job, signing up for utilities or leasing, getting home owner's insurance, getting a cellphone plan, or getting approval to rent an apartment (credit checks usually are required for renters).

8.2 First order conditions for the household problem

$$\frac{\partial L^\alpha}{\partial q^\alpha_{01}} : -U^{\alpha\prime}\left(e^\alpha_{01}-q^\alpha_{01}\right)-\eta^\alpha_{01}p_{01} = 0 \tag{3}$$

$$\frac{\partial L^\alpha}{\partial q^\alpha_{11}} : -\pi_G U^{\alpha\prime}\left(e^\alpha_{11}-q^\alpha_{11}\right)-\eta^\alpha_{11}p_{11} = 0 \tag{4}$$

$$\frac{\partial L^\alpha}{\partial q^\alpha_{21}} : -\pi_B U^{\alpha\prime}\left(e^\alpha_{21}-q^\alpha_{21}\right)-\eta^\alpha_{21}p_{21} = 0 \tag{5}$$

$$\frac{\partial L^\alpha}{\partial b^\alpha_{02}} :\ U^{\alpha\prime}\left(\frac{b^\alpha_{02}}{p_{02}}\right)\frac{1}{p_{02}}+\pi_G U^{\alpha\prime}\left(\frac{b^\alpha_{02}}{p_{02}}+\frac{b^\alpha_{G2}}{p_{G2}}\right)\frac{1}{p_{02}}+\eta^\alpha_{02}-\phi\eta^\alpha_\phi+\frac{\lambda^\alpha\pi_B p_{B2}}{\bar{\mu}p_{02}} = 0 \tag{6}$$

$$\frac{\partial L^\alpha}{\partial b^\alpha_{G2}} :\ \pi_G U^{\alpha\prime}\left(\frac{b^\alpha_{02}}{p_{02}}+\frac{b^\alpha_{G2}}{p_{G2}}\right)\frac{1}{p_{G2}}+\eta^\alpha_{G2} = 0 \tag{7}$$

$$\frac{\partial L^\alpha}{\partial b^\alpha_{B2}} :\ \pi_B U^{\alpha\prime}\left(\frac{b^\alpha_{B2}}{p_{B2}}\right)\frac{1}{p_{B2}}+\eta^\alpha_{B2} = 0 \tag{8}$$

$$\frac{\partial L^\alpha}{\partial \mu^\alpha_0} :\ \eta^\alpha_{01}-\frac{\eta^\alpha_{02}}{1+r_0} = 0 \tag{9}$$

$$\frac{\partial L^\alpha}{\partial \mu^\alpha_1} :\ \eta^\alpha_{11}-\frac{\eta^\alpha_{G2}}{1+r_1} = 0 \tag{10}$$

$$\frac{\partial L^\alpha}{\partial \mu^\alpha_2} :\ \eta^\alpha_{21}-\frac{\eta^\alpha_{B2}}{1+r_2} = 0 \tag{11}$$

$$\frac{\partial L^\alpha}{\partial \bar{\mu}^\alpha} :\ -\frac{\eta^\alpha_{02}}{1+\bar{r}}+\eta^\alpha_{G2}+\frac{\eta^\alpha_\phi}{1+\bar{r}}-\frac{\lambda^\alpha\pi_B p_{B2}b^\alpha_{02}}{\bar{\mu}^2 p_{02}} = 0 \tag{12}$$

$$\frac{\partial L^\alpha}{\partial \eta^\alpha_{01}} :\ \mu^\alpha_0-p_{01}q^\alpha_{01} = 0 \tag{13}$$

$$\frac{\partial L^\alpha}{\partial \eta^\alpha_{02}} :\ b^\alpha_{02}-\frac{\mu^\alpha_0}{1+r_0}-\frac{\bar{\mu}^\alpha}{1+\bar{r}}-m^\alpha_0 = 0 \tag{14}$$

$$\frac{\partial L^\alpha}{\partial \eta^\alpha_\phi} :\ \frac{\bar{\mu}^\alpha}{1+\bar{r}}-\phi b^\alpha_{02} = 0 \tag{15}$$

$$\frac{\partial L^\alpha}{\partial \eta^\alpha_{11}} :\ \mu^\alpha_1-p_{11}q^\alpha_{11} = 0 \tag{16}$$

$$\frac{\partial L^\alpha}{\partial \eta^\alpha_{21}} :\ \mu^\alpha_2-p_{21}q^\alpha_{21} = 0 \tag{17}$$

$$\frac{\partial L^\alpha}{\partial \eta^\alpha_{G2}} :\ b^\alpha_{G2}+\bar{\mu}^\alpha-\frac{\mu^\alpha_1}{1+r_1}-m^\alpha_1 = 0 \tag{18}$$

$$\frac{\partial L^\alpha}{\partial \eta^\alpha_{B2}} :\ b^\alpha_{B2}-\frac{\mu^\alpha_2}{1+r_2}-m^\alpha_2 = 0 \tag{19}$$

8.3 Proof of Proposition 1

From (5),

$$U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}\right)\frac{1}{p_{02}}+\pi_1 U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}+\frac{b_{G2}^{\alpha}}{p_{G2}}\right)\frac{1}{p_{02}}-\phi\eta_{\phi}^{\alpha}+\frac{\lambda^{\alpha}\pi_2 p_{B2}}{\bar{\mu}p_{02}}=-\eta_{02}^{\alpha}$$

From (2), we have that $\frac{-U^{\alpha\prime}\left(e_{01}^{\alpha}-q_{01}^{\alpha}\right)}{p_{01}}=\eta_{01}^{\alpha}$, from (7) $\eta_{01}^{\alpha}(1+r_0)=\eta_{02}^{\alpha}$. Combining these results,

$$\frac{U^{\alpha\prime}\left(e_{01}^{\alpha}-q_{01}^{\alpha}\right)}{p_{01}}(1+r_0)=U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}\right)\frac{1}{p_{02}}+\pi_1 U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}+\frac{b_{G2}^{\alpha}}{p_{G2}}\right)\frac{1}{p_{02}}-\phi\eta_{\phi}^{\alpha}$$

$$\frac{-U^{\alpha\prime}\left(e_{01}^{\alpha}-q_{01}^{\alpha}\right)}{p_{01}\phi}(1+r_0)+\frac{1}{p_{02}\phi}\left[U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}\right)+\pi_1 U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}+\frac{b_{G2}^{\alpha}}{p_{G2}}\right)+\frac{\lambda^{\alpha}\pi_2 p_{B2}}{\bar{\mu}}\right]=\eta_{\phi}^{\alpha}$$

On the other hand, using (6),$-\pi_1 U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}+\frac{b_{G2}^{\alpha}}{p_{G2}}\right)\frac{1}{p_{G2}}=\eta_{G2}^{\alpha}$, we can replace this expression with (5) in (11) and it yields,

$$\frac{-1}{p_{02}(1-\phi)}\left[U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}\right)+\pi_1 U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}+\frac{b_{G2}^{\alpha}}{p_{G2}}\right)+\frac{\lambda^{\alpha}\pi_2 p_{B2}}{\bar{\mu}}\right]+\frac{\lambda^{\alpha}\pi_2 p_{B2}b_{02}^{\alpha}(1+\bar{r})}{\bar{\mu}^2 p_{02}(1-\phi)}+ \pi_1(1+\bar{r})U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}+\frac{b_{G2}^{\alpha}}{p_{G2}}\right)\frac{1}{p_{G2}(1-\phi)}=\eta_{\phi}^{\alpha}$$

