CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Center for Economic Research and Graduate Education

Dissertation Thesis

2024 **Ante Šterc**

CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Center for Economic Research and Graduate Education

Ante Šterc

Essays on Frictions in Financial Decisions

Dissertation Thesis

Prague 2024

Author: **Ante Šterc**
Supervisor: **Ctirad Slavík, Ph.D.**
Year of the defense: 2024

References

ŠTERC, Ante. *Essays on Frictions in Financial Decisions* Praha, 2024. 127 pages. Dissertation thesis (PhD.). Charles University, Faculty of Social Sciences, Center for Economic Research and Graduate Education – Economics Institute. Supervisor Prof. Ctirad Slavík, Ph.D.

Abstract

This dissertation analyzes individual financial decisions and their implications for wealth heterogeneity. In Chapter 1, I build a structural framework of a discrete investment fund choice. Using the Survey of Consumer Finances (SCF) data, I show that households exhibit limited consideration when choosing an investment fund. Specifically, the structural model estimates show that households make their fund choice based only on a subset of available options. Conditional on wealth, monetary losses from limited consideration are higher for less financially literate households, suggestive of their choice simplification.

In Chapter 2, we focus on household mortgage take-up and refinancing decisions. Our novel U.S. data estimates show that the variation in mortgage rates depends on individual financial skill level and search effort. Specifically, we implement stochastic record linkage and find that households with low financial literacy are up to 4% less likely to consider more lenders and lock in at 15-20 b.p. higher rates. Upon origination, unskilled borrowers face a 35-45% higher mortgage delinquency and end up with a 30% lower likelihood of refinancing. We proceed to quantify monetary losses due to ineffective search, and we show that households with low financial skills pay more than 10% extra at the time of origination.

Chapter 3 extends to the general equilibrium and develops a Heterogeneous Agents New Keynesian model with a detailed outline of financial intermediation and plausible marginal propensities to consume (MPC). To motivate the model, we explore household survey data for the Euro area and document substantial heterogeneity in wealthy and poor hand-to-mouth (HtM) shares and in households' liquid and illiquid asset holdings. Accounting for heterogeneous MPCs allows plausible predictions of the effectiveness of fiscal policy in the short and long term. Using the model, we show that financing government debt with debt and government transfers has the largest positive long-term effect on output. To explain aggregate responses to fiscal stimulus, we introduce a new quantitative decomposition of aggregate consumption based on households' HtM status and wealth.

Abstrakt

Tato disertační práce analyzuje individuální finanční rozhodování a jejich důsledky pro heterogenitu v bohatství. V kapitole 1 vytvářím strukturální rámec diskrétní volby investičního fondu. Na základě údajů z Průzkumu spotřebitelských financí (Survey of Consumer Finances, SCF) ukazuji, že domácnosti zvažují výběr investičního fondu omezeně. Odhady strukturálního modelu konkrétně ukazují, že domácnosti se při výběru fondu rozhodují pouze na základě podmnožiny dostupných možností. V závislosti na bohatství jsou peněžní ztráty z omezeného zvažování vyšší u méně finančně gramotných domácností, což naznačuje, že jejich volba je zjednodušená.

V kapitole 2 se zaměřujeme na rozhodování domácností o čerpání hypoték a refinancování. Naše nové odhady na amerických datech ukazují, že rozdíly v hypotečních sazbách závisí na úrovni finančních dovedností a úsilí jednotlivců při hledání. Konkrétně implementujeme stochastické propojení záznamů a zjišťujeme, že domácnosti s nízkou finanční gramotností mají až o 4% nižší pravděpodobnost, že budou zvažovat více věřitelů a zafixují se na sazbách vyšších o 15-20 bazických bodů. Po poskytnutí hypotéky se nekvalifikovaní dlužníci potýkají s o 35-45% vyšší delikvencí a nakonec mají o 30% nižší pravděpodobnost refinancování. Pokračujeme v kvantifikaci peněžních ztrát způsobených neefektivním vyhledáváním a ukazujeme, že domácnosti s nízkými finančními dovednostmi zaplatí při poskytnutí hypotéky o více než 10% více.

Kapitola 3 se rozšiřuje na obecnou rovnováhu a rozvíjí novokeynesiánský model s heterogenními agenty s podrobným popisem finančního zprostředkování a věrohodnými mezními sklony ke spotřebě (MPC). Pro motivaci modelu zkoumáme údaje z šetření domácností v Eurozóně a dokumentujeme značnou heterogenitu v podílech bohatých a chudých domácností, které se chovají jako ‚hand-to-mouth' (HtM, z ruky do úst), a v držbě likvidních a nelikvidních aktiv. Zohlednění heterogenních MPC umožňuje věrohodně předpovědět účinnost fiskální politiky v krátkodobém i dlouhodobém horizontu. Pomocí modelu ukazujeme, že financování vládního dluhu dluhem a vládními transfery má největší pozitivní dlouhodobý vliv na HDP. Pro vysvětlení agregátní reakce na fiskální stimuly zavádíme nový kvantitativní rozklad agregátní spotřeby založený na statusu HtM a bohatství domácností.

Keywords

investment fund choice, limited consideration, financial literacy, mortgage refinancing, mortgage search, fiscal multipliers, wealth heterogeneity

Klíčová slova

výběr investičního fondu, omezené zohlednění, finanční gramotnost, refinancování hypotéky, hledání hypotečního úvěru, fiskální multiplikátory, heterogenita bohatství

Length of the work: There are 143718 characters.

Declaration

1. I hereby declare that I have compiled this thesis using the listed literature and resources only.
2. I hereby declare that my thesis has not been used to gain any other academic title.
3. I fully agree to my work being used for study and scientific purposes.

In Prague on 20.03.2024. Ante Šterc

Acknowledgement

I would like to express my gratitude to my supervisor, Ctirad Slavík, for guidance and encouragement throughout the process. Moreover, I would like to thank my dissertation committee, Marek Kapička, Vasily Korovkin, and Veronika Selezneva, for their support and constant feedback throughout my studies. I extend my appreciation to Christopher Phelan and Jaroslav Borovička for inviting me to their institutions, the University of Minnesota and New York University, respectively, which allowed me to pursue my research.

I am grateful to the participants of reading groups, workshops, and seminars at CERGE-EI, the University of Minnesota, and New York University, and my co-authors Othman Bouabdallah and Pascal Jacquinot for stimulating discussions, as well as many conference participants for their feedback.

I want to thank my parents for their unwavering support and patience and for believing in me. I want to thank my friends from Croatia for their support throughout the process. Special thanks to Lidia Cruces and Nicolò Russo for their support, encouragement, and advice and for many memorable moments in Minneapolis and Frankfurt.

Most of all, I want to thank my friend, partner, and co-author, Marta Cota, for her endless support. You made every day of this process better, funnier, interesting, and easy. Without you, it would not have been the same.

The first chapter was supported by the NPO "Systemic Risk Institute" [grant number LX22NPO5101]. The third chapter uses data from the Eurosystem Household Finance and Consumption Survey. The views expressed here are those of the authors and do not necessarily represent the views of the European Central Bank or the Eurosystem.

Table of Contents

Introduction

This dissertation analyzes households' financial decisions, the implications for their well-being, and the effectiveness of fiscal policy. Chapters 1 and 2 take an econometric approach and outline determinants and potential losses of investment fund choice and mortgage choice, respectively. Chapter 3 develops a general equilibrium model to understand the effectiveness of fiscal policy while accounting for individual frictions in asset management.

Chapter 1, titled *Limited Consideration in the Investment Fund Choice*, focuses on households' investment fund choice. In contrast to standard investment decisions, choosing an investment fund that acts on behalf of households is a common way of investing for many U.S. households. This chapter examines the role of limited consideration in household investment fund choice. I develop the Limited Consideration Model that quantifies the losses from not considering all available investment fund options. The SCF data maximum likelihood estimates statistically reject the standard full consideration Random Utility Model in favor of the proposed Limited Consideration Model. Losses due to limited consideration are significant and heterogeneous across household wealth. Conditional on wealth, the effects of limited consideration are stronger for less financially literate households, suggestive of underlying agent's choice simplification. These findings suggest that financial education policies may moderate choice simplification in investment fund choice.

In Chapter 2, titled *Mortgage Shopping Behavior in the U.S. - Stochastic Record Linkage* (Co-authored with Marta Cota) we aim to explain the difference in mortgage rate attainment among otherwise similar borrowers in the U.S. The paper provides novel insights into the interaction between individual financial literacy and shopping behavior and its effect on the mortgage interest rate in the U.S. market. In this regard, we merge two publicly available U.S. data sets and employ statistical methods that account for the uncertainty in the merging procedure. The merged data set contains mortgage and borrower characteristics, followed by survey responses on shopping behavior and individual objective financial literacy measures. First, we find that financial literacy changes with age and exhibits a hump-shaped life cycle profile. Second, we find that the interaction of financial literacy and search effort explains a part of the mortgage variation among otherwise similar borrowers. Specifically, financially skilled borrowers who consider multiple lenders get 13.4 b.p. lower mortgage rates at origination. This finding translates to over $9,329 of overpayment for a $100,000 loan in the U.S. over the mortgage term. We also show that the interaction coefficient increases over the 2014-2020 period, simultaneously with a steady increase in non-bank lenders in the U.S. mortgage market. Third, our findings suggest that three years after the mortgage originated, financially unskilled borrowers are 35-45% more likely to become delinquent.

Chapter 3, titled *Tax Structures and Fiscal Multipliers in HANK Models* (Co-authored with Othman Bouabdallah and Pascal Jacquinot), develops a heterogeneous-agents model with liquid and illiquid assets to analyze the fiscal multiplier quantitatively. Implementing a rich set of fiscal policy rules, including consumption, capital, progressive income taxes, and government transfers, allows us to quantify the size of fiscal multipliers while accounting for empirical heterogeneity in wealth. Specifically, we emphasize the role of households and their individual frictions in liquidity transformation for the potency of fiscal stimulus. Moreover, we compare the fiscal multiplier when the government spending is financed directly from one of the tax instruments to when the spending is financed through the government deficit. Next, we focus on deficit-financed spending and compare fiscal multipliers depending on the government's source of financing. We show that deficit-financed government spending with lump-sum transfers to households has the largest long-term impact on output. Lump-sum transfers circumvent individual frictions in liquidity transformation and increase demand among liquidity-constrained households. In this respect, accounting for individual frictions in asset management shows that direct transfers to households have a sizable effect on the aggregate.

1 Limited Consideration in the Investment Fund Choice

1.1 Introduction

Existing studies model the way households manage their financial assets by building a standard portfolio optimization problem. The usual assumption of those models is the symmetry in household information and the decision-making process across various household characteristics. Empirical findings suggest that returns to financial wealth exhibit significant persistence, further amplifying wealth inequality (Fagereng et al., 2020). Currently, with the development of the financial industry and access to information, households are able to compare the potential costs and benefits of choosing a specific financial asset. However, it may be that some households simplify their option sets and do not utilize full information.

Given the disproportionality in direct stock ownership across the wealth distribution, I model the choice between discrete fund types akin to fund options presented by the financial intermediary. Therefore, this paper focuses on simplified portfolio choice constrained by the financial intermediary, resulting in mutual fund share choices. The narrative of banks is that they offer to navigate their clients' options toward growth. However, not all households fully consider their investment choices but rather invest passively (Chetty et al., 2014; Andersen et al., 2020). For instance, Chalmers and Reuter (2020) find that a substantial amount of workers remain at the default fund choice when allocating their retirement savings.

In the data analysis, I use the Survey of Consumer Finances (SCF) that includes objective measure of financial literacy.[1] The first part of the data analysis focuses on the intensive margin, outlying household characteristics that induce households to buy shares in the investment fund. The second part evaluates the consequences of the limited consideration in investment fund choice. Together, the two parts outline target sample groups for relevant policies aimed at improving financial well-being.

In the first part of the analysis, I use the standard Two-Step Heckman Model (Heckman, 1979). Model estimates reveal household characteristics that affect the likelihood of the investment (selection equation) and characteristics that affect the investment size (outcome equation). It is more likely that more educated, financially savvy, or wealthier households opt into fund investing. At the same time, older households invest less. Expectedly, investment size decreases with household debt level. Specifically, the investment size decreases with financial literacy, suggestive of a cautious approach to investing with financially savvy households

[1] Financial literacy is measured by the standard three questions proposed by Lusardi and Mitchell (2014) covering inflation, interest rates, and riskiness. The list of questions is in section 1.3.1 of the appendix.

(aligning with Bhutta et al. (2022a) who find that liquidity level increases with financial literacy).

In the second part of the analysis, I introduce the limited consideration framework and focus on discrete investment fund types. I separate between funds based on their return, volatility, and expenses[2]. In this context, I relate the household investment fund choice to simple options outlined by the household intermediary. I evaluate households' fund choice likelihood and compare two models; the Random Utility Model assumes that households understand their options, and the Limited Consideration Model incorporates narrow sets of options from which households may choose.

Barseghyan et al. (2021) are the first to define the econometric framework for the Limited Consideration Model from Manzini and Mariotti (2014), incorporating household choice over insurance policies represented by discrete lotteries. I extend their model and represent a fund as a continuous random variable. This way, households are informed about expected return, volatility, and expenses, which are standard fund attributes households are acquainted with. Therefore, this paper is the first one to bring the limited consideration framework to the investment fund type choice. The Limited Consideration Model fits households' choices much better, and the Vuong (1989) test rejects the Random Utility Model in favor of the Limited Consideration Model at all usual significance levels. This finding suggests that, possibly, taking all available options is too costly for households, and they do not make optimal choices, i.e., achieve the first best allocation for investment.

As I reject the standard full consideration setting in favor of a limited consideration framework, I estimate average monetary loss under limited consideration. Losses are heterogeneous across sample groups. Specifically, conditional on wealth, high school graduates lose more than college and post-college graduates. Given the increasing importance of financial literacy in household finance decisions (Lusardi and Mitchell, 2014), I evaluate monetary losses across financial skill levels conditional on wealth. I find that households with a low level of financial literacy face significantly larger monetary losses due to fund type choice simplification.

On aggregate, I find that all households across the wealth distribution face statistically significant average monetary loss. This result is in contrast with results by Campbell (2006), who uses only the full consideration framework and finds that only a small fraction of households make investment mistakes. However, Campbell (2006) examines aggregate household investments, and I look specifically at investments in investment funds.

[2]For example, stock market funds invest primarily in equity (implying high returns and high volatility), and the money market fund invests mainly in short-term government T-bills (which implies low returns and low volatility).

Finally, the last part of my methodological contribution combines results from two econometric models and calculates the elasticity of marginal utility of investing in investment funds to financial literacy, wealth, and other household characteristics. These elasticities could be especially relevant for households' financial education. The results of my analysis imply that including financial courses in the curriculum that educate households on interest rate accumulation, risk, and inflation could benefit households. Specifically, the marginal elasticity of investing increases in financial literacy, and conditional on investing, average monetary losses are lower for households with higher levels of financial literacy.

The rest of the paper is organized as follows. Section 2 relates this paper to the literature. Section 3 describes the data used for all model estimations. Section 4 outlines important household characteristics that determine investment decisions. Section 5 develops the Limited Consideration framework and evaluates the losses once abstracting from rational behavior. Section 5 compiles all model estimates and discusses target household groups for policies that incentivize efficient investment decisions. Section 6 concludes.

1.2 Related Literature

This study builds on the large body of literature that examines asset market participation, putting forth the effects of heterogeneity in attention span, and thus, different information sets. Specifically, the paper contributes to two streams of literature.

First, this study adds to the literature on households' decision-making under risk with the assumption of constrained choice sets. Jung et al. (2019) and Caplin et al. (2019) show that rational inattention defines heterogeneous consideration sets across agents. Thus, agents use only a subset of available alternatives when optimizing their choice. Andersen et al. (2020) estimate the probability of active mortgage refinancing and link it to households' attention to financial well-being. Manzini and Mariotti (2014) define the departure from full consideration maximization, denoting it with *limited consideration*. Barseghyan et al. (2021) build on Manzini and Mariotti (2014) and develop an econometric framework for a discrete choice model with heterogeneous consideration sets and risk aversion. Coughlin (2019) uses a framework developed in Barseghyan et al. (2021) and explores limited consideration in the medical insurance setting. I contribute to their setting-specific findings and build on the econometric framework that allows the modeling agent's investment fund type choice. Specifically, I expend on discrete lottery framework in the insurance market by Barseghyan et al. (2021), and model expected utility with continuously distributed returns. Other studies that explore preference and discrete choice model estimation over unobserved choice sets include panel data and time variation, including Crawford et al. (2021), Aguiar and Kashaev (2021), and Aguiar et al. (2023).

Second, in line with asset market participation literature, I use the Two-Step Heckman Model (Heckman, 1979) to explore household decisions on whether to invest or not and on the size of the investment while accounting for selection bias. To this extent, I contribute by focusing on investment fund participation. In this respect, I add to the household finance stream of literature. Campbell (2006) finds that less educated households are less likely to participate in the asset market. Similarly, Calvet et al. (2009a) and Calvet et al. (2009b) find that financially sophisticated households with greater income, wealth, and education are more likely to enter the market. My estimates on investment fund participation correspond to previous findings on asset market entrance.

Kacperczyk et al. (2019) use a proxy for financial sophistication using wealth deciles and find that financially unsophisticated households increase their share of liquid instruments. Agarwal and Mazumder (2013) relate consumers' low math test scores to recurring financial mistakes in the credit card market. Moreover, Mani et al. (2013) relate poverty to lower cognition and provide experimental evidence that poor people often behave in less capable ways. Using Swedish data, Calvet et al. (2007) find that financially sophisticated investors invest more efficiently and more aggressively. Using an experiment, Nieddu and Pandolfi (2021) show that financial illiteracy can induce investors to inefficiently under-invest in risky assets such as stocks, because of their inability to understand the associated risks and returns that may affect their choice sets. I complement previous findings without adhering to wealth distribution dependence and use the available objective financial literacy score based on questions suggested by a well known strand of literature (Van Rooij et al., 2011; Lusardi and Mitchell, 2014).

1.3 Structural Econometric Model of Investment Decision

All model estimations use the Survey of Consumer Finances (SCF) data that contain a novel (objective) measure of financial literacy score and the investment fund type. In contrast to previous studies, the objective measure for financial sophistication yields novel findings on household characteristics' effects on investment fund choice. Eliciting financial literacy effects requires using all household characteristics relevant to investment choice. Moreover, the next subsection contains a description of the dataset and variables used in the analysis.

	Values					
Total sample	49377					
Education	no HS 5686	HS 12086	Some College 13756	College Degree 17849		
Age group	< 35 9312	$55 - 64$ 9329	$35 - 44$ 9811	$65 - 74$ 9801	$45 - 54$ 6950	$>= 75$ 4174
Occupation	Managerial/ Professional 15061	Tech/Sales/ Services 10598	Other 8958	Not Working 14760		
Income	$0 - 20\%$ 9678	$20 - 39.9\%$ 9515	$40 - 59.9\%$ 9563	$60 - 79.9\%$ 10011	$80 - 89.9\%$ 5586	$90 - 100\%$ 5024
Wealth	$0 - 24.9\%$ 12928	$25 - 49.9\%$ 11566	$50 - 74.9\%$ 11072	$75 - 89.9\%$ 8860	$90 - 100\%$ 4951	
Financial Literacy	0 2046	1 8287	2 17145	3 21899		
Home-Ownership Category	Owns Ranch/ Mobile Home/House /Condo/etc. 30061	Otherwise 19316				
Direct Stock-ownership	No 42311	Yes 7066				

Table 1: Descriptive statistics and overview of household data from SCF.

1.3.1 Data

For the analysis and estimation of models in this paper, I use the Survey of Consumer Finances. This dataset is suitable for the analysis since it contains the objective measure of financial literacy. Questions related to financial literacy asked in the Survey of Consumer Finances, proposed by Lusardi and Mitchell (2014) are:

- Suppose you had $100 in a savings account and the interest rate was 2 percent per year. After 5 years, how much do you think you would have in the account if you left the money to grow: [more than $102; exactly $102; less than $102; do not know; refuse to answer.]
- Imagine that the interest rate on your savings account was 1 percent per year and inflation was 2 percent per year. After 1 year, would you be able to buy: [more than, exactly the same as, or less than today with the money in this account; do not know; refuse to answer.]
- Do you think that the following statement is true or false? "Buying a single company stock usually provides a safer return than a stock mutual fund." [true; false; do not know; refuse to answer.]

The survey's objective measure of financial literacy is a score of 0 - 3, that is, the number of correctly answered questions. Further, as the first year that includes these questions is 2016, the available sample covering the measure is 2016 - 2019. The sample is a repeated cross-section, which restricts my analysis to a static framework with year controls.

In the analysis, this paper focuses on the household investment into investment funds. From the SCF dataset, i.e., survey questions, I construct indicator variables for each household and investment fund type they have invested into. Categories/types of investment funds in the SCF questions are Money Market, Stock Market, Government Bond, Other Bond, Combined, Tax Free Bond, and Other. Moreover, I use these indicators to estimate the Limited Consideration and the Random Utility models. For estimation of the Heckman Two-Step Model, I create only one one indicator for each household if they invested in any type of investment funds. Furthermore, for the size of the investment, I use the natural logarithm of the investment size.

Table 1 presents an overview of the main characteristics that may affect investment choice, both at the extensive and intensive margin. I standardize age by defining variable $Age_{\text{std}} = \frac{Age - \text{mean}(Age)}{2 * \text{sd}(Age)}$. Next, I use education, wealth, home ownership, stock ownership, and occupation variables described in Table 1. The income in SCF is the household income for the previous calendar year. Moreover, it includes wages, self-employment, and business income, taxable and tax-exempt interest, dividends, realized capital gains, food stamps and other support programs provided by the government, pension income and withdrawals from retirement accounts, Social Security income, alimony, and other support payments, and miscellaneous sources of income. In my analysis, I use indicators for income percentile groups. For variable wealth, I use the net worth of the household variable in the SCF, i.e., the difference between assets and debt. In my analysis, I use net-worth percentiles as described in Table 1. Finally, I use the debt-to-income ratio, defined as debt divided by income.

Investment Fund Type	Year	Mean	Variance	Expense Ratio
Money Market	2019	0.76%	0.000052712	0.001
Money Market	2016	0.094%	0.0000001144	0.001
Stock Market	2019	5.91%	0.009905728	0.0034
Stock Market	2016	12.898%	0.01063693	0.0034
Government Bond	2019	6.402%	0.01006727	0.0007
Government Bond	2016	8.664%	0.025230282	0.0007
Other Bond (i.e., Corporate Bond)	2019	5.952%	0.007935626	0.0022
Other Bond (i.e., Corporate Bond)	2016	7.786%	0.009949778	0.0022
Combined (Balanced)	2019	6.06%	0.003889336	0.0007
Combined (Balanced)	2016	8.876%	0.00368355	0.0007
Other	2019	6.47%	0.010196012	0.0041
Other	2016	11.386%	0.03158025	0.0041
Tax Free Bond	2019	5.142%	0.001673982	0.0017
Tax Free Bond	2016	6.574%	0.003090646	0.0017

Table 2: Approximated expected returns, variance, and expense ratio for investment fund types. Based on data from https://investor.vanguard.com/investment-products/list/mutual-funds.

In the second part of the analysis, I expand the Limited Consideration model (Barseghyan et al., 2021) for a continuous random variable instead of simple lottery. As it is in reality, when households choose an investment fund, they make a discrete choice. Thus, I represent the investment fund by three variables: return, volatility, and expense ratio. As these values are not available in the SCF, each type of investment fund, I approximate by three values using values for Vanguard investment funds presented in Table 2. Therefore, in my model, households will make a discrete choice between different types of investment funds, by choosing the highest expected value dependent on their size of the investment and three characteristics of the investment fund.

The object of interest is the probability of a fund type (j) choice, conditional on household i's characteristics. For example, in equation (1), the fund type corresponds to riskiness level conforming to risk averse (RA), risk neutral (RN), and risk loving (RL) investor. That is, I model the probability

$$p_i(s_i = j | z_i; \theta_i; \phi), \quad j \in \{RA, RN, RL\}, \tag{1}$$

where s is the household's fund choice, z are characteristics and θ_i is the household's specific, and ϕ is the general parameter vector.

I analyze the investment fund choice in two parts. The first part outlines those margins relevant to opting into fund investing, and the ones explaining the investment size. After informing about fund investors, the second part of the analysis zooms in on how the fund type choice is made, separating between standard, rational approach, and limited consideration framework.

1.4 Investment Fund Market Participation and Exposure

To identify households who invest (market participation) and how much they invest (exposure), I use the standard Heckman Two-Step Model. Additionally, Kline and Walters (2019) show that, under certain conditions, the Heckman Two-Step Model estimator is equivalent to the LATE estimator and, therefore, does not suffer from sensitivity critique. I check whether my model specification and the SCF data satisfy conditions in Kline and Walters (2019) and obtain the equivalence of the two estimators. For this reason, the estimates in this paper are robust to the sensitivity critique of the Heckman (1979) estimator. As the modeling approach in the analysis of the investment fund market participation is standard as for the analysis of the general asset market participation, the description of the econometric model and detailed analysis is in the Appendix A.1.

1.4.1 Investment fund participation-who participates?

The first column in Table 16 in the Appendix A.1.4 informs about marginal effects for the selection equation, calculated in percentage points. I discuss my results and compare my findings with other studies that use investor microdata. I focus on the extensive margin (deciding to invest) and discuss sample subgroups as potential targets for policies relevant to investment fund participation.

Estimation results show that older households are less likely to participate, in line with average age differences between asset market participants and non-participants (Calvet et al., 2007). Interestingly, renters are more likely to buy a share in the investment fund. Combining these two facts adheres to the life-cycle narrative: asset accumulation with the purpose of house down payment (Brandsaas, 2021). Clearly, stock owners are more likely to participate in the investment fund, while debt reduces the likelihood of participation.

Higher wealth implies a higher likelihood of investment, with a magnitude of almost four times as large as other household characteristics. In comparison to the middle wealth quantile, the top wealth quantile is 20% more likely to participate in investment funds. Correspondingly, households in managerial and professional occupations are more likely to invest. These results are in line with stock market participation (Campbell, 2006; Calvet et al., 2007, 2009a,b; Calvet and Sodini, 2014), and speak to persistent wealth inequality through fund participation channel.

Households with no high school relative to households with some college are 4% less likely to invest, while households with a college degree are 3% more likely to invest in investment funds. While similar studies Calvet et al. (2009b) and Van Rooij et al. (2011) resort to defining a measure of financial skill, I discuss my findings based on the direct measure of financial literacy. Households with a high degree of financial skill are 5% more likely to participate in the fund, which underlines limited understanding of fund options for the low level of financial skill (Nieddu and Pandolfi, 2021).

While education and wealth effects align with direct stock market participation (Calvet et al., 2007), model estimates inform about the use of financial skill in trusting the fund management. These results are in line with Kacperczyk et al. (2019), where low levels of study-defined financial skill imply shifting from intermediated products to standard liquid assets.

1.4.2 Investment Fund Participation-How Much do Investors Allocate?

The Inverse Mills Ratio is significant, which implies the selection of the data. Thus, both estimated coefficients and marginal effects presented account for the bias.

In the rest of the section, outcome equation marginal effects estimates are reported conditional on investment fund participation, thus informing about relevant margins for the investment size. Table 17 in the Appendix A.1.4 reports all marginal effect coefficients, whereas Figures 30, 31a, 31b, and 32 provide a visual representation.

Even though older households are less likely to participate, older investors allocate more to funds of choice. On the other hand, with the increase in debt-to-income ratio, households invest less in investment funds.

The education effect could be interpreted with student debt effects. College graduates allocate their funds to student debt repayment, therefore, buy smaller fund shares. In contrast, high-school graduates invest approximately 40% more. Financial knowledge effects show substantial variation, suggestive of under-diversification with investors of low degree of financial sophistication, in line with Swedish microdata and study-specific measure of financial knowledge (Campbell, 2006; Calvet et al., 2007). At the same time, households with a higher level of education and financial literacy invest more in other financial and non-financial assets (i.e., liquid savings and housing), according to the breakdown in Brandsaas (2021).

Finally, wealth effects on the fund investment size and supports conventional wisdom in household finance. The wealth effect is substantially larger than others, separating the investment size between the top and middle wealth quantile by more than double. These results align with Calvet et al. (2007), who find that wealthier households invest more.

1.5 Investment Fund Type Choice

Besides the investment size, the fund type choice contributes to heterogeneity in investment return. Each fund type is, in turn, characterized by its performance, realized return, and volatility. In this respect, households that consider all possible options automatically choose their fund optimally. However, some households consider a narrow set of options, potentially shifting away from the optimal choice under full consideration.

In the second part of the paper, I show that monetary losses due to limited consideration are significant and heterogeneous across household characteristics. Specifically, holding wealth fixed, losses are more significant for households with low financial skills and education. Combined with previous findings, adverse effects of financial skills and knowledge on the investment size, I discuss how limited consideration effects may be mitigated from the household side.

Figure 1 depicts fund type distribution for different levels of financial literacy. Fund types such as money market funds are not represented for lower levels of financial skill, whereas the stock market fund choice is the most frequent choice among all levels of financial skills. The Limited Consideration Model developed in this paper takes each fund type sample representativeness into account, allowing different consideration sets among households. To the best of my knowledge, this is the first study that explores limited consideration in investment fund

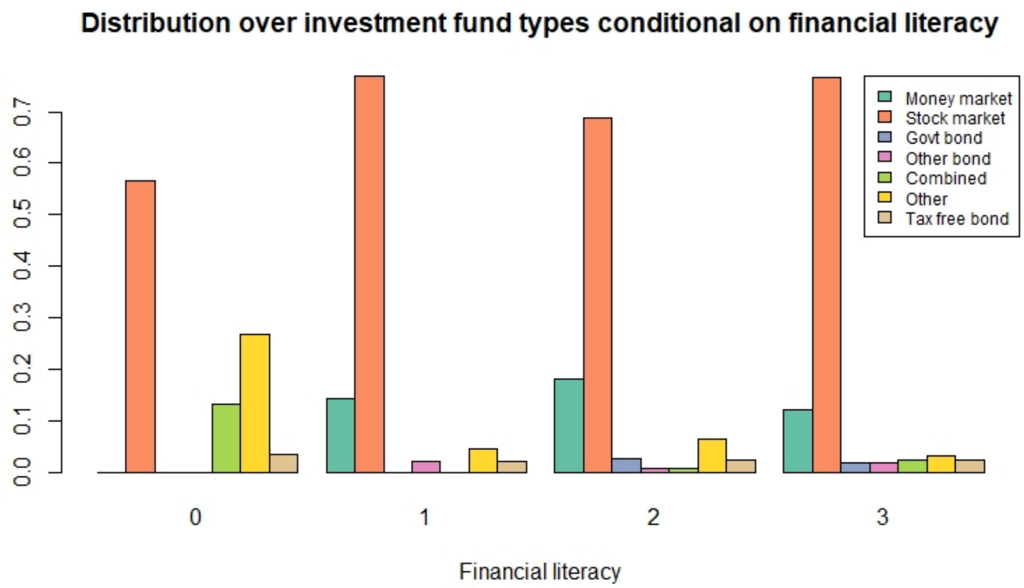


Figure 1: Distribution over investment fund types for different levels of financial literacy.

choice.

1.5.1 Utility Specification

Fund investors have CARA (constant absolute risk aversion) preferences for fund options defined with

$$u(c) = \begin{cases} \frac{1-\exp(-\nu c)}{\nu}, & \text{if } \nu \neq 0 \\ c, & \text{otherwise} \end{cases},$$

where ν is the parameter of the risk aversion. The model builds on Barseghyan et al. (2021), extending utility maximization over discrete lotteries to the maximization over fund choice with continuously distributed returns.

Specifically, I assume that given the size of investment W_i, household i chooses the investment fund j, characterized with a return $r_j \sim \mathcal{N}(\mu_j, \sigma_j^2)$, and the expense ratio ξ_j such that expected utility is maximized. The utility of choosing an investment fund, assuming heterogeneity in preferences $\nu_i \in [0, \bar{\nu}]$ is

$$u_i(r_j, \xi_j) = u_{ij} = \frac{1 - \exp(-\nu_i(W_i r_j(1-\xi_j)))}{\nu_i}.$$

Since returns are normally distributed, for the (narrow) fund choice set $\mathcal{J}$, it follows

$$\begin{aligned}\max_{j\in\mathcal{J}}\mathbb{E}[u_{ij}] &\Leftrightarrow \max_{j\in\mathcal{J}}\mathbb{E}[-\exp(-\nu_i(W_i r_j(1-\xi_j)))]\\ &\Leftrightarrow \min_{j\in\mathcal{J}}\mathbb{E}[\exp(-\nu_i(W_i r_j(1-\xi_j)))]\\ &\Leftrightarrow \min_{j\in\mathcal{J}}\exp(-\nu_i(W_i\mu_j(1-\xi_j))+\frac{\nu_i^2}{2}\sigma_j^2 W_i^2(1-\xi_j)^2)\\ &\Leftrightarrow \max_{j\in\mathcal{J}}\mu_j-\frac{\nu_i}{2}\sigma_j^2 W_i(1-\xi_j).\end{aligned}$$

The final expression allows for easier and faster evaluations of the objects of the model. Estimations use Vanguard's corresponding fund type data to approximate returns, volatility, and expense ratios. Approximations are given in Table 2.

