THE FIFTEENTH
YOUNG ECONOMISTS'
SEMINAR

TO THE TWENTY-EIGHTH DUBROVNIK ECONOMIC CONFERENCE

Organized by the Croatian National Bank

Marta Cota and Ante Šterc

Search and Skills in the Mortgage Market

Royal Hotels & Resort Dubrovnik

July 2, 2022

Draft version

Please do not quote

CROATIAN NATIONAL BANK

Search and Skills in the Mortgage Market

Marta Cota* Ante Šterc†

June 15, 2022

Abstract

The size of the U.S. mortgage market has been steadily increasing ever since 2012, reaching all-time highs during the pandemic due to the increase in demand. While other studies focus on the home price expectations, we focus on the effect of borrowers' financial knowledge on their mortgage performance. Our objective is to capture the difference in borrowers' search efficacy in the mortgage refinance market. Using a novel public data set on mortgage originations, we show that search efficacy accounts for a significant part of the interest rate dispersion. That is, holding personal and loan characteristics fixed, the interaction of borrower's education and search effort explains a part of the interest rate obtained at the mortgage origination. To set grounds for human capital spillovers in the mortgage market, we exploit public survey data and argue that education implies financial knowledge, while holding other characteristics fixed. Moreover, we use machine learning techniques to impute financial literacy based on borrower's observables and infer the consequences for the interest rate attained in the mortgage market. Finally, we develop a search model with endogenous financial skill accumulation to capture mortgage refinance interest rate dispersion. Complementary to typical models in the literature, we set focus on the agent's effort, while incorporating lender's market power via observable default probability. While higher skills imply higher wage and lower default probability, accumulating knowledge takes time from work. Our aim is to capture consumption inequality across financial skills in the economy and propose policies that alleviate search costs through financial education.

Keywords: mortgage refinancing, search, financial literacy, financial skills accumulation.

*Email: marta.cota@cerge-ei.cz; CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Politickych veznu 7, 111 21 Prague, Czech Republic.

†Email: ante.sterc@cerge-ei.cz; CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Politickych veznu 7, 111 21 Prague, Czech Republic. We would like to thank Marek Kapička and Ctirad Slavík for guidance and Anmol Bhandari, V. V. Chari, Mariacristina De Nardi, Axel Gottfries, Fatih Guvenen, Ole Jann, Larry E. Jones, Loukas Karabarbounis, Jeremy Lise, Filip Matějka, Christopher Phelan, Kathrin Schlafmann, Petr Seldáček, Kjetil Storesletten, as well as participants of workshops and seminars at CERGE-EI and at the University of Minnesota for useful comments.

1 Introduction

During the period of low interest rates, the mortgage market share in the U.S. has been increasing steadily, with refinance origination numbers almost twice as high as first origination mortgages. Moreover, throughout the years, most of mortgages were taken up under a fixed rate contract, with a term of 20 or 30 years. Consequently, dollar amounts in the U.S. mortgage market boomed in the last 10 years. Specifically, using the data from the National Mortgage Database(NMDB) data, which serves as a 5% sample of the total mortgage universe, the total amount of dollars shows that first origination amount has doubled. Moreover, mortgages taken up for refinancing purposes have more than doubled from 2010.

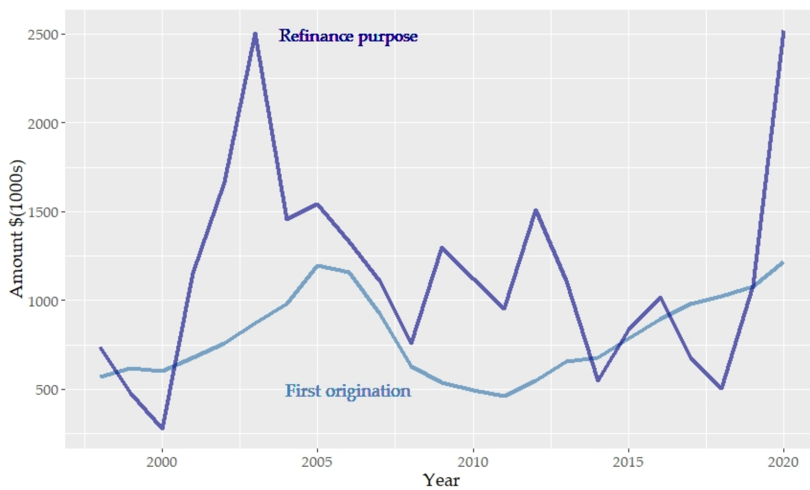


Figure 1: Mortgage amount, in 1000$, own calculations. Source: *Annual New Residential Mortgage statistics, NMDB.*

Even though we take a look at the representative sample in our analysis, it is worth mentioning that the population mortgage amounts have climbed up from 2 trillion to 4 trillion U.S. dollars[1]. Besides, first look at the Survey of Consumer Finances show that mortgage ownership fluctuates around 48% from 1990s onward.

Therefore, taking up mortgage is one of the main financial decisions in the average U.S. household. We analyze the quality of borrower's decision process in terms of the amount of effort given into shopping for mortgage. That is, this paper zooms in on the mortgage contracts across households and explains a part the mortgage interest rate dispersion observable in the data.

[1]Mortgage Bankers Association reports, 2020.

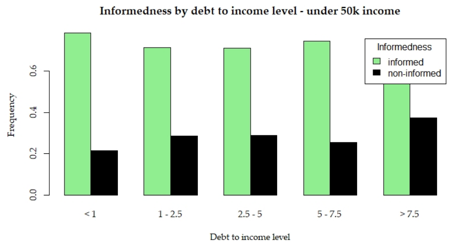
Taking up a mortgage with a fixed interest rate is a 30-year long commitment that affects borrowers' credit scores, quality of living (mortgage monthly payment puts a budget constraint on consumption) and ultimately, other investments. Given the interest rate spread of mortgage contracts, we disentangle between efficient and inefficient search, basing our findings on borrowers' human capital. We think of human capital as a joint measure of education and financial knowledge, as we find that education and financial literacy correlate[2]. Ultimately, we build a model with human capital accumulation where human capital does not only imply borrower's income but also affects credit responsibility that further translates into search efficacy upon entering the mortgage market.

Clearly, as education is not the sole driver of the interest rate spreads in the data, we use empirical estimations to determine the amount of spread explained by education and search efficacy. It turns out that the interaction between the two is significant upon mortgage refinancing. Even though the average loan amounts at first and refinancing origination remain similar over the years, the average refinancing age is 10 years higher than at the first origination (NMDB data), implying that borrowers choose to refinance after some time spent in the mortgage market. The usual understanding is that it may be optimal to refinance in times of low interest rates, and given that the interest rate lingered at record lows over the last decade, we should see big amounts of refinancing originations. Our data analysis shows that, among borrowers who refinanced, almost 65% have at least finished college.

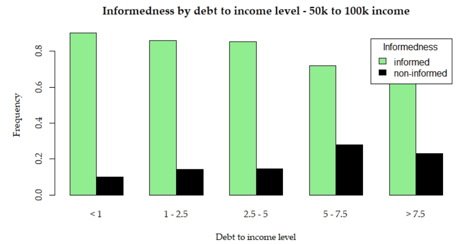
Given the the effect of mortgage contract on household finance, we use the Survey of Consumer Expectations (SCE) data to analyze potential caveats underlying the decision to refinance. Due to the nature of questions within, SCE survey sheds some light on as to why borrowers do not refinance. Our findings on household awareness of opportunities and benefits of refinancing serve as a motivating fact for the model we implement. Put simply, over 40% of low-income households with substantial debt-to-income ratios do not know how to get around refinancing or are not aware of the opportunities provided by changing the mortgage lender (Figure 2a). By contrast, households with low debt exposure and higher income understand the process and the opportunities at hand (Figure 2c).

Our focus is set on the effect of borrower's knowledge accumulation and search effort on the new mortgage interest rate that ultimately benefits borrowers via the period budget constraint. In the model, human capital (skill) accumulation requires time investment and guarantees higher wage. Simultaneously, skill accumulation decreases search cost that borrowers need to endure when opt-

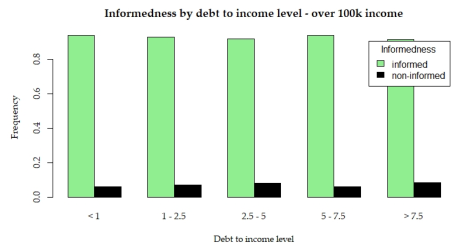
[2] SCF data, 2016.-2020.



(a) Under $50k$ income.



(b) $50k$ to $100k$ income.



(c) Over $100k$ income.

Figure 2: Share of non-informed households by debt to income ratio levels over the income distribution. Source: SCE, authors' calculation.

ing in for mortgage refinancing with other lenders in the market. Consequently, the lender considers skilled borrowers as financially savy, and bargains over the search effort the borrowers need to sustain and the interest rate at which they repay.

Time to acquire skill takes time off work, so low-income workers face a trade-off between setting grounds for better financial conditions with respect to mortgage contracts available and earning their wage to pay off their current debt (conditional on undertaking mortgage before). In contrast, high-skilled workers use their skills and may refinance later.

Even though current research includes adversarial effects in mortgage contract, this paper basis grounds on digitalization of mortgage offers. It has become fairly easy to go online and search for mortgage opportunities by putting in basic information such as credit score and previous credit responsibilities. Furthermore, new algorithms evaluate potential clients based their evaluation on applicants' observables, including individual credit score. With the ever-growing amount of data that trains FinTech algorithms (Bartlett et al., 2022), the adversarial effect of evaluating unobservables becomes less important[3]. As opposed to being random, this paper models search costs as a consequence

[3] As it is automated and evaluated based on 8×8 matrix of borrower's characteristics.

of borrowers' skills, as they know their skills can affect their mortgage contract via decrease in default probability and search efficacy.

As we evaluate the effects of financial skill accumulation on the mortgage performance, we aim to assess the implications of search cost reductions via the decrease in time spent in skill accumulation or a straightforward reduction in search costs. Considering the amount of mortgage interest rate dispersion our model would be able to generate, making borrower's similar with respect to search efforts may imply reduction in consumption inequality propagated by differences in budget constraints across borrowers. Policy tests we aim to implement are motivated by the increasing interest in improving financial literacy among households (Lusardi and Mitchell, 2014; Van Rooij et al., 2011; Bhutta et al., 2020), as it is shown that financial skills affect household liquidity and other financial investments.