Equating for η_{ϕ}^{α}, we have,

$$\frac{-U^{\alpha\prime}\left(e_{01}^{\alpha}-q_{01}^{\alpha}\right)}{p_{01}\phi}(1+r_0)=\frac{-1}{p_{02}(1-\phi)\phi}\left[U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}\right)+\pi_1 U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}+\frac{b_{G2}^{\alpha}}{p_{G2}}\right)+\frac{\lambda^{\alpha}\pi_2 p_{B2}}{\bar{\mu}}\right]+ \frac{\lambda^{\alpha}\pi_2 p_{B2}b_{02}^{\alpha}(1+\bar{r})}{\bar{\mu}^2 p_{02}(1-\phi)}+\pi_1(1+\bar{r})U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}+\frac{b_{G2}^{\alpha}}{p_{G2}}\right)\frac{1}{p_{G2}(1-\phi)}$$

$$\frac{-U^{\alpha\prime}\left(e_{01}^{\alpha}-q_{01}^{\alpha}\right)}{p_{01}}(1+r_0) = \frac{1}{p_{02}(1-\phi)}\left[-U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}\right)+\frac{\pi_1(\phi p_{02}(1+\bar{\mu})-p_{G2})}{p_{G2}}U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}+\frac{b_{G2}^{\alpha}}{p_{G2}}\right)-\frac{\lambda^{\alpha}\pi_2 p_{B2}}{\bar{\mu}}\right]$$
$$\frac{\lambda^{\alpha}\pi_2 p_{B2} b_{02}^{\alpha}(1+\bar{r})\phi}{\bar{\mu}^2 p_{02}(1-\phi)}$$

$$\frac{-b_{02}^{\alpha}p_{B2}}{\bar{\mu}p_{02}} = \frac{\bar{\mu}}{p_{02}\lambda^{\alpha}\pi_2(1+\bar{r})\phi}\left[-U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}\right)+\frac{\pi_1(\phi p_{02}(1+\bar{\mu})-p_{G2})}{p_{G2}}U^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}+\frac{b_{G2}^{\alpha}}{p_{G2}}\right)-\frac{\lambda^{\alpha}\pi_2 p_{B2}}{\bar{\mu}}\right]+$$
$$\frac{-U^{\alpha\prime}\left(e_{01}^{\alpha}-q_{01}^{\alpha}\right)}{p_{01}}\frac{(1+r_0)(1-\phi)\bar{\mu}}{\lambda^{\alpha}\pi_2(1+\bar{r})\phi}$$

Adding 1 and regrouping, we have:

$$1-\frac{b_{02}^{\alpha}p_{B2}}{\bar{\mu}^{\alpha}p_{02}} = \omega_0+\omega_1\bar{\mu}^{\alpha}u^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}\right)+\omega_2\bar{\mu}^{\alpha}u^{\alpha\prime}\left(\frac{b_{02}^{\alpha}}{p_{02}}+\frac{b_{G2}^{\alpha}}{p_{G2}}\right)+\omega_3\bar{\mu}^{\alpha}u^{\alpha\prime}\left(e_{01}^{\alpha}-q_{01}^{\alpha}\right) \quad (20)$$

Where, $\omega_0 = 1-\frac{\lambda^{\alpha}\pi_B p_{B2}}{p_{02}(1+\bar{r})\phi}$, $\omega_1 = \frac{-1}{p_{02}\lambda^{\alpha}\pi_B(1+\bar{r})\phi}$, $\omega_2 = \frac{\pi_G(\phi p_{02}(1+\bar{r})-p_{G2})}{p_{G2}p_{02}\lambda^{\alpha}\pi_B(1+\bar{r})\phi}$, $\omega_3 = \frac{-(1+r_0)(1-\phi)}{p_{01}\lambda^{\alpha}\pi_B(1+\bar{r})\phi}$

Q.E.D

8.4 Tables and Figures

Table 1: Distribution of Households by Income Group (%)

	2007	**2008**	**2009**	**2010**	**2011**	**Total**
Stratum 1	26.38	31.28	25.88	29.46	30.45	28.65
Stratum 2	24.97	29.03	31.09	29.7	29.1	28.1
Stratum 3	48.64	39.69	43.03	40.84	40.45	43.25

Table 2: Mortgage Loans and Delinquency (%)

	2007	**2008**	**2009**	**2010**	**2011**	**Total**
Mortgage holders	16.77	13.17	13.70	18.85	15.00	15.90
Defaulted mortgages	8.26	13.82	9.2	8.07	8.87	8.92
Delinquent mortgages (SBIF)	0.97	1.29	1.95	2.01	1.70	1.58

Table 3: Distribution of Mortgage Characteristics

	p25	**p50**	**p75**
Current Loan to Value	24.6 %	45.1 %	67.9 %
Initial Loan to Value	63.6 %	85.0 %	100.0 %
Monthly Intallment	$ 95,000	$ 185,000	$ 320,000
Term of Credit (in years)	19	20	20
Age of Debt (in years)	3	6	11

Table 4: Default and Renegotiation in the Sample

	Did not renegotiated	**RENEGOTIATED**	**Total**
Paying	74.5 %	16.6 %	91.1 %
DEFAULTED	6.0 %	2.9 %	8.9 %
Total	80.5 %	19.5 %	100 %

Note: Percentages are calculated over the complete group of mortgagors in the sample.

Table 5: Estimation Results 1: No interactions

Dep. Var.: Mortgage Default Dummy	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Idiosyncratic - Demographic Variables				
Number of persons in house	0.211***	0.235***	0.214**	0.213**
	(0.0715)	(0.0826)	(0.0844)	(0.0856)
Income (in logs)	-0.827***	-0.563***	-0.734***	-0.524***
	(0.133)	(0.153)	(0.155)	(0.162)
Primary Education	0.348	0.0446	0.215	0.109
	(0.333)	(0.424)	(0.440)	(0.466)
Tertiary Education	-0.633***	-0.417	-0.554*	-0.397
	(0.241)	(0.291)	(0.308)	(0.309)
Gender	-0.285	-0.214	-0.285	-0.301
	(0.199)	(0.229)	(0.244)	(0.242)
Age 18-35	-0.0218	-0.155	0.0266	0.105
	(0.260)	(0.295)	(0.306)	(0.308)
Age 55-99	-0.699	-0.0802	-0.0578	-0.0759
	(0.603)	(0.675)	(0.696)	(0.683)
Idiosyncratic - Finance Variables				
Negative Shock	1.745***	1.715***	1.668***	1.683***
	(0.207)	(0.245)	(0.255)	(0.256)
Credit Applications Rejected	0.276			
	(0.398)			
Renegotiation	1.352***	1.052***	1.016***	1.073***
	(0.244)	(0.309)	(0.332)	(0.336)
Systemic Variables				
Current Loan to Value		0.253**		
		(0.117)		
Initial House Price (in logs)		-0.416***		-0.450***
		(0.113)		(0.152)
Initial Loan to Value			0.0407	-0.0835
			(0.0249)	(0.0794)
Constant	6.125***	9.095***	4.663**	9.413***
	(1.761)	(2.302)	(2.074)	(2.631)
Observations	1,894	1,446	1,301	1,337

