Title
Behavioral Responses of Household Financial Decision-Making

ISBN
9798189270420

Author
Lyu, Junru

Publication Date
2026-05-01

Peer reviewed|Thesis/dissertation

Behavioral Responses of Household Financial Decision-Making

by

Jun Ru Lyu

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Stefano DellaVigna, Chair
Professor Ulrike Malmendier
Professor Dmitry Taubinsky

Spring 2026

Abstract

Behavioral Responses of Household Financial Decision-Making

by

Jun Ru Lyu

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Stefano DellaVigna, Chair

This dissertation studies how behavioral forces shape household financial decision-making. Across three chapters, we combine administrative credit-panel data, structural modeling, and a controlled laboratory experiment to show that departures from the frictionless, fully-informed benchmark of standard models are not peripheral: they govern how households respond to debt relief, how they form beliefs about asset returns, and how they invest in financial knowledge. Taken together, the chapters argue that the behavioral channel is central to understanding, and designing policy around, household finance.

The first chapter examines the long-run consequences of temporary debt relief, using the federal student loan repayment pause from 2020 to 2023 as a natural experiment. Exploiting variation in loan eligibility through an instrumental variables strategy applied to the University of California Consumer Credit Panel (UC-CCP), we show that the pause initially reduced delinquency and raised consumption—but that the consumption increase persisted and grew gradually, consistent not with a neoclassical liquidity response but with the slow formation of new spending habits. When payments resumed, delinquency rose sharply across credit categories. A life-cycle model with adaptive reference-dependent consumption reproduces these dynamics and, in counterfactual simulations, shows that a shorter pause would have curtailed habit formation and reduced post-pause delinquency. The chapter establishes a recurring theme of the dissertation: temporary changes in household circumstances can leave durable behavioral imprints.

The second chapter shifts from debt management to homeownership, asking how lived experiences of local housing markets shape the decision to buy a home. We develop a life-cycle model in which households form beliefs about both the mean and the volatility of future housing returns from their personal histories, and use the UC-CCP linked to zip code-level Zillow price data to construct individual-level measures of experienced returns and volatility since age 18. Consistent with the model, a one percentage point increase in experienced returns raises the probability of transitioning into homeownership by roughly seven percent, while a one percentage point increase in experienced volatility lowers it by roughly ten percent.

Volatility effects are detectable only at the zip-code level and attenuate with age, underscoring that perceptions of risk are formed at a granular, local scale.

The third chapter turns to the acquisition of financial knowledge itself. Combining administrative data on UC Berkeley students enrolled in a large personal-finance course with an incentivized online experiment, we document strong positive selection into financial literacy education on baseline ability, family income, and quantitative preparation. A controlled lab experiment identifies the mechanism: low-ability individuals systematically overestimate their starting knowledge and underestimate the benefits of education, producing demand that tracks perceived rather than actual returns. Varying the cognitive cost of the material through AI-based, video, and lecture formats raises aggregate demand but does not repair the selection pattern, suggesting that the binding constraint is belief miscalibration rather than access or difficulty.

A common thread runs through the three chapters. Households do not move smoothly along the optimal paths that standard models prescribe: they form habits from temporary windfalls, extrapolate from the local markets they happen to have lived in, and misjudge the returns to knowledge they do not yet have. Each of these departures carries first-order implications for debt policy, housing policy, and the design of financial education. The chapters share a common empirical backbone—linked administrative credit-panel data and a careful mapping from behavioral models to causal estimates—and together they make the case that behavioral mechanisms deserve a central place in the analysis of household finance.

Acknowledgements

This dissertation would not exist without the guidance of my advisors. Stefano DellaVigna has shaped my thinking since my undergraduate years at Berkeley, and his steady, generous mentorship has been the single greatest intellectual influence on the work that follows. Ulrike Malmendier has likewise supported me since I was an undergraduate, and the second chapter of this dissertation carries the clear imprint of her work on experience effects and her patient encouragement to pursue the questions that most interested me. Ricardo Perez-Truglia, in addition to serving as a letter writer and offering sharp feedback on every project here, has been an extraordinary mentor and friend—on the research side and on the soccer pitch, where we have played together throughout his time at Berkeley. Peter Maxted rounds out my letter-writing committee, and I am grateful for his careful reading and unwavering support.

Beyond my formal committee, I owe deep thanks to Dmitry Taubinsky, Ned Augenblick, and Gautam Rao. Their office hours and the Berkeley behavioral economics therapy sessions were the forge in which many of the ideas in this dissertation were tested and sharpened. Their willingness to engage seriously with early-stage work, and their honesty about what was and was not working, has made me a better economist.

I have been extraordinarily lucky in my collaborators and friends. Gene Kang, Elaine Shen, Laila Voss, Xinmei Yang, and Lucio Wasserman have been coauthors, critics, and companions throughout graduate school. The second chapter of this dissertation is joint work with Gene Kang and Sharath Sonti, and the third chapter is joint work with Elaine Shen and Laila Voss; I am grateful to all of them for allowing these papers to appear here and for the countless hours of thinking and writing that went into them.

Sharath Sonti deserves a paragraph of his own. He has been one of the great friends of my graduate school years—a coauthor on the experience-effects work, a partner in many long discussions about economics and everything else, and a steady companion through innumerable video game sessions that I will remember long after I have forgotten the details of any regression.

Soccer has been, throughout, the thing that kept me sane. I am grateful to the cohort of faculty and students who made that possible: Ricardo Perez-Truglia, Jon Kolstad, and Michael Jansson on the faculty side, and Lucio Wasserman, Sharath Sonti, Gene Kang, Sydney Costantini, Elaine Shen, and many others on the student side. The pickup games, the intramural matches, and the after-game conversations have been as much a part of this PhD as the seminars.

To Leo Song, now finishing his own PhD in biology at Berkeley: we have been friends for thirteen years, from the same high school to the same undergraduate campus to the same graduate one. Very few people get to walk the same long road alongside the same friend, and I count myself fortunate to be one of them. Thank you for every conversation, every meal, and every reminder that there is a life outside of research.

To Sijia Zhang, my girlfriend: thank you for your patience through the late nights, the canceled plans, the half-present dinners where I was still thinking about a regression. You have made the hard parts bearable and the good parts immeasurably better. I am grateful beyond what I can write here.

To my parents, Yanzheng Lyu and Liping Sun: everything I have I owe to you. You believed in this path long before there was any evidence it would lead anywhere, and you have supported me across every distance and every setback with a steadiness I can only hope to one day return in kind. This dissertation is, in the most meaningful sense, as much yours as it is mine.

Finally, to everyone I have not named here but who has listened, read drafts, sent encouragement, or simply believed this would eventually get done—thank you. Any errors that remain in the pages that follow are, of course, my own.

Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Motivation: A Behavioral View of Household Financial Decision-Making

Household financial decision-making is a foundational domain in economics. The choices households make about how much to consume, how much to save, when to take on debt, when to buy a home, and how much to invest in their own financial knowledge collectively shape both individual welfare and aggregate economic outcomes. The standard framework for studying these decisions assumes households are forward-looking, fully informed about their financial environment, and able to translate that information into optimal choices. This benchmark has produced a rich and useful body of theory; it also generates predictions that, in many empirical settings, fail to fit the data.

Where standard models fall short, behavioral economics has supplied a growing catalog of alternatives. Households extrapolate from recent experience rather than from full distributions [Malmendier and Nagel, 2011, Kuchler and Zafar, 2019]; they form habits over consumption that anchor future decisions [Chetty and Szeidl, 2016]; they hold systematically biased beliefs about returns to investment in their own knowledge [Lusardi et al., 2017, Ambuehl et al., 2022]. These departures from the frictionless benchmark are not noise. They are systematic, predictable, and often first-order: they govern when households respond to policy interventions, how they perceive the risks they face, and whether they invest in the information that would help them choose well.

This dissertation studies three such departures, each in a different domain of household finance, and asks what each implies for the design of policy. The three chapters share a common framing: each begins with a clean neoclassical benchmark, documents a specific behavioral deviation from that benchmark using administrative data and, in one case, a controlled experiment, and traces the deviation to a measurable consequence for household outcomes. Chapter 2 examines how households respond when the structure of their debt obligations changes temporarily, using the federal student loan repayment pause as a natural experiment. Chapter 3 examines how households' lived experiences of local housing markets shape their decisions to enter homeownership. Chapter 4 examines how households decide whether to invest in financial knowledge in the first place. Together, the three chapters argue that behavioral mechanisms are central to understanding, and to designing policy around, household financial decision-making.

1.2 Preview of the Chapters

Chapter 2: Habit Formation and the Hidden Costs of Temporary Debt Relief

The first chapter studies the long-run consequences of a large, well-defined episode of temporary debt relief: the federal student loan repayment pause in effect from March 2020 to September 2023. During this period, monthly payments, interest accrual, and collections were suspended for tens of millions of federal borrowers, while underlying debt balances were left unchanged. The pause therefore offers an unusually clean window onto how households respond to a temporary expansion of liquidity that does not reduce permanent indebtedness.

A neoclassical model of consumption smoothing yields sharp predictions in this setting: relieved households should raise consumption during the pause and reduce delinquency mechanically, and both responses should reverse cleanly when payments resume. Using the University of California Consumer Credit Panel (UC-CCP) and an instrumental variables strategy that exploits cross-borrower variation in the share of loans eligible for the pause, we test these predictions. The early-period results are consistent with the neoclassical benchmark and with prior work studying the same episode [Dinerstein et al., 2024]. As the pause continued, however, two patterns emerged that the standard model cannot accommodate. First, consumption rose only gradually rather than immediately. Second, delinquency, which initially declined, began to rise during the later quarters of the pause and rose sharply once payments resumed.

We rationalize these dynamics with a life-cycle model in which households form an adaptive reference point over their own consumption. Temporary liquidity raises consumption, which raises the reference point, which in turn supports further increases in consumption; once payments resume, the elevated reference point makes the required cutbacks costly enough that some borrowers default rather than absorb them. Counterfactual simulations show that ending the pause earlier—before habits had time to solidify—would have substantially reduced post-pause delinquency without sacrificing the bulk of the welfare gains from the relief itself. The chapter establishes a recurring theme of the dissertation: temporary changes in household circumstances can leave durable behavioral imprints that shape outcomes long after the original change has reversed.

Chapter 3: Local Housing Market Experiences and Homeownership Decisions

The second chapter, joint with Gene Kang and Sharath Sonti, turns from debt management to homeownership and asks how households' lived experiences of local housing markets shape their decisions to buy a home. Stocks and bonds are traded in essentially national markets, but housing returns vary sharply across locations. This makes housing an unusually well-suited domain in which to study how local, personal histories shape long-horizon financial decisions.

We develop a life-cycle model in which households derive utility from consumption and housing services, choose each period whether to rent or own, and form beliefs about both the mean and the volatility of future house-price returns from the price histories they have personally observed since age 18. The model yields two clear comparative statics: higher subjective expected returns raise the probability of becoming a homeowner, while higher subjective volatility lowers it.

To test these predictions, we again turn to the UC-CCP, which links individuals' residential zip codes to their credit records over time. By merging this with zip-code-level price data from Zillow, we construct, for each individual in our panel, a recency-weighted history of returns and a corresponding history of volatility experienced since age 18. The empirical results confirm both predictions: a one percentage point increase in experienced returns raises the probability of transitioning into homeownership by roughly seven percent relative to the sample's quarterly transition rate, while a one percentage point increase in experienced volatility lowers it by roughly ten percent. These effects are robust to a comprehensive set of controls and fixed effects. The volatility effect attenuates with age, consistent with the life-cycle structure of the model, and is detectable only when experiences are constructed at the zip-code level—a finding that underscores how granular, local experiences are the operative input to belief formation in this market.

The chapter contributes to a literature that has documented how lived macroeconomic experiences shape household financial behavior [Malmendier and Nagel, 2011, Malmendier and Wellsjo, 2024, Bailey et al., 2018, Agarwal et al., 2015]. Its distinctive contribution is to bring both the mean and the volatility of experienced returns into a unified framework, to estimate both effects at the zip-code level, and to do so on a nationally representative panel that lets us track each individual's actual residential history.

Chapter 4: Self-Selection in Financial Literacy Education

The third chapter, joint with Elaine Shen and Laila Voss, takes a step back. The first two chapters take households' beliefs and knowledge about their financial environment as given. A natural prior question is how households come to acquire that knowledge in the first place. Standard human capital theory predicts that those with the most to learn—and therefore the highest marginal returns to education—should be the most likely to invest. Our findings challenge this prediction.

We study self-selection into financial literacy education using two complementary settings. The first is observational data on undergraduate enrollment in a large, highly rated personal finance course at UC Berkeley (UGBA 135). Combining administrative records from the Berkeley Registrar's Office with a financial literacy survey administered through the Associated Students of the University of California, we document strong positive selection on observables: students who enroll in the course are disproportionately drawn from quantitative majors, from families with higher income, and from groups already well-represented at the university. Hypothetical willingness to pay for the course is, if anything, higher among students who already perform better on financial decision-making questions.

To isolate the mechanisms behind these patterns, we move to a controlled online experiment with over a thousand U.S. participants on Prolific. Participants complete an incentivized baseline task on compound interest, are given the opportunity to purchase one of three versions of an educational module—an AI-based learning assistant, a video lesson, and a text-based lecture—and complete an incentivized endline. By eliciting beliefs about their own baseline performance and their expected gains from each version of the material, we can directly measure the gap between perceived and actual returns to education.

The experiment delivers two main findings. First, the positive correlation between baseline ability and willingness to pay replicates in the lab: the participants who would gain the most from financial education are the least willing to pay for it. Second, this misalignment is driven by systematic belief errors. Low-ability participants overestimate their starting knowledge and underestimate their potential learning gains; the result is that demand for education tracks perceived returns rather than actual returns. Varying the cognitive cost of the material—from text-heavy lecture to interactive AI—raises aggregate demand but does not change the selection pattern. The binding constraint is not access or difficulty; it is belief miscalibration. The chapter's policy implication is that supply-side interventions designed to make education easier or cheaper will, on their own, do little to reach the households who would benefit most.

1.3 A Common Methodological Backbone

Although the three chapters span debt, housing, and human capital, they share a methodological structure. Each begins with a clean theoretical benchmark—a neoclassical model in Chapter 2, a life-cycle model with experience-based learning in Chapter 3, and standard human capital theory in Chapter 4—and uses that benchmark to derive sharp empirical predictions. Each then takes those predictions to data drawn from large administrative or experimental sources. Chapters 2 and 3 both rely on the UC-CCP, a panel that combines comprehensive coverage of California residents' credit records with a nationally representative two percent sample of U.S. adults. Chapter 4 pairs administrative records from a large university with a controlled online experiment that lets us measure beliefs and treatment effects directly.

This combination of theory and administrative microdata has two virtues for the questions we ask. The administrative panels give us the geographic granularity, the time horizon, and the sample size needed to construct individual-level histories of experience and to detect the gradual dynamics that distinguish behavioral from neoclassical responses. The structural and reduced-form models give us a disciplined map between behavioral mechanisms and the moments of the data we use to identify them. In each chapter, the main behavioral mechanism is identified through variation that the standard model and the behavioral alternative would predict to look different: an instrument exploiting variation in policy eligibility in Chapter 2, cross-individual variation in residential histories in Chapter 3, and randomized assignment to an educational treatment in Chapter 4.

1.4 Relation to the Literature

Each chapter situates itself within its own substantive literature, and the chapter introductions provide more detailed surveys of that work. Three broader literatures, however, run through the dissertation as a whole.

The first is the broader behavioral household finance tradition, which has documented systematic departures from the standard frictionless benchmark in consumption smoothing, asset returns, and the role of information frictions in financial decision-making. The dissertation contributes to this tradition by documenting three new behavioral patterns—habit-driven debt dynamics, experience-based housing choice, and belief-driven adverse selection into financial education—using administrative data on a scale that has only recently become available.

The second is the literature on experience effects pioneered by Malmendier and Nagel

[2011] and extended by, among others, Kuchler and Zafar [2019], Malmendier et al. [2020], and Malmendier and Wellsjo [2024]. Chapter 3 contributes most directly to this work by extending the experience-effects framework from returns alone to returns and volatility, and by exploiting fine geographic variation that the experience-effects literature has typically lacked. Chapter 2 relates to this tradition through a different channel: the temporary student loan pause is itself an experience that, our model suggests, durably alters the consumption norms households carry forward.

The third is the literature on consumer credit and financial fragility [Ganong and Noel, 2020, Agarwal et al., 2007, Parker et al., 2013, Dinerstein et al., 2024], which Chapter 2 engages most directly. Our finding that temporary forbearance can produce durable post-policy delinquency through habit formation contributes to this literature a behavioral mechanism that purely liquidity-based accounts miss.

1.5 Roadmap

The remainder of the dissertation proceeds as follows. Chapter 2 studies the federal student loan pause and the role of habit formation in shaping post-pause delinquency. Chapter 3 examines how local housing market experiences shape entry into homeownership. Chapter 4 studies self-selection into financial literacy education and the belief errors that drive it. Chapter 5 concludes by drawing together the three chapters and discussing what they collectively imply for the design of household-finance policy.