2 Related Literature

Due to the financial crisis in 2008 starting in sub-prime markets, a lot of work has been done in evaluating regulatory inefficiencies in the mortgage market. Specifically, empirical studies incorporate data prior to the crisis to develop models of adverse selection in the mortgage markets to explain the dispersion among interest rates in mortgage contracts (Woodward and Hall, 2012; Mayer et al., 2013; Campbell, 2013). However, regulatory changes put in effect afterwards prevent lenders and mortgage brokers from passing on their fees through borrowers.

Following the changes in regulations, lenders needed to develop effective screening tools for potential borrowers. In this respect, optimal contract design literature models the price origination in the market while imposing random (or non-existent) search cost assumptions (Yannelis and Zhang, 2021; Chatterjee et al., 2020). Recently, Agarwal et al. (2020) argue that mortgage market data conforms to searching behavior implied by fear of being rejected, and develop the adverse selection model that generates more search with risky borrowers. Notably, Agarwal et al. (2020) incorporate data on formal mortgage applications which exclude simple search methods such as internet search and "listening to opportunities".

We contribute by adding to search literature oriented towards mortgage markets, while incorporating refinancing decision conditional on amount of search effort endured in the process. Search frictions induce sub-optimal behavior in mortgage undertaking, as shown in Argyle et al. (2020) and

Andersen et al. (2020). While Andersen et al. (2020) contribute frictions to behavioral factors, studies closer to ours attribute interest rate dispersion to data on shopping effort. Specifically, Alexandrov and Koulayev (2018) incorporate costly search with borrowers who hold beliefs about the interest rate dispersion. Our study, in contrast, incorporates digitization advancements in the mortgage market and assumes that borrowers are informed about the interest rates they could attain. We add to Alexandrov and Koulayev (2018) by adding more structure to the search costs and abstracting from oligopoly assumptions in the market. Argyle et al. (2020) estimate search costs directly from the price dispersion in the auto-loan markets to decompose borrowers into low- and high- mark-up clients, without relating costs (i.e., mark-ups charged) to borrowers' observable characteristics.

Our model incorporates search costs that depend on individual skill level, possibly in relation to financial literacy that is getting traction in the household finance literature. Since this paper incorporates both first mortgage origination as well as the decision to refinance, we contribute to existing research. Keys et al. (2016) find that over 20% of U.S. borrowers did not refinance when it was optimal, during the time of low interest rates. Importantly, the authors do not find any relation of the sub-optimality with the options borrowers face. Put differently, Keys et al. (2016) conclude that the main driver of sub-optimal behavior lies within search frictions. Similarly, Andersen et al. (2020) document the failure to refinance with wealthier individuals and attribute it to costs of time devoted to the process outweighing the benefits. On the other end, Andersen et al. (2020) explain *the mortgage inaction* with the use of heuristics based on periodic time intervals when households refinance.

Finally, financial literacy is widely used in explaining empirical results obtained from private lender datasets (Bhutta et al., 2020; Agarwal et al., 2016; Koszegi, 2014). This paper takes individual skill as a choice that affects borrower's budget constraints from two sides - through wages and credit responsibility ($\approx$ financial literacy). We motivate our specification by combining findings in multiple datasets. The fact that skill affects households savings through two channels elicits findings on interaction between individual financial literacy and education, which in turn affects interest rate obtained at the mortgage market. In this respect, we contribute to recent studies on importance of financial literacy in mortgage markets and passing fees on financially unskilled borrowers (Gurun et al., 2016; Bhutta et al., 2020; Gathergood and Weber, 2017).

3 Data - facts and figures

In our data analysis we use few publicly available surveys to infer important measures for the mortgage search we implement in the second part of the paper.

3.1 The National Survey of Mortgage Originations

The National Survey of Mortgage Originations (NSMO) contains details on the mortgage contract upon origination, and includes borrower's characteristics. We set focus on the search effort proxy variable that is borrowers'answer to the question:

- *How many different mortgage lenders/brokers did you seriously consider before choosing where to apply for this mortgage?*

Instead of looking at the number of formal application with lenders, we think of the answer to the former question as self-reported effort **before** handling the documentation. The number of lenders that borrowers end up applying to is significantly lower than the number of lenders they actually considered. Our argument is that, due to the complexity of the documentation set-up, it would be only the borrowers who fear getting reject that end up applying with more lenders, which is a driving mechanism of mortgage decision in the literature (Agarwal et al., 2020).

Abstracting from the formal choice only, we think of the consideration as current way that borrowers search, via online applications that compare different lenders and "recommend" the best choice, given the borrower's credit score and income. Given that the number of lenders considered varies much more across our sample, the survey answer captures not only monetary costs endured in the application process, but also cognitive effort facing mortgage undertaking. That is, we take a novel approach by using the consideration set size, that is subject to knowledge effects (Figure 3). Later, we present the evidence in the data that imply assumptions on knowledge effect in the search model.

Mortgages in the NSMO data originate from 2013 till 2021 the latest. Randomly chosen households report the specifics of their contracts, reasons, and experiences regarding mortgage undertaking. Details about their mortgage origination combined with demographic characteristics allow us to capture the effect of borrowers' characteristics on the mortgage interest rate acquired, while controlling for the mortgage specifics. Firstly, we consider respondent's attitude towards the mortgage market and beliefs regarding the optimality of their lender choice. Secondly, we quantify education and effort

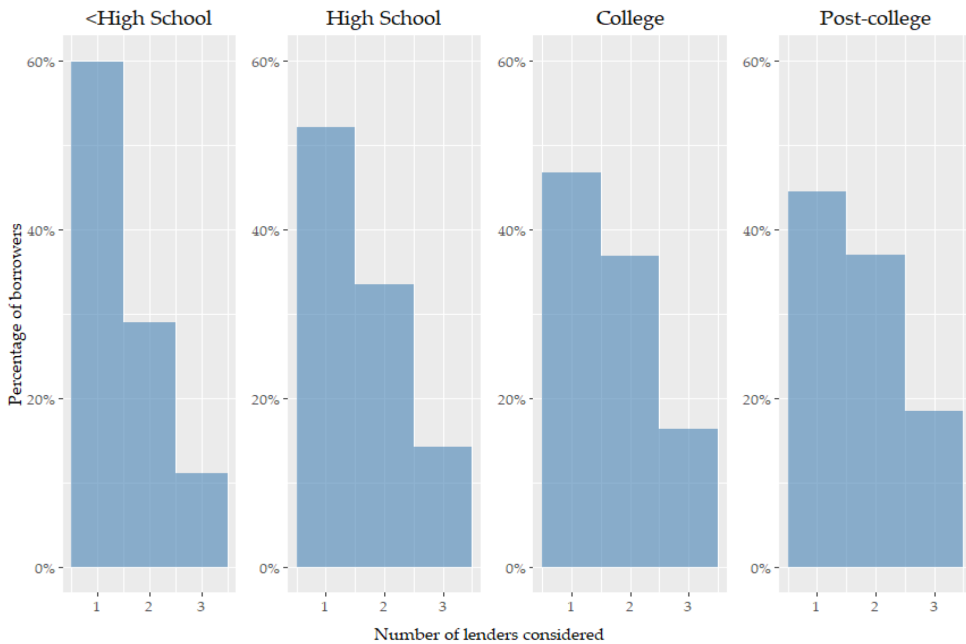


Figure 3: Number of lenders considered upon origination, differences by education level. Source: NSMO data, authors' calculations.

variation effect on the mortgage interest rate. Thirdly, we impute financial literacy from the Survey of Consumer Finances to establish correlation between financial savvyness and the interest rate obtained after searching for the mortgage[4].

Interestingly, almost 70% of the borrowers believe that they would be getting the same interest rate regardless of the choice of the lender. Moreover, 50% consider only one lender/mortgage broker when taking up a mortgage. Consequently, 77% apply to one lender. However, the number of lenders considered varies with education level (Figure 3). In the data, mortgage search is usually done by contacting a lender first and then figuring out the options. This implies that the borrower focuses on the lender type and agrees on repayment schedule after. Borrowers who apply to multiple lenders usually do so in search for better contract terms.

Our latter findings suggest that education and effort simultaneously affect the mortgage interest rate. Hence, the main model assumption is that human capital acts as a proxy for both education and financial knowledge. To support our idea, we control for individual and loan characteristics and show that education may serve as financial responsibility proxy. In Figure 4, the differences in credit score

[4] As we're the first one to be matching the NSMO and the SCF to impute financial literacy score in the NSMO, the imputation will be given more space in the paper.

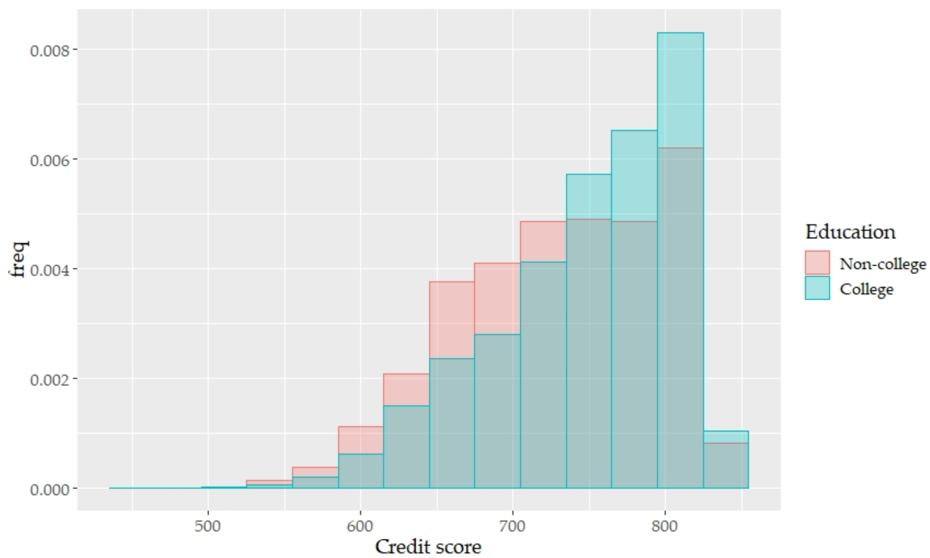


Figure 4: Credit score, difference by education level, NSMO data, authors' calculation.

between college graduates and lower educated borrowers are significant.

