

Mortgage Availability and Homeownership in the Low Income Housing Market: The Role of Credit Scoring Technologies*

Jaime Luque

ESCP Europe

Timothy Riddiough

University of Wisconsin-Madison

Abstract

We study the role of credit scoring technologies (CSTs) in affecting the flow of capital into the low-income housing market. Local traditional banks (TBs) and remote “originate-to-distribute” non-bank lenders (NBs) compete for creditworthy non-prime borrowers. Because NBs apply CSTs based on hard credit information only, whereas TBs utilize soft as well as hard information, screening becomes increasingly lax as the NB lending sector gains market share vis-à-vis perceived improvements in CSTs. A new characterization of housing boom and bust is provided, occurring when creditworthy non-prime consumers migrate *en masse* into and out of the NB lending market.

Key words: non-prime mortgage lending; credit scoring technology; originate-to-distribute; hard and soft information; homeownership

JEL Classification numbers: D4, D53, G21, R21.

*We thank for comments and suggestions the participants and discussants at the annual meetings of UCLA Conference on Housing Affordability, ASSA-AREUEA, HULM, NUS-IRES Annual Research Symposium, Fed Atlanta Real Estate Finance conference, SAET, Finance Forum, EWGET, DePaul conference in Economics and Finance, and seminars at the University of California-Irvine, University of Southern California, University of Wisconsin-Madison, Southern Methodist University, University of Bern, IE Business School, Bank of Spain, ESCP Europe and ESSEC Business School. Authors' email addresses: jluque@escpeurope.eu and timothy.riddiough@wisc.edu. Send correspondence to Timothy Riddiough, Wisconsin School of Business, Grainger Hall, 975 University Ave; Madison, WI 53706. Jaime Luque acknowledges the support of the BNP Paribas Real Estate Professorship. This research did not receive any other specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

1 Introduction

Rajan (2010) argues that, beginning in the early 1990s, concerns over increasing income inequality and a bedrock faith in the welfare improving benefits of home ownership led U.S. policymakers to expand mortgage lending to lower income, “non-prime” households. Technological developments furthered these efforts, as personal computing allowed for the application of statistically-based default models and credit risk analytics to generate “scientific” credit risk underwriting assessments of loan applicants. These early developments were further enhanced in the early 2000s by the rapid development of internet-based mortgage lending platforms as well as an increasing demand for private-label mortgage-backed securities. The combination of hard – written, quantifiable and verifiable – credit risk information analytics with an originate-to-distribute approach to mortgage production helped change the lending landscape from a “relational” to a “transactional” business.¹

A critical vulnerability in this shift was that there was little historical subprime mortgage loan servicing data with which to empirically estimate the automated credit underwriting models. Instead, prime mortgage loan data and subjective assessment were used, resulting in models that were noisy (low precision) and possibly biased (excessively optimistic). This critical vulnerability became less concerning to most market participants as subprime mortgages performed well on a credit basis through the 1990s and into the middle 2000s. In other words, the risk of model misspecification and error became increasingly neglected as confidence in credit scoring technologies (CSTs) rose (see Gennaioli et al. 2015 and Gennaioli and Shleifer 2018).

The unexpected increase in subprime mortgage default rates in late 2006 and 2007, and then the collapse of the financial market in 2008, brought non-prime mortgage originations to a standstill, with a loss of confidence in the CSTs applied during pre-crisis period. Yet, income inequality problems in the U.S. remain and have in fact worsened, with housing in many U.S. cities becoming increasingly unaffordable (Gabriel and Painter 2018).² This has generated a renewed policy

¹See Stein (2002), Keys et al. (2010), Rajan et al. (2015), Miller (2015) and Liberti and Petersen (2017). Also note that the Fannie-Freddie dominated prime mortgage market has for decades been an originate-to-distribute loan production process, with automated underwriting introduced in the middle 1980s. This was not the case for non-prime mortgage lending, as, until the middle 1990s, lending generally occurred on a local or regional basis, with most of the loans held on the lender’s balance sheet. Finally, see our on-line Appendix, Section A.1, for additional historical background on non-prime mortgage market and credit scoring model development.

²For the U.S., see the National Association of Realtors’ Housing Affordability Index, which doubled in the period

interest in non-prime mortgage lending and the emergence of newer non-bank financial firms such as FinTech lending platforms that apply sophisticated CSTs of their own to underwrite and price mortgage loans (see the 2017 Bank of International Settlements Report and the 2015 March 3 Goldman Sachs Report).³

With this background in mind, the goal of this paper is to construct a model that highlights and captures the general equilibrium effects of how changes in the credit scoring technologies affect access of lower-income non-prime consumers to mortgage loan and owner-occupied housing markets. In our model, the non-prime mortgage market structure is determined by the relative sizes of two distinct loan funding sources for consumers. One source is *traditional banks* (TBs). TBs typically have a local presence and function like “relationship” lenders, having acquired soft information on customer credit quality based on face-to-face interactions and the provision of other financial services such as checking accounts, credit cards and other consumer loans. TBs are subject to constraints that limit the quantity of mortgages they can originate to hold in portfolio in any given period. The second potential source of mortgage loan funding is *non-bank* (NB) *lenders*, which are typically more remote and “transactional” in nature, and which largely “originate-to-distribute” (OTD) their loans into the secondary mortgage market. These non-banks enjoy fewer constraints on their origination activities, due largely to their loan sales activities and reduced regulatory scrutiny.⁴ NB lenders rely on *hard* credit information that is input into a CST to make *yes-no* underwriting decisions. Secondary market investors who buy mortgages from NB lenders also rely on the same hard credit information.

CST is modeled as a Bayesian prior that measures the likelihood of making a credit related classification error in the *yes-no* loan funding decision.⁵ A posterior *belief* is generated using Bayes rule, which combines the CST-based prior with the endogenous proportion of lower- versus

2006-2013 (from 100 to 210 index points) and then has remained until now at historically high levels (around 160 index points). See Luque et al. (2018) for evidence and discussion of the current housing affordability crisis in the US.

³See also “Subprime mortgages make a comeback — with a new name and soaring demand”, CNBC, April 12, 2018.

⁴We follow the definition of the Financial Stability Board (FSB) and use the term *non-bank* (non-depository) *lender* to refer to what has been labeled a shadow bank. FinTech firms are a sub-category of shadow banks, also following our definition of non-banks. Other common labels for non-banks prior to the financial crisis include conduit lenders, distance lenders, non-integrated banks, narrow banks, and arm’s length lenders.

⁵Statistically-based metrics such as model goodness-of-fit as well as more qualitative factors can affect priors, where qualitative factors are likely to play a more important role when historical lending experience is limited, as it was in non-prime mortgage lending prior to the onset of the recent financial crisis.

higher-risk types in the population of non-prime loan applicants. The resulting posterior measures the credit quality of the NB lender's mortgage loan pool, and thus provides a basis for calculating the credit risk adjusted mortgage loan rate.

The source of mortgage funding, and hence the loan market configuration, is endogenously determined. The fundamental trade-off with respect to the source, cost and quantity of mortgage funding is, on the one hand, the availability of soft credit risk information in the case of TBs versus access to greater secondary market liquidity in the case of NBs. Due to being endowed with superior credit quality information, TBs lend to lower-risk borrowers and thus enjoy lower rates of default, which reduces the cost of mortgage credit for the consumer. On the other hand, depending on the perceived classification precision of available CST, NBs expect to experience a higher rate of default on the loans they originate. This causes the pooled mortgage loan rate to increase relative to the loan rate offered by the TB. But secondary market investors are more patient than originating lenders, which lowers the cost of funds for NBs to partially or wholly offset the credit risk spread caused by CST-based classification error.

TBs and NBs post mortgage rates in their attempt to attract creditworthy loan applicants. Consumers see those posted rates and, conditional on their risk profile, endogenously sort to their preferred mortgage funding source. A particularly interesting equilibrium mortgage lending structure results when NBs are the preferred source of loan funds. Due to capacity limits that constrain TB origination volume, some creditworthy consumers are rationed out of the TB loan market. Once rejected, depending on the cost of NB mortgage capital, these borrowers will either reapply for a NB loan or decide to rent. If home ownership is preferred to renting, and the creditworthy loan applicant is classified as such, the consumer obtains a mortgage loan. But this loan is at a higher cost than a loan obtained from a TB, resulting in a smaller house and lower consumer welfare. When the creditworthy borrower is erroneously classified as high risk, and thus denied credit from both mortgage loan sources, the consumer has no other option but to rent. Thus, as part of the same equilibrium regime, creditworthy applicants might either borrow from a TB and own a larger home, borrow from a NB and own a smaller home, or rent.

With a focus on the credit scoring, we use our two-period model to characterize equilibrium market configurations as three alternative mortgage funding-house price regimes. In a first regime, traditional bank lending is dominant, as credit scoring technologies are sufficiently prim-

itive so as to make the NB mortgage loan market unattractive to creditworthy borrowers. In this regime, credit worthy borrowers either own a house with mortgage funding obtained from a TB or rent, where the rental outcome results from binding TB capacity constraints.

With improvements in CST associated with updated beliefs as to its credit risk classification accuracy, a NB lending market emerges in a second regime as the back-up choice for creditworthy borrowers that are rationed out of the preferred TB loan market. Imperfect classification precision creates the opportunity for high-risk borrowers to apply for a mortgage loan in this regime, hoping to be misclassified as a low-risk type. The loan application acceptance rate and pooled mortgage loan rate are endogenously determined in the NB mortgage lending sector. As noted earlier, creditworthy households' first preference in this case is to own by borrowing from a TB, then to own by borrowing from a NB, and rent as a last resort.⁶ House prices accelerate when the NB loan market emerges, with home ownership rates expanding rapidly due to availability of secondary market mortgage capital. This direct link between aggregate non-prime mortgage supply and house prices is in accordance with the empirical findings of Mian and Sufi (2009) and Gabriel and Rosenthal (2015).

A third regime occurs when CST precision improves sufficiently so that the NB loan is preferred over the TB loan by creditworthy borrowers. At the regime boundary, lower-risk consumers migrate *en masse* from the TB to NB lending sector. This migration endogenously increases the proportion of low-risk borrowers applying for a NB loan, causing the cost of mortgage capital (quantities) to experience large discrete decreases (increases). House prices also increase rapidly at the regime boundary, registering a large discrete jump that signifies a housing market "boom". The TB sector does not completely vanish, however, since there will be credit-worthy borrowers that are rejected in the NB market and who then use the TB market as their back-up funding choice. These results are consistent with Ashcraft et al.'s (2012) finding that, at the onset of the financial crisis, the volume of credit intermediated by the NB lending system was nearly twice that of credit intermediated by the traditional banking system.

A housing market bust is characterized by reversing the order of the identified regimes. Busts happen when NB/secondary market investors recalibrate downward the perceived classification

⁶Our result that an improvement in hard credit scoring technology can lead to increases in the quantity of lending and also more lending to relatively opaque risky borrowers is similar to the effects of the small business credit scoring on commercial bank lending, as empirically documented by Berger et al. (2005).

accuracy of the CST. In general, the cause of the downward revision in CST classification precision can come from several sources, including new information indicating that their models failed to accurately predict failure (Case 2008, Rajan et al. 2015).⁷ Separately or together with downward revisions to CST precision, a recalibration could also occur in the assessed proportion of lower versus higher risk consumers in the potential loan applicant pool. This occurs, for example, when overly optimistic expectations regarding upward income mobility of non-prime consumers are revised downward (Mian and Sufi 2017). This bust scenario matches rather closely the market's reaction when early term defaults hit the subprime mortgage market in late 2006 and early 2007 (see Brunnermeier 2009, pp. 81-83), raising, among other things, concerns over the possible misrepresentation of borrower incomes stated in loan applications (Ambrose et al. 2016, Mian and Sufi 2017).

Our characterizations of belief formation, the resulting effects of the credit scoring channel on mortgage and housing markets, the centrality of borrower income and employment to credit assessment, and the post-crisis decline of subprime lending all tie closely to the empirical findings of Cotter et al. (2015). Our model also endogenizes “lax screening” of the type identified by Keys et al. (2010), Purnanandam (2011) and Rajan et al. (2015), where increasingly lax screening follows from the introduction and growth of a secondary mortgage market that relies on hard information only in the credit risk assessment process.

Our model posits limited lender recourse in the case of borrower default, allowing the borrower to keep a minimum subsistence rent if default occurs.⁸⁹ In lieu of debt repayment, the lender seizes the house as pledged security as well as income, if any, in excess of that needed to meet the subsistence rent payment. Limited recourse functions as a credible commitment

⁷As Chip Case explained in the fall of 2008, “the problem was that the regressions on which the automated underwriting systems were based had been run with data from a 20-year period of continuously rising national home prices... [T]he model concluded that as long as a portfolio was regionally diversified and pricing was based on credit scores, loan-to-value ratios, and so forth, the business would be profitable. When instead home prices declined everywhere and regional cycles became more synchronized, the model no longer fit the data.” See also Cotter et al. (2015) for more on the issue of regional house price synchronization.

⁸Limited recourse mortgage contracts are representative of the non-prime mortgage market in most states in the US. According to Rao and Walsh (2009), only three states out of 50 exclude any form of recourse to help lenders recover mortgage loan balance repayment shortfalls associated with residential mortgage default and foreclosure.

⁹Our large economy allows us to prove equilibrium existence even in the presence of this non-convexity in the consumer's budget constraint. See also Poblete-Cacenabe and Torres-Martinez (2013) for a general equilibrium model with a continuum of consumers and limited recourse contracts, where equilibrium exists for any continuous garnishment rule and multiple types of reimbursement mechanisms.

device for creditworthy consumers, who basically pledge their excess income to the lender by purchasing a larger house. This increases mortgage proceeds and hence consumer welfare, and follows because these consumers can afford to repay a larger loan. In contrast, non-creditworthy consumers that (erroneously) receive mortgage funding lack the income to meet their repayment promise, and thus default on the loan.¹⁰ The introduction of limited recourse lending thus allows creditworthy consumers to better identify themselves in secured credit markets, where CST can be applied to further aid in credit risk type identification.¹¹

Because non-prime loan contracts are limited recourse, house prices *per se* – which are endogenously determined in general equilibrium and changing with the assessed classification precision of the CST – do *not* determine the net worth of a consumer type. This feature departs from Brueckner et al. (2012) in that the CST becomes income-only driven in our model (it screens good and bad types by looking at their second period endowments only), which is sufficient to generate the collapse of the subprime conduit mortgage market (the collapse occurs because good type consumers prefer renting over owning a house when NB loan terms are prohibitive). In Brueckner et al. mortgage contracts are non-recourse, with borrower’s default risk being observable.¹² This allows mortgage lenders’ expectations of a high house price growth to increase the consumer’s FICO score through current prices, which in turn spurs subprime lending.

In a related paper, Brueckner et al. (2016) study the link between house price expectations and the use of alternative mortgage products. They argue that, as future price gains are perceived as more likely, the riskiness of alternative contracts lessen to encourage their use. In our model, the choice is between a TB loan and a NB loan, not the type of mortgage contract. As perceived CST classification precision improves and secondary market liquidity increases, house prices increase (prices are simultaneously endogenously determined in general equilibrium) and consumers find NB loans more attractive. The mix of loans by originator type (rather than mortgage contract

¹⁰Full non-recourse (losing only the house in default), without an ability to pledge additional income or assets as security, would be incapable of generating credible commitment on the part of low-risk consumers. This in turn would significantly reduce loan proceeds - even when credit risk type is known by the lender with certainty - and hence housing consumption, to neuter incremental benefits associated with the application of CST by NBs. See Geanakoplos and Zame (2014) for a general equilibrium model with non-recourse collateralized loans.

¹¹Our setting with limited recourse mortgage lending also departs from those models that consider uncollateralized lending, such as credits cards; see, for example, Athreya et al. (2012) and Chatterjee et al. (2011).

¹²See Brueckner (2000) for an adverse-selection model of mortgage lending with two types of borrowers that differ in “default costs”.

type) changes as a result.

Another feature of our model is that it endogenizes financial markets to affect the local urban economy. In particular, structural details of the mortgage contract design and financial market industrial organization are shown to feed back to affect the consumer's rent versus own decision. We also show how local housing construction quality and minimum lot size regulations, which effectively impose a minimum house size constraint, create formidable barriers to home ownership for non-prime households. Housing affordability problems in lower income areas are thus, in part, the result of local regulation, as without minimum house size constraints, lower income households that are excluded from home ownership could otherwise purchase smaller houses with smaller mortgage loan amounts.

We also consider the effects of surreptitious acquisition of soft information by NBs, which is used by the NB loan originator to select against secondary mortgage market investors. Soft information acquisition by the NB loan originator improves credit risk classification precision, thus diminishing classification error. Loans that are reclassified from accepts to rejects (reclassified lemons) provide the opportunity for adverse selection, since secondary market investors believe those loans to be acceptable based on hard information only. We show that, in certain cases, the credit quality of the resulting sold loan portfolio can actually exceed that based on hard information only. We also demonstrate how adverse selection of this type exacerbates boom and bust in housing markets, due to the increased aggregate quantity of mortgage loans originated. Another insight from this extension of the model is that the endogenously determined quantity of soft information acquired decreases in the effective mortgage distribution rate.

As noted earlier, our paper contributes to the literature of lax screening in secondary mortgage markets – see, e.g., Keys, Mukherjee, Seru, and Vig (2010), Purnanandam (2011) and Rajan et al (2015). In particular, we show how (perceived) improvements to CST expand the market share of NBs which rely on only hard information in making loan underwriting decisions. Thus greater market share of non-bank lenders implies increasingly lax screening as defined in the literature. When we add the possibility of adverse selection by NB loan originators on loans sold into the secondary market, lax screening by secondary market investors implies an opportunity for NB loan originators to (potentially) further dilute the credit quality of a loan pool below that which would occur based on underwriting with hard information only.

The rest of this paper is organized as follows. In Section 2, we present the model. Section 3 provides the equilibrium definition and states the equilibrium existence result. Section 4 discusses the characteristics of equilibrium mortgage rates and characterizes equilibrium mortgage and housing market configurations as they depend on the available CST. Section 5 examines the role of CST and other key parameters in describing the evolution of non-prime mortgage lending in the US, and characterizes house price boom and bust in relation to equilibrium regime changes. Section 6 elaborates on the role of lax screening and income misrepresentation under the lens of our model. This section also examines other channels that can trigger a regime change, including secondary market liquidity and the fundamental proportion of low-risk consumers. Section 7 extends the model to allow for non-bank lenders to surreptitiously acquire soft information regarding loan credit quality, while secondary market investors limit their attention to mortgage prices and loan quantities they expect to see in the secondary market. Section 8 concludes. The Appendix is reserved for additional explanations, results, and proofs.

2 Model

In our economy there is a continuum of lenders, investors and consumers that we collectively refer to as agents. All agents live for two periods, $t = 1, 2$, with agent choices and market clearing occurring at the beginning of each period. We will also refer to consumers as households. Our focus is on what we call “non-prime” households. Non-prime households do not meet “prime” credit quality underwriting standards set by Fannie Mae and Freddie Mac. More generally, non-prime loans can be thought of and categorized as “below investment grade” in terms of overall credit quality. For simplicity, we ignore the housing and mortgage markets of prime households, and assume that non-prime households cluster together in what we call “non-prime neighborhoods”.

There are two consumption goods: owner-occupied housing and the numeraire good. There are positive endowments of the numeraire good in both periods, although possibly in different amounts. We take the aggregate supply (demand) of the owner-occupied housing good in the first (second) period as exogenously given, and equal to $\bar{H} = 1$.¹³ Later, in Section 4, we will explain

¹³The assumption of inelastic housing supply is realistic given that most non-prime neighborhoods are located in

that our baseline model naturally extends to an overlapping generations economy under specific assumptions on the consumer's utility function.

In period 1 all households are endowed with a subsistence rent (SR) equal to $\omega^{SR} > 0$ units of the numeraire good. One can think of ω^{SR} as a provision that ensures all households access to basic shelter in the rental housing market. For modeling purposes assume the endowment is fungible, in the sense that it can also be used to fund a down payment on a owner-occupied house should the non-prime household qualify for a mortgage.

We follow Mian and Sufi (2009), Ambrose et al. (2016), and others with a focus on household income as it affects a household's ability to obtain a non-prime mortgage and to successfully make the required mortgage payments. In particular, assume in the second period a certain proportion of consumers are expected to experience a positive income shock (e.g., get a better job/salary), with $\omega^+ > \omega^{SR}$, while the rest of the pool remains at their current income level ω^{SR} . Label the consumers that experience an increase in their second period endowment as G-type and those who don't as B-type. Consumers know their type in period 1, but are unable to verifiably convey that information to outside parties on their own. The measures of types G and B consumers in the economy are λ_G and λ_B , respectively. We refer to the ratio $\lambda_G/(\lambda_G + \lambda_B)$ as the *fundamental proportion* of G-type consumers in the economy.

The owner-occupied housing market is subject to a minimum house size H^{\min} that eliminates small owner-occupied houses that could otherwise rely exclusively on the numeraire good endowment to fund ownership.¹⁴ We shall see that, in equilibrium, the minimum house size restriction implies that, in order to buy a house in the first period, a consumer needs a mortgage (the subsistence rent endowment is not enough to fund a house purchase).

Before presenting the remaining details of our model, some useful notation is introduced. We denote an agent by a and the set of all agents by \mathbf{A} . If the agent is a household/consumer (independently of its type), we write h , if it is a lender (independently of its type), we write l , and if it is a secondary market investor, we write i . We further represent the non-atomic measure space of agents in our economy by $(\mathbf{A}, \mathcal{A}, \lambda)$, where \mathcal{A} is a σ -algebra of subsets of the set of agents

previously developed urban areas; see Mian and Sufi (2009) for further discussion and evidence. For an equilibrium model of affordable housing development with financial subsidies to developers, see Luque (2018).

¹⁴See Green and Malpezzi (1996) for a discussion of how minimum house size regulations are applied in the U.S. and other countries, and how consumption standards, such as minimum lot sizes, exclude low-income groups. See also NAHB Research Center (2007) and the Wharton Housing Regulation Index for measures of housing regulation.

in \mathbf{A} , and λ is the associated Lebesgue measure. We will also denote the subsets of households, lenders, and investors by \mathcal{H} , \mathcal{L} , \mathcal{I} , respectively; the subsets of G- and B-type households by \mathcal{G} and \mathcal{B} , respectively; and the subsets of TBs and NBs by \mathcal{TB} and \mathcal{NB} , respectively. For simplicity, we will assume $\lambda(\mathcal{TB}) = \lambda(\mathcal{NB}) = \lambda(\mathcal{I}) = 1$. For the sake of presentation, we will also write $\lambda(\mathcal{G}) \equiv \lambda_G$ and $\lambda(\mathcal{B}) \equiv \lambda_B$.

2.1 Lenders, Credit Scoring, and Mortgage Distribution

Funding sources: To obtain a mortgage, consumers may attempt to borrow from two different types of funding sources: a local *traditional bank* (TB) or a remote *non-bank lender* (NB). The main differences between the two are the following:

- TBs are local lenders with access to soft information on customer credit quality based on time-consuming personal interactions (those personal interactions that generated memorable soft information occurred sometime in the past, with acquisition cost being sunk). NBs are remote lenders that have had no documented historical relations and that rely on only hard information to make lending decisions. Hard information is measurable as well as easy to store and transmit in impersonal ways available to all agents at no cost.¹⁵ Soft information cannot be reduced to a numerical score nor is easily transferable to outsiders (Liberti and Petersen 2017).¹⁶ We assume that the combination of hard and soft information allows TBs to know the consumer's type with certainty.¹⁷
- TBs are subject to a significantly higher regulatory burden than NBs (see Buchak et al. 2018 for empirical evidence, as well as Fuster et al. 2018 for analysis in the context of

¹⁵Examples of hard information are FICO scores, current borrower employment and income information, fixed payment obligations, certain known neighborhood characteristics, and publicly available information on local economic growth prospects. These factors and others may also interact with one another to affect the credit quality of the loan applicant.