Chapter 2 Deferred Burden: Habit Formation and the Hidden Costs of Temporary Debt Relief

2.1 Introduction

This chapter opens our study of behavioral responses in household financial decision-making with a question at the heart of consumer finance: how do households respond when the structure of their debt obligations changes? Debt management is a central feature of household financial decision-making. From mortgages and auto loans to credit cards, households rely on debt to finance both large, infrequent purchases and everyday consumption. As a result, understanding how households manage debt is essential to understanding household finances more broadly. Changes in debt obligations—through the accumulation or relief of debt—directly affect required monthly payments and, in turn, household liquidity. For this reason, many fiscal policies seek to influence consumption and enhance financial stability by altering household liquidity through debt-based interventions[Dobbie and Song, 2015, Agarwal et al., 2007].

Standard economic models predict that temporary relief from debt obligations should raise short-run utility without significant long-term costs [Zeldes, 1989]. Yet the behavioral literature suggests a more complex reality: liquidity may alter spending habits, reference points, or mental accounting in ways that persist beyond the relief period [Chetty and Szeidl, 2016]. If so, policies designed to stabilize borrowers in the short run could inadvertently undermine their long-run financial health.

A central challenge in studying the long-run effects of temporary debt relief is the absence of a clean policy environment: most programs either lack well-defined start and end dates or coincide with major macroeconomic shocks, making it difficult to isolate liquidity expansions from confounding factors such as unemployment, income volatility, or fiscal transfers. Many interventions also bundle interest subsidies, refinancing, or partial forgiveness, blurring the line between temporary relief and permanent changes in indebtedness. Standard comparisons between borrowers and non-borrowers are further complicated by selection, since student loan holders differ systematically in age, education, and financial status. Researchers must therefore construct credible counterfactuals that account for heterogeneity, and link credit data to models of intertemporal choice. Without these methodological and conceptual advances, the long-term consequences of temporary liquidity relief remain difficult to identify.

This paper exploits the federal student loan repayment pause during the COVID-19 pandemic as a natural experiment to study the long-run effects of temporary debt relief. From March 2020 to September 2023, repayments, interest accrual, and collections were suspended for over 40 million federal borrowers—nearly $200 billion in annual payments. Unlike forgiveness, the pause expanded liquidity without reducing underlying debt, creating a sharp, time-bound expansion and contraction of household budgets. A neoclassical model with consumption smoothing and delinquency costs predicts higher consumption and lower delinquency during the pause, followed by full reversion once payments resume. Empirically, however, we find that while delinquency initially declined, sustained increases in consumption led to rising delinquency among eligible borrowers after repayment restarted. A habit-formation framework with adaptive reference-dependent consumption explains these dynamics: temporary liquidity raised spending norms that proved costly once obligations returned. Counterfactual simulations show that ending the pause earlier—before habits solidified—would have reduced post-pause delinquency and improved welfare.

We begin with a neoclassical model in which households choose consumption and repayment subject to an intertemporal budget constraint. Utility depends only on contemporaneous consumption, while delinquency entails a pecuniary penalty. The implication of this framework is generally applicable to all agents, provided they are forward-looking, regardless of their degree of patience: When payments are suspended, households increase consumption as required payments decline, while delinquency falls mechanically. Once repayment resumes, delinquency should return to pre-pause levels, and consumption should fall back accordingly. Thus, the neoclassical model predicts a temporary consumption boost during the pause, coupled with lower delinquency, but no lasting effect after repayment obligations are reinstated.

To test these predictions, we use panel data from the University of California Consumer Credit Panel (UC-CCP), a nationally representative 2% sample of U.S. adults drawn from anonymized credit bureau records. Individuals are randomly selected such that we observe roughly 100,000 individuals per quarter. We use quarterly information from 2018 onward on all loans held by each individual, including balances, payments, and delinquency status.

This comprehensive coverage allows us to observe not only student loans but also mortgages, auto loans, auto leases, credit cards, retail credit cards, utilities, and miscellaneous credit lines. These categories form the basis of our key outcome variables. In our regressions, we aggregate these raw categories into six outcomes of interest: credit card, mortgage, auto loan, the sum of all debts, the sum of large-balance debts (mortgage and auto loan), and the sum of daily-use debts (credit card, retail credit, miscellaneous, and utilities). We measure *consumption* using payments and balances on revolving credit accounts, which provide a high-frequency proxy for household spending. *Financial strain* is captured through delinquency indicators across loan types, defined as any account that is past due. Together, these measures allow us to track how temporary relief in student loan obligations affected broader debt repayment, consumption behavior, and delinquency patterns across the credit portfolio.

We use an instrumental variables strategy by constructing an instrument based on the share of a borrower's student loan portfolio that was federally eligible—the "paused ratio." This instrument captures exogenous variation in payment relief due to policy eligibility, enabling us to isolate causal effects of reduced payment obligations on other credit outcomes. Together, these approaches allow us to overcome the central challenges of confounding macroeconomic shocks and heterogeneous borrower characteristics.

The empirical results show that the student loan pause initially improved financial health in line with neoclassical predictions, consistent with the early-period findings of Dinerstein et al. [2024], who study the impacts of this policy from 2019 to 2022. Borrowers used freed-up liquidity to sustain or increase consumption and, at the same time, reduced delinquencies across mortgages, auto loans, and credit cards. In the early quarters of the pause, these patterns suggest that temporary relief provided meaningful short-run stabilization by lowering financial strain and helping households smooth consumption during an uncertain period. Over time, however, the behavioral dynamics shifted. Consumption—proxied by revolving balances and payments—continued to rise steadily, reflecting the formation of new spending habits. As repayment obligations resumed in January 2024, delinquency rates spiked across nearly all forms of debt.

To capture the dynamics observed in the data, we introduce an adaptive reference point for consumption. Households form habits based on recent consumption, and utility losses arise when realized consumption falls below this reference point. With this modification, the model predicts that some borrowers will choose to be delinquent even during the pause in order to sustain consumption above the evolving reference point. Once repayment resumes, the heightened reference level amplifies financial strain, leading to persistently higher delinquency rates than before the pause. This structural model highlights their welfare consequences: sustained relief encouraged spending habits that proved costly once liquidity contracted. Counterfactual simulations demonstrate that ending the pause earlier would have limited habit formation, reduced post-pause delinquency, and delivered higher long-run welfare.

This chapter is related to a growing literature on student debt relief and liquidity constraints. Dinerstein et al. [2024] study the federal student loan pause and find large improvements in consumption but limited effects on financial distress. The chapter extends on some results of these authors by examining the policy for the entire duration and post-policy implications. Boutros et al. [2024] provide a structural analysis of repayment plans, showing how repayment flexibility shapes welfare and behavior, while Mezza et al. [2020] document that student debt depresses homeownership. Consistent with this work, we also find a consumption increase during the pause. However, because we observe data over a longer horizon, we uncover a sharp post-pause rise in delinquency—highlighting risks of habit formation and repayment resumption. Our study builds on this evidence by showing how repayment suspension policies affect long-term financial trajectories through behavioral channels.

This work also connects to research on household responses to temporary income changes. Studies of tax rebates [Agarwal et al., 2007, Johnson et al., 2006, Parker et al., 2013] show that liquidity-constrained households raise consumption significantly in response to transitory windfalls, while Ganong and Noel [2020] demonstrate that income timing shocks can cause financial strain. These findings parallel the patterns observed during the student loan pause: households benefited from immediate liquidity but were left vulnerable once the relief expired.

Finally, the chapter contributes to the behavioral economics and financial fragility literatures. Models of habit formation and reference-dependent preferences [Chetty and Szeidl, 2016] explain why temporary relief can generate persistent changes in spending behavior. Evidence from consumer credit markets [Skiba and Tobacman, 2014, Dobbie et al., 2017, Livshits et al., 2007] highlights how short-term liquidity shocks can translate into lasting delinquency and default risk. By combining rich credit bureau data with a structural model of habit formation, we show that temporary debt relief can unintentionally erode repayment capacity, underscoring the importance of incorporating behavioral dynamics into the design of forbearance policies.

Overall, the chapter shows that the welfare impact of loan forbearance programs hinges not only on their immediate effects but also on how they shape borrower expectations, habits, and financial discipline. In particular, extended pauses–while politically appealing and administratively simple–may inadvertently erode repayment capacity and financial health over time. The pattern identified here—that temporary changes in household circumstances can leave durable imprints on behavior—reappears in the next chapter in a different domain, where we examine how households' lived experiences of local housing markets shape their decision to become homeowners.

The remainder of the chapter is organized as follows. Section 2.2 provides background on the federal student loan pause. Section 2.3 develops a simple model to predict changes in consumption and delinquency during and after the pause. Section 2.4 describes the data and outlines the empirical strategy used for estimation. Section 2.5 presents the main empirical findings on the impact of the pause. Section 2.6 discusses potential behavioral mechanisms underlying the results and evaluates alternative policy designs. Section 2.7 concludes.

2.2 Institutional Background

In response to the economic and public health crisis caused by the COVID-19 pandemic, the U.S. federal government implemented an unprecedented pause on federal student loan payments. Beginning March 13, 2020, and initially authorized through the CARES Act, this policy suspended monthly payments, set interest rates to zero, and halted collections on defaulted loans. These provisions were extended through executive actions under both the Trump and Biden administrations, ultimately lasting through September 2023.[1]

The pause automatically applied to all loans *owned by the U.S. Department of Education*. This includes:

- All Direct Loans;
- Some Federal Family Education Loan (FFEL) and Perkins Loans that had been consolidated or transferred into federal ownership.

By contrast, loans not federally owned—including most legacy FFEL loans, institutionally held Perkins loans, and all private student loans—were ineligible.[2]

Eligibility Criteria and Borrower Selection

Eligibility for pause benefits was based entirely on the *administrative ownership* of a borrower's loan at the time of policy implementation. Borrowers could not easily select into or out of eligible loan types in response to the policy. For example:

- Two borrowers attending similar institutions with similar financial need could end up in different programs depending on whether they applied before or after the transition from FFEL to Direct Loans;
- Eligibility for Direct versus FFEL loans was often determined by institutional participation and timing, not borrower choice;
- Most borrowers were unaware of the ownership status of their loans, and consolidation into eligible loans was logistically burdensome and rare.

Ideally, differences between borrowers holding Direct versus Indirect Federal Education Loans (FFEL) can be viewed as plausibly exogenous, as eligibility for the federal pause was determined by loan type rather than borrower choice. However, borrowers with private loans may differ systematically from federal borrowers in demographic and financial characteristics. To mitigate these concerns, we show event-study plots tracing outcomes over the pre- and

[1]See CARES Act, Public Law 116–136, Sec. 3513; and U.S. Department of Education (2022), "Student Loan Payment Pause Update."

[1]See Federal Student Aid, "COVID-19 Emergency Relief and Federal Student Aid," `https://studentaid.gov/announcements-events/covid-19`.

post-pause periods in Section 2.5 and include rich controls for observable characteristics such as age, education, homeownership, and baseline debt structure. This design aims to isolate the causal effect of pause eligibility on financial behavior while accounting for residual heterogeneity.

2.3 Model

In this section, we develop a model that captures the borrowing and saving constraints faced by individual agents. Each quarter, agents are allowed to make decisions about whether to save or borrow based on their financial circumstances. However, these decisions are subject to a minimum required payment. This payment represents the threshold needed to avoid more severe financial consequences, such as incurring late fees, triggering delinquencies, or experiencing a deterioration in credit scores. By incorporating this constraint, the model aims to reflect the trade-offs and frictions that shape household financial behavior over time.

We model an agent who lives for T discrete periods, indexed by $t = 0, 1, \ldots, T$. In each period, the agent earns income w_t, chooses consumption c_t, and manages both debt $B_t \geq 0$ and savings $S_t \geq 0$. Debt accrues interest at rate r_b, and savings accrue at rate r_s, where $r_b > r_s$, reflecting institutional borrowing frictions.

Budget Constraints and Repayment

The agent may borrow additional amount $b_t \geq 0$, save $s_t \geq 0$, and withdraw savings $x_t \in [0, S_t]$. The minimum required repayment is mB_t, where $m \in (0, 1)$, and actual repayment p_t may fall below this, incurring delinquency. The constraints are:

$$c_t + p_t + s_t = w_t + x_t + b_t, \tag{2.1}$$

$$d_t = \mathbf{1}_{p_t < mB_t}, \tag{2.2}$$

$$B_{t+1} = (1 + r_b)(B_t - p_t + b_t), \tag{2.3}$$

$$S_{t+1} = (1 + r_s)(S_t - x_t + s_t), \tag{2.4}$$

$$c_T + B_T = w_T + S_T. \tag{2.5}$$

Utility and Objective

The agent derives utility from consumption and incurs disutility from delinquency. The period utility is:

$$U(c_t, d_t) = u(c_t) - kd_t, \tag{2.6}$$

where d_t is the delinquency indicator and $k > 0$ is the penalty. Total utility over the lifetime is:

$$U = \sum_{t=0}^{T-1} \delta^t \left[u(c_t) - k\mathbf{1}_{p_t < mB_t}\right] + \delta^T u(w_T + S_T - B_T). \tag{2.7}$$

Remark 2.1. The agent never simultaneously saves and borrows. When both $B_t > 0$ and $S_t > 0$, the agent reallocates savings to pay debt.

Because the borrowing rate r_b exceeds the saving rate r_s, individuals optimally either save or borrow, but never do both simultaneously.

Delinquency and Payment Thresholds

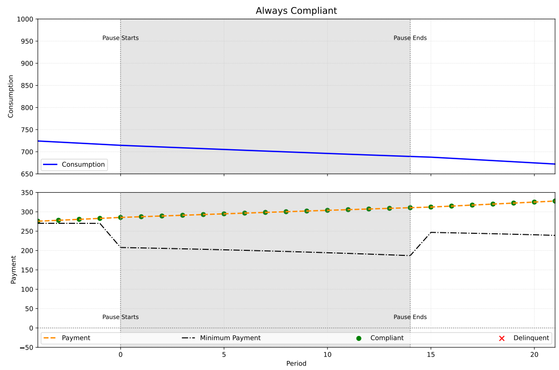
Let us define the optimal repayment in absence of delinquency penalty via the Euler equation:

$$\delta^s(1 + r_b)^s = \frac{u'(c_t)}{u'(c_{t+s})}, \quad c_t^* = f\Big(w_t, w_{t+1}, \ldots, w_T,\, r_b,\, \delta,\, B_t\Big), \qquad p_t^* = w_t - c_t^*, \tag{2.8}$$

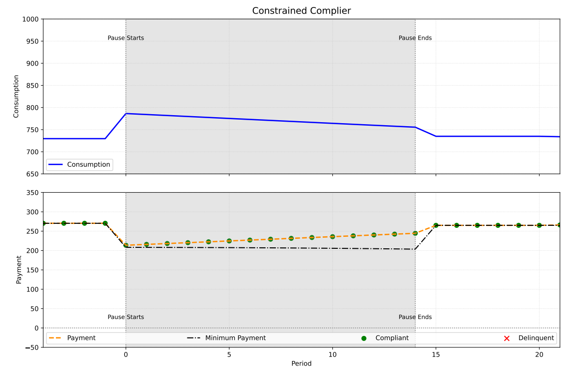
where f depends on inverse of $u'(c_t)$. Choices of consumption and repayment depends on the relationship between their Euler-optimal payment and the minimum required payment. Agents whose Euler-optimal payment exceeds the minimum simply pay their preferred amount, unconstrained by delinquency concerns. Those whose Euler-optimal payment lies below the minimum comply only if the delinquency penalty outweighs the utility gain from consuming more under noncompliance; otherwise, they choose to underpay.

To illustrate heterogeneous responses to lower payment requirement, we classify agents into four types based on their preferred repayment relative to the minimum requirement: (i) **Always Compliant**, whose preferred payment is above the minimum; (ii) **Constraint Compliant**, whose preferred payment is below the minimum but comply to avoid delinquency; (iii) **Strategic Delinquent**, who prefer to underpay and accept delinquency prior to the pause; and (iv) **Always Delinquent**, who never meet repayment thresholds. Figure 2.1 shows simulated trajectories of consumption and delinquency for each type.

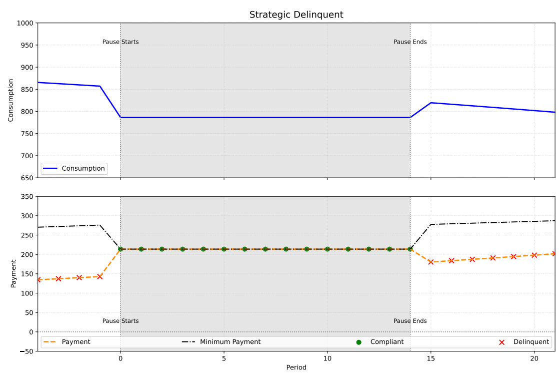
Prior to the pause, repayment behavior falls into three states: overpaying relative to the minimum (Always Compliant), exactly meeting the minimum (Constraint Compliant), or underpaying (Strategic/Always Delinquent). During the pause, the first group is unaffected, the second reallocates payments to consumption without changing delinquency, and the third group faces two possibilities—some begin complying if the reduced threshold is near their preferred level, while others remain delinquent. Once the pause ends, each type reverts to its pre-pause state.