Note: Robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1

Table 6: Estimation Results 2: Including Interactions

Dep. Var.: Mortgage Default Dummy	**Model 1**	**Model 2**
Idiosyncratic - Demographic Variables		
Number of persons in house	0.237***	0.231***
	(0.0820)	(0.0819)
Income (in logs)	-0.578***	-0.582***
	(0.150)	(0.149)
Primary Education	0.0564	0.0804
	(0.421)	(0.414)
Tertiary Education	-0.423	-0.425
	(0.290)	(0.290)
Gender	-0.260	-0.247
	(0.228)	(0.227)
Age 18-35	-0.189	-0.195
	(0.298)	(0.298)
Age 55-99	-0.0714	-0.0950
	(0.675)	(0.675)
Idiosyncratic - Finance Variables		
Negative Shock	1.678***	
	(0.245)	
Credit Applications Rejected	0.687	0.677
	(0.437)	(0.439)
Renegotiation	1.319***	1.325***
	(0.277)	(0.276)
Systemic Variables		
Initial House Price (in logs)	-0.417***	-0.443***
	(0.112)	(0.111)
Interaction Variables		
Income and current loan to value	0.0199**	0.0195**
	(0.00810)	(0.00814)
Initial House Price and Negative shock		0.102***
		(0.0147)
Constant	9.326***	9.817***
	(2.265)	(2.243)
Observations	1,446	1,446

Note: Robust standard errors in parentheses. Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

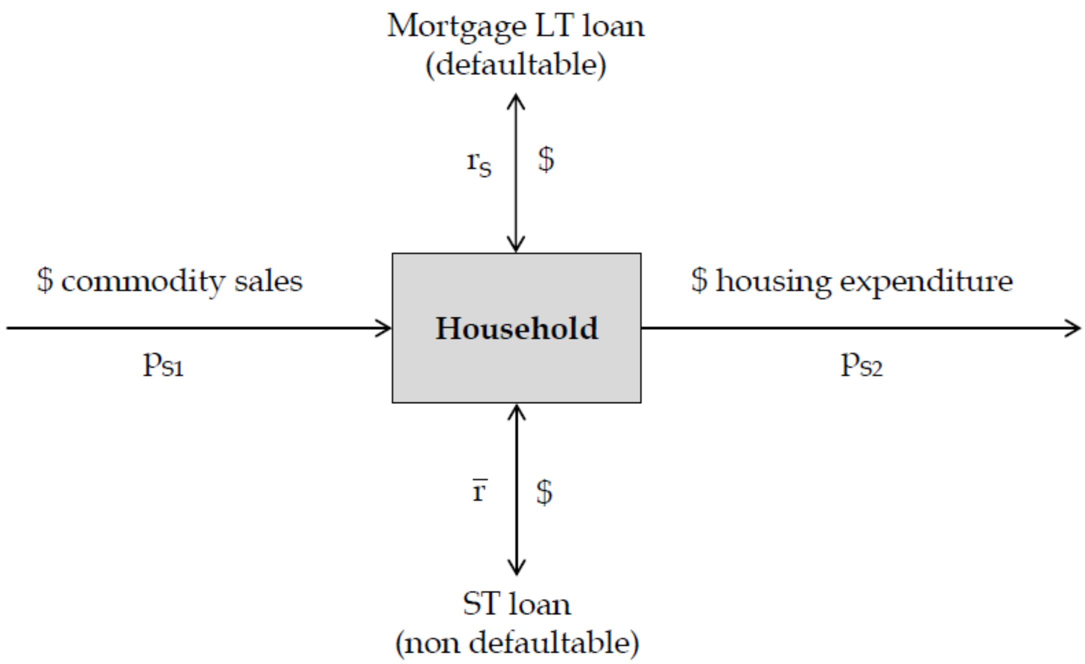


Figure 1: Nominal flows

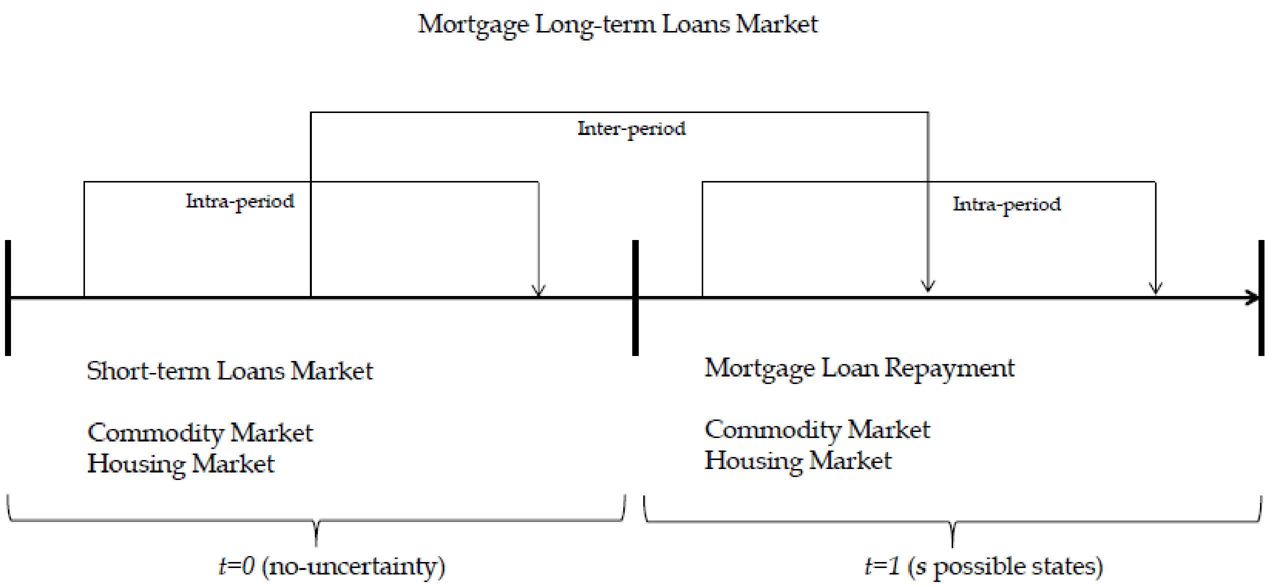


Figure 2: Timing of markets and transactions.