1.5.2 Limited Consideration Model

Similar to Barseghyan et al. (2021) and Manzini and Mariotti (2014), households exhibit limited consideration. In contrast to the standard assumption that investors choose the best alternative among all available, in this model, household i evaluates options within individual consideration set $J_i \subseteq \mathcal{J}$. Indicator $y_{ij} = 1$ denotes if household i prefers option j among other options within their consideration set J_i,. The corresponding probability of choosing fund j is (leaving out conditioning notation):

$$\mathbb{P}(y_{ij}=1)=\sum_{J\subseteq\mathcal{J}:j\in J}\mathbb{P}(J_i=J)\mathbb{P}(\mathbb{E}[u_{ij}]>\mathbb{E}[u_{ik}],\quad \forall k\in J). \tag{2}$$

Investment fund j appears in the household's consideration set with probability φ_j, independently of other alternatives. Further, I assume that consideration probabilities of investment funds are homogeneous across agents who face the same feasible choice set[3]. Thus, the probability of any consideration set $J_i = J \subseteq \mathcal{J}$ is the intersection of individual consideration alternatives:

$$\mathbb{P}(J_i=J)=\prod_{j\in J}\varphi_j\prod_{j\notin J}(1-\varphi_j). \tag{3}$$

Standard to the limited consideration framework, I assume $\varphi_j > 0$ to omit *never-considered* alternatives from the choice problem. This is because the option for which $\varphi_j = 0$ is never considered or compared to other alternatives and as such, does not affect the choice problem. Combining equations (2) and (3) results in the following equation for the probability of

[3] Barseghyan et al. (2021) offer more general consideration probabilities that could be modeled as functions of the agent's characteristics

$y_{ij} = 1$:
$$\mathbb{P}(y_{ij} = 1) = \sum_{J \subseteq \mathcal{J}: j \in J} \prod_{j \in J} \varphi_j \prod_{j \notin J} (1 - \varphi_j) \mathbb{P}(\mathbb{E}[u_{ij}] > \mathbb{E}[u_{ik}], \quad \forall k \in J). \tag{4}$$

Using equation (4) to evaluate the probability of a choice y_{ij}, requires enumeration of all possible consideration sets, which is computationally unfeasible. However, the model feature outlined below does not necessitate approximations. Since equation (2) does not include an error term, the choice-based expected utility can be ranked for a fixed parameter of the risk aversion

$$\mathbb{E}[u_{i1}] < \cdots < \mathbb{E}[u_{ij}] < \mathbb{E}[u_{i|\mathcal{J}|}],$$

where $|\mathcal{J}|$ denotes cardinal number of set $\mathcal{J}$. Therefore, if $y_{ij} = 1$, it means that household i chooses optimally and options ranked higher than fund j cannot be in the consideration set. Thus, for fixed $\nu_i = \nu$, for all alternatives $k \in J$ that are preferred over chosen alternative j, $\mathbb{P}(\mathbb{E}[u_{ij}] > \mathbb{E}[u_{ik}]) = 1$ and for all $k \notin J$ $\mathbb{P}(\mathbb{E}[u_{ij}] > \mathbb{E}[u_{ik}]) = 0$. All together, denoting

$$\mathcal{B}_\nu(y_j = 1, x) = \{k : \mathbb{E}[u_k|\nu, x] > \mathbb{E}[u_j|\nu, x]\},$$

which, in combination with the previous derivation, yields the following form of the conditional probability of choosing a fund j for a fixed value of risk aversion

$$\mathbb{P}(y_j = 1|\nu, x) = \varphi_j \prod_{k \in \mathcal{B}_\nu(y_j = 1, x)} (1 - \varphi_k),$$

and corresponding probability of choosing a fund j across all households, conditional on fund j characteristics:
$$\mathbb{P}(y_j = 1|x) = \int \mathbb{P}(y_j = 1|\nu, x) dF. \tag{5}$$

1.5.3 Maximum Likelihood Estimation

Limited Consideration Model estimation requires distributional assumptions regarding preference parameters. Similar to Barseghyan et al. (2021) and Coughlin (2019), I assume the Beta distribution for the parameter of the risk aversion. Specifically, for each household i with characteristics $\mathbf{X}$:
$$\log \frac{\beta_{1i}}{\beta_2} = \mathbf{X}_i \gamma, \tag{6}$$

where γ is an unknown vector of coefficients to be estimated. Parameters β_{1i} and β_2 are the parameters of the Beta distribution, where β_{1i} is household-specific and β_2 is common across agents. By assumption, preference coefficients are random draws from a distribution with the mean represented as a function of the observable characteristics given with the following equation

$$\mathbb{E}[\nu_i] = \frac{\beta_{1i}}{\beta_{1i} + \beta_2}\bar{\nu} = \frac{\exp(\mathbf{X}_i\gamma)}{1 + \exp(\mathbf{X}_i\gamma)}\bar{\nu}. \tag{7}$$

Note, for consideration probabilities $\{\varphi_j\}_{j\in\mathcal{J}}$, joint product $\prod_{k\in\mathcal{B}_\nu(y_j=1,x)}(1-\varphi_k)$ is piecewise constant over alternatives. Thus, equation (5) can be written in the following form:

$$\mathbb{P}(y_j|x) = \varphi_j \sum_{h=0}^{D-1} \left((F(\nu_{h+1}) - F(\nu_h)) \prod_{k\in\mathcal{B}_{\nu_h}(y_j=1,x)} (1-\varphi_k) \right), \tag{8}$$

where ν_h are the sequentially ordered breakpoints augmented by the integration endpoints $\nu_0 = 0$ and $\nu_D = \bar{\nu}$, and $F(\cdot)$ is a CDF of the Beta distribution. I estimate the equation (8) using a Riemann integral approximation for the CDF. I assume that preferences depend on wealth, education, and financial literacy, eliciting data patterns in household preferences (Ameriks et al., 2003; Sutter et al., 2020; Mudzingiri, 2021). The estimation results are given in section A.1.4 of the appendix in Table 19. In addition, appendix A.1.4 contains estimation results of the Limited Consideration Model without any *ad hoc* assumptions on preference heterogeneity. For both models, estimation results are reported with 95% bootstraped confidence intervals for $B = 1000$ replications.

1.5.4 Full Consideration - Random Utility Model (RUM)

The Random Utility Model, commonly used in consumer choice literature, serves as the benchmark for model comparison since it uses the assumption of full consideration. Hence, I compare the LCM (presented above) with RUM, which incorporates additively separable unobserved heterogeneity (e.g., Mixed Logit). Using standard derivations (McFadden and Train, 2000) and the assumption that the utility error is iid Type 1 Extreme Value distributed, the probability of choosing alternative j, conditional on risk aversion parameter ν, is given with

$$\mathbb{P}(y_j|x,\nu) = \frac{\exp(V_j(x,\nu))}{\sum_k \exp(V_k(x,\nu))}, \forall j \in \mathcal{J}.$$

where $V_j(x,\nu) = \mathbb{E}[u_j|x,\nu] + \varepsilon_j$. Similarly, the risk aversion parameter follows Beta distribution such that parameters satisfy equations (6) and (7). Therefore, risk aversion parameters depend on the same characteristics as in the limited consideration case. Finally, averaging over household preferences yields the final expression for the probability of choosing the fund j, conditioning only on fund characteristics

$$\mathbb{P}(y_j|x) = \int \mathbb{P}(y_j|x, \nu)dF.$$

In 1.5.4, $F(\cdot)$ is the CDF of the Beta distribution, suitable for the Riemann integral approximation within the Maximum Likelihood estimation procedure. The estimation results are given in section A.1.4 of the appendix in Table 20. Estimation results are presented with 95% bootstrapped confidence intervals for $B = 1000$ replications.

1.5.5 Model Comparison

Barseghyan et al. (2021) show that the Limited Consideration Model and RUM generate several contrasting implications. For example, RUM generally implies that each alternative has a positive probability of being chosen and satisfies a generalized dominance property. In contrast, the Limited Consideration Model can generate zero shares (consideration probabilities for some funds can be zero) and does not necessarily abide by generalized dominance. Hence, the limited consideration assumption appropriates non-existent shares with sample groups in the SCF data (Figure 1).

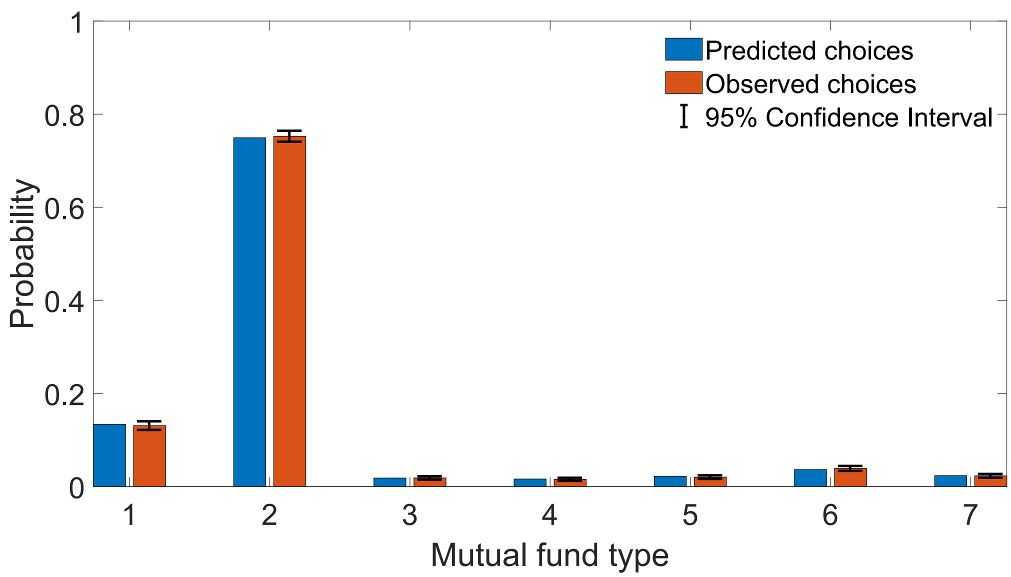


Figure 2: Predicted choices for the Limited Consideration Model and observed choices with 95% confidence intervals. Investment fund types include the money market, stock market, and government bond fund, with other bond funds (i.e., corporate bonds), combined funds, tax-free bond funds, and others (specifically, hedge or growth).

Figure 2 depicts predicted choice probabilities using the estimated Limited Consideration Model in comparison with observed household choice. The predicted probabilities fit observed

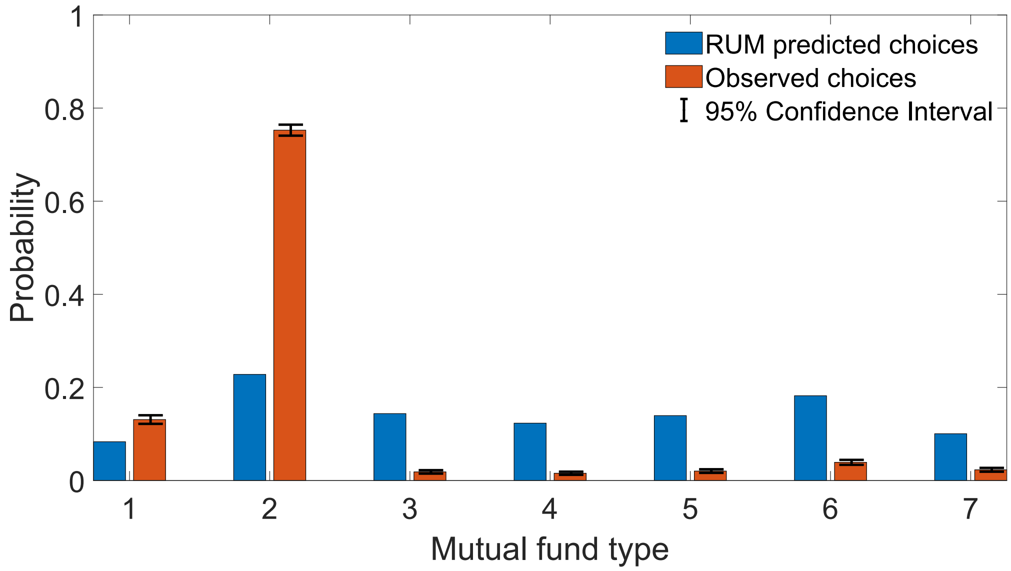


Figure 3: Predicted choices for the Random Utility Model and observed choices with 95% confidence intervals. Investment fund types are: money market, stock market, government bond, other bond (i.e., corporate bond), combined, other (i.e., hedge or growth), and tax-free bond.

choices well across all fund types, even when restricting preference parameters.

In contrast, the Random Utility Model predicts homogeneous probabilities across all fund types (shown in Figure 3), thus does not fit as well. For each type of fund, predicted choice probabilities do not correspond to the 95%-confidence intervals.

Overall, standard full information setting does not reproduce choice patterns in the SCF data, whereas incorporating a narrow consideration set appropriates non-existent shares in some type of funds for sample subgroups. Moreover, calculating Likelihood Ratio test statistics for non-nested models (Vuong, 1989) is 53.4949. Comparing the value with critical values of the Normal distribution implies that the Likelihood Ratio test rejects the Random Utility Model in favor of the Limited Consideration Model at all reasonable significance levels. Therefore, including limited consideration appropriates the household decision process while choosing the investment fund.

1.5.6 Conditional Probabilities Comparison

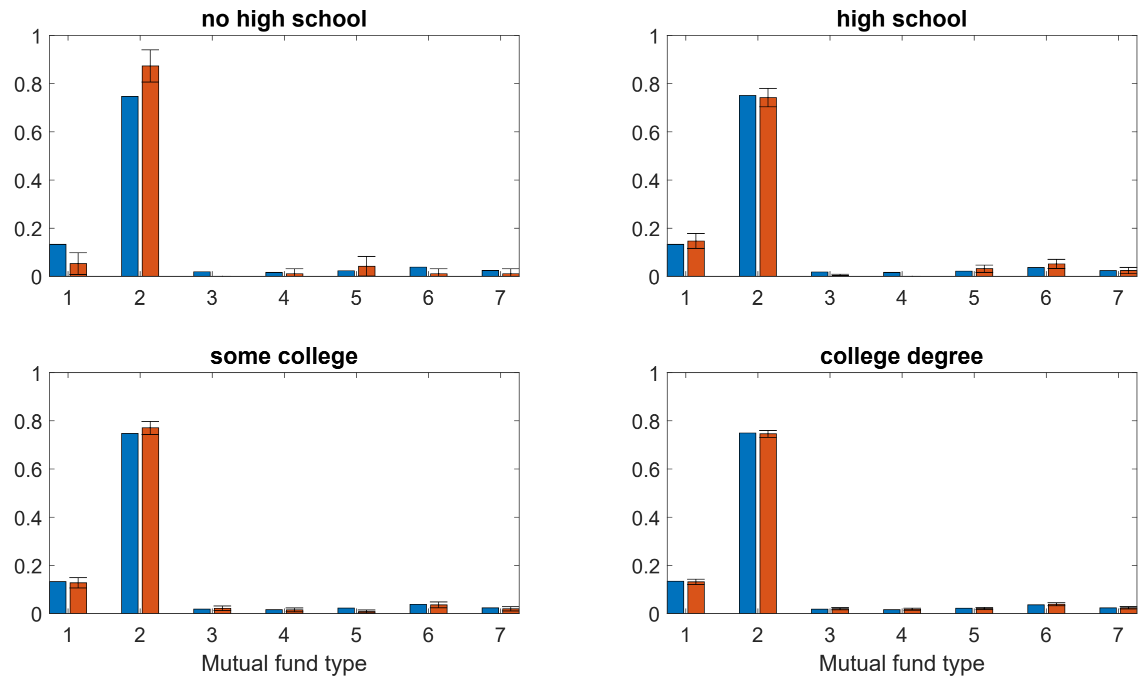
Given that aggregate sample, predictions show that the Limited Consideration Model performs better as a household fund choice model, the next part of the paper estimates choice predictions conditional on key household observables. Marginal effects estimates in the first part of my analysis outline important margins for investment size: education, financial literacy, and wealth (Figures 28, 29a, 31, and 32). In combination with the limited consideration effect, monetary loss estimates, conditional on education, wealth, or financial literacy, may be substantial.

Therefore, I outline conditional choice predictions across education, wealth, and financial literacy levels. Group-based differences between the two model predictions show that the Limited Consideration Model outperforms the Random Utility Model, even though agents' preferences depend on these characteristics. Consequently, cluster-based predictions point to target groups for policy recommendations that could mitigate the effects of limited consideration.

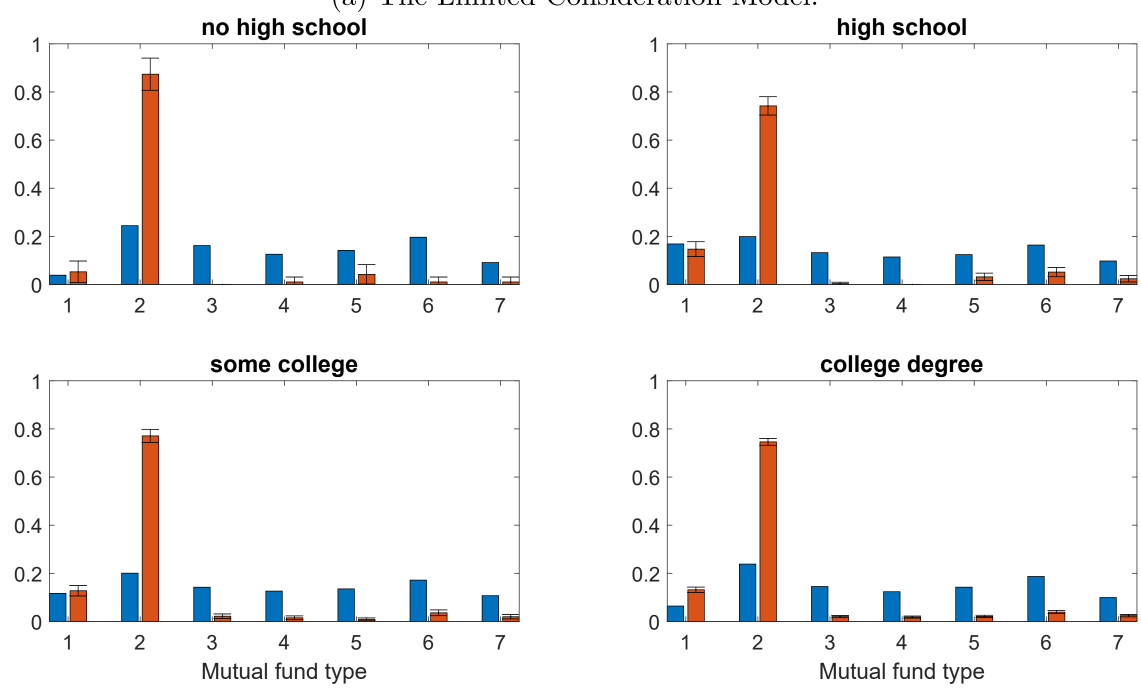
Figure (4a) depicts the quality of the Limited Consideration Model fit, conditional on education. Pertaining to the adverse effects of education on investment size (Figure 28a), model predictions yield unambiguous estimates of average monetary losses.

Even though the Random Utility Model assumes risk aversion dependent on education, it fails to capture the investment fund choice along the education margin. Conditional on education, the Random Utility Model fit yields a similar conclusion as in unconditional predictions; the Limited Consideration Model outperforms the Random Utility Model.

A similar conclusion follows from Figure 5, where observed choices are compared to predicted choices for two models conditional on the level of financial literacy. Again, the Random Utility Model fails to match the choice distribution across all financial literacy levels.

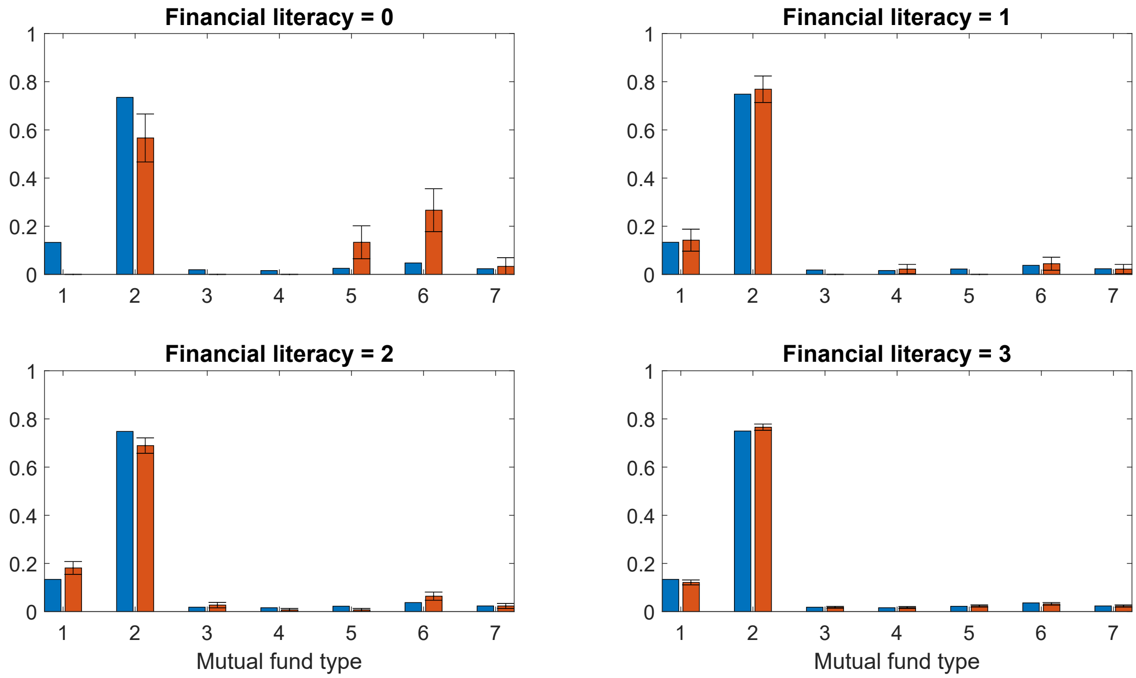


(a) The Limited Consideration Model.

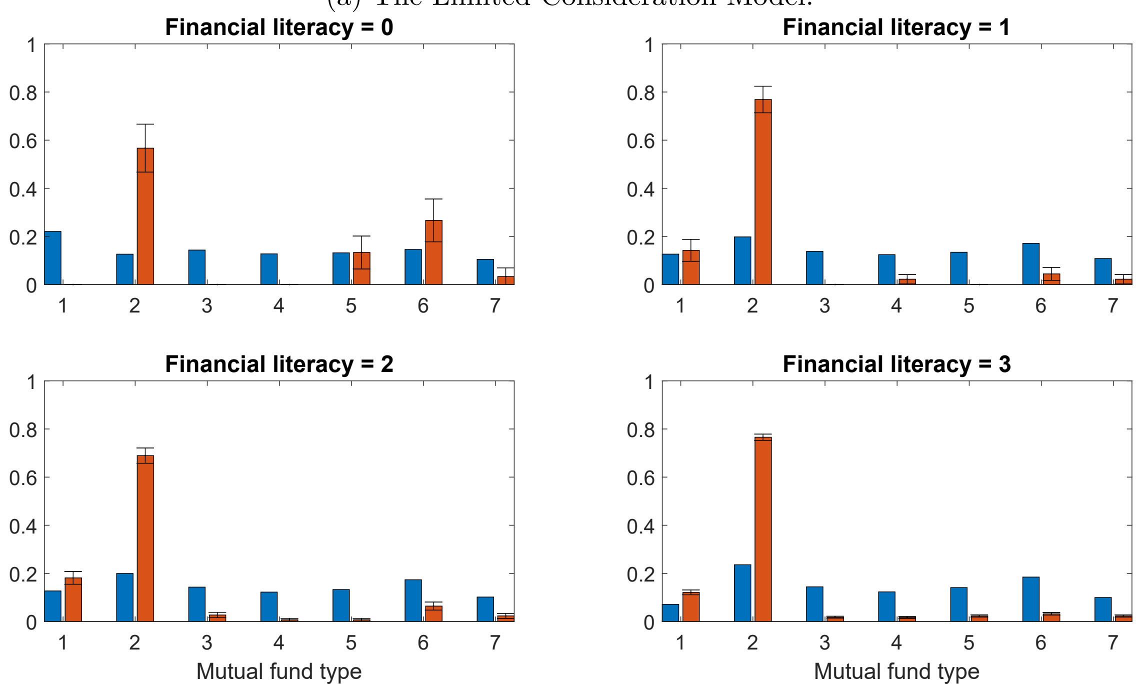


(b) The Random Utility Model.

Figure 4: Distribution of choices for the Limited Consideration Model and for the Random Utility Model compared to observed choices, conditional on the level of education.



(a) The Limited Consideration Model.



(b) The Random Utility Model.

Figure 5: Distribution of choices for the Limited Consideration Model and for the Random Utility Model compared to observed choices, conditional on the level of financial literacy.

1.5.7 Results From the Limited Consideration Model

Limited Consideration and Random Utility model estimates imply the Likelihood Ratio test statistic for non-nested models (Vuong, 1989), which rejects the Random Utility Model in favor of the Limited Consideration Model on all usual significance levels. Thus, I can conclude that some households optimize their fund investment based on a narrow consideration set instead of considering all available options.

Table 19 in section A.1.4 of the Appendix contains maximum likelihood estimates, with the average risk aversion for CARA, $\bar{\nu}_{MLE} = 0.0094$, in line with insurance choice revealed preference in Barseghyan et al. (2021) and Coughlin (2019), and experimental data estimates in Rabin (2013). Shifting away from unrestricted risk aversion to functional dependence on household characteristics does not distort the mean of the risk aversion distribution.

Observable characteristics, such as financial literacy and education, significantly affect a household's risk aversion. In contrast, wealth estimates show insignificant effects on risk aversion, in line with the CARA utility assumption. Due to the high degree of model non-linearity, I abstract from giving an interpretation of the signs or the size of coefficient estimates.

Nevertheless, I analyze the average monetary losses due to limited consideration, thus the losses from not choosing the first best. The model does not allow disentangling the underlying mechanism that prevents households from considering all options. However, Figure 1 suggests that, for the same amount of wealth, financially sophisticated households consider broader sets of options, thus choosing options not chosen by less financially literate households. It may be that financially savvy households find it less costly to learn about the options at hand.

1.5.8 Monetary Loss Due to Limited Consideration

In order to evaluate the effect of agents' limited consideration, I calculate monetary losses conditional on education, financial literacy, and wealth-based household groups. I calculate the gains from switching to full-consideration behavior and discuss policy-relevant household groups.

Because I assume CARA utility, household i is willing to accept ce_{ij} instead of investing in the chosen fund, either under limited or full consideration, j_{LC} and j_{RUM}, respectively (certainty equivalent)

$$ce_{ij} = -\frac{1}{\nu_i}\log(1 - \nu_i \mathbb{E}[u_{ij}]), \quad j \in j_{LC}, j_{RUM}.$$

I take the difference $ce_{ij}^{LC} - ce_{ij}^{RUM}$ and average them across the household and household subgroups.

Table 21 shows average losses in measured $\$10,000$ across education and financial literacy levels. Results from the first column imply that, on average, households lose around $\$2,727$ because of narrow consideration sets. Specifically, households with high school education at most lose more than the average. However, evidence in Campbell (2006) suggests that many households invest effectively and a minority make significant mistakes, whereas I find that all groups of households face significant monetary losses.

Next, re-evaluate monetary losses across education and financial literacy levels for fixed wealth quantiles. While 6 reveals small differences between education levels, Figure 7 shows that financial literacy-based differences are substantial, especially in the lower half of the wealth distribution.

Combining adverse knowledge effects on the investment size with losses in Table 22 speaks to subsamples appropriate for fund type information provision. Namely, wealthier and less educated households face larger losses due to limited consideration. According to Heckman model estimates (Figures 31a,31b and 29a), these households invest more and under-diversify due to limited consideration. The results are presented in Table 22.

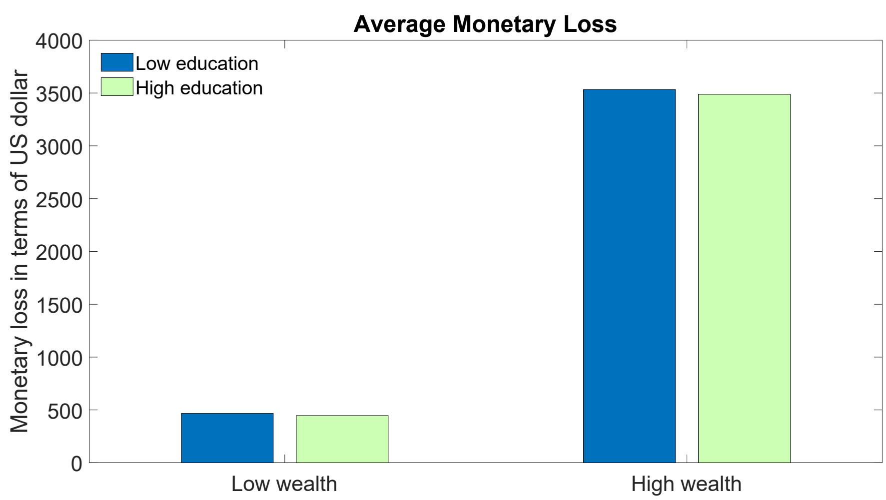


Figure 6: Average monetary loss for households' low and high level of education grouped by wealth category.

To summarize, even though I do not specify the mechanism behind the limited consideration behavior of agents in their choice of investment fund, my estimates imply that agents with a lower level of education or financial literacy face greater losses. Adverse effects of financial literacy and education are potentially attributable to higher information acquisition costs, which prevent the household from evaluating all alternatives in the choice set.

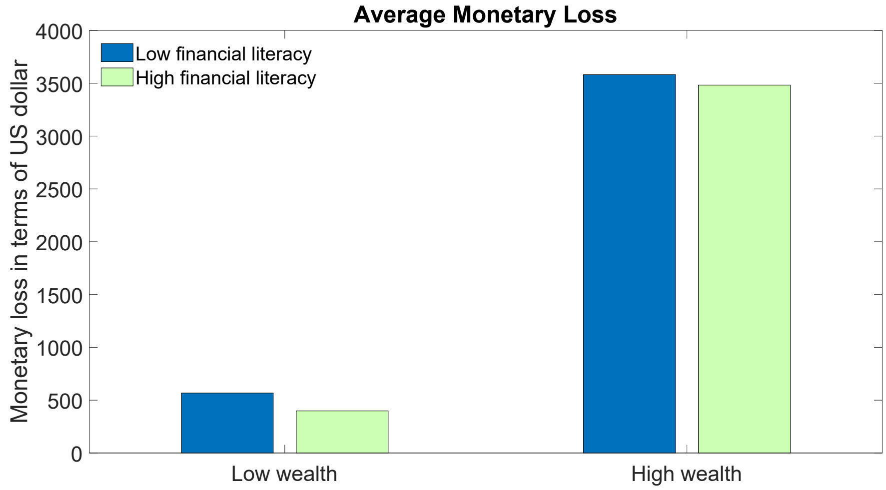


Figure 7: Average monetary loss for households' low and high level of financial literacy grouped by wealth category.

1.5.9 Source of Limited Consideration, Heterogeneous Returns and Wealth Inequality

In the paper, I cannot disentangle the mechanism underlying the limited consideration behavior. Both theoretical and empirical studies on consumer choice discuss potential sources behind abstracting from some options.

Caplin et al. (2019) show that, because it is too costly to consider all available options, rational inattentive agents take actions from constrained choice sets (i.e., they limit their consideration sets). Similarly, households with a lower level of education and financial literacy in the SCF data potentially face higher attention costs and correspondingly larger monetary losses from limited consideration.

In contrast to modeling information acquisition frictions, the source of the limited consideration could be attributable to the information provision from the supply side, with a financial advisor who allocates their clients how they see fits (Mullainathan and Shleifer, 2005; Mullainathan et al., 2012; Gil-Bazo and Imbet, 2020).

Financially skilled households may explore their options on their own, whereas households with low financial literacy may follow financial advice easily, aligning with the model in Gennaioli et al. (2015), where financial intermediaries reduce the perception of the riskiness of a proposed investment. In this regard, their finding is in line with another result of this paper - investment fund choices are concentrated towards stock market investment funds (depicted in Figure 1), which imply a higher return but higher volatility at the same time.

1.5.10 Connecting Two Estimated Models

In this section, I use estimated parameters from the Limited Consideration Model and regress model-implied expected utilities on household characteristics. As a result, I obtain elasticities of (expected) marginal utility across relevant margins. That is, I define

$$Y_i = \begin{cases} \mathbb{E}[u_i], & \text{if investment occurred,} \\ 0, & \text{otherwise.} \end{cases}$$

I keep the regressors the same as in the Heckman Model to draw a parallel between coefficients with (LCM) and without (Heckman) modeling utility. In general, estimated coefficients overlap in sign and size. Results of the estimation are given in section A.1.4 of the appendix in Table 23.

Using the results from Table 23, I am now able to calculate (semi) elasticity of marginal utility of investing in variables such as wealth and financial literacy[4]. I find that the expected utility of investing increases with education level, amounting to 33.5% higher utility for college graduates in comparison to households with lower education. Similarly, the expected utility of investment is 40.9% higher when comparing households with financial literacy equal to 3 to those with financial literacy equal to 2. In addition expected utility monotonically increases with wealth.

Estimated elasticities of the marginal utility of investing in financial literacy imply welfare-increasing effects of financial education. Consequently, improving financial literacy not only increases utility but also reduces monetary losses attributable to limited consideration. Thus, financial education policy implies welfare improvement across the wealth distribution.

1.6 Conclusion

In this paper, I take a novel approach to modeling participation in the financial asset market. Instead of using a standard portfolio model, I consider investment fund choice as a discrete consumer decision problem. Using the SCF data, I employ the Heckman Two-Step Model to elicit household characteristics important for the extensive (opting in) and intensive (investment size) margins of fund investing. Results on the likelihood of participation show that wealthier and financially sophisticated households choose to invest in a fund. Controlling for wealth, investment size decreases with education and financial

[4]That is, I calculate the percentage change of the left-hand side variable corresponding to a change in the categorical variable of the right-hand side variable of the regression.

literacy, potentially contributing to diversification.

Next, I analyze specific investment fund choice. I take a novel approach and explore limited consideration in the type of investment fund choice. Consequently, I build on the lottery-based framework in Barseghyan et al. (2021) by accounting for returns behavior, i.e., incorporating continuous random outcomes of a choice at hand. As a result of my estimates, I reject the full consideration behavior of households (RUM) in favor of limited consideration behavior. In contrast to previous literature, I show that households do not achieve first-best allocation because they consider only a constrained set of available investment options.

Given that the usual approach to investment choice in the literature is a full consideration setting, I evaluate the monetary losses accrued to limited consideration behavior. I find that, under limited consideration, all households make mistakes in their fund choice, which contradicts the findings within the full consideration framework (Campbell [2006] finds that most households invest effectively and a minority makes mistakes). In addition, I find that, across the wealth distribution, households with a lower level of education or financial literacy face larger monetary losses than households with higher levels.

Finally, I estimate the elasticity of the marginal utility of fund investing and find that utility increases in both education and financial literacy. Overall, this study highlights the importance of financial literacy interaction with limited consideration in households' investment decisions, putting forth financial education as a prospective policy that could mitigate the effects.

2 Mortgage Shopping Behavior in the U.S. - Stochastic Record Linkage

Co-authored with Marta Cota (CERGE-EI)

2.1 Introduction

Following the increase in data availability, the literature on financial behavior moved towards empirical estimates of cognitive and monetary costs of individual investing and saving. In an effort to unify cognitive costs and differences in understanding, Lusardi et al. (2010) proposed to measure individual financial knowledge using a set of three survey questions (*"The Big Three"*). These questions define the **objective financial literacy score** and are related to differences in saving and consumption behavior.

Whereas most of the literature focuses on the correlation between financial literacy score and debt or asset level, our paper aims to uncover the mechanism underlying the positive correlation between financial knowledge and individual debt management. In this regard, we focus on the mortgage rate attainment in the U.S. and, using a stochastic imputation procedure, show that individual shopping behavior and financial skill level interact and explain a part of the residual mortgage rate variation after accounting for observables.

The U.S. mortgage market has undergone significant structural changes and advancements in digital mortgage advertising and undertaking. With a steady increase in non-bank online lenders, the mortgage market sustained an increase in competition, elicited through the increase in mortgage eligibility to modest credit score borrowers (Zhou, 2022; Bhattacharya et al., 2021). Following the increase in options, individual shopping behavior and financial knowledge became significantly more important for mortgage attainment. In this respect, we focus on the demand side while controlling for the other contract specifics.

Limited data availability does not allow connecting individual financial knowledge to shopping behavior. To circumvent public data limitations, we employ the Stochastic Record Linkage (Enamorado et al., 2019) and impute individual financial literacy scores for borrowers in the National Survey of Mortgage Originations (NSMO). The stochastic linking method allows us to control for the uncertainty in the financial skill level obtained from the external data set. In this way, for every borrower in the NSMO, we estimate a distribution of the financial skill level that depends on her respective match to a record in the Survey of Consumer Finances (SCF).

The objective measure of financial skills provides unique insights for individual mortgage attainment. In this regard, our findings surpass subjective perceptions of financial knowledge and risk aversion. Our first line of findings uses the SCF sample and suggests that financial literacy exhibits a hump-shaped profile over the life cycle. Moreover, we show that financially skilled borrowers are 20-30% more likely to refinance their mortgage, irrespective of their income, education, and mortgage size.

Next, we turn to our new merged data set and measure the borrower's effort using the survey question on the number of mortgage lenders considered in the mortgage shopping process. Our estimates show that, among similar mortgage applicants, financially savvy ones are 5% more likely to consider one additional lender. Moreover, we show that the search effort effectiveness increases with skill level and predicts a 13.4 b.p. lower mortgage rate for financially savvy borrowers who exert more effort in the mortgage acquisition process.

The sample period from 2014 to 2021 provides a window to observe variations in financial skills and search effects within each origination year. We find that the interaction effect increases over this timeframe. This period aligns with a simultaneous increase in the presence of non-bank lenders in the U.S. mortgage market. Our findings indicate that, when controlling for year effects, the influence of search efforts among financially savvy borrowers increases over time. Consequently, we argue that financial skills and search activity are increasingly pivotal in explaining mortgage rate disparities among U.S. households.

In our estimates, we go beyond the mortgage origination and observe borrowers' loan performance scores over time. We find that financially illiterate borrowers are 35-45% more likely to be late with their payments three years after the mortgage originated. Given that our estimates control for the mortgage amount, credit score, and payment-to-income ratio of every mortgage, we interpret this result as a consequence of ill budgeting with low saving buffers in case of individual payment shocks.

Our findings represent a set of stylized facts for the mortgage attainment process in the U.S. In our subsequent work, we introduce a set of assumptions that correspond to our findings on the importance of individual search behavior and financial knowledge for mortgage rate attainment in the U.S.

2.2 Related Literature

This paper contributes to empirical studies on mortgage undertaking, refinancing, and financial literacy effects on individual mortgage performance. Our paper leverages the current way U.S. households face the mortgage process.