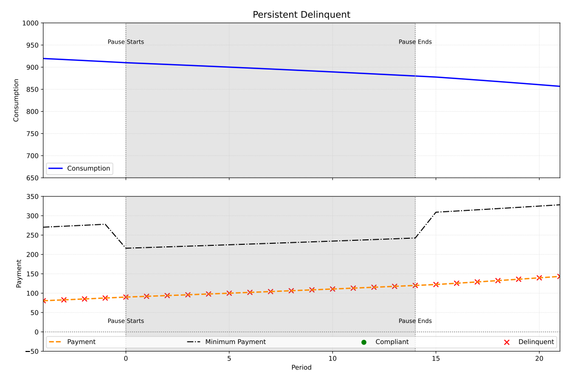
(a) Always Compliant



(b) Constraint Compliant



(c) Strategic Delinquent



(d) Always Delinquent

Figure 2.1: Simulated consumption, repayment, and delinquency trajectories across agent types.[3]

The aggregate implication is clear: delinquency falls mainly through temporary compliance of some Strategic Delinquents, while consumption rises across most types, with the exception that some Strategic Delinquents may see lower consumption when choosing compliance.

This theoretical framework rationalizes the empirical results that many would hypothesize: during the pause, borrowers reallocate payments to consumption, and some reduce delinquency. When the pause ends, higher m restores households' financial status to the pre-pause levels.

2.4 Data and Empirical Strategy

2.4.1 Data Description

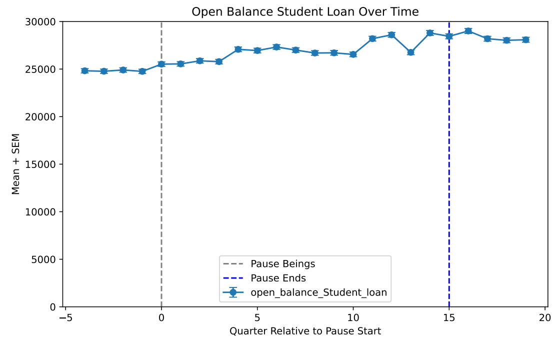
This subsection describes the panel dataset we construct, the criteria we use to form the analytical sample, and the key variables we derive to capture the operational status of student loans during the pause.

To empirically evaluate the theoretical predictions, we utilize a panel dataset derived from one of the three nationwide consumer credit reporting agencies from the University of California Consumer Credit Panel (UCCCP), provided by the California Policy Lab (CPL). The UCCCP includes anonymized credit records for a nationally representative 2% sample of U.S. adult consumers with credit histories, in addition to comprehensive coverage (100%) of Californians with credit records.

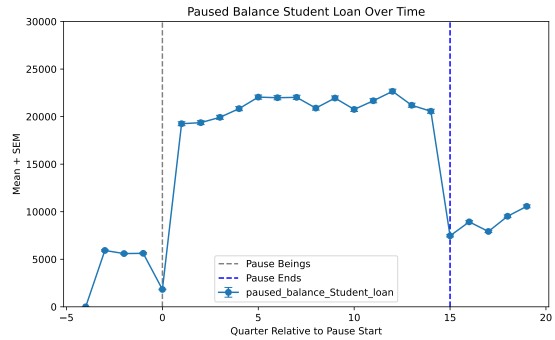
The UCCCP provides detailed, account-level information on a wide range of credit products, including student loans, credit cards, auto loans, and mortgages. These data are reported at the account $\times$ quarterly level, allowing us to observe changes in individual credit behavior over time. To facilitate person-level analysis and to balance computational feasibility with representativeness, individuals are randomly selected such that we observe about 30,000 individuals per quarter for the comparison of borrowers with eligible and ineligible student loans. We reshape the data into a panel format from 2018 onward, in which each observation corresponds to an individual in a given quarter. We aggregate account-level information into individual-level measures of total balances, payment amounts, delinquency rates, and other key credit indicators by loan type and quarter.

Within the sample, student loans represent a substantial component of household indebtedness. Before the federal pause, student loan payments made up about 16% of total loan repayment burdens across the six debt categories we track. During the pause, this share dropped to 5%, underscoring the immediate financial relief the policy provided. After the pause ended, the share recovered to 9%, but remained below pre-pause levels, suggesting a gradual adjustment in borrower behavior or potential challenges in resuming repayment.

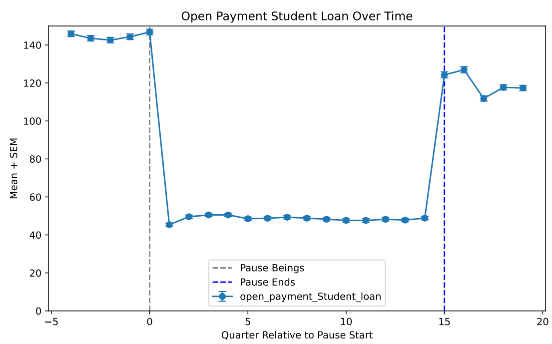
To study the effect of the federal student loan pause, we construct several new variables to capture the operational status of student loans. First, we identify all open student loan accounts—those not fully paid and still active—using a binary indicator present in the data. Closed loans are excluded from our analysis, as they are no longer subject to repayment or interest accrual. We then utilize balance and payment information to classify loans as "paused." Specifically, a loan is deemed paused if the reduction in the loan balance exceeds the reported payment amount. For instance, if an individual makes no payment yet the balance remains unchanged or decreases, it suggests that the loan is not accruing interest—indicative of the federally paused status.



(a) Open Balance



(b) Paused Balance



(c) Open Payment

Figure 2.2: Trends in Student Loan Balances and Payments Over Time. The figures display the evolution of student loan balances and payments for borrowers between 2019 and 2024. Open balance and payment reflect actively repaid loans, while paused balance and payment reflect loans that are paused by the federal student loan pause, which temporarily halted required payments and interest accrual. Quarter 0 shows the start of the pause at 2020 quarter 1, and quarter 15 shows the end of the pause at 2023 quarter 4.

The total balance of student loans remains relatively stable throughout the federal student loan pause period, as shown in Figure 2.2a. While one might expect a decline in balances due to payment relief, the average student loan balance actually increases gradually over time. This rise is largely driven by the accumulation of new student debt. The upward trend in balances confirms that no widespread loan forgiveness occurred during this period.

In contrast, student loan payments exhibit a dramatic decline during the pause. Average monthly payments drop from approximately \$140 to just \$40 per borrower, corresponding to a sharp increase in paused balances (Figures 2.2b and 2.2c). These patterns highlight the effectiveness of the payment suspension: borrowers are permitted to stop making payments without facing interest accrual or penalties. After the pause ends, open payments rebound significantly.

2.4.2 Empirical Strategy

We estimate the causal effect of student loan relief on financial behavior using an instrumental variables (IV) strategy within a difference-in-differences (DiD) framework. A central challenge is the endogeneity of payment behavior: individuals who pay more may differ systematically from those who pay less, in both observable and unobservable ways. We exploit the federal student loan pause as a quasi-experiment. During the pause, eligible borrowers were not required to make payments and interest accumulation was suspended. We do not directly observe eligibility, but proxy it using loan status: paused balances (no interest accrual) indicate eligible accounts, while the rest of open balances reflect ineligible.

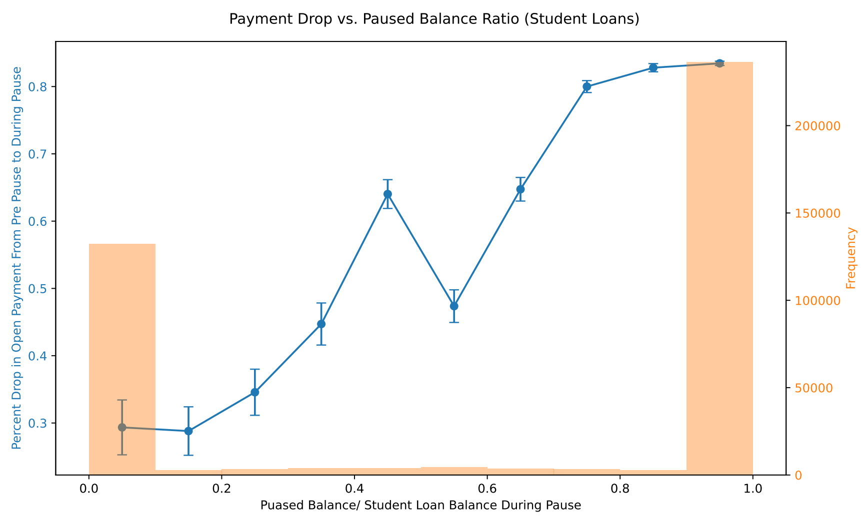


Figure 2.3: Pause Ratio and Drop in Student Loan Payments

This figure plots average drops in open payments against pause ratios (paused balance / total balance). Each point represents a 10 p.p. bin averaged across quarters 1–14 (pause period).

Our main endogenous variable is the open student loan payment of individual i in quarter t, OpenPayment_{it}. We instrument this using the share of balance not paused, $\frac{\text{OpenBalance}_{it} - \text{PausedBalance}_{it}}{\text{OpenBalance}_{it}}$, which captures the portion of loans that remained active and is plausibly exogenous conditional on controls and time effects.

$$\text{asinh}(\text{OpenPayment})_{it} = \pi_0+\pi_1\frac{(\text{OpenBalance}_{it} - \text{PausedBalance}_{it})}{\text{OpenBalance}_{it}}+\pi_2\text{TimePeriod}_t+\pi_3X_{it}+u_{it}, \tag{2.9}$$

where X_{it} includes controls such as age, occupation, and education; TimePeriod_t captures quarter fixed effects; and $\text{asinh}(x) = \log(x + \sqrt{1 + x^2})$, which behaves similarly to $\log(2x)$ for $x > 1$ and is approximately linear for $x \in [0, 1]$. The interpretation of the first-stage coefficient, π_1, is that a 100% reduction in the portion of the loan balance requiring payment (i.e., a full pause of eligible balances) leads to an approximate $\pi_1 * 100\%$ decrease in student loan payments.

Figure 2.3 shows that higher ratio of student loan balances being paused strongly predicts lower payment. The first stage table is also presented in Appendix Table 2.1. Given the strength of the first stage, we proceed with IV regressions.

Table 2.1: Regression of Student Loan Payments on Paused Ratio

Dependent Variable:	**asinh(Open Student Loan Payment)**
Paused Ratio of Student Loan	-1.3052***
	(0.0074)
Observations	877,984
R-squared	0.363
Adjusted R-squared	0.363
F-statistic	5958.0
Prob (F-statistic)	0.000
Covariates	Yes
State FE	Yes
Quarter FE	Yes

Notes: This table reports the OLS estimate of the relationship between the inverse hyperbolic sine of open student loan payments and the paused ratio of student loan balances. Robust standard errors are in parentheses. All specifications include quarter and state fixed effects and individual-level controls.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

The second stage estimates the effect of predicted student loan payments on outcomes:

$$\begin{aligned} asinh(Y_{it}) = \beta_0 + \beta_1\, asinh(\widehat{\text{OpenPayment}}_{it}) + \beta_2\, \text{TimePeriod}_t \\ + \beta_3 \left(asinh(\widehat{\text{OpenPayment}}_{it}) \times \text{TimePeriod}_t\right) + \gamma X_{it} + \epsilon_{it} \end{aligned} \tag{2.10}$$

where Y_{it} denotes balances, payments, or delinquency, and $asinh(\widehat{\text{OpenPayment}})_{it}$ is the fitted value from the first stage. The interaction with TimePeriod_t captures the elasticity of DiD effect over time. That is, β_3 can be interpreted as the effect of a 1% change in student loan payment—corresponding to a β_3% change in payments and balances, or a β_3 percentage-point change in delinquency.

To capture dynamic effects, we further estimate quarter-by-quarter DiD regressions:

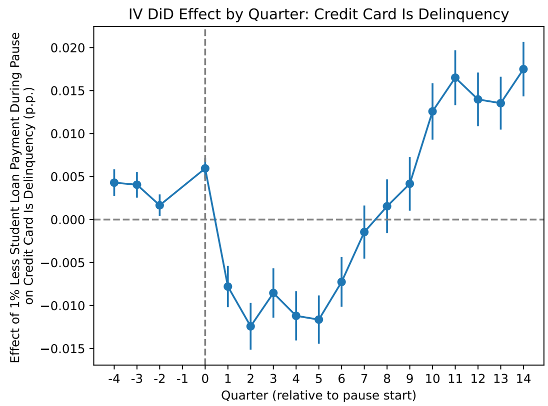
$$asinh(Y_{it}) = \alpha_i + \gamma_t + \sum_{q \in \mathcal{Q}} \delta_q \cdot (asinh(\widehat{\text{OpenPayment}})_i \times \mathbb{1}_{t=q}) + X'_{it}\beta + \epsilon_{it} \tag{2.11}$$

where α_i and γ_t are individual and time fixed effects, and δ_q traces dynamic treatment effects for eligible borrowers. This specification allows us to test parallel trends pre-policy and visualize how impacts evolve during and after the pause.

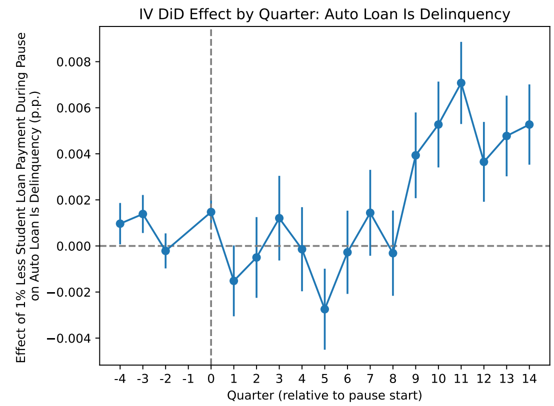
Chen and Roth [2024] highlight limitations of the *asinh* transformation, noting that percentage interpretations can be sensitive to the scale of the data. In this chapter, we focus primarily on the trends and statistical significance of the estimated effects rather than on precise percentage magnitudes.

2.5 Results and Analysis

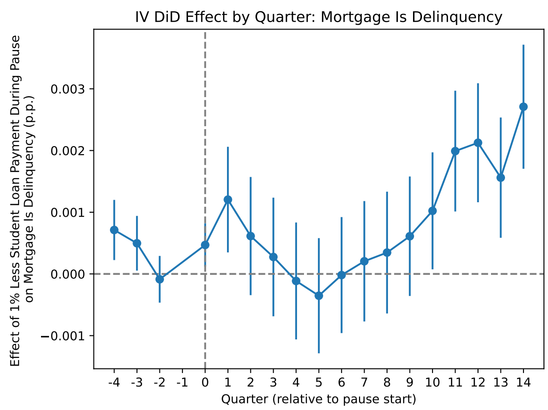
To rigorously evaluate the causal impact of the federal student loan pause, we use an instrumental variables (IV) difference-in-differences (DiD) framework. By leveraging quasi-random variation in paused loan balances as an instrument for actual student loan payment changes, we aim to isolate the exogenous component of payment variation driven by policy. We focus on six categories of debt: credit cards, mortgages, auto loans, sum of all accounts, large-balance debts, and daily-use debts. Our regressions incorporate demographic controls, baseline loan statuses, and individual fixed effects. The IV DiD specification is estimated both in pooled form and quarter-by-quarter to capture dynamic effects.



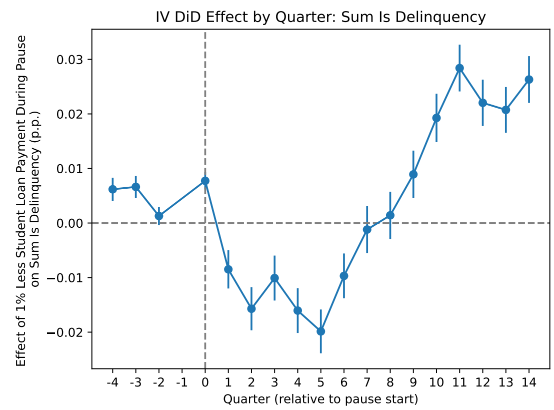
(a) Credit Card Delinquency



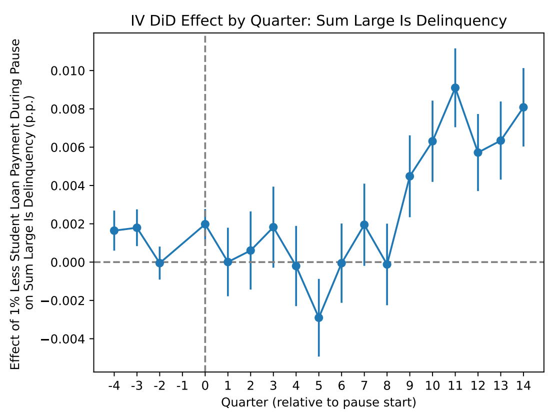
(b) Auto Loan Delinquency



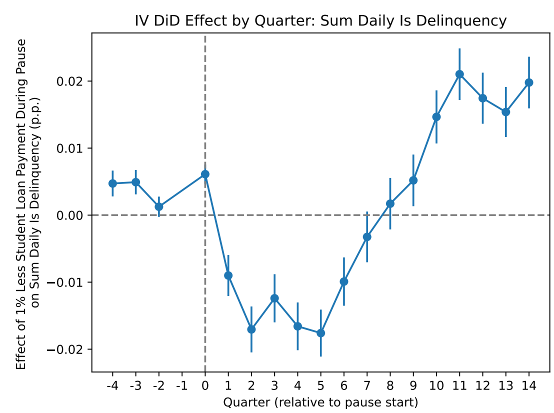
(c) Mortgage Delinquency



(d) Sum Delinquency



(e) Sum (Large) Delinquency



(f) Sum (Daily) Delinquency

Figure 2.4: Quarter by Quarter Regression Coefficient for 6 types of loans for Impact of Payment Reduction on Other Debt Delinquency. Individual fixed effects are added. Standard error bars represent 95% confidence interval

We first present the event-study design of delinquency using Equation 2.11. The patterns deviate from the theoretical prediction that delinquency levels should remain lower when repayment obligations are reduced. While we observe some initial relief—reflected in flat or even slightly declining delinquency rates across most debt types—this effect is short-lived. Over the course of the pause period, delinquency rates begin to rise steadily across all categories of debt, despite the fact that overall minimum repayment requirements have decreased due to the suspension of student loan payments.