3.1.1 Regression comparison

The interest rate is comprised of two components, PMMS determined by borrower's characteristic[5] and the rate spread assigned to each borrower at origination. We combine the two rates and obtain a full interest rate that is our dependent variable in the analysis. Then, we evaluate the effect of search effort and search efficacy implied by the number of lenders considered on the mortgage interest rate.

Since almost $\%50$ of reported mortgages are used in refinancing purposes, we estimate the linear regression separately. Our focus is set on the effects of interaction between the search effort and education of borrowers. Therefore, the interaction coefficients represent the variation in search efficacy with education. Controlling for other demographic characteristics, we argue that highly educated borrowers that consider multiple lenders end up with significantly lower interest rates. Given that we use a novel measure with both cognitive and monetary costs, our estimates account for the unprecedented part of the interest rate dispersion (Table 1).

Next we took an average, college-graduate borrower with income $\$75,000 - \$100,000$, average loan amount of $\$150,000$, and calculated the effect of search effort: the differences in the interest rate are

[5]Freddie Mac's Primary Mortgage Market Survey® (PMMS®) surveys lenders each week on the rates and points for their most popular 30-year fixed-rate, 15-year fixed-rate and other mortgage products.

	interest rate	
	(rate under refinance)	(first origination rate)
#lenders considered: 2	−0.041	0.039
	(0.025)	(0.024)
#lenders considered: 3+	0.050	0.055
	(0.034)	(0.036)
Technical school	0.035^{**}	−0.025
	(0.018)	(0.017)
College	0.014	−0.020
	(0.018)	(0.017)
Post-college	0.024	−0.022
	(0.019)	(0.018)
Technical school; considered 2	0.038	−0.031
	(0.030)	(0.028)
Technical school; considered 3	-0.127^{***}	−0.052
	(0.041)	(0.042)
College; considered 2	0.032	-0.060^{**}
	(0.030)	(0.027)
College; considered 3	-0.118^{***}	-0.079^{*}
	(0.040)	(0.040)
Post-college; considered 2	−0.011	-0.071^{**}
	(0.031)	(0.029)
Post-college; considered 3	-0.133^{***}	-0.129^{***}
	(0.042)	(0.042)
Constant	3.712^{***}	3.649^{***}
	(0.069)	(0.064)
Observations	12,431	14,975
R^2	0.390	0.306
Adjusted R^2	0.388	0.304
Residual Std. Error	0.459 (df = 12392)	0.451 (df = 14936)
F Statistic	208.448^{***} (df = 38; 12392)	173.222^{***} (df = 38; 14936)

Note: Controlled for year-effects, loan and other characteristics. *p<0.1; **p<0.05; ***p<0.01

Table 1: Interest rate under refinancing and for first origination regressed on personal and loan characteristics.

substantial.

Seeing that the mortgage interest rate in fact varies with search effort, it serves us to show which borrower's characteristics affect the amount of search borrowers are willing to take up. Hence, controlling for loan characteristics, we estimate a quasi Poisson model and compare model predictions for different education levels.

	Dependent variable:	
	Number of lenders considered	
	(whole sample)	(mortgage refinance)
age	-0.001^{***}	0.0004
	(0.0002)	(0.0004)
sex	-0.083^{***}	-0.101^{***}
	(0.005)	(0.008)
Education: High-school	0.061^{***}	0.064^{***}
	(0.009)	(0.013)
College	0.096^{***}	0.094^{***}
	(0.009)	(0.013)
Post-college	0.121^{***}	0.108^{***}
	(0.009)	(0.014)
$\$35,000 \leq$ income $\leq 49,999$	0.002	0.006
	(0.013)	(0.021)
$\$50,000 \leq$ income $\leq 74,999$	-0.003	-0.013
	(0.013)	(0.020)
$\$75,000 \leq$ income $\leq 99,999$	-0.004	-0.019
	(0.013)	(0.020)
$\$100,000 \leq$ income $\leq 174,999$	-0.004	-0.032
	(0.013)	(0.021)
income $\geq 175,000$	-0.016	-0.041^{*}
	(0.015)	(0.023)
Constant	0.451^{***}	0.303^{***}
	(0.035)	(0.052)
Observations	39,364	17,810
Note:Controlling for year effects, loan and other demographic characteristics.		*p<0.1; **p<0.05; ***p<0.01

Table 2: Number of lenders considered against education level.

Both sample estimates produce significance with education level coefficients, suggesting that education does affect the search effort in mortgage undertaking. Moreover, other characteristics such as income, employment state, metropolitan area indicator produce either insignificant or substantially lower effect size. Therefore, our count predictions dominantly rely on the level of borrower's education.

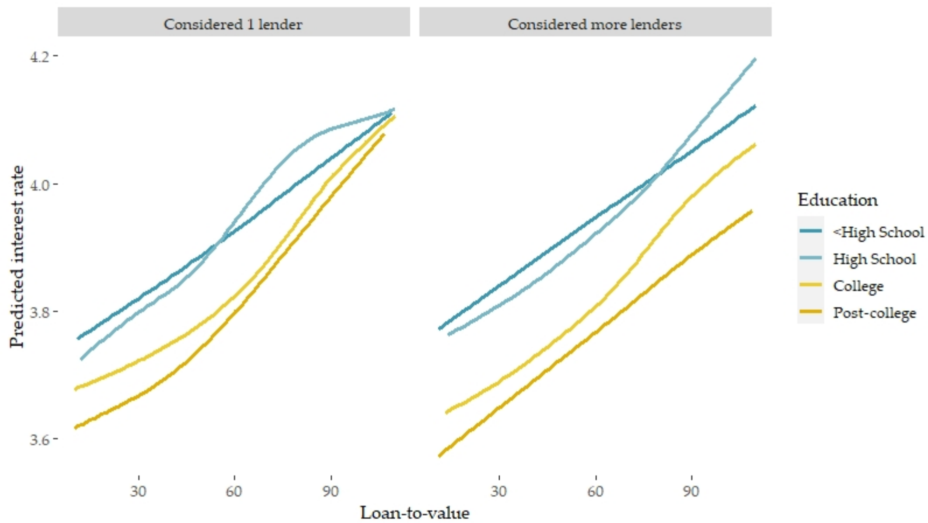


Figure 5: Predicted interest rate, differences in education level and search effort.

Interest rate predictions that are controlled for income, year-effects and other borrower's characteristics show significant discrepancy between the interest rate, both with respect to education and the effort measure (Figure 5). It is worth noting that the differences we estimate account for a part of the interest rate dispersion, whereas the full picture requires data on down-payment amounts, discount points and lender fees. However, the interest rate difference on a typical loan of $\$100,000$ for *low loan-to-value* ratios amounts to $\$210$ yearly or, all else fixed, $\$6,300$ over the mortgage term. Moreover, we find that the number of lenders considered affects the cost of the loan, wherein considering one instead of more lenders yield additional \$1800 increase over the mortgage term. That is, education and search effort may decrease the cost of a standard loan up to \$8000. Our estimates conform to the empirical findings (Keys et al., 2016; Bhutta et al., 2020).

3.2 The Survey of Consumer Finances

Having set grounds for search efficacy differences across education levels, we use additional data source, the Survey of Consumer Finances (SCF) to motivate the fact that human capital in our model serves as a measure of both education and financial knowledge.

Our findings and model assumptions rely on the correlation and functional dependency of financial literacy on education, while controlling for other borrower's characteristics. To this extent, we contribute existing studies on the drivers of financial literacy (Lusardi et al., 2010; Lusardi and Mitchell, 2014) using specific datasets.

For motivating purposes, we use the discrete measures of financial literacy given in the newest two survey waves of SCF data, and estimate the ordered logistic model. Model predictions contain financial literacy likelihoods, evaluated for different borrower's characteristics. Since the focus is set on education, we compare likelihoods between college and non-college respondents. Predicted probabilities show that college graduates answer all of the financial literacy questions with 77% probability, whereas high-school graduates do so with 52% probability[6]. Essentially, taking percentage score from three-questions-answers and plotting the score against highest education level serves as suggestive evidence of correlation (Figure 6).

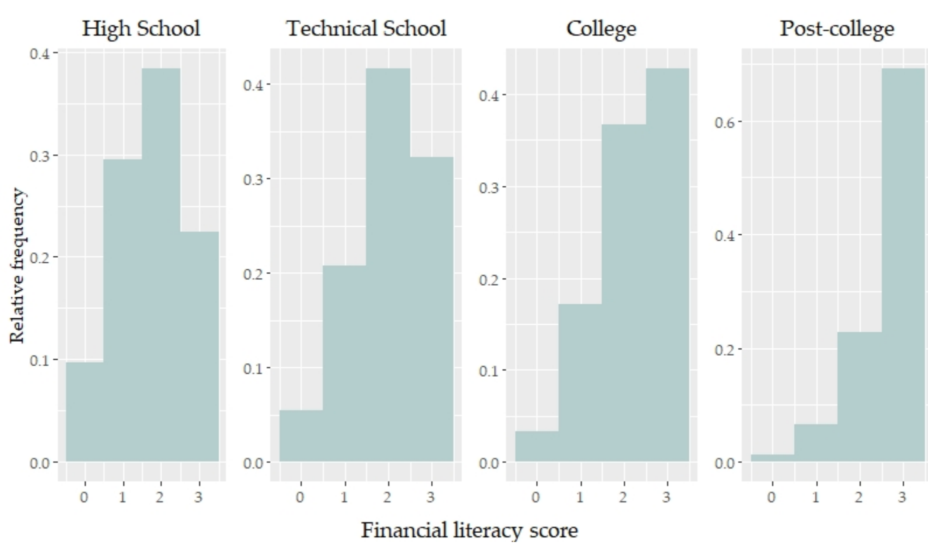


Figure 6: Financial literacy distribution by education level. Source: SCF, 2016-2019, authors' calculations.

Since our model includes financial knowledge accumulation, we provide evidence for our model

[6]Analysis details can be found in the appendix.

