The Credit Scoring and Transmission Channels in the Non-Prime Mortgage Market *

Jaime Luque Timothy Riddiough

University of Wisconsin - Madison

Abstract

We provide a theory that rationalizes how credit scoring technologies control the flow of capital into the non-prime mortgage market. The home ownership rate, the source of mortgage capital, mortgage quantities, portfolio credit quality, and house prices are all determined endogenously in a general equilibrium model with embedded credit scoring. Lending regimes are identified and characterized in the context of lending-house price boom and bust. Adverse selection against secondary market investors and income misrepresentation are analyzed as distortions to the transmission of credit quality information.

Key words: non-prime lending; credit scoring technology; traditional banks; shadow banks; hard and soft information

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

*We thank for comments and suggestions Shane Auerbach, Henrique Basso, Marcus Berliant, Elliot Anenberg, Zahi Ben-David, Jan Brueckner, Briana Chang, Gonzalo Fernandez de Cordoba, Scott Frame, Bulent Guler, Barney Hartman-Glaser, Ben Keys, Erwan Quintin, Steve Malpezzi, Cyril Monnet, Wayne Passmore, Thomasz Piskorski, Marzena Rostek, Ricardo Serrano-Pardial, Shane Sherlund, Juan Pablo Torres-Martinez, Stijn Van Niewerburgh, Kerry Vandell, Abdullah Yavas, Yiro Yoshida, and the participants and discussants at the annual meetings of ASSA-AREUEA (San Francisco), SAET (Cambridge, UK), Finance Forum (Madrid, Spain), EWGET (Naples, Italy), NUS -IRES Annual Research Symposium (Singapore), DePaul conference in Economics and Finance (Chicago), UCLA Conference on Housing Affordability (Los Angeles), HULM (Chicago Fed), Fed Atl Real Estate Finance conference, and seminars at the UC-Irvine, Bank of Spain, IE School of Business, U. Bern and U. Wisconsin-Madison. We are particularly grateful to Xudong An and Vincent Yao for many discussions and their insights into the workings of the subprime secondary mortgage market. Authors email addresses: jluque@wisc.edu and timothy.riddiough@wisc.edu. Send correspondence to Jaime Luque at jluque@wisc.edu, Wisconsin School of Business, 975 University Avenue, Madison, WI 53706-1324. This work was not supported by any funding agency or institution.

1 Introduction

The rapid development of non-prime consumer lending channels has been one of the most important as well as most controversial financial market developments of the past few decades. Non-prime mortgage lending has in particular garnered much attention due to its proximity to the financial crisis of 2007-08, although other non-prime consumer debt markets such as auto and credit card lending have experienced parallel market developments. Proponents of the development of such markets tend to highlight market completeness and access to credit that had been previously denied to less advantaged segments of the population, while detractors generally emphasize agency frictions and resource misallocations that result with the introduction of complex contracting-market structures and when lower-income consumers become over-extended due to "easy" credit availability.

While there has certainly been theoretically motivated work developed to explain non-prime consumer mortgage lending, it seems fair to say that empirical work in the area has outpaced theory and that much of the received wisdom regarding non-prime mortgage loan market characteristics has derived from headline empirical findings (see, for example, Mian and Sufi 2009 and Keys et al. 2010). With this paper we seek to contribute to the theory literature by incorporating primary empirical findings into the development of a general equilibrium model of non-prime (higher-risk) mortgage lending. In doing so we aim to explain historical market development, while also accounting for more recent (post-crisis) innovations in the market (Buchak, Matvos, Piskorski, and Seru 2017). The model is specifically tailored to account for possibly biased beliefs formed regarding the classification accuracy of available credit scoring technologies, as well as to address adverse selection and the potential misrepresentation of borrower information transmitted to secondary market investors.

The main structural features of our model are: 1) Tenure choice (rent v. own decisions) by consumers; 2) Two mortgage loan funding sources for consumers, namely, traditional banks and shadow banks, where shadow banks originate-to-distribute loans into a secondary market; 3) A direct link between aggregate non-prime mortgage supply and house prices at the neighborhood level; 4) An imperfect credit scoring technology that controls the flow of mortgage capital into non-prime housing markets; 5) The possibility of soft information acquisition and adverse selection in the secondary loan market, with implications for measuring "laxness" in loan underwriting; and 6) The misrepresentation of hard information that is transmitted to secondary market investors. Home ownership rates, mortgage loan rates and loan amounts, house prices and portfolio loan credit quality are all endogenously determined in equilibrium, implying the endogeneity of loan acceptance rates and loan screening “laxness” as measured by the exclusive reliance on hard information in making loan funding decisions. Because consumers can choose among owning with a traditional bank loan, owning with a shadow bank loan, or renting, the non-prime mortgage market structure is also endogenous, a unique feature of our model.

In the rest of this Introduction, we highlight the main features and results of our model, and also explain how it contributes to the understanding of some important empirical facts reported in the recent literature.

Non-prime households and asymmetric borrower credit quality information

The empirical literature has identified non-prime neighborhoods as geographically delineated areas with defined populations, typically at the zip code level (Mian and Sufi 2009, Agarwal, et al. 2012). Consumers who populate these neighborhoods are characterized by relatively low current incomes, few financial resources (little or no savings) and low credit scores. Consumers living in non-prime neighborhoods can either rent or own. Because of their low net worth and minimum house size constraints, consumers must borrow to finance the purchase of a house. Accordingly, non-prime households living in non-prime neighborhoods constitute the starting point in our model.

Consistent with Mian and Sufi (2014, 2017), Ambrose, Conklin, and Yoshida (2016), and others, we focus on income as a key input into the loan qualification decision, with shocks to income driving lending outcomes. In particular, a fixed and known proportion of consumers in the defined population are expected to experience a positive income shock in the next period, while the rest remain at their subsistence income levels (G- and B-types in the model, respectively).

Borrower type at the individual level is not known with certainty to shadow banks and secondary market security investors, thus creating the fundamental source of credit risk in the model. For G-type consumers, limited recourse (to excess income) in the case of default functions as a commitment device that increases mortgage proceeds and hence utility. With this commitment, G-type consumers repay their loans, whereas B-type consumers that receive mortgage funding lack the income to meet their repayment promise and thus default.

Traditional banks versus shadow banks

Consideration of alternative loan funding sources for consumers is important to understand the non-prime mortgage market structure. Such a theoretical model seems to be missing in the literature, as pointed out by Agarwal, Amromin, Ben-David, Chomsisengphet and Evano (2011). In our model, we allow for two potential sources of mortgage funding. One source is traditional banks, which are characterized as brick and mortar operations that function like "relationship" lenders. They retain loans on their balance sheet (originate-to-own), and are constrained in quantity of mortages they can offer in any given period (e.g., they have an inelastic deposit base or internal allocation constraints). Importantly, their local presence and integrated nature creates soft information that allows for precise identification of the borrower's type. As a result, only G-type consumers apply to traditional banks should the traditional bank be the preferred source of mortgage funding, with G-types qualifying for a loan until the capacity constraint is met.

The second potential source of mortgage loan funding, shadow banks, are specialized financial service organizations that do not necessarily have a local presence and that are "transactional" in nature. They may have a limited amount of their own capital, largely or exclusively relying on pooled loan asset sales into a secondary mortgage market for funding (they largely originate-to-distribute). This results in an exogenously specified loan sale distribution rate into the secondary mortgage market. In our baseline model, shadow banks rely on only freely available hard consumer credit information to make *yes-no* underwriting decisions and to price the loans. Shadow banks and secondary market investors prefer to lend only to G-type consumers; but, because of the imprecise nature of the credit scoring technology, classification errors occur, which result in a certain proportion of B-type consumers receiving mortgage loan funding.

The fundamental tradeoff with respect to the source, cost and quantity of mortgage funding is the availability of soft information in credit risk assessment in the case of the traditional bank versus liquidity and credit risk transfer in the case of the shadow bank. Due to being endowed with superior credit quality information, traditional banks enjoy low rates of default (strictly zero in the model), which reduces the cost of mortgage capital. On the other hand, depending on the classification precision of available credit scoring technology, as well as the proportion of G- versus B-types in the defined population that apply for a loan, shadow banks experience a higher relative default rate on the loans they originate. This causes the pooled mortgage loan rate to increase relative to the loan rate offered by the traditional bank. But secondary market investors are more patient than originating lenders, which lowers the cost of funds to partially or wholly offset the credit risk spread caused by imprecise loan underwriting. The cost of borrowing (mortgage loan proceeds) can be lower (higher) in either the traditional bank or shadow bank mortgage loan market depending on market conditions, and is a critical factor in generate equilibrium outcomes.

Credit scoring technology

A primary focus and distinguishing feature of our paper is how credit scoring technology affects the consumer's mortgage funding application decision and relative loan pricing. Credit scoring technology is modeled as a Bayesian prior, and hence a belief, that places a probability on the likelihood of classification error in the *yes-no* funding decision. Statistically-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 in non-prime mortgage markets is limited, as it was prior to the onset of the recent financial crisis.

A Bayesian posterior is then generated by combining the credit scoring technology with the proportion of G- versus B-types in the defined population of loan applicants. This posterior determines the credit quality of the shadow bank mortgage loan pool, and thus the pooled mortgage loan price. Perceived improvements in credit scoring technology, or increases in the proportion of G-types (those who are expected to experience a positive income shock), are shown to increase the competiveness of the shadow banking sector relative to traditional bank lending, causing loan amounts and house prices to increase. Excess optimism in the classification accuracy of the available credit scoring technology (possibly based on hindsight), with a subsequent reversal in that optimism, is considered as it operates through the credit scoring channel (see Brunnemeier 2009, p. 81 for more on the role of optimistic beliefs in the subprime mortgage meltdown and broader financial crisis). So are the effects of shadow bank acquisition of soft information used to select against secondary market participants, which serves to distort the transmission of credit scoring information to secondary market participants. Lastly, the misrepresentation of hard information, particularly the exaggeration of past or current income, is shown to potentially operate through both the credit scoring technology channel (affecting *yes-no* funding outcomes) as well as through the estimated proportion of G- versus B-types in the defined non-prime consumer population. Upward biasing as a result of income misrepresentation distorts Bayesian posteriors to cause resource misallocations vis-a-vis mortgage funding decisions and house prices.

Equilibrium market configurations

The proportion of G- versus B-types applying for a loan in the shadow banking sector and home ownership rates embedded in tenure (own-versus-rent) decisions are determined endogenously within our general equilibrium model. Consumer preferences with respect to owning versus renting, and where to borrow, depend on the relative cost of mortgage debt available in the traditional bank versus shadow bank loan markets, which in turn depends on the intensity of the borrower adverse selection problem. Thus, the cost and availability of non-prime mortgage funding affects house prices, consistent with the empirical findings of Mian and Sufi (2009).

With a focus on the credit scoring and transmission channels, we characterize equilibrium market configurations as three alternative mortgage funding-house price regimes that aptly describe the evolution of non-prime mortgage lending in the US. In a first regime, traditional bank lending is dominant, as credit scoring technologies are sufficiently primitive so as to make the shadow banking mortgage loan market uncompetitive. Then, with improvements in credit scoring technology associated with updated beliefs as to its classification accuracy, possibly combined with overly optimistic forecasts of non-prime borrower income gains (Rajan 2010, Mian and Sufi 2017), a shadow banking market emerges in a second regime as the back-up choice for G-type borrowers that are rationed out of the preferred traditional bank loan market. Imperfect classification precision creates the opportunity for B-types to apply for a mortgage loan in this regime, hoping to be misclassified as a G-type. Loan application acceptance rate and the pooled mortgage loan rate are endogenously determined, where we show that loan application acceptance rates can vary in either direction as they depend on the available credit scoring technology.[1] House prices accelerate upward when the shadow bank loan market emerges, with home ownership rates expanding rapidly due to availability of secondary market mortgage capital. A third regime occurs when perceived credit scoring precision improves sufficiently so that the shadow bank loan is preferred over the traditional bank loan market by G-types. House prices increase rapidly at the regime boundary due to all G-types migrating into the shadow bank loan market for lower-cost

[1]Our 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, Frame and Miller (2005).

financing. This migration endogenously increases the proportion of G-types applying for a loan, causing mortgage rates (quantities) to experience large discrete decreases (increases).[2]

Endogenous house prices and the credit cycle

Our model captures the ebbs and flows of shadow bank activity, which often peaks just prior to a downturn. In our model the peak in shadow bank activity corresponds with a bottom in the utilization of soft information in loan screening, and with a peak in secondary market sales. This is consistent with Purnanandam's (2010) evidence that the lack of screening incentives coupled with leverage-induced risk-taking behavior significantly contributed to the recent subprime mortgage crisis. We note that boom-bust equilibrium outcomes are generated in our baseline model without appealing to mispresentation or *lender*'s adverse selection. In this respect our work is different from previous theory that considered *lender*'s adverse selection as the main reason of the expansion and collapse of lending (see e.g. Fishman and Parker 2015, Frankel and Jin 2015, and Gorton and Ordonez 2014).

Our equilibrium mechanism also links non-prime mortgage lending standards to the run-up and eventually collapse in home-prices, and thus 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 booms in our model occur when G-types migrate *en masse* into the shadow bank market, with regime boundaries identifying tipping points. And what goes boom can also go bust when shadow bank/secondary market participants recalibrate downward the assessed classification accuracy of the credit scoring technology in the face of new information indicating that their models failed to accurately predict failure (Case 2008, Rajan, Seru and Vig 2015).[3] The recalibration might also be the result of revising overly optimistic expectations regarding upward income mobility of non-prime consumers. This bust scenario matches rather closely the market's reaction when early term defaults hit the subprime

[2]As pointed out by Ashcraft, Adrian, Boesky and Pozsar (2012), at the onset of the financial crisis the volume of credit intermediated by the shadow banking system was close to $20 trillion, or nearly twice as large as the volume of credit intermediated by the traditional banking system at roughly $11 trillion.

[3]As 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 syncronized, the model no longer fit the data". See also Cotter et al. (2005) for more on the issue of regional price syncronization.

mortgage market in late 2006 and early 2007 (see Brunnemeier 2009, pp.81-83), raising, among other things, concerns over the possible misrepresentation of borrower incomes stated in loan applications (Purnanandam 2011, Ambrose et al. 2016, Mian and Sufi 2017).[4]

Lax Screening, Lending Standards, and Loan Acceptance Rates

Our baseline model endogenizes "lax screening" of the type identified by Keys, Mukherjee, Seru, and Vig (2010), Purnanandam (2011) and Rajan, Seru and Vig (2015), where the introduction of a secondary mortgage market implies increasing reliance on hard information only in the loan screening process. That is, because shadow banks screen based on hard information only, whereas traditional bank lenders utilize soft as well as hard information, screening becomes increasingly lax as the shadow banking sector gains market share vis-a-vis perceived improvements in credit scoring technology.

Exclusive reliance of secondary market investors on hard information introduces the opportunity for originating lenders to exploit those investors. To address this possibility, we extend our baseline model to account for the endogenous (surreptitious) acquisition of soft information by shadow banks, used to select against investors in the secondary mortgage market. The retention of some proportion of originated loans is necessary for the shadow bank to have incentives to engage in these actions, since otherwise there would be no source of revenues to offset the costs of information acquisition. Soft information acquisition improves the precision of the credit scoring model, 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.

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 shadow bank sells the lemons and retains a loan portfolio whose credit quality exceeds that based on hard information only loan pricing. Importantly, we show that 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

[4] Brueckner, Calem and Nakamura (2012) show that mortgage lenders' expectations of a high house price growth increase the consumer's FICO score through current prices and spur subprime lending. In our paper the credit scoring technology is income-only driven (it screens good and bad types by looking at their second period endowments only), and this is enough to generate the collapse of the subprime conduit mortgage market.

quantity of mortgage loans originated. Finally, we show that the endogenously determined quantity of soft information acquired decreases in the mortgage distribution rate, which extends lax screening findings of Keys, Mukherjee, Seru, and Vig (2010), Purnanandam (2011) and Rajan et al (2015) to the case of secondary market adverse selection.