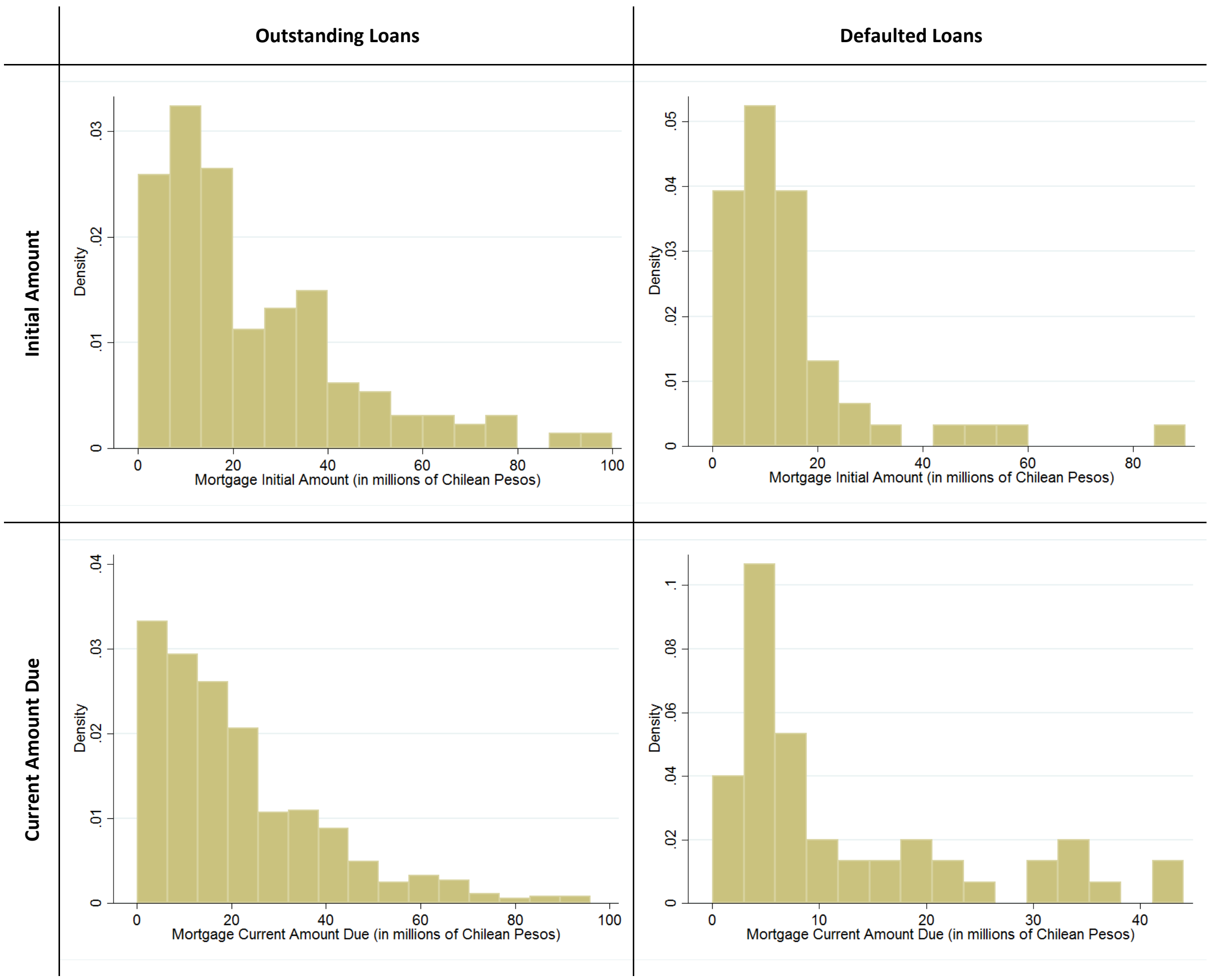


Figure 3: Distribution of the amount of mortgage loans, 2007

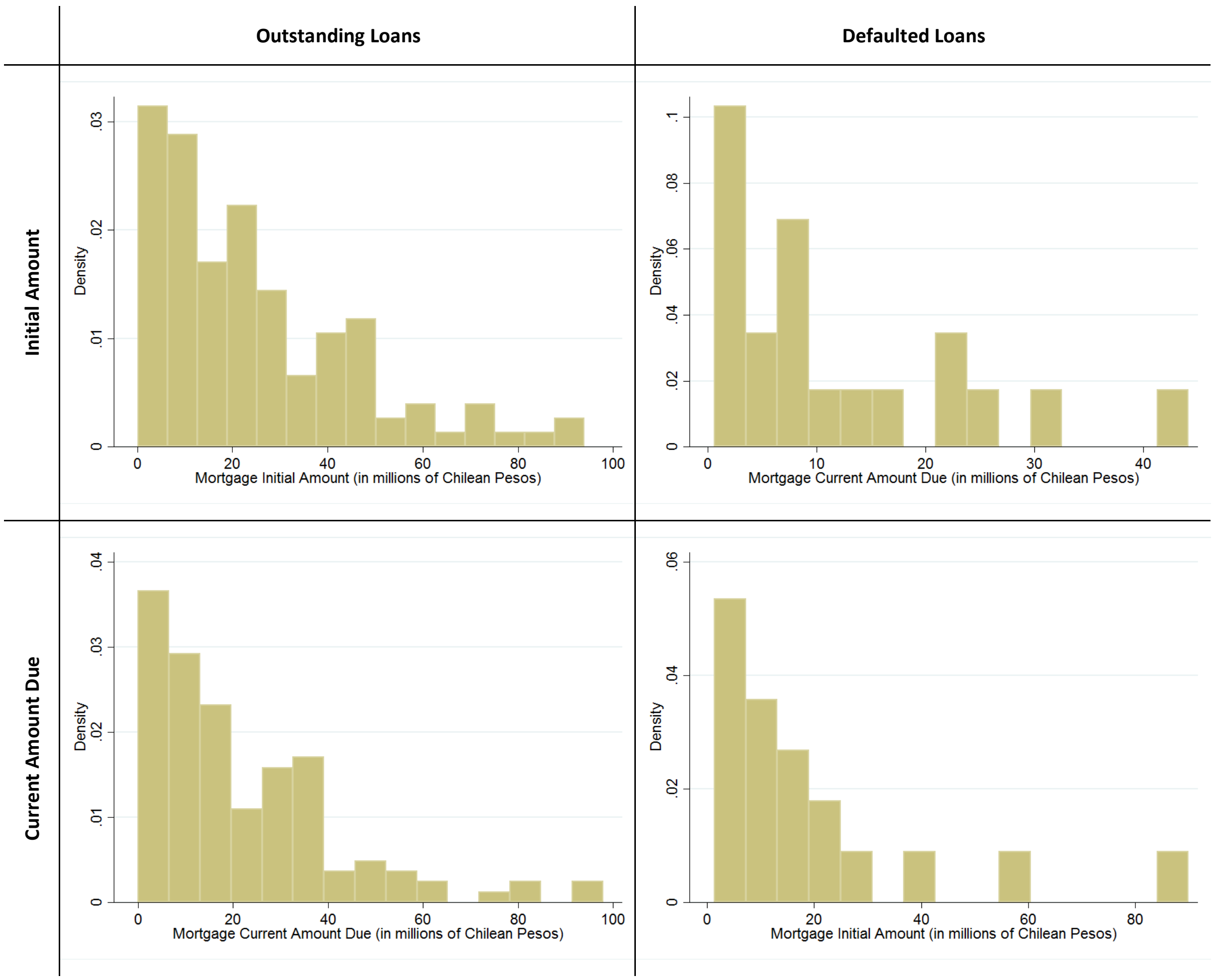


Figure 4: Distribution of the amount of mortgage loans, 2008

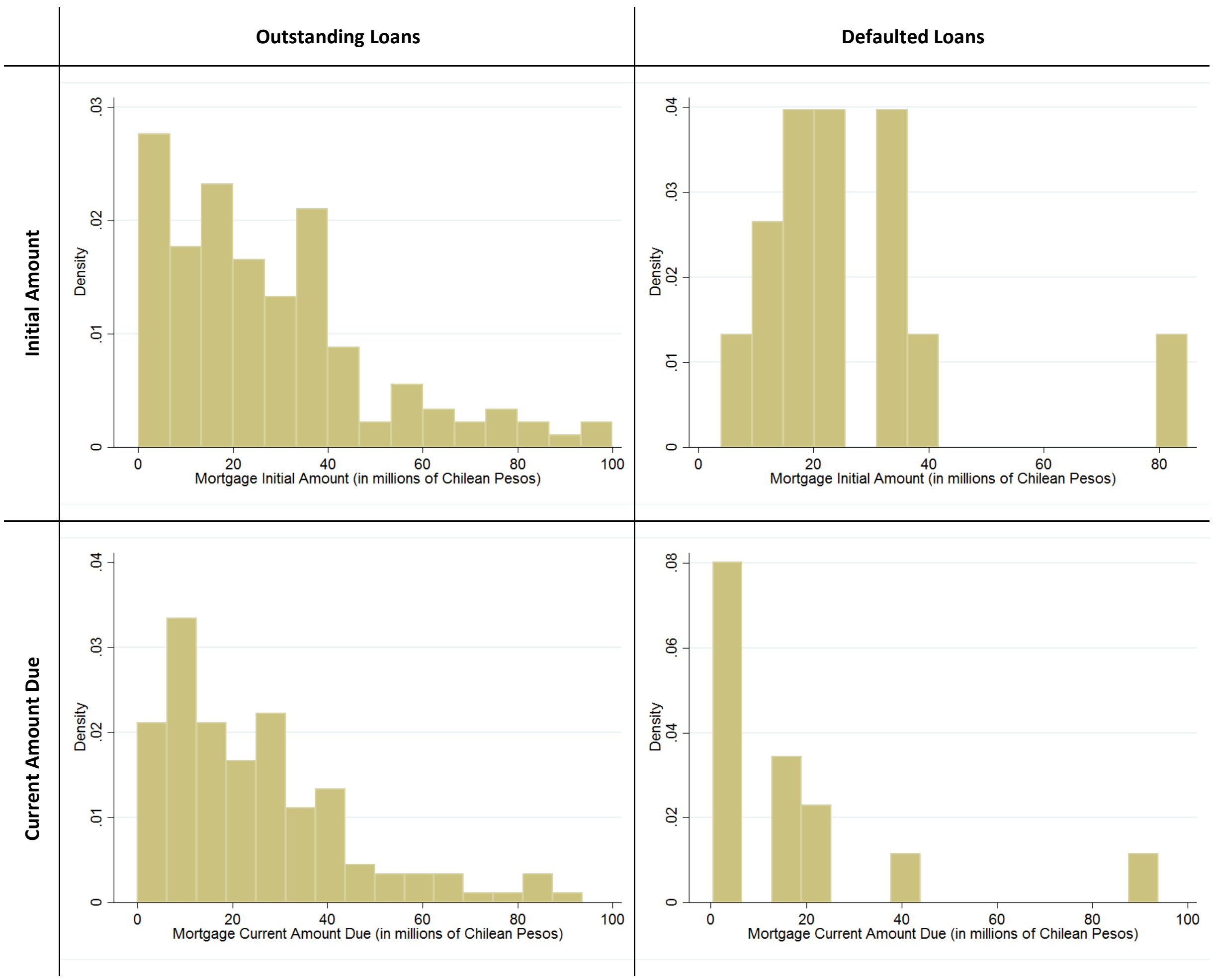


Figure 5: Distribution of the amount of mortgage loans, 2009

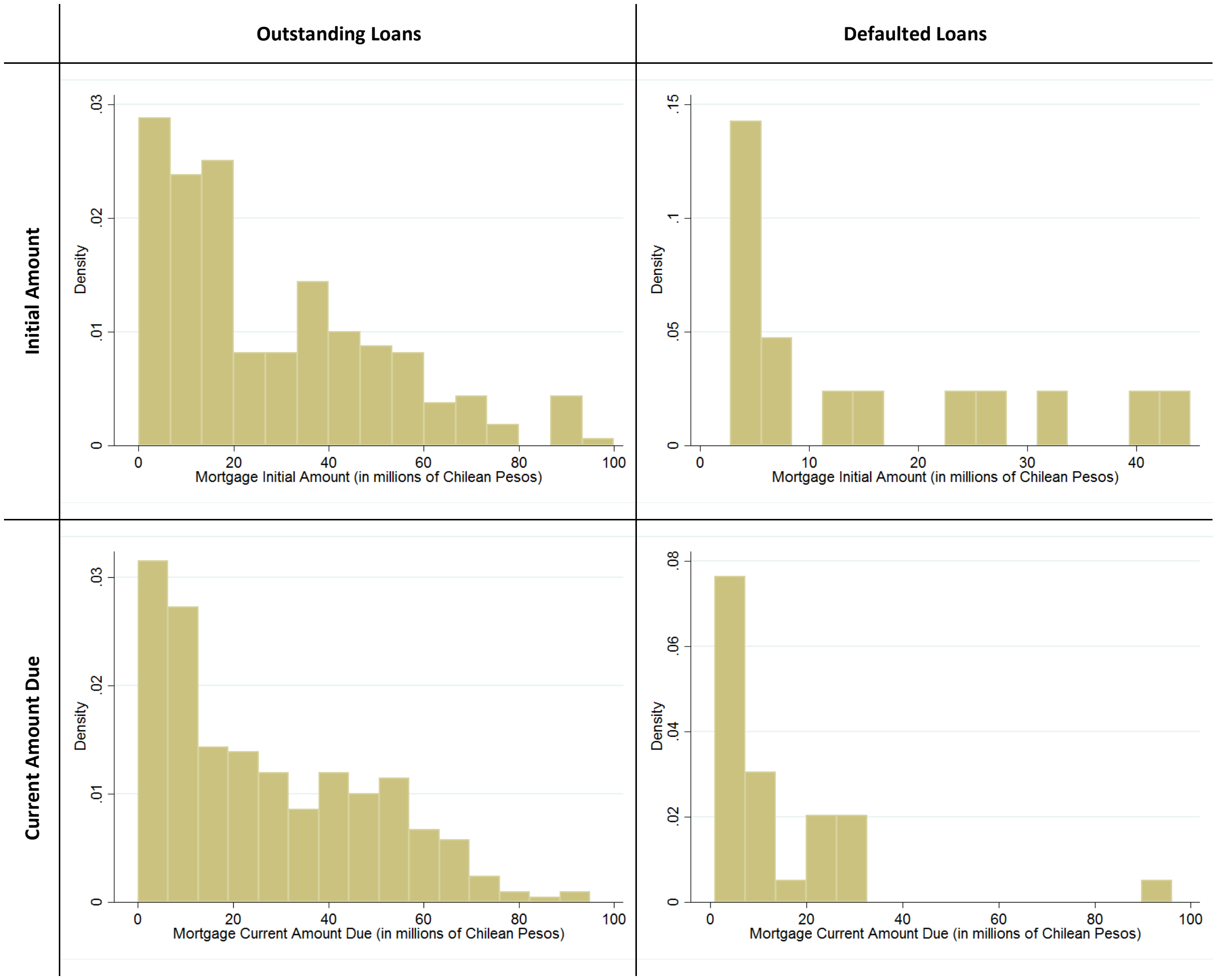


Figure 6: Distribution of the amount of mortgage loans, 2010

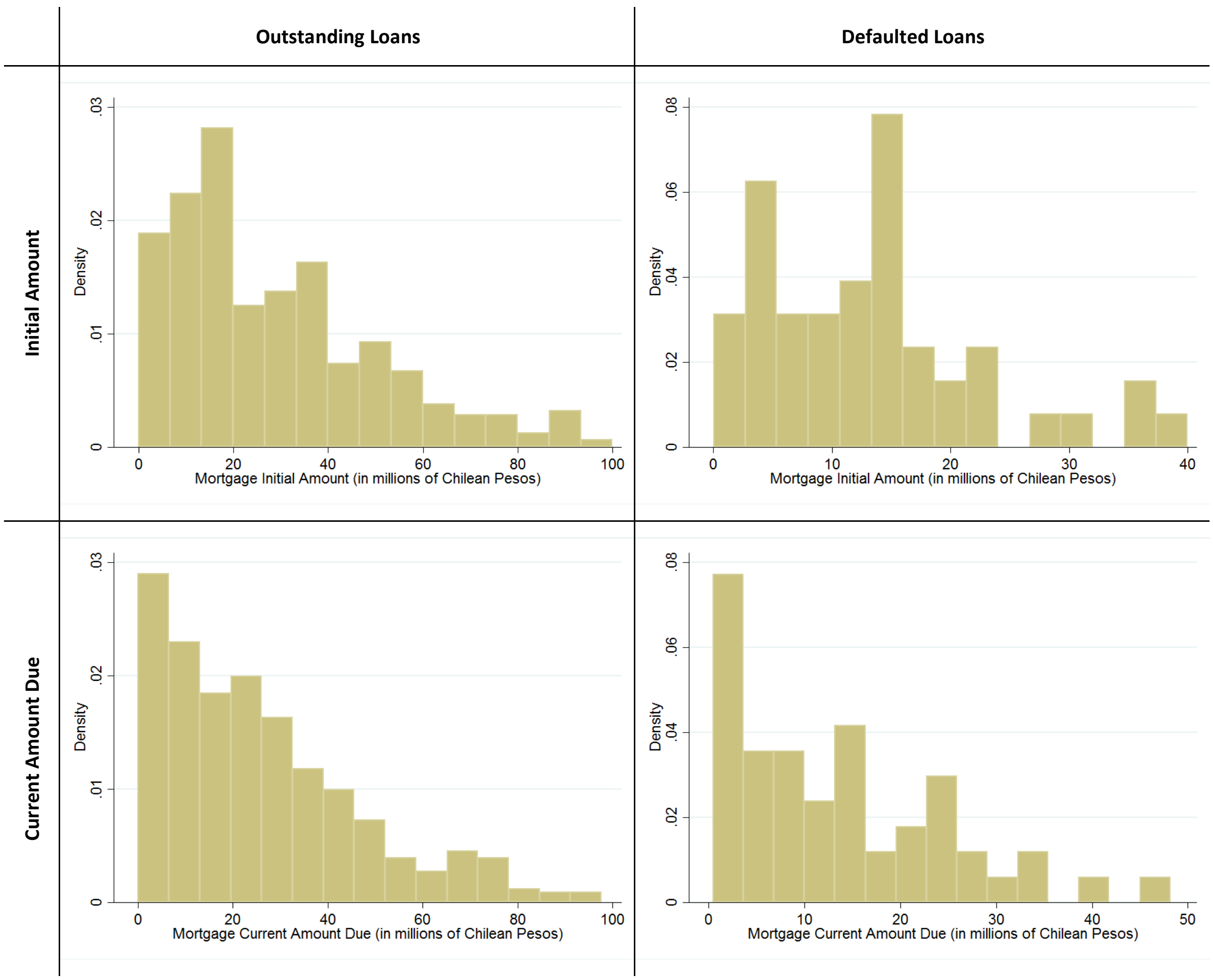


Figure 7: Distribution of the amount of mortgage loans, 2011-12

Irving Fisher Committee on
Central Bank Statistics

BANK FOR INTERNATIONAL SETTLEMENTS

IFC workshop on *"Combining micro and macro statistical data for financial stability analysis. Experiences, opportunities and challenges"*

Warsaw, Poland, 14-15 December 2015