¹⁶See also Stein (2002), Keys et al. (2010), Rajan et al. (2015), and Miller (2015) for more on the role of soft information in non-prime mortgage lending. See Adams et al. (2009) for work that focuses on how different lenders' information sets affect mortgage loan outcomes, borrowers' default, and market unraveling.

¹⁷We note that complete information on the part of the TB is not required for our results to go through - all that is required is that the TB possesses at least some soft information that the NB and secondary market investors do not have. We make the complete information assumption with no real loss of generality and only to simplify the analysis.

FinTech mortgage lending¹⁸). As a result, TBs face capacity constraints and therefore may not be able to serve all mortgage applicants.

- NB lenders are “transactional” in nature – they largely “originate-to-distribute” (OTD) –, whereas TBs keep a significant higher share of their mortgages in portfolio. For example, according to Buchak et al. 2018, TB mortgage loan retention share has declined since 2007, but has never declined below 30%).

Credit scoring: Both TBs and NBs only originate a mortgage if the consumer is identified as a G-type, i.e., rating=G. This loan origination policy is endogenously determined, and follows because of the presence of limited recourse mortgage contracts and a minimum house size (discussed below). We define the *credit scoring technology* (CST) that is applied by lender $l = TB, NB$ as $CST^l = \Pr^l(\text{rating=G}|\text{G})$. Generally speaking, CST^l is a subjective probability that is founded on credit experience, possibly combined with statistical as well as qualitative modeling analytics, to form a belief that quantifies the technology’s predictive power in generating a *yes-no* lending decision. As such, CST^l is not necessarily purely statistically based, where qualitative factors such as “confidence” or other behavioral factors can exert first-order effects in belief formation (see Gennaioli and Shleifer 2018 for insights on the role of beliefs on the recent financial crisis, and how beliefs shaped financial markets, leading to expansions of credit and leverage). To simplify the analysis, we assume that $\Pr^l(\text{rating=G}|\text{G}) = \Pr^l(\text{rating=B}|\text{B})$.

Mortgage loan credit quality based on classification accuracy is determined using Bayes’ rule, which is a posterior probability that measures the lender’s belief as to the credit quality of the originated mortgage loan pool. This posterior probability can be written as follows:

$$\Pr^l(\text{G}|\text{rating=G}) = \frac{CST^l \cdot \hat{\pi}_G^l}{CST^l \cdot \hat{\pi}_G^l + (1 - CST^l) \cdot \hat{\pi}_B^l} \quad (1)$$

where $\hat{\pi}_G^l$ denotes the proportion of G-type consumers among all consumers that attempt to borrow from lender l . This posterior probability provides the basis for loan pricing conditional on loan approval.¹⁹ By assumption, NBs have incomplete information regarding consumer type

¹⁸Much of the growth in the post-crisis non-prime market has been attributed to “Fintech”, which are firms that generally operate as on-line mortgage originators and that leverage alternative sources of information and the “big data” approaches inherent in technology-based lending.

¹⁹This measure corresponds to Rajan et al.’s (2015) interest rate formulation as stated in their equations (2) and (4), where ours is structured to depend specifically on the credit scoring technology and the relative proportion of

in the soft information domain, but complete information in the hard information domain. For simplicity, as noted earlier, we assume $CST^{TB} = 1$ to capture the idea that the TB's access to soft information allows it to perfectly screen the type of borrower. Because non-banks lack soft information, we assume $CST^{NB} < 1$.

To simplify our notation, we write the lender l 's posterior measure of credit quality of the mortgage pool

$$\pi^l \equiv \Pr^l(G|\text{rating}=G).$$

Importantly, this belief π^l is endogenous in our model, as it depends on the endogenous variable $\hat{\pi}_G^l$. This is a fundamental difference from Frankel and Jin (2005) and Chatterjee et al. (2011), where the lender's belief (or "credit scoring function" in Chatterjee et al.'s terminology) is taken as a parameter and thus independent of the market structure. Our use of the Bayes' rule is essentially static, where the credit scoring technology (rather than the credit score itself) enters as a parameter. Chatterjee et al.'s dynamic model of reputation acquisition follows instead a "Bayes inference" approach, where the individual's type score is updated every period given his actions in the credit market. The novelty in our approach is that the proportion of lower-risk consumers applying for a loan in a particular market segment is endogenous, which in turn makes lenders' posterior beliefs on the credit quality of their loan pools, and therefore credit scoring, endogenous.

Given the credit scoring technology CST^l , the measure of consumers that are able to borrow from lender $l = TB, NB$ is given by

$$\lambda(G\text{-Rating}^l) = CST^l \cdot \hat{\lambda}_G^l + (1 - CST^l) \cdot \hat{\lambda}_B^l,$$

where $\hat{\lambda}_G^l$ and $\hat{\lambda}_B^l$ are the measures of G-type and B-type consumers that apply for a loan in market l . For each mortgage market $l = TB, NB$, measures $\hat{\lambda}_G^l$ and $\hat{\lambda}_B^l$ are endogenously determined in equilibrium, consistent with the (endogenous) proportion of G-type consumers in market l .

Mortgage distribution: A lender type $l = TB, NB$ is not only characterized by its CST, but also by its mortgage distribution rate d^l . This parameter is such that $0 \leq d^{TB} < d^{NB} \leq 1$.

G-types in the loan application pool.

The second inequality corresponds to a model restriction that TBs keep a higher share of their mortgages in portfolio than NBs (Buchak et al. 2018). Also notice that in practice, d^{NB} is often equal to 1 for NBs (their business model is originate-to-distribute). But it can be smaller than 1 if, for example, NBs promise to keep some “skin-in-the-game” or are subject to regulation that requires such a practice.

By letting $\varphi^l \geq 0$ denote the face value of mortgages originated by lender l and $z^l \geq 0$ the face value of mortgages originated by lender l that are sold into the secondary market, we can write:

$$z^l = d^l \varphi^l, \quad (2)$$

Roughly speaking, equation (2) says that lender l distributes the promised fraction d^l of its originated mortgage payments into the secondary mortgage market. Equation (2) together with parameter values $0 \leq d^{TB} < d^{NB} \leq 1$, $CST^{TB} = 1$, and $CST^{NB} < 1$ define the two possible loan origination models.

Lender’s optimization problem: We assume that all lenders are risk neutral with time discount factor $\theta^l < 1$. Lenders’ objective function is of the form $\Phi^l = x_1 + \theta^l x_2$, where x_1^l and x_2^l represent the lender l ’s consumption of the numeraire good at dates 1 and 2, respectively. The assumption of lender risk neutrality is common in the literature; see e.g. Arslan et al. (2015), Chatterjee et al. (2011), Dell’Ariccia and Marquez (2006), Frankel and Jin (2015), Guler (2015), and Fishman and Parker (2015). Lender l is subject to budget constraints $x_1^l \leq \omega_1^l - q^l \varphi^l + \tau z^l$ and $x_2^l \leq (1 - d^l)(\pi^l \varphi^l + (1 - \pi^l)\lambda(G-Rating^l)\delta p_2 H_1^G)$ at dates 1 and 2, respectively. Because the lender’s objective function is strictly increasing in the consumption of the numeraire good, we can write the budget constraints with equality. Thus,

$$\Phi^l(\varphi^l) \equiv (\omega_1^l - q^l \varphi^l + \tau z^l) + \theta^l(1 - d^l)(\pi^l \varphi^l + (1 - \pi^l)\delta p_2 H_1 \lambda(G-Rating^l)) \quad (3)$$

Here q^l denotes the discounted mortgage price for lender l , which reflects the cost of mortgage loan capital, φ^l is the lender l ’s mortgage face value (which is promised to be repaid in period

2) and $q^l\varphi^l$ is the mortgage loan amount that is funded in period 1.²⁰ Notice that, because the measure of each type of lender is 1, φ^l and $q^l\varphi^l$ represent the aggregate loan repayment promise required and amount of credit offered by lenders of type l . The discount price of the mortgages sold by lender l in the secondary mortgage market is τ^l .

The lender's first period endowment of the numeraire good, ω_1^l , is assumed to be positive, and, for simplicity, we assume the second period endowment equal to $\omega_2^l = 0$.

The lender incurs direct foreclosure costs associated with repossessing the house and reselling it. This results in a loss to the lender of $(1 - \delta)p_2H_1$ per foreclosure, where $\delta \in [0, 1]$ denotes the foreclosure recovery rate, p_2 is the price of owner-occupied housing in the second period, and H_1 is the house size purchased by a G-rated consumer in the first period. Both H_1 and p_2 are endogenously determined in the model.

The term $\pi^l\varphi^l + (1 - \pi^l)\lambda(G\text{-Rating}^l)\delta p_2H_1$ represents the expected payoff corresponding to the lender's pool of mortgages kept in its own portfolio. With probability $1 - \pi^l$ the lender expects to have originated a (misclassified) mortgage loan to a B-type, so $(1 - \pi^l)\lambda(G\text{-Rating}^l)$ is the measure of borrowers that lender l expects to be of B-type. B-types always (endogenously) default on their mortgage. For each of these B-type consumers, the lender recovers only the depreciated value of the house, δp_2H_1 , in default. Also, notice that in the expression for x_2^l , the lender keeps a fraction $1 - d^l$ of its originated mortgages. Consequently, we weight any income received in the second period by this term.

Type l lenders take $\lambda(G\text{-Rating}^l)$ and π^l as given, as well as prices q^l , τ^l and p_2 , and parameters θ^l and d^l . Lender l 's optimization problem then consists of choosing φ^l to maximize $\Phi^l(\varphi^l)$.

2.2 Secondary Market and the Credit Score Transmission Process

Denote the secondary market (SM) investor i 's consumption bundle of the numeraire good by $x^i = (x_1^i, x_2^i) \in \mathbb{R}_+^2$. Investor i is subject to the following budget constraints:

$$(x_1^i, x_2^i) \leq (\omega_1^i - \tau^{TB}z_{TB}^i - \tau^{NB}z_{NB}^i, z_{TB}^i + \pi^i z_{NB}^i + (1 - \pi^i)\lambda(G\text{-Rating}^{NB})d^{NB}\delta p_2H_1^G),$$

²⁰For example, if the discount price is $q^l = 0.8$ and the mortgage face value is $\varphi^l = \$100,000$, the loan amount is $\$80,000$.

where $\omega_1^i > 0$ indicates the SM investor's endowment of the numeraire good in period 1 (again, for simplicity, we assume $\omega_2^i = 0$). An investor can buy both risk free TB loans (originated with soft information) and risky NB loans (originated with hard information only). The term $\tau^{TB} z_{TB}^i$ corresponds to the cost of purchasing an amount z_{TB}^i of TB loans, which are sold at discount price τ^{TB} in period 1 and pay z_{TB}^i in the second period. The term $\tau^{NB} z_{NB}^i$ is the cost for purchasing an amount z_{NB}^i of NB loans at discount price τ^{NB} . The SM investor's expected second period revenue from purchasing NB mortgages in the first period is $\pi^i z_{NB}^i + (1 - \pi^i) \lambda(G\text{-Rating}^{NB}) d^{NB} \delta p_2 H_1^G$. Consistent with the NB's expectations, the investor anticipates a fraction π^i of the NB loans purchased being G-type loans with total required payment z_{NB}^i . The term $(1 - \pi^i) \lambda(G\text{-Rating}^{NB})$ is the measure of NB borrowers with a G-rating that the investor expects to be of B-type. As before, $d^{NB} \delta p_2 H_1^G$ denotes the recovery that an investor gets for each of these NB mortgages.

Investors also have linear objective functions with functional form $\Lambda^i = x_1^i + \theta^i x_2^i$, where θ^i is the investor's time discount factor. Because this function is strictly increasing in the consumption of the numeraire good, we can write the SM investor i 's optimization problem as follows: investor i chooses z_{TB}^i and z_{NB}^i that maximize

$$\Lambda^i(z^i) \equiv \omega_1^i - \tau^{TB} z_{TB}^i - \tau^{NB} z_{NB}^i + \theta^i (z_{TB}^i + \pi^i z_{NB}^i + (1 - \pi^i) \lambda(G\text{-Rating}^{NB}) d^{NB} \delta p_2 H_1^G), \quad (4)$$

We assume that $\pi^{NB} = \pi^i < 1$, implying credit scores are based on hard information only and that they are transmitted to SM investors truthfully and accurately.²¹ In addition, we assume that investors assign a smaller relative weight to period 1 consumption than lenders do, i.e., $\theta^l < \theta^i$. Greater patience attributable to SM investors can be interpreted as a measure of liquidity in the secondary market (see also Frankel and Jin 2015; see our discussion in the Appendix for anecdotal evidence). Greater liquidity implies higher relative prices paid for mortgage loans sold into the secondary market.

²¹Later in Section 6 we will address issues associated with the misrepresentation of hard information and the possibility of adverse selection by NBs into the secondary market based on soft information acquisition.

2.3 Households

In each period $t = 1, 2$, consumers decide whether to buy the numeraire good (R_t) or owner-occupied housing (H_t), or conceivably some combination of the two.²² If a consumer buys a house, he will borrow from either a TB or a NB in the amount of $q^{TB}\psi^{TB}$ or $q^{NB}\psi^{NB}$, respectively, where $\psi^l \geq 0$ denotes consumer's mortgage face value (i.e., promise to pay) in period 2 to lender $l = TB, NB$. As usual, equilibrium existence requires an upper bound $\Psi > 0$ on ψ^l :

$$\psi^l \leq \Psi \quad (5)$$

Thus, in period 1 a non-prime consumer has three possible choices: (1) borrow in the TB loan market to own a house, (2) borrow in the NB loan market to own a house, or (3) not borrow and consume his endowment of the numeraire good (e.g., subsistence housing). We denote these possibilities by m^{TB} , m^{NB} , and m^\emptyset , respectively, and the set of consumer's "market choices" by $\mathbf{M} = \{m^{TB}, m^{NB}, m^\emptyset\}$. The consumer's market choice is consumer-type ($c(h) = G, B$) and market-type ($l \in \mathbf{L} \equiv \{TB, NB, \emptyset\}$) specific, and is denoted by $m_{c(h)}^l \equiv (c(h), l)$. When a consumer's market choice is m^\emptyset , we write $\psi^\emptyset = 0$.

The period 1 budget constraint of a consumer with market choice m^l , $l = TB, NB, \emptyset$, is:

$$p_1 H_1 + R_1 \leq q^l \psi^l + \omega^{SR} \quad (6)$$

where p_1 is the per unit house price in period 1. Observe that the consumer's mortgage down payment is endogenous in our model; for example, if $R_1 = 0$, then the maximum down payment is equal to $\omega^{SR}/p_1 H_1$.

The second period budget constraint depends on two things. First, we assume that if a consumer buys a house in period 1, then the same house enters in period 2 budget constraint as an asset endowment evaluated at market price p_2 (i.e., owner-occupied housing equals to a long term contract that once signed is valid for two periods). However, a consumer that buys numeraire good R_1 can only consume it for one period (the numeraire good is perishable). Second, loans are subject to a *limited recourse mortgage* contract that stipulates that a borrower is allowed to

²²As it is usual in general equilibrium theory, we allow for any convex combination between choice variables to guarantee equilibrium existence, which requires that the consumer's consumption budget set has convex values.

consume his subsistence income, ω^{SR} , and no more than his subsistence income, if default occurs.²³ Accordingly, we write the second period budget constraint of a consumer with market choice m^l , $l = TB, NB, \emptyset$, as follows:

$$p_2 H_2 + R_2 \leq \max\{\omega^{SR}, \omega_2^c + p_2 H_1 - \psi^l\} \quad (7)$$

where ω_2^c denotes the period 2 endowment. The endowment depends on borrower type $c = G, B$, where $\omega_2^G = \omega^+$ and $\omega_2^B = \omega^{SR}$. The term $p_2 H_1$ in the right hand side of the inequality (7) reflects the value of the house purchased in the previous period, and is interpreted as a sale at market price p_2 per house size unit. The consumer can then use the proceeds of this sale for consumption after repaying his mortgage.²⁴ The maximum operator in (7) determines whether the household defaults, in which case he only consumes the minimum subsistence income ω^{SR} , or honors the loan promise, in which case he consumes at least the minimum subsistence income ω^{SR} .²⁵ There is no default if p_2 , H_1 , and ψ^l are such that $\omega^{SR} \leq \omega_2^c + p_2 H_1 - \psi^l$.

A contribution of this research is that the mortgage market structure is endogenous in our model. In order to incorporate consumer market choices, we need to introduce the following notation. Define the consumer's market choice function by $\mu : \mathbf{A} \rightarrow \mathbf{D}$, where $\mathbf{D} = \{\iota(m_{c(h)}^l) : \iota(m_{c(h)}^l) = 1 \text{ if consumer } h \text{ chooses } m_{c(h)}^l, \text{ and } 0 \text{ otherwise}\}$. We require that a consumer can only choose one market in \mathbf{M} (i.e., $\sum_{l=TB, NB, \emptyset} \iota(m_{c(h)}^l) = 1$). With this notation, we can write the measures of G-type and B-type consumers that attempt to borrow from lender $l \in \{TB, NB\}$ as $\hat{\lambda}_G^l \equiv \lambda(\mathcal{H} : c(h) = G, \mu^h(m_G^l) = 1)$ and $\hat{\lambda}_B^l \equiv \lambda(\mathcal{H} : c(h) = B, \mu^h(m_B^l) = 1)$. Notice that $\hat{\lambda}_G^l$ and $\hat{\lambda}_B^l$ are endogenously determined in our model, and therefore so is $\lambda(G\text{-Rating}^l)$.

Also, denote the household h 's consumption bundle by $x^h = (H_1^h, R_1^h, H_2^h, R_2^h) \in \mathbb{R}_+^4$. We

²³Mortgage laws in most states in the US allow for partial recourse. This means that, in addition to the pledge of the house as security for the loan, lenders secure additional borrower income or assets to fill any remaining gap between the net value of the house in foreclosure and the amount owed on the mortgage. In our model, in addition to the house, period 2 income is the one available form of security that can be pledged in support of the loan. See the Appendix for additional discussion and also for the difference between limited recourse and non-recourse mortgage contracts.

²⁴A consumer with an owner-occupied house at the beginning of period 2 decides whether to sell it at market price, or to consume it. The latter is equivalent to the joint transactions of selling the house the consumer owns at the beginning of period 2 and then immediately buying it back.

²⁵This optimality condition in which the borrower, subject to the relevant recourse requirements, decides whether mortgage loan payoff to retain ownership of the house or default with house forfeiture generates greater utility. See Davila (2015) for an exhaustive analysis of exemptions in recourse mortgages.

will say a pair (x^h, ψ^h) is feasible for consumer h if it satisfies constraints (5), (6) and (7). Now, we can define the consumer h 's choice vector $(x^h, \psi^h, \mu^h) \in \mathbf{X}^h \subset \mathbb{R}_+^5 \times \mathbf{D}$ as the feasible set of elements (x^h, ψ^h, μ^h) that are consistent with his market choice m^l and constraints (5), (6) and (7). This consumption set correspondence $h \rightarrow \mathbf{X}^h$ is assumed to be measurable.

Since consumption depends on the consumer's access to credit, we write the consumer h 's utility function in terms of his consumption and market choice, i.e., $u^h(x^h, \mu^h(m))$.²⁶ We assume that the mapping $(h, x, \mu) \rightarrow u^h(x, \mu)$ is a jointly measurable function of all its arguments, and that $u^h(\cdot, \mu)$ is continuous, strictly increasing and strictly quasiconcave.

The consumer h 's optimization problem consists on choosing a vector $(x^h, \psi^h, \mu^h) \in \mathbf{X}^h \subset \mathbb{R}_+^5 \times \mathbf{D}$ that maximizes his utility function $u^h(x^h, \mu^h(m))$ subject to his market-choice-dependent constraints (5), (6) and (7).

3 Equilibrium Definition and the Existence Result

In this section we define the equilibrium of our economy with endogenously segmented mortgage markets, and show that such an equilibrium exists.

We use a standard notion of a competitive equilibrium with the additional condition that lenders' beliefs must be consistent with the distribution of consumers in the TB and NB mortgage markets. That is, if we define by $\hat{\mu}(G, l) \equiv \int_{\{\mathcal{H}:c(h)=G\}} \mu^h(c(h), l) d\mu$ and $\hat{\mu}(B, l) \equiv \int_{\{\mathcal{H}:c(h)=B\}} \mu^h(c(h), l) d\mu$ the aggregate of type (G, l) - and (B, l) -choices, respectively, we require beliefs $\pi \equiv (\pi^{TB}, \pi^{NB})$ consistent with the market aggregate choice function $\hat{\mu}$. We formally define this condition before providing the definition of equilibrium.

Let a continuous function $f : (\hat{\mu}(G, l), \hat{\mu}(B, l)) \rightarrow [0, 1]^2$ be such that $f(\hat{\mu}(G, l), \hat{\mu}(B, l)) = (f_G, f_B)(\hat{\mu}(G, l), \hat{\mu}(B, l))$, where $f_G(\hat{\mu}(G, l), \hat{\mu}(B, l)) = \hat{\mu}(G, l) / (\hat{\mu}(G, l) + \hat{\mu}(B, l))$ and $f_B(\hat{\mu}(G, l), \hat{\mu}(B, l)) = \hat{\mu}(B, l) / (\hat{\mu}(G, l) + \hat{\mu}(B, l))$. In addition, let another function $g : (f(\hat{\mu}(G, l), \hat{\mu}(B, l)), CST^l) \rightarrow [0, 1]$ be such that it mimics the belief expression (1), where instead of $(\hat{\pi}_G^l, \hat{\pi}_B^l)$ we use $(f_G, f_B)(\hat{\mu}(G, l), \hat{\mu}(B, l))$. Then, given CST^l , we say that the aggregate market choice vector $\hat{\mu}^l \in \mathbb{R}^M$ is consistent with lender l 's belief π^l if $\pi^l = g(f_G(\hat{\mu}(G, l), \hat{\mu}(B, l)), CST^l)$.

²⁶This modelling approach is common in the club theory literature, where consumers choose the club they want to belong to and the consumption vector constrained to the club choice (Ellickson et al. 1999).

Definition 1: Given the triplet $(CST^{TB}, CST^{NB}, CST^i)$, an equilibrium for this economy consists of a vector of market choices μ , prices $(p_1, p_2, q^{TB}, q^{NB}, \tau^{TB}, \tau^{NB})$ and allocations $((x^i, z_{TB}^i, z_{NB}^i)_{i \in \mathcal{I}}, (x^l, \varphi^l, z^l)_{l \in \{\mathcal{T}B, \mathcal{N}B\}}, (x^h, \psi^h)_{h \in \{\mathcal{G} \cup \mathcal{B}\}})$ such that:

(1.1) Agents solve their respective optimization problems.

(1.2) $\hat{\mu}$ is consistent with π .