Lastly, we address the issue of income misrepresentation in non-prime mortgage lending. In the context of our model, income misrepresentation equates to the miscoding of hard credit information. When income is exaggerated, credit scoring model classification outcomes are biased towards the *yes* side, leading to systematic loan underpricing and underperformance. As stressed by Mian and Sufi (2017), such actions will likely affect the estimated proportions of G- versus B-types in the defined non-prime consumer population, as exaggerated past and current income can lead to excess optimism with respect to upward income mobility. When misrepresentation is revealed to have occurred, it is likely that downward revisions occur with both the credit scoring model and the proportion of G-types in the population, resulting in sharp declines in non-prime mortgage supply and house prices (Cotter, Gabriel, and Roll 2015).

Skin-in-the-game bank regulation requires the loan originator to "eat some of their own cooking" through risk retention (Jaffee, Lynch, Richardson and Van Nieuwerburgh 2009). With respect to this issue, we first note that our model suggests that risk retention would increase the cost of capital for the shadow bank simply due to reductions in credit risk and liquidity transfer, but that does not seem to be the stated underlying goal of the regulation. 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 secondary market adverse selection incentives are shown to actually increase in the retention rate. 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 exposure to the loan originator. 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 strong deterrence mechanisms in place, there is nothing preventing that lender from stuffing the secondary market loan asset pool full of affected loans, and keeping the better stuff for itself, thus undermining the regulation.

The rest of this paper is organized as follows. In Section 2, we further explain the relationship between our paper and the literature. In Section 3, we present the baseline model. Section 4 provides the equilibrium definition, states the equilibrium existence result, and discusses the characteristics of equilibrium mortgage rates. Section 5 characterizes equilibrium mortgage market configurations as they depend on the available credit scoring technology. Section 6 examines the role of credit scoring technology 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 7 extends the baseline model to accommodate for soft information acquisition and adverse selection against secondary market investors, exploring implications on portfolio credit quality, endogenous loan sale distribution rates, and house prices. Section 8 addresses issues associated with income misrepresentation and concludes. The Online Appendix is reserved for additional explanations, results, and proofs.

2 Further Relationship with the Literature

To better rationalize our model structure as well as to isolate and highlight our contributions to the literature, in this section we compare and contrast our paper to a selected subset of papers in what is fast becoming a vast literature. The starting point for this paper is, as noted earlier, Mian and Sufi (2009), who focus on subprime consumers residing in subprime neighborhoods. Over the critical 2002-05 time period, the mortgage supply channel is identified, with subprime consumers helping themselves to increasingly cheap and available non-prime mortgage debt to purchase homes. Disproportionate house price increases are documented in subprime neighborhoods. Mian and Sufi (2014, 2017) highlight similar themes, where in addition they focus specifically on the role of income in mortgage lending decisions and repayment outcomes, showing that income exaggeration by subprime consumers occurred in subprime neighborhoods, further distorting mortgage flows and house prices.

Our specific focus on the credit scoring and transmission channels is motivated by the work of Keys, Mukherjee, Seru, and Vig (2010), Purnanandam (2011), and Rajan, Seru and Vig (2015), who highlight differences in hard versus soft information in the lender screening process.[5] In doing so they argue that the introduction of a secondary market for non-prime mortgages led to greater reliance on hard information in the loan underwriting process to result in what they term "lax screening." Like Rajan, Seru and Vig (2015), we highlight the credit scoring and transmission channels by placing a clearly articulated structure on the loan screening-mortgage pricing process, where our primary contribution is to embed this structure into a larger general equilibrium framework.[6] We further create a clear structural role for income in the loan qualification process and for income shocks as they effect repayment outcomes, which we then highlight in our analysis of lending regimes and regime shifts, secondary market adverse selection and income misrepresentation.

Stroebel (2016) empirically analyzes asymmetric information in home development real estate lending, distinguishing between integrated versus non-integrated lenders. He finds that integrated lenders have better information regarding the cost quality of homes than non-integrated lenders, which integrated lenders exploit to lend against higher quality collateral. The performance of non-integrated lender loans is significantly worse than integrated lender loans, where borrowers rejected by integrated lenders that end up borrowing from non-integrated lenders do particularly poorly. Non-integrated lenders recognize they are competing against more informed lenders, and price their loans accordingly. Stroebel's focus on collateral rather than borrower credit quality, and the fact that there are no secondary market investors in his analysis, limits the scope of comparison. Nevertheless, this paper provides an empirical foundation upon which we build the basic information structure of our theory.

Our characterizations of belief formation, the resulting effects of the credit scoring channel

[5]For recent work that focuses on how different lenders' information sets affect mortgage loan outcomes, borrowers' default, and market unraveling, see, e.g., Karlan and Zinman (2009), Adams et al. (2009), and Einav et al. (2013). See Miller (2015) for a related analysis of the importance of information provision to subprime lender screening. More generally, see Stein (2002) for a description of how private information includes soft information, and how difficult is to communicate soft information to other agents at a distance.

[6]Our treatment of the credit scoring technology is different than Chatterjee, Corbae, and Rios-Rull (2011) and Guler (2015) in that they do not distinguish between hard information and soft information, nor between traditional banks and shadow banks. Importantly, the equilibrium structure of the subprime mortgage market is endogenous, a unique feature of our model. Because we distinguish between shadow bank and traditional bank funding models, and relate their change in market share to different equilibrium non-prime mortgage configuration regimes that result from changes in the credit scoring technology and securitization, our work also differs from previous theoretical models on shadow banking and subprime lending (see, e.g., Makarov and Plantin 2013 and Piskorski and Tchistyi 2011).

on mortgage and housing markets, the centrality of borrower income and employment to credit assessment, and the post-crisis decline of subprime lending ties closely to the empirical findings of Cotter, Gabriel and Roll (2015). As noted in the Introduction, prior to the financial crisis there were widely held beliefs that price changes in spatially dispersed housing markets were only weakly correlated. This belief in the effects of geographical diversification strongly mitigated investment risk concerns of mortgage originators and investors, which, in turn, helped offset concerns as to the dearth of data on the performance of subprime loans over the cycle and therefore model goodness-of-fit.

Cotter et al. show that returns to housing became increasingly integrated starting in 2004 (during the heart of the subprime lending boom), continuing to increase through 2010. Increased price syncronization is shown to be strongly associated with the securitization of non-prime mortgages and the ease of mortgage underwriting. Further, consistent with the prominent role we ascribe to downward revisions in non-prime consumer income levels in explaining the bust in house prices, and as a direct complement to the recent findings of Mian and Sufi (2017) with regard to income exaggeration during the boom period, the authors find that, "employment and income fundamentals contributed importantly to the ongoing trending up in house price return integration."

Drozd and Serrano-Padial (2017) closely complements our work, and provides a useful contrast in its study of the "debt collection channel." These authors consider the role of information technology (IT) in addressing distressed unsecured consumer loans, where the focus is on the back-end of the loan's life rather than on the front end. The authors document that "depository and non-depository financial institutions are the most IT-intensive industry in the U.S.," and how IT-enabled credit assessment emerged in unsecured consumer lending in the 1990s. The main argument and model prediction is that:

> "IT-driven improvements in debt collection – which in our model take the form of improved precision – lead to a more prevalent use of risky loans... As signals become more precise, lenders adopt them to target delinquent consumers that are more likely to pay back their debt. This lowers the cost of sustaining risky loans, making such loans more prevalent in equilibrium..."

Their focus and findings mirror in part our arguments on the loan origination side, where increases in the precision of credit assessment technology lead to an increased market share of shadow banking and hence riskier lending 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 lending as opposed to more risky lending, since classification error is reduced as precision increases. Interestingly, the debt collection channel works through a reduction in the severity of loss to the lender, not the frequency of default. The credit scoring channel in our model is complementary, working primarily through default frequency. Ensuring a sufficiently low level of loss severity to the lender vis-a-vis partial lender recourse is, however, embedded in our model structure and is essential in order to generate a commitment by G-types to repay the loan. Without partial recourse, loan amounts would otherwise decline significantly and risky lending would be curtailed.

Our analysis of adverse selection in the secondary market has certain similarities to Fishman and Parker (2015) (see also Bolton et al. 2016). In their model there are sophisticated and unsophisticated investors, which are analogous to shadow banks that can acquire soft information at a cost and secondary market investors that purchase mortgage loans from shadow banks. Sophisticated investors can acquire information to increase their precision of loan value, buying more good loans and leaving more bad loans for unsophisticated investors. 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 value estimation increases asset (house) prices in our model, as more loans are originated by shadow 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 equilibria exist in which unsophisticated investors are selected against, yet the credit quality of their portfolio improves to improve allocative efficiency.

The empirical literature on the misrepresentation of hard information associated with loans sold into the secondary mortgage market has expanded rapidly in recent years. We refer the interested reader to Ambrose, Conklin and Yoshida (2016) for a good summary of the literature.

Piskorskit et al. (2015) and Griffin and Maturana (2016a) are of particular interest due to their focus on misrepresentation originating from issuers of mortgage-backed securities (MBS), and Griffin and Maturana (2016b) and Mian and Sufi (2014, 2017) in their focus on causally linking mispresentation and house prices. We note that our model is capable of addressing the mispresentation of hard information originating from either the borrower or the (shadow bank) lender, and in the case of borrower misrepresentation, where the shadow bank lender is either unaware or aware that the mispresentation occurred. Our model can also address the causative effects of misrepresentation of loan application information on house prices, with mispresentation resulting in increased loan application acceptance rates and higher resulting house prices (Mian and Sufi 2014, pp.76-79).

3 Baseline Model

3.1 Basic Model Set-Up

In our baseline 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, which live in homogeneous neighborhoods. Neighborhood delineation in empirical analysis is often defined by zip code, as in Mian and Sufi (2009), Agrawal et al. (2012), and others. By non-prime we mean that the observed credit quality of the borrower-loan is such that it fails to 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 credit quality.

There are two consumption goods: owner-occupied housing and rental housing. We take rental housing herein as the numeraire. 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 $\bar{H} = 1$.[7] Later, in Section 5, we will explain that our baseline model naturally extends to an

[7]The assumption of inelastic neighborhood housing supply is realistic given that most non-prime neighborhoods are located in previously developed urban areas. See Mian and Sufi (2009) for further discussion and evidence.

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 further assume that rental and owner-occupied housing districts within a non-prime neighborhood are sufficiently segmented and that they are subject to separate market clearing conditions.

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 some consumers experience a positive income shock (e.g., get a better job), 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 fundamental proportion measures the aggregated credit quality of residents residing in a non-prime neighborhood. This measure depends on currently available information regarding the households' employment and income growth prospects, among other factors. Some of this information is common knowledge at the local-neighborhood level, but some is not, requiring factual reporting and assessment at the household level. Following the literature on income misrepresentation (e.g., Jiang, Nelson and Vytlacil 2014, Ambrose, Conklin and Yoshida 2016, Mian and Sufi 2017), our model can distinguish between reported versus true information used in forming beliefs regarding borrower type. *Misrepresentation* of borrower income may affect not only specific assessments of borrower type based on hard (transmittable) information, but also the fundamental proportion of G and B types in the loan application pool. While the former effect has received significant attention in the empirical literature to date, the latter issue has not and will be taken up in some detail later in the paper (Mian and Sufi 2017 is a recent exception).

In the baseline model, however, we assume no misrepresentation of borrower income or other credit-relevant information.

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 rental endowment to fund ownership. The threshold $H^{\min}$ is motivated by the cost of regulation imposed throughout the original construction and subsequent housing maintenance accounts.[8] 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).

To obtain a mortgage, consumers may attempt to borrow from two different types of funding sources: a *traditional bank lender* or a *shadow bank lender*. The basic distinction between the two types is that traditional bank lenders are recognized as such, whereas shadow bank lenders operate as mortgage loan origination specialists outside the traditional banking structure. Traditional bank lenders (TBs) can be further characterized as relational, originating mortgages to be held in their own asset portfolio (they "originate-to-own"). In contrast, shadow bank lenders (SBs) are transactional, selling most if not all of their mortgages into a secondary loan market (they primarily "originate-to-distribute"). SB access to secondary mortgage markets can reduce the cost of mortgage capital due to differences in valuations between SBs and secondary market (SM) mortgage investors.

TBs can also be thought of as brick-and-mortar enterprises that know their borrowers and their communities, offering deposit, checking, credit card and other financial services. A physical presence by the TB may lead to relatively higher operating costs, but in the process also generates soft information, available at no additional cost, regarding the borrower's type. This soft information that is in possession of the TB is not transferable to outsiders, and is on top of available hard credit information (e.g., FICO scores, borrower employment/income growth prospects, certain known neighborhood characteristics, publicly available information on local economic growth prospects). Hard information is available to all agents at no cost. The combi-

[8]According to S&P Global Ratings SF Research, in 2011 the cost of regulation imposed throughout the construction and land development process accounted, on average, for $85,000 of the cost of a new single family home. See Malpezzi and Green (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 and Duraton, Henderson and Strange (2015).

nation of hard and soft information allows TBs to know the consumer's type with certainly.[9]

Prior to the financial crisis, SBs were often referred to as "conduit lenders".[10] Post-crisis much of the growth in the non-prime market has been attributed to "Fintech," which are firms that generally operate as on-line mortgage originators (see Buchak, Matvos, Piskorski, and Seru 2017). SBs in our baseline model rely only on hard credit information, which as noted is accurately recorded and transmitted. Reliance on hard information alone does not allow for a perfect assessment of consumer type, however. As such, in the baseline model, SBs are incapable of fully resolving asymmetric information over and above the available with hard information as it is input into their credit scoring technology.[11]

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 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 traditional bank lenders and shadow bank lenders by $\mathcal{TB}$ and $\mathcal{SB}$, respectively. For simplicity, we will assume $\lambda(\mathcal{TB}) = \lambda(\mathcal{SB}) = \lambda(\mathcal{I}) = 1$. For the sake of presentation, we will also write $\lambda(\mathcal{G}) \equiv \lambda_G$ and $\lambda(\mathcal{B}) \equiv \lambda_B$.

3.2 Lenders, Credit Scoring, and Limited Recourse in Default

Both TBs and SBs only originate a mortgage if the consumer is identified as a G-type (i.e., rating=G). Lending only to G-types is endogenously determined in our model, following because

[9] See Keys et al. (2010) and Rajan et al. (2015) for more on the role of soft information in non-prime mortgage lending. We 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 SB does not have. We make the complete information assumption with no real loss of generality and only to simplify the analysis.

[10] Other common labels include "distance" lenders, "non-integrated" or "narrow" banks, and "arm's length" lenders.

[11] Chatterjee, Corbae, and Rios-Rull (2011) allow consumers to borrow multiple times to study the role of reputation acquisition where the individual's type score is updated every period according to Baye's rule. This setting better describes characteristics of prime borrowers who build some credit reputation over time by borrowing in multiple occasions. In our paper we study non-prime consumers who obtain mortgage loans on a one-off or an occassional basis.

of the presence of limited recourse mortgage contracts and a minimum house size. However, given an imperfect credit scoring technology applied by SBs, it will be the case that some B-type consumers are misclassified as G-types when the SB mortgage lending market is active, implying that they receive mortgage funding and end up defaulting on their mortgage loans.

Given that TBs always classify borrowers correctly by their type, they will only attract G-types as part of the loan application process. This is not the case with SBs, however, who are vulnerable to adverse selection at the loan application stage. In order to improve the likelihood of a correct classification, the SB applies what we term a *credit scoring technology*.

We define the credit scoring technology that is applied by the shadow bank as $CST^{SB} = \Pr^{SB}(\text{rating=G}|\text{G})$ and assume, for simplicity, that $\Pr^{SB}(\text{rating=G}|\text{G}) = \Pr^{SB}(\text{rating=B}|\text{B})$. This subjective probability is, in general, founded on credit experience that is incorporated into 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.

It is useful to contrast our definition of CST^{SB} with that of Rajan, Seru and Vig (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. A credit scoring technology is a statistically-based multivariate model specification that embeds a set of (more or less) continuous cost-benefit tradeoffs to more precisely assess compensating factors. This is in contrast to traditional loan underwriting procedures that used bright-line cut-off values, which are considered one-at-a-time across a set of risk factors, to screen the mortgage loan application.[12]

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 baseline model, we take a somewhat different evolutionary view of model estimation in

[12]Note 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 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).

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.[13] 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, Gabriel and Roll 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.[14] Second, in a closely related way, during this time there was significant buy-in to the Great Moderation, a belief that monetary policymakers 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 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.[15]

[13]See, 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".

[14]As 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."

[15]See 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.

This also exerted a subtle but critically important influence on model application and assessment.