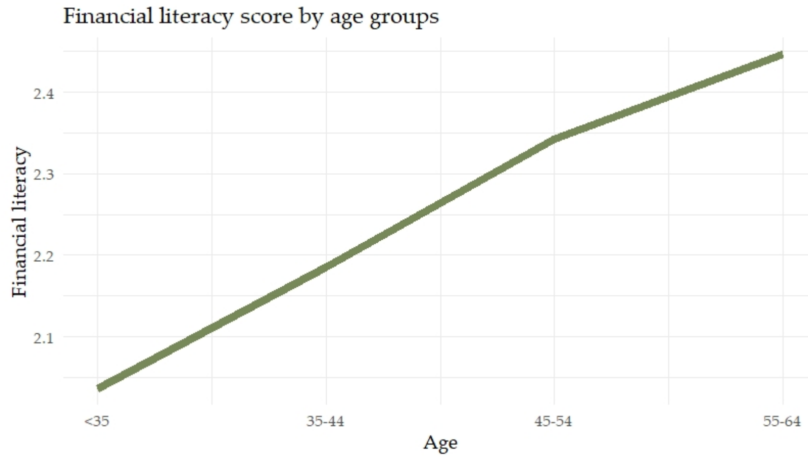


Figure 7: Average financial literacy by age groups.

assumption by plotting financial literacy against age. That is, borrowers develop their financial skills over time.

The SCF also asks whether the respondent has ever refinanced a mortgage on primary residence. We use these responses and evaluate binomial logistic model, with individual search effort elicited through self-reported amount of shopping in borrowing decisions. Controlling for demographic characteristics, both education and financial literacy imply higher likelihood of mortgage refinancing. Interestingly, the way of gathering information shows to be significant - getting information from friends at work incentivizes borrowers to refinance (Table 3 of the appendix). Probability differences with respect to financial knowledge amount to just below 10% across all education levels. However, higher education level implies substantial increase in refinance probabilities, even with income controls. Ultimately, the SCF evidence supports our assumptions with respect to search costs variation with in human capital level, as effects of education and financial literacy are separately unidentifiable[7]. More details on our analysis can be found in the appendix.

[7]Separating financial literacy from education has been subject to discussions regarding cognitive skill inheritance and person fixed-effects(Mitchell and Lusardi, 2015; Lusardi and Mitchell, 2014).

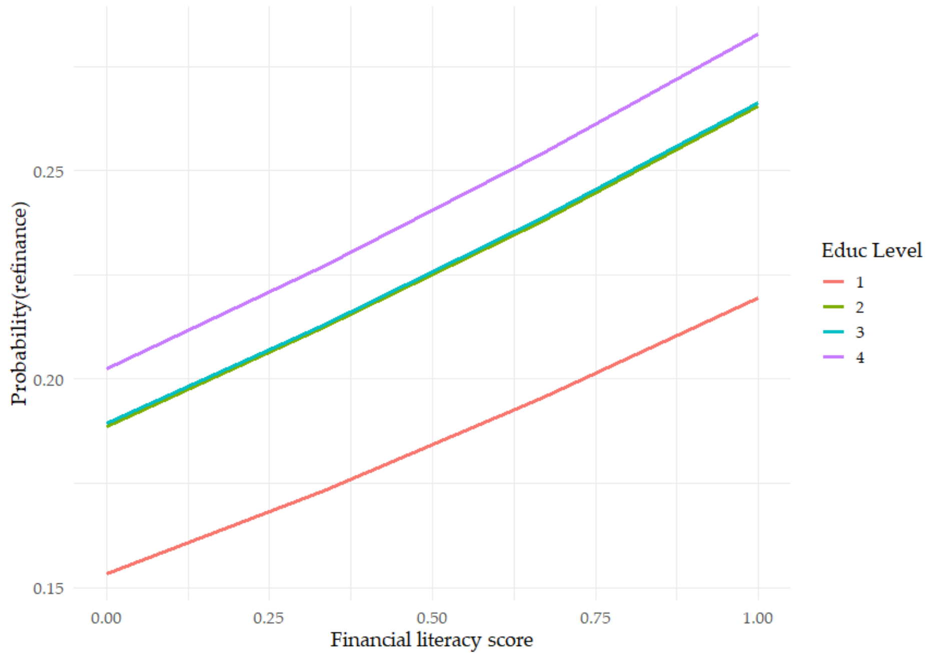


Figure 8: Refinance probabilities predictions across education levels and financial literacy scores.

3.3 Imputing financial literacy into the NSMO data

Instead of relying on education measures only, we impute borrower's financial literacy by comparing individual characteristics with the SCF respondents. To our knowledge, we are the first study to be matching these two datasets based on observables. Our first-hand solution (under revision) is to use the Random Forrest Algorithm that uses machine learning techniques to train the classification model that eventually predicts the borrowers financial literacy in the NSMO dataset.

The algorithm is given with following steps:

1. n random records for a random forest are sampled from the data set having k records.
2. For each sample, individual decision trees are constructed.
3. Each decision tree generates an output.
4. Based on Majority Voting (Average in case of regression) on all outputs generated in previous step, final output is constructed.

Using the imputed financial literacy, we argue that financial literacy correlates with the number of lenders considered, and that it does so in a similar way to education. That is, re-plotting the bar plot from the NSMO data analysis (3) produces similar conclusions - borrowers that know their way about their finances consider more lenders, on average 9. That is, while more than 60% of high-school graduates consider only one lender (9, first bar plot), the breakdown for college graduates is more disperse. That is, 45% of college graduates consider one lender and approximately 20% consider three or more lenders (9, fourth bar plot). Therefore, financially savvy borrowers consider more options while searching for mortgage.

Figure 9: Number of lenders considered by the financial literacy score.

4 Dynamic Model With Heterogeneous Agents

Following our data findings, we develop a mortgage search model with heterogeneous agents (borrowers). The model builds on the idea of search efficacy variation by education level from our estimates. There is a continuum of agents that solve an infinite horizon problem in continents time. Agents are heterogeneous with respect to initial skills[8] $f_0 \sim \Gamma(f_0)$ that ultimately implies their wage. Initially, all agents are assigned a mortgage which can be refinanced later on. Of course, more general version would include first origination costs, but given the increasing number of home ownership rates, we see it as a good first version of the model. Moreover, agent with ability f_0 repays mortgage $M = \xi f_0$.

When deciding to refinance, borrowers search for better interest rate among continuum of lenders.

[8]We still somewhat did not pin-point the name for this. Suggestions are always appreciated.

Lender type is given with the interest rate on debt held outside[9] Exogenous r^L is subject to generalizations, typical interest rate models etc., and would eventually serve for counterfactual analysis.

4.1 Financial skills accumulation

Each period, agent may invest in financial skills $0 \leq i \leq 1$.[10] Benefits of attaining skills may reflect in productivity (i.e., wage), but at this point they solely guarantee credit score increase. Given investment i, skill f evolves according to

$$\dot{f} = \frac{\mu}{\eta}(if)^{\eta}.$$

The remaining amount of time $(1-i)$ agent with productivity $z \sim \Phi(z)$, earns wage $w = (1-i)zf$. Later on, we would like to assume some kind connection such as:

$$w = (1-i)z, \quad dz = \mu(f_t)dt + \sigma(f_t)d\mathcal{B}_t.$$

Also, as previously stated, we would aim for the lender repayment interest as

$$dr_t = -\eta_r(r_t - \bar{r}) + \sigma_r\sqrt{r_t}d\mathcal{B}_t,$$

and obtain a variation with respect to low/high interest rate times. Given that all the data time span incorporated low interest rate period, we'd like to see what the implications would be for potential upcoming increase in the interest rates.

4.2 Refinancing - decision and options

Each period, the borrower chooses whether to refinance, and consequently endures effort that guarantees more offer arrivals. We borrow on-the-job search framework to model this, so the search intensity translates into higher offer arrival number. Additionally, borrowers may work on their financial skills to ensure better credit score. Ultimately, search costs are an increasing function of the search effort s, but vary with financial skill f. I.e., the cost of search changes slope depending on how skilled the

[9] We refer to it as lender repayment rate.

[10] Motivated by increasing financial literacy profile over the lifecycle (Figure 7).

borrower is:

$$c(s, f) = c_0 \frac{s^{1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}} \frac{1}{f}.$$

This assumption will allow us to match the amount of consideration preceding the formal application process that we see in the data. Thus, financial skills investment implies benefits attributed to higher labor earnings and lower search cost.

4.2.1 Probability of a financial shock

As stated previously, accumulating financial skills benefits the credit score [11]. We approximate the borrower's credit score with the probability of facing a financial shock. Since getting a financial shock corresponds to losing resources, the probability of facing a financial shock approximates the credit score in our model. Therefore, assumption is that the default probability $\delta(f)$ decreases with financial skills. Moreover, $\delta(f)$ is observable both to the client and the lender. Thus, skills serve the borrower in terms of interest rate span. At the same time, the lender incorporates default event in setting the mortgage rate offer. If the borrower defaults, she loses the house and needs to start paying rent costs, parameterized by κ.

4.2.2 Mortgage interest rate - bargaining

We assume bargaining procedure as in (Dey and Flinn, 2005; Cahuc et al., 2006), which is currently used as a benchmark in the labor literature, and serves the interest rate negotiation process we see in the data. Borrowers bargain for their mortgage rate with their respective lenders. This setting corresponds to NSMO survey, where more than 80% of borrowers state that they usually contact the lender first, then figure out the final interest rate in the second part of the mortgage deal. Moreover, dominant proportion of borrowers choose lender's reputation as highly important when browsing for mortgage options. In this way, lenders take observable skills of their borrowers into account and attract them as prospective clients. Lastly, even though both commit to the contract, the contract is negotiable later on, if the borrower meets a new lender that offers a better interest rate conditional on her (accumulated) financial literacy.

Let $W = W(f, r(f, r^L, (r^L)'), (r^L)')$ denote the flow value for the borrower with financial skills

[11] Eliciting what we show using NSMO credit score estimates.

f and current mortgage rate r. The mortgage rate is a function of skills, current lender type r^L and prospective lender type $(r^L)'$. Given the new lender's type $(r^L)'$, the borrower will either stay with the current lender or switch to a new one. Should the borrower stay, the contract is renegotiated conditional on the new lender type, which serves as an outside option for the borrower.

Meeting a new lender starts a game between the borrower and two lenders. If the borrower meets a lender of low-enough type (correspondingly, lower interest rate option), she uses that as a threat to her old lender that can offer a lower interest rate, while still retaining profits from the contract. If the lender cannot go any lower with the interest rate offer, the borrower switches to a new lender $(r^L)'$ and bargains over the new interest rate.