(1.3) The following market clearing conditions hold:

$$(MC.1) \int_{\mathcal{H}} \psi^h(m_{c(h)}^{TB}) \mu^h(m_{c(h)}^{TB}) dh = \int_{\mathcal{T}B} \varphi^{TB} dl$$

$$(MC.2) \int_{\mathcal{H}} \psi^h(m_{c(h)}^{NB}) \mu^h(m_{c(h)}^{NB}) dh = \int_{\mathcal{N}B} \varphi^{NB} dl$$

$$(MC.3) \int_{\mathcal{N}B} z^l dl = \int_{\mathcal{I}} z_t^i di, \text{ for } l = TB, NB$$

and, for $t = 1, 2$,

$$(MC.4) \sum_{l \in \mathcal{L}} \int_{\mathcal{H}} R_t^h(m_{c(h)}^l) \mu^h(m_{c(h)}^l) dh + \int_{\mathcal{N}B \cup \mathcal{T}B} x_t^l dl + \int_{\mathcal{I}} x_t^i di = \int_{\mathbf{A}} \omega_t^a da$$

$$(MC.5) \sum_{l \in \mathcal{L}} \int_{\mathcal{H}} H_1^h(m_{c(h)}^l) \mu^h(m_{c(h)}^l) dh = \sum_{l \in \mathcal{L}} \int_{\mathcal{H}} H_2^h(m_{c(h)}^l) \mu^h(m_{c(h)}^l) dh = \bar{H}$$

Theorem 1: An equilibrium, as defined in Definition 1, exists.

We leave the proof of Theorem 1 for the Appendix A.2. This proof is not straightforward, for the following reasons. First, in contrast with standard general equilibrium models where individuals are exogenously assigned to a market, in our model the sizes of the TB and NB mortgage markets are endogenous as they depend on the consumers' preferred mortgage market choices. Second, there are two non-convexities in our model: the maximum operator in the consumer's second period budget constraint and the consumers' discrete choice of mortgage market. Our large economy allows us to deal with these non-convexities, however. Our solution approach is as follows. We construct a generalized game and show that there is a mixed strategies equilibrium. Then we claim that because the auctioneers' payoff functions depend on a profile of mixed strategies only through finitely many indicators, there is a degenerate equilibrium profile of the generalized game. And finally, we show that the equilibrium of the generalized game is in fact an equilibrium in the sense of Definition 1.

Remark 1: In our model default risk is the result of the NB's inability to perfectly screen consumers by type, and thus it can be attributed to the endogenous behavior of consumers with whom they are matched in equilibrium. We treat the risk of misclassifying consumers by type

(i.e., classifying a B-type consumer as a G-type) as idiosyncratic, in the sense that the sorting and then assignment of consumers into the NB loan market depends on the independent and uniform application of a credit scoring model, applied on a case-by-case basis, with classification outcomes governed by the law of large numbers.

Remark 2: Our notion of equilibrium assumes that lenders and SM investors form beliefs about the size and credit quality of the lenders' pools of originated loans. These beliefs are common, degenerate and governed by the lenders' respective CST^l . Equilibrium condition (1.2) guarantees that these beliefs are consistent with the distribution of consumers into the respective mortgage markets.²⁷

Remark 3: Unlike in traditional general equilibrium models where the market structure is exogenous, in our model the measure of borrowers in the TB and NB mortgage markets is endogenous and therefore so is the market structure. To see this, notice that each lender type is assigned an exogenous distribution rate parameter (d^{TB} for traditional banks and d^{NB} for NB lenders). However, consumers are choosing their preferred mortgage funding source, given the constraints of the economy, and therefore the sizes of the TB and NB sectors are endogenously determined. Thus, although the mortgage distribution rates d^{TB} and d^{NB} are exogenously given, their corresponding market sizes are endogenously determined.

4 Equilibrium Characterization

To characterize our equilibrium results, we focus on a more analytically tractable setting where owner-occupied housing and the numeraire good are perfect substitutes for consumers. In particular, consider the following linear separable utility function:

$$u^h(R_1, H_1, R_2, H_2) = R_1 + \eta H_1 + \theta^h(R_2 + H_2), \quad (8)$$

where θ^h (with $\theta^h < \theta^l$) denotes the consumer's time preference parameter and $\eta > 1$ denotes a preference parameter. $\eta > 1$ indicates that in the first period young households prefer to consume

²⁷Lenders and investors optimize using their beliefs, but without taking into account the consumers' choice of mortgage market. This is similar to general equilibrium models of firm formation where agents optimize without taking the supply of jobs into account (Ellickson et al. 2006).

owner-occupied housing over the perishable numeraire good. For the sake of exposition, we interpret the numeraire good as subsistence “rental housing” (one period contract and thus a perishable good).²⁸ In this case, η can be possibly justified by a better access to schools; see, for example, Corbae and Quintin (2015) for a model that incorporates an “ownership premium” in preferences. When households are “old” in period 2, the utility from consumption of owner-occupied housing H_2 and the utility from consumption of rental housing R_2 are the same, however.

4.1 TB Capacity Constraint

Traditional banks which hold non-prime mortgage loans on their balance sheets are subject to capital requirements and regulatory burden. This, together with limited access to capital, serves to constrain TB loan originations. In particular, we establish a one-to-one mapping from ω_1^l to $v(TB)$, where $v(TB)$ denotes the TB’s lending capacity constraint on the number of loan applicants. See Buchak et al.’s (2018) for evidence that TBs reduced their origination activity in markets where they faced more capital and regulatory constraints, and that the reduced origination activity was partly filled by shadow (non-bank) lenders.²⁹

We require that $\lambda_G \geq v(TB)$. Roughly speaking, we constrain the traditional banking sector to meet the demands of some but not all good type consumers should that demand exist. The assumption of capacity constrained TBs implies that, when traditional bank loans are the first choice among consumers, capacity-rationed G-type consumers turn to the NB loan market if they still prefer owning over renting.³⁰ In this case, a mass $\lambda_G - v(TB)$ of G-type consumers attempt to borrow from NBs, with the resulting endogenous proportion of G-types in that market as $\hat{\pi}_G^{NB} = (\lambda_G - v(TB))/(\lambda_G - v(TB) + \lambda_B)$.³¹ If, on the other hand, conditions are such that

²⁸We assume that rental and owner-occupied housing districts within a non-prime neighborhood are sufficiently segmented so that they are subject to separate market clearing conditions. When we say “non-prime neighborhood” within the model, we mean household consumption as it applies in the entire modeled non-prime economy.

²⁹See also Fuster et al. (2018) who find that technology-oriented non-bank lenders (“FinTech”) respond more elastically to changes in mortgage demand, suggesting the existence of capacity constraints in mortgage lending among traditional banks.

³⁰For simplicity, we impose the lending capacity constraint on the number of loan applicants $v(TB)$ and not on the TB’s income in the first period ω_1^{TB} . Observe that there is a one-to-one mapping between $v(TB)$ and ω_1^{TB} .

³¹If the regulatory burden on TBs increases – i.e., $v(TB)$ decreases – the credit quality of NB’s pool of borrowers improves in our model because there are more G-type consumers available to NBs when TB loans are preferred to NB loans.

NBs are universally preferred to TBs by G-type consumers, and home owning with NB mortgage loan financing is preferred to renting, then $\hat{\pi}_G^{NB} = \lambda_G / (\lambda_G + \lambda_B)$.

Interestingly, given the TB's capacity constraint and the NB's imperfect CST^{NB} , consumers of the same type may end up with different loan amounts, and thus realize different housing consumption. For example, there exists an equilibrium configuration in which G-types prefer first to borrow from TBs. In this case, because of the capacity constraint, some G-type consumers are rationed into the NB mortgage loan market. If those G-types still prefer owning over renting and thus apply for a loan from a NB, it will happen that some G-types will obtain a mortgage loan at a higher price, with lower loan proceeds that allows for the purchase of a relatively smaller house, than in the TB market. And finally, some of the remaining G-type consumers will be incorrectly screened by NBs – i.e., denied credit altogether – with no other option but to rent.

Lastly, we note that the capacity constraint $\lambda_G \geq v(TB)$ on TBs is not required for us to characterize a competing role for TBs and NBs in mortgage lending. What would be lost by eliminating the constraint is the second of three equilibrium regimes characterized in Section 5.2.

4.2 Adverse Selection, Housing Market Regulation, and House Prices

Before examining the role of beliefs on mortgage market configurations, we first discuss the equilibrium for our economy in terms of adverse selection in the NB loan market, house sizes, and house prices.

Adverse selection in the NB loan market: We use the concept of pooling equilibrium to address borrower adverse selection in the NB mortgage market. In equilibrium, conditional on consumers that attempt to borrow in the NB loan market and receive a good rating, the loan amount equals $q^{NB}\psi^{NB} = \pi^{NB}\bar{\theta} / (1 - \bar{\theta}(\pi^{NB}(1 - \delta) + \delta))$. The promise to repay is $\psi^{NB} = \omega_2^+ - \omega^{SR} + p_2 H_1^{G,NB}$, where $H_1^{G,NB}$ is the house size that a consumer with a NB loan can buy in the first period (see the Appendix for all details of the equilibrium closed form solution). This contract is designed for a good type consumer who is able to honor his promise. However, B-type consumers that are misclassified as a G-type end up defaulting because their income is ω^{SR} (instead of ω_2^+), and therefore are only able to pay back $p_2 H_1^{G,NB}$ (i.e., the house value). Default is such that foreclosure costs are incurred by the lender/investor, which in our model are captured

by parameter $\delta < 1$. The NB lender chooses a pooling discount price given its belief π^{NB} .

Minimum house size regulation: Motivated by our discussion on the presence of a minimum house size and its impact on housing affordability, we define the minimum owner-occupied house size as follows:

$$H_{NB}^{\min} \equiv \frac{\bar{H}(\omega^{SR} + \bar{L})}{2\omega^{SR} + \bar{L} + L^G} \quad (9)$$

$\bar{L} \equiv \theta^i \delta \omega^{SR} / (1 - \theta^i \delta)$ is the maximum loan amount that a NB lender would give to a B-type consumer being compatible with non-negative profits for the lender, and $L^G \equiv \bar{\theta}(\omega^{SR} + \bar{L}) / (1 - \bar{\theta})$ is the loan amount that a consumer classified as a G-type would obtain from a NB lender when mortgage markets are segmented (using the NB's first order condition and G-type consumer's first period budget constraint). In the Appendix we show that H_{NB}^{\min} rules out a separating equilibrium.³² Also, in the Appendix, we identify the threshold H_{TB}^{\min} that rules out a TB mortgage market specific for B-type consumers.

Threshold $H^{\min} \equiv \max\{H_{TB}^{\min}, H_{NB}^{\min}\}$ captures how a local minimum house size regulation affects the bottom of the housing market by excluding non-prime borrowers of B-type (who are identified as such) from the mortgage market.

The literature on the effects of land-use regulations on house prices is vast. It is now well known that zoning constraints such as minimum lot-size restriction and building-height limits tend to raise housing prices (see Bertaud and Brueckner 2005 for a theoretical analysis). Our model provides a different perspective by showing how, in the presence of credit scoring and alternative lending sources, minimum house size regulations prevent the least well-endowed non-prime consumers from obtaining a loan, which in turn prevents them from purchasing a house. In our equilibrium analysis below, we will see that the minimum house size H^{\min} is a function of the credit scoring technology. Thus, changes in this technology may trigger changes in the non-prime mortgage structure, in turn changing the house price level and number of consumers that have access to owner-occupied housing. The structural details underlying mortgage contract design and market organization consequently feedback to affect the rent versus own decision and

³²A separating equilibrium can exist in terms of lending to a B-type consumer that is misclassified as a G-type. This happens even though B-types always default in the second period. As such, the NB would get income $\delta p_2 \bar{H}_1$ by foreclosing the house, which is profitable for the NB when the loan amount is small, since by foreclosing the house the NB is able to seize the B-type borrower's house equity (ω^{SR}). Formally, the NB gets $\delta p_2 \bar{H}_1$ in $t = 2$ and gives $q^B \varphi^B$ to the borrower in $t = 1$. We have worked out the separating equilibrium and found qualitatively similar thresholds as those found in Section 4.4 below.

the house price level in our model.

House prices: We model the aggregate supply of owner-occupied housing in the *first* period and the aggregate demand for owner-occupied housing in the *second* period as inelastic, both equal to $\bar{H} = 1$. A constant stock of owner-occupied housing is convenient to generate simple closed form equilibrium solutions, with market clearing house prices such that $p_1 = p_2 = p$.³³ Constant intertemporal house prices then allow us to isolate the credit scoring channel's direct influence in the housing sector, whereby mortgage defaults occur in our model due to the imperfect screening of borrowers by type in the NB market.³⁴ Also notice that when $p > 1$ (as it is the case in our numerical examples of equilibrium below), old households with a mortgage will sell their house in the second period and move to rental housing, as the benefits to owning go away as the younger household transitions to older age (Hochguertel and van Soest 2001).³⁵ In the first period, young consumers exhibit a preference for home ownership provided that the credit scoring technology parameter π^{NB} exceeds certain thresholds, as analyzed below.

The setting just described is also convenient because we can conceive our model as characterizing an overlapping generations (OLG) economy, where households in the second period choose to sell their houses to a new generation of younger households, directly implying the stock of owner-occupied housing changes hands from old households in a previous cohort to young households in a new cohort.³⁶

4.3 Mortgage Discount Prices and Excess Premium

Equilibrium mortgage pricing is characterized by an optimality condition that depends on loan source parameters CST^l and d^l . For the NB lender, the determination of mortgage price involves

³³The owner-occupied market clearing equations in periods 1 and 2 and the households' optimal choice $H_2^h = 0$ (shown in the Appendix) imply that $p_1 = p_2 = p$.

³⁴For a model where default is triggered by a fall in house prices, see e.g. Arslan et al. (2015) in which mortgages are non-recourse. Also see Brueckner et al. (2012).

³⁵Hochguertel and van Soest (2001) provide evidence that young households buy a house to accommodate the new family members and possibly to get access to better schools, but when they are old and the family size decreases, these households often sell their houses and move to smaller rental houses.

³⁶To incorporate lenders and investors into this extended setting, we must assume that they live for two periods (as consumers do), or that they cannot share risk across time among different generations of households. Also, notice that extending the OLG model to a more general setting with infinitely lived agents and more than one good is subtle because the presence of such agents may preclude equilibrium existence due to the possibility of Ponzi schemes (see Seghir 2006).

a trade-off between borrower adverse selection and investor liquidity in the secondary market.³⁷ We leave for the Appendix the details of the equilibrium computations. Here we report our findings on the relationship between the TB and the NB mortgage loan rate.

Based on pooling, the NB finds it optimal to add a credit spread to the base loan rate to account for borrower adverse selection risk. This reduces the loan proceeds. The greater the proportion of B-types applying for a mortgage loan in the NB market and the worse the assessed classification accuracy of the CST^{NB} , the lower are the funding proceeds of a NB loan. Offsetting this upward rate pressure is the lower cost of capital found in the secondary mortgage market. Depending on which effect dominates, the implied mortgage loan price in the NB market can be higher or lower than the implied mortgage loan price in the TB market.³⁸

Define the excess premium (EP) as the difference between the implied mortgage loan rate in the NB market and the implied mortgage loan rate in the TB market, as follows:

$$EP \equiv (1/q^{NB}) - (1/q^{TB}) \quad (10)$$

Using the TB and NB pricing expressions derived in the Appendix, we conclude the following:

Proposition 1: *The EP increases in the default rate $(1 - \pi^{NB})$, loss given default δ , and the lender's patience (inverted cost of capital) parameter (θ^l) , and decreases in the NB's credit scoring technology (CST^{NB}) , the differences in SM mortgage distribution rates $(d^{NB} - d^{TB})$, and secondary market liquidity $(\theta^i - \theta^l)$.*

4.4 Beliefs and Mortgage Market Configurations

Next, we explain the mortgage market configuration that follows from our general equilibrium framework when: 1) TBs' access to hard and soft information allows them to know the consumer's type with certainty, while NBs only rely on an imperfect credit assessment based on hard information ($CST^{NB} < CST^{TB} = 1$) and 2) NBs distribute a greater proportion of loans into

³⁷See Gande and Saunders (2012) for empirical evidence of the positive impact of selling a mortgage loan to secondary market investors on the borrower's financial constraints. Another relevant paper is Stroebel (2016), which finds that integrated lenders have better information regarding the cost quality of homes than non-integrated lenders. See also Frankel and Jin (2015) for a result where securitization can lead to remote lending in a game theoretical setting with information differences between bank types.

³⁸See the pricing condition (14) in the Appendix for a mathematical expression of this trade-off in the NB loan market.

the secondary market than TBs ($0 \leq d^{TB} < d^{NB}$). We are particularly interested in examining CST's influence on the flow of capital into non-prime lending markets vis-à-vis the NB lending sector. This will then allow us to analyze the credit scoring channel in terms of its effects on the cost and availability of mortgage funding, and ultimately house prices. In the process, we provide a new characterization of housing market price boom and bust that is the result of equilibrium regime changes.

We now proceed to identify three thresholds, π_0 , π_1 and π_2 , corresponding to the NB lender's belief π^{NB} that delineate alternative lending regimes. Recall that π^{NB} measures the posterior probability that a consumer is a G-type given that the NB classifies the consumer as a G-type. This quantity depends on two factors: the CST and the endogenous proportion of G- and B-types applying for loans from NBs. The three denoted thresholds determine different non-prime mortgage market configurations, or equilibrium regimes, and all can be expressed as a function of the parameters of our economy. In particular, the thresholds, π_0 , π_1 and π_2 , are characterized as follows:

1. *In presence of a minimum house size constraint, the NB lending market shuts down (is inactive) if belief π^{NB} is sufficiently small.* That is, there is a threshold π_0 that solves the following equation:

$$H_1^{G,NB}(\pi_0) = H^{\min} \quad (11)$$

such that when $\pi^{NB} < \pi_0$, NB loans are so small that borrowers cannot afford to buy a house with size above H^{\min} . Here $H_1^{G,NB}(\pi_0)$ denotes G-type owner-occupied housing consumption when borrowing from a NB and belief π^{NB} equals π_0 . The expression for $H_1^{G,NB}$ as a function of π^{NB} , as well as the equilibrium loan amounts, can be found in the Appendix.

2. *There is an inactive NB mortgage lending market as long as G-type consumers prefer to rent in the first period rather than borrow from NBs.* When π^{NB} falls below a given threshold π_1 , the implicit NB mortgage rate is so high that G-type consumers prefer to rent in both periods ($R_1 = \omega^{SR}$ and $R_2 = \omega_2^+$) rather than borrow from a NB to buy a house in the first period. Threshold π_1 , at which indifference between buying a house with a NB

loan and renting in both periods occurs, solves the following equation:³⁹

$$\eta H_1^{G,NB}(\pi_1) + \theta^h \omega^{SR} = \omega^{SR} + \theta^h \omega^+ \quad (12)$$

Thus, when $\pi^{NB} < \pi_1$, NB loans are so small that G-type consumers prefer to rent in both periods. When this occurs B-types cannot borrow in the NB market.

Lemma 1: *The NB mortgage market is inactive when $\pi^{NB} < \max\{\pi_0, \pi_1\}$.*

3. *G-type consumers prefer to borrow from NBs if funding proceeds from the NB loan exceeds that of the TB loan.* There is a threshold π_2 at which a G-type consumer is indifferent between a NB loan and a TB loan. This threshold solves the following expression:

$$\eta H_1^{G,NB}(\pi_2) + \theta^h \omega^{SR} = \eta H_1^{G,TB} + \theta^h \omega^{SR} \quad (13)$$

The left hand side term in equation (13) identifies the G-type consumer's utility from buying a house in the first period with a NB loan and then renting (in a setting where only the NB loan market is active). The right hand side term in equation (13) shows the G-type consumer's utility from buying a house in the first period with a TB loan and then renting (in a setting where both TB loans and NB loans markets are active). Notice that $H_1^{G,TB}$ is not a function of π^{NB} , since, by assumption, TBs have a perfect screening technology and only lend to G-types. Further observe that when $\pi^{NB} > \pi_2$, G-type consumers prefer NB loans over TB loans even though NBs risk-price classification error (lending to some B-types). In this case, the proportion $\hat{\pi}_G^{NB}$ of G-type consumers that attempt to borrow from NBs experiences a discrete increase, as now NB loans are the best option for G-type consumers. In addition, when the SM mortgage distribution rate d^l increases, threshold π_2 decreases and the NB loan market expands. This happens because an increased rate of distribution into the secondary market reduces the implied mortgage rate due to the SM investors' greater patience level ($\theta^i > \theta^l$).

³⁹In the left hand side term of equation (12) both TB loan and NB loan markets are active and the market clearing house price is computed accordingly. See the price function stated in the Appendix.

Lemma 2: *The NB mortgage market becomes the first choice for G-type consumers when $\pi^{NB} > \pi_2$.*

Below we summarize the three different possible market configurations as they depend on the NB's belief π^{NB} , as well as summarize participation rates in the alternative home ownership-rental markets under each of these configurations.

Proposition 2 (Market configurations):

- **Dominant TB Loan Market:** *If $\pi^{NB} < \max\{\pi_0, \pi_1\}$, the NB mortgage market is inactive, and only a mass $v(TB)$ of G-type consumers can borrow to buy a house. The rest of consumers, with mass $\lambda_G - v(TB) + \lambda_B$, rent in both periods.*
- **Dominant NB Loan Market:** *If $\pi^{NB} > \pi_2$, G-type consumers prefer the NB mortgage market. A mass $CST^{NB}\lambda_G + (1 - CST^{NB})\lambda_B$ of consumers receive a good rating and are able to borrow at the NB loan rate to buy a house. A mass $\min[(1 - CST^{NB})\lambda_G, v(TB)]$ of G-type consumers who are incorrectly classified as B-types will borrow from their second best option, the TB loan market. The rest of consumers will rent in both periods.*
- **Coexisting TB-NB Loan Market:** *When $\pi^{NB} \in [\max\{\pi_0, \pi_1\}, \pi_2]$, TBs lend to a mass $v(TB)$ of G-type consumers. The remaining pool of G- and B-types apply for a loan in the NB market. A mass $(1 - CST^{NB})(\lambda_G - v(NB)) + CST^{NB}\lambda_B$ of consumers are rejected by NBs and have no option but to rent in both periods.*

The proof follows immediately from our previous analysis and is thus omitted. Next, we explain how thresholds π_0 , π_1 and π_2 change as a function of key model parameters. First, should the perceived precision of NB's credit scoring technology (CST^{NB}) deteriorate, π^{NB} decreases. There is, as a result, greater asymmetric information between consumers and NBs, and all else equal the NB market is closer to or actually enters into the inactive region. Second, when the SM investor's time preference parameter θ^i and/or the NB's distribution rate d^{NB} increase, all else equal, the active region for a NB loan market expands (as threshold values π_0 , π_1 and π_2 decrease). This is because NB loans become less expensive, with greater loan proceeds, as SM investors are willing to pay more for NB mortgages. We finish this section with a remark on the possible equilibrium mortgage structure configurations in absence of the TB's capacity constraint.