We further note that this same set of factors influenced estimates of the fundamental proportion of G-type households in defined neighborhoods, but did so with somewhat different effects. Optimistic beliefs regarding housing and labor market outcomes can skew the fundamental proportion upwards, which operates like a common shock if beliefs are subsequently revised downwards in light of new information about fundamentals. These revision effects can be exacerbated by concerns over misrepresentation and adverse selection in the secondary market – the kind of effect described by Mian and Sufi (2017) in their analysis of upwardly-biased income misrepresentation in subprime neighborhoods. Again, we postpone analysis of adverse selection and misrepresentation until after the baseline model is fully presented and characterized.

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

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

where $\hat{\pi}_G^{SB}$ denotes the proportion of G-type consumers among all consumers that attempt to borrow from the SB. This posterior probability provides the basis for loan pricing conditional on loan approval. It 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 Bayesian priors and the relative proportion of G-types in the loan application pool.[16]

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

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

This belief π^l is endogenous in our model, as it depends on the endogenous variable $\hat{\pi}_G^l$. When

[16] To better understand the mechanics of our formulation, suppose that the traditional bank loan market is such that G-types prefer to borrow from TBs over SBs. Further suppose that there is a TB capacity constraint of 100 loans. If there are 150 G-type and 100 B-type consumers in total, the fundamental proportion of G-types, $\lambda_G/(\lambda_G + \lambda_B)$, is 0.60. However, once 100 G-types are served by TBs, the impaired loan pool left for SBs is 50 G-type consumers and 100 B-type consumers. The corresponding proportion is $\hat{\pi}_G^{SB} = 0.33$ for SBs. If, on the other hand, G-types prefer to borrow first from SBs, the credit quality of the pool is not impaired, with $\hat{\pi}_G^{SB} = 0.6$. Thus, in our baseline model, $\hat{\pi}_G^{SB}$ is an endogenous variable, whereas CST^{SB} enters as a parameter. Changes in $\hat{\pi}_G^{SB}$ and CST^{SB} will be a focus in our theory of the credit scoring channel.

$l = TB$ we write π^{TB}, and when $l = SB$ we write π^{SB}. Notice that by assumption $CST^{TB} = 1$, and therefore $\hat{\pi}_G^{TB} = 1$ and $\pi^{TB} = 1$.

Shadow banks are subject to an "originate-to-distribute" allocation constraint, which restricts them to distribute no more than a promised fraction d^l of its originated mortgage payments. In particular,

$$z^l \leq d^l \varphi^l, \tag{2}$$

where $\varphi^l \geq 0$ denotes the face value of mortgages bought by lender l, $z^l \geq 0$ is the face value of mortgages issued by lender l that are passed onto investors, and d^l is the fraction of mortgages originated for distribution. φ^l and z^l are choice variables for lender l, while d^l is a parameter that takes value $d^{TB} = 0$ if the lender is a TB, and $d^{SB} \in (0, 1]$ if the lender is a SB.[17] We say that a pair (φ^l, z^l) is feasible if it satisfies (2).[18]

We assume limited lender recourse in default. 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 SB market and successfully obtains funding to purchase a house (it is misclassified as a G-type). Mortgage laws in most states in the US allow for partial recourse, in the sense that, in addition to the pledge of the house as security for the loan, lenders can access certain additional borrower assets to help plug any remaining gap between the net value of the house in foreclosure and the amount owed on the mortgage.[19] In our model, in addition to the house, period 2 income is the only available form of security the consumer can pledge in support of the loan.

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

[17]In practice, d^{CL} is often equal to 1. But notice that it can be smaller than 1 if, for example, SBs promise to keep "skin-in-the-game" or are subject to regulation that requires such a practice.

[18]Observe that the "originate-to-distribute" constraint (2) has the flavor of a box constraint as introduced in Bottazzi, Luque, and Pascoa (2012) to model repurchase agreements, where the amount of a security short sold is bounded by the total amount of securities possessed by an agent. However, we do not allow for security borrowing in our model (i.e., borrowing mortgages to short sell), and thus constraint (2) is closer to the type of no-short sale constraint introduced in Faias and Luque (2016).

[19]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.

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 this requirement is that it functions as an ex-ante incentive compatability 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). In contrast, full 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 has consumed housing services from the larger house in period 1. Thus, in the end, B-type consumers cannot credibly commit to pay back the loan, but a G-type consumer can with a limited recourse condition.[20]

Foreclosure is costly for the lender. There are direct foreclosure costs and other indirect costs associated with repossessing the house and reselling it, and possibly some foreclosure delays. This results in a loss of $(1 - \delta)p_2H_1$ to the lender, 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.

Choices φ^l and z^l determine the lender l's consumption, expressed as a two dimensional vector $x^l \equiv (\omega_1^l - q^l\varphi^l + \tau z^l, (1 - d^l)(\pi^l\varphi^l + (1 - \pi^l)\delta p_2 H_1^G)) \in \mathbb{R}_+^2$. The first and second components in this equation represent the consumption of the numeraire good in the first and second periods, respectively. 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.[21] The discount price of the mortgages sold in the secondary mortgage market is τ. The lender's first period endowment ω_1^l is assumed to be positive, and, for simplicity, we assume the second period endowment equal to $\omega_2^l = 0$. The term $\pi^l\varphi^l + (1-\pi^l)\delta p_2 H_1$ represents the expected payoff corresponding to the lender's pool of mortgages kept in its own portfolio. With probability $1 - \pi^l$

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

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

the lender expects to have originated a mortgage loan to a B-type. The loan defaults at the beginning of period 2 with a recovery of $\delta p_2 H_1$. Also, notice that the lender only keeps a fraction $1 - d^l$ of its mortgages originated, and thus we weight any income received in the second period by this term.

All lenders are risk neutral with time discount factor $\theta^l < 1$.[22] The lender l's profit function can now be stated as

$$\Phi^l(\varphi^l, z^l) \equiv (\omega_1^l - q^l\varphi^l + \tau z^l) + \theta^l(1 - d^l)(\pi^l\varphi^l + (1 - \pi^l)\delta p_2 H_1) \tag{3}$$

The lenders optimization problem is as follows. Each lender l chooses φ^l and z^l to maximize $\Phi^l(\varphi^l, z^l)$ subject to the originate-to-distribute constraint (2). We denote the lender l's choice set by $\mathbf{C}^l \subseteq \mathbb{R}_+^2$, which is composed of all pairs (φ^l, z^l) that are budget feasible.

Notice that the interaction between the originate-to-distribute constraint (2) and the profit function (3) determines the two possible loan origination models. On the one hand, shadow banks distribute a fraction $d^{SB} > 0$ of the originated mortgages in exchange for some revenue today. This exchange with secondary market investors works because investors are more patient than SBs. But SBs lack soft information (so $\pi^{SB} < 1$). Traditional banks, on the other hand, possess soft information (so $\pi^{TB} = 1$ by assumption), but don't sell their mortgages to the investors ($d^{TB} = 0$). Consequently they lack revenue in the first period, and are willing to operate this way because they are more patient than consumers. [23]

[22]The assumption of lender's risk neutrality is common in the literature. See e.g. Arslan, Guler, and Taskin (2015), Chatterjee, Corbae and Rios-Rull (2011), Guler (2015), and Fishman and Parker (2015).

[23]Traditional banks and shadow banks are strongly differentiated in our model, with TBs unable to sell loans into the secondary mortgage market. We note that in practice, TBs, 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 in which loans were originated for distribution. It was the case that secondary market participants generally required a "firewall" to be established between the TB and SB parts of the business in order to prevent the TB from underwriting based on a fuller set of information than was available to secondary market investors. Our model fully accomodates this kind of setting, where later, when we analyze the acquisition of soft information to select against the secondary market. With the extended model strucutre, we are able to further address potential issues associated with the sale of mortgages by TBs.

3.3 Secondary Market and Credit Score Transmission Process

Denote the secondary market (SM) investor i's consumption bundle by a vector $x^i \equiv (\omega_1^i - \tau z^i, \pi^i z^i + (1-\pi^i) d^l \delta p_2 H_1^G) \in \mathbb{R}_+^2$, where $\omega_1^i > 0$ indicates the SM investor's endowment of the numeraire good in period 1 (again, for simplicity, we assume $\omega_2^i = 0$). The term $\pi^i z^i + (1-\pi^i) d^l \delta p_2 H_1^G$ captures the SM investor's expected second period revenue from purchasing mortgages in the first period. Consistent with SB expectations, the investor anticipates a fraction π^i of the loans purchased being G-type loans with promised payment z^i, and a fraction $(1 - \pi^i)$ being B-loans with recovery $\delta p_2 H_1^G$. Of this recovery, the investor is entitled a fraction d^l, whereas the remaining foreclosure proceeds go to the SB. The secondary market investor i's optimization problem consists of choosing a promise z^i that maximizes the following profit function:

$$\Lambda^i(z^i) \equiv \omega_1^i - \tau z^i + \theta^i(\pi^i z^i + (1-\pi^i) d^l \delta p_2 H_1^G), \tag{4}$$

We assume that investors assign a smaller relative weight to period 1 consumption than lenders do, i.e., $\theta^l < \theta^i$.

In this section we have assumed that $\pi^{SB} = \pi^i < 1$, implying credit scores are based on hard information only and that they are transmitted to SM investors truthfully and accurately without misrepresentation or adverse selection, i.e., $CST^{SB} = CST^i$. In practice, the SM sale information production process can be characterized as the generation of a securities prospectus that accurately documents credit scores of loans included in the mortgage pool along with a scenario analysis that produces expected mortgage loan default rates and loss severities.

Greater patience attributable to SM investors can be interpreted as a measure of liquidity in the secondary market for non-prime mortgage loans.[24] Greater liquidity implies higher relative

[24] 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

prices paid for mortgage loans, which feeds back to result in a larger relative loan amount for the borrower. However, because some borrowers default on their mortgages, pooling occurs such that the mortgage proceeds at origination (implied mortgage rate) adjusts downward (upward).

3.4 Households

In each period $t = 1, 2$, consumers decide whether to rent (R_t) or buy owner-occupied housing (H_t), or possibly some combination of the two. If a consumer buys a house, he will borrow from either a TB or a SB in the amount of $q^{TB}\psi^{TB}$ or $q^{SB}\psi^{SB}$, respectively, where $\psi^l \geq 0$ denotes consumer's mortgage face value (i.e., promise to pay) in period 2 to lender $l = TB, SB$. Equilibrium existence requires an upper bound $B > 0$ on ψ^l:[25]

$$\psi^l \leq B \tag{5}$$

Thus, in period 1 a non-prime consumer has three possible choice: (1) Borrow in the traditional bank loan market to own, (2) Borrow in the shadow bank loan market to own, or (3) Rent (i.e., not borrow). We denote these possibilities by m^{TB}, m^{SB}, and $m^{\emptyset}$, respectively, and the set of consumer's "market choices" by $\mathbf{M} = \{m^{TB}, m^{SB}, m^{\emptyset}\}$. The consumer's market choice is consumer-type ($c(h) = G, B$) and market-type ($l \in \mathbf{L} \equiv \{TB, SB, \emptyset\}$) specific, and is denoted by $m^l_{c(h)} \equiv (c(h), l)$. 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, $l = TB, SB, \emptyset$, is:

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

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 $\omega^{SR}/p_1 H_1$ percent.

The second period budget constraint depends on two things. First, we assume that if a con-

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.

[25]Notice that in our examples below, once we compute the equilibrium for a specific set of parameter values, we choose the upper-bound B on the face value of the mortgage loan in such a way that this constraint is non-binding.

sumer 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 is like a long term contract that once signed is valid for two periods). However, a consumer that buys good R_1 can only consume it for one period (i.e., rental housing is a one-period contract). Second, as noted earlier, loans are subject to a *limited recourse mortgage* contract that stipulates that a borrower is allowed to consume his subsistence income, ω^{SR}, and no more than his subsistence income, if default occurs.[26] Accordingly, we write the second period budget constraint of a consumer with market choice m^l, $l = TB, SB, \emptyset$, as follows:

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

where ω_2^c denotes the period 2 endowment of a type $c = G, B$ consumer and is such that $\omega_2^G = \omega^+$ and $\omega_2^B = \omega^{SR}$. The term $p_2 H_1$ in the right hand side of the inequality (7) captures 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.[27] 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 he minimum subsistence income ω^{SR}.[28] There is no default if p_2, H_1, and ψ^l are such that $\omega^{SR} \leq \omega_2^c + p_2 H_1 - \psi^l$. Loan payment is (partially) enforced by the nature of the limited recourse loan in our model, where, as noted earlier, G-types use limited recourse as credible commitment (incentive compatability constraint) in order to increase consumption.

An important feature of our general equilibrium model is that the mortgage market structure is endogenous. In order to incorporate consumer market choices in our setting, we need to introduce the following notation. Define the consumer's market choice function by $\mu : \mathbf{A} \to \mathbf{D}$, where

[26] See 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.

[27] A 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.

[28] The maximum operator can be seen as an 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.

$\mathbf{D} = \{\iota(m^l_{c(h)}) : \iota(m^l_{c(h)}) = 1$ if consumer h chooses $m^l_{c(h)}$, and 0 otherwise$\}$. We require that a consumer can only choose one market in $\mathbf{M}$ (i.e., $\sum_{l=TB,SB,\emptyset} \iota(m^l_{c(h)}) = 1$). Also, denote the household h's consumption bundle by $x^h = (H^h_1, R^h_1, H^h_2, R^h_2) \in \mathbb{R}^4_+$. We 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 $h \to \mathbf{X}^h$ 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., $u^h(x^h, \mu^h(m))$.[29] We assume that the mapping $(h, x, \mu) \to u^h(x, \mu)$ 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).

Finally, the measure of those consumers that receive a good rating in mortgage market $l \in \{TB, SB\}$ is

$$\lambda(G\text{-}Rating^l) \equiv CST^{SB} \cdot \hat{\lambda}^l_G + (1 - CST^{SB}) \cdot \hat{\lambda}^l_B$$

where $\hat{\lambda}^l_G \equiv \lambda(\mathcal{H} : c(h) = G,\, \mu^h(m^l_G) = 1)$ and $\hat{\lambda}^l_B \equiv \lambda(\mathcal{H} : c(h) = B,\, \mu^h(m^l_B) = 1)$ are the measure of G-type and B-type consumers that attempt to borrow from lender $l \in \{TB, SB\}$. Notice that $\hat{\lambda}^l_G$ and $\hat{\lambda}^l_B$ are endogenously determined in our model, and therefore so is $\lambda(G\text{-}Rating^l)$.

4 Equilibrium and Mortgage Pricing

In this section we define the equilibrium of our economy with endogenously segmented mortgage markets. After showing that such an equilibrium exists, we examine some of the primary implications with respect to the price of mortgage debt.

[29] This 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 (see, e.g., Luque 2013).

4.1 Equilibrium Definition and the Existence Result

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 SB 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 $\pi \equiv (\pi^{TB}, \pi^{SB})$ consistent with the market aggregate choice function $\hat{\mu}$. We formally define this condition before providing the definition of equilibrium.

Let a continuous function $f : (\hat{\mu}(G,l), \hat{\mu}(B,l)) \to [0,1]^2$ 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 $g : (f(\hat{\mu}(G,l), \hat{\mu}(B,l)), CST^l) \to [0,1]$ be such that it mimics the belief expression (1), where instead of $(\hat{\pi}^l_G, \hat{\pi}^l_B)$ 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 $\hat{\mu}^l \in \mathbb{R}^{\mathbf{M}}$ is consistent with lender l's belief π^l if $\pi^l = g(f_G(\hat{\mu}(G,l), \hat{\mu}(B,l)), CST^l)$.