Specifically, $W = W(f, r(f, r^L, (r^L)'), (r^L)')$ is flow value of the borrower with financial skills f, who currently repays mortgage M with interest rate r to lender $r^L \sim \Phi$. $L(f, r(f, r^L, (r^L)'), (r^L)')$ denotes the value of the lender of type $r^L \sim \Phi$, whereas $J(f, r(f, r^L, (r^L)'), (r^L)')$ denotes the joint value of the match ($J = W + L$). As mentioned previously, meeting a new lender $(r^L)'$ starts a game that involves the borrower and both lenders. Similar to Dey and Flinn (2005) and Cahuc et al. (2006), the game ends with the mortgage rate $r(f, r^L, (r^L)')$ that solves

$$W(f, r(f, r^L, (r^L)'), (r^L)') = \underbrace{J(f, r^L)}_{\text{threat point}} + \beta \big[\underbrace{J(f, (r^L)') - J(f, r^L)}_{\text{surplus from new match}} \big].$$

Thus, borrower's value is determined via Nash bargaining with a bargaining power parameter β. Resulting value is therefore a convex combination between the old and a new interest-rate driven contracts:

$$W(f, r(f, r^L, (r^L)'), (r^L)') = (1 - \beta) J(f, r^L) + \beta J(f, (r^L)'). \tag{1}$$

Denote with $Q = Q(f, r, r^L)$ the type of a lender such that agent's current value is

$$W(f, r, r^L) = \beta J(f, r^L) + (1 - \beta) J(f, Q). \tag{2}$$

Conditional on the type Q that represents the threshold value of switching (Cahuc et al., 2006), there are three possible outcomes:

- staying with the current lender if $(r^L)' > r^L$,

- refinancing with current lender if $Q < (r^L)' < r^L$,
- refinancing with new lender if $(r^L)' < Q$.

Therefore, if the new lender is of a higher rate type $(r^L)' > r^L$, the agent stays with the current lender. Although, it is possible that the mortgage contract is renegotiated, resulting in the interest rate given by the minimum of current rate r and $r'(f, (r^L)', r^L)$. Outside offer and current switch places in the bargaining process, as outside offer in the renegotiation with the same lender is the threat of leaving with current lender, and outside offer in the refinancing with new lender is threat of staying with the current lender.

!UNDER CONSTRUCTION

4.2.3 Matching borrowers and lenders

Let V denote vacant mortgage lenders and set the mass of agents to be 1. Then the rate at which mortgage contracts are created is defined with $m = M(d, v)$, where d is the default rate and v is the mortgage vacancy rate. As usual in search literature, we assume that matching function M is increasing in both arguments and homogeneous of degree 1. Therefore, mortgage finding and (mortgage vacancy) filling are given with

$$\lambda = \lambda(\theta_t) = \frac{M(d_t, v_t)}{d_t} = M(1, \theta_t) \quad \text{and} \quad q = q(\theta_t) = \frac{M(d_t, v_t)}{v_t} = M(\theta_t^{-1}, 1). \tag{3}$$

Notice that both mortgage finding and filling are function of θ_t - the mortgage market tightness and that $\lambda = \theta q$.[12]

[12] In the rest of derivations we drop subscript t since we are interested in steady state and to simplify notation.

4.3 Value Functions

4.3.1 Agent's value

We derive the agent's value in the appendix E.1, and do an overview of the main equations that are key for the model solution. The agent's value is given with

$$\rho W(r,\cdot) = u + \frac{\partial W}{\partial f}\dot{f} + \frac{\partial W}{\partial t} + \delta(f)(D - W(r)) \\ + \lambda s \int_{\underline{r}}^{r^L} (W(r') - W(r))d\Phi(r') + \lambda s \int_{r^L}^{Q} (W(r') - W(r))d\Phi(r'), \tag{4}$$

where u represents agent's utility net of the mortgage repayment and search cost:

$$u = (1-i)fz - rM - c(s,f),$$

and $Q = Q(f, r, r^L)$ defined with equation (2).

The agent's value incorporates two values driven by the decision to refinance. First, $W(r')$ in the first integral in (4) is the bargain achieved in negotiations with a new lender (1), thus equals

$$W(r') = W(f, r'(f, r^L, (r^L)'), (r^L)') = J(f, r^L) + \beta\big[J(f, (r^L)') - J(f, r^L)\big],$$

and $W(r')$ in the second integral is a result of staying with the old lender, albeit with a renegotiated interest rate (2). The renegotiated rate would be lower than the starting, but still higher than the lenders repayment rate. Hence, similar to the first case,

$$W(r') = W(f, r'(f, (r^L)', r^L), r^L) = J(f, (r^L)') + \beta\big[J(f, r^L) - J(f, (r^L)')\big].$$

Even though the agent "affects"[13] the amount of lenders at disposition by intensifying their search, the prevailing measure is λ, the contract arrival rate. Naturally, λ is implied by the mortgage market tightness, that is in turn implied by the number of mortgage contracts available. That is $\lambda = \lambda(\theta)$, where θ is the market tightness. In sum, the value of the agent is equal to the sum of current wage net of cost and mortgage payments and potential changes either due to the default rate $\delta(f)$ or due to

[13]Everything is derived from discrete time, so the intuition works better in the appendix

mortgage refinancing.

4.3.2 Lender's value

The value for lender r^L is given with

$$\rho L(r,\cdot)=\pi_t+\frac{\partial L}{\partial f}\dot{f}+\frac{\partial L}{\partial t}-\lambda s\Phi(r^L)+\lambda s\int_{r^L}^{Q}(L^*(r'')-L(r))d\Phi(r'')-\delta(f)L, \tag{5}$$

where $\pi_t=rM-r^LM$ denote lender's profits from the difference in mortgage repayment rM acquired and repayment r^LM made to the exogenous lender. Further, $L^*(r'')$ is implied by the bargaining outcome (2) and equals

$$(1-\beta)(J(f,r^L)-J(f,(r^L)')). \tag{6}$$

Equations (5) and (6) imply that lender receives the same profits in case that agent meets a new lender that charges higher price, i.e. is of type $(r^L)'>r$. In contrast, if the agent finds a new lender that charges a lower price, i.e. is of type $r^L<(r^L)'<r$, the agent stays with the current lender. As stated previously, they potentially bargain on the new rate, which is lower than current rate r, but still ensures positive profits for current lender r^L. Consequently, the agent refinances the mortgage with current lender (which in turn faces reduction in profits). Conditional on finding a better mortgage rate offer with new lender $(r^L)'<r^L$, the agent refinances with the new lender, and the current lender no longer receives profits in next period. In case of agent's default, current lender looses the profits as well.

Combining two values, agent's value (4) and lender's value (5) we derive joint value J that satisfies

$$\rho J=u+\pi_t+\frac{\partial J}{\partial f}\dot{f}+\frac{\partial J}{\partial t}+\lambda s\int_{\underline{r}}^{r^L}(J(r')-J)d\Phi(r')+\delta(f)(D-J), \tag{7}$$

where the $\int_{r^L}^{r}$ disappears since $W(r')$ and $L^*(r'')$ in sum give $J(r)$ that cancels out. Note that joint value (7) does not depend on r and that the only choice variables are i and s. In this respect, one can think of s as the amount of the search effort that is bargained over. That is, if the lender offers lower mortgage rate, agent can offer lower search effort in the future and thus reduce her chances of refinancing with a different lender (by lowering the arrival of outside offers voluntarily).

4.3.3 Default value

The agent's value in default is given with

$$\rho D = (1-i)zf - c(s,f) - \kappa + \frac{\partial D}{\partial f}\dot{f} + \frac{\partial D}{\partial t} + \lambda s\phi \int_{\underline{r}}^{\bar{r}} (\widetilde{W}(r') - D)d\Phi(r'). \tag{8}$$

The value of default incorporates two exogenous parameters, κ and ϕ. $\widetilde{W}$ is the bargaining outcome in case of negotiation with the sole agent's outside option is to take the highest rate attainable, $\bar{r}$. Exogenous $\phi < 1$ captures the severity of entering into the mortgage market upon default[14]. Φ difficulties in new mortgage origination, i.e., agent has to prepare all documents and find a house, whereas for refinancing that is not necessary as the lender already has all documents. The last difference is that agent's after default no longer have mortgage payments rM, but agent faces cost κ. Costs κ in our model could be connected to costs of rent.

4.4 Equilibrium

In order to define equilibrium we impose the free entry condition. Let $g_t(f, r^L)$ denote the distribution of non-defaulted borrowers at time t, over human capital and mortgage rates. Correspondingly, the distribution defaulted agents is denoted with $g_t^d(f)$. Finally, d denotes the aggregate default rate. The free entry condition implies the cost that is equal to lender's share of surplus, i.e.

$$\begin{aligned} c^{FE} = (1-\beta)q(\theta)\int_{\underline{r}}^{\bar{r}} \Bigg(d \int_f J(t,f,r^L)dg_t^d(f) \\ + \phi(1-d)\int_f \int_{\underline{r}}^{r^L} \big(J(t,f,r^L) - J(t,f,(r^L)')\big)dg_t(f,(r^L)')\Bigg)d\Phi(r^L). \end{aligned} \tag{9}$$

Law of motions for the market distributions in imply the Kolmogorov Forward Equation (KFE). First, we set $J(t,f,\bar{r}) = D(t,f)$, i.e, the model assumes that the value of renting is equal to paying the mortgage at the highest possible rate. Given this assumption, the distribution of agents over time,

[14]We plan to set f to a fixed outside value upon default. This way we account for a decrease in the credit score and do not need to set κ due to immediate change in wage.

mortgage rates, and accumulated financial literacy $g_t(f, r^L)$ solves the following KFE,

$$\begin{aligned}\frac{\partial g_t(f,r^L)}{\partial t} &= -\big(\delta(f) + \lambda(\theta)s(f,r^L)\Phi(r^L)\big)g_t(f,r^L) - \frac{\mu}{\eta}(i(f,r^L))^{\eta}\frac{\partial g_t(f,r^L)}{\partial f} \\ &\quad + \phi\lambda(\theta)s^d(f)\frac{g_t^d(f)}{nd} + \lambda(\theta)s(f,r^L)\int_{r^L}^{\bar{r}} g_t(f,(r^L)')d(r^L)', \quad \forall t \text{ and } r^L \leq \bar{r}, \qquad (10)\end{aligned}$$

where $nd = 1 - \int g_t^d(f)$ are mortgage repaying agents (non-defaulted), s^d is the policy for defaulted agents, and for default the following equation holds

$$\frac{\partial g_t^d(f)}{\partial t} = -\lambda(\theta)\phi s_d(f)g_t^d(f) + \int_{r^L} \delta(f)g_t(f,r^L)dr^L. \qquad (11)$$

To summarize, flow-out in the equation (10) is due to agents who defaulted on current mortgage or agents who refinance the mortgage with more preferable lender ($(r^L)' < r^L$) and mortgage rate. Flow-in is due to agents who take the mortgage after earlier default or agents who previously had less preferable mortgage rate and refinance with lender r^L. Similarly, flow-out from default is due to agents who find a lender, and flow-in is due to agents who default with current lender.