This pattern is puzzling under a simple consumption-smoothing framework, which would predict improved financial stability with lower debt servicing costs. One possible explanation is that the liquidity created by the pause was not used solely for deleveraging or precautionary savings, but rather fueled additional consumption and borrowing. As individuals took on more debt, they may have overextended themselves financially. In this case, the pause may have inadvertently enabled a temporary increase in living standards that could not be sustained without student loan obligations, leading to increased financial strain and higher rates of delinquency on other debts.

Further examining the impact of delinquency level on aggregate during the pause using Equation 2.10, Table 2.2 presents the results of the IV DiD regressions estimating the effect of a 100% decrease in federal student loan payments—driven by eligibility for the COVID-era payment pause—on delinquency amounts across various credit types. The most robust finding is for credit card delinquency, but also strong in general for aggregate delinquency levels. Effects on other loan types (mortgage, auto loan) are generally positive but not statistically significant, possibly due to changing delinquency impact through time as shown in Figure 2.4. However, the delinquency among all debts, especially among daily purchases, rise significantly during the pause.

A plausible explanation of rising delinquency is increasing payment across other debts during the pause. A quarter-by-quarter analysis (Figure 2.5) shows rising payment across debt over time. Initially, repayment fell across nearly all credit types, especially daily-use debts, as eligible households do not need to rely on borrowing to sustain consumption. As the pause persisted, repayment gradually rose, with the strongest rebound in daily-use categories, while secured or large-ticket debts responded more slowly. These dynamics suggest that liquidity was first met with precaution but later reinforced ongoing consumption.

The pattern of payment rising demonstrates another notable deviation from the neoclassical prediction: the gradual, rather than immediate, increase in payment levels across other types of debt. In a standard model with forward-looking agents, an exogenous increase in liquidity—such as the suspension of student loan payments—should lead to an immediate adjustment in financial behavior. Individuals would be expected to immediately reallocate this freed-up cash flow toward other uses, such as increased consumption, accelerated repayment of other debts, or higher savings. Yet, the payment only rises gradually.

Complementing the event study design, Table 2.3 shows the aggregate impact of lower student loan repayment on other debts' payments. Reductions in student loan repayment translated into higher payments on other debts, particularly credit cards and auto loans, with most of the increase concentrated in daily-use categories. This indicates that households primarily redirected liquidity toward everyday spending rather than large purchases.

Table 2.2: Effect of Student Loan Payment Reduction on Loan Delinquency during Pause

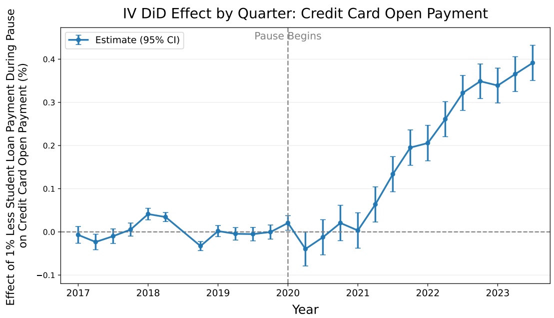
	Credit Card	**Mortgage**	**Auto Loan**	**Sum**	**Sum (Large)**	**Sum (Daily)**
Per 100% Drop	0.392***	0.042	0.009	0.673***	0.059	0.626***
in Loan Payment	(0.080)	(0.026)	(0.046)	(0.113)	(0.054)	(0.099)
Observations	657,981	657,981	657,981	657,981	657,981	657,981
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each column reports results from a separate IV DiD regression using eligibility for the federal student loan payment pause as the instrument. Coefficients and standard errors have been multiplied by 100 and represent the effect of a 100% reduction in student loan repayment. Signs are flipped for interpretation. Standard errors (clustered at the individual level) are in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

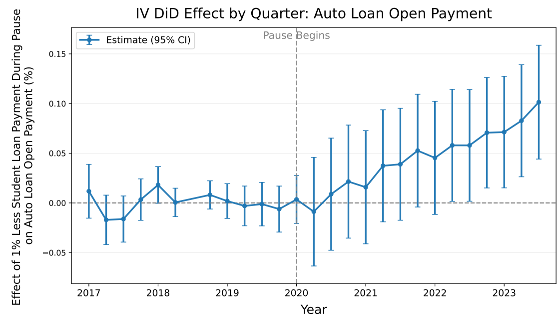
Taken together, patterns from event-study designs and aggregate regressions have several misalignments from theoretical predictions in Section 2.3: 1) rising delinquency during the pause is the complete opposite of prediction using a neoclassical model and 2) gradual, rather than immediate, increasing consumption also does not align well with a model solely based on liquidity relief.

Further examining the impact of repayment after the policy ends, Table 2.4 presents difference-in-differences estimates of the effect of resumed student loan repayment on delinquency across different types of debt in the post-pause period. The results indicate a statistically significant increase in delinquency following the end of the pause, particularly for auto loans, mortgages, and the sum of large-balance debt.

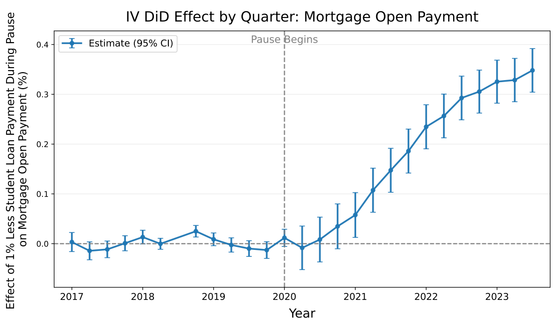
The total delinquency (sum) also rises significantly. These findings offer strong support for the notion that individuals are financially worse off after the pause ends. The resumption of student loan payments appears to crowd out their ability to stay current on other debt obligations. This result is consistent with earlier findings on rising credit card usage and declining large debt payments, indicating that households are likely operating under binding liquidity constraints. The increase in delinquency suggests that many borrowers are unable to absorb the added repayment burden without falling behind on other financial obligations, raising concerns about the long-term stability of household balance sheets as student debt servicing resumes.



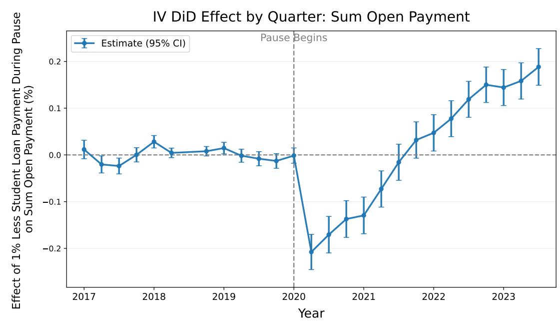
(a) Credit Card Payment



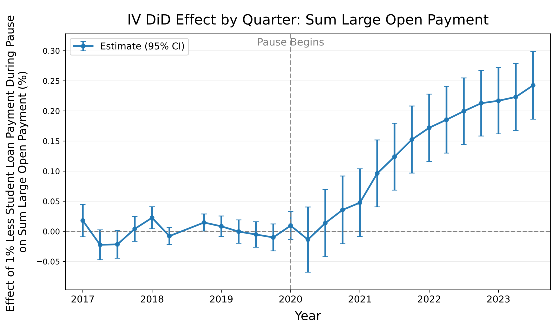
(b) Auto Loan Payment



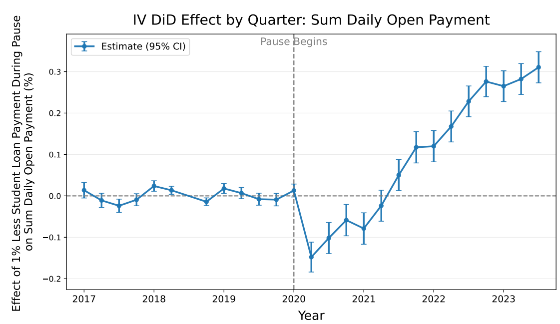
(c) Mortgage Payment



(d) Sum Payment



(e) Sum (Large) Payment



(f) Sum (Daily) Payment

Figure 2.5: Quarter by Quarter Regression Coefficient for 6 types of loans for Impact of Payment Reduction on Other Debt Payments. Individual fixed effects are added. Standard error bars represent 95% confidence interval

Individuals also suffer from decreasing consumption as the pause ends as evident in Table 2.5, which presents difference-in-differences estimates of how changes in student loan repayment affect payments toward other types of debt. Payments toward all other spendings, both daily spending like credit cards and large fixed obligations decline, with the decrease in all payments statistically significant at the 5% level. This finding likely suggests that households are unable to sustain payment levels across almost all debt types prior to the end of the pause.

Table 2.3: Effect of Student Loan Payment Reduction on Debt Payment

	Credit Card	**Mortgage**	**Auto Loan**	**Sum**	**Sum (Large)**	**Sum (Daily)**
Per 1% Drop	0.0486***	-0.009	0.053***	0.059***	0.049***	0.047***
in Loan Payment	(0.007)	(0.008)	(0.010)	(0.007)	(0.010)	(0.006)
Observations	657,981	657,981	657,981	657,981	657,981	657,981
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each column reports results from a separate IV DiD regression using eligibility for the federal student loan payment pause as the instrument.Signs are flipped for interpretation. Standard errors (clustered at the individual level) are in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 2.4: Effect of Increase in Student Loan Repayment on Delinquency

	Credit Card	**Mortgage**	**Auto Loan**	**Sum**	**Sum (Large)**	**Sum (Daily)**
Effect of 100%	0.386***	0.123***	0.318***	0.865***	0.428***	0.500***
Increase in Repayment	(0.123)	(0.040)	(0.068)	(0.164)	(0.080)	(0.146)
Observations	300,712	300,712	300,712	300,712	300,712	300,712
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each column reports a separate DiD regression using the post-pause period. The dependent variable is delinquency amount in each category. The main independent variable is the resumption of student loan repayment. Standard errors clustered at the individual level are shown in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 2.5: Effect of a Increase in Student Loan Repayment on Other Debt Payments after Pause

	Credit Card	**Mortgage**	**Auto Loan**	**Sum**	**Sum (Large)**	**Sum (Daily)**
Effect of 1%	-0.030**	-0.019	-0.047***	-0.090***	-0.070***	-0.056***
Increase in Repayment	(0.012)	(0.012)	(0.017)	(0.011)	(0.016)	(0.011)
Observations	300,712	300,712	300,712	300,712	300,712	300,712
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each column reports a separate difference-in-differences regression where the outcome is the change in payment amount for the respective debt category. The main independent variable is the change in federal student loan repayment (due to the pause ending). Standard errors clustered at the individual level are in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

As individuals resume their student loan payments, they cut back on all their spending, suffering utility loss from decreasing consumption. Yet, even the reduction in spending is not enough to prevent them from being delinquent on their debt, as evidenced by Table 2.4. This behavioral inertia implies that financial habits are sticky—agents may initially maintain pre-pause spending levels out of caution or inertia, and only later adjust upward as the pause persists. It also suggests that temporary policies, if extended or perceived as long-lasting, can reshape financial behavior in persistent ways, with implications for both short- and long-term financial outcomes.

These findings raise important questions about the long-term financial resilience of borrowers following the end of the payment pause—and whether reliance on revolving credit could lead to worsening debt cycles post-policy.

2.6 Behavioral Mechanism and Policy Implication

Section 2.5 documents two main deviations from the neoclassical predictions of the model in Section 2.3: first, an increase in delinquency during the later stages of the pause that continues after the pause ends, across all forms of debt; and second, a gradual rather than immediate increase in consumption, reflected in rising debt balances and payments over time. In this section, we first explore a behavioral mechanism that can explain these patterns, and then use the resulting model to evaluate alternative policy designs that may improve welfare.

2.6.1 Habit Formation with Liquidity-Driven Consumption

Model Setup

We consider a hand-to-mouth consumer who allocates current income between consumption and debt repayment. By Remark 2.1, we assume borrowers do not save. Thus, consumption in period t is determined by:

$$c_t = w_t - p_t, \tag{2.12}$$

where w_t is income and p_t is the loan repayment in period t.

Habit Formation with Expectation-Based Reference Point. We now specify the evolution of the reference point r_{t+1} as a combination of realized consumption in period t and the agent's expectation of future consumption:

$$r_{t+1} = \rho c_t + (1-\rho)\mathbb{E}_t[c_{t+1}], \quad \rho \in [0,1), \tag{2.13}$$

where $\mathbb{E}_t[c_{t+1}]$ reflects the consumer's internal forecast of their next-period consumption.

Crucially, we allow expectations to incorporating not only the current consumption level but also recent trends in consumption:

$$\mathbb{E}_t[c_{t+1}] \text{ depends on } \{c_s\}_{s\leq t}, \text{ including both levels and growth of past consumption.}$$

This formulation captures the idea that individuals form upwardly drifting reference points not only due to high recent consumption, but also from anticipating continued growth—especially during periods such as a repayment pause when constraints are relaxed.

Utility. Utility in each period is given by:

$$U_t(c_t, d_t; r_t) = u(c_t) + v(c_t; r_t) - kd_t, \tag{2.14}$$

where:

- $u(c_t)$ is increasing and concave,
- $v(c_t; r_t)$ is a reference-dependent gain-loss utility with a kink at $c_t = r_t$,
- $d_t \in \{0,1\}$ is a delinquency indicator, incurring fixed cost $k > 0$.

We assume $v(c_t; r_t)$ satisfies:

1. Continuity and differentiability for all $c_t \neq r_t$;
2. $v'(c_t; r_t) > 0$ if $c_t < r_t$; $v'(c_t; r_t) \approx 0$ if $c_t \geq r_t$;

3. $v''(c_t; r_t) \leq 0$ (diminishing sensitivity);
4. Kink at r_t: $\lim_{c_t \to r_t^-} v'(c_t; r_t) \gg \lim_{c_t \to r_t^+} v'(c_t; r_t)$ (loss aversion).

Naive Borrower's Problem. The naive borrower does not forecast future reference point changes and treats r_t as given. They choose whether to comply with repayment or become delinquent:

Delinquency:

$$U_t^{\text{del}} = u(w_t) + v(w_t; r_t) - k.$$

Repayment: They choose repayment $p_t \in [mB_t, w_t]$, resulting in consumption $c_t = w_t - p_t$, and solve:

$$U_t^{\text{rep}} = \max_{p_t \in [mB_t, w_t]} \{u(w_t - p_t) + v(w_t - p_t; r_t)\}.$$

Choice: The borrower compares utility from delinquency versus repayment and chooses:

$$\max \left\{ u(w_t) + v(w_t; r_t) - k, \max_{c \in [0, w_t - mB_t]} \{u(c) + v(c; r_t)\} \right\}. \tag{2.15}$$

Prediction. This formulation can generate increases in delinquency after repayment resumes, even among previously compliant borrowers. During a pause:

- $mB_t \downarrow \Rightarrow p_t \downarrow \Rightarrow c_t \uparrow \Rightarrow r_t \uparrow$.
- Hence, reference point r_{t+1} drifts upward due to both high consumption and growth expectations.

After repayment resumes, required payment $mB_t > 0$ reduces feasible consumption. If the post-resumption consumption falls below the elevated r_t, borrowers may experience strong reference-dependent losses, inducing delinquency even if they were previously compliant.

Proposition 2.1 (Habit-Driven Amplification of Consumption)**.** *Suppose the required minimum repayment falls from m to $\bar{m} < m$ at period $t = \underline{t}$, relaxing the constraint $p_t \geq mB_t$ to $p_t \geq \bar{m}B_t$. Let income $w_t = w$ and balance $B_t = B$ be constant, and assume the borrower always complies (i.e., no delinquency). Then, there exists $t' > \underline{t}$ such that:*

$$c_{t'} < c_{t'+1}, \quad r_{t'+1} < r_{t'+2}.$$

That is, the combination of relaxed repayment and reference-dependent preferences can endogenously generate a dynamic rise in consumption and reference points over time.

Proof idea. The initial drop in repayment increases consumption. This raises the reference point in the next period, which then increases optimal consumption further. A positive feedback loop forms until the repayment hits the lower constraint or the loop stabilizes.

Proposition 2.2 (Habit-Driven Delinquency During Pause)**.** *Suppose the required minimum repayment drops from* m *to* $\bar{m} < m$ *at time* $t = \underline{t}$*, and remains at* $\bar{m}$ *until* $t = \bar{t}$*. Assume the borrower was fully compliant prior to the pause (i.e.,* $d_t = 0$ *for all* $t < \underline{t}$*). Then there exist parameter values such that:*

- *the borrower remains compliant immediately after the pause begins,*
- *but due to habit-driven consumption growth, becomes delinquent during the pause (i.e.,* $d_t = 1$ *for some* $t \in (\underline{t}, \bar{t})$*).*

That is, an agent with full prior compliance may endogenously become non-compliant during a repayment pause if consumption and reference points increase sufficiently.