The empirical literature argues that financial literacy explains financial behavior in the credit market. Bhutta et al. (2020) use mortgage origination platform data and show that, even within the specific loan officer, there is a considerable amount of dispersion in interest rates among otherwise comparable borrowers[5]. Moreover, Gerardi et al. (2023) find significant race differences in mortgage prices, pertaining to more than income and education differences.

The losses from the mortgage contract go beyond the choice at origination and may come from refinancing mistakes. Our estimates from the SCF data corroborate findings in Agarwal et al. (2016) and show that financially unskilled households do not refinance as often. In the Danish environment, Andersen et al. (2020) attribute the mistakes to refinancing to individual inattention. Keys et al. (2016) find that more than 20% of U.S. borrowers did not refinance at the optimal time, when interest rates were low, and relate individual sub-optimality to procrastination and financial sophistication. We estimate individual refinancing probability differences across financial literacy scores while controlling for other observables.

Owing to the series of seminal papers (Lusardi et al., 2010; Lusardi and Mitchell, 2014; Lusardi et al., 2020), the correlation between individual financial literacy and portfolio choice and saving behavior has been well documented. Bhutta et al. (2022b) focus on liquid savings and show that financially unskilled households more often face liquidity constraints due to their low liquid buffers. Through the lens of our estimates, lower buffers may be coming from poor practice in mortgage choice.

Following empirical findings, Jappelli and Padula (2017) and Lusardi et al. (2017) introduce financial literacy in the portfolio allocation model while assuming that individual returns depend on the level of financial literacy. Our estimates suggest a mechanism that relates mortgage rate attainment and individual financial literacy through search effort. In this way, we introduce a search mechanism that we model in our subsequent paper.

In the European contexts, where the number of potential lenders is significantly lower, Damen and Buyst (2017) show that borrowers can save more than €7,078 over the mortgage term by shopping and comparing different mortgage products. Additionally, U.K. estimates show that young and inexperienced borrowers make costly mortgage choices (Coen et al., 2023).

Our estimates underscore the effectiveness of the mortgage search depending on individual financial literacy scores, as low-income borrowers may be searching out of fear and make costly choices (Agarwal et al., 2020). In this regard, the sign of the interaction between search and financial skills changes as individual incentives change.

[5] Specifically, Bhutta et al. (2020) compare borrowers with similar credit scores and characteristics searching for the same loan amount.

2.3 Data analysis and stylized facts

The empirical part of our paper stochastically merges two publicly available survey data sets, effectively defining a novel data set on U.S. mortgage originations. Leveraging on the robustness of stochastic imputations, we outline the set of estimates that highlight the importance of financial skills and search behavior in mortgage attainment. Whereas most of our inference is correlational, a novel dataset provides a causational explanation for mortgage performance a couple of years after the mortgage originated. First, we introduce the SCF data and present three stylized facts important for our model assumptions. Next, we introduce the second data source (NSMO) and later proceed to present the findings of the novel U.S. dataset (NSMO+) generated using the stochastic merging method.

2.3.1 The Survey of Consumer Finances

The SCF, a triennial survey of randomly chosen U.S. households, captures data on investment, housing, and debt. These responses construct a comprehensive balance sheet for typical U.S. households, which is vital for empirical household finance studies. Our analysis focuses on a SCF subset with a "financial literacy score," from the 2016 and 2019 waves, comprising 60,125 responses. By incorporating data on credit search behavior and mortgage refinancing, akin to the NSMO data, we explore credit shopping patterns among 41,788 first-lien mortgage holders and renters, aligning with NSMO standards.

2.3.1.1 Financial literacy

Financial literacy score is based on a set of three questions (*The Big Three*) that are shown to be efficient in comprehensively evaluating individual financial skills (Lusardi et al., 2010; Lusardi and Mitchell, 2014; Bhutta et al., 2022b). The set of questions tests individual understanding of inflation, risk diversification, and compounding:

1. Suppose you had $100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow?
   - More**/Exactly/Less than $102
   - Do not know/Refuse to answer
2. Imagine that the interest rate on your savings account was 1% per year and inflation

was 2% per year. After 1 year, how much would you be able to buy with the money in this account?

   - More/Exactly/Less** than today
   - Do not know/Refuse to answer

3. Please tell me whether this statement is true or false. "Buying a single company's stock usually provides a safer return than a stock mutual fund."

   - True
   - False**
   - Do not know
   - Refuse to answer

Unlike perceived financial knowledge, which signifies confidence, these objective scores provide insight into actual financial planning and behavior (Bhutta et al., 2022b; Lusardi et al., 2010). To explore this, we employ a stochastic merging procedure, integrating mortgage data with the SCF. This approach allows us to discern collective patterns in *objective financial skills*, search effort, and mortgage rates among comparable borrowers.

First, we highlight essential household characteristics pertaining to financial literacy. Utilizing an ordered logistic model, we predict financial literacy scores based on borrower attributes. Table 3 presents personal attributes associated with financial literacy. Model-generated probabilities indicate that college graduates correctly respond to all financial literacy questions with a probability of 77%, while high-school graduates do so with a probability of 52%. Additionally, Figure 8 offers empirical evidence demonstrating a positive correlation between educational attainment and financial literacy.

Although education explains a considerable portion of the variation in financial literacy, as evident from the significant and substantial coefficients in Table 3, income, age, and race also play significant roles. These factors highlight additional dimensions crucial for skills and, consequently, individual saving and borrowing behaviors. We consider financial skills as a dimension that encompasses these conventional explanatory variables, albeit imperfectly, due to the impacts of learning by doing and unexpected expense shocks, as discussed in studies such as Agarwal et al. (2007) and Lusardi and Mitchell (2014).

Table 3: Ordered logistic model, personal characteristics correlating with financial literacy. Source: SCF, 2016-2019, authors' calculations.

	Dependent variable:
	Financial literacy score
Worker	0.041*
	(0.025)
Married	0.111***
	(0.024)
Non-white	−0.392***
	(0.019)
Female	−0.474***
	(0.025)
Education: High-school	0.211***
	(0.031)
Some college	0.599***
	(0.031)
College degree	1.123***
	(0.033)
Income percentile: 20^{th} - 40^{th}	0.049*
	(0.028)
40^{th} - 60^{th} 3	0.073**
	(0.031)
60^{th} - 80^{th}	0.179***
	(0.035)
80^{th} - 90^{th}	0.349***
	(0.043)
90^{th} - 100^{th}	0.649***
	(0.048)
Observations	60,125
Note: Controlling for age and asset amount.	*p<0.1; **p<0.05; ***p<0.01

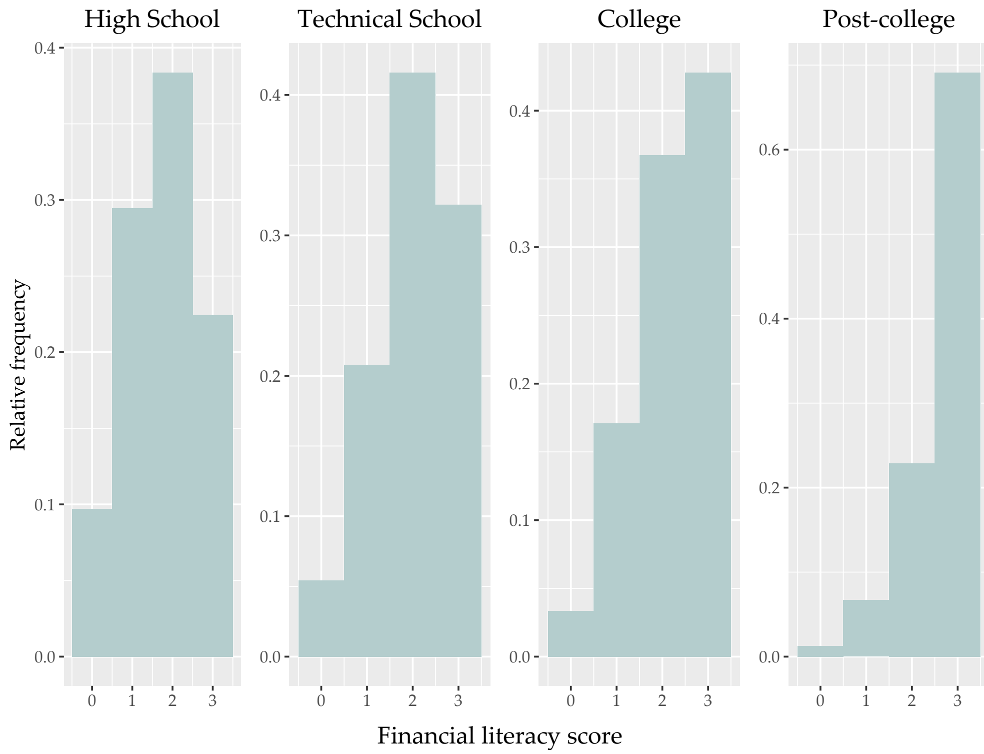


Figure 8: Financial literacy distribution by education level. Source: SCF, 2016-2019, authors' calculations.

2.3.1.2 Stylized facts from the SCF

While the separation of financial literacy from other household characteristics falls beyond the scope of this paper, we present key data patterns shedding light on individual financial skills and their potential impacts on mortgage shopping behavior.

First, we document that financial skills vary with age. We apply a polynomial fit to the standardized skill score across age groups. Although Figure 9 can not account for cohort effects, the hump-shaped fit corresponds to panel data estimates depicting skill variations over time (see Agarwal et al. (2007) and Lusardi et al. (2010)). Indicative of a decline in consumer finance knowledge with approaching retirement, Figure 9 illustrates skill depreciation, corroborating findings from panel-data studies on financial sophistication.

The second empirical fact underscores the positive correlation between refinancing probability and financial literacy. Our analysis reveals that the likelihood of mortgage refinancing increases with higher financial skills and mortgage payments, holding other characteristics constant. Variations in these probabilities are illustrated in the heatmap depicting predicted refinancing probabilities in Figure 10.

We evaluated the likelihood of mortgage refinancing among borrowers based on their self-

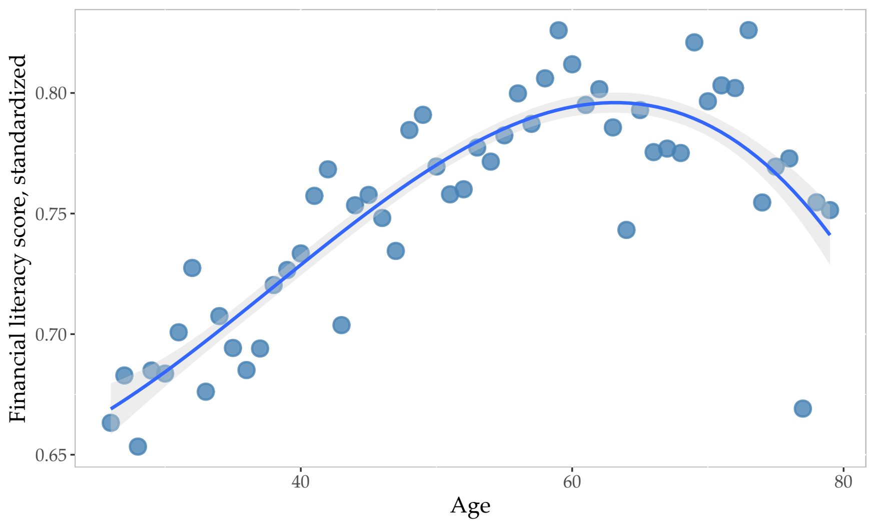


Figure 9: Average financial literacy by age groups, polynomial fit. Source: SCF 2016-2019, authors' calculations.

reported search efforts in making borrowing decisions. With borrower attributes and mortgage size held constant, greater financial literacy, income, and effort imply a higher likelihood of mortgage refinancing (as illustrated in Table 27 in the Appendix). In contrast, Table 4 demonstrates that education does not significantly influence refinancing. Thus, financial skills emerge as a distinct dimension significantly impacting refinancing decisions within the SCF dataset.

Overall, coefficients in Table 4 imply that, across all income categories, financially savvy borrowers are 20%-30% more likely to refinance their mortgage.

Our third finding highlights a positive correlation between financial skills and the time households dedicate to credit shopping. Employing an ordered logistic model, we find that financially savvy renters and homeowners invest a significant amount of time in credit shopping, regardless of their housing expenses. The coefficient estimates are detailed in Table 5, and Figure 11 illustrates a heatmap showing model-predicted probabilities of spending a considerable amount of time searching for credit among renters. Households with strong financial skills tend to allocate more time to exploring credit opportunities, with a 15% increase in the likelihood of spending additional time for mortgage owners and a 10% increase for renters. Furthermore, our estimates indicate that renters, on average, dedicate less time to search efforts, and their search intensity shows a more gradual growth with higher levels of financial skills[6].

In the SCF, an average homeowner has over 70% of their total monthly debt obligations dedi-

[6] The heatmap of predicted probabilities for homeowners is available in Appendix B.3, Figure 41.

Table 4: Binary regression estimates, likelihood of refinancing. Source: SCF 2016-2019, authors' calculations.

	Dependent variable:
	Ever refinanced their mortgage
Financial literacy score: low	0.093
	(0.122)
medium	0.262**
	(0.116)
high	0.478***
	(0.115)
Search effort, borrowing: medium	0.055
	(0.056)
high	0.125**
	(0.058)
Education: high school	−0.106
	(0.081)
some college	−0.222***
	(0.081)
college degree	−0.089
	(0.080)
Female	0.103*
	(0.057)
non-white	−0.280***
	(0.037)
Mortgage size: $83,000 - $159,000	−0.170***
	(0.047)
$159,001 - $ 297,000	−0.360***
	(0.049)
$ 297,001 - $ 1,450,000	−0.394***
	(0.054)
Constant	−0.869***
	(0.175)
Observations	18,702
Note: Controlled for age, income, family structure and survey wave effects.	*p<0.1; **p<0.05; ***p<0.01

Table 5: Ordinal logistic regression, time spent shopping for credit. Source: SCF 2016-2019, authors' calculations.

	Low-to-great deal of spent in shopping for credit(1-3)	
	Homeowners	Renters
Low|Medium	-15.343^{***}	0.439^{***}
	(0.236)	(0.086)
Medium|Great	-18.042^{***}	-1.748^{***}
	(0.237)	(0.090)
Mort. payment per month: -\$750-\$1150	-0.017	
	(0.049)	
\$1150-\$1700	0.038	
	(0.053)	
\$1700-\$2700	0.0314	
	(0.060)	
\$2700+	0.071^{***}	
	(0.056)	
Rent payment per month: \$500-\$690		-0.132^{**}
		(0.046)
\$690-\$920		-0.058
		(0.047)
\$920-\$1300		0.029
		(0.048)
\$1300+		0.0385
		(0.052)
Education: HS	0.421^{***}	0.373^{***}
	(0.074)	(0.048)
some college	0.436^{***}	0.612^{***}
	(0.074)	(0.048)
college degree	0.437^{***}	0.565^{***}
	(0.075)	(0.053)
Wage percentile: 20-40	-0.0368	0.147^{**}
	(0.059)	(0.051)
40-60	-0.016	0.140^{*}
	(0.061)	(0.056)
60-80	-0.051	0.122^{*}
	(0.063)	(0.058)
80-100	-0.097	0.260^{***}
	(0.068)	(0.062)
Financial literacy: level 1	0.256	0.090
	(0.112)	(0.065)
level 2	0.400^{***}	0.161^{***}
	(0.106)	(0.062)
level 3	0.350^{***}	0.360^{***}
	(0.105)	(0.064)
Observations	22,178	19,610

Note: Controlled for gender, race, age, debt-to-income, risk attitudes, assets, and survey wave effects. *p<0.1; **p<0.05; ***p<0.01

Figure 10: Mortgage refinance likelihood across income percentiles and financial literacy scores. Source: SCF 2016-2019, authors' calculations.

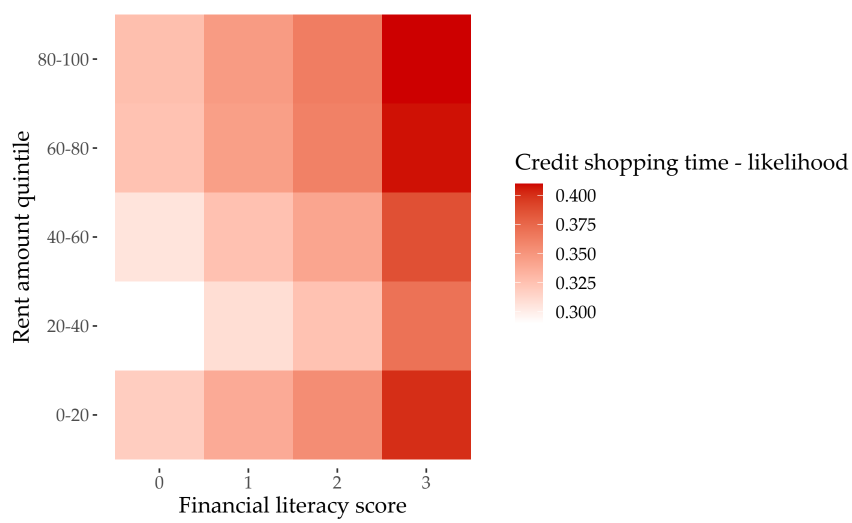


Figure 11: Great deal of time spent shopping for credit, ord. logit predictions, renters only. Source: SCF 2016-2019, authors' calculations.

cated to mortgage repayments. Consequently, the specifics of a mortgage contract significantly influence expenditure and savings patterns throughout their working years, deeply impacting available liquidity. In this context, we obtain a dataset that is comprehensive, encompassing detailed information on both the mortgage contract and household characteristics. Shifting our attention to mortgage data, we gain insights into individual mortgage shopping behavior. Individual shopping behavior, coupled with a standard set of observable factors, determines the mortgage interest rate, which frequently remains fixed over the mortgage term.

2.3.2 The National Survey of Mortgage Originations (NSMO)

Our novel data set leverages the amount of information within the NSMO. For a representative sample of the U.S. population, NSMO connects mortgage registry data to the survey on mortgage acquisition experience, spanning mortgage originations from 2013 to 2021. This survey includes newly originated first-lien residential mortgages, covering both initial acquisitions and refinances. Important for our paper, the survey inquires about loan shopping behavior and the overall consumer experience during the mortgage process. All survey responses are matched with institutional lender data, providing specific details of the mortgage contract, including locked-in mortgage rates, government sponsorship, low-income area indicators, loan-to-value ratios (LTVs), borrower's payment-to-income ratio, credit score, education, and income. We limit the data to home purchases and refinancing, resulting in a survey sample of 43,094 mortgages, each weighted to ensure representativeness in our analysis.

Our focus revolves around borrowers' search behavior prior to the mortgage application. We use the question

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

The individual survey responses serve as a proxy variable for the cognitive search effort. Instead of relying on the number of formal mortgage applications, we analyze the number of lenders considered. We argue that the response reveals the variation in the cognitive search effort **prior to the application process**.

While the majority of borrowers tend to submit formal applications to a single lender – resulting in over 35,000 mortgages being obtained from that chosen lender – the number of lenders seriously taken into account varies across the sample. We assert that, due to the expense associated with the application process, borrowers concerned about rejection are more likely to apply to multiple lenders, driven by fear of being declined. This phenomenon has been discussed in works such as Agarwal et al. (2020). Consequently, the number of lenders considered reveals shopping behavior that provides deeper insights into cognitive efforts invested into the attainment process. Important for our paper, approximately 70 percent of the survey respondents undergo the mortgage process without the use of a mortgage broker.

Furthermore, the number of lenders considered reflects the contemporary approach to mortgage exploration. Online applications typically compare various lenders and "recommend" the optimal choice, considering the borrower's credit score, income, and down payment options [7].

In Figure 12, we depict the raw data estimates to give a preview of search effort variation across different financial skill levels. Low-skilled borrowers predominantly concentrate on a single lender, while high-skilled borrowers frequently consider two, three, or more lenders. While our paper's foundation leverages financial skills data acquired through stochastic matching, the appendix demonstrates how locked-in mortgage rates fluctuate in relation to education and search effort. Leveraging the matched dataset, we introduce the concept of **effective search** among borrowers with higher skills and education. Thus, the rest of our analysis remains concentrated on financial skills.

After the mortgage origination, the NSMO tracks individual mortgage performance until loan closure. Conditional on averages in other borrower characteristics, our estimates underline financial skills and search behavior as being significant in predicting meeting payment due dates.

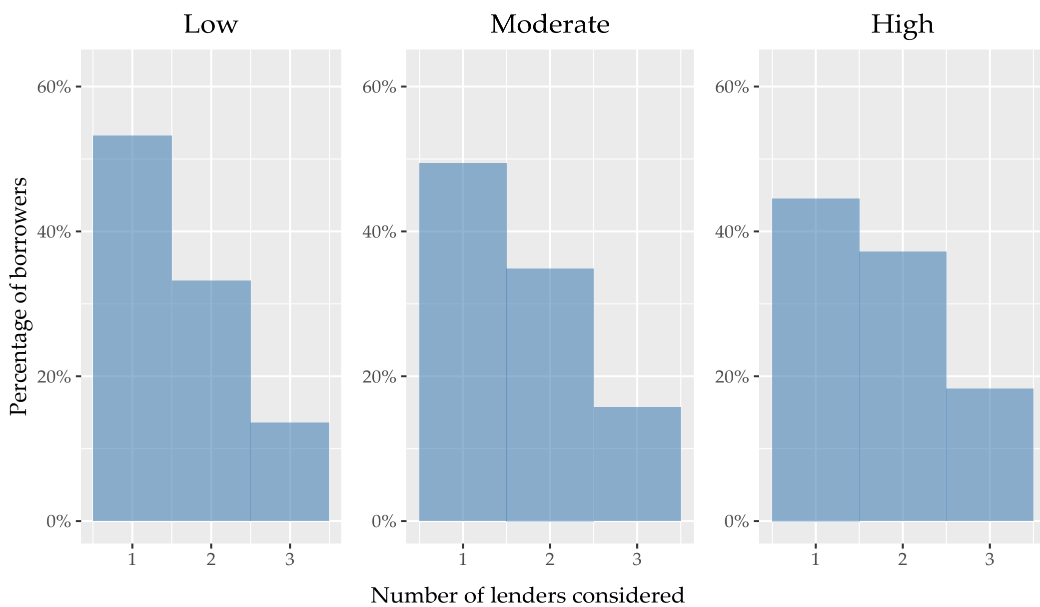


Figure 12: The number of lenders considered at the time of loan origination, across financial skill level, left-to-right panel. Source: NSMO+ data, authors' calculations.

[7]For instance, a consumer can visit https://www.bankrate.com/mortgages/mortgage-rates/ and input their current or desired mortgage amount to compare rates across lenders.

2.3.3 Stochastic imputation, mortgage data extended (NSMO+)

Information regarding individual mortgages is limited within the SCF. Beyond mortgage payments and past refinancing behavior, data on a mortgage contract is unavailable. To overcome this limitation, we employ stochastic matching to integrate the two datasets. By doing so, we maximize the utility of publicly accessible information about mortgage contract specifics and individual skills, and account for the uncertainty inherent in the matching process.

Instead of imputing financial literacy scores deterministically, the BRL method estimates the distribution of financial skill level for every borrower in the NSMO. Based on the set of mutual observables, we obtain Bayesian weights for every match between NSMO and the SCF, and use them later for statistical inferences. This method has been analytically shown to reduce the biases in coefficient estimates in linear models and preserve asymptotic normality and consistency in non-linear estimation (Enamorado et al., 2019). We outline the BRL assumptions and likelihood formulation in section B.4 of the Appendix.

Our paper is the first to link SCF and NSMO. Record matching allows us to estimate the financial skill distribution for every NSMO borrower. While Bayesian weights control for the imputation-driven bias, details of the mortgage contract allow us to control our estimates for other borrowers and mortgage specifics. In this way, our estimates reflect potential sources of the mortgage rate dispersion among otherwise similar borrowers who apply for similar contracts. Table 6 outlines population shares in respective data sources. The selection of common observables we base our matches on are measures relevant to individual financial skills, including income, education, gender, age, race, occupation, family characteristics, and retirement plan and asset holdings. Once we have a borrower-specific skill distribution, our estimates separate skilled and unskilled borrowers who search more or less, keeping the lender's side of the contract fixed (term, amount, government sponsorship, origination year, etc.)

NSMO+ data findings

In this section, we outline joint patterns in mortgage rates, individual search effort, and financial skills and discuss individual mortgage performance across skill levels. Initially, we discuss the importance of financial skills and their role in how much search effort is exerted prior to mortgage attainment. Next, we delve into the interplay between financial skills, search effort, and mortgage rates and introduce the concept of **effective** search among skilled borrowers. Lastly, we focus on repayment behavior heterogeneity across different skill levels.

Table 6: Population shares in the respective sample. Source: NSMO 2013-2022 and SCF 2016-2019, authors' calculations.

	Data set NSMO	SCF
income brackets	[6%, 9% , 18%, 19%, 30%, 18%]	[13%, 8%, 13% ,11%,20%, 35%]
education brackets	[1%, 10%, 5%, 20%, 35%, 29%]	[6%, 18%, 9%, 15%, 27%, 25%]
gender (Female,Male)	[44%, 55%]	[17%,83%]
age (<35,35-44,45-54,55-64,65-74,>=75)	[18%, 22%, 22%, 21%, 14% ,3%]	[8%, 14%, 20%, 26% , 20%, 12%]
race (Caucasian, African-American, other)	[84%, 6%, 10%]	[82%, 7%, 11%]
occupation (Employed, Self-employed, Retired/Student, Other)	[68%, 10%, 19% ,2%]	[47%, 26%, 25%, 2%]
has children (Yes, No)	[64%, 36%]	[60% , 40%]
owns financial assets (Yes, No)	[57%, 43%]	[58% 42%]
retirement plan participation (Yes, No)	[86%, 14%]	[62%, 38%]

2.3.3.1 Search, financial skills and locked-in mortgage rates

Using imputed financial skills, we find that financially savvy borrowers consider more lenders on average, and show that search effort variation patterns resemble the breakdown by education level (see Figure 38 in section B.2 of the Appendix). Moreover, we find that savvy applicants search more effectively and generally secure lower mortgage rates in comparison to their comparable counterparts.

2.3.3.2 Search effort and financial skills

In our sample, we redefine the number of lenders considered and bin 3, 4, and 5+ together and represent it with 3+. Our estimates show that while 60% of low-skilled borrowers focus on only one lender and only 10% on three or more lenders, 58% of financially savvy borrowers consider multiple lenders (Table 7).

Next, we estimate an ordinal logistic model that assumes latent thresholds for every observation ij in the merged data set

$$\mathbb{P}(\text{num\_cons}_{ij} = k) = p_{ij,k} = \mathbb{P}\Big(-\kappa_{k-1} < \beta X_i + \beta^f \text{fin\_skills}_j + u_{ij,k} < \kappa_k\Big), \quad k \in \{1, 2, 3+\}.$$

We adjust our estimates with borrower-skill specific distributional weights that account for

Table 7: Number of lenders considered across financial skills, weighted frequencies. Source: merged dataset, authors' calculations.

	Number of lenders considered		
	1	2	3+
Financial Literacy			
Low	58.48%	41.52%	0
High	41.37%	36.42%	22.21%

match uncertainty in the inflated set of 155,500 observations[8].

Table 8 depicts the explanatory power of each borrower characteristic. Important to our narrative, our estimates imply that financially skilled borrowers (top tercile) are 4% more likely to consider more lenders, i.e., search more. Moreover, we find that females and borrowers living in non-metropolitan areas are 30 and 5 percent less likely to consider multiple lenders. Additionally, education significantly affects search effort, as we find that college graduates and post-college borrowers are 40% and 50% more likely to search more, respectively.

Search effort correlates negatively with low-to-moderate non-metropolitan areas, known as low-shopping areas, which are often subject to mortgage overpricing (Bartlett et al., 2022). Notably, the effect of financial skills is of the same magnitude as income or credit score, or the geo-location effect[9]. Abstracting from all standard observables leaves a significant residual effect of financial skills. However, the skills effect in our estimates remains conservative due to the nature of our merging process and strong correlations between skills and gender, income, education, etc., outlined in the SCF data analysis.

2.3.3.3 Residual mortgage rate dispersion and repayment costs heterogeneity

Next, we turn to the mortgage rate dispersion, controlled for mortgage specifics. We focus on differences in mortgage rates across individual financial skills and search effort.

Controlling for the loan amount, term (30 years), borrower's credit score (*"Very good" and "excellent"*), and the origination year (fixed to 2016), we compare the residual mortgage rate dispersion across different levels of financial skills. Even though these borrowers are comparable to mortgage lenders, financially savvy ones tend to lock in at lower rates. Figure 13 shows that the interest rate density for the savviest borrowers (denoted with the blue curve) has a lower mean and is thicker towards lower interest rates. On the other hand,

[8] We repeat the analysis with the linear probability model that does not require weights inclusion and obtain similar results

[9] In addition, our SCF analysis shows significant variation of credit search effort with financial literacy, with 20% higher likelihood for high-skilled borrowers to spend more time in loan shopping.

	Dependent variable: # of lenders considered		
	Coefficient	SE	z score
(Intercept):1|2	-0.4515^{***}	0.0947	-4.7665
(Intercept):2|3	-2.1960^{***}	0.0950	-23.1239
Financial literacy	0.0444^{**}	0.0216	2.0616
Age	-0.1603^{***}	0.0143	-11.1923
Credit score	0.0515^{***}	0.0146	3.5298
Female	-0.2904^{***}	0.0141	-20.5282
Race: non-white	0.2426^{***}	0.0198	12.2247
Income:			
$\$35,000 - \$49,999$	-0.0262	0.0379	-0.6922
$\$50,000 - \$74,999$	-0.0312	0.0356	-0.8767
$\$75,000 - \$99,999$	-0.0172	0.0364	-0.4734
$\$100,000 - \$174,999$	-0.0351	0.0362	-0.9685
$\$175,000+$	-0.0227	0.0401	-0.5659
Metropolitan area:			
Low-to-moderate income	-0.0176	0.0215	-0.8195
Non-metropolitan area	-0.0517^{*}	0.0237	-2.1834
Loan Amount:			
$\$100,000$-$\$199,999$	0.0852^{***}	0.0231	3.6859
$\$200,000$-$\$299,999$	0.1864^{***}	0.0260	7.1664
$\$300,000$-$\$399,999$	0.2337^{***}	0.0305	7.6579
$> \$400,000$	0.3157	0.0324^{***}	9.7351
Education:			
some college	0.2657^{***}	0.0249	10.6772
college	0.4228	0.0247^{***}	17.1297
post-college	0.5302^{***}	0.0264	20.0973
Observations			155,500
Note: controlled for year effects.			$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 8: Ordered logit with imputed financial literacy and weights.

unskilled borrowers are more likely to end up with higher interest rates, as shown in Figure 13 with the red density graph.

Using the 2020 origination subsample, we show that, for a \$200,000 loan, the top tercile of financially skilled borrowers secured mortgages with a **20 percent lower spread in the mortgage rate distribution**, underscoring the larger variation in interest rates obtained by low-skilled borrowers. This pattern holds consistently over time, with the usual spread difference ranging between 15% and 20%.

Next, we regress the locked-in interest rate on a set of borrower characteristics X_i, mortgage

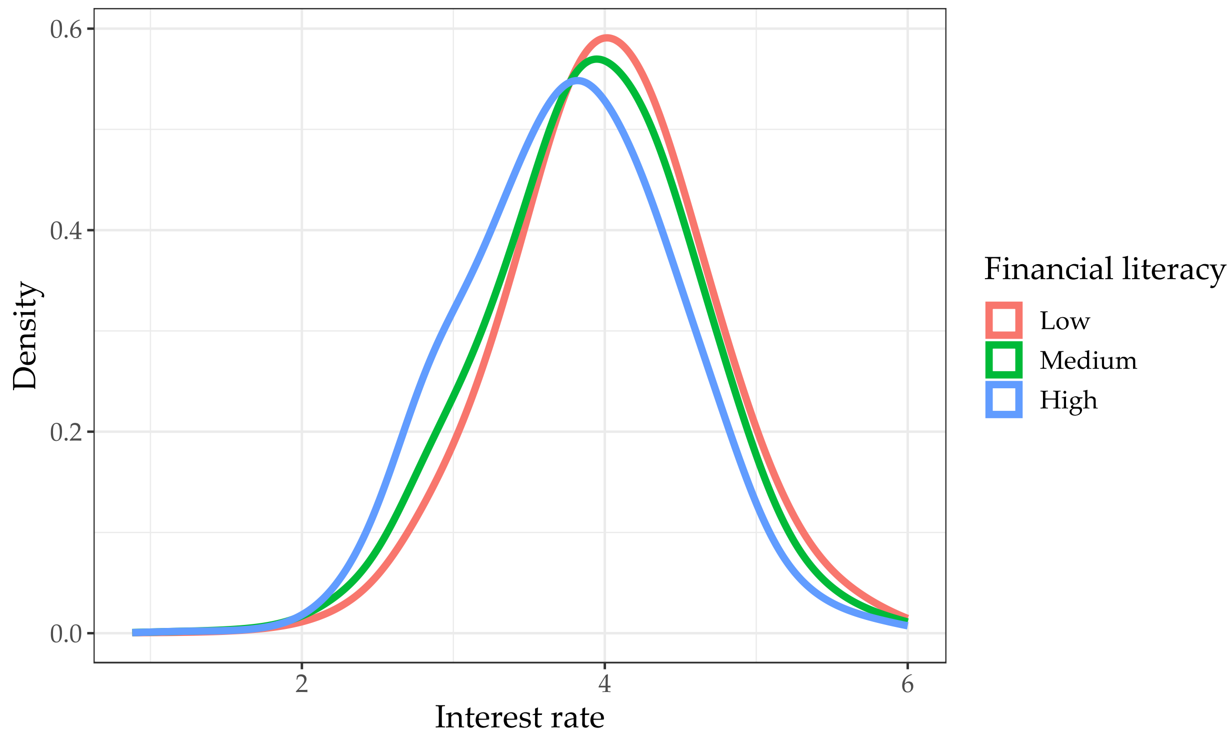


Figure 13: Residual mortgage rate across financial skills. Source: merged data set, authors' calculation.

contract specifics M_i and match-based financial skills fin_skills_i:

$$\text{rate}_i = \alpha + \beta X_i + \beta^m M_i + \beta^f \text{fin\_skills}_i + \gamma \text{fin\_skills}_i \times \text{num\_len}_i + \varepsilon_i,$$

and estimate the rate-based losses over the mortgage duration.

Table 9 displays coefficients for two sets of estimates, with the first column focusing solely on first originations. In both regressions, we account for mortgage specifics, including loan type, amount, term, sponsorships, number of underwriters, and loan-to-value ratios. Notably, both sets of estimates reveal an interaction between financial literacy and search effort, significantly contributing to the explanation for locked-in mortgage rates.

Initially, our findings align with those of Agarwal et al. (2020), showing that fear of application rejection mechanically amplifies search efforts among first originations, ultimately leading to higher average rates. This is highlighted in Table 9, which reveals a significant and positive coefficient of 0.220 for search effort within the context of first originations. Upon interaction with skills, the intensity of the search assumes the role of an informed mortgage search. Financially skilled borrowers who explore a wider range of lenders tend to secure lower mortgage rates. This translates to an average rate difference of 13.4 basis points (with a

Table 9: Mortgage rate regression, controlling for loan and borrower characteristics. Source: merged data set, authors' calculations.

	mortgage rate	
	(First origination)	(All mortgages)
#Lenders considered: two	0.034	−0.006
	(0.087)	(0.062)
#Lenders considered: three	0.220*	0.125
	(0.120)	(0.083)
Financial skills	0.017	−0.016
	(0.088)	(0.060)
Considered 2 lenders× fin skills	−0.072	−0.023
	(0.113)	(0.080)
Considered 3 lenders × fin skills	−0.354**	−0.220**
	(0.153)	(0.106)
Age	0.044***	0.062***
	(0.010)	(0.007)
Metro area - LMI tract	0.033**	0.022**
	(0.013)	(0.009)
Non-metro area	−0.018	0.003
	(0.015)	(0.010)
Female	0.032***	0.030***
	(0.009)	(0.006)
African-American	−0.005	0.007
	(0.019)	(0.013)
Asian	−0.021	−0.036***
	(0.020)	(0.013)
Other (including hispanic)	0.069***	0.051***
	(0.025)	(0.017)
Income: $35,000-$50,000	0.007	−0.043**
	(0.024)	(0.017)
$50,000-$75,000	0.036	−0.018
	(0.023)	(0.016)
$75,000-$100,000	0.034	−0.011
	(0.024)	(0.017)
$100,000-$175,000	0.064***	0.004
	(0.024)	(0.017)
$175,000 and more	0.054**	−0.00004
	(0.027)	(0.019)
Education: high-school	−0.054***	−0.033***
	(0.017)	(0.011)
college graduate	−0.105***	−0.071***
	(0.017)	(0.012)
post-college graduate	−0.131***	−0.090***
	(0.019)	(0.012)
Refinancing		−0.074***
		(0.007)
Credit score	−0.263***	−0.247***
	(0.010)	(0.007)
Constant	5.269***	4.955***
	(0.099)	(0.066)
Observations	21,461	43,084
R^2	0.369	0.440
Adjusted R^2	0.368	0.439
Residual Std. Error	23.662 (df = 21412)	22.325 (df = 43034)
F Statistic	260.809*** (df = 48; 21412)	689.013*** (df = 49; 43034)

Note: Controlled for loan type, government-sponsored enterprise, loan amount, number of borrowers, time effects, LTV and term. *p<0.1; **p<0.05; ***p<0.01

corresponding coefficient of 0.220-0.354=-0.134).