4.4.1 Definition

The stationary recursive equilibrium consists of a set of values $\{W(f,r^L), J(f,r^L), D(f)\}$; a set of policy functions $\{s(f,r^L), i(f,r^L), s^d(f), i^d(f)\}$; a distribution over financial literacy and mortgage rates $g(f,r^L)$; a distribution over financial literacy in default $g^d(f)$; a set of prices $\{r(f)\}$; financial literacy stock per agent f; and mortgage market tightness θ such that:

- Mortgage repayment:

  Given prices and mortgage market tightness, decision rules $\{s(f,r^L), i(f,r^L)\}$ maximizes joint value (7).

- Default:

  Given prices and mortgage market tightness, decision rules $\{s^d(f), i^d(f)\}$ maximizes default value (8).

- Consistency of stationary distribution: Given the decision rules and tightness, the distributions $g(f,r^L)$ and $g^d(f)$ satisfy equations (10) and (11).

- Mortgage market:

  The mortgage finding and filling probabilities are given by tightness according to (3). Default can be calculated as: $d = \int_f g^d(f)df$. The measure of vacancies is then given by $v = \theta d$.

- Free entry:

  Given prices, tightness, and the stationary distribution, free entry condition is given with (9)

- Mortgage rate setting:

  Given an allocation, mortgage repayment rates are determined by the value of a agent (4) and the wage policies (1)–(2) under the optimal financial literacy investment and search effort policies.

4.5 Numerical Solution

coming soon

4.6 Calibration

5 Conclusion

References

Agarwal, S., Grigsby, J., Hortaçsu, A., Matvos, G., Seru, A., and Yao, V. (2020). Searching for approval. Working Paper 27341, National Bureau of Economic Research.

Agarwal, S., Rosen, R. J., and Yao, V. (2016). Why do borrowers make mortgage refinancing mistakes? *Management Science*, 62(12):3494–3509.

Alexandrov, A. and Koulayev, S. (2018). No shopping in the u.s. mortgage market: Direct and strategic effects of providing information. Working Paper 2017-01, Consumer Financial Protection Bureau Office.

Allgood, S. and Walstad, W. B. (2016). The effects of perceived and actual financial literacy on financial behaviors. *Economic inquiry*, 54(1):675–697.

Andersen, S., Campbell, J. Y., Nielsen, K. M., and Ramadorai, T. (2020). Sources of inaction in household finance: Evidence from the danish mortgage market. *American Economic Review*, 110(10):3184–3230.

Argyle, B., Nadauld, T. D., and Palmer, C. (2020). Real effects of search frictions in consumer credit markets. Technical report, National Bureau of Economic Research.

Bartlett, R., Morse, A., Stanton, R., and Wallace, N. (2022). Consumer-lending discrimination in the fintech era. *Journal of Financial Economics*, 143(1):30–56.

Bhutta, N., Fuster, A., and Hizmo, A. (2020). Paying Too Much? Price Dispersion in the U.S. Mortgage Market. Finance and Economics Discussion Series 2020-062, Board of Governors of the Federal Reserve System (U.S.).

Cahuc, P., Postel-Vinay, F., and Robin, J.-M. (2006). Wage bargaining with on-the-job search: Theory and evidence. *Econometrica*, 74(2):323–364.

Campbell, J. Y. (2013). Mortgage market design. *Review of finance*, 17(1):1–33.

Chatterjee, S., Corbae, D., Dempsey, K. P., and Ríos-Rull, J.-V. (2020). A quantitative theory of the credit score. Technical report, National Bureau of Economic Research.

Dey, M. S. and Flinn, C. J. (2005). An equilibrium model of health insurance provision and wage determination. *Econometrica*, 73(2):571–627.

Gathergood, J. and Weber, J. (2017). Financial literacy, present bias and alternative mortgage products. *Journal of Banking and Finance*, 78:58–83.

Gurun, U. G., Matvos, G., and Seru, A. (2016). Advertising expensive mortgages. *The Journal of Finance*, 71(5):2371–2416.

Keys, B. J., Pope, D. G., and Pope, J. C. (2016). Failure to refinance. *Journal of Financial Economics*, 122(3):482–499.

Koszegi, B. (2014). Behavioral contract theory. *Journal of Economic Literature*, 52(4):1075–1118.

Lusardi, A. (2019). Financial literacy and the need for financial education: evidence and implications. *Swiss Journal of Economics and Statistics*, 155(1):1–8.

Lusardi, A. and Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of economic literature*, 52(1):5–44.

Lusardi, A., Mitchell, O. S., and Curto, V. (2010). Financial literacy among the young. *Journal of consumer affairs*, 44(2):358–380.

Mayer, C., Piskorski, T., and Tchistyi, A. (2013). The inefficiency of refinancing: Why prepayment penalties are good for risky borrowers. *Journal of Financial Economics*, 107(3):694–714.

Mitchell, O. S. and Lusardi, A. (2015). Financial literacy and economic outcomes: Evidence and policy implications. *The journal of retirement*, 3(1):107–114.

Van Rooij, M., Lusardi, A., and Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial economics*, 101(2):449–472.

Woodward, S. E. and Hall, R. E. (2012). Diagnosing consumer confusion and sub-optimal shopping effort: Theory and mortgage-market evidence. *American Economic Review*, 102(7):3249–76.

Yannelis, C. and Zhang, A. L. (2021). Competition and selection in credit markets. Working Paper 29169, National Bureau of Economic Research.

A Motivating Findings From SCE

Motivating findings based on the data from the Survey of Consumer Expectations. Figure 10 that the largest mass of non-informed households is from the lowest income group. Moreover, figure shows that mass of non-informed households decreases with higher income. Figure 11 shows that the households from the lowest income group, have highest debt to income ratios. In addition, Figure 12 shows that largest shares of highest debt to income ratios are in the lowest part of the income distribution. To summarize, findings from these figures imply that most exposed households are the ones who are the least informed about credit possibilities.

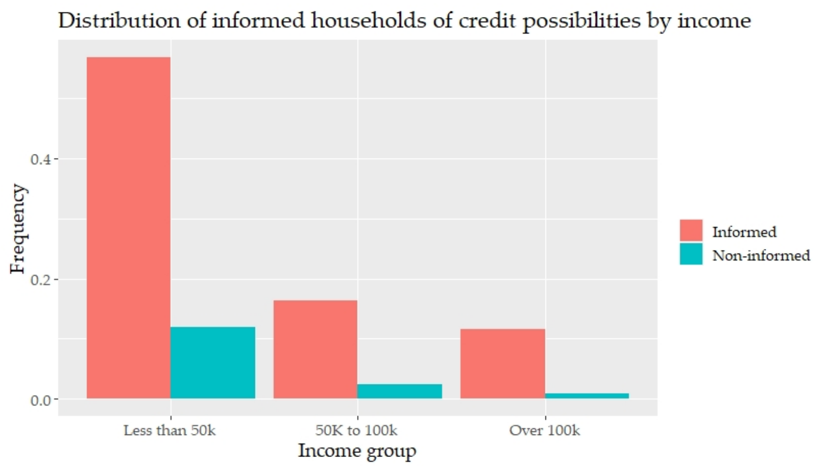


Figure 10: Share of non-informed households by income group. Source: SCE, authors' calculation.

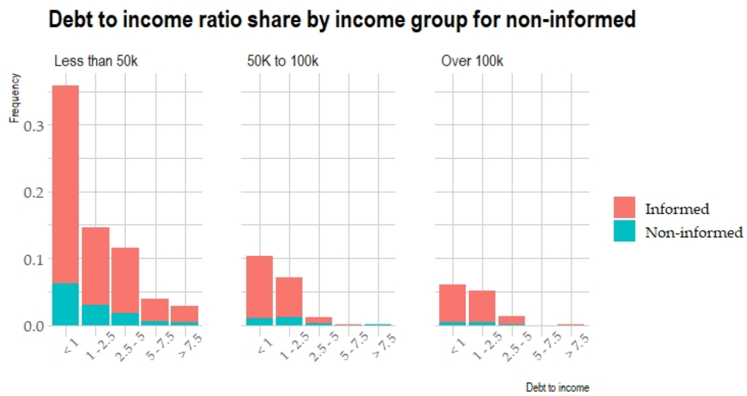


Figure 11: Share of non-informed households for each debt to income level over the income distribution. Source: SCE, authors' calculation.

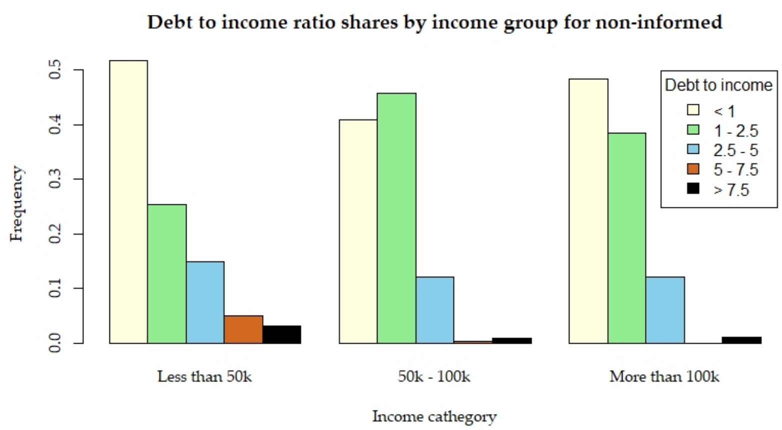
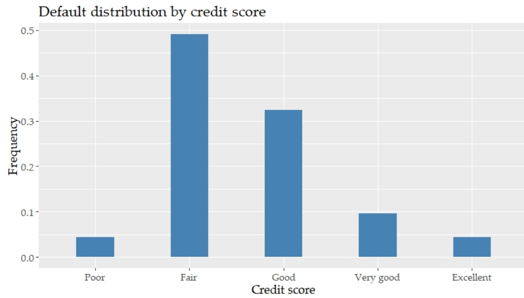


Figure 12: Debt to income ratio distributions for each income group. Source: SCE, authors' calculation.