A micro-powered model of mortgage default risk for full recourse economies, with an application to the case of Chile[1]

Diego Avanzini, Juan Francisco Martínez and Víctor Pérez, Central Bank of Chile

[1] This presentation was prepared for the meeting. The views expressed are those of the authors and do not necessarily reflect the views of the BIS or the central banks and other institutions represented at the meeting.

A micro-powered model of mortgage default risk for full recourse economies, with an application to the case of Chile[1]

D. Avanzini, J. F. Martínez and V. Pérez

Central Bank of Chile

December, 2015

[1] DISCLAIMER: The views expressed here are my own and do not necessarily represent those of the Central Bank of Chile or its Board.

The paper in a nutshell

The question:
Which are the determinants of mortgage default in a full-recourse economy?

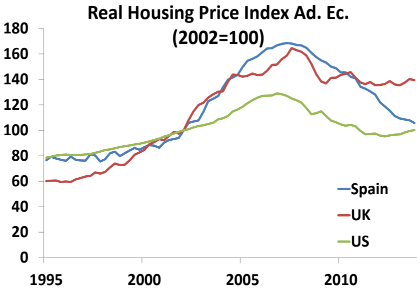
- Full-recourse vs. non-recourse regulatory frameworks
- Systemic vs. idiosyncratic factors
- Application: the case of Chile

What we do:

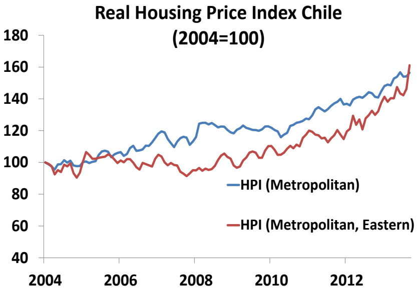
1. A theoretical model of the determinants of mortgage default under a full-recourse credit regulation
2. A suitable estimation strategy for mortgage default
3. Results from a micro-powered model estimation

Context for the question

- Real estate prices are growing fast in Chile
- These prices follow economic growth and fundamentals
- However, advanced economies had difficulties keeping up with high growth levels in the past...

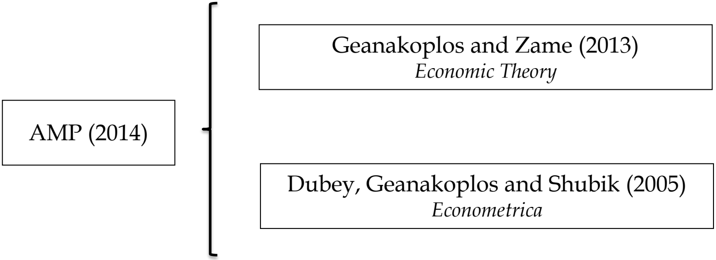


Quarterly data. Source: Own calculations based on Dallas FED International Housing Prices database.



Monthly data. Source: CChC database for houses.