Definition 1: *Given the triplet $(CST^{TB}, CST^{SB}, CST^i)$, an equilibrium for this economy consists of a vector of market choices μ, prices $(p_1, p_2, q^{TB}, q^{SB}, \tau)$ and allocations $((x^i, z^i)_{i\in\mathcal{I}}$, $(x^l, \varphi^l, z^l)_{l\in\{\mathcal{TB},\mathcal{SB}\}}$, $(x^h, \psi^h)_{h\in\{\mathcal{G}\cup\mathcal{B}\}})$ 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) $\int_{\mathcal{H}} \psi^h(m^{TB}_{c(h)})\mu^h(m^{TB}_{c(h)})dh = \int_{\mathcal{TB}} \varphi^{TB} dl$

(MC.2) $\int_{\mathcal{H}} \psi^h(m^{SB}_{c(h)})\mu^h(m^{SB}_{c(h)})dh = \int_{\mathcal{SB}} \varphi^{SB} dl$

(MC.3) $\int_{\mathcal{SB}} z^l dl = \int_{\mathcal{I}} z^i di$

and, for $t = 1, 2$,

(MC.4) $\sum_{l\in\mathbf{L}} \int_{\mathcal{H}} R^h_t(m^l_{c(h)})\mu^h(m^l_{c(h)})dh + \int_{\mathcal{SB}\cup\mathcal{TB}} x^l_t dl + \int_{\mathcal{I}} x^i_t di = \int_{\mathbf{A}} \omega^a_t da$

(MC.5) $\sum_{l\in\mathbf{L}} \int_{\mathcal{H}} H^h_1(m^l_{c(h)})\mu^h(m^l_{c(h)})dh = \sum_{l\in\mathbf{L}} \int_{\mathcal{H}} H^h_2(m^l_{c(h)})\mu^h(m^l_{c(h)})dh = \bar{H}$

In the Appendix A.1, we show that an equilibrium, as defined in Definition 1, exists.

4.2 Remarks

Remark 1: We use the concept of pooling equilibrium to address borrower adverse selection in the shadow bank mortgage market. 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:

$$H_{SB}^{\min} \equiv \frac{\bar{H}(\omega^{SR} + \bar{L})}{2\omega^{SR} + \bar{L} + L^G} \tag{8}$$

where $\bar{L} \equiv \theta^i \delta \omega^{SR}/(1 - \theta^i \delta)$ is the maximum loan amount that a SB 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 G-type consumer would obtain from a SB when mortgage markets are segmented (using the SB's first order condition and G-type consumer's first period budget constraint). In the Appendix we show that $H_{SB}^{\min}$ rules out a separating equilibrium.[30] Also, in the Appendix, we identify the threshold $H_{TB}^{\min}$ that rules out a TB mortgage market specific for B-type consumers. Threshold $H^{\min} \equiv \max\{H_{TB}^{\min}, H_{SB}^{\min}\}$ 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. This result illustrates how housing regulations prevent the least well-endowed non-prime consumers from purchasing a house, implying the structural details underlying mortgage contract design and market organization consequently feed back to affect the rent versus own decision in our model.

Remark 2: In our model default risk is the result of the SB'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 SB 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 3: Our notion of equilibrium assumes that lenders and SM investors form beliefs

[30] A separating equilibrium can exist even when B-types always default in the second period, because the SB would get income $\delta p_2 H_1$ by foreclosing the house. This is profitable for the SB when the loan amount is small because by foreclosing the house the SB is able to seize the B-type borrower's house equity (ω^{SR}). Formally, the SB gets $\delta p_2 H_1$ 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 Sections 4 and 5 below.

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.[31]

Remark 4: Given the TB's capacity constraint and the SB's imperfect CST^{SB}, 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 and where the TB capacity constraint is binding. In this case, some G-type consumers are rationed into the SB mortgage loan market. If those G-types still prefer owning over renting and thus apply for a loan from a SB, 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 SBs – i.e., denied credit altogether – with no other option but to rent.

Remark 5: In equilibrium consumers that attempt to borrow in the shadow bank loan market, and receive a good rating, obtain a loan equal to $q^{SB}\psi^{SB} = \pi^{SB}\bar{\theta}/(1-\bar{\theta}(\pi^{SB}(1-\delta)+\delta))$ and promise to repay $\psi^{SB} = \omega_2^+ - \omega^{SR} + p_2 H_1^{G,SB}$, where $H_1^{G,SB}$ is the house size that a consumer with a SB 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 with good rating end up defaulting because their income is ω^{SR} (instead of ω_2^+), and therefore are only able to pay back $p_2 H_1^{G,SB}$ (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 shadow bank chooses a pooling discount price given its belief π^{SB}. In the next subsection, we elaborate on the determinants of mortgage pricing.

[31]Lenders 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.

4.3 Mortgage Discount Prices and Excess Premium

Risk-neutrality implies that the lender's first order condition determines the competitive mortgage prices q^l, $l = TB, SB$. In the Appendix we write these pricing conditions. Here we comment on the relative mortgage prices observed in each market, where determination of SB mortgage price involves a trade-off between borrower adverse selection and investor liquidity in the secondary market, specifically depending on parameters such as CST^{SB} and the SM sale distribution rate, d^l.

As a result, the SB finds it optimal to tack on a pooling rate premium to the base loan rate to account for borrower adverse selection risk. The greater the proportion of B-types applying for a mortgage loan in the SB market and the worse the classification accuracy of the CST^{SB}, the lower are the funding proceeds of a SB loan. Offsetting this downward price 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 SB market can be higher or lower than the implied morgage loan price in the TB market. See the pricing condition (22) in the Appendix for a mathematical expression of this trade-off in the conduit loan market.

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

$$EP \equiv (1/q^{SB}) - (1/q^{TB}) \tag{9}$$

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

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

5 Equilibrium Regimes

In this section we explain the mortgage market configuration that follows from our general equilibrium framework. We are particularly interested in examing credit scoring technology's influence on the flow of capital into non-prime lending markets vis-a-vis the shadow banking 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 equlibrium regime changes.

We make two assumptions before presenting the results. First, we constrain the traditional banking sector to meet the demands of some but not all good type consumers should that demand exist. That is, $\lambda_G \geq v(TB)$, where $v(TB)$ denotes the TB's lending capacity constraint.[32] The assumption of capacity constrained TBs then implies that, when traditional bank loans are the first choice among consumers, capacity-rationed G-type consumers turn to the shadow bank loan market if they still prefer owning over renting. In this case, a mass $\lambda_G - v(TB)$ of G-type consumers attempt to borrow from SBs, with the resulting endogenous proportion of G-types in that market as $\hat{\pi}_G^{SB} = (\lambda_G - v(TB))/(\lambda_G - v(TB) + \lambda_B)$. If, on the other hand, SBs are preferred to TBs by G-type consumers, and home owning with SB mortgage loan financing is preferred to renting, then $\hat{\pi}_G^{SB} = \lambda_G/(\lambda_G + \lambda_B)$.

Second, we focus on a more analytically tractable setting where owner-occupied housing (H) and rental housing (R) are perfect substitutes for consumers, and consider the following linear separable utility function:

$$u^h(R_1, H_1, R_2, H_2) = R_1 + \eta H_1 + \theta^h(R_2 + H_2), \tag{10}$$

where θ^h (with $\theta^h < \theta^l$) denotes the consumer's time preference parameter and $\eta > 1$ denotes a preference parameter which reflects that, all else equal, in the first period young households prefer to consume owner-occupied housing over rental housing (this can be possibly due to 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

[32] This assumption is motivated by additional constraints faced by traditional banks, such as the time constraint to originate loans that require soft information acquisition through relational interactions. Other considerations may also apply, such as internally imposed capital allocation constraints or the TBs' inability in the short run to raise external capital to finance new mortgages (due to an inelastic supply of low-cost deposits, for example). In any case, a capacity constraint on TBs is not required for us to characterize a competing role for TBs and SBs in mortgage lending. What would be lost is the second of three equilibrium regimes we characterize in this section.

from consumption of owner-occupied housing H_2 and the utility from consumption of rental housing R_2 are the same, however. To get simple closed form solutions, we assume throughout that $\omega^{SR} = 0.5$, $\omega_2^+ = 1$, $\lambda_G = 1.5$, and $v(TB) = 1$. The first two parameters respectively quantify the subsisdence rent endowment and the G-types' second period income, and the last two parameters are the respective measure of G-types and the TB capacity constraint.

Before examining the role of beliefs on mortgage market configurations, we first discuss the effect of the owner-occupied housing price on consumers' housing choices.

5.1 House Prices

We model the aggregate demand for owner-occupied housing in the *first* period and the aggregate supply of 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 primal infuence in the housing sector, whereby mortgage defaults occur in our model due to the imperfect screening of borrowers by type in the SB 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 to young households in the second period and move to rental housing, as the benefits to owning go away as the younger household transitions to older age.[35] In the first period young consumers will generally find it optimal to buy a house, provided that the credit scoring technology parameter π^{SB} 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-occuppied housing changes hands from old households in a previous cohort to young

[33] The 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 $p_1 = p_2 = p$.

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

[35] See Hochguertel and van Soest (2001) for 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.

households in a new cohort.[36]

5.2 Beliefs and Mortgage Market Configurations

This sub-section identifies three thresholds, π_0, π_1 and π_2, that correspond to the shadow bank's belief π^{SB}. We say that these thresholds delineate alternative lending regimes. Recall that π^{SB} measures the posterior probability that a consumer is a G-type given that the SB classifies the consumer as a G-type. This quantity depends on two factors: the credit scoring technology and the proportion of G- and B-types applying for loans from SBs. The three denoted thresholds determine different subprime 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 shadow banking market shuts down (is inactive) if belief π^{SB} is sufficiently small.* That is, there is a threshold π_0 that solves the following equation:
$$H_1^{G,SB}(\pi_0) = H^{\min} \tag{11}$$
such that when $\pi^{SB} < \pi_0$, shadow banking loans are so small that borrowers cannot afford to buy a house with size above $H^{\min}$. Here $H_1^{G,SB}(\pi_0)$ denotes G-type owner-occupied housing consumption when borrowing from a SB and belief π^{SB} equals π_0. The expression for $H_1^{G,SB}$ as a function of π^{SB}, as well as the equilibrium loan amounts, can be found in the Appendix.

2. *There is an active shadow banking mortgage market as long as G-type consumers prefer to borrow from SBs rather than rent in the first period.* When π^{SB} falls below a given threshold π_1, the implicit SB 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 SB to buy a house in the first period. Threshold π_1, at which indifference between buying a house with a conduit

[36] To 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).

loan and renting in both periods occurs, solves the following equation:[37]

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

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

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

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

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

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 SB loan and then renting (in a setting where only the SB 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 SB loans markets are active). Notice that $H_1^{G,TB}$ is not a function of π^{SB}, since, by assumption, TBs have a perfect screening technology and only lend to G-types. Further observe that when $\pi^{SB} > \pi_2$, G-type consumers prefer SB loans over TB loans even though SBs risk-price classification error (lending to some B-types). In this case, the proportion $\hat{\pi}_G^{SB}$ of G-type consumers that attempt to borrow from SBs improves, as now SB loans are the first best option for G-type consumers. In addition, when the SM mortgage distribution rate d^l increases, threshold π_2 decreases and the shadow bank loan market in effect 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 ($\theta^i > \theta^l$).

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

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

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

Proposition 2 (Mortgage market configurations)**:**

- Dominant TB Loan Market: *If* $\pi^{SB} < \max\{\pi_0, \pi_1\}$, *the shadow bank 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 SB Loan Market: *If* $\pi^{SB} > \pi_2$, *G-type consumers prefer the shadow bank mortgage market. A mass* $CST^{SB}\lambda_G + (1 - CST^{SB})\lambda_B$ *of consumers receive a good rating and are able to borrow at the shadow bank loan rate to buy a house. A mass* $\min[(1 - CST^{SB})\lambda_G, 1]$ *of G-type consumers who are incorrectly classified as B-types will borrow from their second best option, the traditional bank loan market. The rest of consumers will rent in both periods.*

- Coexisting TB-SB Loan Market: *When* $\pi^{SB} \in [\max\{\pi_0, \pi_1\}, \pi_2]$, *traditional banks lend to a mass* $v(TB)$ *of G-type consumers. The remaing pool of G- and B-types apply for a loan in the SB market. A mass* $(1 - CST^{SB})(\lambda_G - v(SB)) + CST^{SB}\lambda_B$ *of consumers are rejected by SBs 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 SB's credit scoring technology (CST^{SB}) deteriorate, π^{SB} decreases. There is, as a result, greater asymmetric information between consumers and SBs, and all else equal the SB 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 distribution rate d^l increase, all else equal, the active region for a SB loan market expands (as threshold values π_0, π_1 and π_2 decrease). This is because SB loans become less expensive, with greater loan proceeds, as SM investors are willing to pay more for mortgages.

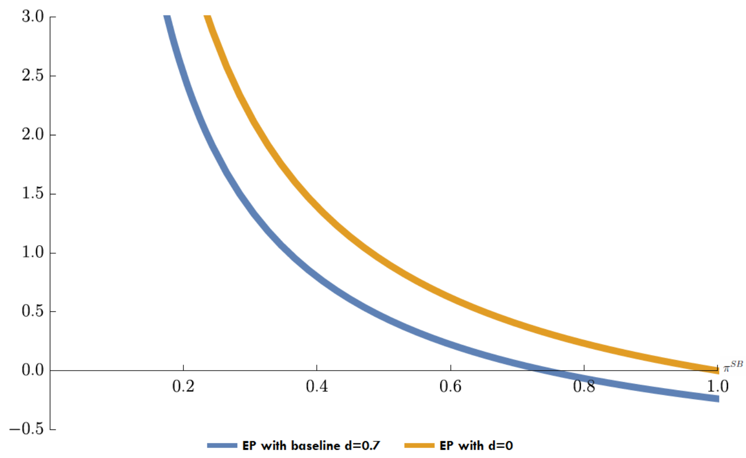


Figure 1: $EP \equiv (1/q^{SB}) - (1/q^{TB})$

In Figure 1 we illustrate how the excess premium (EP) changes as a function of the SB's belief, π^{SB}. Figure 1 and all following figures in this paper assume the following parameter values as a baseline: $d^l = 0.7$, $\theta^h = 0.4$, $\theta^l = 0.7$, $\theta^i = 0.9$, $\eta = 4$, $\delta = 0.5$, $\lambda_G = 1.5$, $\lambda_B = 1$ and $v(TB) = 1$. The equilibrium π-thresholds for these parameters are $\pi_0 = 0.16$, $\pi_1 = 0.18$ and $\pi_2 = 0.74$. In this figure we observe two EP-lines that vary inversely with π^{SB}. The line $d = 0$, which represents a change in the baseline parameter value $d = 0.7$, computes EP when SBs (counterfactually) do not distribute mortgages to investors. In this case, the G-type consumers always prefer TB loans over SB loans, due to the implied SB mortgage rate, $1/q^{SB}$, always exceeding $1/q^{TB}$ (so that $EP > 0$). The EP-line for the baseline $d = 0.7$ changes from positive to negative at $\pi^{SB} = \pi_2 \equiv 0.74$. At this point, the positive effect of SM investor liquidity (which decreases the implied SB mortgage rate) exactly offsets the negative adverse selection effect (which increases the implied SB loan rate), and the shadow bank is able to offer the same mortgage rate (and loan proceeds) as the traditional bank. When $\pi^{SB} > 0.74$, the SB's implied mortgage rate is lower than the TB rate, with SB loan proceeds exceeding TB loan proceeds, so G-types consumers prefer SB loans to TB loans in equilibrium. Notice that in this case that the TB mortgage loan market does not completely collapse, as misclassified G-types will approach TBs as the second-best option. All rejected G-types will obtain a TB mortgage unless the capacity constraint is met. For G-types that are rationed out of the TB market, there is no other choice but to rent.

6 The Credit Scoring Channel

6.1 Overview

In this section we highlight the credit scoring channel's role as a governor that controls the flow of capital into the non-prime mortgage market. As we discussed in Section 3, improvements in credit scoring model technology combined with increased (over-)confidence in the classification ability of the technology to result in upward revisions in CST^{SB} over time. 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 their beliefs as they realized that their credit scoring models were failing to work as expected.[38] This failure of the credit scoring models to accurately predict failure consequently led to wholesale downward revisions in both prior and posterior credit assessment probablities (CST^{SB} and π^{SB} in our model), resulting in the collapse of non-prime mortgage lending and house prices.

6.2 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^{SB}, can trigger changes in the equilibrium structure of the mortgage market. In particular, we will show how, starting from an inactive shadow banking region, incremental improvements in CST^{SB} can trigger the emergence of the shadow bank mortgage market. Then with further improvements in CST^{SB}, the shadow banking sector comes to dominate traditional bank lending as all G-type consumers migrate from the traditional bank to shadow bank market as their first choice for mortgage funding. This migration tipping point triggers 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.