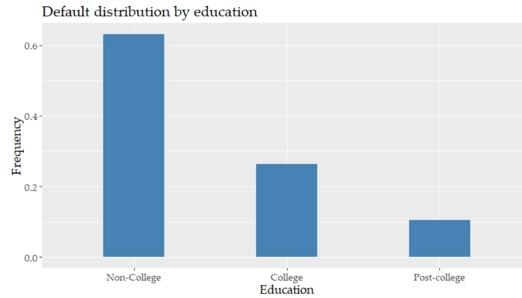
B NSMO

B.1 Who are agents who default on mortgage

Distributions in Figure 13 shows that households who default on they mortgage and face bankruptcy are associated with lower credit score and lower education. Only exemption is the lowest credit score, but households' mortgage request with "Poor" credit score are usually denied.



(a) Default by credit score.



(b) Default by education.

Figure 13: Share of households who default by credit score and education. Source: NSMO, authors' calculation.

C SCF data, financial knowledge and education

As we discuss the effects of human capital accumulation both on labor earnings and financial knowledge simultaneously, we set grounds for this assumption using SCF data. First, to motivate investing in, and accumulation of financial skills in the model, we look at the average financial literacy score over the lifecycle. Figure 7 shows increasing average financial literacy score by age groups. Further, we argue that education remains a significant measure of worker's financial knowledge, once controlled for other personal characteristics. Specifically, we use the measure of financial literacy given in the 2016 and 2019 waves of SCF survey, and build an ordered logit model to acquire predictions for different education levels. Keeping other covariates at the mean of the respective distribution, we render higher literacy level as more likely with more educated respondents. Moreover, we control for self-reported measures of financial knowledge as it has been shown to co-effect important financial decisions (Allgood and Walstad, 2016). The objective measure of financial literacy corresponds to the number of questions answered correctly (Lusardi, 2019).

Even though we focus on the interaction between education and financial literacy, first suggestive evidence show significant differences in refinancing decisions based on financial knowledge only. That is, while 40% of financially knowledgeable respondents refinance their mortgages, only 12% of their counterparts decide to refinance. Using a simple binary regression, we endorse the idea of thinking about human capital as more than just education. Namely, the decision to refinance is affected by financial knowledge, even while for controlling for education, mortgage monthly payments and income. Moreover, the effect is of larger size than for other characteristics. Therefore, our view of human capital complements refinancing patterns in the data.

Table 3: Binary regression, mortgage refinance choice.

	Dependent variable:
	Ever refinanced a mortgage
Financial literacy score	0.441***
	(0.062)
Education: High-school	0.252***
	(0.071)
College	0.256***
	(0.070)
Post-College	0.338***
	(0.070)
mortgage payment/monthly income	7.366***
	(0.127)
Borrowing shopping amount : Moderate	0.338***
	(0.049)
Great deal	0.424***
	(0.052)
Information source: Friend at work	0.059**
	(0.029)
Internet	0.085***
	(0.032)
Financial advisor	0.133***
	(0.028)
Wage income quartile: q_2	0.507***
	(0.052)
q_3	0.983***
	(0.051)
q_4	1.284***
	(0.052)
Constant	−5.421***
	(0.122)
Observations	38,850
Akaike Inf. Crit.	31,455.860
Note: Controlled for demographic characteristics and survey wave.	*p<0.1; **p<0.05; ***p<0.01

Breaking down financial literacy score across the sample yields fairly intuitive results. While locations in the income and asset distributions imply higher financial literacy, education level remains significant as a predictor of financial knowledge. Specifically, controlling for asset position we abstract from *learning-by-doing* explanations and argue that education still explains financial knowledge.

Table 4: Financial Literacy level

	Financial literacy level
mortgage payment/monthly income	-0.237^{***}
	(0.059)
non-white	-0.361^{***}
	(0.019)
Female	-0.455^{***}
	(0.026)
Technical school	0.204^{***}
	(0.031)
College	0.552^{***}
	(0.032)
Post college	1.020^{***}
	(0.034)
Income percentile: $20 - 39.9$	0.024
	(0.029)
$40 - 59.9$	0.039
	(0.031)
$60 - 79.9$	0.144^{***}
	(0.035)
$80 - 89.9$	0.337^{***}
	(0.044)
$90 - 100$	0.614^{***}
	(0.049)
Asset quartile: q_a^2	0.162^{***}
	(0.026)
q_a^3	0.429^{***}
	(0.033)
q_a^4	0.871^{***}
	(0.044)
Observations	52,931

Note: Controlled for family structure, age and survey wave. *p<0.1; **p<0.05; ***p<0.01

D Stylized Two Period Model

In this section we develop two period stylized model to motivated by empirical findings from the previous section. Economy consists of one agent with ability f_0 and a mortgage $M = \xi f_0$ and two lenders. The agent initially is matched with a lender with a type r^L (the rate at which lender borrows money) repays mortgage at repayment rate $r \geq r^L$. Condition $r \geq r^L$ ensures non-negative profits for a lender. Further, the agent decides how much time to invest in financial literacy $i \leq 1$ such that

$$f_1 = f_0 + \frac{\mu(if_0)^\eta}{\eta}.$$

Moreover, the agent chooses search intensity s such that s maximizes joint surplus under cost

$$c(s) = c_0 \frac{s^{1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}}.$$

Thus, search intensity is the outcome of the bargaining between the agent and the lender.

In the second period, agent defaults with probability $\delta(f_1)$. We assume default probability dependence on the accumulated financial literacy that serves as a signal to lenders that the agent is financially responsible. Therefore, higher amount of accumulated financial literacy benefits both agent and the lender. First benefit for the agent is that it brings higher wage in the second period and the second is that it reduces default probability. Lender can view this as increased financial responsibility and thus reduce the rate at which the agent repays the mortgage. Further, as mentioned, is the second period, the agent receives wage f_1 and repays the rest of the mortgage M_1. According the the search intensity, agent can remain at the same lender or switch to repayment with the new lender $(r^L)'$. New lender's type is rate $(r^L)' < r^L$ and thus the new lender can offer a better rate to the agent. Therefore, the repayment in the second period is the following:

- With probability s agent is matched with a better lender $(r^L)' < r^L$. The repayment rate at the new lender is determined as an outcome of the bargaining where β is agent's bargaining power.
- With probability $1 - s$ agent remains at the old lender and repays the rest of the mortgage at the same rate as in the first period.

f_0	μ	η	r^L	$(r^L)'$	r	β	ξ	γ	λ	c_0
1	0.01	0.1	1.1	1.05	1.11	0.5	0.3	1	5	0.1

Table 5: Parameters for the stylized two period model.

All together, agent's value for 2 periods is given with

$$
\begin{aligned}
W(i,s) = \; & \underbrace{(1-i)f_0}_{1^{st}\text{period wage}} - \underbrace{\frac{r^2}{r+1}M}_{1^{st}\text{annuity}} - \underbrace{c(s)}_{\text{cost of search}} + \\
& + \underbrace{(1-\delta(f_1))s}_{\text{no default and a new match}} \Big[\underbrace{f_1 - [(1-\beta)r^L + \beta(r^L)']M_1}_{\text{wage-}2^{nd}\text{ annuity at new rate}} \Big] + \\
& + \underbrace{(1-\delta(f_1))(1-s)}_{\text{no default and no new match}} \; [\; \underbrace{f_1 - \frac{r^2}{r+1}M}_{\text{wage-}2^{nd}\text{ annuity at old rate}} \;],
\end{aligned}
$$

where $M_1 = \frac{r}{r+1}M$, is amount left to repay in the second period. The second participant who's value we model is the lender that the agent is matched in the first period. In first period the value for the ledner is given with the difference between the payment that lender receives from the agent. In the second period, the value is the same, but only if the agent does not switch to the new lender and if the agent does not default. Therefore, lender's value is given with

$$
\begin{aligned}
L(i,s) = \; & \underbrace{\frac{r^2}{r+1}M}_{1^{st}\text{annuity}} - \frac{(r^L)^2}{r^L+1}M + \\
& + \underbrace{(1-\delta(f_1))(1-s)}_{\text{no default and agent stays}} \Big[\underbrace{\frac{r^2}{r+1}M}_{2^{nd}\text{annuity}} - \frac{(r^L)^2}{r^L+1}M \Big].
\end{aligned}
$$

To solve the model we maximize the joint surplus by the agent and the lender that is given with

$$J(i,s) = W(i,s) + L(i,s).$$

In addition, we hold fixed rate $r \geq r^L$, and assume that the default is exponentially distributed. For this motivating model, we set parameters of the model as in the Table 5.

Resulting $s(i)$ and $i(s)$ are presented in Figure 14. As the investment in financial literacy increases, the search intensity increases as well which (in our specification) indicates that the benefit of higher

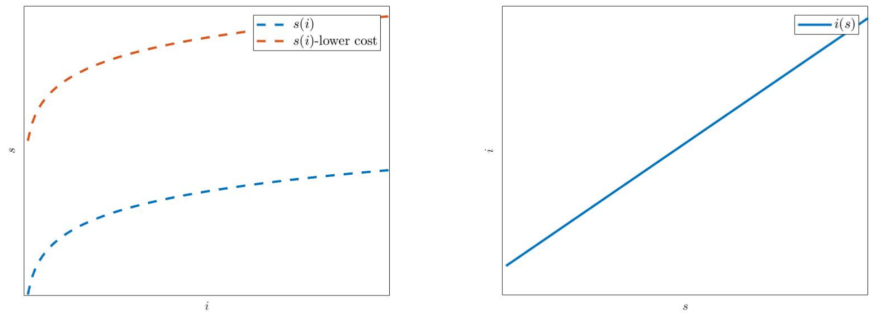


Figure 14: Search intensity as a function of the investment in financial literacy and investment in financial literacy as the function of search intensity.

search out-weights costs of increase in search. Further, as agent searches more resulting investment in financial literacy increases. Thus, as the likelihood of the new match increases, agent invests more in financial literacy to decrease value of default and increase the value of the joint surplus.