Remark 4: *The capacity constraint $\lambda_G \geq v(TB)$ on TBs is not required for us to characterize a competing role for TBs and NBs in mortgage lending. What would be lost by eliminating the constraint is the third of three equilibrium regimes characterized in Proposition 2. Also, the second equilibrium regime in this proposition would slightly change by allowing all G-type consumers that are incorrectly classified as B-types by NB lenders (with mass $(1 - CST^{NB})\lambda_G$ instead than $\min[(1 - CST^{NB})\lambda_G, v(TB)]$) to borrow from their second best option, the TB loan market.*

5 The Credit Scoring Channel

5.1 Overview

In this section we highlight the credit scoring channel’s effect on the flow of capital into the non-prime mortgage market. As we discussed in Section 2, improvements in credit scoring model technology or increased confidence in the classification accuracy of the technology result in upward revisions in CST^{NB} . We highlight these effects, arguing they played a prominent role in explaining the rise of non-prime mortgage lending during the critical 1995 to 2006 time period. Then, with a surge of unexpected defaults and foreclosures occurring in 2006 and 2007, market participants started questioning and then revising their beliefs as they realized that their credit scoring models were failing to work as expected.⁴⁰ This failure of the credit scoring models to accurately predict failure consequently led to wholesale downward revisions in both prior and posterior credit assessment probabilities (CST^{NB} and π^{NB} in our model), resulting in the collapse of non-prime mortgage lending and house prices.

5.2 A Credit Scoring Technology Shock Triggers the “Boom”

With this background, we now examine how changes in beliefs embedded in our model’s credit scoring technology, captured by parameter CST^{NB} , can trigger changes in the equilibrium structure of the mortgage market. In particular, we will show how, starting from an inactive NB lending region, incremental improvements in CST^{NB} can trigger the emergence of the NB mortgage

⁴⁰See Figure 1 of Brunnermeier (2009) and the associated discussion.

market. Then with further improvements in CST^{NB} , the NB lending sector comes to dominate traditional bank lending, as all G-type consumers migrate from the TB to NB market as their first choice for mortgage funding. These migration tipping points trigger sharp changes in the source as well as quantity of non-prime mortgage funding, and consequently in house prices, providing a new characterization for the underlying causes of asset market boom and bust.

We illustrate these changes in equilibrium market structures by considering the following parameter values: $d^{TB} = 0$, $d^{NB} = 0.7$, $\theta^h = 0.4$, $\theta^l = 0.7$, $\theta^i = 0.9$, $\eta = 4$, $\delta = 0.5$, $\omega^{SR} = 0.5$, $\omega_2^+ = 1$, $\lambda_G = 1.5$, $\lambda_B = 1$ and $v(TB) = 1$. Figure 1 illustrates the equilibrium aggregate mortgage quantities in the TB versus NB markets, respectively, for different equilibrium regimes identified above. Figure 2 shows resulting equilibrium owner-occupied house prices as a function of CST^{NB} . The thresholds for CST^{NB} follow from expression (1), the π -thresholds identified above,⁴¹ and the corresponding proportion of G-type consumers in the NB loan market. With this parameter set, it follows that $CST^{NB} < 0.31$ establishes Regime 1, $0.31 \leq CST^{NB} < 0.85$ establishes Regime 2, and $CST^{NB} \geq 0.85$ establishes Regime 3.⁴²

In Regime 1 the NB market is dormant, caused by borrower type classification errors of sufficient size so that high pooled mortgage loan rates result. The high implied NB mortgage rates are such that G-type households which are rationed out of the TB market prefer to rent rather than own. With no G-types applying for mortgage loans in the NB sector, B-types are unable to represent themselves as G-types, so the NB loan market is dormant. And with only a limited number of G-type households gaining access to mortgage financing and an inactive NB loan market, equilibrium house prices are relatively low (Figure 2).

In Regime 2 the NB mortgage market emerges because beliefs regarding the precision of the CST have improved sufficiently. As a result, NBs offer mortgages at prices that are attractive to G-type households that have been rationed out of the preferred TB market. Mortgage loan amounts from TBs remain constant as a function of CST^{NB} in Regime 2, as the credit quality of the mortgage pool remains constant (only G-types obtain loans in the TB market). In contrast, while pooled loan rates are lower in Regime 2 than in Regime 1, implied risk-adjusted mortgage rates are nevertheless high in the NB market relative to the TB market due to a relatively high

⁴¹The equilibrium π -thresholds for the parameter values are $\pi_0 = 0.16$, $\pi_1 = 0.18$ and $\pi_2 = 0.74$.

⁴² $\hat{\pi}_G^{NB}$ increases from $(\lambda_G - v(NB))/(\lambda_G - v(NB) + \lambda_B) = 0.33$ to $\lambda_G/(\lambda_G + \lambda_B) = 0.6$ when $CST^{NB} \geq 0.85$.

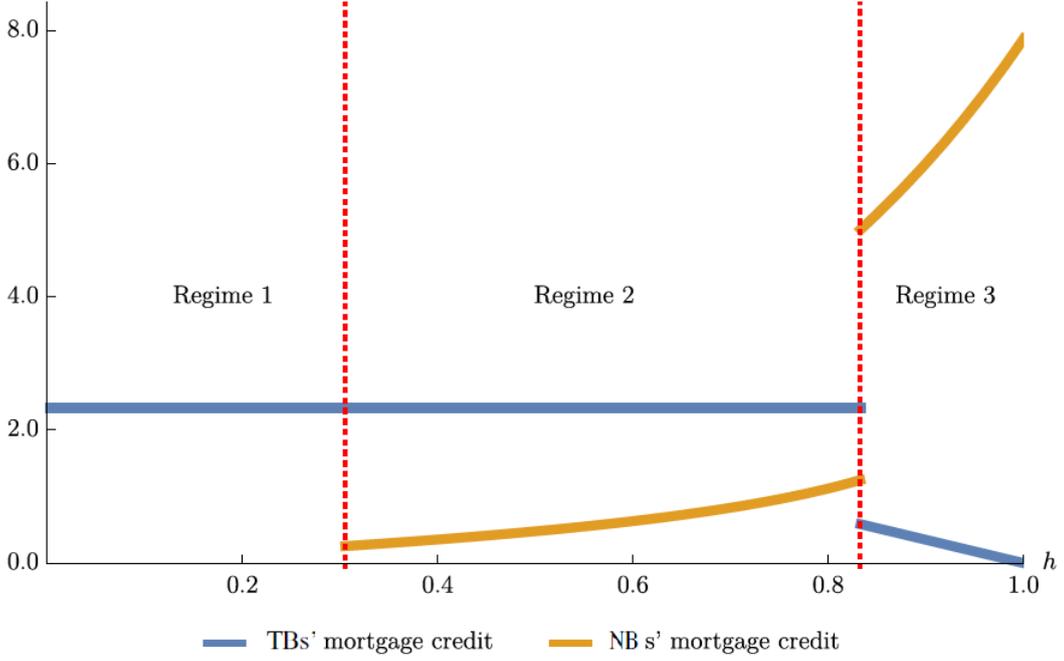


Figure 1: This figure plots the total amount of TB and NB mortgage credit as a function of $CST^{NB} \equiv h$. The thresholds for CST^{NB} follow from expression (1), the π -thresholds identified above and the corresponding proportion of G-type consumers in the NB loan market ($\hat{\pi}_G^{NB} = 0.33$ if $CST^{NB} \leq CST_2^{NB} \equiv 0.85$ and $\hat{\pi}_G^{NB} = 0.6$ otherwise). In particular, thresholds for CST are $CST_0^{NB} = 0.28$, $CST_1^{NB} = 0.31$, and $CST_2^{NB} = 0.85$.

rate of borrower type classification errors. Aggregate mortgage quantity increases (implied credit spread decreases) in Regime 2 due to a perceived decrease in borrower type classification errors. House prices experience a discrete increase at the lower bound of Regime 2, signifying a boom, as the number of households entering the owner-occupied market jumps due to the emergence of the NB market. House prices increase as a function of CST^{NB} in Regime 2, because aggregate mortgage funding proceeds increase to drive up the demand for a fixed stock of owner-occupied homes.

In Regime 3 a wholesale migration of G-type consumers into the NB market occurs. Because of this, there is a large discrete jump in the proportion of G-types applying for a NB loan, an effect that by itself causes a discrete downward adjustment in the pooled mortgage loan rate. Aggregate mortgage loan quantity in the NB market also experiences a large jump, while aggregated TB mortgage quantities experience a large decline. The decline occurs because fewer G-type households apply for a TB mortgage loan, only doing so when they are misclassified as a B-type in the NB mortgage market. House prices continue to increase as a function of CST^{NB} as aggregate

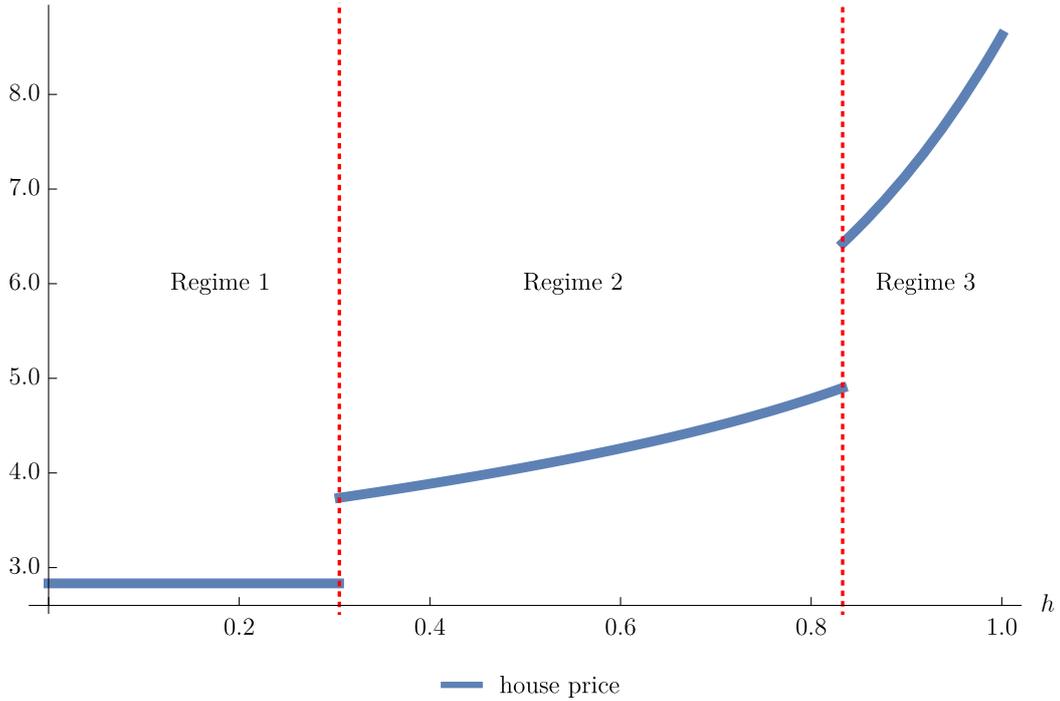


Figure 2: This figure illustrates the equilibrium house price p as a function of $CST_G^{NB} \equiv h$. Thresholds for CST are $CST_0^{NB} = 0.28$, $CST_1^{NB} = 0.31$ and $CST_2^{NB} = 0.85$.

mortgage quantities increase due to lower subjective probabilities of misclassification outcomes.

This described financing-to-house price channel is consistent with Mian and Sufi (2009, 2014) and others in its emphasis on funding liquidity through the private-label Mortgage-Backed Securities market, where our contribution to the literature lies in highlighting the credit scoring channels role in facilitating those capital flows. At the regime boundaries changes in home ownership, mortgage amounts and house prices are discrete and large in magnitude. This is especially true when transitioning from Regime 2 to 3, in which all households prefer to borrow in the NB market. Our model thus provides a new explanation as well as alternative characterization for house price booms, implying the existence of large price increases occurring over short time periods due to perceived improvements in credit scoring technology (the *credit scoring channel*).

Similar logic applied in reverse can be used to characterize the housing bust. As early term mortgage loan defaults spiked unexpectedly starting in the second half of 2006, confidence in CST was initially shaken and then shattered by its failure to predict failure.⁴³ The credit scoring

⁴³In Rajan et al. (2015) the model's failure to predict failure is not a surprise to the issuer given the use of only hard information to assess applicant credit quality. In contrast, in our baseline model the issuer relies on hard information

channel in our model, where negative shocks to CST^{NB} imply moving from region 3 to region 2, and then to region 1, illustrates the resulting substantial declines in NB mortgage lending volume, house prices and home ownership rates.

6 Further Remarks

6.1 Secondary Market Funding Liquidity and the Fundamental Proportion of G-borrowers

We start this section by dissecting two complementary channels that likely contributed to the boom and bust in mortgage and housing markets: 1) the SM investor's discount factor (a proxy for secondary market funding liquidity, e.g., demand for US mortgage-backed securities by foreign investors) as measured by θ^i , and 2) the fundamental proportion of G-type (higher income) consumers as measured by the ratio $\lambda_G/(\lambda_G + \lambda_B)$. Highlighted results are summarized as follows.

Proposition 3:

1. *When the secondary market investor's patience parameter θ^i increases, marking a decrease in capital costs and an increase in SM liquidity, the threshold CST_2^{NB} indicates an increased Regime 3 (dominant NB sector) size.*
2. *For a given CST^{NB} , a negative shock $\varepsilon > 0$ of sufficient size to the fundamental proportion of G-type consumers (i.e., $\lambda'_G = \lambda_G - \varepsilon$ and $\lambda'_B = \lambda_B + \varepsilon$) causes a transition from Regime 3 to Regime 2 or 1.*

For the sake of brevity, we leave the proof of Proposition 3 for the Appendix.

Proposition 3.1 states that when the investor's patience parameter θ^i increases, the NB's implied mortgage rate decreases to increase the mortgage loan amount and house price. This happens because a larger fraction of the mortgage loans originated by NBs ($d^{NB} > d^{TB}$) are now

only to form what is, in hindsight, overly optimistic beliefs regarding loan applicant credit quality (see Brunnermeier 2009). Period beliefs are consequently revised downward based on the arrival of new information.

subject to a lower cost of capital.⁴⁴ Such an effect goes in the opposite direction as well, where a decline in the investor’s patience parameter θ^i can lead to a shift from Regime 3 to Regime 2 and possibly then to Regime 1.

Concurrent with the perceived improvements in credit scoring models was increased secondary market securities purchase activity by GSEs and other large institutional investors, as well as the introduction of capital reserve regulation (Basel II) that increased the attractiveness of owning the higher-credit rated securities. There were also financial shocks (the Asian and Russian financial crises) that shifted foreign capital flows towards dollar-denominated U.S. Treasuries and close substitutes. This shift in demand decreased yields of riskless bonds, causing fixed-income investors to move further out the credit risk curve in search for higher yields. The search for higher yields and favorable capital treatment combined to cause demand for AAA-rated securities, and therefore NB originated mortgage product, to increase significantly (again see Brunneimeier 2009 for additional detail).

Proposition 3.2 considers a negative shock to the fundamental proportion of G-types, $\lambda_G/(\lambda_G + \lambda_B)$. The intuition for the result is that, although a shock to the fundamental proportion of G-types does not affect the regime thresholds identified above, it does change the likelihood of classification error as measured by $\hat{\pi}_G^{NB}$ (see equation (1)). As the proportion of G-types in the NB pool of approved mortgage loans decline, the assessed credit quality of the pool declines to increase the risk-adjusted credit spread required on the mortgage loan. This negative (aggregate) shock can be interpreted most directly as a downward shift in the future employment prospects of non-prime households, and more broadly as a deterioration in household’s net worth, as documented by Mian and Sufi (2014, 2017).

6.2 The Boom-Bust of the Great Financial Crisis

In this section we make further remarks explaining how our paper contributes to the theoretical literature that puts emphasis on information problems to rationalize the great financial crisis. First, our boom-bust equilibrium outcomes are generated in a setting with adverse selection in the primary mortgage (origination) market, and thus depart from Gorton and Ordonez (2014) and Fishman and Parker (2015), who consider adverse selection in the secondary mortgage (resale)

⁴⁴See the pricing equation (14) and the equilibrium loan amount expression in the Appendix A.3.

market as the a prominent reason of the expansion and collapse of lending. Also, our approach differs from other theoretical models of sophisticated banking and subprime lending, such as Makarov and Plantin (2013), because we distinguish between NB and TB funding models, relating their changes in market share to different equilibrium non-prime mortgage configuration regimes that result from changes in the credit scoring technology and securitization.

Second, our model considers a general equilibrium framework of perfect competition and thus departs from Ruckes (2004) and Dell’Ariccia and Marquez (2006) sequential strategic decision making models that emphasize the different degrees of bank competition to rationalize variations in credit standards over the business cycle. Our paper also departs from Dell’Ariccia and Marquez (2006) in that, instead of allowing banks to choose between collateralized and uncollateralized loans, we consider both traditional and NB loan markets subject to the same collateralized, limited recourse mortgage contract. In our setting, variations in the credit scoring technology may result in the consumers’ preferred funding source changing from one to another, in turn changing the NB lenders’ loan portfolio quality and the associated mortgage rates, credit amounts, and house prices.

Third, because our equilibrium mechanism links non-prime mortgage lending standards and changes in the market structure to the run-up and eventually collapse in home-prices, our paper also fills a gap in the literature that studies mortgage leverage and the foreclosure crisis - see, e.g., Corbae and Quintin (2015) work on foreclosure dynamics with exogenous house prices.

House price expectations are the main ingredient in Brueckner et al. (2012), who show that subprime lending is both a consequence and a cause of house-price escalation. Our model also captures these effects, but focuses instead on income shocks. In addition, we consider a new channel - credit scoring - that relates subprime lending and house prices. Because we embed this effect in a general equilibrium model, we also capture the possibility that house prices determine the amount of non-prime credit that TBs and NBs originate. Importantly, in our model it is possible that a smooth small increase in the assessment of the classification accuracy of a credit scoring model can lead to an abrupt change in house prices and mortgage credit quantities.

Lastly, our focus on the credit scoring technology is also related to Drozd and Serrano-Padial (2017), whose focus and findings mirror in part our arguments on the loan origination side. In both models, increases in the precision of credit assessment technology lead to an increased NB

market share and hence riskier lending (in terms of mortgage rates) as measured by reliance on hard credit information. Our model differs, however, in that increased precision in credit assessment, to the extent that it is real, leads to less risky NB lending, since classification error is reduced as precision increases.

6.3 Lax Screening

Since the seminal works of Keys et al. (2010), Purnanandam (2011) and Rajan et al. (2015), there has been significant focus in the mortgage lending literature on what has become known as “lax screening”. The meaning of this term is not simply a relaxation in observable lending standards, but rather a devolution from incorporating available hard and soft information into the *yes-no* underwriting decision to relying on a reduced set of hard information only. The underlying cause of lax screening during the early and middle 2000s is generally attributed to an increased volume of secondary market loan sales to investors that were only able to verify hard information in their assessment of loan pool credit quality, with loan originators having no incentive to generate soft information as a result. Because information is lost or ignored in the loan underwriting process, there is a loss of efficiency due to an increased likelihood of bad lending outcomes resulting from credit quality misclassification error.

Using our model we can make two relevant points in the context of declining lending standards and lax screening. First, preliminarily, our assumption of no soft information acquisition by NBs is consistent with findings in this literature. And more importantly, as CST^{NB} increases in our model (based on inputting hard information only), in moving from a dormant NB loan market (in Regime 1) to increasing activity (in Regime 2) and finally to domination (in Regime 3), lax screening as measured by the volume of NB-originated non-prime mortgage loans outstanding relative total outstanding non-prime mortgages, becomes increasingly prevalent (see Figure 2). This increasing prevalence is explained in our model by increased confidence, and hence perceived classification precision, in the applied credit scoring technology.

Thus, somewhat unintuitively, according to our model, increases in classification accuracy result in increasingly lax screening outcomes as defined in the literature. But, to the extent that increases in CST^{NB} do in fact reflect increasing precision in classification accuracy, lax screen-

ing does not necessarily imply an incremental loss in economic efficiency. Alternatively, to the extent that overconfidence in CST^{NB} or the misrepresentation of loan credit information are responsible for the increasingly lax screening outcomes, an important inefficiency problem remains or emerges.

Second, the phrase “lax screening” has traditionally been associated with a relaxation of observable lending standards, which in turn are often inferred from increases in loan acceptance rates – see, e.g., Dell’Ariccia et al. (2012) and Mian and Sufi (2014, pp. 76-79).⁴⁵ In our model, changes in loan acceptance rates can be caused simply by changes in the CST^{NB} applied in the NB lending sector, with no change in observable loan underwriting standards. For example, in Regime 2, based on the chosen parameter values, acceptance rates decline with increases in CST^{NB} due to the relatively low proportion of G-types in the NB loan application pool (the proportion of G-types applying for a NB loan is less than 0.50). At the transition from Regime 2 to Regime 3, all G-types migrate to the NB loan market as their first funding choice, dramatically increasing the proportion of G-types in the NB loan application pool (the proportion of G-types applying for a NB loan now exceeds 0.50) and, as a result, the loan application acceptance rate increases in CST^{NB} without any changes to observable lending standards.

Thus, we show that an analysis of loan acceptance rates to infer relaxation (or tightening) in lending standards can be misleading. Instead, in our model, increases or decreases in loan acceptance rates depend importantly on changes in the screening technology, with the possibility that an increase in the application acceptance rate occurs simply because of (perceived) improvements in the classification precision of the applied CST^{NB} .

6.4 Income Misrepresentation

The seminal works of Bolton and Dewatripont (1994), Garicano (2000), Stein (2002), and Vayanos (2003) have emphasized the economics of information transmission. Stein (2002) makes a clear distinction between hard and soft information in relation to communications costs in hierarchical

⁴⁵The OCC formally assesses trends in underwriting practices by surveying national banks. According to the OCC’s survey of credit underwriting practices, “the term ‘underwriting standards’, as used in this report, refers to items such as loan maturities, facility pricing, and covenants that banks establish when originating and structuring loans ... A conclusion that the underwriting standards for a particular loan category have eased or tightened does not indicate that all the standards for that particular category have been adjusted... It indicates that the adjustments that did occur had the *net effect* of easing or tightening underwriting criteria”.

organizations. Traditional banks are hierarchical, in the sense that they are relational and specialize in the generation of soft information that improves the precision of credit assessment. But this increased precision comes at a cost, having two related components. First, the acquisition of soft information inhibits the marketability of the loans because the information is not easily nor credibly transmittable to outside agents. The immobility problem is exacerbated by the very existence of the asymmetric non-transmittable soft information, which raises concerns over adverse selection in the secondary market. Second, due to the relatively slow production of soft information through relational means, realizing scale economies in mortgage loan production can be difficult. Indeed, diseconomies resulting from relational soft information production can be a central cause of capacity constraints that limit TBs from serving its entire market of credit worthy consumers.

In contrast, the development of hard information-based credit scoring technologies that automate loan underwriting decisions and pricing is very much about relaxing capacity constraints and realizing scale economies in loan production (see Fuster et al. 2018 for recent evidence). Credit decisions are admittedly less exact, but the process works as long as none of the participants along the securitization supply chain tamper with the transmitted hard information.⁴⁶ Indeed, the misrepresentation of borrower credit information as a means of tampering with the hard information has been held up as *the* central flaw of the securitization process, exemplifying how the information transmission process breaks down when certain market participants take advantage of the process.

In this vein, we point out that our model can fully capture a particular form of loan misrepresentation: borrower misrepresentation of their income prospects, in which B-types represent themselves as G-types when applying to NBs for mortgage loans. When credit evaluation is imperfect, as it is in our model of NB lending and secondary market loan sales, a certain amount of misrepresentation is inevitable in the realization of scale economies in lending markets. Misrepresentation risk is rationally anticipated and priced into the mortgage loan in our model, with income misrepresentation increasing both the total supply of mortgage loans and house prices. Market participants recognize the relevance of these effects, understanding that inflated prices and quantities naturally emerge in connection with the NB originate-to-distribute lending model.