Using the same parameter values considered in the previous section, Figure 2 illustrates the equilibrium aggregate mortgage quantities in the TB versus SB markets, respectively, for differ-

[38] See Figure 1 of Brunnermeier(2009) and the associated discussion.

ent equilibrium regimes identified above. Figure 3 shows resulting equilibrium owner-occupied house prices as a function of CST^{SB}. The thresholds for CST^{SB} follow from expression (1), the π-thresholds identified above and the corresponding proportion of G-type consumers in the SB loan market, where for this parameter set it follows that $CST^{SB} < 0.31$ establishes Regime 1, $0.31 \leq CST^{SB} < 0.85$ establishes Regime 2, and $CST^{SB} \geq 0.85$ establishes Regime 3.[39]

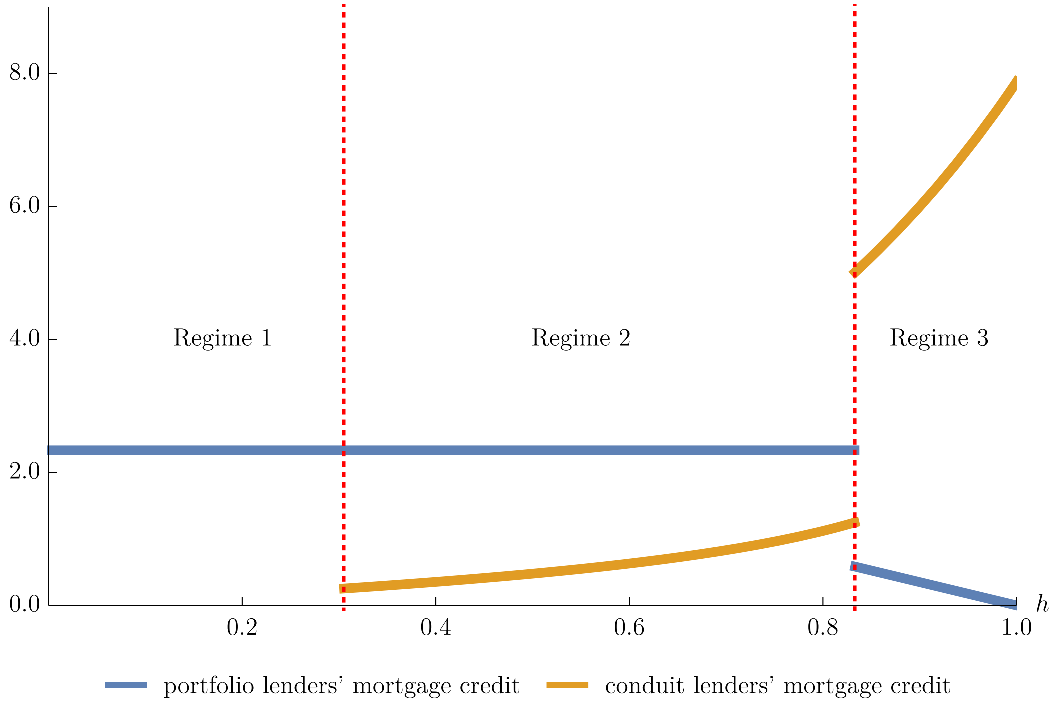


Figure 2: This figure plots the total amount of TB and SB mortgage credit as a function of $CST^{SB} \equiv h$. The thresholds for CST^{SB} follow from expression (1), the π-thresholds identified above and the corresponding proportion of G-type consumers in the conduit loan market ($\hat{\pi}_G^{SB}= 0.33$ if $CST^{SB}\leq CST_2^{SB}\equiv 0.85$ and $\hat{\pi}_G^{SB}= 0.6$ otherwise). In particular, thresholds for CST are $CST_0^{SB}= 0.28$, $CST_1^{SB}= 0.31$, and $CST_2^{SB}= 0.85$.

In Regime 1 the SB market is dormant, caused by borrower type classification errors of sufficient size so that high pooled mortgage loan rates result (recall Figure 1 in which EP is large when π^{SB} is small). The high implied SB mortgage rates are such that G-type households which are rationed out of the TB market prefer to rent rather than own. With only a limited number of G-type households gaining access to morgage financing and an inactive SB loan market, equilibrium house prices are relatively low (Figure 3).

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

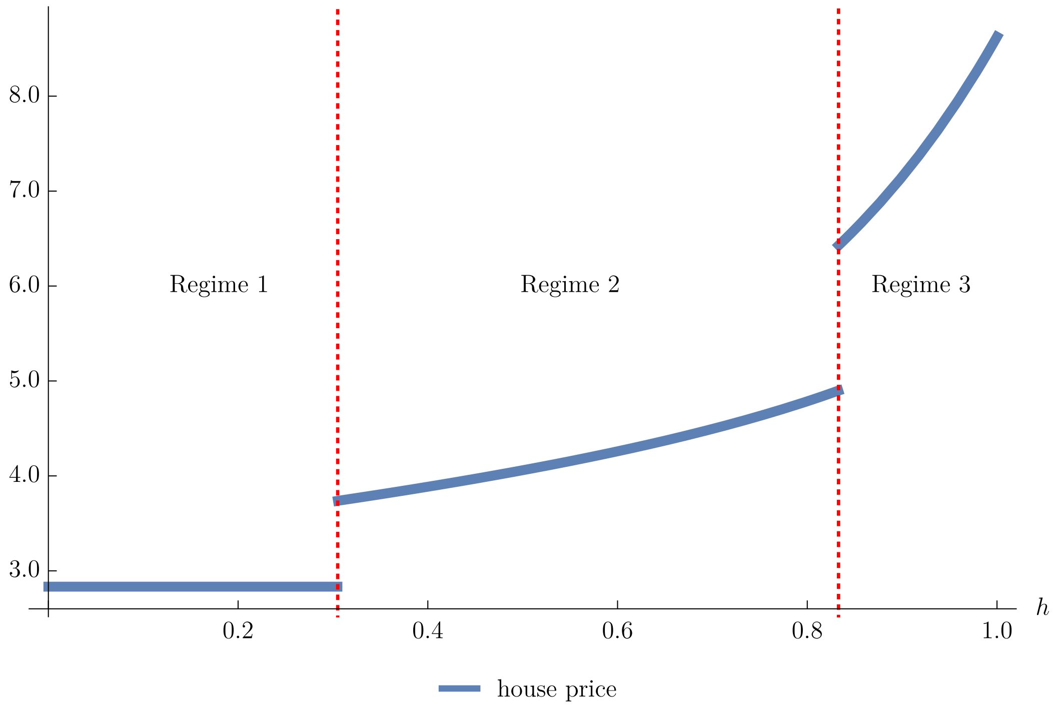


Figure 3: This figure illustrates the equilibrium house price p as a function of $CST_G^{SB} \equiv h$. Thresholds for CST are $CST_0^{SB} = 0.28$, $CST_1^{SB} = 0.31$and $CST_2^{SB} = 0.85$.

In Regime 2 the SB mortgage market emerges because beliefs regarding the precision of the credit scoring technology have improved sufficiently. As a result, SBs 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^{SB} 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 SB market relative to the TB market due to a relatively high 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 SB market. House prices increase as a function of CST^{SB} 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 to the SB market occurs. Because of this, there is a large discrete jump in the proportion of G-types applying for a SB loan, an effect that by itself causes a discrete downward adjustment in the pooled mortgage loan rate. Aggregate mortgage loan quantity in the SB 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 SB mortgage market. House prices continue to increase as a function of CST^{SB} as aggregate 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 MBS 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 SB 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*).

We note that what can go boom can also go bust. As early term mortgage loan defaults spiked unexpectedly starting in the second half of 2006, confidence in credit scoring technology was initially shaken and then shattered by its failure to predict failure.[40] The credit scoring channel in our model, where negative shocks to CST^{SB} imply moving from region 3 to region 2, and then to region 1, illustrates the resulting substantial declines in SB mortgage lending volume, house prices and home ownership rates.

Finally, as documented by Buchak, Matvos, Piskorski and Serue (2017), we note that in more recent years the shadow bank loan market has reemerged to coexist with the traditional bank loan market. Hard information-based credit scoring technology has thus apparently recovered to some extent, but the secondary market investor clientale has changed to a more dedicated and sophisticated set of secondary market investors consisting of large banks, insurance companies

[40] In 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 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.

and specialty finance firms such as mortgage REITs. Of the next-generation SB loan originators, according to Buchak et al., fintech firms seem to be gaining market share. Interestingly, these firms focus on offering complementary on-line financial services and use sophisticated data mining and online marketing research techniques to learn about consumer tastes and preferences in an attempt to enhance profitability.

6.3 Lax Screening, Lending Standards, and Loan Acceptance Rates

Since the seminal works of Keys, Mukherjee, Seru, and Vig (2010), Purnanandam (2011) and Rajan, Seru and Vig (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 as a result of credit quality misclassification error.

Using our baseline 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 SBs is consistent with findings in this literature. And more importantly, as CST^{SB} increases in our model (based on inputting hard information only), in moving from a dormant SB 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 SB-originated non-prime mortgage loans outstanding relative total outstanding non-prime mortgages, becomes increasingly prevalent (see Figure 2). This increasing prevalance 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^{SB} do in fact reflect increasing precision in classification accuracy, lax screening does not necessarily imply an incremental loss in economic efficiency. On the contrary, to the extent that overconfidence in CST^{SB} 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, Igan and Laeven (2015) and Mian and Sufi (2014, pp. 76-79).[41] In our model, changes in loan acceptance rates can be caused simply by changes in the CST^{SB} applied in the shadow banking 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^{SB} due to the relatively low proportion of G-types in the SB loan application pool (the proportion of G-types applying for a SB loan is less than 0.50). In Regime 3, where all G-types have migrated to the SB loan market as their first funding choice, which dramatically increases the proportion of G-types in the SB loan application pool (the proportion of G-types applying for a SB loan now exceeds 0.50), the loan application acceptance rate increases in CST^{SB} 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 improvements in the classification precision of the applied CST^{SB}.

[41]The 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".

7 The Credit Scoring Transmission Channel: Soft Information Acquisition and Adverse Selection in the Secondary Loan Market

Previously we assumed that shadow banks and secondary market investors relied on the same (hard) information in their assessment of mortgage loan credit quality, with accurate and credible transmission of that information from the SB 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 SBs and used to adversely select against secondary market investors.

In the model, adverse selection in the secondary loan investor market occurs when a SB acquires soft information, in secret and at a cost, over and above the hard information that is already available to the SB and secondary market investors. With a fixed SM distribution rate, $d^{SB} < 1$, the additional soft information allows the SB 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^{SB} may include soft information on top of hard information, we write $CST^{SB}_{Soft} \equiv h + f(s)$, 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^{SB}_{Hard} 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 SBs. 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 (22) in the Appendix), where

$$q^{SB} = \frac{\pi_H^{SB}(d^l\theta^i + (1-d^l)\theta^l)}{1-\delta(1-\pi_H^{SB})(d^l\theta^i + (1-d^l)\theta^l)} \tag{14}$$

with π_H^{SB} denoting the posterior Bayesian probability using *hard* credit information. The mortgage loan price condition is empirically supported by Rajan, Seru and Vig (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 SB 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 7.2 for further details on this constraint. The constraints imposed on SB 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.

7.1 The Mechanics of Soft Information Acquisiton and Adverse Selection in the Secondary Loan Market

In this subsection we specifically examine what it means for the SB to acquire soft information and use that information to select against the secondary market.[42] The *hard* CST^{SB} 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 quanity of soft information that is optimally acquired at a cost.

[42]In our model, the SB 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.

From the SB's perspective, conditional on CST^{SB}_{Hard} 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 SB'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 SB when only hard information is considered was determined previously to be:

$$N^{SB}_{Hard} = CST^{SB}_{Hard}\lambda_G + (1 - CST^{SB}_{Hard})\lambda_B \tag{15}$$

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 SB 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 $\Pr^{SB}(\text{rating=G}|\text{G}) = \Pr^{SB}(\text{rating=B}|\text{B})$ throughout.

The number of loans originated by the SB 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^{SB}, 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.[43] Importantly, reconfirmed cherries are G-types with higher

[43]That is, there will exist a unique critical d^* in $[0, 1)$ such that, for d^{SB} less than or equal to critical d^*, only

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 SB will want to originate all loans which it rates as G-type (reconfirmed cherries as well as upgraded loans). The SB 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 SB with soft information acquisition in Regime 3 is:

$$N^{SB}_{Soft} = \underbrace{\Big(CST^{SB}_{Soft}\lambda_G + (1 - CST^{SB}_{Soft})\lambda_B\Big)}_{\text{measue of loans with ``}rating\ soft{=}G\text{''}} + \underbrace{\Big((1 - CST^{SB}_{Hard})\lambda_B - (1 - CST^{SB}_{Soft})\lambda_B\Big)}_{\text{lemons}} \quad (16)$$

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 $(1 - CST^{SB}_{Hard}) - (1 - CST^{SB}_{Soft}) > 0$. These are the respective probabilities that a SB 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 SB, 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 SB retains all of the remaining reconfirmed cherries as well as the upgraded loans. Plugging the above expressions into equation (18), we obtain a expression for the endogenous SM distribution rate:

$$\Delta(s) = \frac{\Theta^{SB}_{Hard}}{N^{SB}_{Soft}} \quad (17)$$

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

$$\Theta^{SB}_{Hard} = d^{SB} N^{SB}_{Hard} \quad (18)$$

where d^{SB} is the exogenously specified distribution rate used in that regime and N^{SB}_{Hard} is the total number of loans originated by SBs when loan underwriting decisions are made based on

lemons are sold into the secondary market. When this occurs, the SB will originate only enough lemons to satisfy the sales quantity condition.

hard information only. Below, in the optimization problem for the SB, we shall assume that the SB adheres to requirement (17).

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

Remark 6: *With soft information acquisition, the increase in the total number of originated loans is $f(s)\lambda_G$.*

Remark 7: *With soft information acquisition, the number of upgraded and downgraded loans are $f(s)\lambda_G$ and $f(s)\lambda_B$, respectively.*

Remark 8: *When the SB has no quantity contstraint of its own, the actual distribution rate of loans sold into the secondary market depends on the amount of soft informationa acquired by the SB, and is decreasing in s.*

7.2 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 shadow bank'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 SB 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 SBs 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 SB mortgage price, q^{SB}, and the total number of mortgages distributed, Θ^{SB}_{Hard}, as indicated by expressions (14) and (17).

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

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

subject to the price and quantity restrictions (14) and (17), and where

$$\pi_S^{SB} \equiv \frac{CST^{SB}_{Soft}\hat{\pi}^l_G}{CST^{SBL}_{Soft}\hat{\pi}^l_G + (1-CST^{SB}_{Soft})(1-\hat{\pi}^l_G)} \tag{19}$$

denotes the SB'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 SB 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.[44]

Figure 4 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 propogated 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.

7.3 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 SB due to the fact that lemons are sold first and not retained by the SB.

Define the SM investor's *ex ante* (expected) portfolio quality to be $\pi^i_{\text{ex-ante}} \equiv \Pr{}_{Hard}[G|\text{rating=G}]$ 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,

[44]Given the limited recourse nature of the subprime mortgage contract, $\psi^{CL} = \omega_2^G - \omega^{SR} + p_2 H_1$. Thus, consumers get the same loan amount $q^{CL}\psi^{CL}$ than in the no-soft information setting; however, as shown previously, there are more loans originated under soft information. Since the SB's mortgage discount price q^{SB} doesn't change, market clearing for the SB loan market follows by accommodating φ^{SB} 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.