Theory

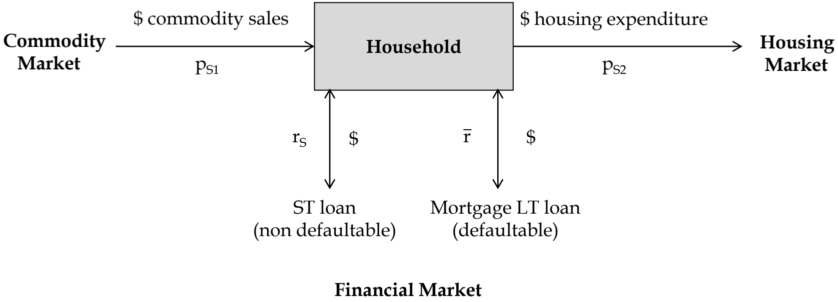

AMP (2014)
Geanakoplos and Zame (2013)
Economic Theory
Dubey, Geanakoplos and Shubik (2005)
Econometrica

Empirics

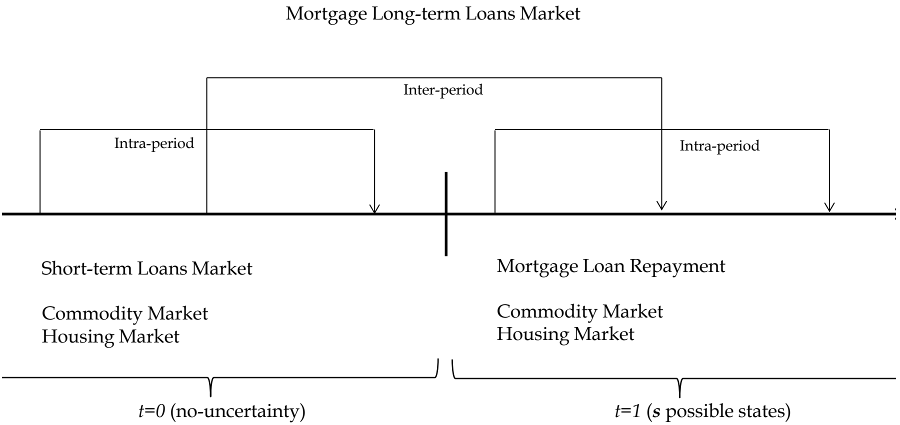
AMP (2014)

- Foster and Van Order (1984)
  Housing Finance Review
- Bhutta, Dokko and Shan (1999)
  Finance and Economics Discussion Series, Federal Reserve (USA)
- Harrison, Noordewier, and Yavas (2004)
  Real Estate Economics
- Elul, Souleles, Chomsisengphet, Glennon and Hunt (2010)
  Journal of Housing Research
- Alfaro and Gallardo (2012)
  Economic Analysis Review

Nominal Flows of the Household



Timing of the household decisions



Household optimization problem

$\max_{\mu_s, \bar{\mu}, b_{s2}, q_{s1}}$

$$
\begin{aligned}
& U\left(e_{01}-q_{01}\right)+U\left(\frac{b_{02}}{p_{02}}\right) && t=0 \\
& +E_G\left\{U\left(e_{G1}-q_{G1}\right)+U\left(\frac{b_{02}}{p_{02}}+\frac{b_{G2}}{p_{G2}}\right)\right\} && t=1,\ s\in G \\
& +E_B\left\{U\left(e_{B1}-q_{B1}\right)+U\left(\frac{b_{B2}}{p_{B2}}\right)-\lambda\left(1-\frac{b_{02}p_{B2}}{p_{02}\bar{\mu}}\right)\right\} && t=1,\ s\in B
\end{aligned}
$$

Household budget constraint

Period 0 (Deterministic):

- The short term loans must not exceed the revenues from commodity sales
- The housing expenditure must be lower than or equal to its long and short term credits and monetary endowment
- There is a LTV limit (i.e. ϕ) required for a mortgage loan

Period 1 (Stochastic):

- The short term loans must not exceed the revenues from commodity sales
- Good state: The repayment of the mortgage loans plus the new housing consumption of the household must not exceed the agent's short-term borrowing and monetary endowment
- Bad state: The new housing consumption of the household must not exceed the agent's short-term borrowing and monetary endowment

Household budget constraint

$$\mu_0 \leq p_{01} q_{01}$$

ST loan repayment $\leq$ Sales of commodities at t=0.

$$b_{02} \leq \frac{\mu_0}{1+r_0} + \frac{\bar{\mu}}{1+\bar{r}} + m_0$$

Money spent in houses$\leq$ ST loan + mortgage + monetary endowment.

$$\frac{\bar{\mu}}{1+\bar{r}} \leq \phi b_{02}$$

Mortgage Money spent in houses$\leq$ LTV*Money spent in houses.

$$\mu_s \leq p_{s1} q_{s1}$$

ST loan repayment $\leq$ Sales of commodities at t=0.

$$b_{s2} + \bar{\mu} \leq \frac{\mu_s}{1+r_s} + m_s \quad /\forall s \in S_1$$

Money spent in houses$\leq$ ST loan + mortgage + monetary endowment.

$$b_{s2} \leq \frac{\mu_s}{1+r_s} + m_s \quad /\forall s \in S_2$$

Money spent in houses $\leq$ ST loan + mortgage + monetary endowment.

- In a non-recourse mortgage economy we would only have that defaulters are enforced to repay by the threat of their collateral being confiscated. This approach includes three modelling devices within the framework:
  1. Utilities:

$$\sum_{s\in S_\alpha^1} \pi_s \left\{ U\left(\frac{b_{02}}{p_{02}} + \frac{b_{s2}}{p_{s2}} \right) \right\} + \sum_{s\in S_\alpha^2} \pi_s \left\{ U\left(\frac{b_{s2}}{p_{s2}} \right) \right\}$$

  2. Budget constraint

$$b_{s2} + \bar{\mu} \leq \frac{\mu_s}{1+r_s} + m_s \quad /\forall s \in S_1$$
$$b_{s2} \leq \frac{\mu_s}{1+r_s} + m_s \quad /\forall s \in S_2$$

  3. Interest rates (hence expectations)

$$1 + \bar{r}_s = \frac{\min\left\{ \frac{b_{02}}{p_{02}} p_{s2}, \bar{\mu} \right\}}{\bar{l}^\theta}$$

- In a full-recouse economy, we propose to add a reputational cost that further discourages default

$$-\lambda \sum_{s\in S} \pi_s \max\left\{ \left(1 - \frac{b_{02} p_{s2}}{p_{02}\bar{\mu}}\right), 0 \right\}$$

Household's Default Decision

$$\underbrace{1 - \frac{b_{02} p_{22}}{\bar{\mu} p_{02}}}_{Default} = \omega_0 + \omega_1 \bar{\mu} \underbrace{U'^{\alpha}\left(\frac{b_{02}}{p_{02}}\right)}_{Ut.Mg\ \ Houses\ \ t=0} + \omega_2 \bar{\mu} \underbrace{U'^{\alpha}\left(\frac{b_{02}}{p_{02}} + \frac{b_{12}}{p_{12}}\right)}_{Ut.Mg\ \ Houses\ \ s=1} + \omega_3 \bar{\mu} \underbrace{U'^{\alpha}\left(e_{01} - q_{01}\right)}_{Ut.Mg\ \ Commodities}$$

Where,

$$\omega_0 = 1 - \frac{\lambda \pi_2 p_{22}}{p_{02}(1+\bar{r})\phi}$$

$$\omega_1 = \frac{-1}{p_{02} \lambda \pi_2 (1+\bar{r})\phi}$$

$$\omega_2 = \frac{\pi_1(\phi p_{02}(1+\bar{r}) - p_{12})}{p_{12} p_{02} \lambda \pi_2 (1+\bar{r})\phi}$$

$$\omega_3 = \frac{-(1+r_0)(1-\phi)}{p_{01} \lambda \pi_2 (1+\bar{r})\phi}$$

Household's Default Decision

$$Default = \omega_0 + \sum_{i=1}^{3} \omega_i U_i'$$

- Where U_i for $i = 1, 2, 3$ are *Idiosyncratic Default Incentives*

$$U_1' = \bar{\mu}\, U' \left(\frac{b_{02}}{p_{02}} \right)$$

$$U_2' = \bar{\mu}\, U' \left(\frac{b_{02}}{p_{02}} + \frac{b_{G2}}{p_{G2}} \right)$$

$$U_3' = \bar{\mu}\, U' \left(e_{01} - q_{01} \right)$$

- And ω_i stand for Systemic factors

Household's Default Decision

$$Default = F\left(\underbrace{\lambda, \phi, \pi_s, p_0, p_s, \bar{r}, r_0}_{\substack{Systemic \\ Factors \\ (Regulation, \\ Prices, \\ Expectations) \\ (\omega' s)}}, \underbrace{\bar{\mu}, e_0, q_0}_{\substack{Idiosyncratic \\ Factors \\ (Income, \\ Indebtedness) \\ (U'\text{'s})}}\right)$$