Our supplementary findings align with existing research employing loan-level data, underscoring that female and Hispanic borrowers often encounter higher mortgage rates. On the flip side, individuals with higher education enjoy, on average, a reduction of 13.1 basis points in rates during initial originations, though this effect decreases during refinancing. As we consider the intricate interplay among skills, gender, race, and education, our estimates concerning skill disparities present a cautious estimate of the minimum divergence in mortgage repayments, subsequently impacting differences in consumption after accounting for mortgage payments.

Nevertheless, when we analyze the variations in search effort and interest rate regressions, it becomes evident that the extent and effectiveness of search effort differs based on financial skills. This implies the likelihood of lower mortgage payments among financially skilled yet comparable borrowers.

2.3.3.4 Effective search

We emphasize the role of effective search and compare our predicted distributions of locked-in rates between borrowers who engage in extensive searches and those who consider one lender only. Figure 14 depicts mortgage rate distributions across two scenarios. Low-skilled borrowers that search more effectively do not gain from the search, as the mortgage rate distribution stays the same (left panel in Figure 14). In contrast, high-skilled borrowers who search more end up with lower rates (depicted by the blue curve in the right panel of Figure 14), rendering their search as effective. Our findings on search effectiveness, coupled with a significant and positive search coefficient in the interest rate regression (Table 9), align with the fear of rejection mechanism among low-income borrowers in Agarwal et al. (2020). Less financially savvy borrowers search more because they fear rejection. As a result, this does not significantly change their mortgage rates compared to those who put in less effort.

The disparities observed in lock-in rates during the origination phase ultimately translate into compounded losses over the entire mortgage term[10]. To illustrate, for a $\$100,000$ loan with a standard duration, an average borrower with high financial skills can secure a rate of approximately 3.8%, compared to 4.05% for those with lower financial skills. This sets the lower boundary for cumulative losses at $\$6,693$ over the mortgage term. Moreover, the additional impact of low search effort introduces more than $\$2,636$ in costs throughout the mortgage term. These estimates, though not accounting for other correlations among

[10]Over 75% of mortgages in our sample are 30 years fixed-rate mortgages.

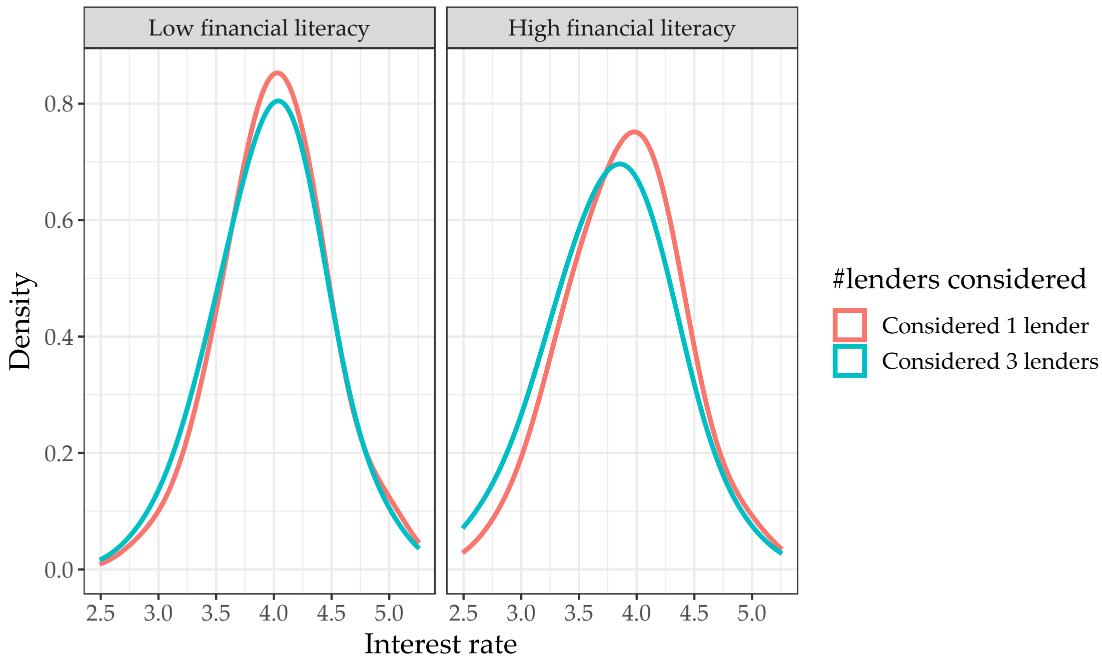


Figure 14: Mortgage rate dispersion; interaction of search effort and financial skills. High skilled borrowers who exert more search effort generally lock in at lower mortgage rates. Source: merged data set, authors' calculation.

borrower characteristics, stand as conservative approximations for losses in the mortgage market, amounting to at least $9,329. Notably, this represents a significant proportion of the losses derived from institutional data and subjective insights into the mortgage process (Bhutta et al., 2020). Given that mortgage repayments accounts for over 70% of monthly debt payments, addressing these losses is an imperative for bolstering liquidity for all households, especially those with lower incomes.

Figure 15 represents the year and financial skills interaction coefficient over the sample period. Relative to the first year in the sample, 2013, later mortgage origination years show signs of increasing significance of both financial skills and search effort for mortgage rate attainment. Our sample period is marked by the steady increase in non-bank lenders share in the mortgage market. As these lenders turn to online advertising and borrowing (Bhattacharya et al., 2021), our findings are suggestive of increasing effects of skilled search effort amidst the mortgage options expansion.

2.3.3.5 Mortgage performance after origination

NSMO+ tracks the individual mortgage performance until the loan closure, with scores denoting missing repayment due dates up to and over 180 days, bankruptcy levels based on

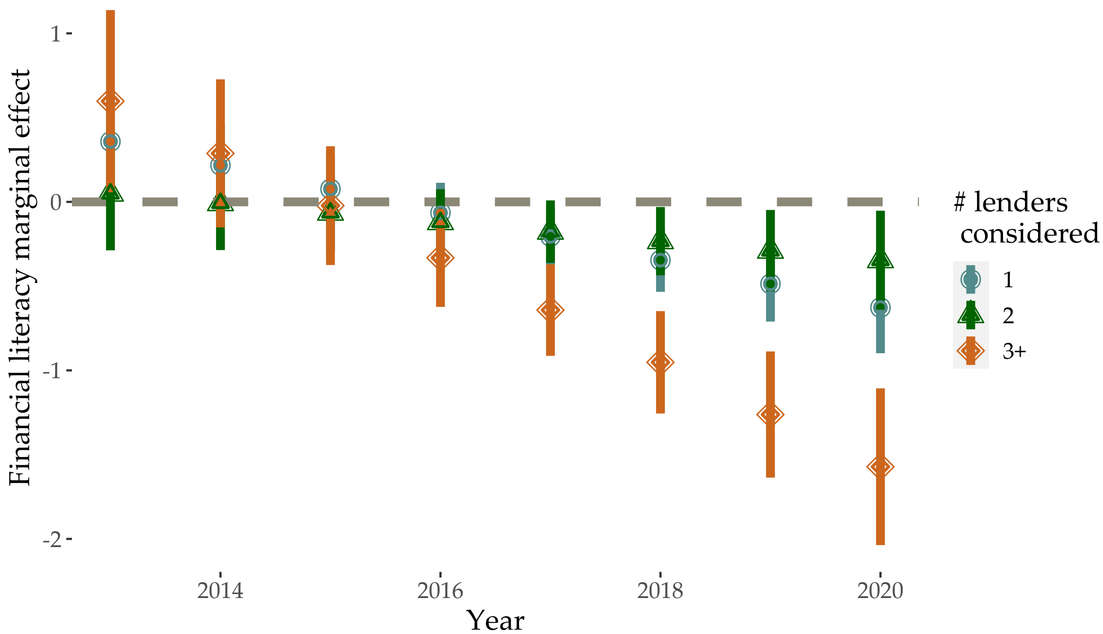


Figure 15: Financial skill coefficient in the mortgage rate regression, differences over the sample period. Source: merged data set, authors' calculations.

U.S. law, and regular payments made on time. Specifically, the data set separates scores for late payments up to 150 days, and the worst scores indicates mortgage payments later than 150 days and defaults.[11].

The sample size constrains our analysis of the default and late payment indicators, so we separate the score values for late payments and defaults from regular payments and define the indicator variable $\mathbf{1}_{\{\text{late payments or defaults}\}}$. We quantify the effect of individual financial skills and search effort at the time of origination using the linear probability model estimation that controls for other observables.

We model the probability as

$$\mathbb{P}(\text{late with payments}) = \alpha + \beta X_i + \beta^f \text{fin\_lit}_i + \beta^s \text{search\_effort}_i + \varepsilon_i,$$

where fin_lit_i is the average skill amount across all matches[12]. We regress the indicator on a set of borrower observables, mortgage characteristics, individual financial skills, and search effort at the time of origination.

[11]According to the Home Mortgage Disclosure Act data, delinquency rates are reliable indicators of mortgage default. https://www.consumerfinance.gov/data-research/mortgage-performance-trends/mortgages-30-89-days-delinquent/

[12]We perform a separate, score-based analysis that shows the significance and similar effect size.

We standardize all continuous regressors (age, credit score, payment-to-income ratio) and compare the size of the coefficients. Our estimates are presented in Table 10.

Table 10 conforms to the standard intuition regarding household characteristics prevalent for mortgage performance. While borrowers with greater payment-to-income ratio are more likely to be late, those with higher credit scores are more likely to meet their payment due dates. In line with Gerardi et al. (2023) and Bhutta et al. (2020), we find that non-white borrowers are more likely to be late with payments. Importantly for our paper, financially skilled borrowers who exerted more effort are less likely to have been late on payments two years after mortgage origination.

Figure 16 plots default prediction differences across different skill and search levels. Specifically, our predictions state that financial unskilled face a 1.6 p.p. higher likelihood of being late with mortgage payments. Added to this, borrowers who considered one lender are 0.2 p.p. more likely to be late with payments, possibly because they secured their mortgages at higher rates. Put differently, getting one more question wrong in the financial literacy test corresponds to being 40%-50% more likely to not meet mortgage repayments dates three years after the origination.

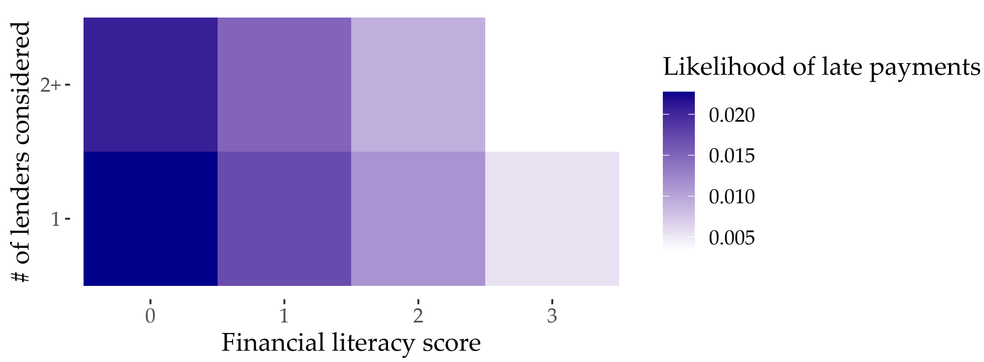


Figure 16: Likelihood of late payments across effort and financial skills. Source: Probability model predictions, merged data set, authors' calculation.

The patterns identified through our analysis of the SCF and NSMO+ serve as the foundation for a mortgage search model that accounts for the variation in search costs contingent on individual financial skills. We revisit each of these findings within the framework of our model setup and explore their implications in our analysis of the steady state.

Table 10: Late payment probability, linear model. *Source:* merged data set, authors' calculation.

	$\mathbb{P}$*(Late payment)*
Loan Amount: \$100,000-\$199,999	0.0001
	(0.002)
\$200,000-\$299,999	-0.004^{**}
	(0.002)
\$300,000-\$399,999	-0.004^{**}
	(0.002)
> \$400,000	-0.005^{***}
	(0.002)
Financial literacy	-0.017^{**}
	(0.007)
Multiple lenders considered	-0.002^{**}
	(0.001)
Female	0.002^{*}
	(0.001)
Education: high-school	0.003
	(0.002)
college	-0.0001
	(0.002)
post-college	-0.0002
	(0.002)
Race: non-white	0.005^{***}
	(0.001)
Age	0.002^{*}
	(0.001)
Payment-to-income	0.005^{***}
	(0.001)
Credit Score	-0.020^{***}
	(0.001)
Constant	0.023^{***}
	(0.005)
Observations	43,084
Adjusted R^2	0.017
F Statistic	54.783^{***} (df = 14; 43069)
Note: all variables are standardized to preserve interpretability.	*p<0.1; **p<0.05; ***p<0.01

2.4 Conclusions

Our paper contributes to the empirical literature on mortgage undertaking in two ways. First, we employ the stochastic record linkage procedure and merge the National Survey of Mortgage Originations with the Survey of Consumer Finances, effectively creating a new data set on mortgages that incorporates objective financial literacy scores. Second, we leverage the statistical properties of the merging procedure and investigate the joint correlation between individual financial literacy and search effort in the mortgage undertaking process while accounting for specific record link uncertainty. Third, our findings introduce a novel search mechanism that connects individual financial literacy and mortgage rate attainment.

Our data estimates show that financially skilled households seriously consider multiple lenders more often, showing signs of an effective search procedure. Moreover, we show that financial literacy and search interact and explain a part of the mortgage rate variation. Specifically, skilled borrowers who search more end up getting a 13.4 b.p. lower interest rate at the time of the origination. Using back-of-the-envelope calculations, we estimate the lower bound for potential losses from unskilled search - for a $100,000 loan, financially unskilled borrowers lose at least $9,329 dollars over the thirty-year mortgage span.

Our paper speaks to behavior after the mortgage was originated. Using our novel data set, we show that financially unskilled households face a 34-45% higher likelihood of becoming delinquent three years after the mortgage originated, irrespective of their payment-to-income ratio. This finding, coupled with our findings on lower refinancing probability among financially unskilled households, motivates the importance of the mortgage search mechanism for consumption differences across similar borrowers.

3 Tax Structures and Fiscal Multipliers in HANK Models

Co-authored with Othman Bouabdallah (ECB) and Pascal Jacquinot (ECB)

3.1 Introduction

There is a long tradition in the literature of assessing the effect of an increase in government spending on aggregate economic responses. However, standard DSGE models do not capture heterogeneity, i.e., important distributional aspects such as inequality. Using the Eurosystem Household Finance and Consumption Survey, we document heterogeneity in HtM (hand-to-mouth) status and household asset holdings in liquid and illiquid accounts for a set of European countries.

Following Kaplan et al. (2014) and Kaplan and Violante (2014), we make an important distinction between different types of HtM households with respect to asset holdings. On the one hand, poor hand-to-mouth (pHtM) households have little or no liquid wealth and no illiquid wealth. On the other hand, the wealthy HtM (wHtM) also hold little or no liquid wealth but hold positive amounts of illiquid assets. The third group of households, non-HtM households, hold positive amounts in their liquid accounts. Both pHtM and wHtM households have a large marginal propensity to consume (MPC) out of small transitory income fluctuations. However, Kaplan et al. (2014) show that wHtM households are similar to non-HtM households along several dimensions. They emphasize the distinction between three groups of households, which we also show in our analysis as an important component. In this paper, we analyze how the fiscal multiplier (elasticity of output with respect to government spending) depends on household heterogeneity in HtM status and asset holdings.

Usually, to answer questions regarding fiscal multipliers, the literature analyzes the U.S. economy. However, the case in Europe is different and more granular. For example, EU countries are heterogeneous in their debt-to-GDP ratios. On the one hand, some countries, including Italy and France, have debt-to-GDP ratios well over 100% (Bezhanova et al., 2023). On the other hand, smaller countries, such as Luxembourg and Estonia, have a debt-to-GDP ratio below 30% (Bezhanova et al., 2023).

Moreover, Qiu and Russo (2023) show that European countries are heterogeneous in tax levels and income tax progressivity. Therefore, in this paper, we pose multiple questions. First, we aim to answer how the fiscal multiplier depends on different taxation schemes and debt levels. Second, we answer the question of how the fiscal multiplier depends on different sources of financing of government spending. Next, we are interested in the role of household heterogeneity and distributional moments in explaining aggregate movements.

We build a quantitative Heterogeneous Agent New Keynesian (HANK) model with liquid and illiquid assets and a rich set of fiscal policy instruments to answer these questions. We show that financing government spending through deficit, in general, implies higher fiscal multipliers. Moreover, financing deficit with non-distortionary government transfers implies the highest positive long-term impact on output. More specifically, lump-sum transfers circumvent individual frictions in liquidity transformation and increase demand among liquidity-constrained households. Aligned with literature, e.g., Hagedorn et al. (2019) and Auclert et al. (2018), we show that RANK (Representative Agent New Keynesian) and Two Agent New Keynesian (TANK) models cannot produce the size of fiscal multipliers or consumption response consistent with the data.

Broer et al. (2023) highlight the fact that the transmission of fiscal shocks in the (New Keynesian) NK setting is rather different from that of monetary shocks for at least two reasons. First, since a fiscal shock directly affects households' budgets, its effect directly depends on other sources of income and their endogenous dynamic responses over time. Assumptions about the distribution of factor incomes thus have a first-order effect on the propagation of fiscal shocks.

Second, it is well known that the effect of fiscal shocks depends on the response of real interest rates. Moreover, Broer et al. (2023) note that accounting for wage rigidity dampens the inflation response to fiscal shocks and, thus, the endogenous reaction of monetary policy that typically counteracts the demand effect of fiscal shocks. This raises the fiscal multiplier relative to the standard version of the model with only price rigidities but also makes it less sensitive to the current stance of monetary policy. A recent paper supports this view: Auclert et al. (2023) show that it is impossible for NK models with flexible labor markets to simultaneously match empirical estimates for marginal propensities to earn, marginal propensities to consume, and fiscal multipliers.

Kaplan and Violante (2022) show that the HANK model with liquid and illiquid assets matches the empirical MPCs much better than the one-asset HANK model. In addition, Kaplan and Violante (2014) introduce two-asset models, and Kaplan et al. (2018) and Luetticke (2021) highlight the ability of the two-asset model to match the differential portfolio response to monetary policy shocks and provide new evidence for the importance of modeling both liquid and illiquid assets. We build on that and implement two-asset HANK model with adjustment costs à la Kaplan et al. (2018). We rely on fast and accurate sequence-space

Jacobian method implementation by Auclert et al. (2021) for the solution method.

The rest of the paper is organized as follows. Section 2 connects and differentiates our paper to the most relevant literature. Section 3 presents findings on HtM shares and asset composition for a set of European countries. In Section 4, we introduce our model as well as the calibration and show model performance. Section 5 contains the quantitative analysis of the fiscal multiplier. In Section 6, we calibrate two models and compare aggregate responses for core and periphery European countries. Section 7 concludes.

3.2 Related Literature

In this paper, we use the definition of three types of HtM households (pHtM, wHtM, and non-HtM) from Kaplan et al. (2014). Kaplan et al. (2014) and Slacalek et al. (2020) use HFCS and show heterogeneity in HtM status for four large countries (France, Germany, Italy, and Spain). Using the updated HFCS, we complement their analysis by estimating shares for all available countries. Moreover, we document heterogeneity between countries in liquid and illiquid asset holdings. Using the HFCS, Carroll et al. (2014) show heterogeneity in liquid assets and wealth across countries. We use the most recent HFCS and show heterogeneity in liquid and illiquid asset holdings.

We build on a large body of literature exploring HANK models. See Kaplan and Violante (2018) for a recent overview of the literature.

The three papers that are most closely related to ours are Bayer et al. (2023), Hagedorn et al. (2019), and Auclert et al. (2018). All three papers study fiscal multipliers, and their models include a two-asset structure and rigid wages. However, Bayer et al. (2023) consider only one tax rate and simple government problem. Similarly, the government in Auclert et al. (2018) collects only progressive income taxes. The government problem in Hagedorn et al. (2019) is more elaborate and includes dividend taxes as we do.

In contrast to all three papers, we combine progressive income taxes and dividend taxes. In addition, we include distortive consumption taxes in our analysis. In addition, in our analysis, we explore how fiscal multipliers vary with different levels of debt and tax structures, whereas all three mentioned papers concentrate only on the U.S. case. Thus, none of the papers above offer answers to how the fiscal multiplier changes in the case of government spending in a highly indebted country or in the case of highly progressive taxes, which are not rare in many European countries. In addition, we analyze how the fiscal multiplier depends on the HtM share of households in the economy.

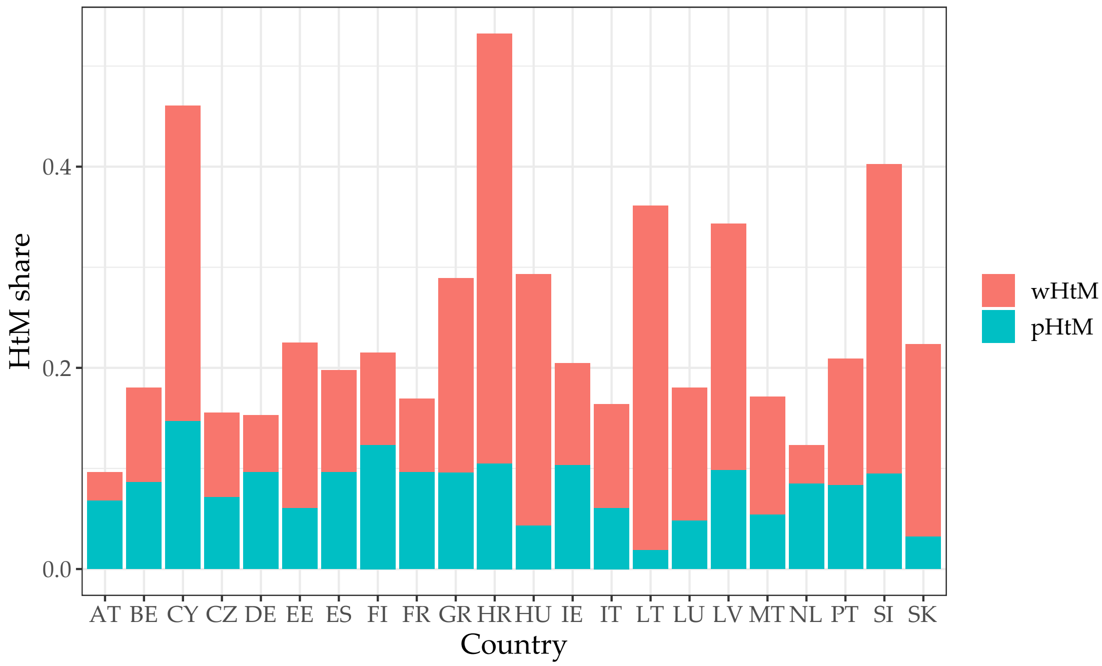


Figure 17: wHtM and pHtM shares for a set of European countries; Kaplan et al. (2014) definition. Source: Eurosystem Household Finance and Consumption Survey, wave 4.

3.3 HtM Status and Household Portfolio Comparison

This section highlights household heterogeneity by comparing households' asset holdings by asset type and HtM status across a set of European countries using the HFCS. The analysis details additional results, and variable definitions are in Appendix C.1.

Figure 17 shows HtM, pHtM, and wHtM shares across European countries. We observe large heterogeneity in shares across countries. First, HtM shares range from low shares in Austria and the Netherlands to high shares in smaller European countries such as Croatia and Slovenia.

Second, we can also observe heterogeneity in pHtM and wHtM shares across countries with similar shares of HtM households. For example, Germany and Italy have similar shares of HtM households, around 16%. However, Germany has a larger share of pHtM households (around 10%) while Italy has a larger share of wHtM households (around 10%).

Figure 18 and Figure 19 document heterogeneity in net liquid and illiquid asset holdings, respectively, across European countries. Both amounts are measured relatively to net income. Figure 18 shows heterogeneity in net liquid asset holdings relative to net income for smaller European countries. However, heterogeneity is also present for larger European countries, i.e., France (around 1.05), Germany (around 1.26), and Italy (around 1.4). The same holds

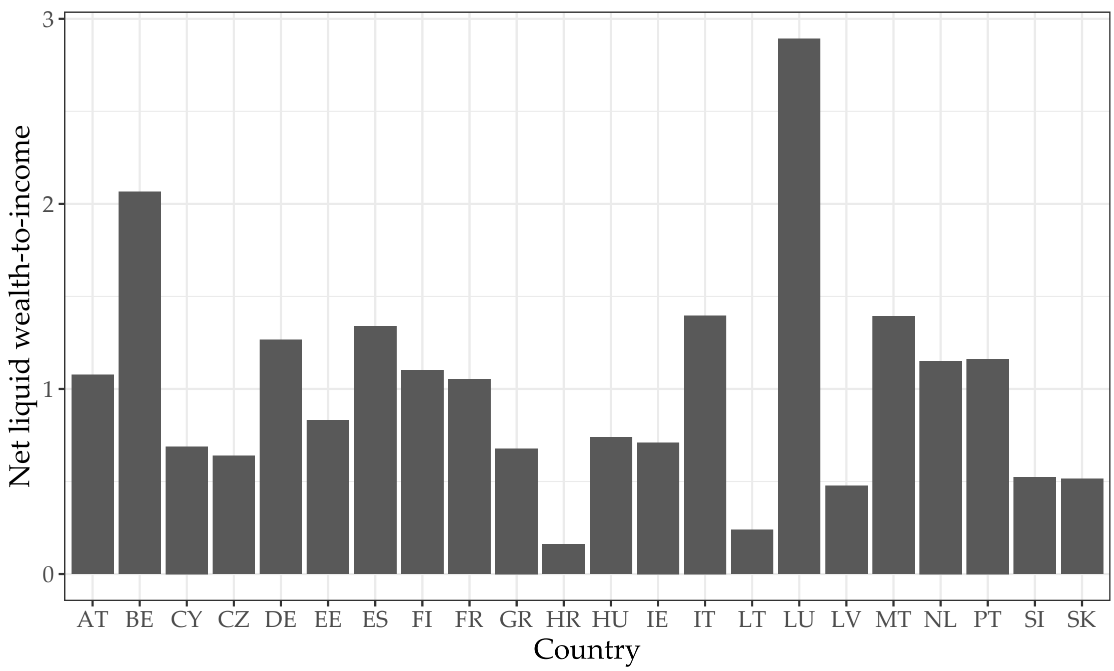


Figure 18: Net liquid wealth to income for a set of European countries. Source: Eurosystem Household Finance and Consumption Survey, wave 4.

for net illiquid asset holdings (Figure 19).

These findings further motivate our two-asset HANK model. They highlight the heterogeneity across HtM status as well as in asset holdings. With our model, we account for differential HtM status and asset types and highlight their importance for the fiscal multiplier analysis.

3.4 Quantitative HANK Model

In this section, we present our model blocks. The model consists of households who can save in two types of assets. The agents in the economy can save in liquid assets accounts that they can tap into in every period at no cost. However, returns on liquid assets are lower than returns on illiquid assets. Accumulating illiquid assets brings higher returns, but when adjusting illiquid assets, agents face monetary costs. The rest of the economy consists of separate blocks. The first block are financial intermediaries who manage agents' assets and provide agents with returns. Other blocks are more standard in the literature and consist of intermediate and final goods-producing firms of unions and labor packers who manage labor in the economy. The last is the government block, in which the government collects taxes, supplies bonds, and controls government spending and transfers.

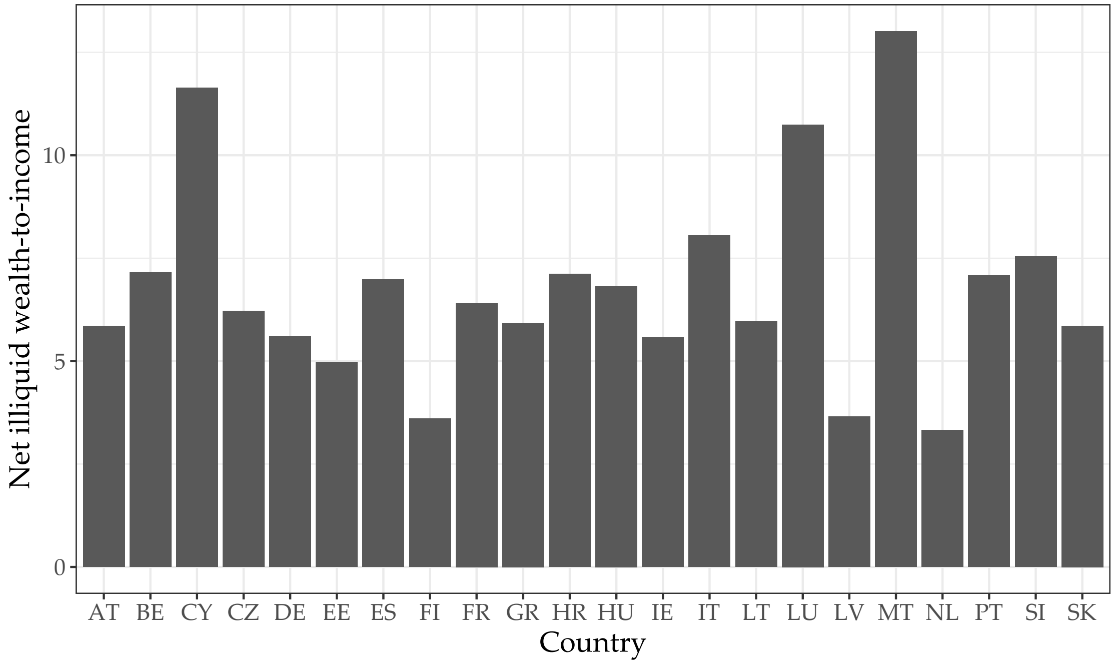


Figure 19: Net illiquid wealth to income for a set of European countries. Source: Eurosystem Household Finance and Consumption Survey, wave 4.

3.4.1 Households

The following Bellman equation characterizes the household problem:

$$\begin{aligned} V_t(e_{i,t}, b_{i,t-1}, a_{i,t-1}, \beta_{i,t}) &= \max_{c_{i,t}, b_{i,t}, a_{i,t}} \left\{ \frac{c_{i,t}^{1-\sigma}}{1-\sigma} - v(n_{i,t}) + \beta_i \mathbb{E}_t V_{t+1}(e_{i,t+1}, b_{i,t}, a_{i,t}, \beta_{i,t+1}) \right\} \\ c_{i,t}(1+\tau_t^c) + a_{i,t} + b_{i,t} &= z_{i,t} + (1+r_t^a)a_{i,t-1} + (1+r_t^b)b_{i,t-1} - \Psi(a_{i,t}, a_{i,t-1}) \qquad (9) \\ a_{i,t} \geq \underline{a}, &\quad b_{i,t} \geq \underline{b}, \end{aligned}$$

where $z_{i,t} = \tau_t(w_t n_{i,t} e_{i,t})^{1-\theta} + T_{i,t}$ is after-tax labor income and $\beta_i \in \{\beta_1, \beta_2\}$ such that $\beta_1 < \beta_2$. Households choose consumption, illiquid, and liquid assets and face adjustment costs for managing illiquid assets à la Kaplan et al. (2018). The adjustment cost function is specified as

$$\Psi(a, a_-) = \frac{\chi_1}{\chi_2} \left| \frac{a - (1+r_t^a)a_-}{(1+r_t^a)a_- + \chi_0} \right|^{\chi_2} \left[(1+r_t^a)a_- + \chi_0\right],$$

with $\chi_0, \chi_1 > 0$ and $\chi_2 > 1$, and $v(n_t) = \gamma n_t^{1+\frac{1}{\phi}}$.

3.4.2 Financial Intermediary Problem

A representative risk-neutral financial intermediary takes liquid and illiquid deposits from households and invests them in government bonds B_t^g and firm equity p_t. The financial intermediary's objective is to maximize the expected real rate of return r_{t+1}. It performs liquidity transformation at proportional cost $\omega \int b_{i,t} di$. No arbitrage requires that the ex-ante return $\mathbb{E}_t = [1 + r_{t+1}]$ equals the expected returns on nominal government bonds and on equity. The competitive financial intermediary passes these returns on to households subject to intermediation costs:

$$\mathbb{E}_t[1 + r_{t+1}] = \frac{1 + i_t}{\mathbb{E}_t[1 + \pi_{t+1}]} = \frac{\mathbb{E}_t[d_{t+1} + p_{t+1}]}{p_t} = \mathbb{E}_t[1 + r_{t+1}^a] = \mathbb{E}_t[1 + r_{t+1}^b] + \omega,$$

where $d_t = \widetilde{d}_t(1 - \tau_t^k)$, are after tax dividends. The ex-post returns r_t, r_t^a, r_t^b however, are subject to surprise inflation and capital gains. Assuming that capital gains accrue to the illiquid account, we have

$$1 + r_t = \frac{1 + i_{t-1}}{1 + \pi_t} = 1 + r_t^b + \omega$$

and

$$1 + r_t^a = \Theta_p \left(\frac{d_t + p_t}{p_{t-1}} \right) + (1 - \Theta_p)(1 + r_t)$$

where Θ_p denotes the share of equity in the illiquid portfolio.

3.4.3 Wage Setting

The labor sector in our model consists of multiple levels. On the first level, the labor sector is composed of unions that differentiate raw labor and labor packers who buy differentiated labor and then sell labor services to intermediate goods producers.

At any time t, union k sets its wage W_{kt} to maximize, on behalf of all the workers it employs, utility facing Rotemberg (1982) adjustment costs,

$$J_t^U = \max_{W_{k,t}} \int \Big(u(c_{i,t+t'}) - v(n_{i,t+t'}) \Big) d\Psi_{i,t+t'} - \frac{\psi}{2} \left(\frac{W_{k,t+t'}}{W_{k,t+t'-1}} - 1 \right)^2 + \frac{1}{1 + r_t} J_{t+1}^U$$

taking as given the initial distribution of households over idiosyncratic states $\Psi_{i,t}$ as well as the demand curve for tasks coming from the labor packers, which is

$$N_{k,t} = \left(\frac{W_{k,t}}{W_t} \right)^{-\varepsilon} N_t,$$

where

$$W_t = \left(\int W_{k,t}^{1-\varepsilon} dk \right)^{\frac{1}{1-\varepsilon}}$$

is the price index for aggregate employment services. Solving the unions' problem implies wage NKPC

$$(1+\pi_t^w)\pi_t^w = \kappa^w \left(\gamma N_t^{1+\frac{1}{\phi}} - \frac{(1-\theta)}{(1+\tau_t^c)\mu^w} Z_t u'(\tilde{C}_t) \right) + \beta(1+\pi_{t+1}^w)\pi_{t+1}^w.$$

where $\kappa^w = \frac{\varepsilon}{\psi}$, $\mu^w = \frac{\varepsilon}{\varepsilon-1}$, $u'(\tilde{C}_t) = \int \frac{e_{i,t}^{1-\theta}}{\int e_{i,t}^{1-\theta} di} u'(c_{i,t}) di$, and Z_t is aggregate income tax (net of transfers).