Proof idea. We build on Proposition 2.1, where a reduction in the repayment floor to $\bar{m}B_t$ leads to increasing consumption c_t and reference point r_t over time via habit formation:

$$r_{t+1} = \rho c_t + (1 - \rho)\mathbb{E}_t[c_{t+1}].$$

Initially, the borrower chooses a compliant repayment $p_t \geq \bar{m}B_t$, allowing higher consumption and setting off the feedback loop:

$$c_t \uparrow \Rightarrow r_{t+1} \uparrow \Rightarrow p_{t+1} \downarrow \Rightarrow c_{t+1} \uparrow .$$

As this cycle continues, the reference point r_t grows rapidly. Eventually, the marginal disutility of reducing consumption (due to kinked loss aversion at $c_t < r_t$) dominates the cost of delinquency. At some time $t^* \in (\underline{t}, \bar{t})$, the utility-maximizing borrower strictly prefers to set $p_{t^*} < \bar{m}B_{t^*}$, entering delinquency (i.e., $d_{t^*} = 1$).

This result holds despite the absence of income shocks or exogenous shocks to debt, purely as a result of **self-reinforcing consumption expectations** under habit formation.

Proposition 2.3 (Post-Pause Delinquency Spike)**.** *Suppose the minimum payment rate returns from* $\bar{m}$ *to* m *at* $t = \bar{t} + 1$*, after a pause period. Then, even if the agent was compliant throughout the pause, there exists* $t > \bar{t}$ *such that:*

$$d_t = 1 \quad \textit{(delinquency occurs).}$$

Proof idea. Elevated reference point due to habit formation during the pause causes large psychological loss once repayment increases and consumption falls. This gap may trigger delinquency, even without income change.

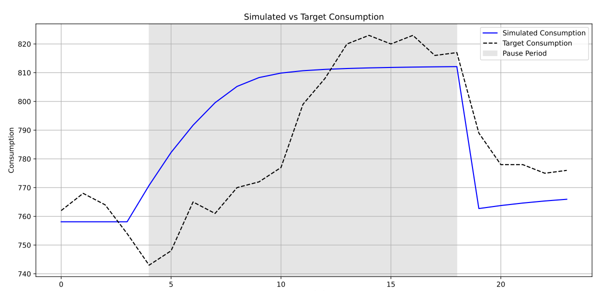
Conclusion. These theoretical results jointly mirror the empirical patterns observed in Section 2.5: gradually increasing consumption (Proposition 2.1), rising delinquency during the pause (Proposition 2.2), and a post-pause spike in default risk (Proposition 2.3).

2.6.2 Policy Implication

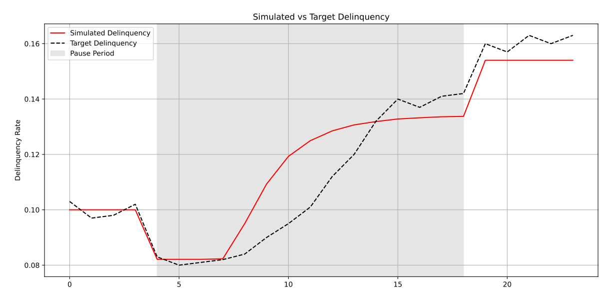
We use the structural model developed in Section 2.6.1 to simulate alternative policy counterfactuals. The model is calibrated to match empirical moments of consumption and delinquency observed before, during, and after the federal student loan pause. With a bi-modal distribution of patience, measured by δ, and loss aversion, measured by λ, we are able to replicate the patterns of the empirical data in Section 2.5.

Simulation Effect

To assess the ability of the model to capture observed dynamics, we compare simulated outcomes under the reference-dependent induced habit formation framework to the empirical series for consumption (proxied by debt payments) and delinquency rates. In Figure 2.6, the model reproduces the key empirical patterns: 1) gradual increase in consumption due to increasing reference points 2) initial decline of delinquency level, 3) gradual increase of delinquency, and 4) further increase in delinquency and decrease in consumption after pause.



(a) Consumption Fit (Payments)



(b) Delinquency Fit

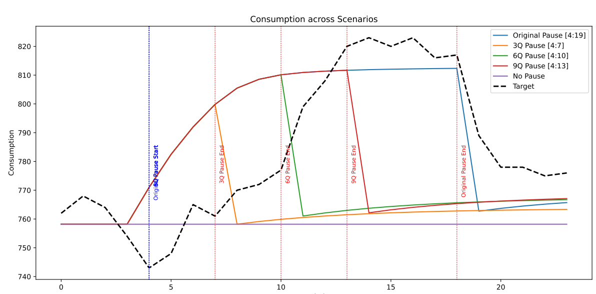
Figure 2.6: Model fit for consumption and delinquency under the reference-dependent framework with moving reference points. The model reproduces the observed increase in consumption and eventual rise in delinquency following the student loan pause.

The dynamics of consumption and delinquency emerge from heterogeneity in patience and loss aversion. The rise in consumption during the pause is driven primarily by patient households that had previously been constrained by minimum payment requirements; the reduction in required payments freed liquidity that they allocated toward additional spending. The initial drop in delinquency comes from a subset of impatient borrowers whose optimal payment choice lay far below the original minimum but near the lowered threshold, making compliance temporarily attractive. Over time, however, higher consumption raised reference points, and for loss-averse households this gradually pushed desired payments below the effective minimum, generating an upward trend in delinquency. Once the pause ended, impatient borrowers who had complied under the temporary terms again fell short of their obligations, leading to a sharp post-pause resurgence in delinquency.

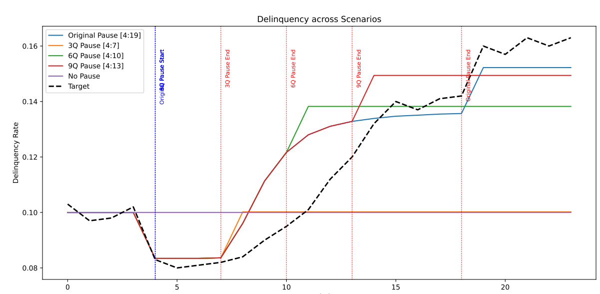
Counterfactual Policy Experiments—Ending the Policy Earlier

To evaluate the welfare implications of alternative designs, we conduct counterfactual simulations varying the length of the repayment pause. Specifically, we compare policies lasting 3, 6, and 9 months, as well as a case with no pause, against the actual four-year suspension. For each scenario, we simulate consumption (proxied by payments) and delinquency trajectories. The model assumes that when the policy ends, expectations reset: borrowers who were non-delinquent remain current, while those already delinquent continue in delinquency. Figure 2.7 presents the results.

The simulations reveal that the eventual decline in consumption occurs regardless of policy duration, as households must ultimately comply with the original minimum payment requirement. By contrast, delinquency dynamics are highly sensitive to policy length. When the pause ends earlier—before borrowers have fully adapted to higher consumption levels—delinquency rates remain substantially lower both during and after the pause. These findings underscore that prolonged forbearance fosters habit formation and amplifies long-run repayment strain, while shorter or partial suspensions mitigate these adverse behavioral responses.



(a) Consumption Trajectories



(b) Delinquency Trajectories

Figure 2.7: Counterfactual simulations comparing alternative policy lengths (3, 6, and 9 months, and no pause) with the actual student loan repayment pause. While consumption eventually converges in all cases, delinquency dynamics diverge: shorter pauses prevent the rise in delinquency that emerges under prolonged forbearance.

The trade-offs in welfare evaluation, thus, is between a prolonged period of elevated consumption and lower delinquency level after the pause. While the cost a delinquency is subject to debate [Greenwald et al., 2014], we estimate the cost of delinquency using a combination of late fees and higher future borrowing costs, which together imply that each delinquency costs a household roughly \$600. This aligns with our calibration of $k = 0.5$ in the utility specification: under log utility, $\log(c + 600) - \log(c)$ with $c \approx 800$ yields 0.5, meaning that the welfare loss from a single delinquency is equivalent to a 50% drop in marginal utility at this consumption level. Table 2.6 summarizes the simulated welfare outcomes across different policy durations.

Although aggregate welfare is relatively stable across scenarios, this masks large heterogeneity across types. The similarity at the group level is largely driven by higher consumption for all households during the pause. However, high δ high λ people suffer the most: their utility losses relative to the more patient group are sizable, particularly under the prolonged pause where habit formation amplifies post-resumption delinquency.

Table 2.6: Welfare Across Pause Lengths

	Aggregate	High δ Low λ	High δ High λ	Low δ
Original Pause	6.6040	6.6456	6.3843	6.3709
3M Pause	6.5880	6.6132	6.6108	6.3619
6M Pause	6.5860	6.6217	6.4388	6.3640
9M Pause	6.5914	6.6306	6.3943	6.3662
No Pause	6.5786	6.6067	6.6067	6.3261

Importantly, the true cost of delinquency likely exceeds our \$600 benchmark. Beyond household financial strain, delinquency imposes externalities in the form of social and governmental costs. As Welsh [2003] and Craig et al. [2011] emphasize, defaults generate administrative burdens, collection expenses, and broader credit market distortions that fall on taxpayers and financial institutions. If the social cost of delinquency is even twice as large as the private cost—say \$1,200 per delinquency—then a 6% increase in delinquency rates, which is the actual increase for people whose student loan payments were paused (Figure 2.6b), translates into billions of dollars in lost welfare when aggregated across households, government expenditures, and social spillovers. These broader costs highlight that the welfare implications of prolonged debt forbearance cannot be ignored. Hence, we propose that ending the policy earlier, ideally before people form unsustainable habits, may yield better long term results for societal welfare.

2.7 Conclusion

This chapter evaluates the effects of the federal student loan payment pause on household financial behavior, focusing on debt repayment and delinquency across various credit types. The results suggest that while the policy initially achieved its goal of relieving financial burden and stimulating consumption, it also led to unintended longer-term consequences. As the pause extended, individuals gradually developed new spending habits, increasing their use of credit and raising both repayment amounts and debt balances—particularly in categories tied to day-to-day consumption and auto loans. These increases, while manageable during the pause, appear to have been unsustainable once repayment resumed.

Following the end of the pause, we observe a clear rise in delinquency, even after accounting for individual and time fixed effects. This suggests that the temporary increase in liquidity may have encouraged financial behaviors that borrowers were unable to maintain under normal repayment obligations, ultimately leaving many in a worse financial position than before. The sharp post-pause increase in delinquency reinforces the concern that some borrowers were not financially prepared for the return of student loan payments.

Using a reference-dependent consumption model with habit formation, we build a simulation that matches our empirical result and explore potential changes in policy that may offer better financial wellbeing after the pause. We find that either ending the pause earlier or reducing the magnitude of the payment drop would lower delinquency after the pause and reduce volatility in financial wellbeing.

Future research could explore the long-term effects of the pause and repayment resumption, including whether delinquency stabilizes, continues to rise, or leads to more severe outcomes such as defaults or bankruptcy. It would also be valuable to study heterogeneous effects—for example, whether lower-income borrowers, first-generation students, or borrowers with high pre-pause debt levels were disproportionately affected. Finally, further work could examine how perceptions of policy permanence affect financial decision-making, particularly when temporary relief measures create lasting behavioral changes.

Chapter 3 Local Housing Market Experiences and Homeownership Decisions

3.1 Introduction

The previous chapter showed that behavioral forces shape how households respond to changes in their debt obligations. This chapter turns to a different, and arguably larger, margin of household financial decision-making: the decision to become a homeowner. While prior work emphasizes macroeconomic or network-mediated experiences, much less is known about how individuals' own local housing market experiences shape whether and when they become homeowners. In contrast to stocks and bonds, housing returns are highly heterogeneous across locations, making local experiences presumably more salient than national averages for households' tenure choices. Studying these effects requires data that track individuals' residential locations at fine geographic granularity (e.g., zip code) over an extended horizon.

Beyond experienced returns, periods of heightened uncertainty also leave deep and lasting marks on economic behavior. When individuals live through sharp swings in prices or economic activity, volatility becomes more salient than average returns, shaping perceptions of risk and willingness to commit to long-term investments [Duxbury and Summers, 2018, Levin and Vidart, 2023]. These experience effects are not abstract: the turbulence of the global financial crisis, the housing boom and bust, the pandemic, and the recent era of inflation and interest-rate uncertainty continue to reverberate through households' financial choices [Malmendier, 2021]. Understanding how such lived experiences of volatility translate into durable shifts in behavior is essential, not only for explaining market dynamics today but also for anticipating the long-run consequences of the uncertainty shocks currently confronting households worldwide.

However, estimating impacts of volatility has been difficult, as measuring not just the mean but the dispersion of experiences requires finely grained data that capture how individuals confront uncertainty at the district or neighborhood level. Aggregate volatility measures—whether at the state or national level—wash out precisely the variation that drives household perceptions of risk and their corresponding choices. This gap is compounded by the lack of theoretical models that integrate housing, experiential learning, and volatility perceptions into a unified framework.

In this chapter, we combine a novel theoretical framework with uniquely detailed microdata. We develop a dynamic life-cycle model in which households form beliefs about housing returns and volatility based on their personal histories, and we derive clear predictions: higher experienced returns encourage ownership, while higher experienced volatility discourages it. We then test these predictions using the University of California Consumer Credit Panel (UC-CCP), which allows us to construct individual-level histories of local housing market experiences. Our empirical analysis confirms the model's implications: households with higher experienced returns are significantly more likely to transition into homeownership, while households with higher experienced volatility are less likely to buy. Specifically, we find that a 1 percentage point (p.p.) increase in experienced returns raises probability of transitioning into homeownership by 7% relative to the quarterly average of homeownership transition rate, and a 1 p.p. increase in experienced volatility decreases this probability by 10%. These results remain robust across a wide set of controls and specifications, underscoring the importance of individuals' local housing market experiences of the mean and volatility of housing returns in shaping their homeownership decision.

To motivate our empirical analysis, we introduce experience-based learning about the mean and variance of future housing returns into an otherwise standard life-cycle model with a rent-own tenure choice. Households derive utility from consumption and housing services, face idiosyncratic income risk, and in each period choose whether to rent or own. Aggregate house prices follow a deterministic trend with aggregate shocks; the return observed by household i is the product of aggregate house-price growth and an idiosyncratic shock, generating cross-sectional heterogeneity in beliefs about future returns. Following Malmendier and Nagel [2011], agents use the history of return realizations observed over their lifetimes to form beliefs about the mean $\mu_{i,t}$ and standard deviation $\sigma_{i,t}$ of future housing returns. The model yields sharp comparative statics: higher subjective mean returns raise the continuation value of ownership and the probability of buying, whereas higher subjective volatility—through risk aversion—reduces the continuation value of ownership and lowers the probability of becoming a homeowner.

To test the predictions of our framework, we use the University of California Consumer Credit Panel (UC–CCP), an administrative dataset covering all California residents with credit records and a 2% nationally representative sample of U.S. individuals from 2004-2025. To maintain tractability while retaining sufficient cross-sectional heterogeneity in individual experiences, we work with a 1% subsample of the national panel, which covers approximately 100,000 individuals per quarter. For each individual in the sample, we compute the experienced housing return as the recency-weighted mean of nominal house-price growth across the zip codes where they have lived since age 18[1]; we calculate the individual's experienced volatility as the corresponding recency-weighted standard deviation of those returns. To understand

[1]Data on nominal house price growth at the zip code level is obtained from Zillow.

how experiences at different geographies affect homeownership choice, we also construct county and state level counterparts to our zip code-level experience measures. Guided by the model's rent-own choice, our primary outcome is transition into homeownership, measured by first-time mortgage origination, i.e., the first quarter an individual appears with a mortgage on record in our credit panel.

Our empirical analysis confirms the model's predictions. Higher experienced returns significantly raise the probability of transitioning into homeownership, whereas greater experienced volatility reduces it. Quantitatively, a 1 p.p. increase in quarterly experienced returns raises the transition probability by 7% relative to the sample's mean quarterly transition rate, while a 1 p.p. increase in experienced standard deviation lowers it by 10%. These effects are robust to a comprehensive set of controls—including demographics; levels and lagged growth of zip code-level house prices and rents; and individual credit characteristics—as well as fixed effects for county-quarter, age, education, and occupation. We also document that the effects decay with age, consistent with life-cycle models. Finally, we also show that our experience measures predict homeownership status at the end of the sample and that conditional on ownership, our experienced returns measure also predicts higher leverage at mortgage origination, highlighting that past experiences can also shape household risk taking.

Together, our framework and evidence demonstrate that personal histories of local housing markets exert a powerful influence on homeownership decisions. By integrating a theoretical structure with uniquely detailed microdata, the chapter establishes local house price return and volatility experiences as a central, yet previously understudied, determinant of household housing choices.