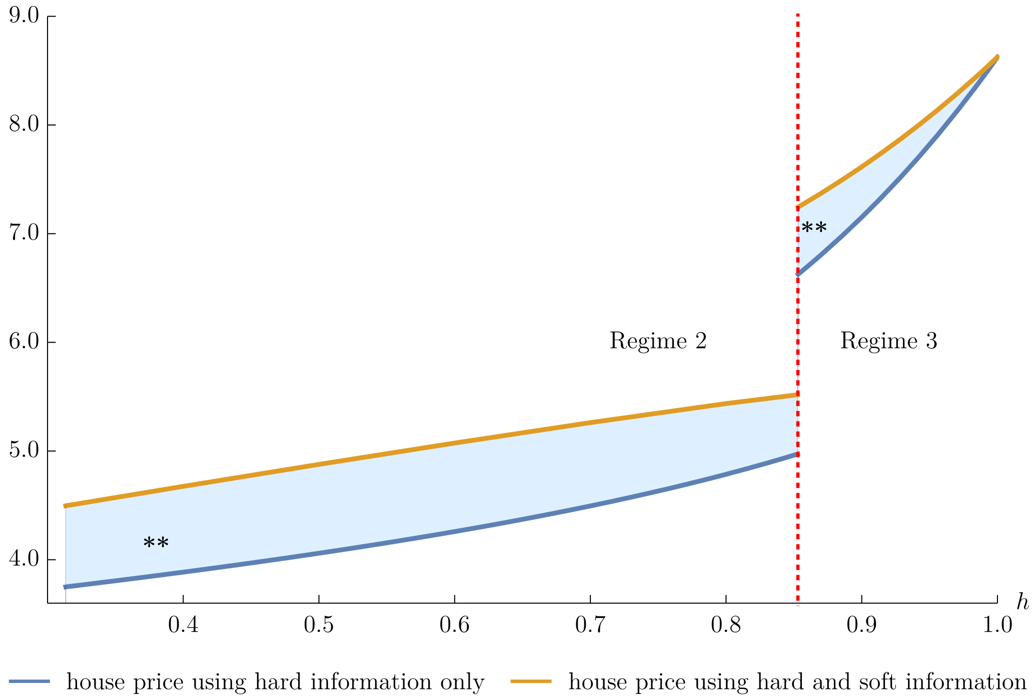


Figure 4: This figure portraits the price level, with and without soft information, as a function of $CST^{CL}_{Hard} = h$, The light-blue**-region captures the price increase due to a higher number of loans downgraded and sold as lemons with soft information acquisition.

ρ^L, as $\rho^L = f(s)\lambda_B/N^{SB}_{Hard}$, where $\rho^C = 1 - \rho^L$. With this, the credit quality of the portfolio actually sold into the SM is expressed as follows:

$$\pi^i_{\text{ex-post}} \equiv \rho^C \Pr{}_{Soft}[G|\text{rating=G}] + \rho^L \Pr{}_{Soft}[G|\text{rating=B}], \tag{20}$$

where $\Pr{}_{Soft}[G|\text{rating=B}] = \pi^{SB}_S$ (as given by (19)) and

$$\Pr{}_{Soft}[G|\text{rating=B}] = \frac{(1 - CST^{SB}_{Soft})\hat{\pi}^l_G}{(1 - CST^{SB}_{Soft})\hat{\pi}^l_G + CST^{SB}_{Soft}(1 - \hat{\pi}^l_G)}.$$

Given these relations, it is straightforward to show that, as CST^{SB}_{Soft} 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 5 we plot the SM investor's portfolio quality (as stated in equation (20)) 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.

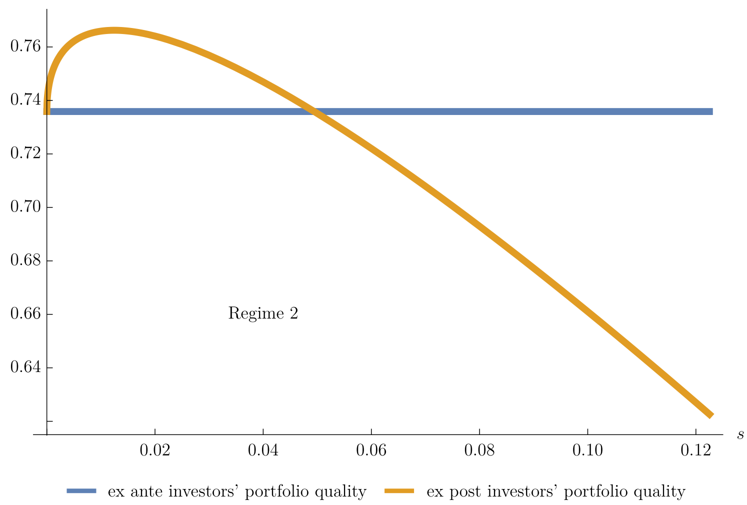


Figure 5: This figure portraits the ex-ante and ex-post investor's portfolio quality as a funciton of soft information.

Notice that Figure 5 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 SB'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 SBs to acquire soft information while the economy remains in Regime 2, consistent with the findings of Keys, Mukherjee, Seru, and Vig (2010) and Dell' Ariccia, Igan and Laeven (2012). This result is intuitive, since a higher distribution rate, d^{SB}, results in relatively fewer retained loans by the SB, 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 SB 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^{SB} lowers soft information acquisition, making SM investors better off than what they expected. Investors only become worse off than expected when d^{SB} is so high that it triggers a transition from Regime 2 to Regime 3.

8 Concluding Remarks: Model Summary and Borrower Income Misrepresentation

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. shadow bank lenders) helps to highlight a fundamental trade-off between access to soft information versus liquidity provided 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. Another important component of our theory is the limited recourse nature of the non-prime mortgages, which functions as an incentive compatability constraint for better credit quality borrowers. This feature helps to ensure 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 shadow banking sector is either inactive, in direct competition with traditional bank lenders, or is dominant.

Credit scoring technology, operating through the shadow banking 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 shadow banks can surrepticiously acquire soft information regarding loan credit quality, and use that information to select against the secondary market. In this version 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. Novel results are presented, showing 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 determined quantity of soft information acquired by the originating lender, consistent with the "lax screening" findings of Keys, Mukherjee, Seru, and Vig (2010), Purnanandam (2011), Rajan, Seru and Vig (2015) and others.

An additional important factor to consider is the misrepresentation of hard information used to make *yes-no* underwriting decisions, and to assess loan credit quality for loan pricing purposes. A number of empirical studies cited earlier have now identified misrepresentation as a pervasive distortion that affected the flow of mortgage debt and, in certain cases, the determination of local house prices (see, in particular, Griffin and Maturana (2016)). Although misrepresentation of loan application information has been shown to take several forms, arguably the clearest and most compelling evidence to date derives from income misrepresentation, particularly as it concerns the use of low- and no-documentation loans (Ambrose et al. 2016, Mian and Sufi 2017). Income misrepresentation in the context of our model is straightforward to accommodate, where misrepresentation takes the form of miscoded hard information that generates too many *yes* lending decisions in relation to the fundamental proportion of good types in the defined population, with mortgage loan credit spreads that are too low relative to the risks. Such practices, assuming they go undetected, serve to fuel mortgage and house price booms, as too many consumers receive loans and own a house in a non-prime neighborhood.

An important uresolved question in the literature is whether income misrepresentation primarily originated from the borrower or (shadow bank) lender, and, if it was primarily the former, whether the lender tacitly participated in the scheme, knowing that the proportion of good types in the population was less than that implied by loan acceptance rates.[45] Our model makes this distinction between posited and actual proportions of accept decisions clear, where the shadow bank originating mortgage loans can observe acceptance rates in relation to the *a priori* assessed credit quality of the population as a whole. The secondary market investor is not in such an advantaged position, and thus is vulnerable to distortionary loan origination practices of this type.

In addition to not observing the actual acceptance rate in relation to fundamental population proportions, the secondary market investor might be misled as to the true credit quality of the population given that household stated incomes are systematically exaggerated. We note that our model could be extended to accommodate such a setting, with the informed originating lender trading off profits from increases in lending volume with costs associated with losses on retained loans (if any) as well as possible future losses associated with reputation and legal challenges. Finally, we note that downward revisions in population incomes in the face of evidence that such incomes were exaggerated (as shown by Mian and Sufi 2017) can be captured in our model structure through sharp downward adjustments in posterior portfolio quality measures, which results in sharp declines in house prices, as shadow bank/secondary market non-prime lending contracts quickly in response.

[45] Piskorski, Seru, and Witkin (2015) present evidence of the misrepresentation of hard information by security issuers, and Ambrose et al. (2016) as well as Piskorski et al. show that originating lenders price-compensate for misrepresentation while secondary market investors do not.

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] Agarwal, S., B.W. Ambrose, S. Chomsisengphet, A.B. Sanders (2012), "Thy Neighbor's Mortgage: Does Living in a Subprime Neighborhood Affect One's Probability of Default?", *Real Estate Econ* 40, 1-22.

[3] Agarwal, S., Y. Chang, and A. Yavas (2012), "Adverse Selection in Mortgage Securitization", *J Financ Econ* 105, 640-660.

[4] Agarwal, S., G. Amromin, I. Ben-David , S. Chomsisengphet, and D. Evanoff (2011), "The Role of Securitization in Mortgage Renegotiation", *J Finan Econ* 102, 559-578.

[5] Ambrose, B., J. Conklin, and J. Yoshida (2016), "Credit Rationing, Income Exaggeration, and Adverse Selection in the Mortgage Market", *J Finan* 71, 2637-2686.

[6] Arslan, Y., B. Guler and T. Taskin (2015), "Joint Dynamics of House Prices and Foreclosures", *J Money, Credit and Bank* (forthcoming).

[7] Ashcraft, A., T. Adrian, H. Boesky, and Z. Pozsar (2012), "Shadow Banking", Federal Reserve Bank of New York, Staff Report No. 458.

[8] Ashcraft, A. and T. Schuermann (2008), "Understanding the Securitization of Subprime Mortgage Credit", Federal Reserve Bank of New York Staff Reports, no. 318.

[9] Ben-David, I. (2011), "Financial Constraints and Inflated Home Prices during the Real-Estate Boom", *Amer Econ J: Applied Econ* 3, 55–78.

[10] 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.

[11] Bernake, B. (2004), "The Great Moderation", Remarks by Governor Ben S. Bernanke at the meetings of the Eastern Economic Association, Washington, DC.

[12] Blanchard, O. and J. Simon (2001), “The Long and Large Decline in U.S. Output Volatility”, *Brookings Papers on Economic Activity* 1, 135-174.

[13] Bolton, P., X. Freixas, L. Gambacorta, P. E. Mistrulli (2016), “Relationship and Transaction Lending in a Crisis”, *Rev Financ Stud* 29, 2643-2676.

[14] Bottazzi, J.M., J. Luque, and M. Pascoa (2012), “Securities Market Theory: Possession, Repo and Rehypothecation”, *J Econ Theory* 147, 477-1500.

[15] Brueckner, J. (2000), “Mortgage Default with Asymmetric Information”, *J Real Estate Finan Econ* 20, 251-274.

[16] Brueckner, J., P. Calem and L. Nakamura (2012), “Subprime Mortgages and the Housing Bubble”, *J Urban Econ* 71, 230-243.

[17] Brunnermeier, M. K. (2009) “Deciphering the Liquidity and Credit Crunch 2007-2008”, *J Econ Perspect* 23, 23, 77-100.

[18] Buchak, G., G. Matvos, T. Piskorski, and A. Seru (2017), “Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks”, NBER Working Paper No. 23288.

[19] 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.

[20] 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.

[21] Chatterjee, S., D. Corbae, and V. Rios-Rull (2011), “A Theory of Credit Scoring and Competitive Pricing of Default Risk”, mimeo.

[22] 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.

[23] Corbae, D. and E. Quintin (2015), “Leverage and the Foreclosure Crisis”, *J Polit Econ* 123, 1-65.

[24] Curtis, Q. (2014), "State Foreclosure Laws and Mortgage Origination in the Subprime", *J Real Estate Fin and Econ* 49, 303-328.

[25] Davila, E. (2015), "Using Elasticities to Derive Optimal Bankruptcy Exemptions", mimeo.

[26] 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.

[27] Downing, C., D. Jaffee, N. Wallace (2009), "Is the Market for Mortgage-Backed Securities a Market for Lemons?", *Rev Financ Stud* 22, 2457-2494.

[28] Drozd, L. and Serrano-Pardial (2017), "Modeling the Revolving Revolution: The Debt Collection Channel", *Amer Econ Rev* 107, 897-930.

[29] Duraton, G., V. Henderson and W. Strange, *Handbook of Regional and Urban Economics,* Elsevier (2015).

[30] Einav, L., M. Jenkins, and J. Levin (2013), "The Impact of Credit Scoring on Consumer Lending", *RAND J Econ* 44, 249–274.

[31] Faias, M. and J. Luque (2016), "Endogenous Formation of Security Exchanges", *Econ Theory*, forthcoming.

[32] Fishman, M. J. and J. A. Parker (2015), "Valuation, Adverse Selection, and Market Collapses", *Rev Financ Stud* 28, 2575-2607.

[33] Frankel, D. and Y. Jin (2015), "Securitization and Lending Competition", *Rev Econ Stud* 82, 1383-1408.

[34] Gan, J. and T. J. Riddiough (2008), "Monopoly and Information Advantage in the Residential Mortgage Market", *Rev Financ Stud* 21, 2677-2703.

[35] Ghent, A. and M. Kudlyak (2011), "Recourse and Residential Mortgage Default: Evidence from US States", *Rev Financ Stud* 24, 3139-3186.

[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 (2016a), "Did Dubious Mortgage Origination Practices Distort House Prices?" *Rev Financ Stud* 29, 1671-1708.

[38] Griffin, J. and G. Maturana (2016b), "Who Facilitated Misreporting in Securitized Loans?", *Rev Financ Stud* 29, 384-419.

[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] Karlan, D. and J. Zinman (2009), "Observing Unobservable: Identifying Information Asymmetries With a Consumer Credit Field Experiment", *Econometrica* 77, 1993–2008.

[44] 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.

[45] Krainer, J., and E. Laderman (2014), "Mortgage Loan Securitization and Relative Loan Performance", *J Financ Serv Research* 45, 39-66.

[46] Luque, J. (2013), "Heterogeneous Tiebout Communities with Private Production and Anonymous Crowding", *Reg Sci Urban Econ* 43, 117-123.

[47] Makarov, I. and G. Plantin (2013), "Equilibrium Subprime Lending", *J Financ* 68, 849-879.

[48] Mian, A. and A. Amir (2017), "Fraudulent Income Overstatement on Mortgage Applications During the Credit Expansion of 2002 to 2005", *Rev Financ Stud* 30, 1831-1864.

[49] Mian A. and A. Sufi (2014a), "What Explains the 2007-2009 Drop in Employment", *Econometrica* 82, 2197-2223.

[50] Mian, A. and A. Sufi (2014b), House of Debt, *The University of Chicago Press.*

[51] Mian, A. and A. Sufi (2011), "House Prices, Home Equity Borrowing, and the US Household Leverage Crisis", *Amer Econ Rev* 101, 2132-56.

[52] 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.

[53] Miller, S. (2015), "Information and Default in Consumer Credit Markets: Evidence From a Natural Experiment", *J Financ Intermed* 24, 45-70.

[54] 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.

[55] Tchistyi, A. and T. Piskorski (2011), "Stochastic House Appreciation and Optimal Mortgage Lending", *Rev Financ Stud* 24, 1407-1446.

[56] Pence, K. M. (2006), "Foreclosing on Opportunity: State Laws and Mortgage Credit", *Rev Econ Stat* 88, 177-182.

[57] Poblete-Cazenave, R. and J.P. Torres-Martinez (2013), "Equilibrium with Limited-recourse Collateralized Loans", *Econ Theory* 53, 181-211.

[58] Piskorski, T., A. Seru, and J. Wikin (2015), "Asset Quality Misrepresentation by Financial Intermediaries: Evidence from the RMBS Market", *J Financ* 70, 2635-2678.

[59] Piskorski, T. and A. Tchistyi (2011), "Stochastic House Appreciation and Optimal Mortgage Lending", *Rev Financ Stud* 24, 1407-1446.

[60] Purnanandam, A.K. (2011), "Originate-to-distribute Model and the Sub-prime Mortgage Crisis", *Rev Financ Stud* 24, 1881-1915.

[61] 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.

[62] Rajan, R. (2010), “Fault Lines”, Princeton University Press, Princeton, NJ.

[63] Rao, J. and G. Walsh (2009), “Foreclosing a Dream: State Laws Deprive Homeowners of Basic Protection”, National Consumer Law Center, Inc.

[64] Sato, Y. (2014), “Opacity in Financial Markets”, *Rev Financ Stud* 27, 3502-3546.

[65] Seghir, K. (2006), “Overlapping Generations Model with Incomplete Markets: The Numeraire Case II”, *Annales d’Economie et de Statistique* 81, 113-139.

[66] Stein, J.C. (2002), “Information Production and Capital Allocation: Decentralized Versus Hierarchical Firms,” *J Financ* 57, 1891-1921.