3.4.4 Firms

In our quantitative model, the firm's sector also consists of multiple levels, i.e., intermediate and final good producers. First, intermediate goods producers hire labor services from labor packers and rent out capital to produce goods. Second, final goods producers aggregate intermediate goods with a constant elasticity of substitution $\frac{\mu_p}{\mu_p - 1} > 1$.

The equations for the model with investment are as follows. The production function of each firm is Cobb-Douglas, $F(k_{t-1}, n_t) = \Omega_t k_{t-1}^{\alpha} n^{1-\alpha}$. Each firm pays out wages, invests in capital that depreciates while facing capital adjustment costs $\phi\left(\frac{k_t}{k_{t-1}}\right) = \frac{1}{2\delta\varepsilon_I}\left(\frac{k_t}{k_{t-1}} - 1\right)^2$, and sets prices facing Rotemberg (1982) adjustment cost function $\xi(\mathcal{P}_t, \mathcal{P}_{t-1}) = \frac{1}{2\kappa^p(\mu^p-1)}\left(\frac{\mathcal{P}_t - \mathcal{P}_{t-1}}{\mathcal{P}_{t-1}}\right)^2$. The Bellman equation for the intermediate good's firm is:

$$\begin{aligned} J_t(\mathcal{P}_{t-1}, k_{t-1}) = \max_{\mathcal{P}_t, k_t, n_t} \Bigg\{ & \frac{\mathcal{P}_t}{P_t} F(k_{t-1}, n_t) - \frac{W_t}{P_t} n_t - i_t - \phi\left(\frac{k_t}{k_{t-1}}\right) k_{t-1} - \xi(\mathcal{P}_t, \mathcal{P}_{t-1}) Y_t + \\ & \frac{1}{1+r_t} J_{t+1}(\mathcal{P}_t, k_t) \Bigg\}, \\ \text{subject to} \quad F(k_{t-1}, n_t) = & \left(\frac{\mathcal{P}_t}{P_t}\right)^{-\frac{\mu^p}{\mu^p - 1}} Y_t, \end{aligned}$$

where $i_t = k_t - (1-\delta)k_{t-1}$.

All intermediate goods firms are identical in the equilibrium and thus make the same choices, i.e., $k_t = K_t$, $n_t = N_t$, and $\mathcal{P}_t = P_t$. Resulting Phillips curve for inflation is given with

$$(1+\pi_t)\pi_t = \kappa_p (\mu_p \cdot mc_t - 1) + \frac{1}{1+r_t} \frac{Y_{t+1}}{Y_t} (1+\pi_{t+1})\pi_{t+1},$$

where $mc_t = \frac{W_t/P_t}{F_{n,t}}$ are marginal costs. Moreover, Tobin's Q and capital satisfy the following two equations

$$Q_t = 1 + \frac{1}{\delta\epsilon_I}\left(\frac{K_t - K_{t-1}}{K_{t-1}}\right)$$

and

$$\begin{aligned}(1+r_t)Q_t = \alpha\Omega_{t+1}\left(\frac{N_{t+1}}{K_t}\right)^{1-\alpha} mc_{t+1} - \left[\frac{K_{t+1}}{K_t} - (1-\delta) + \frac{1}{2\delta\epsilon_I}\left(\frac{K_{t+1}-K_t}{K_t}\right)^2\right] \\ + \frac{K_{t+1}}{K_t}Q_{t+1}.\end{aligned}$$

Lastly, dividends satisfy

$$\tilde{d}_t = F(K_{t-1}, N_t) - w_t N_t - I_t - \phi\left(\frac{K_t}{K_{t-1}}\right)K_{t-1} - \xi(\pi_t)Y_t.$$

3.4.5 Monetary and Fiscal Policy

The monetary authority now follows a Taylor rule:

$$i_t = r + \phi^\pi \pi_t.$$

In addition to labor taxes τ_t, the government collects dividend taxes τ_t^k, and (collects/pays out) lump-sum taxes/transfers T_t. Therefore, the government budget constraint is given with

$$B_t^g + \tau_t^c C_t + Z_t + \tau_t^k \tilde{d}_t = (1+r_{t-1})B_{t-1}^g + G_t + T_t.$$

The government balances its budget and follows an AR(1)-type spending policy, $dG_t = \rho^G dG_0$. We take the following approaches to analyze the impact of different sources of financing government spending. First, to analyze the effect of financing government spending with the deficit, we assume that taxes are chosen such that the path of public debt is given by $dB_t = \rho^B(dB_{t-1} + dG_t)$. Running deficit in the following exercise means that government spending B_t^g solves the following equation every period

$$\begin{aligned}B_t^g &= (1-\rho^B)B_{ss}^g + \rho^B(B_{t-1}^g + G_t - G_{ss}) \quad \Leftrightarrow \\ B_t^g - B_{t-1}^g &= (1-\rho^B)(B_{ss}^g - B_{t-1}^g) + \rho^B(G_t - G_{ss}) \quad \Leftrightarrow \\ B_t^g - B_{ss}^g &= \rho^B(B_{t-1}^g - B_{ss}^g) + \rho^B(G_t - G_{ss}) \quad \Leftrightarrow \\ dB_t &= \rho^B(dB_{t-1} + dG_t),\end{aligned}$$

where the last equation refers to how the debt policy is addressed in the text above. We compare cases when different taxes are chosen for each period to satisfy government budget constraint. To assess the effect of financing the government spending with the deficit, for each tax, we compare the impact to the impact when the taxes are increased to satisfy the government budget each period without an increase in the debt level.

3.4.6 Calibration

This section describes the choice of model parameters and parameters that target moments from the data. Table 11 presents externally set parameters and sources from the literature. Most of the choices for these parameters are standard in the literature, such as inverse intertemporal elasticity of substitution (IES) and inverse Frisch elasticity, which are set to value 2. Moreover, the second block of Table 11 presents external calibration of tax-related parameters in the government block. These values are specific to the U.S. economy. For example, we use the value for the progressivity parameter θ estimated by Heathcote et al. (2017).

In contrast, Table 12 presents calibrated parameters. The last two columns present targets and resulting values in the steady state. Again, these parameters are specified for the U.S. economy.

Parameter	Description	Value	Source
σ	Inverse IES	2	Auclert et al. (2023)
ξ_0	Portfolio adj. cost pivot	0.25	Auclert et al. (2021)
ξ_2	Portfolio adj. cost curvature	2	Auclert et al. (2021)
ρ_e	Autocorrelation of earnings	0.966	Floden and Lindé (2001)
σ_e	Cross-sectional s.d. of log earnings	0.92	Auclert and Rognlie (2018)
ϕ	Inverse Frisch elasticity	2	Chetty et al. (2011)
ϕ^π	Taylor rule coefficient	1.5	standard value
κ^w	Slope of wage Phillips curve	0.03	Hagedorn et al. (2019)
κ^p	Slope of price Phillips curve	0.03	Christiano et al. (2011)
ε_I	Investment elasticity to Q	1	Auclert et al. (2018)
ρ^G	Spending persistence	0.7	
τ	Income tax level	0.325	
θ	Income tax progressivity	0.181	Heathcote et al. (2017)
τ^k	Dividend tax level	0.36	Trabandt and Uhlig (2011)

Table 11: Externally set parameters.

Parameter	Description	Value	Target
β_1	Discount factor	0.950	pHtM share
β_2	Discount factor	0.982	asset mkt. clearing
ξ_1	Portfolio adj. cost scale	14.664	Liquidity B/Y=26%
γ	Disutility of labor	1.219	N=1
μ^w	Steady state wage markup	1.1	10% markup
μ^p	Steady-state markup	1.070	A/Y=292%
Ω	TFP	0.436	Y=1
α	Capital share	0.357	K/Y=2.565 yearly
δ	Depreciation	0.02	8% yearly depreciation
r	Real interest rate	0.0125	5% yearly return
ω	Liquidity premium	0.005	2% yearly spread
G	Government spending	0.2	G/Y=20%
B^g	Bond supply	2.8	Debt-to-GDP 70%
B	Liquid assets	1.04	B/Y=26%
T	Transfers	0.062	govt. budget
τ^c	Consumption tax level	0.08	8% VAT

Table 12: Calibrated parameters.

3.4.7 Model Performance

Quarterly MPC is 0.17, that is 0.52 in annual terms and the range of the annual empirical estimates from the literature (e.g., Johnson et al. (2006); Jappelli and Pistaferri (2014); Carroll et al. (2017)). Table 13 presents non-targeted moments. Results show that our model performs well in matching HtM and wHtM shares without targeting those. Moreover, it performs well in matching Gini coefficients for both types of assets and shares of assets in the bottom 50% of the distributions. However, the model performs less well in matching shares in assets in the top 50% of distributions. Namely, the model understates the share of assets of the top 10% and overstates the share of the next 40% in respective distributions. It is worth noting that we do not allow agents to borrow and, thus, potentially restrict model performance in matching untargeted moments for liquid wealth.

3.5 Fiscal Multipliers

This section compares the heterogeneous agents model with the representative and two-agent models and further motivates the use of HANK models in analyzing fiscal multipliers. Consequently, we explore the fiscal multiplier's dependence on tax structures, government debt level, and household heterogeneity. We use two measures for the fiscal multiplier. The first measure, impact multiplier, is defined with $\frac{dY_0}{dG_0}$, whereas the second measure, cumulative

	Moment	Model	Data	Source
Liquid Assets	top 10% share	65.72	86	Kaplan et al. (2018)
	next 40% share	35.14	18	
	bottom 50% share	0.14	-4	
	Gini coefficient	0.81	0.98	
Illiquid Assets	top 10% share	49.11	70	Kaplan et al. (2018)
	next 40% share	50.74	27	
	bottom 50% share	0.15	3	
	Gini coefficient	0.84	0.81	
HtM	HtM	41.8%	41%	Kaplan and Violante (2022)
	$wHtM$	27.7%	27%	
	$pHtM^*$	14%	14%	

Table 13: Non-targeted moments: model outcomes compared to data counterparts. Note: * denotes the targeted moment used in the calibration.

fiscal multiplier, is defined as $\frac{\sum_t (1+r)^{-t} dY_t}{\sum_t (1+r)^{-t} dG_t}$.

3.5.1 HANK-TANK-RANK Comparison

Consistent with results in the literature (e.g., (Auclert et al., 2018; Hagedorn et al., 2019; Bayer et al., 2023)), RANK and TANK cannot produce responses to government spending or fiscal multipliers as seen in the data. Figure 20 compares RANK and TANK fiscal multipliers and impulse responses to ones resulting from our HANK model, additionally motivating the use of the HANK model in further analysis. The fiscal multiplier corresponding to the RANK model is below 1, resulting from strong crowding out of investments and consumption. Consumption response in the TANK model is positive and stronger than in the RANK model, which results in an impact fiscal multiplier larger than 1. However, consumption drops quickly, resulting in a cumulative fiscal multiplier smaller than one. Conversely, consumption response in our HANK model is larger and declines slowly to the steady state value, resulting in an impact and cumulative fiscal multiplier larger than 1.

3.5.2 Consumption Decomposition

As Figure 20 highlights, the HANK model is the only one able to produce fiscal multipliers larger than one in addition to positive consumption multipliers. To explain the consumption response more intuitively, Figure 21 presents a compact decomposition of different effects relevant to the consumption response. Using the Jacobian structure of Auclert et al. (2021), we can decompose consumption responses into effects due to transfers, income, rate change, and investments coupled with indirect effects. We use this decomposition further to explain and

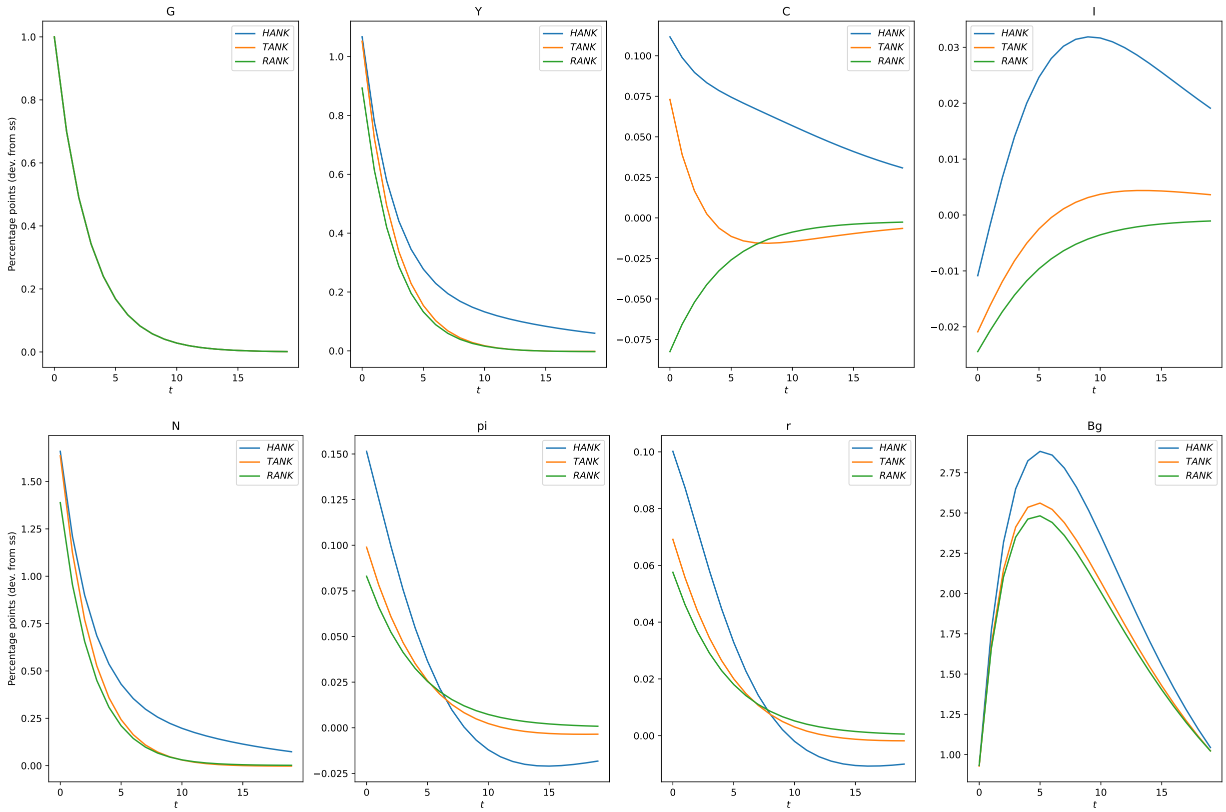


Figure 20: Impulse response functions corresponding to 1% increase in government spending financed with government debt and transfers across representative, two-agents, and heterogeneous agents models.

highlight two dimensions of heterogeneity important for differential consumption responses.

First, Figure 22 decomposes the aggregate consumption response on the response of households' consumption from the bottom 50% and top 10% parts of the wealth distribution. From the bottom part of the wealth distribution, poorer households drive an initial positive consumption response due to an increase in labor income and government transfers. However, the effect is then crowded out. In contrast, wealthy households respond with a small decrease in consumption and turn to investments. Consequently, due to higher capital gains, they drive positive responses in aggregate consumption.

The second important dimension is to see how HtM status affects households' consumption decisions. Figure 23 decomposes consumption on the average consumption response of HtM households and non-HtM households. The figure shows that HtM households, after the initial increase in consumption, based on the effect of income and transfers, turn to savings/investments and reduce consumption. In contrast, non-HtM households are not affected that much by constrained assets and increase their consumption throughout the

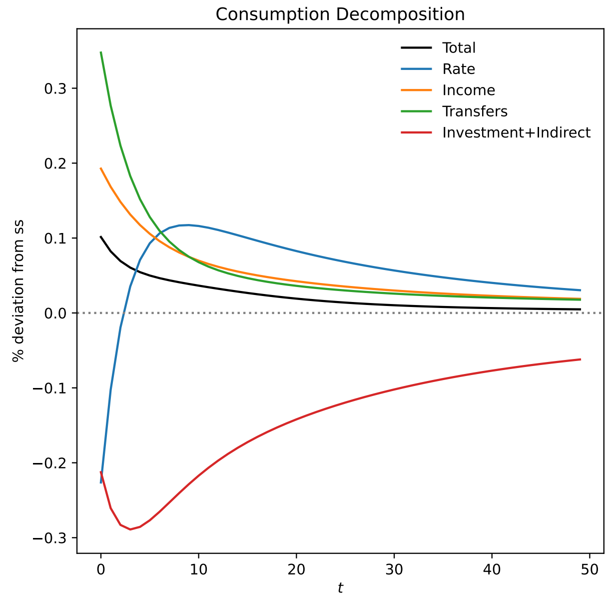


Figure 21: Aggregate consumption responses decomposition.

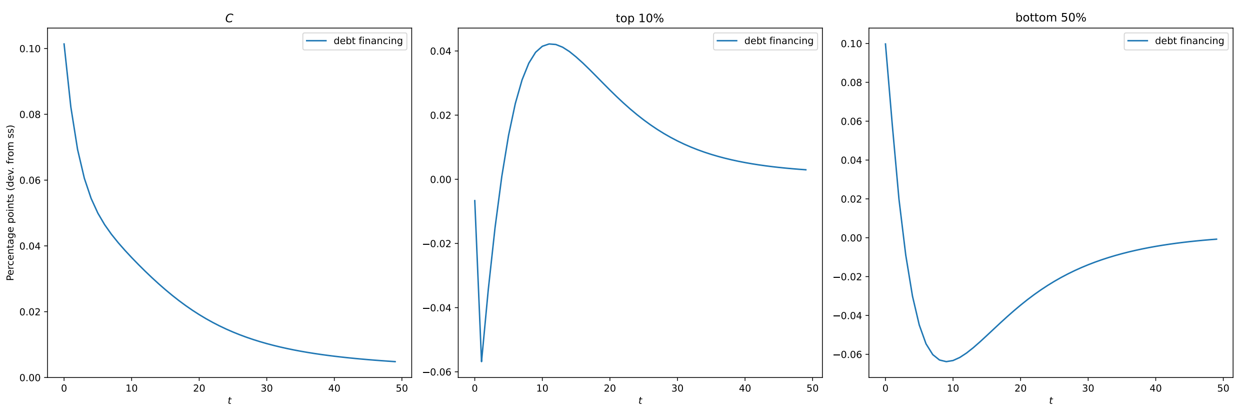


Figure 22: Consumption decomposition based on households' wealth.

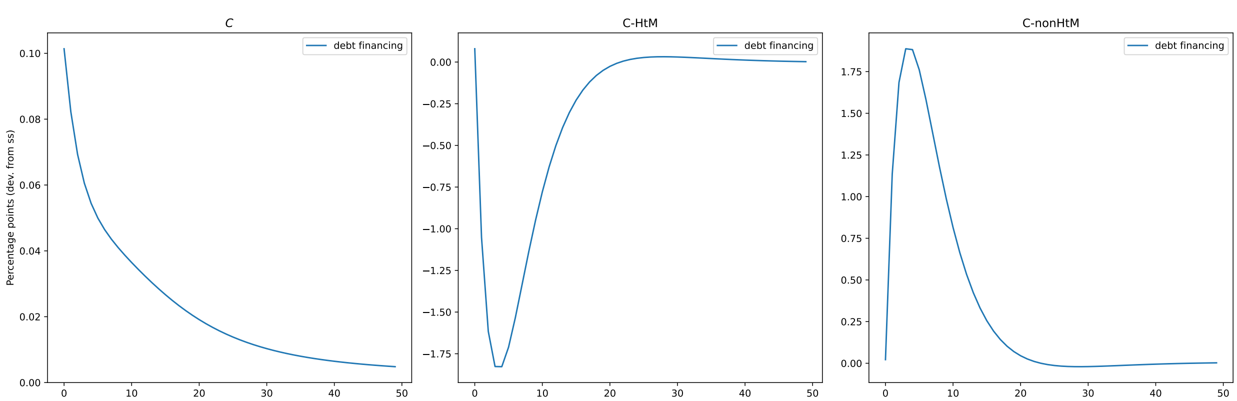


Figure 23: Consumption decomposition based on households' HtM status.

period of government spending.

3.5.3 Sources of Financing of Government Spending: Debt vs. Direct Financing

To compare fiscal multipliers dependent on the source of financing, we compare the case when government spending is financed directly from taxes or transfers to when spending is financed from an increase in government debt. Therefore, we use the path of government debt given with the AR(1) process specified above. When spending is financed with an increase in debt, we compare cases when the residual in government spending is financed by raising taxes or reducing government transfers.

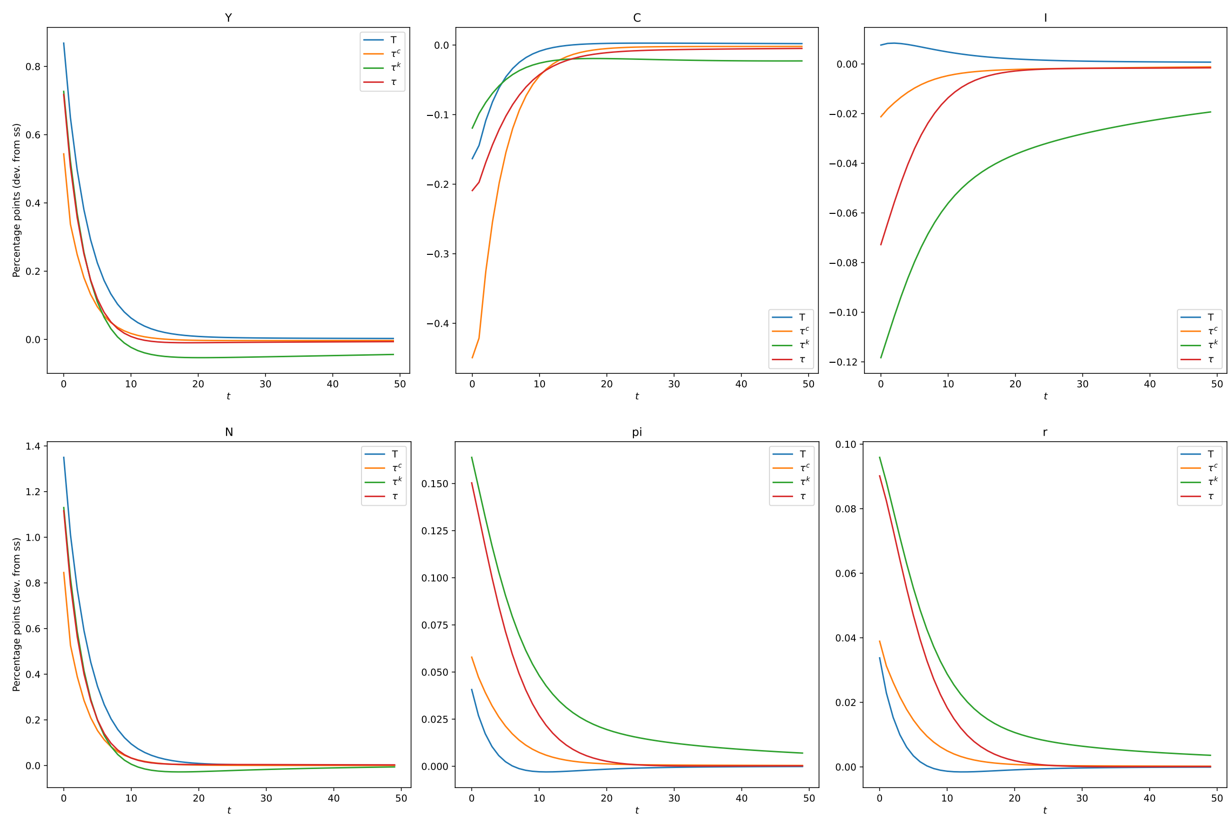


Figure 24: Impulse response functions corresponding to a 1% increase in government spending financed directly from government transfers or from consumption, dividend, and income taxes.

Figure 24 presents responses in four cases when government spending is directly financed with transfers, consumption taxes, dividend taxes, and changes in the income tax level. In all four cases, consumption is completely crowded out with investments, and the consumption response is negative. Moreover, in all cases, both impact and cumulative multipliers are less than one. On the one hand, financing with government lump sum transfers produces both the highest impact (0.87) and cumulative fiscal multiplier (0.92). On the other hand, financing government spending with an increase in consumption taxes produces the lowest impact multiplier (0.54). Lastly, financing spending with dividend taxes produces a negligible negative multiplier.

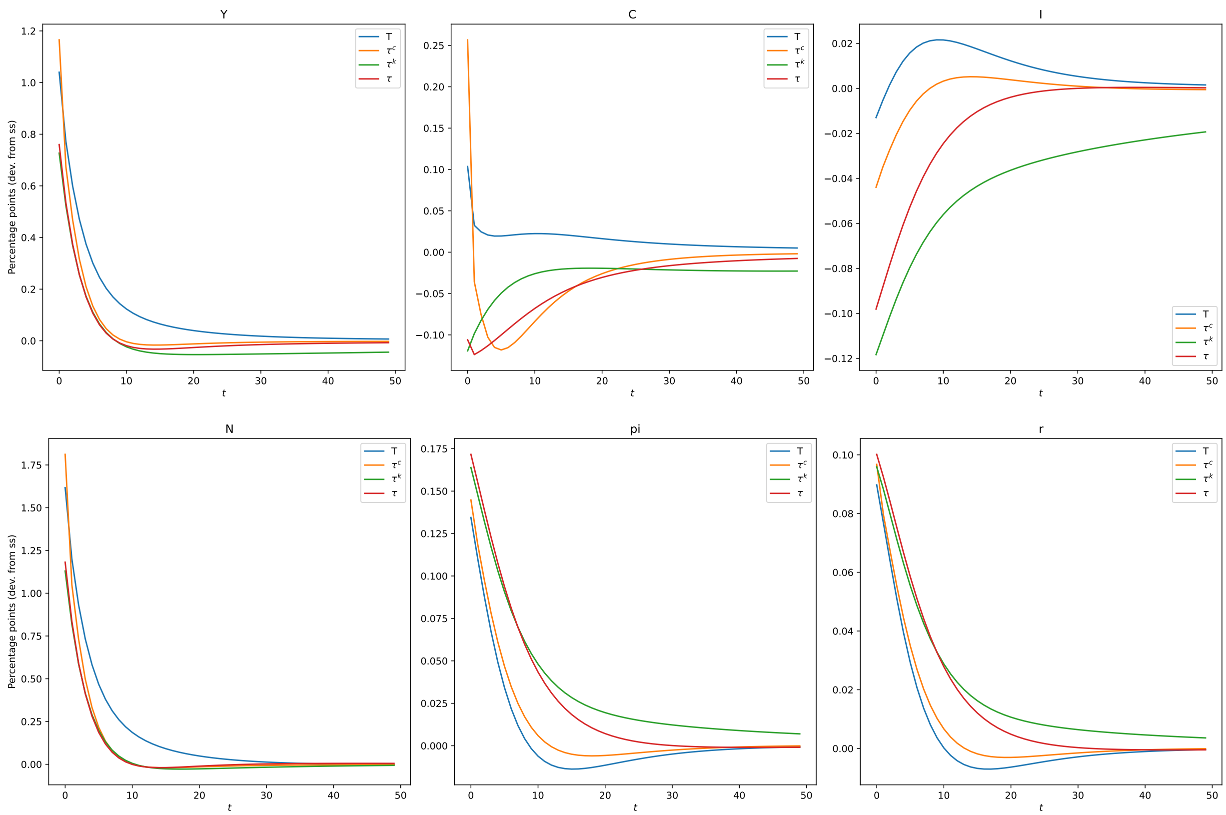


Figure 25: Impulse response functions corresponding to a 1% increase in government spending financed with debt and from government transfers or from consumption, dividend, and income taxes.

In the case of government spending being financed by raising debt (Figure 25), both impact and cumulative fiscal multipliers do not change much when using consumption and income taxes. In contrast, using consumption taxes produces the largest impact fiscal multiplier (1.17) and positive initial consumption response. However, consumption is crowded out with investments, and the output response drops, resulting in a cumulative multiplier of around 0.7. Finally, consumption is not crowded out when the government uses transfers, which results in both impact and cumulative fiscal multipliers higher than one. Table 14 summarizes the impact and cumulative fiscal multipliers for all cases.

	Tax instrument	Impact multiplier	Cumulative multiplier
Direct financing	T	0.87	0.92
	τ^c	0.54	0.39
	τ^k	0.73	-0.06
	τ	0.72	0.48
Debt financing	T	1.04	1.32
	τ^c	1.17	0.70
	τ^k	0.73	-0.06
	τ	0.76	0.40

Table 14: Cumulative and impact fiscal multipliers depending on the source of financing of government spending.

3.5.4 Debt Level and Income Tax Progressivity

In this section, we compare the effectiveness of the fiscal stimulus in low debt and in the economy with not progressive income taxes. To be specific, we calibrate the model to limiting cases, i.e., *low* B^g in which debt-to-GDP is set to be 30% and the *low* θ in which tax progressivity is close to zero. Lastly, we calibrate the model with both low progressivity and low debt-to-GDP and compare it to our baseline model.

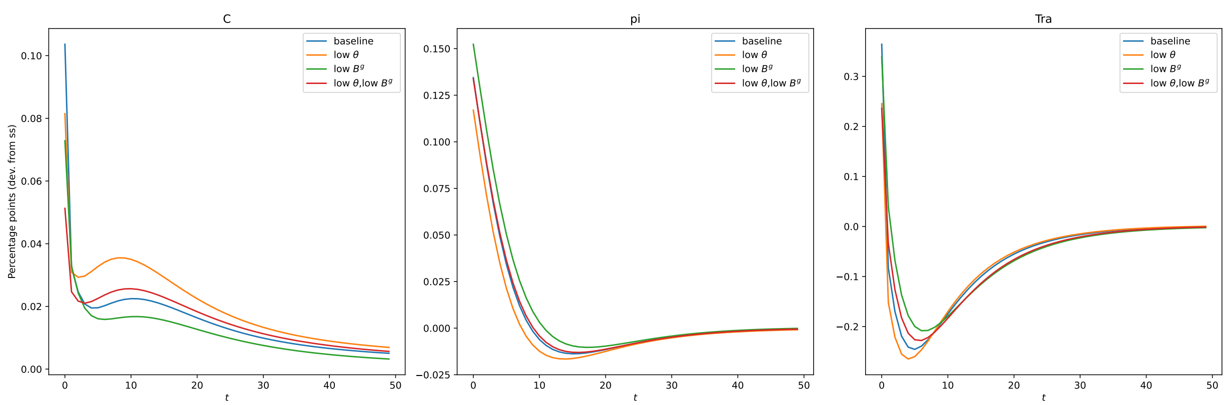


Figure 26: Impulse response functions corresponding to a 1% increase in government spending financed with debt and from government transfers for low debt-to-GDP and income tax progressivity environments.

The impact fiscal multiplier does not change substantially, up to 5 b.p. However, the cumulative fiscal multiplier varies a lot more, in the range of 22 b.p., i.e., different tax structures and debt-to-GDP levels have long-term implications for the effectiveness of the fiscal stimulus. Figure 26 presents impulse responses corresponding to a 1% increase in government spending financed with debt and debt from government transfers. Specifically, *low* B^g equilibrium has a lower impact on consumption, higher inflation, and 12 b.p. lower cumulative multiplier of 1.20 (0.99 impact) than in the baseline economy due to the lower supply of bonds in the economy.

Conversely, the low-progressivity economy has a lower impact but a higher cumulative effect on aggregate consumption than spending in the baseline economy. Moreover, an increase in spending implies a 10 b.p. higher cumulative multiplier than in the baseline economy of 1.42 (impact 1.05). With low progressive income taxes, for constrained households of low wealth and low income, the effect is similar. However, for wealthy households, this increases liquid assets and allows them to consume more, which in turn increases aggregate consumption response.

Lastly, we calculate responses in the economy with low debt-to-GDP and low progressivity of income taxes. As noted above, the effectiveness of the fiscal stimulus is lower in an economy with a lower debt-to-GDP economy. However, imposing less progressive taxes can circumvent the lower effectiveness of the fiscal stimulus. Specifically, the economy with both low debt and low progressivity of income taxes has a cumulative fiscal multiplier of 1.32 and 1.01 impact, close to ones of the baseline economy.

3.6 Conclusion

In this paper, we develop a heterogeneous-agents model with liquid and illiquid assets to analyze the fiscal multiplier quantitatively. Implementing a rich set of fiscal policy rules, including consumption, capital, progressive income taxes, and government transfers, allows us to measure fiscal multipliers in various cases.

First, when we implement the tax structure with all taxes, we show that the RANK and TANK models cannot reproduce aggregate responses as observed in the data. This finding is similar to one already noted in literature (Auclert et al., 2018; Hagedorn et al., 2019) but in an economy with less rich tax structures than ours. Second, using aggregate consumption decomposition, we highlight the role of household heterogeneity in explaining the aggregate consumption response to the increase in government spending. Third, using our HANK model calibrated to the U.S. economy, we compare fiscal multipliers depending on the source of financing. We show that financing government spending with debt and repayment with lump-sum transfers yields the highest long-term effects on output. Moreover, lump-sum transfers circumvent individual frictions in liquidity transformation and increase demand among liquidity-constrained households. Lastly, we show that less indebted economies face lower effectiveness of fiscal policy that can be circumvented by imposing less progressive income taxes

A Limited Consideration in the Investment Fund Choice

A.1 Investment Fund Market Participation

A.1.1 Two-Step Heckman Model and the LATE estimator

In the Two-Step Heckman Model, the outcome is the value of investing representative of household preferences, thus unobserved by the econometrician. The model outcome of household i is the latent variable

$$V_{ik}^* = W_{ik}'\alpha + \eta_{ik}, \tag{10}$$

where k separates between investing and not investing. W_{ik} defines the vector of household observables and the error term η_{ij} contains characteristics that are unobservable in the data. For households who decide to invest, V_{ik} corresponds to the investment size. Taking *log* of the investment size allows interpreting marginal effects in percentage points.

Ultimately, the household i selects to invest if the value of investing is higher than the value of non-investing. The second part of the Heckman Model is the selection equation

$$Y_{ik}^* = X_{ik}'\beta + \varepsilon_{ik}, \tag{11}$$

where X defines households characteristics that correlate with latent variable Y_{ik}^* that affects the model outcome

$$INV_{ik} = \mathbb{1}_{\{Y_{ik}^* > 0\}}. \tag{12}$$

All together, the selection equation (11) and the outcome equation (10) add to model specification

$$INV_{ik} = \mathbb{1}_{\{Y_{ik}^* \geq 0\}} = \begin{cases} 1, & \text{if } Y_{ik}^* \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad V_{ik} = \begin{cases} V_{ik}^*, & \text{if } Y_{ik}^* \geq 0 \\ 0, & \text{otherwise} \end{cases}.$$

Finally, the joint error distribution is assumed to be normal

$$\begin{pmatrix} \varepsilon_{ik} \\ \eta_{ik} \end{pmatrix} \sim \mathcal{N} \begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix}, \tag{13}$$

which corresponds to probit specification. I normalize the scale and set the variance of the error ε to 1.

Table 15 in the Appendix A.1.4 contains estimation results. Inverse Mills Ratio is significant in Table 15, and correlation ρ is negative. Thus, I need to account for the bias in the outcome equation. Adjusted marginal effects are in Tables 16 and 17. As an additional robustness check, in Appendix A.1.5, I include income in the regression. Table 18 shows that other estimates do not change and that income is insignificant in some instances. Thus, the paper analysis does not contain income as an explanatory variable.

Kline and Walters (2019) show that, under certain conditions, the Heckman Two-Step Model estimator is equivalent to the LATE estimator and, therefore, does not suffer from sensitivity critique. I check whether my model specification and the SCF data satisfy conditions in Kline and Walters (2019) and obtain the equivalence of the two estimators. For this reason, the estimates in this section are robust to the sensitivity critique of the Heckman (1979) estimator.

A.1.2 Investment Fund Participation-Who Participates?

The first column in Table 16 in the Appendix A.1.4 informs about marginal effects for the selection equation, calculated in percentage points. I discuss my results using visual representation in Figures 27, 28a, 28b, 29a, and 29b, and compare my findings with other studies that use investor microdata. I focus on the extensive margin (deciding to invest) and discuss sample subgroups as potential targets for policies relevant to investment fund participation.