Literature Review

Our chapter builds on the vast and growing literature in finance that studies the effects of personal experiences on beliefs and choices. Previous work has documented the effects of experiences on beliefs and participation in stock and bond markets [Malmendier and Nagel, 2011, Malmendier et al., 2020, Kaustia and Knüpfer, 2008]. A more recent literature studies how lived experiences shape housing decisions. Botsch and Malmendier [2023] show that greater exposure to high nominal interest rates increases demand for fixed-rate mortgages. Malmendier and Wellsjo [2024] find that higher inflation experiences raise the likelihood of homeownership, as houses are traditionally viewed as assets that hedge against inflation. They further show how inflation experiences explain differences in homeownership rates across countries. Bailey et al. [2018] show that stronger house-price growth experienced by an individual's Facebook friends elevates the individual's perception of housing as a good investment and increases the probability of transitioning into homeownership. We contribute to this literature by focusing on the effects of local housing market experiences—specifically, the mean and standard deviation of nominal house-price growth in an individual's zip code since age 18—on the individual's decision to transition into homeownership.

There is also a substantial literature that studies the effects of past experiences on beliefs about future house prices. It has been shown that households extrapolate from past house price changes to form expectations of future housing returns [Case and Shiller, 2003, Case et al., 2014, Shiller and Thompson, 2022]. Kuchler and Zafar [2019] show that households that have experienced higher local house price growth have higher expectations over future national house prices. They also show that higher experienced volatility of local house prices raises the dispersion in expected future housing returns. Agarwal et al. [2015] show that households that have lived through faster local house price growth are more likely to purchase homes sooner. Armona et al. [2019] show how house price expectations respond to information treatments about past house prices. Adelino et al. [2018] show that perceptions of house-price risk are formed from local house price volatility and recent local house price movements. They also show that perceptions of risk in the housing market strongly predict current tenure choice and future intentions to buy houses. Relative to these studies, while we do not use surveys, the nationally representative University of California Consumer Credit Panel (UC-CCP) allows us to track individuals' residence at the zip code level over time, construct person-specific histories of local nominal house-price returns and volatility since age 18, and observe the exact quarter in which they become homeowners. The precision in geographic regions allows us to estimate at scale how local housing market experiences shape the probability of a first-time transition into homeownership.

The remainder of this chapter proceeds as follows: Section 3.2 introduces a model of homeownership choice under experience-based belief formation, incorporating both return and volatility learning. Section 3.3 describes our data sources, presents summary statistics, and defines our empirical measures of experiences. Section 3.4 presents the empirical strategy and the main results. Section 3.5 concludes.

3.2 Model

In this section, we propose a theoretical framework which guides our empirical analysis. Households live for finite periods and in each period, they choose whether to own or rent a house. Ownership decisions depend on households' expectations of the mean and volatility of future house price returns, which are formed from the realizations of house prices that they have seen during their lifetime.

A continuum of households live for T periods, $t \in \{0, 1, ..., T\}$. Each household i is endowed with initial saving $s_{-1,i}$ and maximizes lifetime utility

$$U_i \;=\; \sum_{t=0}^{T} \beta^t \, \frac{g(c_{i,t}, h_{i,t})^{1-\eta}}{1-\eta}, \qquad 0 < \beta < 1,\ \eta > 0.$$

The within-period aggregator over consumption and housing services is the constant-elasticity-of-substitution (CES) function

$$g(c_{i,t}, h_{i,t}) = \left(c_{i,t}^{\frac{\sigma-1}{\sigma}} + \omega \, h_{i,t}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \qquad \sigma > 0,\ \omega > 0,$$

where $h_{i,t} \in \{h_0, h_1\}$; with h_0 being the utility from housing services when renting and h_1 being the housing utility from owning a house. The budget constraint at time $t \neq T$ is

$$c_{i,t} + s_{i,t} + p_t(h_{i,t+1}) + m_t(h_{i,t}) = w_{i,t} + R_f s_{i,t-1} + p_t(h_{i,t}),$$

where $c_{i,t}$ denotes consumption in period t and $s_{i,t}$ is the savings of agent i in that period. Agent i's real wage is given by $w_{i,t} = w_t e^{\eta_{i,t}}$, where w_t is the aggregate wage and $\eta_{i,t}$ is an idiosyncratic wage shock. The agent also carries forward savings from the previous period $s_{i,t-1}$, which accrue at the risk-free rate R_f. The market price of housing in period t is denoted by p_t, while the rental price is m_t. The value of owning a house is defined as $p_t(h_{i,t}) = ⊮\{h_{i,t} = h_1\} \, p_t$, so that homeowners ($h_{i,t} = h_1$) hold real estate worth the market value p_t, whereas renters hold none. Similarly, the rental cost is defined as $m_t(h_{i,t}) = (1 - ⊮\{h_{i,t} = h_1\}) \, m_t$, so that non-homeowners pay rent m_t at t, while homeowners incur no rental cost. The budget constraint for the last period $t = T$ is different as the agent cannot retain any saving or house value, that is:

$$c_{i,T} + m_T(h_{i,T}) = w_{i,T} + R_f s_{i,T-1} + p_T(h_{i,T}).$$

For an agent i living in period t, the observed house price return is the product of aggregate house price growth and an idiosyncratic term:

$$r_{i,t} = \frac{p_t}{p_{t-1}} e^{\varepsilon_{i,t}} = \exp\Big(\gamma_0 + \gamma_1 t + \epsilon_t + \varepsilon_{i,t} \Big). \tag{3.1}$$

Thus, house prices grow exponentially with a deterministic trend (governed by γ_0 and γ_1) and an aggregate shock ϵ_t. Each individual i also observes an idiosyncratic shock $\varepsilon_{i,t}$. All shocks are assumed to be mean zero with finite variance.

Agents form expectations of the mean $\mu_{i,t}$ and standard deviation $\sigma_{i,t}$ of house price growth between t and $t+1$ in the following way:

$$\mu_{i,t} = \sum_{k=0}^{t} r_{i,k} w(k, \lambda, t), \tag{3.2}$$

$$\sigma_{i,t} = (\sum_{k=0}^{t} r_{i,k}^2 w(k, \lambda, t) - \mu_{i,t})^{0.5} \tag{3.3}$$

Where the weights $w(k, \lambda, t)$ follow Malmendier and Nagel [2011]:

$$w(k, \lambda, t) = \frac{(t+1-k)^\lambda}{\sum_{k'=0}^{t}(t+1-k')^\lambda}. \tag{3.4}$$

For each time $t < T$, household i with state $(s_{i,t-1}, h_{i,t})$ and beliefs $(\mu_{i,t}, \sigma_{i,t})$ chooses $(c_{i,t}, s_{i,t}, h_{i,t+1})$ taking prices $(p_t, m_t, w_{i,t})$ as given. The associated value function is:

$$V_t(s_{i,t-1}, h_{i,t}; \mu_{i,t}, \sigma_{i,t}) = \max_{h_{i,t+1} \in \{h_0, h_1\}} \max_{c_{i,t},\, s_{i,t}} \left\{ \frac{g(c_{i,t}, h_{i,t})^{1-\eta}}{1-\eta} + \beta\, \mathbb{E}_{i,t}\Big[V_{t+1}(s_{i,t}, h_{i,t+1}; \mu_{i,t+1}, \sigma_{i,t+1})\Big] \right\}, \tag{3.5}$$

subject to the budget constraint

$$c_{i,t} + s_{i,t} + p_t(h_{i,t+1}) + m_t(h_{i,t+1}) = w_{i,t} + R_f s_{i,t-1} + p_t(h_{i,t}), \tag{3.6}$$

where $\mathbb{E}_{i,t}$ denotes the subjective expectations of household i at time t.

At $t = T$ the agent cannot carry savings or housing wealth forward, so the problem reduces to

$$V_T(s_{i,T-1}, h_{i,T}) = \max_{c_{i,T}} \left\{ \frac{g(c_{i,T}, h_{i,T})^{1-\eta}}{1-\eta} \right\} \quad \text{s.t.} \quad c_{i,T} = w_{i,T} + R_f s_{i,T-1} + p_T(h_{i,T}) - m_T(h_{i,T}). \tag{3.7}$$

The agent can then optimally choose $(c_{i,t}, s_{i,t}, h_{i,t+1})$ for each period using backward induction.

At time t, let V_t^H be the value from choosing homeownership in the next period ($h_{i,t+1} = 1$) and V_t^R be the value from renting ($h_{i,t+1} = 0$). The household's value function $V_t = \max\{V_t^H, V_t^R\}$. Conditional on $\mu_{i,t}$ and $\sigma_{i,t}$, we define the ownership surplus as $\Delta V_t(\mu_{i,t}, \sigma_{i,t}) = V_t^H - V_t^R$, the household chooses homeownership if $\Delta V_t > 0$.

Building on this setup, we derive three predictions concerning the effects of $\mu_{i,t}$ (expected returns) and $\sigma_{i,t}$ (volatility) on the ownership surplus:

Prediction 3.1 (Expected Appreciation). A higher subjective expected house-price growth $\mu_{i,t}$ (holding $\sigma_{i,t}$ fixed) strictly increases the probability of homeownership.

Intuitively, using first order stochastic dominance, an increase in μ strictly increases agent's belief regarding the value function of homeownership V^H. Thus, higher beliefs about homeownership μ formed through positive past experiences increase homeownership.

Prediction 3.2 (Return Uncertainty). A higher subjective volatility $\sigma_{i,t}$ (holding $\mu_{i,t}$ fixed) strictly lowers the probability of homeownership.

By concavity of utility function U, larger subjective volatility σ decreases the value function V^H, as owners' wealth is exposed to house price risk. Hence, higher experienced volatility σ reduces the ownership surplus and predicts a lower probability of purchasing a house.

Prediction 3.3 (Life-Cycle Sensitivity). Let $t_1 < t_2 < T$. Then the marginal impact of expected appreciation and volatility on the ownership value difference satisfies

$$\frac{\partial}{\partial \mu_{i,r,t}} \Delta V_{t_1}(\mu_{i,r,t}) \ > \ \frac{\partial}{\partial \mu_{i,r,t}} \Delta V_{t_2}(\mu_{i,r,t}) \quad \text{and} \quad \left| \frac{\partial}{\partial \sigma_{i,r,t}} \Delta V_{t_1}(\sigma_{i,r,t}) \right| \ > \ \left| \frac{\partial}{\partial \sigma_{i,r,t}} \Delta V_{t_2}(\sigma_{i,r,t}) \right|.$$

That is, the effects characterized in Predictions 3.1 and 3.2 are stronger for younger households earlier in the life cycle.

Since housing is a long-lived asset and the value of owning depends on the stream of future returns it generates. Early in life, households expect to enjoy the benefits (or costs) of price appreciation and volatility over a longer horizon, so both the upside from higher expected returns and the downside from greater volatility weigh more heavily in their ownership decision. Later in life, the remaining horizon is shorter, so changes in beliefs about returns or volatility have a smaller impact on the continuation value of ownership. Thus, the sensitivity of the homeownership decision to both mean and volatility of housing experiences naturally declines with age.

Linearization of the homeownership decision. Building on these predictions, we linearize the impact of $\mu_{i,t}$ and $\sigma_{i,t}$ on the probability of homeownership to clearly illustrate how experienced returns and volatility shape housing decisions.

Define the ownership outcome $Y_{i,t+1} := \mathbb{1}\{h_{i,t+1} = h_1\}$. Let $P_{i,t}$ denote the probability of choosing $h_{i,t+1} = h_1$:

$$P_{i,t} := \Pr(Y_{i,t+1} = 1 \mid \mu_{i,t}, \sigma_{i,t}) = G\Big(\Delta V_t(\mu_{i,t}, \sigma_{i,t})\Big)$$

Assumption (regularity). The mapping $(\mu, \sigma) \mapsto \Delta V_t(\mu, \sigma)$ is continuously differentiable in a neighborhood of a reference point $(\bar{\mu}, \bar{\sigma})$.[2]

[2]This holds, for instance, when beliefs over next-period log price $g_{t+1} = \log(p_{t+1}/p_t)$ belong to a smooth

Following Predictions 3.1 and 3.2:

$$\frac{\partial \Delta V_t}{\partial \mu_{i,t}}(\bar{\mu},\bar{\sigma}) \equiv \partial_\mu \Delta V_t(\bar{\mu},\bar{\sigma}) \;>\; 0, \qquad \frac{\partial \Delta V_t}{\partial \sigma_{i,t}}(\bar{\mu},\bar{\sigma}) \equiv \partial_\sigma \Delta V_t(\bar{\mu},\bar{\sigma}) \;<\; 0.$$

With the regularity assumption, a first-order Taylor expansion of the conditional mean $P_{i,t} = G(\Delta V_t)$ around $(\bar{\mu},\bar{\sigma})$ gives

$$P_{i,t} = G\Big(\Delta V_t(\bar{\mu},\bar{\sigma})\Big) + g\Big(\Delta V_t(\bar{\mu},\bar{\sigma})\Big)\Big[\partial_\mu \Delta V_t(\bar{\mu},\bar{\sigma})\,(\mu_{i,t}-\bar{\mu}) + \partial_\sigma \Delta V_t(\bar{\mu},\bar{\sigma})\,(\sigma_{i,t}-\bar{\sigma})\Big] + R_{i,t},$$

where g is the derivative of G with respect to ΔV_t, $R_{i,t} = o(\|(\mu_{i,t}-\bar{\mu},\sigma_{i,t}-\bar{\sigma}\|)$ collects second and higher orders. Define the model-implied local coefficients:

$$\alpha_{i,t} \;:=\; g(\overline{\Delta V})\,\partial_\mu \Delta V_t(\bar{\mu},\bar{\sigma}) \;>\; 0, \qquad \beta_{i,t} \;:=\; g(\overline{\Delta V})\,\partial_\sigma \Delta V_t(\bar{\mu},\bar{\sigma}) \;<\; 0,$$

with $\overline{\Delta V} := \Delta V_t(\bar{\mu},\bar{\sigma})$, and set $a_{i,t} := G(\overline{\Delta V}) - \alpha_{i,t}\bar{\mu} - \beta_{i,t}\bar{\sigma}$. Then,

$$P_{i,t} \;=\; a_{i,t} \;+\; \alpha_{i,t}\,\mu_{i,t} \;+\; \beta_{i,t}\,\sigma_{i,t} \;+\; R_{i,t}, \tag{3.8}$$

with $\alpha_{i,t} > 0$ (positive effect from higher price return experience) and $\beta_{i,t} < 0$ (negative effect from higher price volatility experience), with $\alpha_{i,t_1} > \alpha_{i,t_2}$ and $|\beta_{i,t_1}| > |\beta_{i,t_2}|$ for $t_1 < t_2$ (higher sensitivity for younger individuals). In the population, the OLS coefficients of the linear projection of $Y_{i,t+1}$ on $(1,\mu_{i,t},\sigma_{i,t})$ equal $(a_{i,t},\alpha_{i,t},\beta_{i,t})$ when second-order terms are small locally, delivering the desired reduced form with signs disciplined by the model. As a result, we use the linearization of the model to test the predictions in Section 3.4.

The next section presents the administrative credit-panel dataset we compile, details its linkage to zip code level house price data and explains how we construct household-specific measures of experienced returns and volatility to test these predictions.

3.3 Data

3.3.1 Data Sources

To study the effect of housing return experience on purchasing decision, we need the house price index of the area an individual has lived in and the homeownership status of the individual. We obtain these data from the University of California, Consumer Credit Panel

location–scale family indexed by (μ,σ), and the Bellman operator is well-behaved under such parametric perturbations.

(UC-CCP) dataset, which is a comprehensive longitudinal dataset that contains consumer credit behavior and financial health. It includes anonymized data on a variety of credit-related metrics such as outstanding balances across different types of loans (including mortgages, auto loans, credit cards, student loans), payment history, bankruptcies, and foreclosures. Moreover, the dataset contains the living address of an individual to the zip code level; to generate house price experiences, we merge the zip code level house address of the individuals with the Zillow House Price Index (ZHVI) at the zip code level and weight the the experienced returns and variance of each individual according to the method outlined in Equations (3.9) to (3.11).

3.3.2 Summary Statistics

In this section, we provide summary statistics on our sample and document the geographic and temporal variation in house price returns and volatility that underpins our main experience measures.

Sample Demographics

In this section, we first describe the sample restriction rules and the process through which we arrive at the sample of individuals used in our regressions. We then provide various summary statistics on sample demographics, homeownership status, and credit profiles.

We start with UC-CCP's panel data, which originates from one of the three nationwide consumer reporting agencies. The UC-CCP data includes detailed information on various types of credit, including mortgages, credit cards, auto loans, student loans, home equity loans, and other credit products. The sample comprises anonymized credit records of a nationally representative 2% sample of US adult consumers with credit records along with 100% of Californians with credit histories. In this study, we use the 1% subsample of the 2% national sample of US adult consumers, which randomly selects one percent of individuals from the 2% U.S. population who have their credit information registered and tracks approximately 100,000 individuals per quarter.

We impose the following sample restriction rules to ensure data consistency and improve the reliability of our estimates. First, when an individual exits and re-enters the panel, we drop the individual after their first exit rather than interpolating or filling in the gap in the time series. The rationale for this restriction is that, once we lose access to the individual for some period, we lose the geographic location of the individual, so any accounts of their experiences of local house prices would have to be imputed, which tends to be imprecise. This step removes less than 1% of the observations. Next, we retain only individuals that appear for at least ten consecutive quarters following their initial appearance. This criterion ensures that we accurately capture changes in homeownership status and can compute a reliable measure of each individual's house price experience over a sufficient time horizon. After applying these restrictions, our final sample consists of over 10 million individual × quarter observations tracking 218,699 unique individuals from 2004 to 2024[3].