⁴⁶There are of course other forms of agency-based moral hazard problems that have been identified, including inflated credit ratings and over-production incentives.

To the extent other forms of misrepresentation of hard credit quality information exist, critical questions are: i) where does the misrepresentation originate from – borrower or NB lender (or possibly securities underwriter), and ii) do market participants recognize the risk and price it into the mortgage loan, or are they somehow ignorant that the misrepresentation is occurring. The latter issue suggests that identifying anticipated versus unanticipated misrepresentation in mortgage lending is an important distinction that has only begun to be emphasized in the literature (see Ambrose et al. 2016 and Griffin and Maturana 2016). Our model provides a structure to begin to get at these issues, although doing so is beyond the scope of this paper.

7 The Credit Scoring Transmission Channel: When Non-Bank Lenders Acquire Soft Information

So far we have assumed that traditional banks have access to both hard and soft information, allowing them to perfectly screen borrower types, while non-bank lenders only rely on hard information to underwrite mortgages. In this section, we allow non-bank lenders to acquire soft information at a cost. In this context, the exclusive reliance of secondary market investors on hard information introduces the opportunity for originating lenders to surreptitiously acquire soft information to select against investors in the secondary mortgage market. The acquisition of soft information improves the precision of credit quality assessment for the non-bank, where, based on the new information some loans are reclassified from accepts to rejects (reclassified lemons). This reclassification provides the lender the opportunity to adversely select against the secondary market, since secondary market investors believe those reclassified lemons to be acceptable based on hard information only. By pricing all originated loans as if only hard information is utilized, and delivering the quantity of loans expected in a hard information only origination regime, the informed lender can sell the lemons and retain a loan portfolio whose credit quality exceeds that based on hard information only loan pricing. The retention of some proportion of originated loans is necessary for the non-bank lender to have incentives to acquire soft information, since otherwise there would be no source of revenues to offset the costs of information acquisition.⁴⁷

⁴⁷We also note that in practice, some traditional banks, particularly the larger more complex banks such as Wells Fargo, were known to originate mortgage loans for ownership as well as set up their own conduit lending operations

For the sake of brevity we leave the details of this extension for the Appendix, and here highlight the main insights from the analysis. First, rather unintuitively, we show that adverse selection in the secondary market can actually improve the credit quality of the sold loan portfolio in comparison to that based on hard information only. This finding depends on the resulting mix of good loans reconfirmed as good, with greater precision (reconfirmed cherries), and good loans reclassified as bad but sold as good (lemons). Second, adverse selection of this type exacerbates boom and bust in housing markets, due to the increased aggregate quantity of mortgage loans originated. Finally, the endogenously determined quantity of soft information acquired decreases in the mortgage distribution rate, which extends lax screening findings of Keys et al. (2010), Purnanandam (2011) and Rajan et al. (2015) to the case of secondary market adverse selection. This result is intuitive, since a higher distribution rate results in relatively fewer retained loans by the non-bank lender, which in turn decreases incentives to acquire soft information at a positive marginal cost.

8 Concluding Remarks

This paper provides a general equilibrium model of a non-prime economy with endogenous market segmentation, tenure choice, mortgage quantities and prices, and house prices. Our distinction between the two different sources of funding for consumers (traditional vs. non-bank lenders) helps to highlight a fundamental trade-off between access to soft information versus liquidity provision through the secondary loan sale market. The model further illustrates how, depending on market conditions, consumers can migrate from one mortgage market to another, with implications for the sources and sizes of mortgage flows together with their effects on house prices and access to owner-occupied housing.

Another important component of our theory is the limited recourse nature of the non-prime mortgages, which functions as an incentive compatibility constraint for better credit quality borrowers. The limited recourse clause allows for a focus on income rather than collateral in the in which loans were originated for distribution. It was the case that secondary market participants generally required a “firewall” to be established between the “originate-to-own” and the “originate-to-distribute” parts of the business in order to prevent these banks from underwriting based on a fuller set of information than was available to secondary market investors. Our model fully accommodates this kind of setting.

assessment of credit risk,⁴⁸ and is a feature that helps generate greater housing consumption and a pooling equilibrium in the shadow banking mortgage loan sector.

Given this setting, we focus much of our analysis on what we term the credit scoring and transmission channels that operate in the shadow banking/secondary mortgage market for non-prime mortgage loans. A credit scoring technology, which in essence is a Bayesian prior founded on beliefs regarding the classification accuracy of a *yes-no* loan underwriting model, is used in combination with estimated proportions of good versus poor lending credit risks in a defined population, as specifically measured by household income, to generate a posterior that measures the credit quality of the non-prime mortgage pool. This measure of credit quality serves as a basis for mortgage pricing, which in turn determines equilibrium configurations in which the non-bank lending sector is either inactive, in direct competition with traditional bank lenders, or is dominant.

Credit scoring technology, operating through the non-bank lending sector, thus serves as a mechanism that controls the flow of mortgage capital into non-prime housing markets. The boundaries of the alternative equilibrium regimes define tipping points, where consumers abruptly change their loan application migration patterns. These tipping points set off booms or busts in mortgage and housing markets, depending on the direction of the migration pattern changes. The equilibrium configurations we highlight together with tipping point regime boundaries are consistent with stylized empirical facts associated with the boom, bust, and post-bust time periods experienced in non-prime mortgage markets in the US.

We further show how non-bank lenders can surreptitiously acquire soft information regarding loan credit quality, and use that information to select against the secondary market. In this extension of our model, secondary market investors do not anticipate this kind of adverse selection, limiting their attention to mortgage prices and loan quantities they expect to see in the secondary market. We show how such adverse selection can exacerbate housing boom and bust, with the portfolio quality of loans sold into the secondary market possibly improving relative to that which would happen with the use of hard information only. We further show how increases in the secondary market loan distribution rate can cause a reduction in the endogenously deter-

⁴⁸As discussed in the Introduction, this aspect of our model departs from Bruecker et al. (2012) where mortgage contracts are non-recourse and borrower's default risk is observable. See Bruckner (2000) for an adverse-selection model of mortgage lending with two types of borrowers that differ in "default costs".

mined quantity of soft information acquired by the originating lender, consistent with the “lax screening” findings of Keys et al. (2010), Purnanandam (2011), Rajan et al. (2015) and others.

References

- [1] Adams, W., Einav, L., and J. Levin, (2009), “Liquidity Constraints and Imperfect Information in Subprime Lending”, *Amer Econ Rev* 99, 49–84.
- [2] Ambrose, B., J. Conklin, and J. Yoshida (2016), “Credit Rationing, Income Exaggeration, and Adverse Selection in the Mortgage Market”, *J Finan* 71, 2637-2686.
- [3] Arslan, Y., B. Guler and T. Taskin (2015), “Joint Dynamics of House Prices and Foreclosures”, *J Money, Credit and Bank* (forthcoming).
- [4] Ashcraft, A., T. Adrian, H. Boesky, and Z. Pozsar (2012), “Shadow Banking”, Federal Reserve Bank of New York, Staff Report No. 458.
- [5] Athreya, K., X. S. Tam, and E. R. Young (2012), “A Quantitative Theory of Information and Unsecured Credit”, *Amer Econ J: Macro* 4, 153-183.
- [6] Bank of International Settlements (2017), “FinTech credit. Market structure, business models and financial stability implications”, http://www.bis.org/publ/cgfs_fsb1.pdf
- [7] Berger, A.N., S. Frame and N. Miller (2005), “Credit Scoring and the Availability, Price, and Risk of Small Business Credit”, *J Money, Credit and Bank* 37, 191-222.
- [8] Bertaud, A., and J.K. Brueckner (2005), “Analyzing Building-Height Restrictions: Predicted Impacts and Welfare Costs,” *Reg Sci and Urban Econ* 35, 109-125.
- [9] Bolton, P., X. Freixas, L. Gambacorta, P. E. Mistrulli (2016), “Relationship and Transaction Lending in a Crisis”, *Rev Financ Stud* 29, 2643-2676.
- [10] Bolton, P. and M. Dewatripont (1994), “The Firm as a Communication Network”, *Quart J Econ* 109, 809-839.

- [11] Brueckner, J. (2000), “Mortgage Default with Asymmetric Information”, *J Real Estate Finan Econ* 20, 251-274.
- [12] Brueckner, J., P. Calem and L. Nakamura (2012), “Subprime Mortgages and the Housing Bubble”, *J Urban Econ* 71, 230-243.
- [13] Brueckner, J., P. Calem and L. Nakamura (2016), “House-Price Expectations, Alternative Mortgage Products, and Default”, *J Money, Credit and Bank* 48, 81-112.
- [14] Brunnermeier, M. K. (2009) “Deciphering the Liquidity and Credit Crunch 2007-2008”, *J Econ Perspect* 23, 23, 77-100.
- [15] Buchak, G., G. Matvos, T. Piskorski, and A. Seru (2018). “Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks”, *J Fin Econ*, in press.
- [16] Case, K.E. (2008), “The Central Role of Home Prices in the Current Financial Crisis: How Will the Market Clear?”, *Brookings Papers on Economic Activity*, Fall 2008, 161-193.
- [17] Chatterjee, S., D. Corbae, and V. Rios-Rull (2011), “A Theory of Credit Scoring and Competitive Pricing of Default Risk”, mimeo.
- [18] Cotter, J., S. Gabriel, and R. Roll (2015), “Can Housing Risk Be Diversified? A Cautionary Tale from the Housing Boom and Bust”, *Rev Financ Stud* 28, 913-936.
- [19] Corbae, D. and E. Quintin (2015), “Leverage and the Foreclosure Crisis”, *J Polit Econ* 123, 1-65.
- [20] Davila, E. (2015), “Using Elasticities to Derive Optimal Bankruptcy Exemptions”, mimeo.
- [21] Dell’Ariccia, G., D. Igan and L. Laeven (2012), “Credit Booms and Lending Standards: Evidence from the Subprime Mortgage Market”, *J Money, Credit and Bank* 44, 367-384.
- [22] Dell’Ariccia, G. and R. Marquez (2006), “Lending Booms and Lending Standards”, *J Finance* LXI, 2511-2546.
- [23] Drozd, L. and Serrano-Pardial (2017), “Modeling the Revolving Revolution: The Debt Collection Channel”, *Amer Econ Rev* 107, 897-930.

- [24] Ellickson B, Grodal B, Scotchmer S, Zame WR (2006), “The Organization of Production, Consumption and Learning”, In: Vind K (ed), The Birgit Grodal symposium. Springer, Berlin.
- [25] Ellickson, B., Grodal, B., Scotchmer, S., Zame, W.R. (1999), “Clubs and the market”, *Econometrica* 67, 1185–1217.
- [26] Fishman, M. J. and J. A. Parker (2015), “Valuation, Adverse Selection, and Market Collapses”, *Rev Financ Stud* 28, 2575-2607.
- [27] Frankel, D. and Y. Jin (2015), “Securitization and Lending Competition”, *Rev Econ Stud* 82, 1383-1408.
- [28] Fuster, A., M. Plosser, P. Schnabl, and J. Vickery (2018), “The Role of Technology in Mortgage Lending”, Federal Reserve Bank of New York, Staff Report No. 836.
- [29] Gabriel, S. and G. Painter (2018), “Why Affordability Matters”, *Reg Sci Urban Econ*, forthcoming. <https://doi.org/10.1016/j.regsciurbeco.2018.07.001>
- [30] Gabriel, S. and S. S. Rosenthal (2015), “The Boom, The Bust, and the Future of Homeownership”, *Real Estate Econ* 43, 334-374.
- [31] Gande, A. and A. Saunders (2012), “Are Banks Still Special When There Is a Secondary Market for Loans?”, *J Finan* LXVII, 1649-1684.
- [32] Garicano, L. (2000), “Hierarchies and the Organization of Knowledge in Production”, *J Politi Econ* 108, 874-904.
- [33] Geanakoplos, J. and W. R. Zame (2014), “Collateral equilibrium, I: A Basic Framework”, *Econ Theory* 56, 443-492.
- [34] Gennaioli, N. and A. Shleifer (2018), *A Crisis of Beliefs: Investor Psychology and Financial Fragility*, Princeton University Press.
- [35] Gennaioli, N., A. Shleifer, and R. Vishny (2015), “Neglected Risks: The Psychology of Financial Crises”, *American Economic Review: Papers & Proceedings* 105, 310-314.

- [36] Green, R. and S. Malpezzi (1996), “What Has Happened to the Bottom of the US Housing Market?”, *Urban Studies* 33, 1807-1820.
- [37] Griffin, J. and G. Maturana (2016), “Who Facilitated Misreporting in Securitized Loans?”, *Rev Financ Stud* 29, 384-419.
- [38] Goldman Sachs, “The Future of Finance: The Rise of the new Shadow Bank”, Report, March 3, 2015.
- [39] Gorton, G. and G. Ordonez (2014), “Collateral Crises”, *Amer Econ Rev* 104, 343-378.
- [40] Guler, B. (2015) “Innovations in Information Technology and the Mortgage Market”, *Rev Econ Dynamics* 18, 456-483.
- [41] Hochguertel, S. and A. van Soest (2001), “The Relation between Financial and Housing Wealth: Evidence from Dutch Households”, *J Urban Econ* 49, 374-403.
- [42] Jaffee, D., A. Lynch, M. Richardson and S. Van Nieuwerburgh (2009), “Mortgage Origination and Securitization in the Financial Crisis”, in *Restoring Financial Stability: How to Repair a Failed System*, John Wiley and Sons, 2009, edited by V. Acharya and M. Richardson, Chapter 1.
- [43] Keys, B.J., T. Mukherjee, A. Seru, and V. Vig (2010), “Did Securitization Lead to Lax Screening? Evidence From Subprime Loans”, *Quart J Econ* 125, 307–362.
- [44] Liberti, J. M. and M. A. Petersen (2017), “Information: Hard and Soft”, mimeo.
- [45] Luque, J. (2018), “Assessing TIF and LIHTC in an Equilibrium Model of Affordable Housing Development”, *Reg Sci Urban Econ*, forthcoming. DOI: 10.1016/j.regsciurbeco.2018.06.005
- [46] Luque, J., N. Ikromov and W. Noseworthy (2018), *Affordable Housing Development: Financial Feasibility, Tax Increment Financing, and Tax Credits*, Springer International Publishing, forthcoming.
- [47] Makarov, I. and G. Plantin (2013), “Equilibrium Subprime Lending”, *J Financ* 68, 849-879.

- [48] Mian, A. and A. Sufi (2011), “Fraudulent Income Overstatement on Mortgage Applications During the Credit Expansion of 2002 to 2005”, *Rev Fin Stud* 30, 1832-1864.
- [49] Mian, A. and A. Sufi (2014), House of Debt, *The University of Chicago Press*.
- [50] Mian, A. and A. Sufi (2009), “The Consequences of Mortgage Credit Expansion: Evidence from the US Mortgage Default Crisis”, *Quart J Econ*, 124, 1449-1496.
- [51] Miller, S. (2015), “Information and Default in Consumer Credit Markets: Evidence From a Natural Experiment”, *J Financ Intermed* 24, 45-70.
- [52] NAHB Research Center (2007), *Study of Subdivision Requirements as a Regulatory Barrier*, prepared for U.S. Department of Housing and Urban Development Office of Policy Department and Research.
- [53] Poblete-Cazenave, M. and J. P. Torres-Martinez (2013), “Equilibrium with Limited-Recourse Collateralized Loans”, *Econ Theory* 53, 181-211.
- [54] Purnanandam, A.K. (2011), “Originate-to-distribute Model and the Sub-prime Mortgage Crisis”, *Rev Financ Stud* 24, 1881-1915.
- [55] Rajan, R., A. Seru and V. Vig (2015), “The Failure of Models that Predict Failure: Distance, Incentives, and Defaults”, *J Financ Econ* 115, 237-260.
- [56] Rajan, R. (2010), “Fault Lines”, Princeton University Press, Princeton, NJ.
- [57] Rao, J. and G. Walsh (2009), “Foreclosing a Dream: State Laws Deprive Homeowners of Basic Protection”, National Consumer Law Center, Inc.
- [58] Seghir, K. (2006), “Overlapping Generations Model with Incomplete Markets: The Numeraire Case II”, *Annales d’Economie et de Statistique* 81, 113-139.
- [59] Stein, J.C. (2002), “Information Production and Capital Allocation: Decentralized Versus Hierarchical Firms,” *J Financ* 57, 1891-1921.
- [60] Stroebel, J. (2016), “Asymmetric Information about Collateral Values”, *J Financ*, 71(2), June 2016

- [61] Vanasco, V. (2017), “The Downside of Asset Screening for Market Liquidity”, *J Finance* LXXII, 1937-1981.
- [62] Vayanos, D. (2003), “The Decentralization of Information Processing in the Presence of Interactions”, *Rev Econ Stud* 70, 667-695.

A Online Appendix

In this Appendix we further comment on our definition of Credit Scoring Technology and contrast it with other approaches in the literature; show that an equilibrium, in the sense of Definition 1, exists; show the main equilibrium closed form solutions that were used in our numerical simulations; work out the minimum house thresholds that prevent a mortgage market for B-type consumers; and dissect two complementary channels that likely contributed to the boom and bust in mortgage and housing markets: the SM investor’s discount factor (a proxy for secondary market funding liquidity) and the fundamental proportion of G-type (higher income) consumers. In addition, we make some remarks regarding the difference between recourse and non-recourse mortgages and their welfare implications. Finally, we work out the details of the extension of our model that accommodates the possibility of non-bank lenders surreptitiously acquiring soft information to select against investors in the secondary mortgage market.

A.1 On the notion of Credit Scoring Technology

It is useful to contrast our definition of CST with that of Rajan et al. (2015). In both cases there is an implicit notion of using what the mortgage lending industry describes as “compensating factors” to reach a *yes-no* lending decision. In this regard, a credit scoring technology can be thought of as a statistically-based multivariate model specification that embeds a set of (more or less) continuous cost-benefit trade-offs to precisely account for compensating factors. This is in contrast to traditional loan underwriting procedures that used bright-line cut-off values (minimum quality standards), which are considered one-at-a-time across a set of risk factors, to screen the mortgage loan application.⁴⁹

⁴⁹Note that bright-line underwriting cutoffs can be accommodated into a statistically based credit scoring model. The use of compensating factors to justify underwriting exceptions with non-prime mortgage loan applicants has

In Rajan et al. (2015), the loan acceptance function (see their equation (1)) is initially conceived as a statistical model estimated in a “low securitization era” using a full set of variables that produce a relatively precise and unbiased *yes-no* underwriting decision. Then, in a “high securitization era”, the model is misapplied, using a reduced set of variables (transmitted hard information only, see their equation (3)), that biases the acceptance decision towards a *yes* outcome. In our model, we take a somewhat different evolutionary view of model estimation in which we acknowledge that non-prime mortgage lending only emerged in the 1990s without a full lending cycle to estimate a model of credit risk. As time progressed into the 2000s, presumably more information became available and estimation technology improved. But it is unclear whether additional lending experience in-and-of-itself improved the statistical properties of the model, since few mortgages had actually defaulted.⁵⁰ Rather, we recognize how other, more qualitative factors, influenced beliefs regarding the perceived quality of the credit scoring technology.

Although there was a confluence of contributing factors, we have in mind three specific qualitative factors. First, and perhaps most importantly, prior to the crisis there was a belief based on more than 50 years of experience that not all housing markets would decline in price at the same time, implying mortgage default outcomes were less susceptible to common shocks and thus more idiosyncratic (Brunnemeier 2009, Case 2009, Cotter et al. 2015). This belief mitigated and effectively truncated concerns regarding far left tail loss outcomes, and subtly increased confidence in the classification accuracy of the model.⁵¹ Second, in a closely related way, during this time there was significant buy-in to the Great Moderation, a belief that monetary policy makers had substantially “tamed” the business cycle (Blanchard and Simon 2001, Bernanke 2004). This further decreased concerns over the costs of misclassification outcomes to increase confidence in the technology. Third, beginning with the Clinton administration in the 1990s and continuing

been common practice since at least the middle 1990s. Federal real estate lending standards currently state: “Some provisions should be made for the consideration of loan requests from creditworthy borrowers whose credit needs do not fit within the institution’s general lending policy.” (Federal Deposit Insurance Corporation, Appendix A to Subpart A of Part 365–Interagency Guidelines for Real Estate Lending Policies, last amended at 78 Fed Reg. 55597, September 10, 2013).

⁵⁰See, for example, Brunnemeier (2009, p.81), who states that models applied by market participants, “provided overly optimistic forecasts about structured finance products. One reason is that these models were based on historically low mortgage default and delinquency rates”.

⁵¹As Brunnemeier (2009, p.81) relates, “past downturns in housing prices were primarily regional phenomena—the United States had not experienced a nationwide decline in housing prices in the period following World War II. The assumed low cross-regional correlation of house prices generated a perceived diversification benefit that especially boosted the valuations of AAA-rated tranches.”

with the Bush administration in the 2000s, there was a strong push towards housing and mortgage lending policies that facilitated home ownership for lower income households.⁵² This also exerted a subtle but critically important influence on model application and assessment.

Thus, our credit scoring model posits no intentional formation of biased beliefs at the time the model is applying. Based on the arrival of information subsequent to the model application, prior beliefs may appear to be biased in one direction or the other, but according to our model there is no intent to misapply it at the time of loan application. Perceived improvements in model classification precision lead to increases in the number of loans originated. Rajan et al. (2015), in contrast, posit the intentional misapplication of the credit assessment model in order to increase the volume of loans originated, where the credit risk of the pool of originated loans is actually higher than what the model indicates.

A.2 Equilibrium Existence

Proof of Theorem 1: We investigate the problem of equilibrium existence by transforming it first into a problem of existence of a social system equilibrium. Our approach is by simultaneous optimization. There, a player’s payoff function and constraint set are parameterized by the other players’ actions. This second dependence does not occur in games. The extension is a mathematical object referred to as a generalized game by Debreu (1952). We carry out this analysis in the continuum of agents framework. Most of our extensions follow by application of Hildenbrand’s (1974) results.⁵³

The generalized game: In the generalized game a player a chooses his strategy x^a parameterized by the other players’ strategies \bar{x}^{-a} . For our economy this game is played by the con-

⁵²See Rajan (2010), especially chapter 1 entitled “Let Them Eat Credit”, for a detailed critique. Among the most significant initiatives was the Federal Housing Enterprises Financial Safety and Soundness Act of 1992, which led the U.S. Department of Housing and Urban Development (“HUD”) in 1993 to establish the nation’s first affordable housing goals. The new standards required the GSEs to ensure that specified percentages of the loans they purchased complied with affordable lending criteria. From 1993 to 1995, the targeted percentage was 30 percent. The goal was increased to 40 percent in 1996, to 42 percent in 1997, to 50 percent in 2001, and to 56 percent by 2008. The GSEs’ pursuit of HUD’s affordable lending goals has been cited as one factor contributing to gains among low-income and minority families in the mortgage market.

Over time, pursuit of the goals caused the GSEs to adjust their offerings and expand the types of loans they purchased. In 1999, for example, under pressure from the Clinton administration, Fannie Mae announced that it would reduce credit requirements on the loans it purchased, thereby encouraging lenders to offer loans to borrowers with lower credit scores.

⁵³See Luque (2013) for a similar approach in a local public goods non-atomic economy

sumers, the lenders, the investors, and five fictitious auctioneers. To incorporate consumers' market choice decisions into the generalized game, we divide the consumers' optimization problem in two stages.