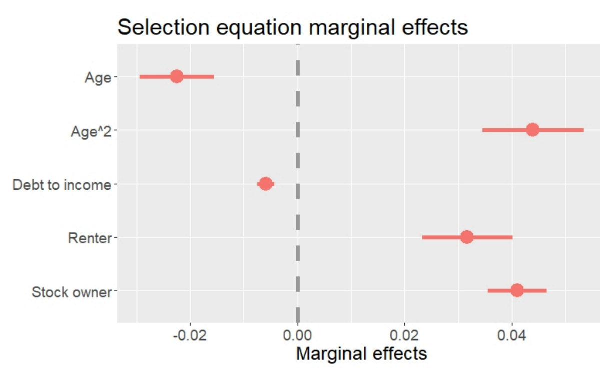
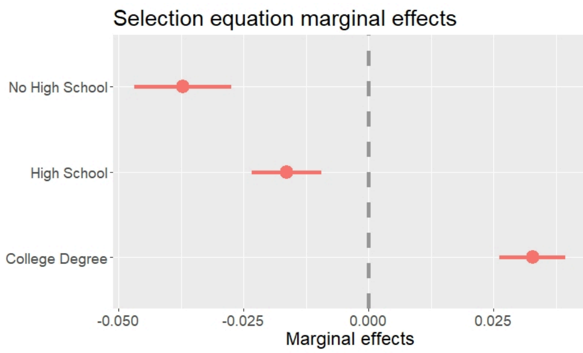
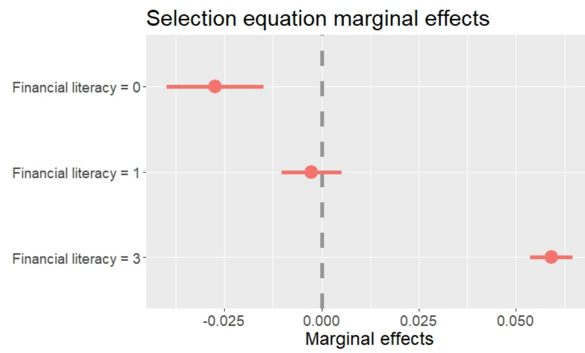


Figure 27: Selection equation marginal effects for age, debt to income ratio, homeownership and stockownership status. Marginal effects are reported with 95% confidence intervals.

Figure 27 shows that older households are less likely to participate, in line with average age differences between asset market participants and non-participants (Calvet et al., 2007). Interestingly, renters are more likely to buy a share in the investment fund. Combining these two facts adheres to the life-cycle narrative: asset accumulation with the purpose of house down payment (Brandsaas, 2021). Clearly, stock owners are more likely to participate in the investment fund, while debt reduces the likelihood of participation.



(a) Base category is "Some College".



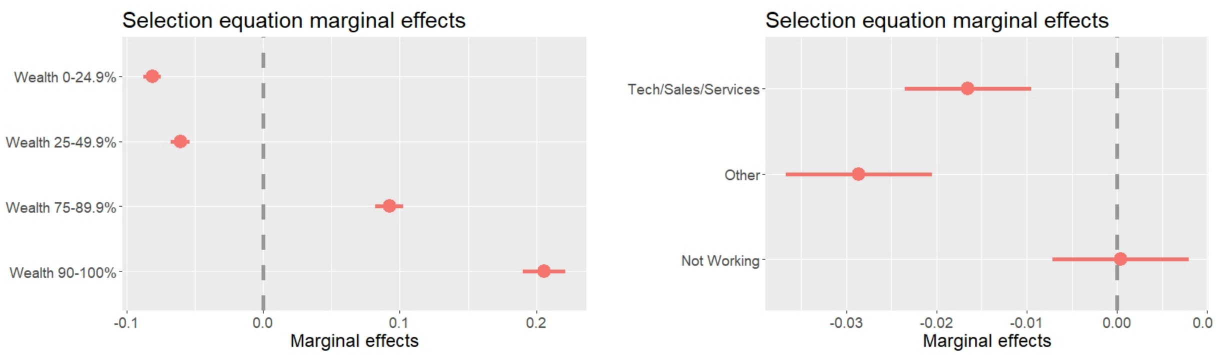
(b) Base category is "Financial literacy = 2".

Figure 28: Selection equation marginal effects for education and financial literacy. Marginal effects are reported with 95% confidence intervals.

Figures 28a, and 28b present marginal effects of education and financial literacy on the likelihood of the investment. Households with no high school relative to households with some college are 4% less likely to invest, while households with a college degree are 3% more likely to invest in investment funds. While similar studies Calvet et al. (2009b) and Van Rooij et al. (2011) resort to defining a measure of financial skill, I discuss my findings based on the direct measure of financial literacy. Households with a high degree of financial skill are 5% more likely to participate in the fund, which underlines a limited understanding of fund options for the low level of financial skill (Nieddu and Pandolfi, 2021).

While education and wealth effects align with direct stock market participation (Calvet et al., 2007), model estimates inform about the use of financial skill in trusting the fund management. These results are in line with Kacperczyk et al. (2019), where low levels of study-defined financial skill imply shifting from intermediated products to standard liquid assets.

Figure 29a shows that higher wealth implies a higher likelihood of investment, with a magnitude of almost four times as large as other household characteristics. In comparison to the middle wealth quantile, the top wealth quantile is 20% more likely to participate in investment funds. Correspondingly, households in managerial and professional occupations are more likely to invest. These results are in line with stock market participation (Campbell, 2006; Calvet et al., 2007, 2009a,b; Calvet and Sodini, 2014), and speak to persistent wealth inequality through fund participation channel.



(a) Base category is "Wealth 50-74.9%". (b) Base category is "Managerial/Professional".

Figure 29: Selection equation marginal effects for wealth and occupation. Marginal effects are reported with 95% confidence intervals.

A.1.3 Investment Fund Participation-How Much do Investors Allocate?

The inverse Mills Ratio is significant, which implies the selection of the data. Thus, both estimated coefficients and marginal effects presented account for the bias.

In the rest of the section, outcome equation marginal effects estimates are reported conditional on investment fund participation, thus informing about relevant margins for the investment size. Table 17 in the Appendix A.1.4 reports all marginal effect coefficients, whereas Figures 30, 31a, 31b, and 32 provide a visual representation.

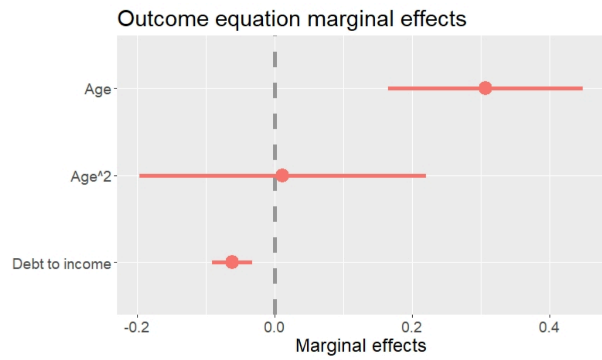
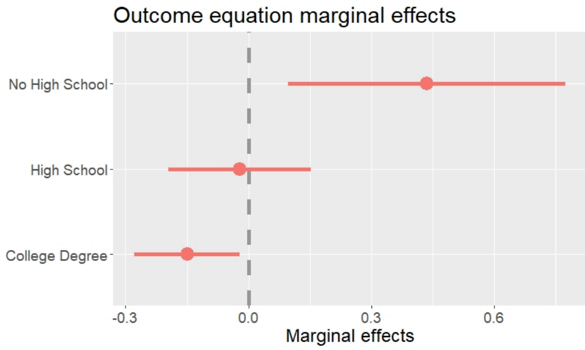


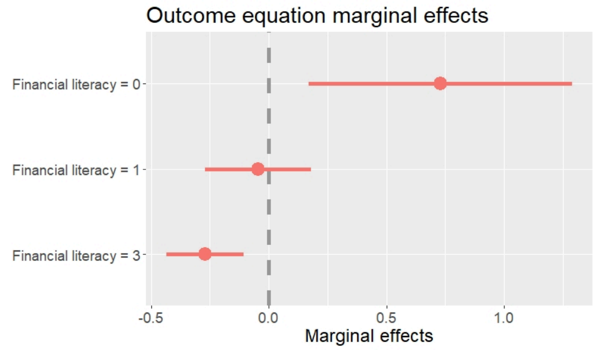
Figure 30: Outcome equation marginal effects for age, and debt to income ratio. Marginal effects are reported with 95% confidence intervals.

Even though older households are less likely to participate, older investors allocate more to funds of choice (Figure 30). On the other hand, with the increase in debt-to-income ratio, households invest less in investment funds.

Figure 31a represents the education effect and could be interpreted with student debt effects. College graduates allocate their funds to student debt repayment and, therefore, buy smaller



(a) Base category is "Some College".



(b) Base category is "Financial literacy = 2".

Figure 31: Outcome equation marginal effects for education and financial literacy. Marginal effects are reported with 95% confidence intervals.

fund shares. In contrast, high-school graduates invest approximately 40% more. Financial knowledge effects show substantial variation, suggestive of under-diversification with investors of low degree of financial sophistication, in line with Swedish microdata and study-specific measure of financial knowledge (Campbell, 2006; Calvet et al., 2007). At the same time, households with a higher level of education and financial literacy invest more in other financial and non-financial assets (i.e., liquid savings and housing), according to the breakdown in Brandsaas (2021).

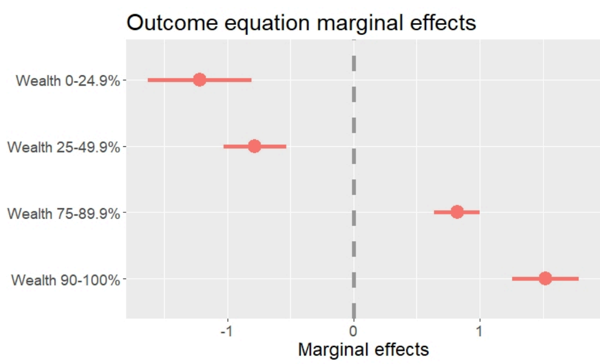


Figure 32: Outcome equation marginal effects for wealth. Base category is "Wealth 50-74.9%". Marginal effects are reported with 95% confidence intervals.

Finally, Figure 32 depicts wealth effects on the fund investment size and supports conventional wisdom in household finance. The wealth effect is substantially larger than others, separating the investment size between the top and middle wealth quantile by more than double. These results align with Calvet et al. (2007), who find that wealthier households invest more.

A.1.4 Estimation Results Tabulated

Table 15: The estimation results for the Two-Step Heckman Model estimated from the SCF.

	Selection Equation	Outcome Equation
	Dependent variable:	
	$\mathbb{1}_{\{Y_{ij}>0\}}$	$\log(invsize)$
Year	-0.089^{***}	-0.164^{***}
	(0.018)	(0.044)
Age	-0.168^{***}	0.307^{***}
	(0.027)	(0.072)
Age^2	0.328^{***}	0.012
	(0.036)	(0.107)
No High School	-0.350^{***}	0.435^{**}
	(0.054)	(0.173)
High School	-0.138^{***}	-0.021
	(0.030)	(0.088)
College Degree	0.230^{***}	-0.149^{**}
	(0.024)	(0.066)
Financial Literacy = 0	-0.314^{***}	0.730^{**}
	(0.086)	(0.285)
Financial Literacy = 1	-0.025	-0.045
	(0.038)	(0.114)
Financial Literacy = 3	0.442^{***}	-0.271^{***}
	(0.022)	(0.084)
Tech/Sales/Services	-0.124^{***}	
	(0.027)	
Other	-0.225^{***}	
	(0.034)	
Not Working	0.003	
	(0.027)	
Owns Stocks	0.306^{***}	
	(0.021)	
Renter	0.227^{***}	
	(0.030)	
Wealth $0 - 24.9\%$	-1.070^{***}	-1.219^{***}
	(0.048)	(0.210)
Wealth $25 - 49.9\%$	-0.586^{***}	-0.781^{***}
	(0.035)	(0.127)
Wealth $75 - 89.9\%$	0.470^{***}	0.821^{***}
	(0.025)	(0.093)
Wealth $90 - 100\%$	0.872^{***}	1.521^{***}
	(0.028)	(0.136)
Debt to Income Ratio	-0.044^{***}	-0.061^{***}
	(0.006)	(0.015)
Constant	-1.696^{***}	11.345^{***}
	(0.036)	(0.345)
Observations	49,377	5,125
R^2		0.338
Adjusted R^2		0.336
ρ		-0.402
Inverse Mills Ratio		-0.637^{***} (0.165)

Note: *p<0.1; **p<0.05; ***p<0.01
Base category for education is "Some College", for financial literacy is "Financial Literacy = 2," for wealth is "Wealth 50 − 74.9%", and for occupation is "Professional/Managerial".

Table 16: Marginal effects for the selection equation of the model.

	estimate	std.error	z-statistic	p-value	conf.low	conf.high
Age^2	0.04391	0.00481	9.12	0.000	0.03447	0.05334
Age	−0.02247	0.00355	−6.33	0.000	−0.02943	−0.01552
Debt to Income Ratio	−0.00583	0.00081	−7.17	0.000	−0.00742	−0.00424
No High School	−0.03710	0.00494	−7.51	0.000	−0.04677	−0.02742
High School	−0.01633	0.00356	−4.59	0.000	−0.02329	−0.00936
College Degree	0.03279	0.00334	9.81	0.000	0.02624	0.03933
Financial literacy = 0	−0.02740	0.00634	−4.32	0.000	−0.03983	−0.01497
Financial literacy = 1	−0.00261	0.00392	−0.67	0.505	−0.01029	0.00507
Financial literacy = 3	0.05906	0.00289	20.96	0.000	0.05354	0.06458
Renter	0.03169	0.00432	7.33	0.000	0.02322	0.04016
Stock owner	0.04096	0.00279	14.66	0.000	0.03548	0.04643
Wealth 0 − 24.9%	−0.08132	0.00329	−24.71	0.000	−0.08777	−0.07487
Wealth 25 − 49.9%	−0.06080	0.00359	−16.95	0.000	−0.06783	−0.05376
Wealth 75 − 89.9%	0.09234	0.00516	17.88	0.000	0.08222	0.10246
Wealth 90 − 100%	0.20576	0.00797	25.82	0.000	0.19014	0.22138
Tech/Sales/Services	−0.01650	0.00357	−4.62	0.000	−0.02351	−0.00950
Other	-0.02861	0.00413	−6.93	0.000	−0.03670	−0.02052
Not Working	0.00038	0.00385	0.10	0.92200	−0.00717	0.00792
Observations	49,377					

Note: *p<0.1; **p<0.05; ***p<0.01

Base category for education is "Some College", for financial literacy is "Financial Literacy = 2," and for wealth is "Wealth 50 − 74.9%".

Table 17: Marginal effects for the outcome equation of the model.

	estimate	std.error	z-statistic	p-value	conf.low	conf.high
Age^2	0.01183	0.10651	0.11	0.912	−0.19692	0.22058
Age	0.30661	0.07228	4.24	0.000	0.16494	0.44827
Debt to Income Ratio	−0.06116	0.01493	−4.10	0.000	−0.09041	−0.03190
No High School	0.43528	0.17261	2.52	0.012	0.09698	0.77358
High School	−0.02135	0.08847	−0.24	0.809	−0.19475	0.15205
College Degree	−0.14925	0.06560	−2.28	0.023	−0.27782	−0.02069
Financial literacy = 0	0.73032	0.28549	2.56	0.011	0.17076	1.28988
Financial literacy = 1	−0.04530	0.11437	−0.40	0.692	−0.26945	0.17886
Financial literacy = 3	−0.27094	0.08374	−3.24	0.001	−0.43508	−0.10680
Wealth 0 − 24.9%	−1.21865	0.20957	−5.81	0.000	−1.62940	−0.80789
Wealth 25 − 49.9%	−0.78148	0.12718	−6.14	0.000	−1.03076	−0.53221
Wealth 75 − 89.9%	0.82095	0.09296	8.83	0.000	0.63874	1.00315
Wealth 90 − 100%	1.52092	0.13565	11.21	0.000	1.25506	1.78678
Observations	5,125					

Note: *p<0.1; **p<0.05; ***p<0.01

Base category for education is "Some College", for financial literacy is "Financial Literacy = 2," and for wealth is "Wealth 50 − 74.9%".

A.1.5 Robustness Check - Income

Table 18: The estimation results for the Two-Step Heckman Model estimated from the SCF with income included.

	Selection Equation	Outcome Equation
	Dependent variable:	
	$\mathbb{1}_{\{Y_{ij}>0\}}$	log(*invsize*)
Year	-0.087^{***}	-0.162^{***}
	(0.018)	(0.044)
Age	-0.164^{***}	0.271^{***}
	(0.027)	(0.073)
Age^2	0.337^{***}	-0.019
	(0.037)	(0.108)
No High School	-0.331^{***}	0.424^{**}
	(0.054)	(0.172)
High School	-0.131^{***}	-0.011
	(0.031)	(0.089)
College Degree	0.221^{***}	-0.133^{**}
	(0.024)	(0.066)
Financial Literacy = 0	-0.297^{***}	0.767^{***}
	(0.087)	(0.286)
Financial Literacy = 1	-0.020	-0.018
	(0.038)	(0.115)
Financial Literacy = 3	0.440^{***}	-0.267^{***}
	(0.023)	(0.083)
Tech/Sales/Services	-0.122^{***}	
	(0.027)	
Other	-0.234^{***}	
	(0.034)	
Not Working	0.015	
	(0.028)	
Owns Stocks	0.306^{***}	
	(0.021)	
Renter	0.239^{***}	
	(0.030)	
Wealth 0 − 24.9%	-1.027^{***}	-1.241^{***}
	(0.049)	(0.208)
Wealth 25 − 49.9%	-0.569^{***}	-0.799^{***}
	(0.036)	(0.127)
Wealth 75 − 89.9%	0.463^{***}	0.852^{***}
	(0.025)	(0.093)
Wealth 90 − 100%	0.865^{***}	1.594^{***}
	(0.031)	(0.137)
Debt to Income Ratio	-0.042^{***}	-0.062^{***}
	(0.006)	(0.015)
Income 20 − 39.9%	0.046	0.145
	(0.044)	(0.130)
Income 40 − 59.9%	0.078^{*}	-0.044
	(0.042)	(0.121)
Income 60 − 79.9%	0.198^{***}	-0.051
	(0.041)	(0.115)
Income 80 − 89.9%	0.079^{*}	0.021
	(0.045)	(0.119)
Income 90 − 100%	0.127^{***}	-0.148
	(0.047)	(0.121)
Constant	-1.812^{***}	11.376^{***}
	(0.052)	(0.363)
Observations	49,377	5,125
R^2		0.339
Adjusted R^2		0.337
ρ		-0.412
Inverse Mills Ratio		-0.655^{***} (0.163)

Note: *p<0.1; **p<0.05; ***p<0.01
Base category for education is "Some College", for financial literacy is "Financial Literacy = 2," for wealth is "Wealth 50 − 74.9%", for occupation is "Professional/Managerial", and for for income is "Income 0 − 19.9%".

A.2 Investment Fund Type Choice - Estimation Results

Table 19: MLE results for the Limited Consideration Model (LCM): Investment Fund Choice

		LCM	LCM with Observables	
Average β_{1i}	8.29	[2.86, 12.3]	4.70	[0.0000, 8.51]
β_2	18.3	[16.9, 21.0]	11.2	[6.52, 11.3]
Mean of ν	0.0094	[0.0058, 0.013]	0.0058	[0.0020, 0.010]
SD of ν	0.0026	[0.0025, 0.0029]	0.0025	[0.0022, 0.0045]
Intercept	-	-	-2.57	[-2.73, -1.83]
Age	-	-	-0.027	[-0.366, 0.026]
Age^2	-	-	0.0008	[-0.0002, 0.068]
Have Stocks	-	-	0.932	[0.889, 1.77]
Debt to income	-	-	-0.209	[-0.399, 0.021]
Year	-	-	-0.212	[-0.551, -0.106]
High School	-	-	-0.202	[-0.393, 0.114]
Some College	-	-	-0.0085	[-0.052, 0.260]
College Degree	-	-	-0.928	[-1.74, -0.883]
Wealth 25 - 49.9%	-	-	0.015	[-0.013, 0.096]
Wealth 50 - 74.9%	-	-	-0.0079	[-0.087, 0.022]
Wealth 75 - 89.9%	-	-	-0.016	[-0.140, 0.043]
Wealth 90 - 100%	-	-	-0.047	[-0.185, 0.022]
Financial Literacy = 1	-	-	0.0006	[-0.030, 0.038]
Financial Literacy = 2	-	-	-0.434	[-0.806, -0.243]
Financial Literacy = 3	-	-	0.424	[0.247, 0.792]
Money Market	0.501	[0.468, 0.530]	0.501	[0.469, 0.530]
Stock Market	0.753	[0.744, 0.761]	0.753	[0.744, 0.761]
Govt Bond	0.0039	[0.0003, 0.0070]	0.0039	[0.0003, 0.0070]
Other Bond	0.0000	[0.0000, 0.0000]	0.0000	[0.0000, 0.0000]
Combined	0.0094	[0.0056, 0.013]	0.0094	[0.0057, 0.013]
Other	0.030	[0.025, 0.035]	0.030	[0.025, 0.035]
Tax Free Bond	0.029	[0.016, 0.041]	0.029	[0.016, 0.041]

Table contains MLE results and 95% bootstrapped confidence intervals (in brackets) for $B = 1000$ repetitions.

Table 20: MLE results for the Mixed Logit: Investment Fund Choice

	Mixed Logit	
Average β_{1i}	1079.8	[123.5, 1807.4]
β_2	113.2	[112.0, 127.9]
Mean of ν	0.011	[0.0092, 0.014]
SD of ν	0.0005	[0.0004, 0.0005]
Intercept	-17.4	[-33.4, -15.7]
Age	0.660	[0.587, 1.24]
Age^2	-0.0064	[-0.011, -0.0056]
Have Stocks	-1.54	[-4.27, -1.22]
Debt to income	-1.85	[-2.02, -1.76]
Year	1.49	[1.11, 2.00]
High School	-1.77	[-2.45, -1.62]
Some College	-1.55	[-2.10, -1.25]
College Degree	1.85	[1.75, 4.10]
Wealth 25 - 49.9%	-0.113	[-0.169, 0.418]
Wealth 50 - 74.9%	-0.512	[-1.17, 0.198]
Wealth 75 - 89.9%	-1.37	[-2.21, -0.953]
Wealth 90 - 100%	1.36	[0.963, 2.16]
Financial Literacy = 1	-1.28	[-1.67, -0.833]
Financial Literacy = 2	-1.43	[-2.84, -1.07]
Financial Literacy = 3	1.60	[1.36, 2.07]
Sigma	0.768	[0.704, 0.827]

Table contains MLE results and 95% bootstrapped confidence intervals (in brackets) for $B = 1000$ repetitions.

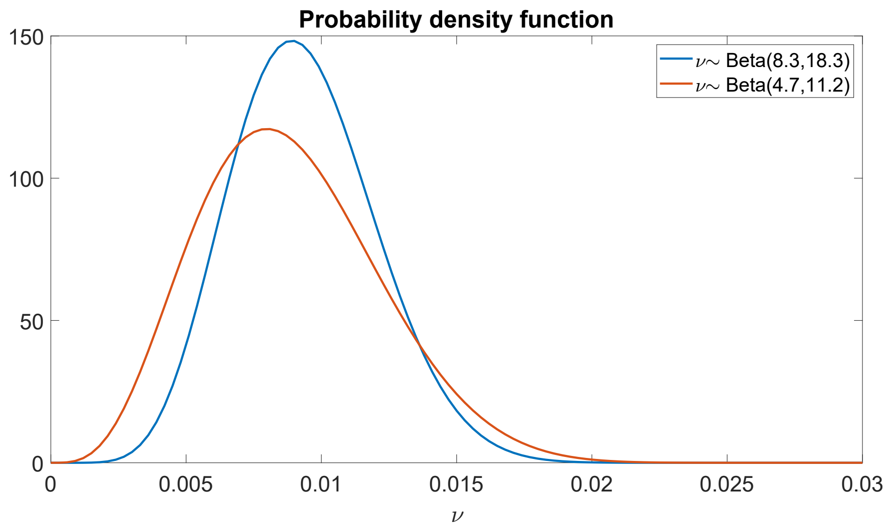


Figure 33: Shift in the estimated average distribution of the risk aversion parameter.

A.2.1 Monetary Loss Estimation Results

Table 21: Average monetary loss by group.

	Average Monetary Loss	
All	-0.2727	[-0.4235, -0.2144]
No High School	-0.4430	[-0.6834, -0.3229]
High School	-0.2744	[-0.4202, -0.2111]
Some College	-0.2034	[-0.3131, -0.1664]
College Degree	-0.2856	[-0.4450, -0.2250]
Financial Literacy = 0	-0.1566	[-0.2209, -0.1188]
Financial Literacy = 1	-0.1520	[-0.2362, -0.1168]
Financial Literacy = 2	-0.2616	[-0.3963, -0.2072]
Financial Literacy = 3	-0.2825	[-0.4408, -0.2229]
Wealth 0 - 24.9%	-0.0128	[-0.0201, -0.0096]
Wealth 25 - 49.9%	-0.0210	[-0.0327, -0.0161]
Wealth 50 - 74.9%	-0.0591	[-0.0905, -0.0464]
Wealth 75 - 89.9%	-0.1887	[-0.2930, -0.1471]
Wealth 90 - 100%	-0.5032	[-0.7835, -0.4002]

The average monetary loss is calculated and reported in $\$10,000$.

Table 22: Average Monetary Loss by Group

	Average Monetary Loss	
Low Financial Literacy & Low Wealth	-0.0568	[-0.0855, -0.0452]
High Financial Literacy & Low Wealth	-0.0398	[-0.0620, -0.0316]
Low Financial Literacy & High Wealth	-0.3583	[-0.5474, -0.2834]
High Financial Literacy & High Wealth	-0.3483	[-0.5435, -0.2753]
Low Education & Low Wealth	-0.0467	[-0.0706, -0.0380]
High Education & Low Wealth	-0.0446	[-0.0695, -0.0353]
Low Education & High Wealth	-0.3533	[-0.5446, -0.2714]
High Education & High Wealth	-0.3488	[-0.5439, -0.2772]
All	-0.2727	[-0.4235, -0.2144]

The average monetary loss is calculated and reported in $\$10,000$.

Table 23: The estimation results for expected utility estimated from the Limited Consideration Model.

	Dependent variable:
	$\mathbb{E}[u_i]$
Year	0.069***
	(0.007)
Age	−0.006
	(0.009)
Age^2	0.037***
	(0.012)
No High School	0.030**
	(0.012)
High School	0.014
	(0.009)
College Degree	0.045***
	(0.009)
Financial Literacy = 0	−0.012
	(0.017)
Financial Literacy = 1	−0.003
	(0.010)
Financial Literacy = 3	0.055***
	(0.008)
Tech/Sales/Services	0.019*
	(0.010)
Other	0.0003
	(0.011)
Not Working	0.051***
	(0.010)
Owns Stocks	0.070***
	(0.010)
Rents	0.059***
	(0.010)
Wealth 0 − 24.9%	−0.056***
	(0.012)
Wealth 24 − 49.9%	−0.029***
	(0.010)
Wealth 75 − 89.9%	0.147***
	(0.011)
Wealth 90 − 100%	0.879***
	(0.013)
Debt to Income Ratio	−0.006***
	(0.001)
Constant	−0.092***
	(0.012)
Observations	49,371
R^2	0.138
Adjusted R^2	0.138
Residual Std. Error	0.721 (df = 49351)
F Statistic	415.887*** (df = 19; 49351)
Note:	*p<0.1; **p<0.05; ***p<0.01

Base category for education is "Some College", for financial literacy is "Financial Literacy = 2, for wealth is "Wealth 50 − 74.9%", and for occupation is "Professional/Managerial".

B Mortgage Shopping Behavior in the U.S. - Stochastic Record Linkage

B.1 Motivating Findings From SCE

Motivating findings based on the data from the U.S. Survey of Consumer Expectations. Figure 34 shows that the largest mass of non-informed households is from the lowest income group. Moreover, the figure shows that the mass of non-informed households decreases with higher income. Figure 35 shows that households from the lowest income group have the highest debt-to-income ratios. In addition, Figure 36 shows that the largest shares of the highest debt-to-income ratios are in the lowest part of the income distribution. The findings from these figures imply that most exposed households are those that are the least informed about credit possibilities.

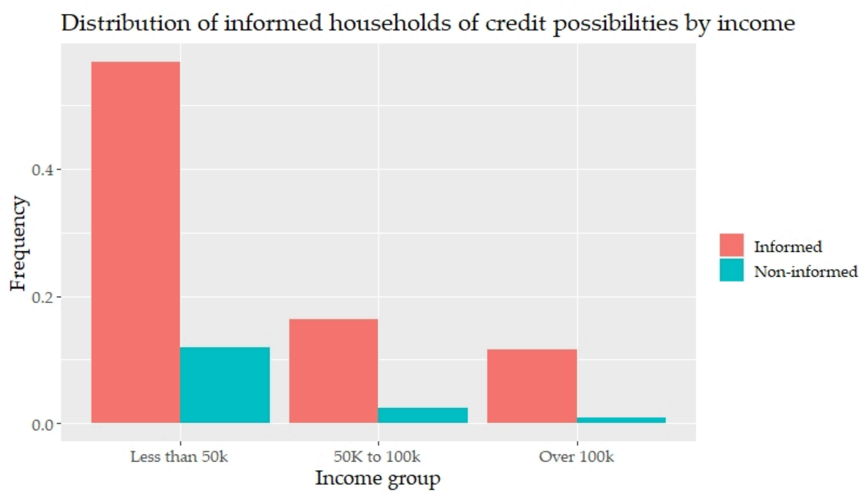


Figure 34: Share of non-informed households by income group. Source: SCE, authors' calculation.

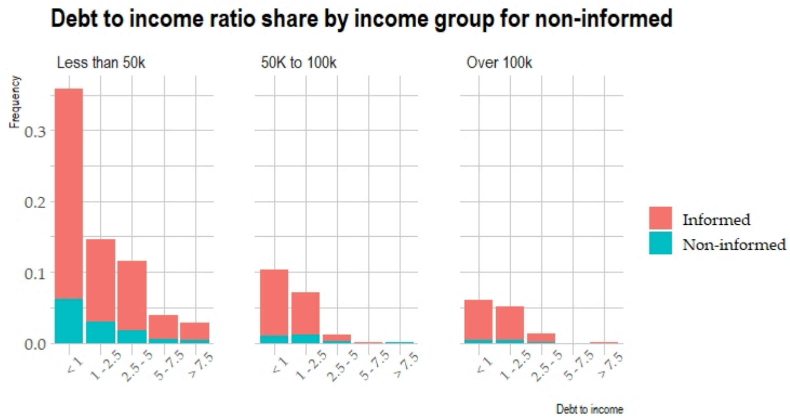


Figure 35: Share of non-informed households for each debt to income level over the income distribution. Source: SCE, authors' calculation.

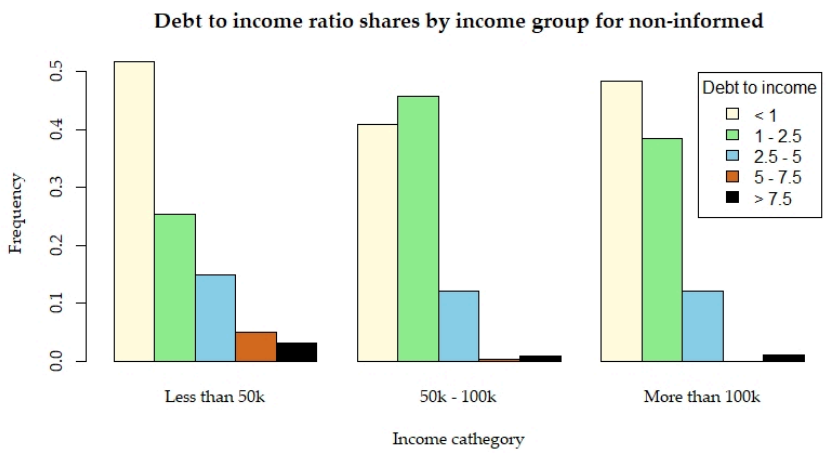


Figure 36: Debt to income ratio distributions for each income group. Source: SCE, authors' calculation.

B.2 The NSMO (2013-2020) analysis

The data on mortgages in the NSMO data range from 2013 to 2021, and tracks mortgages originated during the 2013-2020 period. Households were chosen at random to report the specifics of their mortgage contracts, reasons, and experiences. Details about mortgage origination, combined with demographic characteristics, allow us to estimate the effect of borrowers' characteristics on the acquired mortgage interest rate, controlling for mortgage specifics. First, we consider respondents' attitudes toward the mortgage market and their beliefs about the appropriateness of their lender selection. Second, we quantify the correlation between education and search effort variation and the mortgage rate attained at origination. Third, we extrapolate financial literacy from the Survey of Consumer Finances to find a link between financial skills and the interest rate obtained after the mortgage is locked in.[13]

Interestingly, almost 70% of the borrowers believe that they would be getting the same interest rate regardless of their choice of lender. 86% initiated the contact with the lender themselves. While searching for options, 48% consider only one lender/mortgage broker. Consequently, 77% applies to only one lender. However, the number of lenders considered varies with education level (Figure 37). Borrowers who apply to multiple lenders usually do so in search of better contract terms.

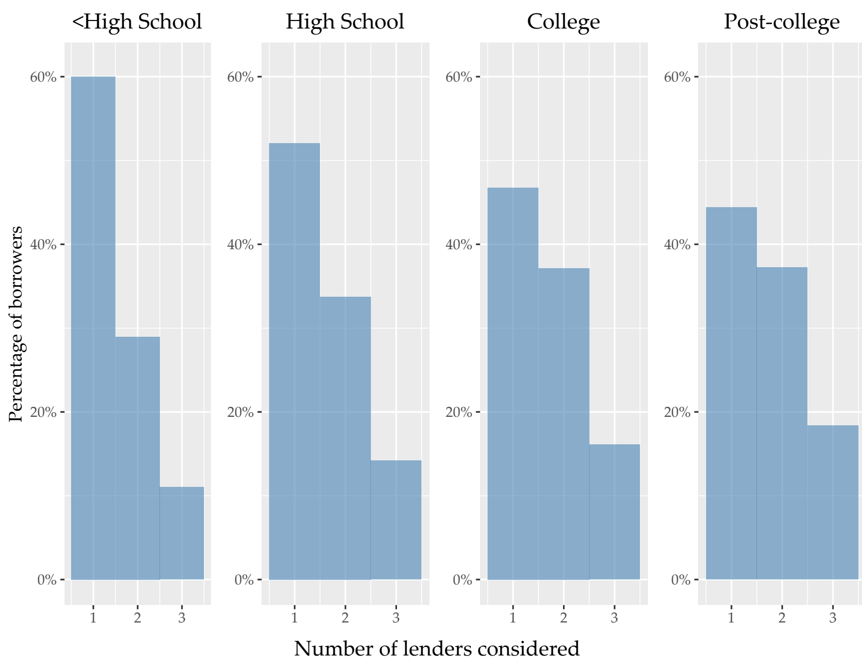


Figure 37: Number of lenders considered by education level. Source: NSMO data set, authors' calculations.

When refinancing, 88% of borrowers found lower interest rates as an important reason to start the process. Moreover, 75% of these borrowers rendered lower monthly payments as equally important. In our paper, the search model conforms to the trade-offs of a homeowner and assigns lower repayments as the benefit. Figure 38 shows that almost 60 percent of high-skilled borrowers consider two or more lenders (the right histogram), which holds for the lower percentage of low-skilled borrowers (the left histogram). In the paper, we show that financial skills remain significant for search effort and that one standard deviation increase in skill leads to a four percent increase in the probability of considering more lenders.

[13]Because we are the first to match the NSMO and the SCF to impute financial literacy scores in the NSMO, the imputation details are in the main part of the paper.