Table 3.1: Sample Demographics - Full Sample

(in 2023Q4)

	N	Mean	Std. dev.	Min	P25	Median	P75	Max
Demographics								
Age	101840	54.785	19.841	18	38	55	69	122
Years of education	79239	13.873	2.593	10	12	14	16	19
Gender(1=Male)	73807	0.505	0.500	0	0	1	1	1
Married(1=Yes)	55448	0.812	0.391	0	1	1	1	1
Has child (1=Yes)	79942	0.235	0.424	0	0	0	0	1
Lives in rich zipcode* (1=Yes)	102664	0.447	0.497	0	0	0	1	1
Homeownership & Mortgage								
Homeowner(1=Yes)	80809	0.536	0.499	0	0	1	1	1
Has open mortgage (1=Yes)	104781	0.283	0.451	0	0	0	1	1
Number of mortgages	104781	1.020	1.684	0	0	0	2	41
Other Credit Information								
Has open credit card (1=Yes)	104781	0.684	0.465	0	0	1	1	1
Credit card utilization (%)	89622	21.142	30.469	0	0	4.879	31.121	100
Has open auto loan (1=Yes)	104781	0.278	0.448	0	0	0	1	1
Has open student loan (1=Yes)	104781	0.146	0.353	0	0	0	0	1
Delinquent (1=Yes)	104781	0.043	0.204	0	0	0	0	1
Experience Parameters**								
Return (Zip code level)	104781	1.283	0.465	-2.904	0.992	1.202	1.494	4.528
Return (County level)	104781	1.259	0.416	-2.135	0.996	1.189	1.440	3.679
Return (State level)	104781	1.284	0.355	0.114	1.038	1.203	1.432	3.311
Variance (Zip code level)	104781	5.261	3.799	0.308	2.609	3.961	6.916	99.363
Variance (County level)	104781	4.494	3.008	0.499	2.360	3.404	6.063	78.182
Variance (State level)	104781	3.547	2.307	0.404	1.810	2.660	4.647	16.899

* A zip code is a rich neighborhood if it is in the top quartile by HPI among all U.S. zip codes.
** Experience parameters presented are calculated using $\lambda = 1$.

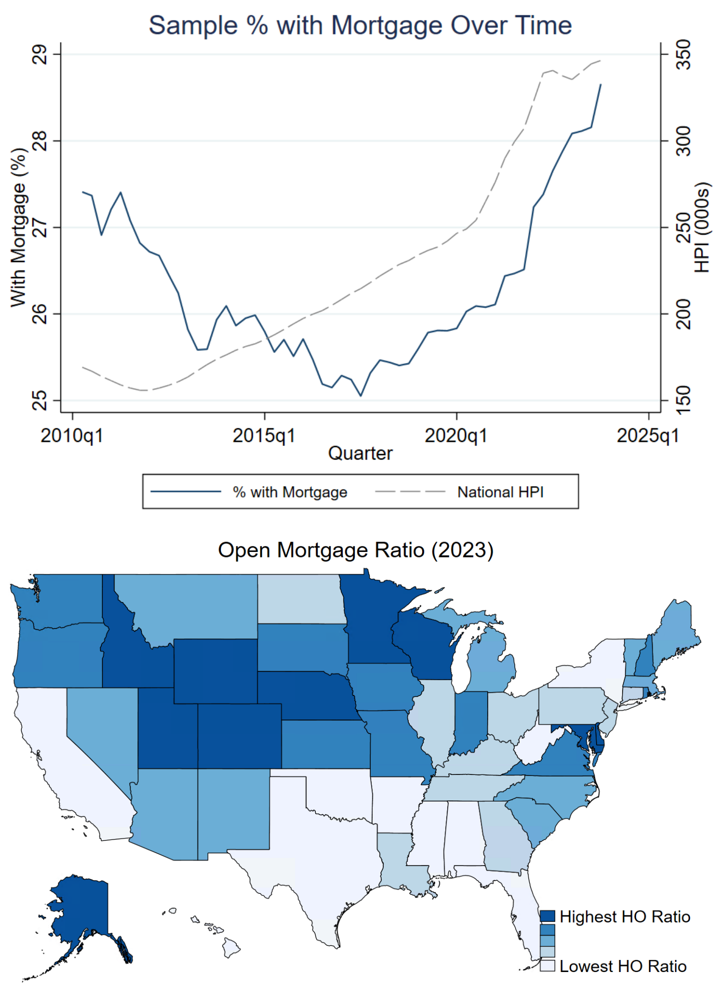
Table 3.1 presents summary statistics for the full sample in 2023Q4. The table reports individual demographics, homeownership and mortgage status, credit information, and experience parameters. Key metrics relevant to this chapter are homeownership and mortgage status: 53.6 percent of the sample are homeowners and 28.3 percent have an active mortgage. Overall, the sample displays significant variation in demographics, credit behavior, and housing market exposure. This heterogeneity highlights the importance of accounting for

[3] The panel is not balanced as individuals may enter and exit the sample in the middle of the sample period.

individual-level experiences when analyzing home-ownership decisions. The construction and interpretation of the experience parameters will be discussed in Section 3.3.3.

Following Agarwal et al. [2015], we identify individuals as homeowners if they have at least one open mortgage account. Figure 3.1 presents the evolution and geographical distribution of the open mortgage ratio in our sample. The top panel shows the percentage of individuals with an open mortgage over time, alongside the national House Price Index (HPI). The open mortgage ratio exhibits a downward trend from 2010 to around 2015, followed by a gradual increase through 2024. The open mortgage ratio remains relatively stable over the entire period, fluctuating within a narrow range of 25 percent to 29 percent.

Figure 3.1: Evolution and Geographical Distribution of Homeownership Rate



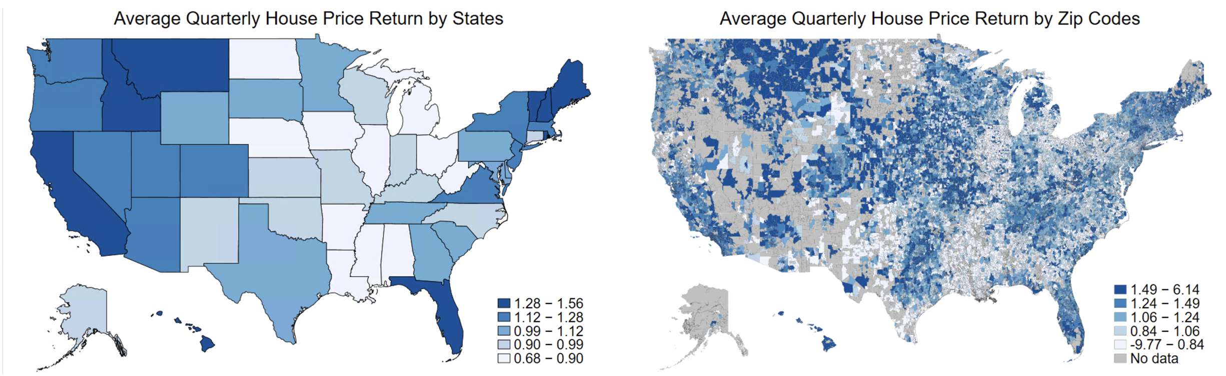
The bottom panel illustrates the geographical distribution of the open mortgage ratio in 2023 at the state level. The highest ratios are concentrated in parts of the Midwest, Mountain West, and Alaska, while the lowest ratios are observed in Texas, California, and several

Southern and Northeastern states. This pattern highlights substantial regional variation, with states in the northern and interior regions generally exhibiting higher open mortgage ratios compared to those in the South and parts of the East Coast.

House Price Returns and Volatility

Households reside in different locations at different times. We construct their house price experiences based on the house prices of the zip codes where they live. To ensure meaningful variation in our house price experience measures across individuals in the sample, we first document that both house price returns and volatility exhibit significant variation across geographic regions and over time. As a result, the house price experiences of sampled households also vary considerably, as we document below in Section 3.3.3.

Figure 3.2: Geographical Variation of House Price Returns



(a) State Level

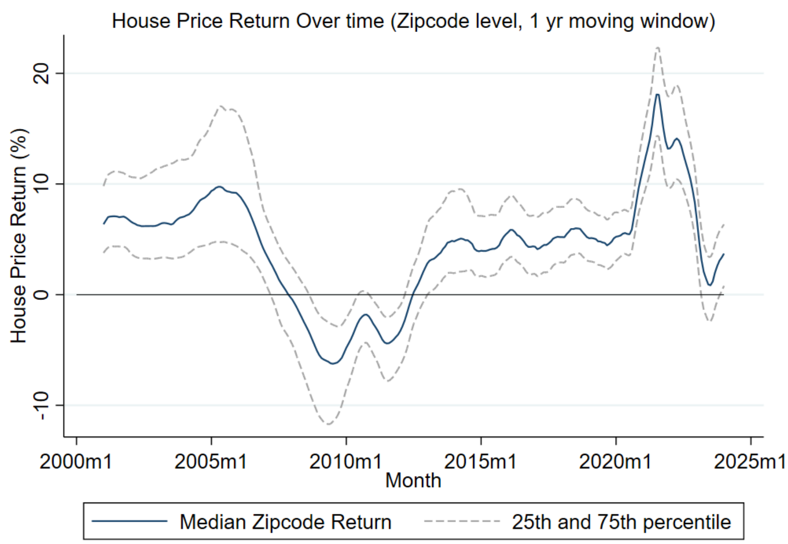
(b) Zip code Level

Figure 3.2 presents state and zip code level maps of average quarterly house price returns. At the state level, Western states such as Nevada, Idaho, and California, as well as parts of the South, including Texas and Florida, exhibit the highest house price returns over the decade. In contrast, Midwestern and Northeastern states, such as Illinois, Connecticut, and West Virginia, show relatively lower returns. This pattern aligns with broader economic trends, where states with rapid population growth and strong housing demand have seen greater appreciation in house prices.

At the zip code level, the variation becomes even more pronounced. Metropolitan areas tend to have a mix of both high- and low-return zip codes, reflecting the uneven nature of housing booms and gentrification. For instance, in California, urban centers such as San Francisco, Los Angeles, and San Diego contain zip codes with some of the highest returns, but also pockets of lower growth. Similarly, regions with robust economic expansion, such as parts of Texas and Florida, exhibit clusters of high-return zip codes. Meanwhile, rural areas generally display more muted house price growth.

Overall, these patterns suggest that local economic factors, migration trends, and housing supply constraints play a crucial role in shaping house price appreciation. The substantial within-state variation highlights the importance of analyzing house price returns at a finer geographic level to fully capture the heterogeneity in homeowners' experiences.

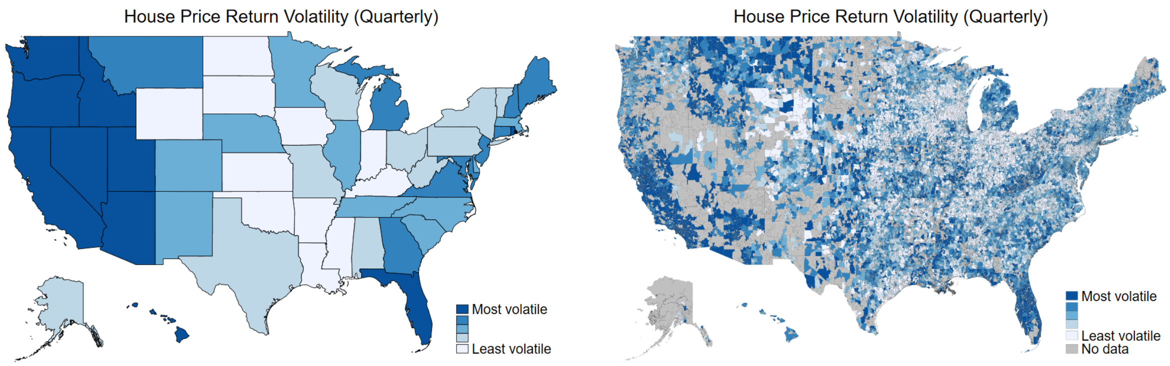
Figure 3.3: Zip Code Level House Price Returns Over Time



Over time, house price returns also fluctuate considerably. Figure 3.3 plots zip code-level returns from 2004 to 2023 using a one-year moving window. The distribution widens during periods of economic instability, highlighting temporal dispersion in housing market outcomes.

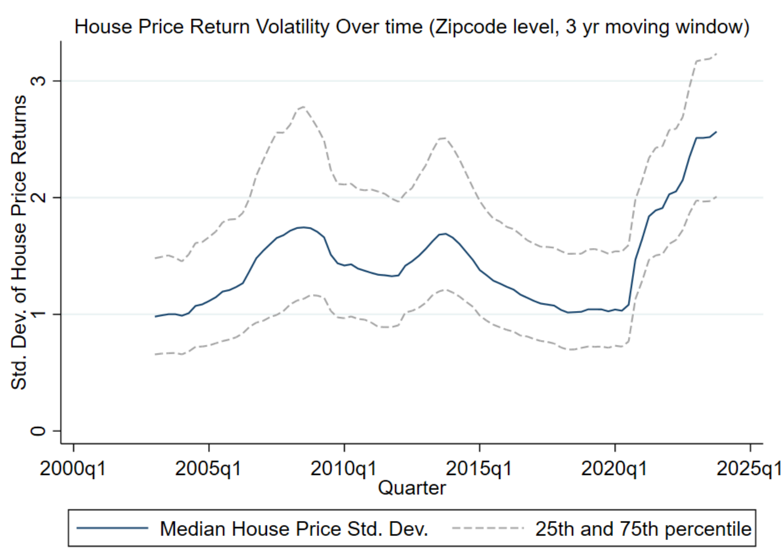
Next, we present a similar set of exhibits to demonstrate that the volatility of the house price return also varies significantly across regions and over time. Volatility is calculated as the standard deviation of quarterly house price returns. Plotting the state and zip code level volatility of house price returns on a map (Figure 3.4), we again find significant variation in house price return volatility not only across states but also within states and between adjacent zip codes. Figure 3.5 demonstrates that housing market volatility exhibits pronounced increases during major economic crises, such as the 2008 financial crisis and the COVID-19 pandemic, highlighting the sharp rise in uncertainty during these periods. Beyond these spikes, the figure reveals substantial fluctuations in volatility over time. Even within a given period, the interquartile range remains wide, indicating large cross-sectional differences in local housing market risk. These patterns underscore the presence of both temporal and regional heterogeneity in the volatility individuals experience, which plays a critical role in shaping their perceived risk and, consequently, their homeownership decisions.

Figure 3.4: Maptile of house price return volatility



House price returns and volatility vary markedly over time and across locations, generating substantial dispersion in individuals' lived experiences. Aggregate measures do not capture precisely the variation that drives beliefs and homeownership choice. By constructing individual-specific histories of local returns and volatility, we show that both dimensions matter for entry into homeownership.

Figure 3.5: Zip Code Level House Price Return Volatility Over Time



3.3.3 Experienced Returns and Volatility

In this section, we describe how we construct our measures of house price experience. Our main measure of housing returns experience for an individual is a weighted average of house price growth in the different zip codes where the individual has lived since age 18. Following

Equation (3.2), we calculate experienced returns as follows:

$$\text{EXP\_HR}_{i,t} = \sum_{k=0}^{t-n_{18}} r_{t-k} w(k, \lambda, t - n_{18}) \tag{3.9}$$

$$w(k, \lambda, t - n_{18}) = \frac{(t - n_{18} + 1 - k)^{\lambda}}{\sum_{k'=0}^{t-n_{18}} (t - n_{18} + 1 - k')^{\lambda}}. \tag{3.10}$$

This specification follows Malmendier and Nagel [2011]. A greater value of λ implies that more recent house price returns receive greater weight in the computation. We use $\lambda = 1$ as our baseline specification, but our results are robust to alternative choices of λ. Here, t denotes the time period in which the experience parameter is computed, n_{18} is the period when the agent turns 18, and k is an index running from 0 to $t - n_{18}$, i.e. the number of periods an agent experiences between age 18 and period t. Thus, the experience parameters incorporate the full history of house price returns observed by the individual.

A novel contribution of our analysis is that we also test for the effects of experienced volatility in house price returns on transitions into homeownership. Following Equation (3.3), we measure experienced volatility as:

$$\text{EXP\_SD}_{i,t} = \left(\sum_{k=0}^{t-n_{18}} r_{t-k}^2 w(k, \lambda, t - n_{18}) - (\text{EXP\_HR}_{i,t})^2 \right)^{1/2} \tag{3.11}$$

When we do not observe the specific zip code where an individual resides in a given quarter, or when zip code level house price data are unavailable, we impute returns using county level house price returns. If county level returns are unavailable, we rely on state level returns, and if those are also unavailable, we use national returns. This approach ensures that we make use of the best available geographical variation in house price returns.

In addition to computing experience measures at the zip code level, we also construct them at the county and state levels to examine how experience measures constructed at broader geographies affect homeownership. Summary statistics of the house price experience measures for the full sample across different geographic levels are presented in Table 3.1. The average quarterly house price return experience at the zip code level is 0.47%. The volatility experience at the zip code level is 3.80%.