Stage 1: Consumer h chooses his most preferred consumption for a given mortgage market choice $m_{c(h)}^l \equiv (c(h), l)$, i.e., taking $\bar{\mu}^h(m_{c(h)}^l) = 1$ as given. The consumer h 's consumption and loan demand when market choice is $m_{c(h)}^l$ is given by

$$\begin{aligned} (x^h(m_{c(h)}^l), \psi^h(m_{c(h)}^l)) &\in \arg \max \{ u^{h\theta}(\cdot, \bar{\mu}^h(m_{c(h)}^l)) : \\ &\bar{p}_1 H_1^h(m_{c(h)}^l) + R_1^h(m_{c(h)}^l) \leq \bar{q} \psi^h(m_{c(h)}^l) + \omega^{SR}, \psi^h(m_{c(h)}^l) \leq B, \\ &\bar{p}_2 H_2^h(m_{c(h)}^l) + R_2^h(m_{c(h)}^l) \leq \max\{\omega^{SR}, \omega_2^c + \bar{p}_2 H_1^h(m_{c(h)}^l) - \psi^h(m_{c(h)}^l)\} \} \end{aligned}$$

Observe that the choice variables in the constrained optimization problem should all be multiplied by $\bar{\mu}^h(m_{c(h)}^l)$, but we chose to omit it since we are already assuming that $\bar{\mu}^h(m_{c(h)}^l) = 1$ (the consumer is evaluating his utility at specific market choice $m_{c(h)}^l$) - e.g., when writing $H_1^h(m_{c(h)}^l)$ we mean $H_1^h(m_{c(h)}^l) \bar{\mu}^h(c(h), l)$ with $\bar{\mu}^h(c(h), l) = 1$.

Let us show that $(x^h(m_{c(h)}^l), \psi^h(m_{c(h)}^l))$ has nonempty compact values and is continuous. First, notice that $h \rightarrow (x^h(m_{c(h)}^l), \psi^h(m_{c(h)}^l))$ has a measurable graph (see Hildenbrand 1974, p. 59, Proposition 1.b). Non-emptiness follows from the positive endowment assumption. Compactness follows because $H^h(m_{c(h)}^l) \leq \bar{H} < \infty$, $R^h(m_{c(h)}^l) \leq \int_{\mathbf{A}} \omega^a da < \infty$, and $\psi^h(m_{c(h)}^l) \leq \Psi$ if prices p_1 and p_2 are uniformly bounded away from 0 (i.e., $p_1, p_2 \geq \alpha, \alpha > 0$).⁵⁴ Continuity of $(x^h(m_{c(h)}^l), \psi^h(m_{c(h)}^l))$ follows if consumer's demand is both upper and lower hemi-continuous. Since the consumer's consumption set and utility function are both continuous in (x, ψ) and endowments are desirable, we can apply Berge's Maximum theorem to show that $(x^h(m_{c(h)}^l), \psi^h(m_{c(h)}^l))$ is upper hemi-continuous. Next, we prove lower hemi-continuity of $(x^h(m_{c(h)}^l), \psi^h(m_{c(h)}^l))$. Denote by $B^h = \{(x^h, \psi^h) : p_1 H_1^h(m_{c(h)}^l) + R_1^h(m_{c(h)}^l) \leq q \psi^h(m_{c(h)}^l) + \omega^{SR}, \psi^h(m_{c(h)}^l) \leq B, p_2 H_2^h(m_{c(h)}^l) + R_2^h(m_{c(h)}^l) \leq \max\{\omega^{SR}, \omega_2^c + p_2 H_1^h(m_{c(h)}^l) - \psi^h(m_{c(h)}^l)\}\}$ the set of consumer h 's consumption and borrowing amounts that are budget feasible.

⁵⁴One can show that prices are indeed positive with a strictly monotonic utility. The argument is standard and thus omitted for the sake of brevity (one should consider a sequence of truncated generalized games by relaxing α and then applying the multidimensional Fatou's lemma (see Hildenbrand 1974, p. 69, to obtain a cluster point of this sequence; see Poblete-Cazenave and Torres-Martinez 2013).

Claim 1: $(x^h(m_{c(h)}^l), \psi^h(m_{c(h)}^l))$ is lower hemi-continuous.

Proof: Fix $\bar{\mu}^h(m_{c(h)}^l) = 1$ and consider consumer h 's correspondence \dot{B}^h that associates to each vector (p_1, p_2, q) the collection of plans $(x^h, \psi^h, \bar{\mu}^h(m_{c(h)}^l)) \in \mathbf{X}^h$ that satisfies consumer's budget constraints in B^h as strict inequalities. \dot{B}^h has non-empty endowments because consumer's endowments are strictly positive. Also, since the constraints that define \dot{B}^h are given by inequalities that only include continuous functions, the correspondence \dot{B}^h has an open graph. Therefore, for any consumer h , \dot{B}^h is lower hemi-continuous (see Hildebrand 1974, Prop. 7, p. 27). Moreover, the correspondence that associates any vector (p_1, p_2, q) to the closure of the set $\dot{B}^h(p_1, p_2, q)$ is also lower hemi-continuous (see Hildebrand 1974, Prop. 7, p. 26). Now define the closure of \dot{B}^h by \bar{B}^h . We affirm that $\bar{B}^h = B^h$. Since for any (p_1, p_2, q) we have $\bar{B}^h(p_1, p_2, q) \subset B^h(p_1, p_2, q)$, it is sufficient to show that $B^h(p_1, p_2, q) \subset \bar{B}^h(p_1, p_2, q)$.

Given $(x^h, \psi^h) \in B^h(p_1, p_2, q)$ and $(\varepsilon, \delta_1, \delta_2) \in [0, 1]^3$, let $\psi^h(\varepsilon, \delta_1) = (1 - \delta_1)\psi^h + \varepsilon$. We first prove that $((1 - \delta_1)x_1^h, (1 - \delta_2)x_2^h, \psi^h(\varepsilon, \delta_1)) \in \dot{B}^h$, where $x_1^h = (H_1^h, R_1^h)$ and $x_2^h = (H_2^h, R_2^h)$. It is not difficult to see that this last property holds if $\delta_1\omega^{SR} > \delta_1\psi^h - \varepsilon > 0$ (C1) and $\delta_2 = \delta_1(\omega^{SR} + p_2H_1^h)/(\psi^h + p_2H_1^h)$ (C2). In fact, when (x^h, ψ^h) is changed to $((1 - \delta_1)x_1^h, \psi^h(\varepsilon, \delta_1))$, a quantity $\delta_1\omega^{SR} + \varepsilon$ becomes available at the first period. Thus, if (C1), the possible lower revenue from modified debt (if $\delta_1\psi^h - \varepsilon > 0$) is covered by a portion δ_1 of period 1 endowment.

It remains to be shown that a consumer can buy $(1 - \delta_2)x_2^h$. This follows in (C2). To see this, notice that the new resources that become available in the second period are $\max\{\delta_2\omega^{SR}, \delta_2\omega_2^c + p_2\delta_2H_1^h - p_2\delta_1H_1 - \delta_2\psi^h + \delta_1\psi^h - \varepsilon\}$. New resources must be greater than $\delta_2\omega^{SR}$ in the event of no-default, i.e., $\delta_2\omega_2^c + p_2\delta_2H_1^h - p_2\delta_1H_1 - \delta_2\psi^h + \delta_1\psi^h - \varepsilon \geq \delta_2\omega^{SR}$. We know that $\omega_2^G > \omega^{SR}$, so by choosing $\omega_2^c = \omega^{SR}$ we immediately see that sufficient condition (C2) follows.

Finally, making $\delta_1 \rightarrow 0$ (so ε and δ_2 vanish too), we conclude that $(x^h, \psi^h) \in \bar{B}^h(p_1, p_2, q)$, as long as consumers can consume their resources. Thus, correspondence \dot{B}^h is lower hemi-continuous for each consumer. \square

Now, let $\int_{\mathcal{G} \cup \mathcal{B}:c(h)=c} (x^h(m_{c(h)}^l), \psi^h(m_{c(h)}^l)) d\lambda$ represent the measurable demand of goods and loan payments by the continuum of a type C consumers in market m_c^l . Because the aggregate consumer demand function

$$\int_{\mathcal{G} \cup \mathcal{B}:c(h)=C} (x^h(m_{c(h)}^l; \bar{p}, \bar{q}), \psi^h(m_{c(h)}^l; \bar{p}, \bar{q})) d\lambda$$

is the integral of upper semi-continuous demands with respect to a nonatomic measure, we have that $\int_{\mathcal{G} \cup \mathcal{B}: c(h)=C} (x^h(m_{c(h)}^l), \psi^h(m_{c(h)}^l)) d\lambda$ is upper hemi-continuous. The compact-valued function $h \rightarrow (x^h(m_{c(h)}^l), \psi^h(m_{c(h)}^l))$ is bounded above and below by $(\int_{\mathbf{A}} \omega(a) da, \bar{H}, \int_{\mathbf{A}} B da)$ and 0, respectively. According to Hildenbrand (1974, p. 62, Theorem 2), the aggregate consumer demands function is nonempty. And according to Hildenbrand (1974, p. 73, Proposition 7) this set, which is bounded below by 0, is also compact. Therefore, $\int_{\mathcal{G} \cup \mathcal{B}: c(h)=C} (x^h(m_{c(h)}^l; \bar{p}, \bar{q}), \psi^h(m_{c(h)}^l; \bar{p}, \bar{q})) d\lambda$ is compact and has nonempty values. Using a similar reasoning, we can show that the measurable aggregate demand $\int_{\mathcal{N} \cup \mathcal{B} \cup \mathcal{T} \cup \mathcal{B}: c(l)=C} (R^l, \varphi^l) d\lambda$ is compact and has nonempty values.

Observe that the consumer's consumption budget set does not have convex values due to the maximum operator in the second period budget constraint, and therefore, we cannot claim that $(x^h(m_{c(h)}^l), \psi^h(m_{c(h)}^l))$ has convex values.⁵⁵ However, Lyapounov's convexity theorem of an atomless finite dimensional vector measure (see Hildenbrand 1974, p. 62, Theorem 3) implies that the aggregate consumer demand is convex-valued.

Stage 2: Given the consumers' optimal consumption bundles in each mortgage market, consumers choose their most preferred mortgage market (recall that $l = \emptyset$ is a possibility). Let

$$U^h(m_{c(h)}^l) \equiv u^h(x^h(m_{c(h)}^l), \psi^h(m_{c(h)}^l), \mu^h(m_{c(h)}^l)).$$

Then, $\mu^h(m_{c(h)}^l) = 1$ if $l \in \arg \max U^h(l)$ and 0 otherwise (as $\sum_{l=TB, NB, \emptyset} \mu^h(m_{c(h)}^l) = 1$). We represent the pure strategy of consumer h by a basis vector $m \in \mathbf{M}$ of dimension M . The vector $\mu^h(m)$ is the vector in \mathbb{R}^M with 1 as $(m)^{\text{th}}$ coordinate and zero otherwise. By a parallel argument as above, there is a measurable selection $h \rightarrow \mu^h(m)$ with an associated aggregate demand vector $\int_{\mathcal{G} \cup \mathcal{B}} \mu^h(m) d\lambda$, which is the integral of upper hemi-continuous demands with respect to a non-atomic measure. Thus, $\int_{\mathcal{G} \cup \mathcal{B}} \mu^h(m) d\lambda$ is upper hemi-continuous, with compact (by the assumption $\sum_{l=TB, NB, \emptyset} \mu^h(m_{c(h)}^l) = 1$), convex (by Lyapounov's convexity theorem) and nonempty values.

Lenders and investors' objective functions are linear and their choice variables belong to nonempty closed compact sets. Thus, their respective optimization problems pin down prices q^{TB} ,

⁵⁵If the budget set had convex values, then we could have used quasiconcavity of u^h to demonstrate that $x(h, s)$ has convex values.

q^{NB} and τ parameterized in the other players' actions.

Auctioneer 1 chooses p_1 to minimize $\left(\sum_l \int_{\mathcal{G} \cup \mathcal{B}} H_1^h(m_{c(h)}^l) \bar{\mu}^h(m_{c(h)}^l) d\lambda - \bar{H}\right)^2$, where \bar{H} stands for the exogenous supply of housing from an older previous generation. *Auctioneer 2* chooses p_2 to minimize $\left(\sum_l \int_{\mathcal{G} \cup \mathcal{B}} H_2^h(m_{c(h)}^l) \bar{\mu}^h(m_{c(h)}^l) d\lambda - \bar{H}\right)^2$, where \bar{H} stands for the exogenous demand of housing from a younger future generation. *Auctioneer 3* chooses φ^{TB} and φ^{NB} to minimize $\sum_l \int_{\mathcal{G} \cup \mathcal{B}} (\bar{\psi}^h(m_{c(h)}^l) \bar{\mu}^h(m_{c(h)}^l) d\lambda - \int_{\mathcal{L}} \varphi^l(m_{c(h)}^l) \bar{\mu}^h(m_{c(h)}^l) d\lambda)^2$. *Auctioneer 4* chooses z^l to minimize $(d^l \bar{\varphi}^l - z^l)^2$, for $l = TB, NB$. *Auctioneer 5* chooses z_i^l to minimize $(z_i^l - \bar{z}^l)^2$, for $l = TB, NB$. Finally, to guarantee the consistency condition (1.2), we introduce *Auctioneer 6*, whose optimization problem consists of choosing $\pi^l \in [0, 1]$ to minimize $(\pi^l - g(f_G(\hat{\mu}(G, l), \hat{\mu}(B, l)), CST^l))^2$, for $l \in \{TB, NB\}$, where functions f and g are as defined in Section 3.

All auctioneers' strategy sets are nonempty, convex, and compact. An equilibrium for the constructed generalized game consists of a vector $(\bar{x}, \bar{\mu}, \bar{\psi}, \bar{\varphi}, \bar{z}, \bar{p}_1, \bar{p}_1, \bar{q}, \bar{\tau})$ such that each player a chooses a strategy \bar{x}^a to solve his respective optimization problem parameterized in the other players' actions \bar{x}^{-a} .

Claim 2: *There exists an equilibrium in mixed strategies for the constructed generalized game.*

Proof: Note that the consumer's strategy set for choosing his most preferred mortgage market in stage 2 has a finite and discrete space domain \mathbf{M} . In order to circumvent this problem, we extend our generalized game to allow for consumers' mixed strategies in the set of group types \mathbf{M} . Let $\Sigma(\mathbf{M}) = \{\sigma = (\sigma(m))_{m \in \mathbf{M}} : \sigma(m) \geq 0, \sum_{m \in \mathbf{M}} \sigma(m) = 1\}$. Then, $\Sigma(\mathbf{M})$ stands for the convex hull of $\{TB, NB, \emptyset\}$, which is the set of mixed strategies for each consumer. A profile of strategies $\rho : \mathcal{G} \cup \mathcal{B} \rightarrow \Sigma(\mathbf{M})$ brings the continuum of consumers into strategies (pure or mixed). Consumer h 's stage 2 optimization problem extended to mixed strategies is such that this consumer randomizes over the possible consumptions in the set of different market choices. We write $U^h(\sigma) \equiv u^h\left(\sum_{m \in \mathbf{M}} \sigma(m)(x^h(m), \psi^h(m)), \sigma\right)$. That is, *consumer randomizes in \mathbf{M} , but not directly in consumption*. Then, consumer h 's stage 2 maximization problem is $\max_{\sigma \in \Sigma(\Omega)} U^h(\sigma)$. Utility function $u^h\left(\sum_{m \in \mathbf{M}} \sigma(m)(x^h(m), \psi^h(m)), \sigma\right)$ is a continuous bounded real valued function on $\sum_{m \in \mathbf{M}} \sigma(m)(x^h(m), \psi^h(m))$, and the mixed strategy σ belongs to the convex compact set $\Sigma(\mathbf{M})$. $\mathbf{K}(h) = \{\sigma \in \Sigma(\mathbf{M}) : \sigma \in \arg \max U^h(\sigma)\}$ denotes the set of mixed strategies that

solve consumer h 's second stage maximization problem.

We must extend the fictitious auctioneers' problems to allow for consumers' mixed strategies. Given a mixed strategy profile $\rho : \mathcal{G} \cup \mathcal{B} \rightarrow \Sigma(\mathbf{G})$, we can rewrite the auctioneers 1, 2 and 3's objective functions extended to mixed strategies as follows: *Auctioneer 1* chooses p_1 to minimize $(\sum_{m \in \mathbf{M}} \int_{\mathcal{G} \cup \mathcal{B}} H_1^h(m) \rho^h(m) d\lambda - \bar{H})^2$; *Auctioneer 2* chooses p_2 to minimize $(\sum_{m \in \mathbf{M}} \int_{\mathcal{G} \cup \mathcal{B}} H_2^h(m) \rho^h(m) d\lambda - \bar{H})^2$; *Auctioneer 3* chooses φ^{TB} and φ^{NB} to minimize $\sum_{m \in \mathbf{M}} (\int_{\mathcal{G} \cup \mathcal{B}} \bar{\psi}^h(m_{c(h)}^l) \rho^h(m_{c(h)}^l) d\lambda - \int_{\mathcal{L}} \varphi^l(m_{c(h)}^l) d\lambda)^2$, for $l = TB, NB$. All the conditions of Debreu's (1952) theorem hold. Thus, we can assert that the extended generalized game has an equilibrium, possibly in mixed strategies.

At this point it remains to observe that auctioneers 1-3' new (extended) objective functions do not depend only on the average of the consumers' profile, as consumers' demands for commodities may be different among consumers of the same type as they can have different access to the mortgage market. Consequently, we cannot apply Schmeidler (1973) to show that a degenerate equilibrium of the extended generalized game is, in fact, an equilibrium of the original game. Instead, we apply a particular result of Pascoa (1998), used by Araujo and Pascoa (2002, Lemma 2) in an incomplete markets economy, which says that purification can be possible if in the extended generalized game, players' mixed strategies depend only on finitely many indicators, one for each type (a statistical indicator).

In particular, auctioneers 1's extended payoff functions depend on the profile of mixed strategies ρ only through finitely many indicators, one for each consumer type $C = G, B$ in $m \in \mathbf{M}$, of the form $\int_{\mathcal{G} \cup \mathcal{B}} \int_{\mathbf{M}} H_1^h(m; \bar{p}_1) d\rho^h(m) d\lambda$.⁵⁶ Given a mixed strategies equilibrium profile ρ , there exists a profile $(h, m)_{h \in \mathcal{G} \cup \mathcal{B}; m \in \mathbf{M}}$ such that the Dirac measure $\hat{\rho}^h$ at m is an extreme point of the set $\mathbf{K}(h)$, which is the consumer h 's best response to the price chosen by Auctioneer 1 in the previous equilibrium in mixed strategies. And moreover, $\int_{\mathcal{G} \cup \mathcal{B}} \int_{\mathbf{M}} H_1^h(m; \bar{p}_1) d\rho^h(m) d\lambda$ is the same as $\int_{\mathcal{G} \cup \mathcal{B}} \int_{\mathbf{M}} H_1^h(m; \bar{p}_1) d\hat{\rho}^h(m) d\lambda$. Hence, we can replace (h, m) by $(h, \hat{\rho}^h(m))$, for all $h \in \mathcal{G} \cup \mathcal{B}$, and keep all the equilibrium conditions satisfied. The indicators that the atomic auctioneer takes as given evaluated at $\hat{\rho}$ are still the same as when evaluated at ρ . The proofs for auctioneers 2 and 3's payoff functions follow the same lines. Therefore, we conclude that $\hat{\rho}$ is a degenerate equilibrium profile. \square

⁵⁶Observe that we could have written $\sum_{m \in \mathbf{M}} H_1^h(m) \rho^h(m)$ instead of $\int_{\mathbf{M}} H_1^h(m) d\rho^h(m)$.

Claim 3: *An equilibrium for our generalized game (in pure strategies) is an equilibrium as defined in Definition 1.*

Proof: Let $(x, \psi, \varphi, z, \mu, p_1, p_1, q, \tau)$ be an equilibrium in pure strategies of the generalized game introduced above. Our construction of consumers' optimization in stage 1 and stage 2 of the above generalized game imply that equilibrium condition (1.1) is satisfied for consumers - otherwise, we would find a smaller consumption bundle and use continuity to get into a contradiction with the proposed optimum. Equilibrium condition (1.1) for lenders and investors follow from the solutions of lender l and investor i ' linear optimization problems, respectively. Equilibrium condition (1.2) follows from the auctioneer 6's optimization problem. Market clearing conditions are satisfied due to the following reasons: (MC.1) and (MC.2) result from the solutions to auctioneer 3's optimization problem. (MC.3) results from the solutions to auctioneer 4 and 5's optimization problems. (MC.5) follows from the solutions to auctioneers 1 and 2 optimization problems. (MC.4) follow by Walras' law in periods 1 and 2. In particular, we can aggregate all agents' resources in period 1, including the exogenous supply of owner-occupied housing in period 1 from a previous old generation of consumers, and obtain:

$$\zeta_1 \equiv \sum_{m \in \mathbf{M}} \int_{\mathcal{G} \cup \mathcal{B}} \left(p_1 H_1^h(m_{c(h)}^l) + R_1^h(m_{c(h)}^l) - q^l \psi^h(m_{c(h)}^l) - \omega^{SR} \right) \mu^h(m_{c(h)}^l) d\lambda - p_1 \bar{H} + \sum_{l=TB, NB} \int_{\mathcal{L}} (\omega_1^l - q^l \varphi^l + \tau z^l) d\lambda + \int_{\mathcal{I}} (\omega_1^i - \tau z^i) d\lambda \leq 0$$

It is easy to see that, when market clearing conditions (MC.1), (MC.2), (MC.3) and (MC.5) hold, there is no excess demand of the numeraire good consumption in period 1 ($\zeta_1 \leq 0$). Otherwise, we would contradict the above aggregation of budget constraints. In fact, the previous inequality holds with equality (i.e., the market of the numeraire good in period 1 clears). Suppose, by contradiction, that $\zeta < 0$. Then, there is a nonnull set of agents with non-binding budget constraints, which contradicts the optimization requirement. Thus, $\zeta_1 = 0$. By a similar argument, we can also prove that $\zeta_2 = 0$. \square

A.3 Equilibrium Closed Form Solutions

In this section we briefly present the closed form solutions of our equilibrium model. First, mortgage discount prices q^{TB} , q^{NB} and τ follow by solving the lenders TB and NB optimization problems. Here we focus on the case where traditional banks originate-to-own, i.e., $d^{TB} = 0$ (it stands to reason that the higher is d^{TB} for a given d^{NB} , TBs have access to more liquidity from secondary markets and TB mortgage prices and loan amounts will increase). A TB with $d^{TB} = 0$ and $\pi^{TB} = 1$ finds it optimal to set its mortgage discount price equal to its discount factor θ^l , i.e., $q^{TB} = \theta^l$. On the other hand, the discount price for NB loans under hard information only is

$$q^{NB} = \frac{\bar{\pi}\bar{\theta}}{1 - \delta(1 - \bar{\pi})\bar{\theta}} \quad (14)$$

where $\bar{\pi} \equiv \pi^{NB} = \pi^i$ and $\bar{\theta} \equiv d^l\theta^i + (1 - d^l)\theta^l$. Expression (14) follows by using the consumer's first period budget constraint, $p_1 = p_2$ and $\lambda(G\text{-rating})\psi^{TB} = \varphi^{TB}$, and then deriving the corresponding optimality conditions from the lender and investor' optimization problems. The equilibrium for our economy with two consumption goods and two potential funding sources must be consistent with these optimal pricing conditions.