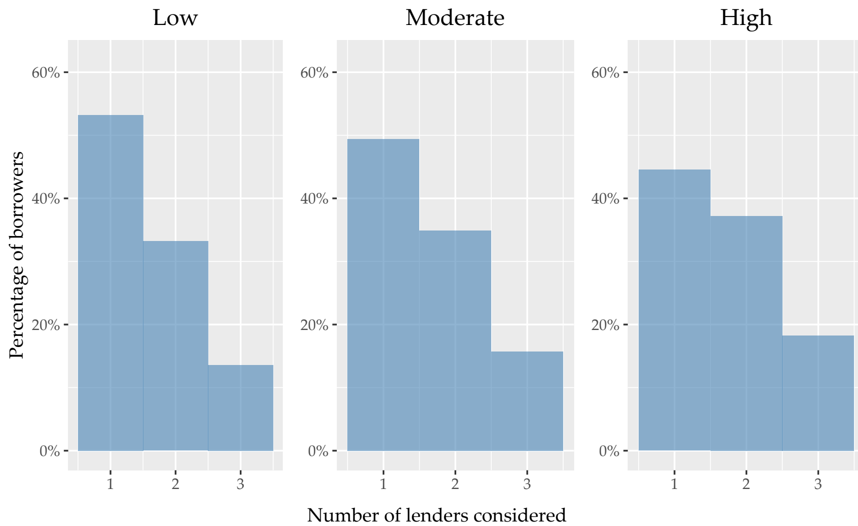


Figure 38: Number of lenders considered by financial skills tercile. Source: merged data set, authors' calculations.

Our latter findings suggest that education and effort simultaneously affect the mortgage interest rate. Using NSMO data only, we control for individual and loan characteristics to support our findings in the merged data set, as financial literacy exhibits a strong, but not perfect, correlation with education.

B.2.1 Mortgage rate regressions

Mortgage interest rates are comprised of two components: PMMS determined by the borrower's characteristics[14] and the rate spread assigned to each borrower at origination. Combining the two yields the mortgage interest rate, which is the dependent variable in the analysis.

Because nearly half of all reported mortgages are for refinancing, we estimate the linear regression separately. Both estimations control for loan-sponsorship types, guarantor enterprises (Fannie Mae, Freddie Mac, or Federal Home Loan Bank), loan amount, metropolitan (low-to-moderate) area, time effects, and the number of borrowers. The rate under refinance estimates control for non cash-out loans.

The variation in search efficacy with education is represented by interaction coefficients. Controlling for other demographic factors, we find that highly educated borrowers who shop around for loans get significantly lower interest rates. Given that we employ a novel measure that includes both cognitive and effort costs, our estimates account for an unprecedented

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

part of the interest rate dispersion (Table 24, highlighted). All interaction coefficients are statistically significant and pass difference tests.

Model predictions allow us to calculate the present value of the difference in mortgage payments over the duration of a mortgage. We think of the payment difference as the additional costs low-educated and low-shopping behavior borrowers pay. For a 30-year loan at \$200,000, high-school graduates pay on average at the 4.43% rate, whereas post-college graduates get 4.26%. The mortgage spread implies a \$9900 mortgage payment difference over the duration of the mortgage. Keeping education fixed, search effort induces the mortgage spread of 8 b.p. and implies an additional \$7500 in mortgage payments, on top of education differences. These estimates serve as a lower bound for mortgage payment losses in the market, as they abstract from additional correlations that substantiate search effort or mortgage process knowledge.

Our predicted rate plots (Figure 39) show that searches are most effective for highly educated borrowers as the predicted interest rate density moves to the left. On the other hand, those low-educated borrowers who search more do so due to the fear of rejection. All plots show that controlling for other characteristics still leaves the residual spread that borrowers face based on their education.

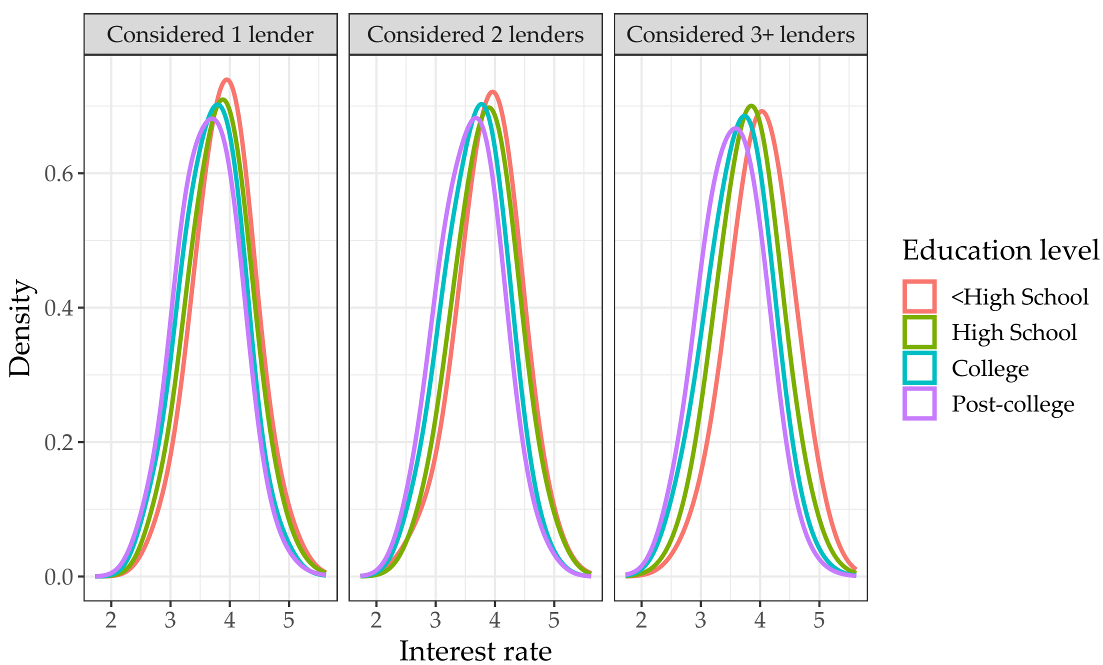


Figure 39: Predicted interest rate by education type. Each plot represents a separate case for the number of lenders considered in the mortgage process. Regression predictions, NSMO.

Table 24: Interest rate upon origination and under refinancing, explanatory characteristics, NSMO data.

	mortgage rate	
	(first origination)	(under refinancing)
Age	0.043***	0.076***
	(0.010)	(0.010)
Female	0.033***	0.033***
	(0.009)	(0.008)
Race: African-American	−0.005	0.026
	(0.019)	(0.018)
Asian	−0.020	−0.049***
	(0.020)	(0.017)
Other	0.068***	0.012
	(0.025)	(0.023)
Income: \$30,000 - \$50,000	0.008	−0.107***
	(0.024)	(0.024)
\$50,000 - \$75,000	0.034	−0.082***
	(0.023)	(0.022)
\$75,000 - \$100,000	0.031	−0.064***
	(0.024)	(0.023)
\$100,000 - \$175,000	0.061**	−0.063***
	(0.024)	(0.023)
\$175,000 or more	0.050*	−0.063**
	(0.026)	(0.025)
Credit Score	−0.264***	−0.218***
	(0.010)	(0.009)
Loan term	0.024***	0.036***
	(0.001)	(0.001)
Loan-to-Value ratio	0.004***	0.004***
	(0.0004)	(0.0003)
Number of lenders considered: 2 lenders	0.038	−0.014
	(0.030)	(0.027)
3 lenders or more	0.115**	0.053
	(0.047)	(0.038)
Education: Some college	−0.037*	−0.001
	(0.022)	(0.019)
college degree	−0.066***	−0.024
	(0.021)	(0.019)
post-college degree	−0.079***	−0.011
	(0.023)	(0.020)
Interaction: some college; considered 2	−0.028	0.005
	(0.036)	(0.033)
some college; considered 3 or more	−0.130**	−0.102**
	(0.055)	(0.045)
college degree; considered 2	−0.076**	−0.011
	(0.034)	(0.031)
college degree; considered 3 or more	−0.177***	−0.088**
	(0.051)	(0.042)
post-college degree; considered 2	−0.085**	−0.053*
	(0.035)	(0.032)
post-college degree;considered 3 or more	−0.234***	−0.131***
	(0.052)	(0.043)
Constant	5.256***	4.578***
	(0.081)	(0.070)
Observations	21,469	21,625
R^2	0.370	0.466
Residual Std. Error	23.650 (df = 21417)	20.678 (df = 21572)
F Statistic	246.159*** (df = 51; 21417)	362.082*** (df = 52; 21572)

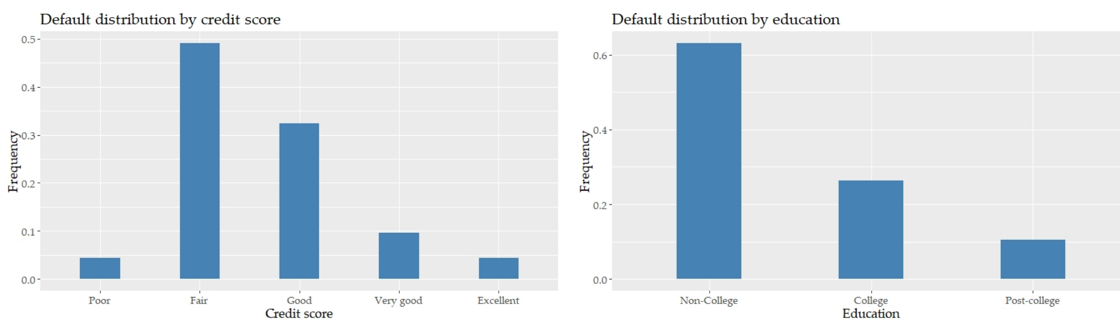
Note: Other regressors are stated in the text. *p<0.1; **p<0.05; ***p<0.01

B.2.2 Education effects in mortgage search

Because the mortgage interest rate varies with search effort, we investigate borrower characteristics that affect the amount of search borrowers are willing to take on. Controlling for loan characteristics, ordered logistic model estimates show that college and post-college graduates are 50% and 65% more likely to search more (Table 25). On the other hand, women and financially inexperienced search less. Both of these characteristics are highly correlated with financial literacy in the SCF data and this strand of literature (Lusardi et al., 2010; Lusardi and Mitchell, 2014; Lusardi, 2019).

B.2.3 What agents are most likely to default on mortgage

The NSMO dataset allows us to track mortgage performance after origination. In the main part of the paper, we show that financially skilled borrowers are 50% more likely to meet the due date of their mortgage payments. Here, we show that low-educated borrowers default more often (Figure 40b).



(a) Default by credit score. (b) Default by education.

Figure 40: Share of households that default by credit score and education. Source: NSMO, authors' calculation.

The distributions in Figure 40 shows that households that default on a mortgage and face bankruptcy are associated with lower credit scores and lower education. The only exception is those with the lowest credit scores, but household mortgage requests with "Poor" credit scores are usually denied.

Table 25: Ordered logistic regression results

	Dependent variable:	
	Number of lenders considered	
	(all originations)	(under refinancing)
Income: \$35,000-\$50,000	-0.018	-0.013
	(0.053)	(0.077)
\$50,000-\$75,0000	-0.024	-0.034
	(0.050)	(0.071)
\$75,000-\$100,000	-0.024	-0.070
	(0.051)	(0.073)
\$100,000-\$175,000	-0.054	-0.157^{**}
	(0.051)	(0.074)
\$175,000 or more	-0.090	-0.162^{**}
	(0.056)	(0.081)
Education: some college	0.267^{***}	0.263^{***}
	(0.035)	(0.049)
college degree	0.408^{***}	0.383^{***}
	(0.035)	(0.048)
post-college degree	0.501^{***}	0.431^{***}
	(0.036)	(0.051)
Female	-0.279^{***}	-0.336^{***}
	(0.019)	(0.027)
Age	-0.177^{***}	-0.040
	(0.019)	(0.030)
Have stocks	-0.097^{***}	-0.103^{***}
	(0.020)	(0.029)
Metro area, low-to-moderate income tract	0.007	-0.036
	(0.029)	(0.041)
Non-metro area	-0.053^{*}	-0.071
	(0.032)	(0.046)
Observations	43,094	21,625
Note: Controlled for time and loan amount effects.		*p<0.1; **p<0.05; ***p<0.01

B.3 SCF data analysis

We use the Bayesian Record Linkage algorithm to impute the financial literacy score from the SCF data into the NSMO data. To begin, we examine the average financial literacy score over the lifecycle to motivate investment in, and accumulation of financial skills in the model. Figure 9 shows increasing average financial literacy scores by age groups.

The first model estimates outline correlations between financial literacy and household characteristics. Our predicted probabilities of the ordered logistic model (Table 26) suggest that high-income level households are 12% more likely to be fully financially skilled, keeping other characteristics fixed. Though education explains the largest part of financial literacy, income-based differences relate to financial skills needed to understand the mortgage refinancing process.

Next, we restrict the SCF sample to borrowers who hold a mortgage on their primary residence and estimate a binary regression model to evaluate their likelihood of refinancing. The estimates pinpoint vital characteristics that explain a household's effort in shopping for credit.

Controlling for income and mortgage size, we find significant and large effects of financial literacy - a high financial literacy score relates to a 60% higher likelihood of refinancing. In contrast, education effects are insignificant (Table 27). Our analysis supports Lusardi (2019) and highlights the relevance of the financial knowledge margin in the decision to refinance.

Using the question about the amount of shopping time allocated to borrowing options, we proxy borrower's search effort and find a 12% higher likelihood of refinancing by borrowers who allocate time to exploring borrowing options (Table 27). Further, keeping other characteristics fixed, financial knowledge, and search effort positively correlate with the decision to refinance. As a result, the mortgage search model with financial skills investment and search effort disentangles the two dimensions relevant to the decision to refinance.

Our estimates on credit shopping behavior emphasize financial skills as an important dimension of heterogeneity (Table 5). While mortgage owners shop more on average, separate analyses for mortgage owners and renters reach the same conclusion: controlling for individual characteristics, including age, income, and education, financially savvy borrowers spend more time searching for credit.

Keeping other characteristics fixed at the mean of each subsample, we plot the likelihood change over financial literacy level and monthly housing expenses. Homeowners are more likely to spend a lot more time shopping for credit than renters. Specifically, financially savvy homeowners are up to 15 p.p. more likely to allocate more time to credit shopping

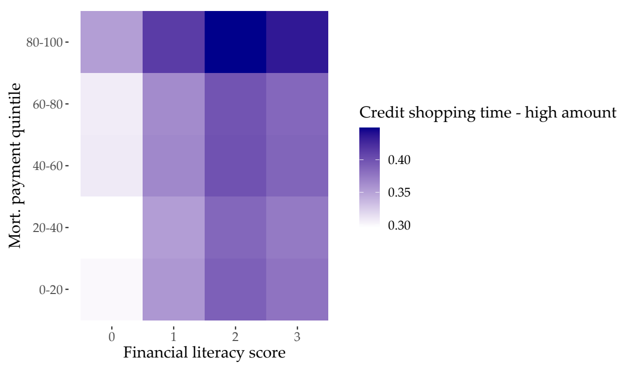
Table 26: Financial Literacy Score, relation to observables. Source: SCF data.

	Dependent variable:
	Financial literacy score
Worker	0.041*
	(0.025)
Married	0.111***
	(0.024)
Non-white	−0.392***
	(0.019)
Female	−0.474***
	(0.025)
Education: High-school	0.211***
	(0.031)
Some college	0.599***
	(0.031)
College degree	1.123***
	(0.033)
Income percentile: 20^{th} - 40^{th}	0.049*
	(0.028)
40^{th} - 60^{th} 3	0.073**
	(0.031)
60^{th} - 80^{th}	0.179***
	(0.035)
80^{th} - 90^{th}	0.349***
	(0.043)
90^{th} - 100^{th}	0.649***
	(0.048)
Observations	60,125
Note: Controlling for age and asset amount.	*p<0.1; **p<0.05; ***p<0.01

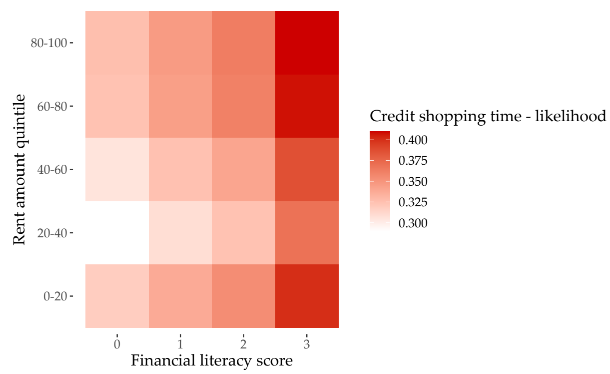
Table 27: Binary regression estimates, likelihood to refinance, SCF data.

	Dependent variable:
	Ever refinanced their mortgage
Financial literacy score: low	0.099
	(0.104)
medium	0.252^{***}
	(0.098)
high	0.400^{***}
	(0.098)
Search effort, borrowing: medium	0.055
	(0.050)
high	0.110^{**}
	(0.052)
Female	0.075
	(0.049)
non-white	-0.247^{***}
	(0.034)
Mortgage size: $83,000 - $159,000	-0.148^{***}
	(0.042)
$159,001 - $ 297,000	-0.285^{***}
	(0.044)
$ 297,001 - $ 1,450,000	-0.304^{***}
	(0.050)
Liquid savings: $\leq$ $4,500	0.145^{***}
	(0.049)
$4,500 - $21,000	- 0.045
	(0.050)
$\geq$$21,000	-0.017
	(0.051)
Income percentile group: 20^{th}-40^{th}	0.242^{***}
	(0.083)
40^{th}-60^{th}	0.260^{***}
	(0.079)
60^{th}-80^{th}	0.482^{***}
	(0.079)
80^{th}-90^{th}	0.874^{***}
	(0.084)
top 10	1.047^{***}
	(0.085)
Constant	-0.961^{***}
	(0.145)
Observations	22,178
Note: Controlled for age, family structure, education, and survey wave effects.	*p<0.1; **p<0.05; ***p<0.01

than low-skilled homeowners (Figure 41, left). The difference in likelihood decreases with the size of their mortgage payment. In contrast, renters allocate their time to credit shopping independently of their rent amount, and financially skilled are 10 p.p. more likely to spend a great deal of time in searching for credit (Figure 41, right).



(a) Likelihood variation, homeowners.



(b) Likelihood variation, renters.

Figure 41: Great deal of time spent shopping for credit, SCF data. Ord. logit predictions.

B.3.1 Rent and mortgage payments as shares of labor income

In the model calibration, we inform the rental rate κ with the share of homeowners in the SCF. When compared to an average mortgage monthly payment, rental payments are twice as high. The averages from the SCF data are computed for the subsample of workers up to age 55 with wage income higher than the yearly amount of retirement benefits. Sample averages show that monthly rental payments are up to two times higher than monthly mortgage payments.

Living arrangement	Financial literacy score			
	0	1	2	3
Homeowner	0.140	0.139	0.142	0.129
Renter	0.257	0.241	0.233	0.222

Table 28: First row: monthly mortgage payment as a share of income - homeowners, second row: monthly rent as a share of income; renters. SCF data, worker subsample.

B.3.2 Homeownership choice and financial literacy

Our model assumes that the homeownership choice depends on individual assets, financial skills, and productivity. As a result, the model's equilibrium generates a positive correlation between mortgage take-up and financial skills, which aligns with the similar positive association we observe in the SCF data. Table 29 presents estimates from the logistic regression, where we regress the choice to rent or own against a set of observable characteristics, including skills, assets, and wage income. To maintain consistency with our model, the estimates are derived from a subsample of workers. The first two rows in the coefficient table 29 show that the likelihood of owning a home increases with skills, with age and wage income showing the same direction. Importantly, education is non-significant and varies in the direction of the correlation. The SCF data reinstate the salience of individual skills in financial behavior and choice.

B.4 Bayesian Record Linkage method (BRL)

Recently developed in Enamorado et al. (2019), Bayesian Record Linkage (BRL) is a probabilistic approach designed to match census data. Unlike deterministic methods such as mean-imputation and cluster-based algorithms commonly used in standard imputation, BRL leverages probabilistic techniques to account for the uncertainty inherent in the merging

Table 29: Binary regression estimates, homeownership choice, SCF data.

	Dependent variable:
	Owns a house or an apartment
Financial literacy score: medium	0.170***
	(0.038)
high	0.146***
	(0.039)
Education: high-school	0.067
	(0.052)
some college	-0.051
	(0.052)
college	-0.039
	(0.056)
Married	-0.852***
	(0.042)
Female	0.176***
	(0.044)
non-white	−0.536***
	(0.029)
Leverage ratio	-0.029***
	(0.003)
Willing to take risk	0.009
	(0.063)
Wage income quartile: \$ 25,800 -\$58,200	0.235***
	(0.041)
\$58,200 - \$117,000	0.778***
	(0.047)
≥\$117,000	1.143***
	(0.061)
Constant	−1.112***
	(0.064)
Observations	40,071
Note: Controlled for age, family structure, occupation category, liquid savings amount, and survey wave effects.	*p<0.1; **p<0.05; ***p<0.01

process. The advantages of employing BRL in this context include its scalability to handle large datasets and its ability to facilitate post-merge analyses through the utilization of match-specific posterior weights.

In the context of Bayesian Record Linkage (BRL), the matching process assigns posterior probabilities of a match for each record pair (i, j), where i represents the records from the NSMO data ($i \in \mathcal{A}$), and j corresponds to the SCF dataset ($j \in \mathcal{B}$). The BRL method employs pairwise comparisons for each distinct record pair (i, j) and computes the probability of a match based on the presence of a specific set of common observables denoted as K. The selection of these common observables focuses on factors generally considered relevant for assessing individual financial skills, including income, education, gender, age, race, occupation, family characteristics, retirement plan, and asset holdings. Table 30 shows the population shares in SCF and NSMO for every common observable used in the matching process. To ensure consistency in the matching procedure, we impose certain restrictions on the SCF sample. Specifically, we only include homeowners who hold a first lien mortgage, while we make no restrictions to the NSMO sample.

Table 30: Population shares in the respective samples. Source: NSMO 2013-2022 and SCF 2016-2019, authors' calculations.

	Data set	
	NSMO	SCF
income brackets	[6%, 9% , 18%, 19%, 30%, 18%]	[13%, 8%, 13% ,11%,20%, 35%]
education brackets	[1%, 10%, 5%, 20%, 35%, 29%]	[6%, 18%, 9%, 15%, 27%, 25%]
gender (Female,Male)	[44%, 55%]	[17%,83%]
age (<35,35-44,45-54,55-64,65-74,>=75)	[18%, 22%, 22%, 21%, 14% ,3%]	[8%, 14%, 20%, 26% , 20%, 12%]
race (Caucasian, African-American, other)	[84%, 6%, 10%]	[82%, 7%, 11%]
occupation (Employed, Self-employed, Retired/Student, Other)	[68%, 10%, 19% ,2%]	[47%, 26%, 25%, 2%]
has kids (Yes, No)	[64%, 36%]	[60% , 40%]
owns financial assets (Yes, No)	[57%, 43%]	[58% 42%]
retirement plan participation (Yes, No)	[86%, 14%]	[62%, 38%]

For each of $card(\mathcal{A}) \times card(\mathcal{B})$ distinct observations, BRL defines an agreement vector $\gamma(i, j)$

of length K. The k-th element $\gamma_k(i,j)$ represents the degree of agreement corresponding to the k-th observable in the set of mutual observables[15]. Following Enamorado et al. (2019), for a given observable k, we assume the agreement degree to be discrete, with a maximum $L_k - 1$.

Based on variable k (for example, income category), $\gamma_k(i,j) = 0$ represents a no-match, whereas agreement level $\gamma_k(i,j) = L_k - 1$ corresponds to a perfect match for a pair of records (i,j). Therefore, two records from SCF and NSMO may be matching in education brackets but may differ in income levels, leading to a lower degree of agreement. The BRL takes every agreement degree into account and evaluates the posterior probability conditional on all agreement degrees for the pair. For each observation in the NSMO, we obtain the distribution of matches across the SCF sample.

BRL builds on the Fellegi-Sunter model (Fellegi and Sunter, 1969): $M_{i,j}$ denotes a latent mixing variable that shows whether distinct records pair (i,j) form a match or not. That is, M_{ij} is Bernoulli-distributed

$$M_{i,j} \overset{i.i.d.}{\sim} \mathrm{B}(\lambda),$$

and k-based agreement level $\gamma_k(i,j)$ has a discrete distribution

$$\gamma_k(i,j)|M_{i,j} \sim \begin{pmatrix} 0 & 1 & \dots & L_k - 1 \\ \pi_{k0} & \pi_{k1} & \dots & \pi_{kL_k-1} \end{pmatrix},$$

where $\pi_{kl}, \quad l \in \{0, \dots, L_k - 1\}$ represents the probability of each agreement degree for the pair (i,j). The vector of probabilities is denoted with $\boldsymbol{\pi}_{km}$.

Record matching probabilities imply the observed-data likelihood $\mathcal{L}_{\text{obs}}$, that we estimate later using the Expectation-Maximization algorithm (suggested by Enamorado et al. (2019)). Using the matched records from the NSMO and SCF data, we apply the Bayesian posteriors $\epsilon_{i,j} = \mathbb{P}(M_{ij} = 1|\gamma(i,j))$ as weights for statistical inference when we use the (imputed) financial literacy score. This way, we incorporate the match procedure uncertainty and avoid biases that emerge in standard deterministic methods.

[15]Income brackets are not listed for compactness; we group income in the SCF according to brackets in the NSMO data: (<$35,000,$35,000-$50,000,$50,000-$75,000,$75,000-$100,000,$100,000-$175,000, >$175,000). Similarly, we take the highest education grade data in the SCF and group them according to education brackets in the NSMO: (Some schooling, High-School graduate, Technical School, Some College, College degree, Post-college degree).

Bayes rule implies the probability of a match which defines the post-merge weight

$$\begin{aligned}\varepsilon_{ij} &= \mathbb{P}(M_{ij} = 1 \mid \gamma(i,j)) \\ &= \frac{\lambda \prod_{k=1}^{K} (\prod_{l=0}^{L_k-1} \boldsymbol{\pi}_{k1l}^{\mathbf{1}\{\gamma_k(i,j)=l\}})}{\sum_{m=0}^{1} \lambda^m (1-\lambda)^{1-m} \prod_{k=0}^{K} (\prod_{l=0}^{L_k-1} \boldsymbol{\pi}_{kml}^{\mathbf{1}\{\gamma_k(i,j)=l\}})},\end{aligned}$$

that we use later for statistical inference. Financial literacy for the borrower i, $\bar{Z}_i$ is the sum of literacy scores of the respective record matches in the SCF Z_j, with corresponding weights ε_{ij}[16]:

$$\bar{Z}_i = \frac{\sum_{j=1}^{N_{\mathcal{B}}} \varepsilon_{ij} Z_j}{\sum_{j=1}^{N_{\mathcal{B}}} \varepsilon_{ij}}.$$

Post-merge analysis includes $\bar{Z}_i$ as the independent variable in linear model estimates.

Non-linear models, such as the ordered logistic and binary regression models we use for inference, need to be adjusted with the posterior weight. Therefore, the maximum likelihood function includes all the record pair matches with the corresponding Bayesian weight. With the assumption $Y_i | X_i, Z_i^* \overset{indep.}{\sim} P_\theta(Y_i | X_i, Z_i^*)$, the ML estimator

$$\hat{\theta} = \sum_{i=1}^{\mathcal{N}_A} \sum_{j=1}^{\mathcal{N}_B} \varepsilon_{ij}^* \log P_\theta(Y_i | X_i, Z = Z_j^*), \quad \varepsilon_{ij}^* = \frac{\varepsilon_{ij}}{\sum_{j=1}^{\mathcal{N}_B} \varepsilon_{ij}}$$

is consistent and asymptotically normal and hence follows standard rules of significance tests. We use these theoretical results derived in Enamorado et al. (2019) and implement our estimators that ensure solid statistical properties.

B.4.1 Number of lenders considered

For every record pair (i, j) with a corresponding match weight ε_{ij}^*, the likelihood of number of lenders considered num_cons is characterized using the borrower's observables $(X_i, \text{fin\_skills}_i)$

$$\mathbb{P}(\text{num\_cons}_{ij} = k) = p_{ij,k} = \mathbb{P}\Big(-\kappa_{k-1} < \beta X_i + \beta^f \text{fin\_skills}_j + u_{ij,k} < \kappa_k\Big), \quad k \in \{1, 2, 3+\},$$

with κ_{k-1} and κ_k representing latent thresholds that define the search effort level. The logistic model assumes

$$p_{ij,k} = \frac{1}{1 + \exp\Big(-\kappa_k + \beta X_i + \beta^f \text{fin\_skills}_i\Big)} - \frac{1}{1 + \exp\Big(-\kappa_{k-1} + \beta X_i + \beta^f \text{fin\_skills}_i\Big)},$$

[16]Our merging procedure uses the standardized literacy score.

which pins down the log-likelihood adjusted by the posterior match weight

$$\ln L = \sum_{i=1}^{\mathcal{N}_A} \sum_{j=1}^{\mathcal{N}_B} \varepsilon_{ij}^* \sum_{k=1}^{3+} \mathbf{1}_{\{\text{num\_cons}_{ij}=k\}} \ln(p_{ij,k} | X_i, \text{fin\_skills}_j).$$

B.4.2 Additional NSMO+ estimates

As an additional counterfactual exercise, we estimate the linear probability model where the dependent variable is the number of lenders considered with our new NSMO+ dataset. We estimate the model when the number of lenders considered equals one versus more than one. Estimates are presented in Table (31). The results imply a strong positive correlation between higher financial skills and the probability of considering more than one lender when searching for a mortgage. In particular, the model predicts that an average borrower who answered zero questions correctly has a probability of considering more than one lender equal to 0.381. On the other hand, for an average financially savvy borrower who answered all questions correctly, our linear probability model predicts a 0.546 probability of considering more than one lender. The model predicts similar probabilities of considering more than one lender for average borrowers upon refinancing the mortgage.

Table 31: Linear probability model for the number of lenders considered one vs. more. Source: NSMO+, own calculation.

	Lenders considered	
	All origination	Refinancing
Age	−0.042***	−0.019**
	(0.005)	(0.008)
Credit Score	0.009*	0.005
	(0.005)	(0.007)
Married	0.020***	0.014
	(0.006)	(0.010)
Female	−0.058***	−0.076***
	(0.005)	(0.007)
Race: Black or African-American	0.055***	0.038**
	(0.011)	(0.015)
Asian	0.055***	0.055***
	(0.010)	(0.014)
other (including hispanic)	0.059***	0.083***
	(0.014)	(0.020)
Financial Literacy	0.164***	0.166***
	(0.038)	(0.056)
Education: high school	0.056***	0.052***
	(0.009)	(0.013)
college graduate	0.090***	0.075***
	(0.009)	(0.013)
post-college graduate	0.107***	0.086***
	(0.010)	(0.014)
Loan amount: $50,000 - $99,999	0.019	0.066**
	(0.019)	(0.029)
$100,000 - $149,999	0.037*	0.130***
	(0.019)	(0.029)
$150,000 -$199,999	0.047**	0.154***
	(0.020)	(0.029)
$200,000 - $249,999	0.066***	0.152***
	(0.020)	(0.030)
$250,000 to $299,999	0.071***	0.167***
	(0.021)	(0.031)
$300,000 -$349,999	0.071***	0.180***
	(0.021)	(0.032)
$350,000 - $399,999	0.088***	0.182***
	(0.022)	(0.033)
≥$400,000	0.099***	0.176***
	(0.021)	(0.031)
Constant	0.271***	0.246***
	(0.047)	(0.068)
Observations	43,084	21,623
R^2	0.024	0.025
Adjusted R^2	0.023	0.023
Residual Std. Error	17.837 (df = 43039)	17.676 (df = 21578)
F Statistic	23.681*** (df = 44; 43039)	12.666*** (df = 44; 21578)
Note:		*p<0.1; **p<0.05; ***p<0.01

Controlled for: Loan type, Year, Government Sponsored Enterprise, Term, LTV, Number of borrowers, and Income.

C Tax Structures and Fiscal Multipliers in HANK Models

C.1 HFCS

In this section, we classify the three groups of households by their hand-to-mouth status using two slightly different definitions. Moreover, we outline the construction of variables used in our analysis. The variable names refer to wave 4 of the Household Finance and Consumption Survey.

We follow Kaplan et al. (2014) and Slacalek et al. (2020), and use the following definition. A household is considered as hand-to-mouth (HtM) if:

- Net liquid wealth $\geq$0 & net liquid wealth $\leq$ biweekly (net) income

  or

- Net liquid wealth $<$ 0 & net liquid wealth $\leq$ biweekly (net) income $-$ credit limit.

Moreover, a household is:

- Poor HtM if it is HtM & net illiquid wealth $\leq$ 0,
- Wealthy HtM if it is HtM & net illiquid wealth $>$ 0,
- Non-HtM if it is not HtM.

We compare the resulting HtM decomposition using the additional definition of Slacalek et al. (2020). They classify all HtM households with some housing assets as wHtM, including households whose mortgage exceeds the house's value. In addition, they classify all HtM households with some self-employment business wealth as wHtM.

As Slacalek et al. (2020) also use the HFCS, we follow their construction of liquid and illiquid asset variables. The variables used in specifying wHtM and pHtM households are defined as follows:

- Net liquid wealth = liquid assets $-$ liquid liabilities
- Liquid assets = sight and saving accounts (deposits), directly held mutual funds, bonds, and stocks

- Liquid liabilities = overdraft debt and credit card debt
- Net illiquid wealth = illiquid assets − illiquid liabilities
- Illiquid assets = illiquid real assets, the value of the household main residence and other properties and the value of self-employment businesses
- Illiquid liabilities = amount of non-collateralized loans for household main residence and other properties, mortgage debt.

We assume that the credit limit is one month of income. Moreover, we use a simplified definition of net income. For each country in the HFCS, we use the average tax wedge from the OECD Tax Database (https://www.oecd.org/tax/tax-policy/tax-database/). Specifically, we define net income as:

- net income = (1-τ)*(employment income + 2/3 self employment income) + non taxable income.

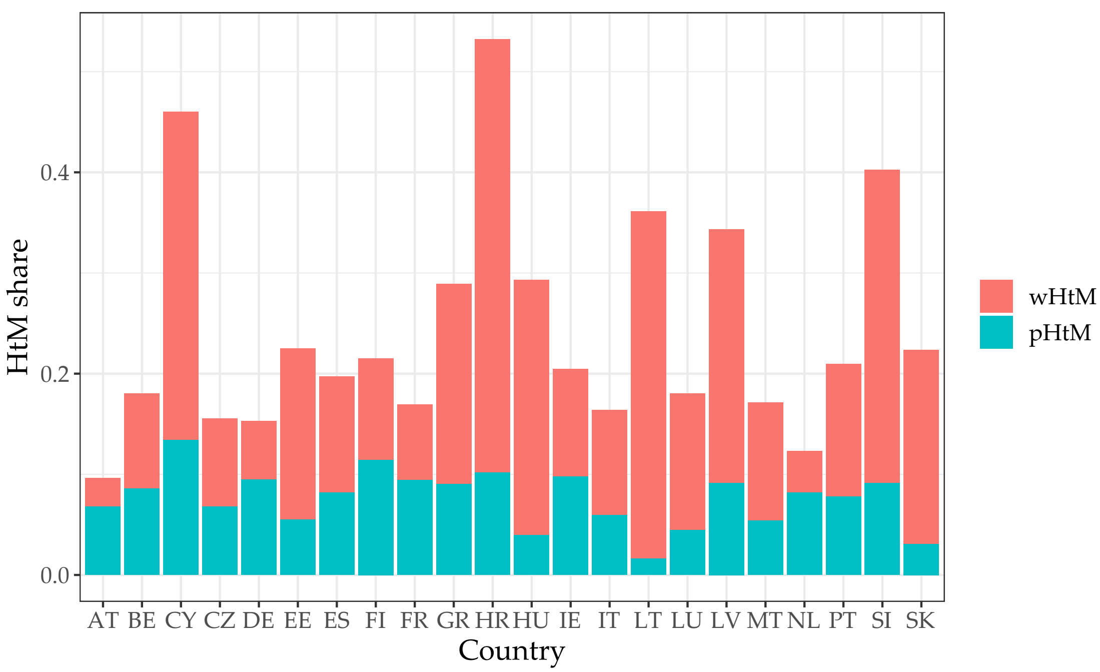


Figure 42: wHtM and pHtM shares for a set of European countries; Slacalek et al. (2020) definition. Source: Eurosystem Household Finance and Consumption Survey, wave 4.