Data Description: Households Situation

Table: Distribution of Households by Income Group (%)

	2007	2008	2009	2010	2011	Total
Stratum 1	26.38	31.28	25.88	29.46	30.45	28.65
Stratum 2	24.97	29.03	31.09	29.7	29.1	28.1
Stratum 3	48.64	39.69	43.03	40.84	40.45	43.25

Note: Stratum 1: percentiles 1-50; Stratum 2: percentiles 51-80; Stratum 3: percentiles 81-100.

Data Description: Distribution of Variables

Table: Mortgage Loans and Delinquency (%)

	2007	2008	2009	2010	2011	Total
Mortgage holders	16.77	13.17	13.70	18.85	15.00	15.90
Defaulted mortgages	8.26	13.82	9.2	8.07	8.87	8.92
Delinquent mortgages (SBIF)	0.97	1.29	1.95	2.01	1.70	1.58

Table: Distribution of Mortgage Characteristics

	p25	p50	p75
Current Loan to Value	24.6 %	45.1 %	67.9 %
Initial Loan to Value	63.6 %	85.0 %	100.0 %
Monthly Installment	CLP$ 95,000 $\sim$ *USD*$ 180	CLP$ 185,000 $\sim$ *USD*$ 350	CLP$ 320,000 $\sim$ *USD*$ 600
Term of Credit (in years)	19	20	20
Age of Debt (in years)	3	6	11

Data Description: Distribution of Variables

Table: Default and Renegotiation in the Sample

	Did not renegotiated	RENEGOTIATED	Total
Paying	74.5 %	16.6 %	91.1 %
DEFAULTED	6.0 %	2.9 %	8.9 %
Total	80.5 %	19.5 %	100 %

Note: Percentages are calculated over the complete group of mortgagors in the sample.

Estimation Methodology

Problems with mortgage delinquency data:

- Defaulting a loan (specially mortgages) is not an usual event
- Statistical procedures can sharply underestimate the probability of rare events
    1. Increasing the size of the sample does not alleviate the bias
    2. The bias of the estimated coefficients tend to underestimate the probability of the rare event
    3. Finite samples aggravate the underestimation problem

Solution:

- Apply *Rare Events Logistic Regression* (King and Zeng, 2001, 2002)
- The procedure corrects bias and variance using auxiliary information (e.g. public records)

Estimation Results 1: No interactions

Dep. Var.: Mortgage Default Dummy	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Idiosyncratic - Demographic Variables				
Number of persons in house	0.211***	0.235***	0.214**	0.213**
	(0.0715)	(0.0826)	(0.0844)	(0.0856)
Income (in logs)	-0.827***	-0.563***	-0.734***	-0.524***
	(0.133)	(0.153)	(0.155)	(0.162)
Primary Education	0.348	0.0446	0.215	0.109
	(0.333)	(0.424)	(0.440)	(0.466)
Tertiary Education	-0.633***	-0.417	-0.554*	-0.397
	(0.241)	(0.291)	(0.308)	(0.309)
Gender	-0.285	-0.214	-0.285	-0.301
	(0.199)	(0.229)	(0.244)	(0.242)
Age 18-35	-0.0218	-0.155	0.0266	0.105
	(0.260)	(0.295)	(0.306)	(0.308)
Age 55-99	-0.699	-0.0802	-0.0578	-0.0759
	(0.603)	(0.675)	(0.696)	(0.683)
Idiosyncratic - Finance Variables				
Negative Shock	1.745***	1.715***	1.668***	1.683***
	(0.207)	(0.245)	(0.255)	(0.256)
Credit Applications Rejected	0.276			
	(0.398)			
Renegotiation	1.352***	1.052***	1.016***	1.073***
	(0.244)	(0.309)	(0.332)	(0.336)

Estimation Results 1: No interactions

Dep. Var.: Mortgage Default Dummy	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Idiosyncratic - Finance Variables				
Negative Shock	1.745***	1.715***	1.668***	1.683***
	(0.207)	(0.245)	(0.255)	(0.256)
Credit Applications Rejected	0.276			
	(0.398)			
Renegotiation	1.352***	1.052***	1.016***	1.073***
	(0.244)	(0.309)	(0.332)	(0.336)
Systemic Variables				
Current Loan to Value		0.253**		
		(0.117)		
Initial House Price (in logs)		-0.416***		-0.450***
		(0.113)		(0.152)
Initial Loan to Value			0.0407	-0.0835
			(0.0249)	(0.0794)
Constant	6.125***	9.095***	4.663**	9.413***
	(1.761)	(2.302)	(2.074)	(2.631)
Observations	1,894	1,446	1,301	1,337

Note: Robust standard errors in parentheses. Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Estimation Results 2: Including Interactions

Dep. Var.: Mortgage Default Dummy	**Model 1**	**Model 2**
Idiosyncratic - Demographic Variables		
Number of persons in house	0.237***	0.231***
	(0.0820)	(0.0819)
Income (in logs)	-0.578***	-0.582***
	(0.150)	(0.149)
Primary Education	0.0564	0.0804
	(0.421)	(0.414)
Tertiary Education	-0.423	-0.425
	(0.290)	(0.290)
Gender	-0.260	-0.247
	(0.228)	(0.227)
Age 18-35	-0.189	-0.195
	(0.298)	(0.298)
Age 55-99	-0.0714	-0.0950
	(0.675)	(0.675)

Estimation Results 2: Including Interactions

Dep. Var.: Mortgage Default Dummy	**Model 1**	**Model 2**
Idiosyncratic - Finance Variables		
Negative Shock	1.678***	
	(0.245)	
Credit Applications Rejected	0.687	0.677
	(0.437)	(0.439)
Renegotiation	1.319***	1.325***
	(0.277)	(0.276)
Systemic Variables		
Initial House Price (in logs)	-0.417***	-0.443***
	(0.112)	(0.111)
Interaction Variables		
Income and current loan to value	0.0199**	0.0195**
	(0.00810)	(0.00814)
Initial House Price and Negative shock		0.102***
		(0.0147)
Constant	9.326***	9.817***
	(2.265)	(2.243)
Observations	1,446	1,446

Note: Robust standard errors in parentheses. Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Final remarks

- We are able to estimate a micro-powered model of mortgage default determinants
- Contrary to the existing literature, we find that interaction between macro and micro factors is key

- Income is an important determinant of the probability of default
- A negative shock in the recent past significantly increases the probability of defaulting a mortgage loan
- Higher housing prices lessen the probability of default (the contrary is problematic)

- A higher value of the interaction between income and current LTV is associated to higher mortgage default
- Also, a higher value of the interaction between origination housing prices and negative budget shocks is associated with higher default rates

- We propose to extend this framework to analyze further financial issues in a more general setting

Appendix: Nominal Flows of the Economy