Since $\theta^i > \theta^l$, a higher mortgage distribution rate d^l implies a higher q^{NB} .⁵⁷ Adverse selection is captured by belief $\bar{\pi} < 1$ and decreases the NB's discount price. The term $1 - \delta(1 - \bar{\pi})\bar{\theta}$ in (14) is the "default loss" that the NB incurs when its pool contains an expected fraction $1 - \bar{\pi}$ of B-type borrowers: the higher the default loss, the lower is the discount price that the NB offers to

⁵⁷Greater patience attributable to SM investors can be interpreted as a measure of liquidity in the secondary market. One of the main sources of secondary market liquidity during the middle 2000s derived from demand from the GSEs (Fannie Mae and Freddie Mac) in pursuit of affordable housing goals. HUD required the GSEs to purchase loans that complied with affordable lending criteria, with targeted percentages starting at 30 percent in 1993 and peaking at 50 percent in 2001 and 58 percent in 2008. Fannie and Freddie were able to count purchases of private-label sub-prime MBS towards meeting those goals, including securities containing reduced and low documentation loans (the loans for which income misrepresentation was most prominent—see Ambrose et al. 2016). From 2003 to 2006, according to data from *Inside Mortgage Finance* and the FHFA, those purchases totaled more than \$533 billion and accounted from 36.3% of all securities purchases during that four year window. Other sources of liquidity included depository banks, foreign investors, mutual funds and life insurers, all of which increased their holdings in mortgage-related investment over the 2003-06 time period according to *Inside Mortgage Finance*. Basle II has been cited as an important reason for the increased demand. Foreign investment played a particularly prominent role, where, for example, China increased their holdings of U.S. securities by six times between 2002 and 2008. Accommodating Fed interest rate policy has also been cited as contributing to secondary market liquidity, where, according to John Taylor (2009), targeted rates were a full 3 percent below target, with the "extra-easy policy" contributing to increases in home prices.

its borrowers. The NB's mortgage rate (or cost of capital) is $1/q^{NB}$.

Using the TB and NB' pricing expressions, we obtain

$$EP = \frac{1}{\pi^{NB}(d^{NB}\theta^i + (1 - d^{NB})\theta^l)} - \delta \frac{1 - \pi^{NB}}{\pi^{NB}} - \frac{1}{\theta^l}$$

where π^{NB} is given by (1) and thus increasing in CST^{NB} .

Prices q^{NB} and q^{TB} can be compared as follows:

$$q^{NB} < q^{TB} \text{ if } \pi^{NB} < \pi_2 \equiv \frac{\theta^l(1 - \delta\bar{\theta})}{\bar{\theta}(1 - \delta\theta^l)}.$$

Threshold π_2 is the same as the one found in Section 4 when we characterized the different equilibrium regimes.

Finally, the discount price that investors pay for the subprime mortgages is

$$\tau = \bar{\pi}\theta^i / (1 - \delta(1 - \bar{\pi})\bar{\theta}).$$

Next, we give the closed form solutions for loan amounts. We refer to the pairs (q^{TB}, φ^{TB}) and (q^{NB}, φ^{NB}) as the pooling contracts offered by TBs and NBs, respectively. First, notice that G-type consumers take prices as given, including the mortgage discount price (determined by the lender's optimization problem), and borrow against all their second period revenue, provided that they consume exactly the subsistence rent ω^{SR} . We find the following equilibrium TB loan amount expression:

$$q^{TB}\psi^{TB} = \frac{\theta^l}{1 - \theta^l} \quad (15)$$

which is an increasing function of the TB's discount factor.

Similarly, we find the following equilibrium NB loan amount expression:

$$q^{NB}\psi^{NB} = \frac{\pi^{NB}\bar{\theta}}{1 - \bar{\theta}(\pi^{NB}(1 - \delta) + \delta)} \quad (16)$$

B-type consumers that receive a good rating by the NB misrepresent their type and borrow under the same terms and conditions than G-type consumers. In the expression above we can see that

the equilibrium NB loan amount increases with the predictive power of the hard credit scoring technology (and thus with the NB's belief π^{NB}), the foreclosure recovery rate δ , and the d^{NB} -weighted discount factor $\bar{\theta}$. The term $\bar{\theta}$ in turn increases with the distribution rate d^{NB} and the investor's discount factor θ^i , and decreases with the lender's discount factor θ^l .

The NB's income from distributing mortgages to investors is given by the following expression:⁵⁸

$$\tau^{NB} z^{NB} = \frac{d^{NB} \mu^{NB}(\text{rating=G})}{1 - \delta \bar{\theta} (1 - \pi^{NB})} \quad (17)$$

A.4 Minimum House Size

A.4.1 NB Mortgage Market Specific to B-type Consumers

We focus on the existence of a pooling equilibrium. This is because, as we argued in Section 4, we can rule out the existence of a mortgage market for B-type consumers, given the presence of a minimum house size H_{NB}^{\min} that prevents B-type consumers with a small NB loan to buy a house with a lot size larger than H_{NB}^{\min} .

We now identify threshold H_{NB}^{\min} as a function of the parameters of our economy. First, notice that the NB would get positive profits by lending to a B-type consumer if $q^B \varphi^B \leq \theta^i \delta p H_1^B$ (here we are assuming $d^{NB} = 1$ as this gives the largest loan amount to the consumer since pricing uses the investor's discount factor θ^i). Then, using $p H_1^B = \omega^{SR} + q^B \varphi^B$ from the first period budget constraint (assuming \bar{H} constant in both periods), we get

$$q^B \varphi^B \leq \frac{\theta^i \delta \omega^{SR}}{1 - \theta^i \delta} \equiv \bar{L}$$

that is, \bar{L} is the maximum loan amount that a NB would give to a B-type consumer being compatible with non-negative profits for the lender. Now, going back to the minimum house size regulation argument, we can rule out a mortgage market for B-type consumers if $H^B < H_{NB}^{\min}$, i.e., if $(\omega^{SR} + \bar{L})/p < H_{NB}^{\min}$. The market clearing price for owner-occupied housing is $p =$

⁵⁸The equilibrium quantity of mortgages originated by NBs is constrained by the investor's wealth because $\tau z^i \leq \omega_1^i$. Thus, our model is also able to capture Gennaioli et al. (2013) result that investors' wealth may drive up securitization. To see this, notice that NBs are constrained by the total amount of credit that can be securitized, i.e., $d^{NB} \varphi^{NB} = z^{NB} = z^i$ where the first inequality obeys the originate-to-distribute equation (2) and the second equality follows from market clearing in the secondary mortgage market.

$(2\omega^{SR} + \bar{L} + L^G)/\bar{H}$, where L^G is the loan amount that a G-type consumer would obtain from a NB when mortgage markets are segmented (using the NB's first order condition and that G-type consumer's first period budget constraint we get $L^G = \bar{\theta}(\omega^{SR} + \bar{L})/(1 - \bar{\theta})$). Then, back to the inequality for H_{NB}^{\min} we can write

$$H_{NB}^{\min} > \frac{\bar{H}(\omega^{SR} + \bar{L})}{2\omega^{SR} + \bar{L} + L^G}$$

Hence, we conclude that a minimum house size policy can rule out the possibility of a separating equilibrium when inequality (9) holds.

A.4.2 TB Mortgage Market Specific to B-type Consumers

TBs can in general lend to G-type consumers or to B-type consumers. Similarly to our discussion on the effect of H_{NB}^{\min} on a conduit mortgage market specific for B-type consumers, we can also find a threshold H_{TB}^{\min} that rules out a TB mortgage market specific for B-type consumers. The TB mortgage contract $(q^{B,r}, \psi^{B,r})$ specific for B-type consumers must satisfy budget constraints $pH_1^{B,r} = \omega^{SR} + q^{B,r}\psi^{B,r}$ and $\omega^{SR} = \omega^{SR} - \psi^{B,r} + pH_1^{B,r}$ (the latter coming from the limited recourse requirement), which implies $\psi^{B,r} = pH_1^{B,r}$ and $\psi^{B,r} = \omega^{SR}/(1 - q^{B,r})$. TB's optimization implies that $q^{B,r} = \theta^l \delta$. Thus, $\psi^{B,r} = \omega^{SR}/(1 - \theta^l \delta)$ and using again equation $\psi^{B,r} = pH_1^{B,r}$ we get $H_1^{B,r} = \omega^{SR}/p(1 - \theta^l \delta)$. Then, set

$$H_{TB}^{\min} \equiv \omega^{SR}/p(1 - \delta\theta^l) \quad (18)$$

This housing policy implies that subprime consumers with a small TB loan (or no loan) have no other option but to rent in the first period, because when $p > 1$ these consumers can only afford buying a house of size ω^{SR}/p , which is below H_{TB}^{\min} .

A.5 House Prices

The equilibrium subprime house price depends on the mass of consumers with access to a non-prime mortgage. In particular, we find that:

- If $CST^{NB} < \max\{CST_0, CST_1\}$,

$$p = v(TB) \left(\omega^{SR} + \frac{\theta^l \omega^+}{1 - \theta^l} \right)$$

- If $CST^{NB} \in [\max\{CST_0, CST_1\}, CST_2]$,

$$p = (v(TB) + \lambda_2(G\text{-Rating}^{NB}))\omega^{SR} + \frac{\theta^l \omega^+}{1 - \theta^l} + \lambda_2(G\text{-Rating}^{NB}) \frac{\omega^+ \bar{\theta} \pi^{NB}}{1 - \bar{\theta}(\pi^{NB}(1 - \delta) + \delta)}$$

- If $CST^{NB} > CST_2$,

$$p = (\min\{\lambda_G(1 - CST^{NB}), v(TB)\} + \lambda_3(G\text{-Rating}^{NB}))\omega^{SR} + \frac{\theta^l \omega^+}{1 - \theta^l} \min\{\lambda_G(1 - CST^{NB}), v(TB)\} + \lambda_3(G\text{-Rating}^{NB}) \frac{\omega^+ \bar{\theta} \pi^{NB}}{1 - \bar{\theta}(\pi^{NB}(1 - \delta) + \delta)}$$

where $\lambda(G\text{-Rating}^{NB})$ is the endogenous measure of consumers that borrow from NBs,⁵⁹ i.e., $\lambda_1(G\text{-Rating}^{NB}) = 0$ when $\pi^{NB} < \max\{\pi_0, \pi_1\}$, $\lambda_2(G\text{-Rating}^{NB}) = CST^{NB}(\lambda_G - v(TB)) + (1 - CST^{NB})\lambda_B$ when $CST^{NB} \in [\max\{CST_0, CST_1\}, CST_2]$, and $\lambda_3(G\text{-Rating}^{NB}) = CST^{NB}\lambda_G + (1 - CST^{NB})\lambda_B$ when $CST^{NB} > CST_2$.

With the above expressions, we can now compute the equilibrium house size using the consumer's first period budget constraint. Because $p > 1$, we have that, whenever the consumer has access to a mortgage lender l , the house size consumption is a corner solution, given by expression:

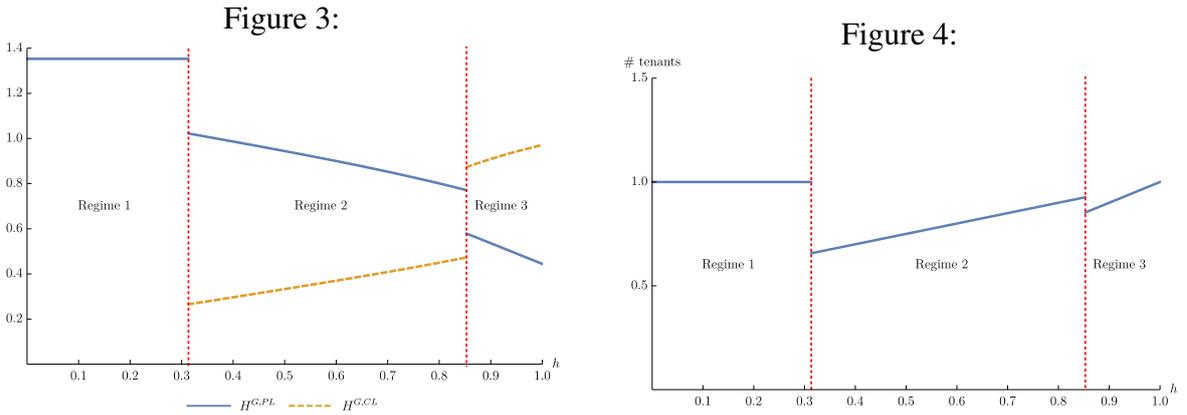
$$H_1^l = \frac{\omega^{SR} + q^l \psi^l}{p}$$

Expression H_1^l depends on the equilibrium regime as p does. In Figure 3 we illustrate the equilibrium values of house sizes H^{TB} and H^{NB} as a function of CST^{NB} . There we see that when the economy enters Regime 3, the house size of borrowers with NB loans is larger than for borrowers with TB loans. This is consistent with the idea that the TB mortgage market is not the

⁵⁹Recall that the measure of consumers that receive a good rating in mortgage market $l \in \{TB, NB\}$ is $\lambda(G\text{-Rating}^l) = CST^l \cdot \hat{\lambda}_G^l + (1 - CST^l) \cdot \hat{\lambda}_B^l$, where $\hat{\lambda}_G^l \equiv \lambda(\mathcal{H} : c(h) = G, \mu^h(m_G^l) = 1)$ and $\hat{\lambda}_B^l \equiv \lambda(\mathcal{H} : c(h) = B, \mu^h(m_B^l) = 1)$ are the measure of G-type and B-type consumers that attempt to borrow from lender $l \in \{TB, NB\}$.

consumers' first option in Regime 3. We also see that the equilibrium house size of consumers with TB loans plummets when the NB loan size enters in Region 3, as the expansion of the NB loan market injects more credit in the economy and house price jumps. Also notice that there is a discontinuity in the equilibrium house size purchased with NB loans when $\hat{\pi}_G^{NB}$ and CST^{NB} are such that $\pi^{NB} = \pi_2$ even when the jump in the NB loan amount is partially offset by the jump in the equilibrium house price at that point.

Figure 4 shows that the size of the rental market is largest in Regime 1 (only the TB mortgage market exists). When Regime 2 starts (CST^{NB} attains π_1), the rental market shrinks as new consumers get (conduit) mortgages. The rental market shrinks again in Regime 3 (CST^{NB} attains π_2), as the NB mortgage market absorbs a substantial larger fraction of G-type and B-type consumers, while the portfolio mortgage market also absorbs those G-type consumers without a conduit loan.



For the sake of brevity, we omit the computation details of thresholds π_0 , π_1 and π_2 (and their corresponding thresholds CST_0 , CST_1 and CST_2).⁶⁰

A.6 Investors' Liquidity and the Distribution of Income

Proof of Proposition 3: The thesis part of Proposition 3.1 follows because a higher θ^i decreases the threshold CST_2^{NB} (as well as thresholds CST_0^{NB} and CST_1^{NB}). For example, using the specified parameters in our simulations above, when θ^i goes from 0.9 to 0.95 all else constant, π_2 falls from 0.74 to 0.69, and CST_2^{NB} falls from 0.85 to 0.81.

⁶⁰The authors can facilitate the algebra details upon request.

To prove Proposition 3.2, first notice that a negative shock to $\lambda_G/(\lambda_G + \lambda_B)$ does not change equilibrium threshold π_2 since

$$\pi_2 = \frac{\theta^l(1 - \delta(d^{NB}\theta^i + (1 - d^{NB})\theta^l))}{(d^{NB}\theta^i + (1 - d^{NB})\theta^l)(1 - \theta^l\delta)}.$$

However, a shock to the fundamental proportion of G-type consumers decreases π^{NB} because it depends on $\hat{\pi}_G^{NB}$ (see expression (1)), and $\hat{\pi}_G^{NB}$ is a decreasing function of $\lambda_G/(\lambda_G + \lambda_B)$. Because of this dependence, a sufficiently large shock to the fundamental proportion will bring π^{NB} below threshold π_2 . ■

A.7 Recourse Versus Non-recourse Mortgages

Limited Recourse in Default: We assumed that mortgage contracts were recourse but subject to limits in consumer liability.⁶¹ In our model payment default will endogenously occur at the beginning of period 2 when a B-type applies for a mortgage loan in period 1 in the NB market and successfully obtains funding to purchase a house (it is misclassified as a G-type).

Limited recourse is achieved as follows. All period 2 consumer income is pledged to the lender in default, subject to a protective “safe harbor” rule that allows the consumer to retain a minimum subsistence level of consumption equal to ω^{SR} . Hence, limited recourse implies that lenders cannot take everything and leave a consumer homeless when he defaults and becomes bankrupt. In fact, homestead rules in bankruptcy are designed to shield consumers from “too much” recourse on mortgage loans. Because B-types have no additional income to pledge in any case, it is costless for the B-type to subject itself to partial recourse. The “teeth” in the limited recourse requirement is that it functions as an ex-ante incentive compatibility condition for G-types that supports a larger loan amount, which in turn increases consumption in period 1 due to an ability to purchase a larger house (see Davila 2015).

Non-recourse mortgage contract: In contrast, strict non-recourse would reduce the mortgage amount and hence consumption for G-types, since there is no credible commitment to fund the mortgage payment in the second period (i.e., strategic default would occur) after the consumer

⁶¹See Poblete-Cazenave and Torres-Martinez’s (2013) for a general equilibrium model with limited-recourse collateralized loans and securitization of debts, where equilibrium is shown to exist for any continuous garnishment rule and multiple types of reimbursement mechanisms.

has consumed housing services from the larger house in period 1.

With a strict non-recourse mortgage, based on previous logic if the house does not sell for at least what the borrower owes, the lender must absorb the difference. Accordingly, in this case the second period budget constraint (7) should be rewritten as follows: $p_2H_2 + R_2 \leq \omega_2^t + p_2H_1 - \min\{p_2H_1, \psi\}$.⁶² If we were to modify our model in such a way, the loan size reduces sufficiently such that the adverse selection problem in the NB loan market would be absent because both types of consumers would be able to always repay their debt using part or all of the proceeds from the house sale, and still consume the subsistence rent ω^{SR} .

A strict non-recourse contract may prevent the consumer from borrowing against all the second period income that is above ω^{SR} , as the promise cannot be larger than the house value (pH_1). Strict non-recourse, by eliminating adverse selection, causes the G-type to delay some consumption until the second period. This is welfare decreasing, since households are impatient and because the younger household derives more utility from owning a house than renting.

Empirical evidence: Partial recourse is available in 40 out of 50 US states. See Ghent and Kudlyak's (2011) table 1 for a summary of the different state recourse laws in the US. According to Rao and Walsh (2009), only three states out of 50 exclude any form of recourse to help lenders recover fees and costs associated with residential mortgage default and foreclosure.

Consistent with our previous discussion, Cao and Liu (2016) find that higher-risk loans are more likely to be originated in recourse states, and Pence (2006) shows that mortgage loan amounts are higher in recourse states.⁶³ Curtis (2014) shows that "lender-friendly" foreclosure law states are associated with larger increases in subprime origination volume.

A.8 Extended Model: The Case of Adverse Selection in the Secondary Mortgage Market

Previously we assumed that non-bank lenders and secondary market investors relied on the same (hard) information in their assessment of mortgage loan credit quality, with accurate and credi-

⁶²See Araujo et al. (2002) for a recursive general equilibrium economy with non-recourse collateral and default.

⁶³It is interesting to observe that some of our predictions appear to be consistent with the data. In particular, Elliot Annenberg used loan performance data to regress the mortgage rate spread of securitized (NB) over portfolio (TB) loans (partialling out observable borrower characteristics) on a state recourse dummy (Ghent and Kudlyak 2011), and found a positive coefficient for the recourse state dummy ($spread = 0.09 + 0.11 \cdot recourse\ dummy$). See Elliot Annenber's discussion at the HULM conference in the Federal Reserve of Chicago (October 23, 2015).

ble transmission of that information from the non-bank loan originator to the secondary market investor. In this section we extend the model to provide a structural analysis of what happens when soft information is acquired by sophisticated non-bank lenders and used to adversely select against secondary market investors. We relabel sophisticated non-bank lenders as SLs.

In the model, adverse selection in the secondary loan investor market occurs when a SL acquires soft information, in secret and at a cost, over and above the hard information that is already available to the unsophisticated non-bank lender (NB) and secondary market investors. With a fixed SM distribution rate, $d^{SL} < 1$, the additional soft information allows the SL to sell lower credit quality loans into the secondary market while retaining higher credit quality loans for itself. The retention of higher credit quality loans, originated at an implied credit spread that reflects risks based on hard information only, provides the incremental profits necessary to pay to acquire soft information.

Because now CST^{SL} may include soft information on top of hard information, we write $CST_{Soft}^{SL} \equiv h + f(s)$, where h denotes the hard information component and $f(s)$ denotes the soft information component. The quantity of soft information acquired is denoted by s . We require f to be any smoothly continuous concave function that satisfies $f(s) \in [0, 1 - h]$. When there is only hard information, we write CST_{Hard}^{SL} which is known to equal h .

Here we consider a setting in which secondary market investors attend to information in the secondary mortgage market that is directly relevant to them, but who are inattentive to other measures such as the total supply of funded mortgage loans. In particular, we assume that secondary market investors believe that loan applications are screened using hard information only given the current credit scoring technology, and pay attention to: 1) the price at which mortgage loans are sold into the secondary market, and 2) the quantity of originated loans sold into the secondary market. The price and quantity constraints are such that they equal price and quantity which obtain in equilibrium when only hard information is acquired by SLs. To sell loans into the SM at any other price and quantity (as specified in the securities prospectus) would otherwise tip off secondary market participants that some unexpected out-of-equilibrium action had been undertaken.

The mortgage price condition that obtains when only hard information is input into the credit

scoring model follows from our previous work (for details, see expression (14)), where

$$q^{SL} = \frac{\pi_H^{SL}(d^l\theta^i + (1-d^l)\theta^l)}{1 - \delta(1 - \pi_H^{SL})(d^l\theta^i + (1-d^l)\theta^l)} \quad (19)$$

with π_H^{SL} denoting the posterior Bayesian probability using *hard* credit information. The mortgage loan price condition is empirically supported by Rajan et al. (2015), who find that subprime conduit lenders set interest rates only on the basis of variables that are reported to investors.

The quantity condition is that the SL sells into the secondary market the same number of loans that would obtain when only hard information is used to assess credit quality. We refer to Section A.8.2 for further details on this constraint. The constraints imposed on SL mortgage loan price and quantity sold into the secondary market allow us to isolate changes in the aggregate supply of originated loans as well as changes in the true credit quality of the mortgage loan portfolio.

Because traditional banks (TBs) only originate mortgages based on a full set of hard and soft information, we assume $d^{TB} = 0$, consistent with the implicit market arrangement that traditional banks establish a “firewall” between the “originate-to-own” and the “originate-to-distribute”. As explained in Section 7, this firewall was imposed by secondary market investors in order to prevent traditional banks from underwriting based on a fuller set of information than was available to secondary market investors.

A.8.1 The Mechanics of Soft Information Acquisition and Adverse Selection in the Secondary Loan Market

In this subsection we specifically examine what it means for the SL to acquire soft information and use that information to select against the secondary market.⁶⁴ The *hard* CST^{SL} h will, as a baseline, determine the precision of borrower type classification. Soft information, on top of hard information, improves precision with respect to classifying loans as G or B. We note that, in general, the classification precision remains imperfect. That is, we do not require that information acquisition results in full information regarding borrower type; rather, we analyze continuous margins to determine the quantity of soft information that is optimally acquired at a

⁶⁴In our model, the SL mortgage price is constrained to mimic the case without soft information acquisition. See Van Nieuwerburgh and Veldkamp (2012) for a more sophisticated framework that solves jointly for investment and information choices, with general preferences and information cost functions.

cost.