[67] Stroebel, J. (2016), “Asymmetric Information about Collateral Values”, *J Financ*, 71(2), June 2016

[68] 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.

[69] Van Nieuwerburgh, S. and L. Veldkamp (2012), “Information Acquisition and Under-Diversification”, *Rev Econ Stud* 77, 779-805.

A Online Appendix

In this Online Appendix we show that an equilibrium, in the sense of Definition 1, exists; work out the minimum house thresholds that prevent a mortgage market for B-type consumers; show the main equilibrium closed form solutions that were used in our numerical simulations; 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.

A.1 Equilibrium Existence

Proving existence of equilibrium, in the sense of Definition 1, is not straightforward. The sizes of the TB and SB mortgage markets are endogenous, as they depend on the consumers' preferred mortgage market choices. In addition, 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.

Our approach is as follows. We construct a generalized game and show that there is a mixed strategies equilibrium. Then 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.

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.[46]

[46] See Luque (2013) for a similar approach in a local public goods non-atomic economy, and Luque (2014) for a review of different approaches to the presence of equilibrium in a continuum of agents framework.

*The generalized game***:** In the generalized game a player a chooses his strategy $\varkappa^a$ parameterized by the other players' strategies $\bar{\varkappa}^{-a}$. For our economy this game is played by the consumers, 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 (Non-convex generalized game with given market choices): Consumer h chooses his most preferred consumption for a given mortgage market choice $m^l_{c(h)} \equiv (c(h), l)$, i.e., taking $\bar{\mu}^h(m^l_{c(h)}) = 1$ as given. The consumer h's consumption and loan demand when market choice is $m^l_{c(h)}$ is given by

$$(x^h(m^l_{c(h)}), \psi^h(m^l_{c(h)})) \in \arg\max\{\ u^{h\theta}\left(\cdot, \bar{\mu}^h(m^l_{c(h)})\right) :$$
$$\bar{p}_1 H^h_1(m^l_{c(h)}) + R^h_1(m^l_{c(h)}) \leq \bar{q}\psi^h(m^l_{c(h)}) + \omega^{SR},\ \psi^h(m^l_{c(h)}) \leq B,$$
$$\bar{p}_2 H^h_2(m^l_{c(h)}) + R^h_2(m^l_{c(h)}) \leq \max\{\omega^{SR}, \omega^c_2 + \bar{p}_2 H^h_1(m^l_{c(h)}) - \psi^h(m^l_{c(h)})\}\}$$

Observe that the choice variables in the constrained optimization problem should all be multiplied by $\bar{\mu}^h(m^l_{c(h)})$, but we chose to omit it since we are already assuming that $\bar{\mu}^h(m^l_{c(h)}) = 1$ (the consumer is evaluating his utility at specific market choice $m^l_{c(h)}$) - e.g., when writing $H^h_1(m^l_{c(h)})$ we mean $H^h_1(m^l_{c(h)})\bar{\mu}^h(c(h), l)$ with $\bar{\mu}^h(c(h), l) = 1$.

Let us show that $(x^h(m^l_{c(h)}), \psi^h(m^l_{c(h)}))$ has nonempty compact values and is continuous. First, notice that $h \to (x^h(m^l_{c(h)}), \psi^h(m^l_{c(h)}))$ has a measurable graph (see Hildenbrand (1974, p. 59, Proposition 1.b)). Non-emptiness follows from the positive endowment assumption. Compactness follows because $H^h(m^l_{c(h)}) \leq \bar{H} < \infty$, $R^h(m^l_{c(h)}) \leq \int_{\mathbf{A}} \omega^a da < \infty$, and $\psi^h(m^l_{c(h)}) \leq B$ if prices p_1 and p_2 are uniformly bounded away from 0 (i.e., p_1,$p_2 \geq \alpha$, $\alpha > 0$).[47] Continuity of $(x^h(m^l_{c(h)}), \psi^h(m^l_{c(h)}))$ follows if consumer's demand is both upper and lower hemi-continous. Since the consumer's consumption set and utility function are both continuos in (x, ψ) and endowments are desirable, we can apply Berge's Maximum theorem to show that $(x^h(m^l_{c(h)}), \psi^h(m^l_{c(h)}))$ is upper hemi-continuous. Next, we prove lower hemi-continuity of

[47]One can show that prices are indeed positive with a strictly monotonic utility. The argument is standard and thus ommitted for the sake of brevety (one should consider a sequence of truncated generalized games by relaxing α and apply 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).

$(x^h(m^l_{c(h)}), \psi^h(m^l_{c(h)}))$. Denote by $B^h = \{(x^h, \psi^h) : p_1 H_1^h(m^l_{c(h)}) + R_1^h(m^l_{c(h)}) \leq q\psi^h(m^l_{c(h)}) + \omega^{SR}, \psi^h(m^l_{c(h)}) \leq B,\ p_2 H_2^h(m^l_{c(h)}) + R_2^h(m^l_{c(h)}) \leq \max\{\omega^{SR}, \omega_2^c + p_2 H_1^h(m^l_{c(h)}) - \psi^h(m^l_{c(h)})\}\}$ the set of consumer h's consumption and borrowing amounts that are budget feasible.

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

Proof: Fix $\bar{\mu}^h(m^l_{c(h)}) = 1$ and consider consumer h's correspondence $\dot{B}^h$ that associates to each vector (p_1, p_2, q) the collection of plans $(x^h, \psi^h, \bar{\mu}^h(m^l_{c(h)})) \in \mathbf{X}^h$ that satisfies consumer's budget constraints of 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 $\overline{\dot{B}^h}$. We affirm that $\overline{\dot{B}^h} = B^h$. Since for any (p_1, p_2, q) we have $\overline{\dot{B}^h}(p_1, p_2, q) \subset B^h(p_1, p_2, q)$, it is sufficient to show that $B^h(p_1, p_2, q) \subset \overline{\dot{B}^h}(p_1, p_2, q)$.

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 want to 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 $\delta_2 = \delta_1(\omega^{SR} + p_2 H_1^h)/(\psi^h + p_2 H_1^h)$ (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 show that a consumer can buy $(1 - \delta_2)x_2^h$ after deciding whether to strategically default or not. This follows by (C2). To see this, notice that the new resources that become available in the second period are $\max\{\delta_2\omega^{SR}, \delta_2\omega_2^c + p_2\delta_2 H_1^h - p_2\delta_1 H_1 - \delta_2\psi^h + \delta_1\psi^h - \varepsilon\}$. New resources must be greater than $\delta_2\omega^{SR}$ in the event of no-default, i.e., $\delta_2\omega_2^c + p_2\delta_2 H_1^h - p_2\delta_1 H_1 - \delta_2\psi^h + \delta_1\psi^h - \varepsilon \geq \delta_2\omega^{SR}$. 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 $\delta_1 \to 0$ (so ε and δ_2 vanish too), we conclude that $(x^h, \psi^h) \in \overline{\dot{B}^h}(p_1, p_2, q)$, as long as consumers can consume their resources. Thus, correspondence $\dot{B}^h$ is lower hemi-continuous for each consumer. □

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

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

is the integral of upper semi-continuous demands with respect to a nonatomic measure, we have that $\int_{\mathcal{G}\cup\mathcal{B}:c(h)=C}(x^h(m^l_{c(h)}), \psi^h(m^l_{c(h)}))d\lambda$ is upper hemi-continuous. The compact-valued function $h \to (x^h(m^l_{c(h)}), \psi^h(m^l_{c(h)}))$ is bounded above and below by $(\int_{\mathbf{A}} \omega(a)da, \bar{H}, \int_{\mathbf{A}} Bda)$ 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, $\int_{\mathcal{G}\cup\mathcal{B}:c(h)=C}(x^h(m^l_{c(h)};\bar{p},\bar{q}),$ $\psi^h(m^l_{c(h)};\bar{p},\bar{q}))d\lambda$ is compact and has nonempty values. Using a similar reasoning, we can show that the measurable aggregate demand $\int_{\mathcal{SB}\cup\mathcal{TB}:c(l)=C}(R^l, H^l, \varphi^l)d\lambda$ 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 $(x^h(m^l_{c(h)}), \psi^h(m^l_{c(h)}))$ has convex values.[48] 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 (Non-convex generalized game with endogenous mortgage market choices): Given the consumers' optimal consumptions in each mortgage market, consumers choose their most preferred mortgage market (recall that $l = \emptyset$ is a possibility). Let

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

Then, $\mu^h(m^l_{c(h)}) = 1$ if $l \in \arg\max U^h(l)$ and 0 otherwise (as $\sum_{l=TB,SB,\emptyset} \iota(m^l_{c(h)}) = 1$). 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

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

argument as above, there is a measurable selection $h \to \mu^h(m)$ 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 $\sum_{l=TB,SB,\emptyset} \iota(m^l_{c(h)}) = 1$), convex (by Lyapounov's convexity theorem) and nonempty values.

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

Auctioneer 1 chooses p_1 to minimize $\left(\sum_l \int_{\mathcal{G}\cup\mathcal{B}} H^h_1(m^l_{c(h)})\bar{\mu}^h(m^l_{c(h)})d\lambda - \bar{H}\right)^2$, where $\bar{H}$ stands for the exogenous supply of housing from an older previous generation. *Auctioneer 2* chooses p_2 to minimize $\left(\sum_l \int_{\mathcal{G}\cup\mathcal{B}} H^h_2(m^l_{c(h)})\bar{\mu}^h(m^l_{c(h)})d\lambda - \bar{H}\right)^2$, where $\bar{H}$ stands for the exogenous demand of housing from a younger future generation. *Auctioneer 3* chooses φ^{TB} and φ^{SB} to minimize $\sum_l \int_{\mathcal{G}\cup\mathcal{B}}(\bar{\psi}^h(m^l_{c(h)})\bar{\mu}^h(m^l_{c(h)})d\lambda - \int_{\mathcal{L}} \varphi^l(m^l_{c(h)})\bar{\mu}^h(m^l_{c(h)})d\lambda)^2$. *Auctioneer 4* chooses z^{SB} to minimize $\left(d^{SB}\bar{\varphi}^{SB} - z^{SB}\right)^2$. *Auctioneer 5* chooses z^i to minimize $(z^i - \bar{z}^{SB})^2$. 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 $l \in \{TB, SB\}$, where functions f and g are as defined in Section 4.

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 $\bar{\varkappa}^a$ to solve his respective optimization problem parameterized in the other players' actions $\bar{\varkappa}^{-a}$.

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 $\{TB, SB, \emptyset\}$, which is the set of mixed strategies for each consumer. A profile of strategies $\rho : \mathcal{G} \cup \mathcal{B} \to \Sigma(\mathbf{M})$ 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 $\max_{\sigma\in\Sigma(\mathbf{\Omega})} U^h(\sigma)$. 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 extent the fictitious auctioneers' problems to allow for consumers' mixed strategies. Given a mixed strategy profile $\rho:\mathcal{G}\cup\mathcal{B}\to\Sigma(\mathbf{G})$, 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 $\left(\sum_{m\in\mathbf{M}}\int_{\mathcal{G}\cup\mathcal{B}}H_1^h(m)\rho^h(m)d\lambda-\bar{H}\right)^2$; *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^{SB} to minimize $\sum_{m\in\mathbf{M}}(\int_{\mathcal{G}\cup\mathcal{B}}\bar{\psi}^h(m^l_{c(h)})\rho^h(m^l_{c(h)})$ $d\lambda-\int_{\mathcal{L}}\varphi^l(m^l_{c(h)})d\lambda)^2$, for $l=TB,SB$. 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 and, therefore, 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 $\int_{\mathcal{G}\cup\mathcal{B}}\int_{\mathbf{M}}H_1^h(m;\bar{p}_1)d\rho^h(m)d\lambda$.[49] 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 pre-

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

vious 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 $\int_{\mathcal{G}\cup\mathcal{B}}\int_{\mathbf{M}} H_1^h(m;\bar{p}_1)d\bar{\rho}^h(m)d\lambda$. 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. □

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:

$$\zeta_1 \equiv \sum_{m\in\mathbf{M}}\int_{\mathcal{G}\cup\mathcal{B}}\left(p_1H_1^h(m_{c(h)}^l)+R_1^h(m_{c(h)}^l)-q^l\psi^h(m_{c(h)}^l)-\omega^{SR}\right)\mu^h(m_{c(h)}^l)d\lambda-p_1\bar{H}+ \sum_{l=TB,SB}\int_{\mathcal{L}}(\omega_1^l-q^l\varphi^l+\tau z^l)d\lambda+\int_{\mathcal{I}}(\omega_1^i-\tau z^i)d\lambda\leq 0$$

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, a contradiction with optimization. Thus, $\zeta_1 = 0$. By a similar argument, we can also prove that $\zeta_2 = 0$. □

A.2 Minimum House Size

A.2.1 Conduit 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 $H_{SB}^{\min}$ that prevents B-type consumers with a small conduit loan to buy a house with a lot size larger than $H_{SB}^{\min}$.

We now identify threshold $H_{SB}^{\min}$ as a function of the parameters of our economy. First, notice that the SB 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 $d^{SB} = 1$ as this gives the largest loan amount to the consumer since pricing uses the investor's discount factor θ^i). Then, using $pH_1^B = \omega^{SR} + q^B \varphi^B$ 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 SB 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 $H^B < H_{SB}^{\min}$, i.e., if $(\omega^{SR} + \bar{L})/p < H_{SB}^{\min}$. The market clearing price for owner-occupied housing is $p = (2\omega^{SR} + \bar{L} + L^G)/\bar{H}$, where L^G is the loan amount that a G-type consumer would obtain from a SB when mortgage markets are segmented (using the SB'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 $H_{SB}^{\min}$ we can write

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

Hence, we conclude that a minimum house size policy can rule out the possibility of a separating equilibrium if inequality (8) holds.

A.2.2 Portfolio 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 $H^{\min}_{SB}$ on a conduit mortgage market specific for B-type consumers, we can also find a threshold $H^{\min}_{TB}$ that rules out a portfolio mortgage market specific for B-type consumers. The portfolio mortgage contract $(q^{B,r}, \psi^{B,r})$ specific for B-type consumers must satisfy budget constraints $pH^{B,r}_1 = \omega^{SR} + q^{B,r}\psi^{B,r}$ and $\omega^{SR} = \omega^{SR} - \psi^{B,r} + pH^{B,r}_1$ (the latter coming from the limited recourse requirement), which implies $\psi^{B,r} = pH^{B,r}_1$ 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 $\psi^{B,r} = pH^{B,r}_1$ we get $H^{B,r}_1 = \omega^{SR}/p(1 - \theta^l\delta)$. Then, set

$$H^{\min}_{TB} \equiv \omega^{SR}/p(1 - \delta\theta^l) \tag{21}$$

This housing policy implies that subprime consumers with a small portfolio 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 certainly below $H^{\min}_{TB}$.

A.3 Equilibrium Closed Form Solutions

In this section we briefly present the closed form solutions of our equilibrium model. We start with the baseline model where the SB does not acquire soft information.

A.3.1 No Soft Information Acquisition

First, mortgage discount prices q^{TB}, q^{SB} and τ follow by solving the system of first order conditions on $\Phi^{TB}(\varphi^{TB}, z^{TB})$, $\Phi^{SB}(\varphi^{SB}, z^{SB})$ and $\Lambda^i(z^i)$ with respect to φ^{TB}, (φ^{SB}, z^{SB}) and z^i, respectively, and subject to the respective mortgage distribution constraints. On the one hand, a TBs, who by assumption has $d^{TB} = 0$ and $\pi^{TB} = 1$, finds 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 conduit loans under hard information only is

$$q^{SB} = \frac{\bar{\pi}\bar{\theta}}{1 - \delta(1-\bar{\pi})\bar{\theta}} \tag{22}$$

where $\bar{\pi} \equiv \pi^{SB} = \pi^i$ and $\bar{\theta} \equiv d^l\theta^i + (1-d^l)\theta^l$. Since $\theta^i > \theta^l$, a higher mortgage distribution rate d^l implies a higher q^{SB}. Adverse selection is captured by belief $\bar{\pi} < 1$ and decreases the SB's discount price. The term $1-\delta(1-\bar{\pi})\bar{\theta}$ in (22) is the "default loss" that the SB 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 SB offers to its borrowers. The SB's mortgage rate (or cost of capital) is $1/q^{SB}$.