Figure 42 shows heterogeneity is shares of HtM, wHtM, and pHtM households accross countries. The majority of countries have HtM share lower than 41%, which is the share of

Country Code	HtM	wHtM	pHtM	wHtM*	pHtM*	Net liquid	Net illiquid
AT	0.10	0.03	0.07	0.03	0.07	42.82	233.03
BE	0.18	0.09	0.09	0.09	0.09	87.06	301.27
CY	0.46	0.33	0.13	0.31	0.15	18.52	312.73
CZ	0.16	0.09	0.07	0.08	0.07	12.16	118.21
DE	0.15	0.06	0.09	0.06	0.10	55.06	244.05
EE	0.22	0.17	0.06	0.16	0.06	21.50	128.80
ES	0.20	0.12	0.08	0.10	0.10	40.34	210.27
FI	0.21	0.10	0.11	0.09	0.12	46.44	152.18
FR	0.17	0.08	0.09	0.07	0.10	35.73	217.24
GR	0.29	0.20	0.09	0.19	0.10	12.79	111.58
HR	0.53	0.43	0.10	0.43	0.10	2.45	108.05
HU	0.29	0.25	0.04	0.25	0.04	9.47	87.45
IE	0.20	0.11	0.10	0.10	0.10	39.29	307.86
IT	0.16	0.10	0.06	0.10	0.06	46.99	270.95
LT	0.36	0.35	0.02	0.34	0.02	3.30	82.20
LU	0.18	0.14	0.04	0.13	0.05	255.75	949.72
LV	0.34	0.25	0.09	0.25	0.10	7.89	60.36
MT	0.17	0.12	0.05	0.12	0.05	38.59	360.68
NL	0.12	0.04	0.08	0.04	0.08	52.87	152.55
PT	0.21	0.13	0.08	0.13	0.08	25.72	156.71
SI	0.40	0.31	0.09	0.31	0.09	11.77	169.43
SK	0.22	0.19	0.03	0.19	0.03	9.49	108.21

Table 32: Htm, wHtM, and pHtM shares and net liquid and illiquid asset positions in thousands of EUR for a set of European countries. Note: * denotes wHtM and pHtM shares using the definition of Kaplan et al. (2014). Source: Eurosystem Household Finance and Consumption Survey, wave 4.

HtM households in the U.S. (Kaplan and Violante, 2022). The figure also shows heterogeneity across countries in both wHtM and pHtM shares.

Figures 43 and Figure 44 show net liquid and illiquid asset holdings in absolute terms, i.e., in thousands of EUR.

Table 32 documents shares of HtM, wHtM, and pHtM households across countries as well as net liquid and illiquid asset positions depicted in Figures 17,43,44, and 42.

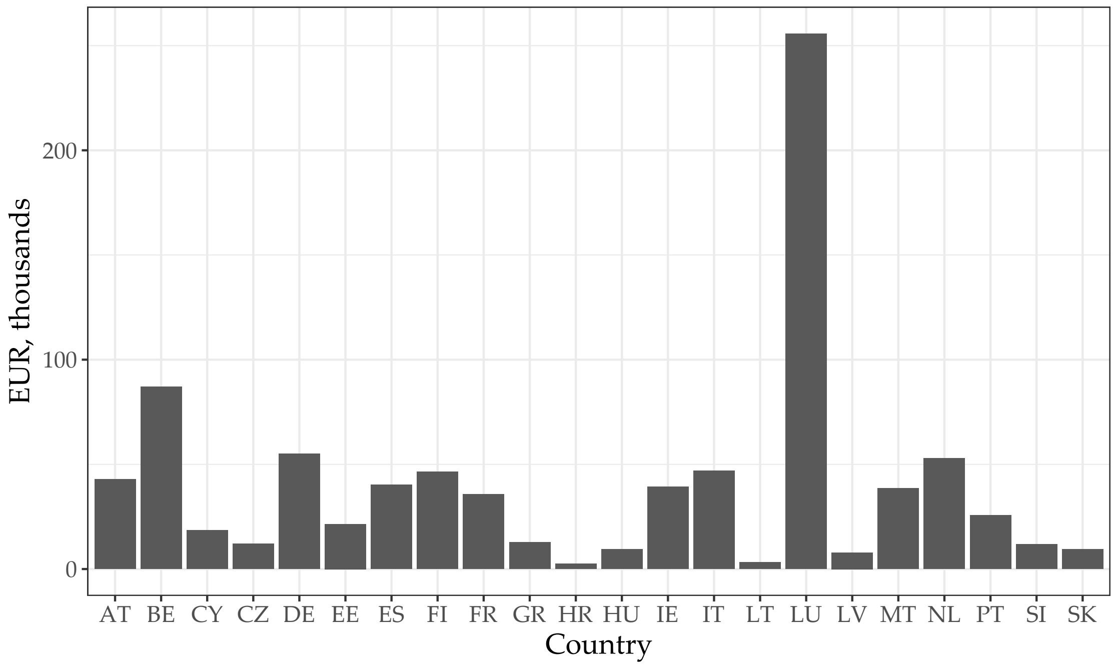


Figure 43: Net liquid asset holdings for a set of European countries. Source: Eurosystem Household Finance and Consumption Survey, wave 4.

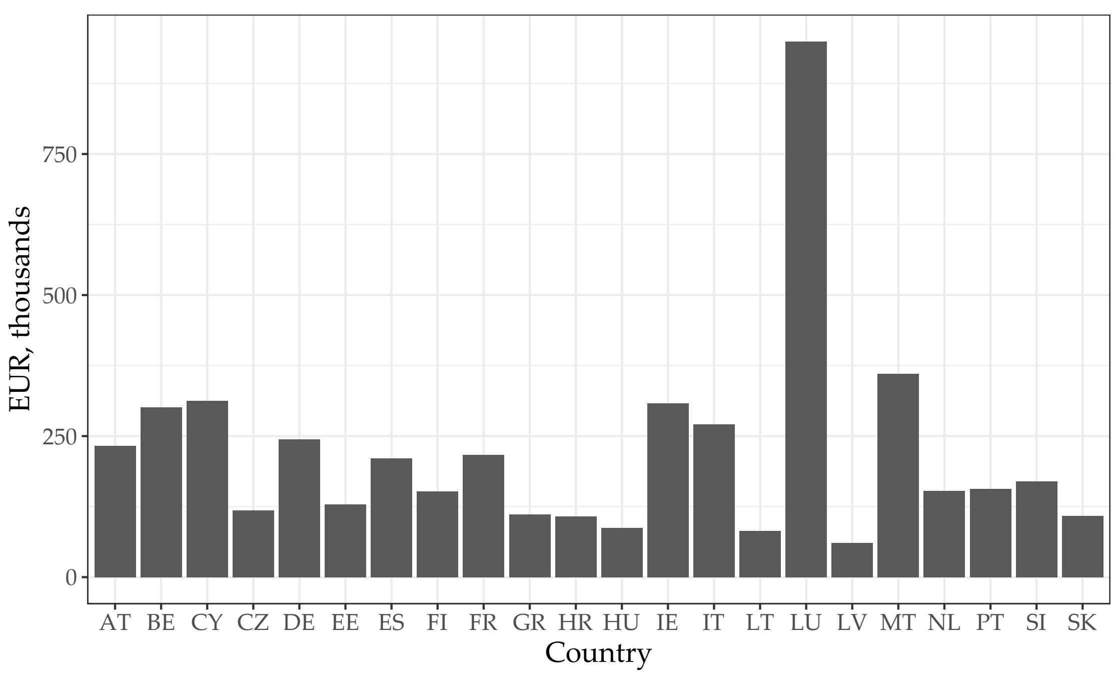


Figure 44: Net illiquid asset holdings for a set of European countries. Source: Eurosystem Household Finance and Consumption Survey, wave 4.

C.2 Household problem description

In this section, we derive the first order and envelope conditions. The Bellman equation can be rewritten as

$$\begin{aligned}
V_t(z_{i,t}, b_{i,t-1}, a_{i,t-1}, \beta_{i,t}) = \max_{b_{i,t}, a_{i,t}} u\Bigg(&\frac{1}{1+\tau^c}(z_{i,t} + (1+r_t^a)a_{i,t-1} + (1+r_t^b)b_{i,t-1} - \\
&- \Psi(a_{i,t}, a_{i,t-1}) - a_{i,t} - b_{i,t})\Bigg) + \\
&+ \lambda_{i,t} b_{i,t} + \mu_{i,t} a_{i,t} + \beta_{i,t}\mathbb{E}V_{t+1}(z_{i,t+1}, b_{i,t}, a_{i,t}, \beta_{i,t+1}),
\end{aligned}$$

where $\lambda_{i,t}$ and $\mu_{i,t}$ are Lagrange multipliers on non-negativity constraints for both types of assets. The first order conditions with respect to $b_{i,t}$ and $a_{i,t}$ are given with

$$u'(c_{i,t})\frac{1}{1+\tau^c} = \lambda_{i,t} + \beta_{i,t}\mathbb{E}\partial_b V_{t+1}(z_{i,t+1}, b_{i,t}, a_{i,t}, \beta_{i,t+1})$$

and

$$u'(c_{i,t})\frac{1}{1+\tau^c}\Bigg(1 + \Psi_1(a_{i,t}, a_{i,t-1})\Bigg) = \mu_{i,t} + \beta_{i,t}\mathbb{E}\partial_a V_{t+1}(z_{i,t+1}, b_{i,t}, a_{i,t}, \beta_{i,t+1}).$$

Lastly, the envelope conditions are

$$\partial_b V_t(z_{i,t}, b_{i,t-1}, a_{i,t-1}, \beta_{i,t}) = (1+r_t^b)u'(c_{i,t})\frac{1}{1+\tau^c}$$

and

$$\partial_a V_t(z_{i,t}, b_{i,t-1}, a_{i,t-1}, \beta_{i,t}) = \Bigg(1 + r_t^a - \Psi_2(a_{i,t}, a_{i,t-1})\Bigg)u'(c_{i,t})\frac{1}{1+\tau^c}.$$

To solve this part of the model, we follow Auclert et al. (2021) and use the endogenous gridpoints method of Carroll (2006). Details of the implementation can be found in the appendix of Auclert et al. (2021).

C.3 Derivation of the nonlinear wage NKPC

To derive the wage NKPC, we first use the definition of the real wage w_t and expression for the demand curve to rewrite $z_{i,t}$:

$$\begin{aligned} z_{i,t} &= \tau_t(w_t N_{k,t} e_{i,t})^{1-\theta} + T_t = \tau_t\left(\frac{W_{k,t}}{P_t} N_{k,t} e_{i,t}\right)^{1-\theta} + T_t \\ &= \tau_t\left(\frac{W_{k,t}}{P_t} e_{i,t}\left(\frac{W_{k,t}}{W_t}\right)^{-\varepsilon} N_t\right)^{1-\theta} + T_t. \end{aligned}$$

Second, we note that applying the Euler theorem to the household's problem (9) implies $\frac{\partial c_{i,t}}{\partial W_{k,t}} = \frac{1}{1+\tau_t^c}\frac{\partial z_{i,t}}{\partial W_{k,t}}$. Using the expression derived above, and exploiting the fact that in equilibrium $W_{k,t} = W_t$ we get that

$$\begin{aligned} \frac{\partial z_{i,t}}{\partial W_{k,t}} &= (1-\theta)\tau_t\left(\frac{W_{k,t}}{P_t} N_{k,t} e_{i,t}\right)^{-\theta} \frac{e_{i,t}}{P_t}\left(N_{k,t} - W_{k,t}\varepsilon\left(\frac{1}{W_t}\right)^{-\varepsilon} N_t W_{k,t}^{-\varepsilon-1}\right) \\ &= (1 - MTR_{i,t})\frac{e_{i,t}}{P_t} N_{k,t}(1-\varepsilon), \end{aligned} \tag{14}$$

where $MTR_{i,t} = 1 - (1-\theta)\tau_t\left(\frac{W_{k,t}}{P_t} N_{k,t} e_{i,t}\right)^{-\theta}$ is marginal tax rate of household i at time t. Lastly, since household i's total hours work equal $\left(\frac{W_{k,t}}{W_t}\right)^{-\varepsilon} N_t$, we have that hours worked also satisfy

$$\frac{\partial n_{i,t}}{\partial W_{k,t}} = -\varepsilon\frac{N_{k,t}}{W_{k,t}}. \tag{15}$$

Now, we take the first order condition of the union k's problem with respect to $W_{k,t}$ as well as the envelope condition and obtain

$$\int\left(u'(c_{i,t})\frac{\partial c_{i,t}}{W_{k,t}} - v'(n_{i,t})\frac{\partial n_{i,t}}{\partial W_{k,t}}\right) d\Psi_{i,t} - \psi\left(\frac{W_{k,t}}{W_{k,t-1}} - 1\right)\frac{1}{W_{k,t-1}} + \\ \frac{1}{1+r_t}\psi\left(\frac{W_{k,t+1}}{W_{k,t}} - 1\right)\frac{W_{k,t+1}}{W_{k,t}^2} = 0.$$

Further, we plug in for expressions (14) and (15):

$$\int\left(u'(c_{i,t})\frac{1}{1+\tau_t^c}(1 - MTR_{i,t})\frac{e_{i,t}}{P_t}(1-\varepsilon)N_t + v'(n_{i,t})\varepsilon\frac{N_{k,t}}{W_{k,t}}\right) d\Psi_{i,t} - \\ \psi\left(\frac{W_{k,t}}{W_{k,t-1}} - 1\right)\frac{1}{W_{k,t-1}} + \frac{1}{1+r_t}\psi\left(\frac{W_{k,t+1}}{W_{k,t}} - 1\right)\frac{W_{k,t+1}}{W_{k,t}^2} = 0.$$

Next, we multiply with $\frac{W_{k,t}}{\psi}$, substitute for $\pi^w = \frac{W_t}{W_{t-1}} - 1$, and, exploiting the fact that in equilibrium $W_{k,t} = W_t$, we obtain:

$$\frac{\varepsilon}{\psi}\left(\int u'(c_{i,t})(1 - MTR_{i,t})w_t e_{i,t}\frac{1-\varepsilon}{\varepsilon}N_t d\Psi_{i,t} + v'(N_t)N_t\right) -$$
$$(1+\pi_t^w)\pi_t^w + \frac{1}{1+r_t}(1+\pi_{t+1}^w)\pi_{t+1}^w = 0.$$

Further, note that

$$(1 - MTR_{i,t})w_t e_{i,t} N_t = (1-\theta)\tau_t\Big(w_t N_t e_{i,t}\Big)^{1-\theta} = (1-\theta)\frac{e_{i,t}^{1-\theta}}{\int e_{i,t}^{1-\theta} di} Z_t,$$

where Z_t is aggregate income tax (net of transfers). Now, we define $\kappa^w = \frac{\varepsilon}{\psi}$, $\mu^w = \frac{\varepsilon}{\varepsilon-1}$, and $u'(\tilde{C}_t) = \frac{e_{i,t}^{1-\theta}}{\int e_{i,t}^{1-\theta} di} u'(c_{i,t})di$, and rearrange to get the final expression for our nonlinear wage NKPC

$$(1+\pi_t^w)\pi_t^w = \kappa^w\left(\gamma N_t^{1+\frac{1}{\phi}} - \frac{(1-\theta)}{(1+\tau_t^c)\mu^w} Z_t u'(\tilde{C}_t)\right) + \frac{1}{1+r_t}(1+\pi_{t+1}^w)\pi_{t+1}^w.$$

C.4 Derivation of the nonlinear price NKPC

Recall the Bellman equation for the intermediate good's firm is:

$$\begin{aligned} J_t(\mathcal{P}_{t-1}, k_{t-1}) = \max_{\mathcal{P}_t, k_t, n_t} & \left\{ \frac{\mathcal{P}_t}{P_t} F(k_{t-1}, n_t) - \frac{W_t}{P_t} n_t - i_t - \phi\left(\frac{k_t}{k_{t-1}}\right) k_{t-1} - \xi(\mathcal{P}_t, \mathcal{P}_{t-1}) Y_t + \right. \\ & \left. \frac{1}{1+r_t} J_{t+1}(\mathcal{P}_t, k_t) \right\}, \\ \text{subject to} \quad & \left(\frac{F(k_{t-1}, n_t)}{Y_t}\right)^{\frac{1-\mu^p}{\mu^p}} Y_t = \left(\frac{\mathcal{P}_t}{P_t}\right) Y_t. \end{aligned}$$

If we denote the Lagrange multiplier on the production constraint with λ_t, the first order condition with respect to n_t is:

$$0 = \frac{\mathcal{P}_t}{P_t} F_{n,t}(k_{t-1}, n_t) - \frac{W_t}{P_t} + \lambda_t \left(\frac{1-\mu_p}{\mu_p}\right)\left(\frac{F(k_{t-1}, n_t)}{Y_t}\right)^{\frac{1-2\mu_p}{\mu_p}} F_{n,t}(k_{t-1}, n_t).$$

Rearranging implies

$$\frac{W_t}{P_t}\frac{1}{F_{n,t}(k_{t-1},n_t)} = mc_t = \frac{\mathcal{P}_t}{P_t} + \lambda_t\left(\frac{1-\mu_p}{\mu_p}\right)\left(\frac{F(k_{t-1},n_t)}{Y_t}\right)^{\frac{1-2\mu_p}{\mu_p}}. \tag{16}$$

Since in equilibrium all firms set the same wage $\mathcal{P}_t = P_t$, and $F(k_{t-1},n_t) = Y_t$, equation (16) simplifies to

$$\frac{W_t}{P_t}\frac{1}{F_{n,t}(k_{t-1},n_t)} = mc_t = 1 - \lambda_t\left(\frac{\mu_p - 1}{\mu_p}\right). \tag{17}$$

Condition (17) has two implications. First, higher Lagrange multiplier λ_t is associated with a lower real marginal cost mc_t, i.e.,

$$\lambda_t = 1 \implies mc_t = \frac{1}{\mu_p} \leq 1,$$

and

$$\lambda_t \to 1 \implies mc_t \to 1.$$

In order to get the price NKPC, we start by taking first order condition with respect to $\mathcal{P}_t$ and get:

$$0 = \frac{1}{P_t}F(k_{t-1},n_t) - \frac{1}{\kappa^p(\mu^p-1)}\left(\frac{\mathcal{P}_t - \mathcal{P}_{t-1}}{\mathcal{P}_{t-1}}\right)\frac{1}{\mathcal{P}_{t-1}}Y_t + \frac{1}{1+r_t}J_{p,t+1}(\mathcal{P}_t,k_t) - \lambda_t\frac{Y_t}{P_t}. \tag{18}$$

The envelope condition implies that the following condition holds:

$$J_{p,t}(\mathcal{P}_{t-1},k_{t-1}) = \frac{1}{\kappa^p(\mu^p-1)}\left(\frac{\mathcal{P}_t - \mathcal{P}_{t-1}}{\mathcal{P}_{t-1}}\right)\frac{\mathcal{P}_t}{\mathcal{P}_{t-1}^2}Y_t. \tag{19}$$

Again, we use the fact that in equilibrium $\mathcal{P}_t = P_t$, and $Y_t = F(k_{t-1},n_t)$. Moreover, by multiplying (18) with $\mathcal{P}_t$, rolling over one period condition (19) and substituting, we get:

$$0 = Y_t - \frac{1}{\kappa^p(\mu^p-1)}\left(\frac{P_t - P_{t-1}}{P_{t-1}}\right)\frac{P_t}{P_{t-1}}Y_t + \frac{1}{1+r_t}\frac{1}{\kappa^p(\mu^p-1)}\left(\frac{P_{t+1} - P_t}{P_t}\right)\frac{P_{t+1}}{P_t}Y_{t+1} - \lambda_t Y_t. \tag{20}$$

Next, we rearrange (20), divide by Y_t, and subsititue for $\pi_t = \frac{P_t}{P_{t-1}} - 1$, to get:

$$1 - \lambda_t = \frac{1}{\kappa^p(\mu^p-1)}(\pi_t)(1+\pi_t) - \frac{1}{1+r_t}\frac{1}{\kappa^p(\mu^p-1)}(\pi_{t+1})(1+\pi_{t+1})\frac{Y_{t+1}}{Y_t}. \tag{21}$$

Rearranging (17)

$$
\begin{aligned}
&mc_t = 1 - \lambda_t\left(\frac{\mu_p - 1}{\mu_p}\right) \implies \lambda_t = (1 - mc_t)\frac{\mu_p}{\mu_p - 1} \implies \\
&1 - \lambda_t = \frac{\mu_p - 1 - \mu_p + \mu_p mc_t}{\mu_p - 1} = \frac{\mu_p mc_t - 1}{\mu_p - 1},
\end{aligned}
$$

and substituting the last expression to (21) yields:

$$
\frac{\mu_p mc_t - 1}{\mu_p - 1} = \frac{1}{\kappa^p(\mu^p - 1)}(\pi_t)(1 + \pi_t) - \frac{1}{1 + r_t}\frac{1}{\kappa^p(\mu^p - 1)}(\pi_{t+1})(1 + \pi_{t+1})\frac{Y_{t+1}}{Y_t}. \tag{22}
$$

Lastly, by multiplying (22) with $\kappa^p(\mu^p - 1)$ and rearranging, we get the final expression for our nonlinear price NKPC:

$$
\pi_t(1 + \pi_t) = \kappa^p(\mu_p mc_t - 1) + \frac{1}{1 + r_t}\pi_{t+1}(1 + \pi_{t+1})\frac{Y_{t+1}}{Y_t}. \tag{23}
$$

Further, we take the first order condition with respect to k_t and get:

$$
0 = -1 - \frac{1}{\delta\varepsilon_I}\left(\frac{k_t}{k_{t-1}} - 1\right)\frac{1}{k_{t-1}} + \frac{1}{1 + r_t}J_{k,t+1}(\mathcal{P}_t, k_t),
$$

which after rearranging

$$
1 + \frac{1}{\delta\varepsilon_I}\left(\frac{k_t}{k_{t-1}} - 1\right) = \frac{1}{1 + r_t}J_{k,t+1}(\mathcal{P}_t, k_t) \equiv Q_t, \tag{24}
$$

and the fact that in the equilibrium $k_t = K_t$, gives us the equation from the text. The envelope condition with respect to k_{t-1} gives us:

$$
\begin{aligned}
J_{k,t}(\mathcal{P}_{t-1}, k_{t-1}) =& \frac{\mathcal{P}_t}{P_t}F_k(k_{t-1}, n_t) + (1 - \delta) - \phi\left(\frac{k_t}{k_{t-1}}\right) + \frac{1}{\delta\varepsilon_I}\left(\frac{k_t}{k_{t-1}} - 1\right)\frac{k_t}{k_{t-1}} + \\
&+ \lambda_t\left(\frac{1 - \mu^p}{\mu^p}\right)\left(\frac{F(k_{t-1}, n_t)}{Y_t}\right)^{\frac{1-2\mu^p}{\mu^p}} Y_t F_k(k_{t-1}, n_t) \\
=& mc_t F_k(k_{t-1}, n_t) + (1 - \delta) - \phi\left(\frac{k_t}{k_{t-1}}\right) + \frac{1}{\delta\varepsilon_I}\left(\frac{k_t}{k_{t-1}} - 1\right)\frac{k_t}{k_{t-1}}.
\end{aligned} \tag{25}
$$

We rearrange (24) and plug it in from both the left- and right-hand side of (25) and get:

$$
(1 + r_{t-1})Q_{t-1} = mc_t F_k(k_{t-1}, n_t) + (1 - \delta) - \phi\left(\frac{k_t}{k_{t-1}}\right) + (Q_t - 1)\frac{k_t}{k_{t-1}},
$$

which we can further simplify, by rearranging and using the Cobb-Douglas property of the production ($F_k(k_{t-1}, n_t) = \alpha \frac{Y_t}{k_{t-1}}$), to:

$$(1 + r_{t-1})Q_{t-1} = mc_t \alpha \frac{Y_t}{k_{t-1}} - \frac{i_t}{k_{t-1}} - \phi\left(\frac{k_t}{k_{t-1}}\right) + Q_t \frac{k_t}{k_{t-1}}.$$

Finally, similar to before, using the fact that in equilibrium $k_t = K_t$, and $n_t = N_t$ yields the final expression from the text.

List of References

Agarwal, S., Driscoll, J. C., Gabaix, X., and Laibson, D. (2007). The Age of Reason: Financial Decisions over the Lifecycle. Working Paper 13191, National Bureau of Economic Research Cambridge, Mass., USA.

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. and Mazumder, B. (2013). Cognitive Abilities and Household Financial Decision Making. *American Economic Journal: Applied Economics*, 5(1):193–207.

Agarwal, S., Rosen, R. J., and Yao, V. (2016). Why do Borrowers make Mortgage Refinancing Mistakes? *Management Science*, 62(12):3494–3509.

Aguiar, V. and Kashaev, N. (2021). Identification and Estimation of Discrete Choice Models with Unobserved Choice Sets. *Available at SSRN 3869963*.

Aguiar, V. H., Boccardi, M. J., Kashaev, N., and Kim, J. (2023). Random Utility and Limited Consideration. *Quantitative Economics*, 14(1):71–116.

Ameriks, J., Caplin, A., and Leahy, J. (2003). Wealth Accumulation and the Propensity to Plan. *The Quarterly Journal of Economics*, 118(3):1007–1047.

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.

Auclert, A., Bardóczy, B., and Rognlie, M. (2023). MPCs, MPEs, and Multipliers: A Trilemma for New Keynesian Models. *Review of Economics and Statistics*, 105(3):700–712.

Auclert, A., Bardóczy, B., Rognlie, M., and Straub, L. (2021). Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models. *Econometrica*, 89(5):2375–2408.

Auclert, A. and Rognlie, M. (2018). Inequality and Aggregate Demand. Working Paper 24280, National Bureau of Economic Research.

Auclert, A., Rognlie, M., and Straub, L. (2018). The Intertemporal Keynesian Cross. Working Paper 25020, National Bureau of Economic Research.

Barseghyan, L., Molinari, F., and Thirkettle, M. (2021). Discrete Choice Under Risk With Limited Consideration. *American Economic Review*, 111(6):1972–2006.

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.

Bayer, C., Born, B., and Luetticke, R. (2023). The Liquidity Channel of Fiscal Policy. *Journal of Monetary Economics*, 134:86–117.

Bezhanova, K., Gapinska, M., Yordanov, Y., Zubimendi Toran, L., and Wahrig, L. (2023). Government Debt Down to 90.3% of GDP in Euro Area. Euro Indicator 119/2023, European Comission.

Bhattacharya, K., Kapoor, A., and Madan, A. (2021). Five Trends Reshaping the US Home Mortgage Industry. Technical report, McKinsey & Company.

Bhutta, N., Blair, J., and Dettling, L. (2022a). The Smart Money is in Cash? Financial Literacy and Liquid Savings Among US Families. *Journal of Accounting and Public Policy*, page 107000.

Bhutta, N., Blair, J., and Dettling, L. (2022b). The Smart Money is in Cash? Financial Literacy and Liquid Savings Among U.S. Families. *Journal of Accounting and Public Policy*, page 107000.

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.

Brandsaas, E. E. (2021). Household Stock Market Participation and Exit: The Role of Homeownership. Technical report, Working Paper.

Broer, T., Krusell, P., and Öberg, E. (2023). Fiscal Multipliers: A Heterogenous-Agent Perspective. *Quantitative Economics*, 14(3):799–816.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2007). Down or Out: Assessing the Welfare Costs of Household Investment Mistakes. *Journal of Political Economy*, 115(5):707–747.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2009a). Fight or Flight? Portfolio Rebalancing by Individual Investors. *The Quarterly Journal of Economics*, 124(1):301–348.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2009b). Measuring the Financial Sophistication of Households. *American Economic Review*, 99(2):393–98.

Calvet, L. E. and Sodini, P. (2014). Twin Picks: Disentangling the Determinants of Risk-Taking in Household Portfolios. *The Journal of Finance*, 69(2):867–906.

Campbell, J. Y. (2006). Household Finance. *The Journal of Finance*, 61(4):1553–1604.

Caplin, A., Dean, M., and Leahy, J. (2019). Rational Inattention, Optimal Consideration Sets, and Stochastic Choice. *The Review of Economic Studies*, 86(3):1061–1094.

Carroll, C., Slacalek, J., Tokuoka, K., and White, M. N. (2017). The Distribution of Wealth and the Marginal Propensity to Consume. *Quantitative Economics*, 8(3):977–1020.

Carroll, C. D. (2006). The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems. *Economics letters*, 91(3):312–320.

Carroll, C. D., Slacalek, J., and Tokuoka, K. (2014). The Distribution of Wealth and the MPC: Implications of New European Data. *American Economic Review*, 104(5):107–111.

Chalmers, J. and Reuter, J. (2020). Is Conflicted Investment Advice Better than No Advice? *Journal of Financial Economics*, 138(2):366–387.

Chetty, R., Friedman, J. N., Leth-Petersen, S., Nielsen, T. H., and Olsen, T. (2014). Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence From Denmark. *The Quarterly Journal of Economics*, 129(3):1141–1219.

Chetty, R., Guren, A., Manoli, D., and Weber, A. (2011). Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins. *American Economic Review*, 101(3):471–475.

Christiano, L., Eichenbaum, M., and Rebelo, S. (2011). When is the Government Spending Multiplier Large? *Journal of Political Economy*, 119(1):78–121.

Coen, J., Kashyap, A. K., and Rostom, M. (2023). Price discrimination and mortgage choice. Working Paper 31652, National Bureau of Economic Research.

Coughlin, M. (2019). Insurance Choice With Non-Monetary Plan Attributes: Limited Consideration in Medicare Part D. Technical report, Cornell University Working Paper.

Crawford, G. S., Griffith, R., and Iaria, A. (2021). A Survey of Preference Estimation With Unobserved Choice Set Heterogeneity. *Journal of Econometrics*, 222(1):4–43.

Damen, S. and Buyst, E. (2017). Mortgage shoppers: how much do they save? *Real Estate Economics*, 45(4):898–929.

Enamorado, T., Fifield, B., and Imai, K. (2019). Using a probabilistic model to assist merging of large-scale administrative records. *American Political Science Review*, 113(2):353–371.

Fagereng, A., Guiso, L., Malacrino, D., and Pistaferri, L. (2020). Heterogeneity and Persistence in Returns to Wealth. *Econometrica*, 88(1):115–170.

Fellegi, I. P. and Sunter, A. B. (1969). A Theory for Record Linkage. *Journal of the American Statistical Association*, 64(328):1183–1210.

Floden, M. and Lindé, J. (2001). Idiosyncratic Risk in the United States and Sweden: Is There a Role for Government Insurance? *Review of Economic dynamics*, 4(2):406–437.

Gennaioli, N., Shleifer, A., and Vishny, R. (2015). Money Doctors. *The Journal of Finance*, 70(1):91–114.

Gerardi, K., Willen, P. S., and Zhang, D. H. (2023). Mortgage Prepayment, Race, and Monetary Policy. *Journal of Financial Economics*, 147(3):498–524.

Gil-Bazo, J. and Imbet, J. F. (2020). Tweeting for Money: Social Media and Mutual Fund Flows. *Available at SSRN 3719169*.

Hagedorn, M., Manovskii, I., and Mitman, K. (2019). The Fiscal Multiplier. Working Paper 25571, National Bureau of Economic Research.

Heathcote, J., Storesletten, K., and Violante, G. L. (2017). Optimal Tax Progressivity: An Analytical Framework. *The Quarterly Journal of Economics*, 132(4):1693–1754.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica: Journal of the Econometric Society*, pages 153–161.

Jappelli, T. and Padula, M. (2017). Consumption Growth, the Interest Rate, and Financial Sophistication. *Journal of Pension Economics and Finance*, 16(3):348–370.

Jappelli, T. and Pistaferri, L. (2014). Fiscal Policy and MPC Heterogeneity. *American Economic Journal: Macroeconomics*, 6(4):107–136.

Johnson, D. S., Parker, J. A., and Souleles, N. S. (2006). Household Expenditure and the Income Tax Rebates of 2001. *American Economic Review*, 96(5):1589–1610.

Jung, J., Kim, J. H., Matějka, F., and Sims, C. A. (2019). Discrete Actions in Information-Constrained Decision Problems. *The Review of Economic Studies*, 86(6):2643–2667.

Kacperczyk, M., Nosal, J., and Stevens, L. (2019). Investor Sophistication and Capital Income Inequality. *Journal of Monetary Economics*, 107:18–31.

Kaplan, G., Moll, B., and Violante, G. L. (2018). Monetary Policy According to HANK. *American Economic Review*, 108(3):697–743.

Kaplan, G. and Violante, G. L. (2014). A Model of the Consumption Response to Fiscal Stimulus Payments. *Econometrica*, 82(4):1199–1239.

Kaplan, G. and Violante, G. L. (2018). Microeconomic Heterogeneity and Macroeconomic Shocks. *Journal of Economic Perspectives*, 32(3):167–194.

Kaplan, G. and Violante, G. L. (2022). The Marginal Propensity to Consume in Heterogeneous Agent Models. *Annual Review of Economics*, 14:747–775.

Kaplan, G., Violante, G. L., and Weidner, J. (2014). The Wealthy Hand-to-Mouth. *Brookings Papers on Economic Activity*, 1:77–153.

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

Kline, P. and Walters, C. R. (2019). On Heckits, LATE, and Numerical Equivalence. *Econometrica*, 87(2):677–696.

Luetticke, R. (2021). Transmission of Monetary Policy with Heterogeneity in Household Portfolios. *American Economic Journal: Macroeconomics*, 13(2):1–25.

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., Michaud, P.-C., and Mitchell, O. S. (2017). Optimal Financial Knowledge and Wealth Inequality. *Journal of Political Economy*, 125(2):431–477.

Lusardi, A., Michaud, P.-C., and Mitchell, O. S. (2020). Assessing the Impact of Financial Education Programs: A Quantitative Model. *Economics of Education Review*, 78:101899.

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.

Mani, A., Mullainathan, S., Shafir, E., and Zhao, J. (2013). Poverty Impedes Cognitive Function. *Science*, 341(6149):976–980.

Manzini, P. and Mariotti, M. (2014). Stochastic Choice and Consideration Sets. *Econometrica*, 82(3):1153–1176.

McFadden, D. and Train, K. (2000). Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics*, 15(5):447–470.

Mudzingiri, C. (2021). The Impact of Financial Literacy on Risk Seeking and Patient Attitudes of University Students. *Development Southern Africa*, 38(5):845–861.

Mullainathan, S., Noeth, M., and Schoar, A. (2012). The Market for Financial Advice: An Audit Study. Working Paper 17929, National Bureau of Economic Research.

Mullainathan, S. and Shleifer, A. (2005). Persuasion in Finance. Working Paper 11838, National Bureau of Economic Research.

Nieddu, M. and Pandolfi, L. (2021). Cutting Through the Fog: Financial Literacy and Financial Investment Choices. *Journal of the European Economic Association*, 19(1):237–274.

Qiu, X. and Russo, N. (2023). Income Taxation: A Cross-Country Comparison. Mimeo.

Rabin, M. (2013). Risk Aversion and Expected-utility Theory: A Calibration Theorem. In *Handbook of the Fundamentals of Financial Decision Making: Part I*, pages 241–252. World Scientific.

Rotemberg, J. J. (1982). Monopolistic Price Adjustment and Aggregate Output. *The Review of Economic Studies*, 49(4):517–531.

Slacalek, J., Tristani, O., and Violante, G. L. (2020). Household Balance Sheet Channels of Monetary Policy: A Back of the Envelope Calculation for the Euro Area. *Journal of Economic Dynamics and Control*, 115:103879.

Sutter, M., Weyland, M., Untertrifaller, A., and Froitzheim, M. (2020). Financial Literacy, Risk and Time Preferences – Results From a Randomized Educational Intervention. Discussion Paper 13566, IZA.

Trabandt, M. and Uhlig, H. (2011). The Laffer Curve Revisited. *Journal of Monetary Economics*, 58(4):305–327.

Van Rooij, M., Lusardi, A., and Alessie, R. (2011). Financial Literacy and Stock Market Participation. *Journal of Financial Economics*, 101(2):449–472.

Vuong, Q. H. (1989). Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica: Journal of the Econometric Society*, pages 307–333.

Zhou, X. (2022). FinTech Lending, Social Networks and the Transmission of Monetary Policy. FRB of Dallas Working Paper 2203, FRB of Dallas.

List of Appendices

List of Figures

List of Tables