From the SL's perspective, conditional on CST_{Hard}^{SL} and the acquisition of soft information, there are four relevant categories of loan classifications. Throughout we assume the law of large numbers applies so that classification precision can be concisely expressed as a probability. Let “*rating hard*” indicate the *yes-no* classification outcome based on hard information only, and “*rating soft*” indicate the SL's own, more precise, classification based on the acquired soft information. The four loan categories are:

$(rating\ soft = G \mid rating\ hard = G):$	Want to own, can sell (reconfirmed as a cherry)
$(rating\ soft = B \mid rating\ hard = G):$	Don't want to own, can sell (downgraded from cherry status, now a lemon)
$(rating\ soft = G \mid rating\ hard = B):$	Want to own, cannot sell (upgraded from lemon status, now a cherry)
$(rating\ soft = B \mid rating\ hard = B):$	Don't want to own, cannot sell (reconfirmed as a lemon)

The total number of loans originated by the SL when only hard information is considered was determined previously to be:

$$N_{Hard}^{SL} = CST_{Hard}^{SL} \lambda_G + (1 - CST_{Hard}^{SL}) \lambda_B \quad (20)$$

where λ_G and λ_B indicate the number of G- and B-types applying for a conduit loan in Regime 3. When Regime 2 applies in equilibrium, the number of G-types applying for a SL loan equal $\lambda_G - v(TB)$ rather than λ_G (see Proposition 2 for a general formulation). Also, as before, to simplify calculations we assume $\Pr^{SL}(\text{rating}=G|G) = \Pr^{SL}(\text{rating}=B|B)$ throughout.

The number of loans originated by the SL is now endogenously determined as a function of the soft information acquired. In order to simplify the analysis, we will assume that the exogenous distribution rate, d^{SL} , is sufficiently large enough so that the quantity of lemons that are available for sale into the secondary market is less than the total number of loans actually sold into that market. This means that, in addition to the lemons, there will be reconfirmed cherries

sold to secondary market investors.⁶⁵ Importantly, reconfirmed cherries are G-types with higher probability than under the hard information only regime.

Now, given a fixed quantity of acquired soft information, and with no internal mortgage quantity constraint of its own, the SL will want to originate all loans which it rates as G-type (reconfirmed cherries as well as upgraded loans). The SL will also originate the downgraded loans and sell them as lemons into the secondary market. That is, the number of loans actually originated by the SL with soft information acquisition in Regime 3 is:

$$N_{Soft}^{SL} = \underbrace{\left(CST_{Soft}^{SL} \lambda_G + (1 - CST_{Soft}^{SL}) \lambda_B \right)}_{\text{measure of loans with "rating soft=G"}} + \underbrace{\left((1 - CST_{Hard}^{SL}) \lambda_B - (1 - CST_{Soft}^{SL}) \lambda_B \right)}_{\text{lemons}} \quad (21)$$

where the appropriate quantity adjustment is made as before if Regime 2 applies in equilibrium. Notice the total quantity of lemons available for sale into the secondary market is determined by the difference $(1 - CST_{Hard}^{SL}) - (1 - CST_{Soft}^{SL}) > 0$. These are the respective probabilities that a SL assigns a good rating to a B-loan under hard and soft information, where this difference is positive when $s > 0$.

Adverse selection in the secondary market occurs when the SL, first, sells all of the downgraded loans (lemons) into the SM, and then fills its quantity sales constraint with reconfirmed cherries. Once the sales quantity constraint is met, the SL retains all of the remaining reconfirmed cherries as well as the upgraded loans. Plugging the above expressions into equation (23), we obtain an expression for the endogenous SM distribution rate:

$$\Delta(s) = \frac{\Theta_{Hard}^{SL}}{N_{Soft}^{SL}} \quad (22)$$

where Θ_{Hard}^{SL} is the number of loans distributed into the secondary market when only hard information is used to assess credit quality. Formally,

$$\Theta_{Hard}^{SL} = d^{SL} N_{Hard}^{SL} \quad (23)$$

⁶⁵That is, there will exist a unique critical d^* in $[0, 1)$ such that, for d^{NL} less than or equal to critical d^* , only lemons are sold into the secondary market. When this occurs, the NL will originate only enough lemons to satisfy the sales quantity condition.

where d^{SL} is the exogenously specified distribution rate used in that regime and N_{Hard}^{SL} is the total number of loans originated by SLs when loan underwriting decisions are made based on hard information only. Below, in the optimization problem for the SL, we shall assume that the SL adheres to requirement (22).

Based on the expressions above, three preliminary remarks follow immediately:

Remark 5: *With soft information acquisition, the increase in the total number of originated loans is $f(s)\lambda_G$.*

Remark 6: *With soft information acquisition, the number of upgraded and downgraded loans are $f(s)\lambda_G$ and $f(s)\lambda_B$, respectively.*

Remark 7: *When the SL has no quantity constraint of its own, the actual distribution rate of loans sold into the secondary market depends on the amount of soft information acquired by the SL, and is decreasing in s .*

A.8.2 Implications for the Equilibrium Mortgage Market Configuration

To incorporate soft information acquisition and adverse selection in the secondary market into our model, we modify the SL's optimization problem as follows. To allow for costly soft information acquisition, let the cost of soft information acquisition increase linearly at a rate of β . In addition, there may be longer run costs to the SL for selling lemons into the secondary market (reputation or legal costs), since repercussions may occur when mortgage loan performance does not line up with expectations. As a result, a penalty parameter, ς , is incorporated into the SLs objective function that is increasing in the number of lemons originated (i.e., total cost is $\varsigma\lambda_B f(s)$). Finally, we recall restricting the SL mortgage price, q^{SL} , and the total number of mortgages distributed, Θ_{Hard}^{SL} , as indicated by expressions (19) and (22).

Formally, the SL chooses $s \in [0, f^{-1}(1 - h)]$ to maximize

$$(\omega_1^l - \beta s - q^{SL}\varphi^{SL} + \tau z^{SL}) + \theta^l(1 - \Delta(s))(\pi_S^{SL}\varphi^l + (1 - \pi_S^{SL})\delta p_2 H_1^G) - \varsigma\lambda_B f(s)$$

subject to the price and quantity restrictions (19) and (22), and where

$$\pi_S^{SL} \equiv \frac{CST_{Soft}^{SL} \hat{\pi}_G^l}{CST_{Soft}^{SL} \hat{\pi}_G^l + (1 - CST_{Soft}^{SL})(1 - \hat{\pi}_G^l)} \quad (24)$$

denotes the SL's belief regarding its retained portfolio quality as a result of soft information acquisition.

Given the quantity and price constraints imposed on loan sales into the secondary market, it follows that the regime boundaries do not change with soft information acquisition and secondary market adverse selection. Further, because SL mortgage price is constrained to equal the equilibrium price that obtains under hard information only, the individual consumer's mortgage size also equates to that obtained with hard information only. But the total quantity of mortgage loans originated increases due to increases in the total number of loans originated, which causes the home ownership rate and house prices to increase in equilibrium.⁶⁶

A.8.3 Soft Information Acquisition

In our extension of the baseline model to allow for SL's soft information acquisition, we assumed that in order to hide the soft information to the secondary market investors, the SL offers the same mortgage discount price and distributes the same number of loans as in the setting with hard information only. Thus, equilibrium regimes do not change. The only difference is on the "number" of mortgages originated, which has the effect of increasing the owner-occupied house price. Therefore, here we only indicate how $\lambda_1(G\text{-Rating}^{SL})$ changes when soft information acquisition is a possibility. The new expressions for Regimes 2 and 3 are the following:

- Regime 2:

$$\begin{aligned} \lambda_{2,Soft}(G\text{-Rating}^{SL}) &= \\ &= \underbrace{\left(CST_{Soft}^{SL}(\lambda_G - v(TB)) + (1 - CST_{Soft}^{SL})\lambda_B \right)}_{\text{measure of loans with "rating soft=G"}} + \underbrace{\left((1 - CST_{Hard}^{SL})\lambda_B - (1 - CST_{Soft}^{SL})\lambda_B \right)}_{\text{lemons}} \end{aligned}$$

⁶⁶Given the limited recourse nature of the subprime mortgage contract, $\psi^{SL} = \omega_2^G - \omega^{SR} + p_2 H_1$. Thus, consumers get the same loan amount $q^{SL} \psi^{SL}$ than in the no-soft information setting; however, as shown previously, there are more loans originated under soft information. Since the SL's mortgage discount price q^{SL} doesn't change, market clearing for the SL loan market follows by accommodating φ^{SL} to the number of loans originated under soft information; see the Appendix for the closed form market clearing equations corresponding to Regimes 2 and 3.

- Regime 3:

$$\lambda_{3,Soft}(G\text{-Rating}^{SL}) = \underbrace{\left(CST_{Soft}^{SL}\lambda_G + (1 - CST_{Soft}^{SL})\lambda_B \right)}_{\text{measure of loans with "rating soft=G"}} + \underbrace{\left((1 - CST_{Hard}^{SL})\lambda_B - (1 - CST_{Soft}^{SL})\lambda_B \right)}_{\text{lemons}}$$

Thus, the market clearing equation is

$$\begin{aligned} \lambda_{2,Soft}(G\text{-Rating}^{SL})\psi^{SL} &= \varphi^{SL} \text{ if Regime 2} \\ \lambda_{3,Soft}(G\text{-Rating}^{SL})\psi^{SL} &= \varphi^{SL} \text{ if Regime 3} \end{aligned}$$

A.8.4 House prices

Figure 5 illustrates the owner-occupied house price level with and without soft information acquisition, showing that adverse selection in the secondary market magnifies housing booms – both within and across regimes. Similarly, a negative shock to house prices propagated by a loss of confidence in credit scoring technology (or through some other complementary channel) may be accompanied by a reduction in or an elimination of soft information acquisition, which further magnifies housing busts.

A.8.5 Secondary Market Investor’s Portfolio Quality

Conditional on soft information acquisition, an important question is whether the credit quality of the portfolio of loans sold into the secondary market deteriorates. Intuition suggests that it would, but in fact we will show that portfolio quality can actually improve. The credit quality of loans sold into the SM are, nevertheless, always inferior to loans retained by the SL due to the fact that lemons are sold first and not retained by the SL.

Define the SM investor’s *ex ante* (expected) portfolio quality to be $\pi_{\text{ex-ante}}^i \equiv \Pr_{\text{Hard}}[G|\text{rating=G}]$ using hard information only as originally defined in equation (1). The actual, *ex post* portfolio quality conditional on acquiring a specified quantity of soft information depends on the proportion of lemons sold into the SM relative to cherries sold. Define the proportion of lemons sold, ρ^L , as $\rho^L = f(s)\lambda_B/N_{\text{Hard}}^{SL}$, where $\rho^C = 1 - \rho^L$. With this, the credit quality of the portfolio

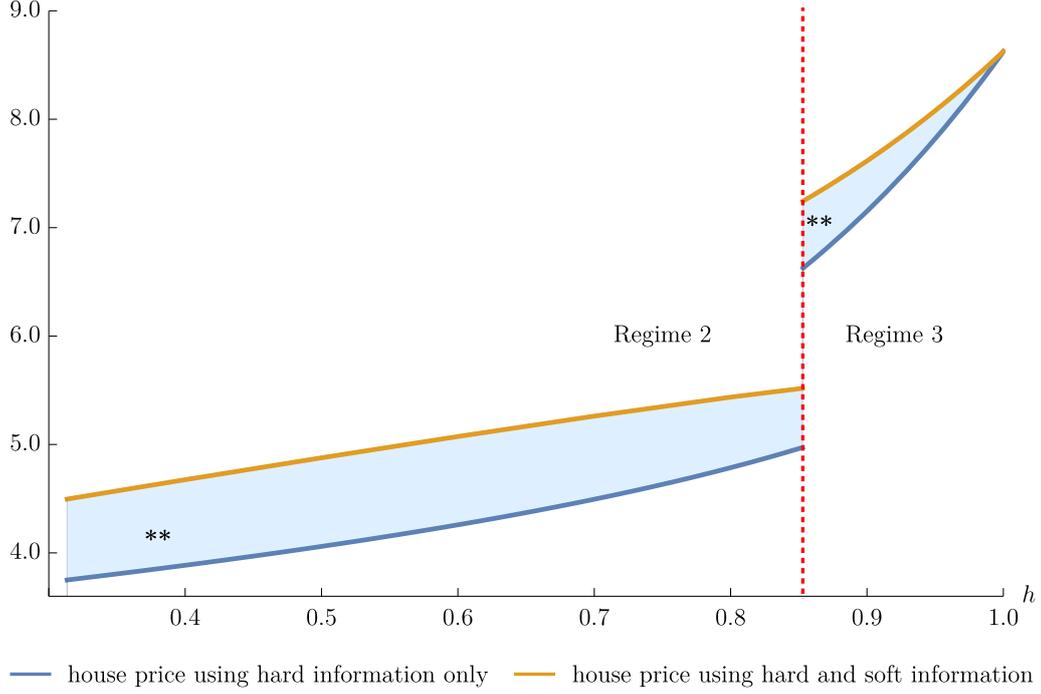


Figure 5: This figure portraits the price level, with and without soft information, as a function of $CST_{Hard}^{CL} = h$, The light-blue** -region captures the price increase due to a higher number of loans downgraded and sold as lemons with soft information acquisition.

actually sold into the SM is expressed as follows:

$$\pi_{\text{ex-post}}^i \equiv \rho^C \Pr_{\text{Soft}}[G|\text{rating=G}] + \rho^L \Pr_{\text{Soft}}[G|\text{rating=B}], \quad (25)$$

where $\Pr_{\text{Soft}}[G|\text{rating=B}] = \pi_G^{SL}$ (as given by (24)) and

$$\Pr_{\text{Soft}}[G|\text{rating=B}] = \frac{(1 - CST_{\text{Soft}}^{SL})\hat{\pi}_G^l}{(1 - CST_{\text{Soft}}^{SL})\hat{\pi}_G^l + CST_{\text{Soft}}^{SL}(1 - \hat{\pi}_G^l)}.$$

Given these relations, it is straightforward to show that, as CST_{Soft}^{SL} approaches perfection ($\Pr(\text{rating=G}|G) = \Pr(\text{rating=B}|B) = 1$), the credit quality of the portfolio of secondary market loans always deteriorates relative to a hard information only regime. This result primarily occurs because lemons default with certainty and therefore don't contribute anything to the portfolio quality measure.

But when credit scoring under soft information acquisition is imperfect, it can be the case that

the credit quality of the portfolio of loans sold into the secondary market actually improves. This is because, with soft information acquisition, the loans that are reconfirmed and sold as cherries possess a higher probability of correct classification than under hard information only. At the same time, however, the downgraded loans have a relatively low (but non-zero) probability of performing well. The relative proportions of reconfirmed cherries and downgraded lemons in the sold portfolio will in combination with posterior probabilities determine the overall credit quality of loans sold into the secondary market.

In Figure 6 we plot the SM investor’s portfolio quality (as stated in equation (25)) as a function of the amount of soft information s . There we confirm that when soft information takes small values ($s < 0.05$), SM investors are actually better off than what they expected, whereas for larger values of s ($s \geq 0.05$) portfolio quality increasingly deteriorates.

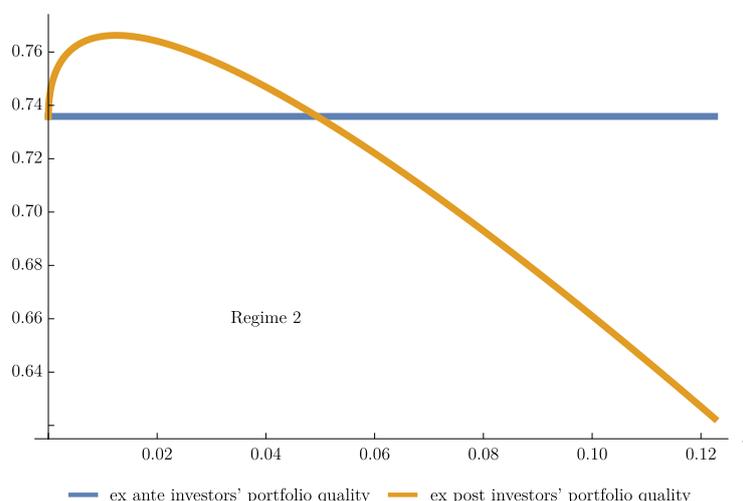


Figure 6: This figure portrays the ex-ante and ex-post investor’s portfolio quality as a function of soft information.

Notice that Figure 6 takes soft information as a parameter. Soft information is in fact endogenously determined in equilibrium, implying that changing the value of any parameter value that affects the SL’s optimal amount of soft information acquired will end up modifying the relationship between SM investor’s portfolio quality before and after the acquisition of soft information.

In unreported simulations we confirm that, in equilibrium, a higher market mortgage distribution rate d reduces incentives for SLs to acquire soft information while the economy remains in Regime 2, consistent with the findings of Keys et al. (2010) and Dell’ Ariccia et al. (2012). This

result is intuitive, since a higher distribution rate, d^{SL} , results in relatively fewer retained loans by the SL, which in turn decreases incentives to acquire soft information at a positive marginal cost. We can also show that in equilibrium a reduction in soft information acquisition by the SL due to increased sales distribution into the secondary market does not necessarily make SM investors worse off. This occurs because a decline in soft information acquisition reduces the quantity of lemons sold into the secondary market. In fact, in unreported simulations we find that in Regime 2 a higher d^{SL} lowers soft information acquisition, making SM investors better off than what they expected. Investors only become worse off than expected when d^{SL} is so high that it triggers a transition from Regime 2 to Regime 3.

A.8.6 Further relationship with the literature

This extension of the model sheds light into the skin-in-the-game bank regulation literature, which requires the loan originator to “eat some of its own cooking” through risk retention (Jaffee et al. 2009). With respect to this issue, we first note that our model implies that risk retention would increase the cost of capital for the non-bank simply due to a reduction in liquidity transfer. Further, our model suggests that the policy could be subject to manipulation should originating lenders decide to engage in adverse selection and/or income misrepresentation. For example, we show that adverse selection incentives of non-bank lenders naturally increase in the loan retention rate, which could be an unintended consequence of the regulation. With respect to income exaggeration and other types of misrepresentation, effects depend on the source of the misrepresentation. If the borrower is the source of the problem, and the originating lender (as well as secondary market) is truly oblivious to the problem, increased retention increases credit exposure and risk of the regulated bank. Perhaps retention increases incentives of the originating lender to detect borrower misrepresentation, but it is unclear whether this was an intended part of the policy. On the other hand, should the originating lender be aware that misrepresentation is occurring, and without enforceable deterrence mechanisms in place, there is nothing preventing that lender from stuffing the secondary market loan asset pool full of lemon loans, and keeping the better stuff for itself, thus undermining the regulation.

The analysis of our extended model has certain similarities to Fishman and Parker (2015) (see also Bolton et al. 2016 and Vanasco 2017). A critical difference in model structure is

that secondary market investors recognize the adverse selection problem in Fishman and Parker, whereas secondary market investors fail to fully anticipate this type of adverse selection in our model (but they are fully cognizant of borrower adverse selection). Increased precision in loan credit quality assessment increases asset prices in our model, as more loans are originated by “sophisticated” banks, whereas asset value is decreasing in information acquisition in their model. Fishman and Parker conclude that markets generally produce too much information, since market responses to information acquisition by better informed agents leads to market breakdown and other related inefficiencies. We in contrast show that in certain circumstances equilibria exist in which unsuspecting secondary market investors are selected against, yet the credit quality of their portfolio improves to improve allocative efficiency.

Online Appendix List of References

- [1] Araujo, A. and M. Pascoa (2002), “Bankruptcy in a Model of Unsecured Claims”, *Econ Theory* 20, 455-481.
- [2] Araujo, A., M. Pascoa, and J. P. Torres-Martinez (2002), “Collateral Avoids Ponzi Schemes in Incomplete Markets”, *Econometrica* 70, 1613-1638.
- [3] Blanchard, O. and J. Simon (2001), “The Long and Large Decline in U.S. Output Volatility”, *Brookings Papers on Economic Activity* 1, 135-174.
- [4] Brunnermeier, M. K. (2009) “Deciphering the Liquidity and Credit Crunch 2007-2008”, *J Econ Perspect* 23, 23, 77-100.
- [5] Cao, Q. and S. Liu (2016), “The Impact of State Foreclosure and Bankruptcy Laws on Higher-Risk Lending: Evidence from FHA and Subprime Mortgage Originations”, *J Real Estate Research* 38, 505-538.
- [6] Case, K.E. (2008), “The Central Role of Home Prices in the Current Financial Crisis: How Will the Market Clear?”, *Brookings Papers on Economic Activity*, Fall 2008, 161-193.

- [7] Cotter, J., S. Gabriel, and R. Roll (2015), “Can Housing Risk Be Diversified? A Cautionary Tale from the Housing Boom and Bust”, *Rev Financ Stud* 28, 913-936.
- [8] Curtis, Q. (2014), “State Foreclosure Laws and Mortgage Origination in the Subprime”, *J Real Estate Fin and Econ* 49, 303-328.
- [9] Debreu, G. (1952), “A Social Equilibrium Existence Theorem”, *Proceedings of the National Academy of Sciences*.
- [10] Gennaioli, N., A. Shleifer, and R. Vishny (2013), “A Model of Shadow Banking”, *J Finance* 68, 1331-1363.
- [11] Ghent, A. and M. Kudlyak (2011), “Recourse and Residential Mortgage Default: Evidence from US States”, *Rev Financ Stud* 24, 3139-3186.
- [12] Hildebrand, W. (1974), *Core and Equilibrium in a Large Economy*, Princeton: Princeton University Press.
- [13] Luque, J. (2013), “Heterogeneous Tiebout Communities with Private Production and Anonymous Crowding”, *Reg Sci Urban Econ* 43, 117-123.
- [14] Pascoa, M. (1998), “Nash Equilibrium and the Law of Large Numbers”, *Int J Game Theory* 27, 83-92.
- [15] Pence, K. M. (2006), “Foreclosing on Opportunity: State Laws and Mortgage Credit”, *Rev Econ Stat* 88, 177-182.
- [16] Poblete-Cazenave, R. and J.P. Torres-Martinez (2013), “Equilibrium with Limited-recourse Collateralized Loans”, *Econ Theory* 53, 181-211.
- [17] Rajan, R., A. Seru and V. Vig (2015), “The Failure of Models that Predict Failure: Distance, Incentives, and Defaults”, *J Financ Econ* 115, 237-260.
- [18] Rao, J. and G. Walsh (2009), “Foreclosing a Dream: State Laws Deprive Homeowners of Basic Protection”, National Consumer Law Center, Inc.

- [19] Taylor, J. B. (2009), *Getting Off Track: How Government Actions and Interventions Caused, Prolonged, and Worsened the Financial Crisis*, Hoover Institution Press. ISBN 0-8179-4971-2.
- [20] Van Nieuwerburgh, S. and L. Veldkamp (2012), “Information Acquisition and Under-Diversification”, *Rev Econ Stud* 77, 779-805.
- [21] Schmeidler, D. (1973), “Equilibrium Points of Non-atomic Games”, *J Statistical Physics* 7, 295-301.