Using the TB and SB' pricing expressions, we obtain

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

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

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

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

Threshold π_2 is the same as the one found in Section 5 when we characterized the different equilibrium regimes. Interesting, we see that as the distribution rate d^{SB} increases, threshold π_2 decreases, and hence more hard information is needed to sustain an environment where the conduit mortgage rate is below the TB's rate.

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^{SB}, φ^{SB}) as the pooling contracts offered by TBs and SBs, 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 portfolio loan amount expression:

$$q^{TB}\psi^{TB} = \frac{\theta^l}{1-\theta^l} \tag{23}$$

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

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

$$q^{SB}\psi^{SB} = \frac{\pi^{SB}\bar{\theta}}{1-\bar{\theta}(\pi^{SB}(1-\delta)+\delta)} \tag{24}$$

B-type consumers that receive a good rating by the SB 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 conduit loan amount increases with the predictive power of the hard credit scoring technology (and thus with the SB's belief π^{SB}), the foreclosure recovery rate δ, and the d^{SB}-weighted discount factor $\bar{\theta}$. The term $\bar{\theta}$ in turn increases with the distribution rate d^{SB} and the investor's discount factor θ^i and decreases with the lender's discount factor θ^l.

The SB's income from distributing mortgages to investors is given by the following expression:[50]

$$\tau z^{SB} = \frac{d^{SB}\mu^{SB}(\text{rating=G})}{1-\delta\bar{\theta}(1-\pi^{SB})} \tag{25}$$

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

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

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

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

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

[50] The equilibrium quantity of mortgages originated by SBs is constrained by the investor's wealth because $\tau z^i \leq \omega_1^i$. Thus, our model is also able to capture Gennaioli, Shleifer, and Vishny (2012) result that investors' wealth may drive up securitization. To see this, notice that SBs are constrained by the total amount of credit that can be securitized, i.e., $d^{SB}\varphi^{SB} \leq z^{SB} = z^i$ where the first inequality obeys the originate-to-distribute constraint (2) and the second equality follows from market clearing in the secondary mortgage market.

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

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

where $\lambda(G\text{-}Rating^{SB})$ is the endogenous measure of consumers that borrow from SBs,[51] i.e., $\lambda_1(G\text{-}Rating^{SB}) = 0$ when $\pi^{SB} < \max\{\pi_0, \pi_1\}$, $\lambda_2(G\text{-}Rating^{SB}) = CST^{SB}(\lambda_G - v(TB)) + (1-CST^{SB})\lambda_B$ when $CST^{SB} \in [\max\{CST_0, CST_1\}, CST_2]$, and $\lambda_3(G\text{-}Rating^{SB}) = CST^{SB}\lambda_G + (1-CST^{SB})\lambda_B$ when $CST^{SB} > CST_2$.

With the above expresssions, 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 (6) we illustrate the equilibrium values of house sizes H^{TB} and H^{SB} as a function of CST^{SB}. There we see that when the economy enters Regime 3, the house size of borrowers with conduit loans is larger than for borrowers with portfolio loans. This is consistent with the idea that the portfolio mortgage market is not the consumers' first option in Regime 3. We also see that the equilibrium house size of consumers with portfolio loans plummets when the conduit loan size enters in Region 3, as the expansion of the conduit 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 conduit loans when $\hat{\pi}_G^{SB}$ and CST^{SB} are such that $\pi^{SB} = \pi_2$ even when the jump in the conduit loan amount is partially offset by the jump in the equilibrium house price at that point.

It remains to give the expressions that determine the size of the rental market. First, notice that TBs exhaust their lending capacity constraint $v(TB) = 1$ by lending to a mass 1 of G-type

[51] Recall that the measure of consumers that receive a good rating in mortgage market $l \in \{TB, SB\}$ is $\lambda(G\text{-}Rating^l) = CST^{SB} \cdot \hat{\lambda}_G^l + (1-CST^{SB}) \cdot \hat{\lambda}_B^l$, 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 $l \in \{TB, SB\}$.

Figure 6:

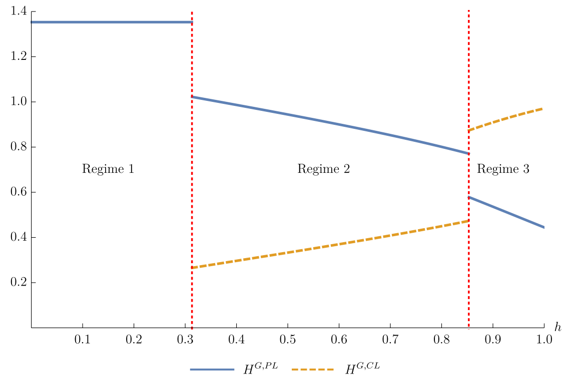
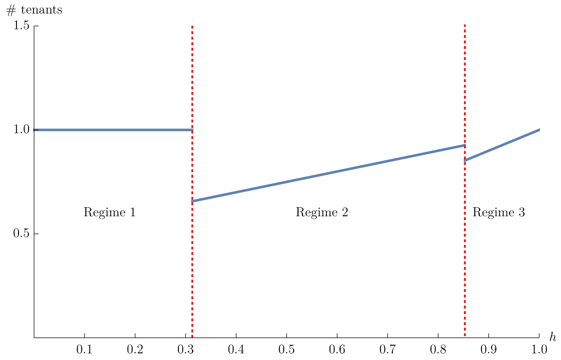


Figure 7:



borrowers. Hence, when $CST^{SB} < CST_1$, there are only portfolio loans issued, and therefore a mass

$$\lambda_G + \lambda_B - v(TB) \tag{26}$$

of households have no other option but to rent. Second, since SBs can absorb all excess demand of consumers with a good rating (mass $\lambda_2(GRating^{SB})$), we have that, when $CST^{SB} \in [CST_1, CST_2]$, a mass

$$\underbrace{(\lambda_G - 1) + \lambda_B}_{\text{Remaining consumers without a portfolio loan}} \quad - \quad \underbrace{\lambda_2(G\text{-}Rating^{SB})}_{\text{Mass of consumers with a conduit loan } (\lambda(GRating^{SB})) \text{ when } \pi^{SB} \in [\pi_1, \pi_2]} \tag{27}$$

have no other option but to rent. Third, when $CST^{SB} \geq CST_2$, all consumers attempt to get a conduit loan first. However, only a mass $\lambda_3(G\text{-}Rating^{SB})$ of consumers get a conduit loan. Those G-type consumers without a conduit loan, with mass $(1 - CST^{SB})\lambda_G$, try to get a portfolio loan, their second option, but not all of them may end up with a portfolio loan if the TB's capacity constraint binds. Thus, the size of the rental market in Regime 3 is

$$\lambda_G + \lambda_B - \underbrace{\lambda_3(G\text{-}Rating^{SB})}_{\text{Mass of consumers with a conduit loan } (\lambda(GRating^{SB}) \text{ when } \pi^{SB} > \pi_2} - \underbrace{\min[(1 - CST^{SB})\lambda_G, 1]}_{\text{G-type consumers wih a conduit loan}}, \tag{28}$$

Figure (7) shows that the size of the rental market is largest in Regime 1 (only the portfolio mortgage market exists). When Regime 2 starts (CST^{SB} attains CST_1), the rental market shrinks as new consumers get (conduit) mortgages. The rental market shrinks again in Regime 3 (CST^{SB}

attains CST_2), as the conduit 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. In Regime 3 a mass $(1 - CST^{SB})\lambda_B$ of B-type consumers are able to get a conduit loan. However, as CST^{SB} gets closer to 1, SBs better differentiate between G-type and B-type consumers and reject more B-type consumers, and, as a result, the size of the rental market converges to the "number" of B-type consumers in the economy ($\lambda_B = 1$).

For the sake of brevety, we decided to omit the computation details of thresholds π_0, π_1 and π_2 (and their corresponding thresholds CST_0, CST_1 and CST_2).[52]

A.3.2 Soft Information Acquisition

Section 7 extended the baseline model to allow for SB's soft information acquisition. There, we assumed that in order to hide the soft information to the secondary market investors, the SB 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 $\lambda_1(G\text{-}Rating^{SB})$ changes when soft information acquisition is a possibility. The new expressions for Regimes 2 and 3 are the following:

- Regime 2:

$$\lambda_{2,Soft}(G\text{-}Rating^{SB}) = \\ = \underbrace{\Big(CST^{SB}_{Soft}(\lambda_G - v(TB)) + (1 - CST^{SB}_{Soft})\lambda_B\Big)}_{\text{measue of loans with ``}rating\ soft{=}G\text{''}} + \underbrace{\Big((1 - CST^{SB}_{Hard})\lambda_B - (1 - CST^{SB}_{Soft})\lambda_B\Big)}_{lemons}$$

- Regime 3:

$$\lambda_{3,Soft}(G\text{-}Rating^{SB}) = \\ \underbrace{\Big(CST^{SB}_{Soft}\lambda_G + (1 - CST^{SB}_{Soft})\lambda_B\Big)}_{\text{measue of loans with ``}rating\ soft{=}G\text{''}} + \underbrace{\Big((1 - CST^{SB}_{Hard})\lambda_B - (1 - CST^{SB}_{Soft})\lambda_B\Big)}_{lemons}$$

[52]The authors can facilitate the algebra details upon request.

Thus, the market clearing equation is

$$\lambda_{2,Soft}(G\text{-}Rating^{SB})\psi^{SB} = \varphi^{SB} \text{ if Regime 2}$$
$$\lambda_{3,Soft}(G\text{-}Rating^{SB})\psi^{SB} = \varphi^{SB} \text{ if Regime 3}$$

A.4 Investors' Liquidity and the Distribution of Income

In this section, we 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) as measured by θ^i and 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^{SB} indicates an increased Regime 3 (dominant SB sector) size.*

2. *For a given CST^{SB}, 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.*

Proof of Proposition 3: The thesis part of Proposition 3.1 follows because a higher θ^i decreases threshold CST_2^{SB} (as well as thresholds CST_0^{SB} and CST_1^{SB}). 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^{SB} falls from 0.85 to 0.81.

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

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

However, a shock to the fundamental proportion of G-type consumers decreases π^{SB} because it depends on $\hat{\pi}_G^{SB}$ (see expression (1), and $\hat{\pi}_G^{SB}$ is a decreasing function of $\lambda_G/(\lambda_G + \lambda_B)$. It stands to reason that a big enough shock to the fundamental proportion can bring π^{SB} below threshold π_2. ■

Proposition 3.1 states that when the investor's patience parameter θ^i increases, the SB's implied risk-adjusted mortgage rate decreases to increase the mortgage loan amount and house price. This happens because a fraction d^{SB} of the mortgage loans originated are now subject to a lower cost of capital.[53] 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, as previously noted, 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 SB originated mortgage product, to increase significantly (again see Brunnemeier 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 proof of this proposition is left for the Appendix. 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 $\hat{\pi}_G^{SB}$ (see equation (1)). As the proportion of G-types decline, the assessed credit quality of the SB mortgage 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 subprime households, and more broadly as a deterioration in household's net worth, as documented by Mian and Sufi (2014, 2017).

[53] See the pricing equation (22) and the equilibrium loan amount expression in the Appendix A.3.

B Recourse Versus Non-recourse Mortgages

In our baseline model we assumed that mortgage contracts were recourse but subject to limited liability. Here we analyze the equilibrium implications of considering non-recourse mortgages instead than (limited) recourse mortgages.

Adverse selection and the nature of the mortgage contract: In a recourse mortgage the borrower can credibly commit to pay back the loan even if the house value is below the debt amount (until the point where paying the promise would involve consuming below the subsistence rent).[54] Adverse selection then arises for (limited) recourse contracts because subprime consumers have different probabilities of receiving a high endowment in the second period. In the good state both consumer types honor the promise, and in the bad state both types default. The probability of occurrence of each state is different between the two types of consumers though. In a non-recourse mortgage, if the house does not sell for at least what the borrower owes, the lender must absorb the difference and walks away.[55] Accordingly, the second period budget constraint (7) should be rewritten as follows: $p_2 H_2 + R_2 \leq \omega_2^t + p_2 H_1 - \min\{p_2 H_1, \psi\}$.[56] Notice that if we were to modify the baseline model in such a way, the adverse selection problem 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}. Notice also that the non-recourse contract does not need to include a limited liability clause, which allows the borrower to consume at least the subsistence rent in the second period, since when $\psi \leq pH_1$, the borrower always has means to repay the loan by selling his house and, therefore, does not need to use his own endowment to satisfy the mortgage payment.

Welfare: A non-recourse contract may prevent the consumer to borrow against all the second period income that is above ω^{SR}, as the promise cannot be larger than the house value (pH_1) in the baseline model. Non-recourse, by eliminating adverse selection, causes the G-type to delay some consumption until the second period. This is welfare decreasing, since households prefer to consume more in the first period. This is both because the household is impatient and because

[54]The way bankruptcy/foreclosure law works is that non-payment results in wage garnishment.

[55]Notice that in both recourse and non-recourse mortgages, the lender would be able to seize and sell the house to pay off the loan if the borrower defaults.

[56]See Araujo, Pascoa and Torres-Martinez (2002) and Araujo, Fajardo and Pascoa (2005) for general equilibrium ecoomies with non-recourse collateral and default.

the younger household derives more utility from owning a house than renting.

In a recourse mortgage the lender can go after the borrower's other assets or sue to have his assets, including his saving accounts, garnished. Limited recourse loans are captured in our baseline model by the second period budget constraint (7): $p_2H_2 + R_2 \leq \max\{\omega^{SR}, \omega_2^t + p_2H_1 - \psi\}$. Under this contract the borrower can credibly commit to pay back the loan even if $p_2H_1 < \psi$ until the point where paying the promise would involve consuming below the subsistence rent (i.e., $\omega_2^t + p_2H_1 - \psi < \omega^{SR}$), which by assumption is protected by the nature of the contract. Adverse selection then arises for (limited) recourse contracts because subprime consumers have different probabilities of receiving a high endowment in the second period. In the good state both consumer types honor the promise, and in the bad state both types default. The probability of occurrence of each state is different between the two types of consumers though.

Empirical evidence: Finally, it 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 over portfolio loans (partialing out observable borrower characteristics) on a state recourse dummy (Ghent and Kudlyak 2011), and found a positive coefficient for the recourse state dummy ($spread = 0.09 + 0.11 \cdot recourse\ dummy$).[57]

Online Appendix List of References

[1] Araujo, A., M. Pascoa, and J. P. Torres-Martinez (2002), "Collateral Avoids Ponzi Schemes in Incomplete Markets", *Econometrica* 70, 1613-1638.

[2] Araujo, A., J. Fajardo, and M. Pascoa (2005), "Endogenous Collateral", *J Math Econ* 41, 439-462.

[3] Araujo, A. and M. Pascoa (2002), "Bankruptcy in a Model of Unsecured Claims", *Econ Theory* 20, 455-481.

[4] Debreu, G. (1952), "A Social Equilibrium Existence Theorem", *Proceedings of the National Academy of Sciences.*

[57] See Elliot Annenmber's discussion at the HULM conference in the Federal Reserve of Chicago (Ocotber 23, 2015).

[5] Gennaioli, N., A. Shleifer, and R. Vishny (2013), "A Model of Shadow Banking", *J Finance* 68, 1331-1363.

[6] Ghent, A. and M. Kudlyak (2011), "Recourse and Residential Mortgage Default: Evidence from US States", *Rev Financ Stud* 24, 3139-3186.

[7] Hildebrand, W. (1974), *Core and Equilibrium in a Large Economy*, Princeton: Princeton University Press.

[8] Luque, L. (2014) "Wages, Local Amenities and the Rise of the Multi-skilled City", *Annals Reg Sci* 52, 457-467.

[9] Luque, J. (2013), "Heterogeneous Tiebout Communities with Private Production and Anonymous Crowding", *Reg Sci Urban Econ* 43, 117-123.

[10] Pascoa, M. (1998), "Nash Equilibrium and the Law of Large Numbers", *Int J Game Theory* 27, 83-92.

[11] Poblete-Cazenave, R. and J.P. Torres-Martinez (2013), "Equilibrium with Limited-recourse Collateralized Loans", *Econ Theory* 53, 181-211.

[12] Schmeidler, D. (1973), "Equilibrium Points of Non-atomic Games", *J Statistical Physics* 7, 295-301.