Policy Experiment

A policy experiment, we propose deduction in search costs. Intuitively, this policy would enable lower income agents and lower ability agents to search more intensively for better mortgage repayment rate. Thus, potentially, this policy could have welfare improving implications for the economy. In our stylized two period model, as a policy experiment, we decrease the search costs function's parameter $(c_0 = 0.0999)$. Left graph in Figure 14, shows how search intensity function increases for lower search costs, implying higher equilibrium search intensity and higher level of investment in financial literacy. Thus, agent benefits from reduction in search cost as it allows agent to increase the probability of a match with a lender with lower rate as well as to increase investment in financial literacy. As before, increase in investment in financial literacy brings the agent higher wage in the second moment and lower probability of default.

To summarize, we develop a stylized two period model of search for mortgage refinancing rate with financial literacy accumulation. Model features increasing search intensity as a function of investment in financial literacy. We propose modelling financial literacy as a general skill that reflects in various outcomes. Higher financial literacy increases agent's wage and at the same time reflects agent's financial responsibility as it reduces agent's default probability. Moreover, investment in financial literacy has similar implications for fully dynamic model

- it is costly to invest (lower wage this period)
- higher wage next period
- lower cost of search next period
- lower probability of default next period.

E Bellman Equation Derivation

E.1 Agent's Value

Agent's value $W = W(f, r(f, r^L, (r^L)'), (r^L)')$ is the value in period t of the agent with accumulated financial literacy f, training i, search effort s, whose current mortgage rate is r, and repays mortgage M at lender r^L. Further, agent's problem (in shorter notation and discrete time) is given with

$$W = \max_{i,s}\left\{\underbrace{((1-i)fz - rM - c(s,f))\Delta t}_{=u\Delta t} + \frac{1}{1+\rho\Delta t}\mathbb{E}W'\right\} = \max_{i,s}\left\{u\Delta t + \frac{1}{1+\rho\Delta t}\mathbb{E}W'\right\}. \tag{12}$$

In first period, the amount to return is $M_0 = M$, and the payment for each period is given with the usual accounting formula

$$rM.$$

We model mortgage repayment as payment's rM in each period.

We start the derivation by using Taylor expansion of the continuation value and rewrite it as

$$\begin{aligned}
\mathbb{E}W' =& \frac{\partial W}{\partial f}\dot{f}\Delta t + \frac{\partial W}{\partial t}\Delta t + (1 - \lambda s\Delta t - \delta(f)\Delta t)W + \\
&+ \lambda s\Delta t \underbrace{\int_{\underline{r}}^{r^L} W(r')d\Phi(r')}_{\text{moving to the new lender}} + \lambda s\Delta t \underbrace{\int_{r^L}^{Q} W(r')d\Phi(r')}_{\text{refinancing with current lender}} + \lambda s\Delta r \int_{Q}^{\bar{r}} W(r)d\Phi(r') + \delta(f)D \\
\overset{\text{W=W(r)}}{=}& \frac{\partial W}{\partial f}\dot{f}\Delta t + \frac{\partial W}{\partial t}\Delta t + (1 - \delta(f)\Delta t)W(r) + \\
&+ \lambda s\Delta t \int_{\underline{r}}^{r^L} (W(r') - W(r))d\Phi(r') + \lambda s\Delta t \int_{r^L}^{Q} (W(r') - W(r))d\Phi(r') + \delta(f)D,
\end{aligned}$$

where $Q = Q(f, r, r^L)$ defined by equation (2). We now plug back in derived expression for continuation value in the equation (12) and rearrange:

$$
\begin{aligned}
(1+\Delta t\rho)W(r) = \max_{i,s}\Big\{ & u\Delta t(1+\Delta t\rho) + \frac{\partial W}{\partial f}\dot{f}\Delta t + \frac{\partial W}{\partial t}\Delta t + (1-\delta(f)\Delta t)W(r) + \delta(f)D \\
& +\lambda s\Delta t\int_{\underline{r}}^{r^L}(W(r')-W(r))d\Phi(r') + \lambda s\Delta t\int_{r^L}^{Q}(W(r')-W(r))d\Phi(r') + o(t)\Big\} \\
& \iff \\
\Delta t(\delta(f)+\rho)W(r) = \max_{i,s}\Big\{ & u\Delta t(1+\Delta t\rho) + \frac{\partial W}{\partial f}\dot{f}\Delta t + \frac{\partial W}{\partial t}\Delta t + \delta(f)D \\
& +\lambda s\Delta t\int_{\underline{r}}^{r^L}(W(r')-W(r))d\Phi(r') + \lambda s\Delta t\int_{r^L}^{Q}(W(r')-W(r))d\Phi(r') + o(t)\Big\},
\end{aligned}
$$

divide by Δt, take the limit $\lim\limits_{\Delta t\to 0}$ and obtain

$$
\begin{aligned}
\rho W(r) = {} & u + \frac{\partial W}{\partial f}\dot{f} + \frac{\partial W}{\partial t} + \delta(f)(D - W(r)) \\
& + \lambda s\int_{\underline{r}}^{r^L}(W(r')-W(r))d\Phi(r') + \lambda s\int_{r^L}^{Q}(W(r')-W(r))d\Phi(r'),
\end{aligned}
$$

where $W(r')$ in the first integral is given by (1) (bargaining and moving to the new lender) and is equal to

$$J(f, r^L) + \beta(J(f, (r^L)') - J(f, r^L))$$

and $W(r')$ in the second integral is given by (2) (bargaining the new rate, but staying with the current lender) and is equal to

$$J(f, (r^L)') + \beta(J(f, r^L) - J(f, (r^L)')).$$

E.2 Lender's Value

Next step is to derive value of lender r^L, where lender's type is the rate at which the lender borrows money. We start by defining lender's period t profits by

$$\pi_t = rM - (r^L)'M.$$

Using π_t we can write lender's problem (in shorter notation) as

$$L = \max\left\{\pi_t \Delta t + \frac{1}{1+\Delta t\rho}\mathbb{E}L'\right\}.$$

We take the similar approach as in the derivation of agent's value and rewrite continuation value using the Taylor expansion

$$\begin{aligned}
\mathbb{E}L' = &\frac{\partial L}{\partial f}\dot{f}\Delta t + \frac{\partial L}{\partial t}\Delta t + \underbrace{(1-\delta(f)\Delta t - \lambda s\Delta t)L(r)}_{\text{agent's search not successful}} + \underbrace{\Delta t\delta(f)0}_{\text{agent defaults}} \\
&+ \underbrace{\lambda s\Delta t\int_{r^L}^{Q}(L^*(r''))d\Phi(r'')}_{\text{agent stays but with new rate}} \pm \lambda s\Delta t L(r)(\Phi(r)-\Phi(r^L)) \\
&+ \underbrace{\lambda s\Delta t L(r)(1-\Phi(r))}_{\text{new lender's rate is too high-agent stays}} + \underbrace{\lambda s\Delta t 0(\Phi(r^L))}_{\text{agent refinances with the new lender}} + o(t) \\
&= \frac{\partial B}{\partial f}\dot{f}\Delta t + \frac{\partial L}{\partial t}\Delta t + (1-\delta(f)\Delta t - \lambda s\Delta t\Phi(r^L))L(r) + \\
&+ \lambda s\Delta t\int_{r^L}^{Q}(L^*(r'') - L(r))d\Phi(r'') + o(t),
\end{aligned}$$

where $L^*(r'')$ is outcome of the bargaining (2) and is equal to

$$(1-\beta)(J(f,r^L) - J(f,(r^L)')).$$

Next, we plug the value for $\mathbb{E}L'$, multiply by $(1+\rho\Delta t)$, and rearrange

$$\begin{aligned}
\Delta t(\delta(f)+\rho)L = \pi_t\Delta t(1+\rho\Delta t) + \frac{\partial L}{\partial f}\dot{f}\Delta t + \frac{\partial L}{\partial t}\Delta t - \lambda s\Delta t\Phi(r^L)L(r) + \\
+ \lambda s\Delta t\int_{r^L}^{Q}(L^*(r'') - L(r))d\Phi(r'') + o(t).
\end{aligned}$$

In the last step, we divide by Δt, take the limit $\lim\limits_{\Delta t\to 0}$, and obtain

$$\rho L = \pi_t + \frac{\partial L}{\partial f}\dot{f} + \frac{\partial L}{\partial t} - \lambda s\Phi(r^L)L(r) + \lambda s\int_{r^L}^{Q}(L^*(r'') - L(r))d\Phi(r'') - \delta(f)L.$$

E.3 Default Value

Let D denotes default value, again by suppressing conditional notation. We proceed similar and write the maximization problem in discrete time

$$D = \max\left\{ \big[\underbrace{(1-i)zf - c(s,f) - \kappa}_{u_\kappa}\big]\Delta t + \frac{1}{1+\rho\Delta t}\mathbb{E}D' \right\}, \tag{13}$$

where κ denotes value od renting or loss on not owning a house. Taylor expansion of the continuation value is

$$\begin{aligned}\mathbb{E}D' = \frac{\partial D}{\partial f}\dot{f}\Delta t + \frac{\partial D}{\partial t}\Delta t + (1 - \lambda s\phi\Delta t)D + \\ + \lambda s\phi \int_{\underline{r}}^{\bar{r}} \widetilde{W}(r')d\phi(r'),\end{aligned}$$

where $\widetilde{W}$ outcome of the bargaining where the outside option for the agent after default is to take the mortgage with the lender with highest rate-type $\bar{r}$. Moreover, $\phi < 1$ stands here as it is more difficult to obtain a new mortgage than to refinance existing one, i.e., an agent has to prepare all documents and find a house, whereas for refinancing that is not necessary. We continue the derivation by multiplying equation (13) with $(1 + \Delta t\rho)$ and substitute for $\mathbb{E}D'$

$$\begin{aligned}(1+\Delta t\rho)D = u_\kappa\Delta t(1+\Delta t\rho) + \frac{\partial D}{\partial f}\dot{f}\Delta t + \frac{\partial D}{\partial}\Delta t + \\ + (1-\lambda s\phi\Delta t)D + \lambda s\phi \int_{\underline{r}}^{\bar{r}} \widetilde{W}(r')d\phi(r').\end{aligned}$$

After rearranging, dividing by Δt, and taking a limit $\lim\limits_{\Delta t\to 0}$, we obtain the final expression

$$\rho D = u_\kappa + \frac{\partial D}{\partial f}\dot{f} + \frac{\partial D}{\partial t} + \lambda s\phi \int_{\underline{r}}^{\bar{r}} (\widetilde{W}(r') - D)d\Phi(r').$$