3.4 Empirical Methodology and Results

In this section, we test the predictions of the model developed in Section 3.2. We find that households with higher lifetime experiences of house price returns at the zip code level exhibit a higher likelihood of becoming homeowners in the current quarter, whereas households who have experienced more volatile house price returns at the zip code level exhibit a lower likelihood. While return experiences at the county and state level also predict a higher likelihood of becoming first time homeowners, volatility matters only at the zip code level. Moreover, the effects of experienced returns and experienced volatility both decline with age, consistent with our life-cycle model.

At longer horizons, we find consistent evidence: households with higher return experiences are more likely to have a mortgage by the end of the sample, while households with greater exposure to zip code level volatility are less likely to hold a mortgage at the end of the sample. We also show that higher experienced returns lead to higher leverage at mortgage origination conditional on a first time entry into homeownership.

3.4.1 Local Housing Market Experiences and First-Time Homeownership

In order to study the effect of local housing market experiences on first-time homeownership motivated by Equation (3.8), we estimate the following equation:

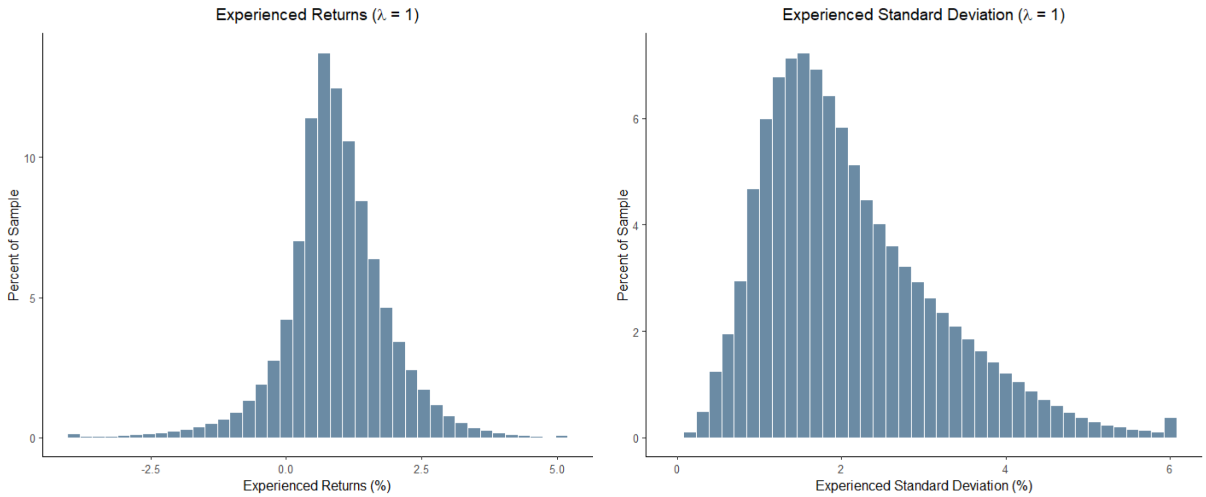
$$Y_{iact} = \alpha + \delta_{ct} + \delta_a + \beta_1 \text{EXP\_HR}_{it} + \beta_2 \text{EXP\_SD}_{it} + \mathbf{X}'_{it}\gamma + \epsilon_{it} \tag{3.12}$$

where the dependent variable Y_{iact} is 1 if individual i, who is a years old in county c at time t opens a mortgage for the first time in our sample. All specifications include county-by-time fixed effects (δ_{ct}) and age fixed effects (δ_a). We restrict the sample to individuals younger than 50, as the factors driving housing choices for older households are likely quite different from those relevant for younger first-time buyers. In our credit panel, we observe mortgage originations but not cash purchases. Since cash purchases account for less than 6% of first-time homebuyer transactions in the U.S.[4], our measure provides a close approximation of first-time homeownership. We further require that individuals are observed without a mortgage for at least 10 quarters prior to their first origination. Individuals are included in the panel until the quarter in which they appear with an open mortgage. The distributions of experienced returns and volatility used in these regressions are shown in Figure 3.6.

Our baseline results from estimating Equation (3.12) are in Table 3.2. Column (1), which includes only demographic controls (marital status, gender, number of adults, and number of children in the household), indicates that higher experienced returns are associated with a significantly greater likelihood of mortgage origination, whereas higher experienced volatility is associated with a significantly lower likelihood. To ensure that our results are not driven

[4]See a recent report by National Association of Realtors [2024].

Figure 3.6: Distribution of Experienced Returns and Volatility Constructed at the Zip Code Level



by momentum in local housing markets, Column (2) adds controls for the zip code level house price and four quarterly lags of zip code level house price growth in the individual's current zip code of residence[5]. Column (3) adds controls for the individual's current credit conditions. Specifically, we include the credit card balance-to-limit ratio as a measure of credit constraints, along with measures of outstanding student loans, auto loans, and other unsecured borrowing. We find that including controls for individual credit conditions raises the estimated coefficient on experienced returns. This suggests that the baseline relationship was partly attenuated by credit constraints, as households with weaker credit conditions are less able to act on stronger lifetime return experiences by taking out a mortgage. We also find that the coefficient on experienced volatility becomes less negative once credit conditions are included, indicating that part of its effect on homeownership operates through weaker credit conditions: higher experienced volatility is associated with worse credit conditions, which reduce the likelihood of mortgage origination. Nevertheless, this coefficient remains negative and statistically significant. Columns (4) and (5) add controls for the individual's education and occupation.[6].

Across all specifications, higher experienced returns are associated with a greater likelihood of first-time mortgage origination in the current quarter, while higher experienced volatility is associated with a lower likelihood. In terms of magnitudes, the estimates in Column (5) imply that a 1 p.p. increase in experienced returns raises the probability of transition into

[5]Our results are robust to including 8 or 12 quarterly lags as well.

[6]The data include roughly 10 education categories and 50 occupation categories.

Table 3.2: Baseline Transition to Homeownership Regressions, $\lambda = 1$

Dependent Variable:	First Time Homeowner in Current Quarter				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Experienced Returns ($\lambda = 1$)	0.064**	0.075**	0.093***	0.106***	0.102***
	(0.028)	(0.030)	(0.030)	(0.030)	(0.030)
Experienced Std. Dev. ($\lambda = 1$)	-0.263***	-0.251***	-0.193***	-0.152***	-0.145***
	(0.037)	(0.034)	(0.030)	(0.026)	(0.026)
Controls					
Demographics	✓	✓	✓	✓	✓
Zip-Level Prices		✓	✓	✓	✓
Credit Vars			✓	✓	✓
Fixed-effects					
Quarter- County	✓	✓	✓	✓	✓
Age	✓	✓	✓	✓	✓
Education				✓	✓
Occupation					✓
Fit statistics					
Observations	1,753,005	1,558,198	1,552,507	1,552,507	1,552,507
R^2	0.092	0.090	0.094	0.094	0.095
Within R^2	0.002	0.003	0.006	0.006	0.006

Table 3.2 reports estimates of Equation (3.12) for $\lambda = 1$ for different sets of controls across columns. Demographics include controls for gender, marital status, and number of adults and kids in the family. Zip code level prices include controls for the level of house prices at the zip code level, along with 4 lagged quarters of zip code level house price growth. Column (3) includes controls for the individual's credit conditions, which include outstanding credit card debt, student debt, auto loans and other unsecured borrowing. Columns (1)-(3) include county-time fixed effects together with age fixed effects. In column (4) we include fixed effects for the individual's level of education and column (5) includes fixed effects for the individual's current occupation. Standard Errors are clustered by quarter, age and county. Significance levels: *** $\Rightarrow$ p $<$ 0.01, ** $\Rightarrow$ p $<$ 0.05, * $\Rightarrow$ p $<$ 0.1.

homeownership by about 0.1 p.p., whereas a 1 p.p. increase in experienced volatility lowers it by about 0.15 p.p. Given that the average likelihood of becoming a homeowner in a given quarter is roughly 1.5 percent, these estimates imply that a 1 percentage point increase in experienced returns raises the transition probability by about 7 percent relative to the baseline, while a 1 percentage point increase in experienced volatility reduces it by about 10 percent.[7]

In order to study how experiences translate into homeownership decisions across the individual's life cycle, we re-estimate (3.12) separately for four age bins (20–30, 30–40, 40–50, 50–60), holding fixed the control set (demographics, zip code level prices and rents, and credit variables) and fixed effects (county×quarter, education, occupation, and age). Figures 3.7 and 3.8 report the coefficients on experienced returns (β_1) and volatility (β_2), respectively.

[7]The standard deviations of experienced returns and experienced volatility are each close to one percentage point, so the magnitudes reported above are approximately equal to the effect of a one-standard-deviation change in these variables.

Experienced returns significantly raise the transition into homeownership for households aged 30–50, while the coefficients for the 20–30 and 50–60 bins are statistically indistinguishable from zero.

Figure 3.7: Effects of Experienced Returns on Transition into Homeownership by Age

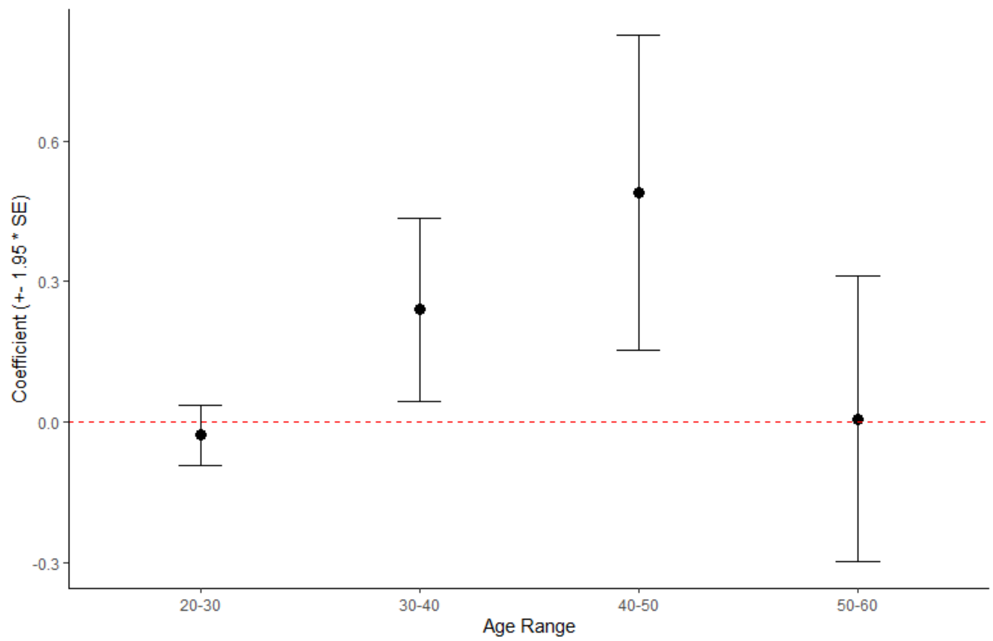


Figure 3.7 reports estimates of β_1, the coefficient on experienced returns from Equation (3.12) for $\lambda = 1$ separately for each age category from 20-60 in 10-year age bins. We include demographic controls, controls for zip code level house prices and rents, and individual credit conditions. We also include age fixed effects and county-time fixed effects. Standard Errors are clustered by quarter, age and county. The coefficient estimate together with the 95% confidence intervals are reported.

This life-cycle pattern is consistent with the Prediction 3.3 in Section 3.2: experience-based beliefs shift expected appreciation, but the translation of those beliefs into purchase decisions depends on both the horizon and constraints. For the youngest group, tighter liquidity/down-payment constraints mute the impact of optimistic beliefs. Because older households face a shorter remaining holding horizon, the present value of capital gains from expected house–price appreciation under ownership is attenuated, weakening the return channel. The fact that the return effect peaks in mid-life reinforces our interpretation that experienced returns affect ownership by shaping expectations about future price growth.

Figure 3.8 shows the effects of experienced volatility on entry into homeownership throughout the life cycle. We observe that, consistent with our life-cycle model, our results in Table 3.2 are driven by younger households, for whom we estimate a negative and precise coefficient. A natural concern with our interpretation of these results is that, despite our rich set of controls

Figure 3.8: Effects of Experienced Volatility on Transition into Homeownership by Age

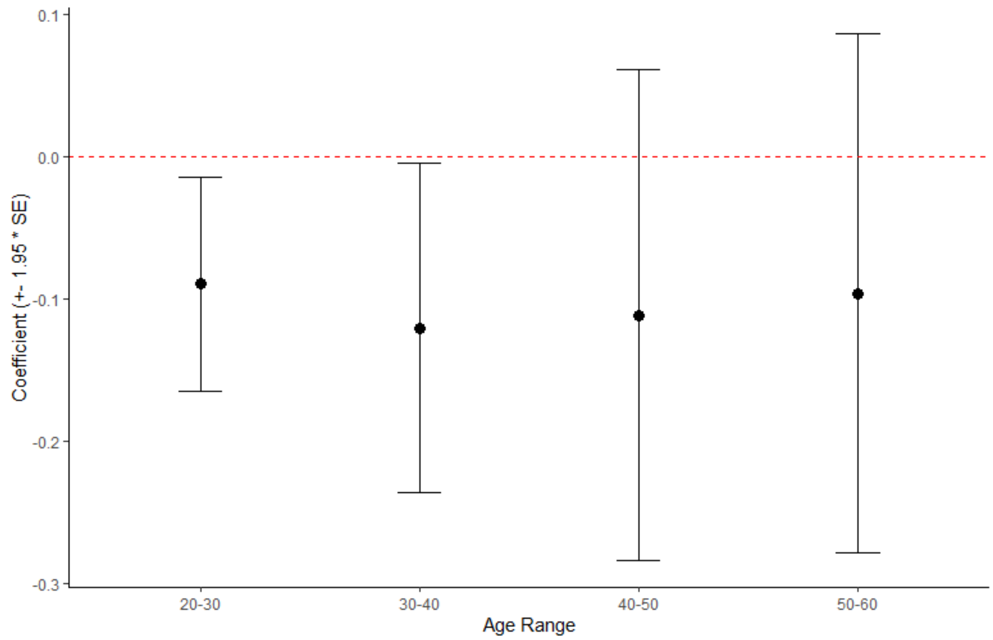


Figure 3.8 reports estimates of β_2, the coefficient on experienced standard deviation from Equation (3.12) for $\lambda = 1$ separately for each age category 20-60. We include demographic controls, controls for zip code level house prices and rents, and individual credit conditions. We also include age fixed effects and county-time fixed effects. Standard Errors are clustered by quarter, age and county. The coefficient estimate together with the 95% confidence intervals are reported.

and fixed effects, higher experienced volatility could be correlated with worse unobserved borrower quality (e.g., lower permanent income or tighter unmeasured liquidity). If this were the case, the volatility effect would be expected to fade sharply as financial positions improve with age; instead, coefficients remain negative even for the 50-60 cohort (albeit less precisely estimated).

Finally, Table 3.3 shows the effect of experience variables constructed at broader geographies on the individual's decision to become a first-time homeowner. Column (1) reports the zip code level specification with the full control set—demographics, zip code level prices and rents, and borrower credit conditions. To preserve variation when experiences are aggregated to the county level, Column (2) replaces county×quarter fixed effects with state×quarter and county fixed effects. For state level experiences, Column (3) includes quarter and state fixed effects. All columns include controls for the individual's education and occupation. Across all specifications, the coefficient on experienced returns is positive and statistically significant, indicating that our results on experienced returns are not driven by house price momentum at the zip code level. By contrast, the coefficient on experienced volatility is negative and significant only when experiences are constructed at the zip code level. This suggests that higher-order moments of past returns are particularly salient at the zip code level.

Table 3.3: TRANSITION TO HOMEOWNERSHIP REGRESSIONS, $\lambda = 1$, ZIP CODE, COUNTY AND STATE LEVEL

Dependent Variable:	First Time Homeowner in Current Quarter		
Model:	(1)	(2)	(3)
Variables			
Experienced Returns, Zip Code ($\lambda = 1$)	0.1151***		
	(0.0311)		
Experienced Std. Dev., Zip Code ($\lambda = 1$)	-0.1257***		
	(0.0245)		
Experienced Returns, County ($\lambda = 1$)		0.1259***	
		(0.0415)	
Experienced Std. Dev., County ($\lambda = 1$)		-0.0195	
		(0.0385)	
Experienced Returns, State ($\lambda = 1$)			0.1439***
			(0.0404)
Experienced Std. Dev., State ($\lambda = 1$)			0.0472
			(0.0537)
Controls			
Demographics	✓	✓	✓
Zip-Level Prices	✓	✓	✓
Credit Vars	✓	✓	✓
Zip-Level Rents	✓	✓	✓
Fixed-effects			
Age	✓	✓	✓
Quarter- County	✓		
Education	✓	✓	✓
Occupation	✓	✓	✓
Quarter-State		✓	
County		✓	
Quarter			✓
State			✓
Fit statistics			
Observations	1,552,507	1,552,507	1,552,507
R^2	0.09499	0.01818	0.01126
Within R^2	0.00598	0.00589	0.00625

Table 3.3 reports estimates of Equation (3.12) for $\lambda = 1$, including experiences at the county and state level in addition to zip code level experiences. Demographics include controls for gender, marital status, and number of adults and kids in the family. Zip code level prices include controls for the level of house prices at the zip code level, along with 4 lagged quarters of zip code level house price growth. "Credit Vars" includes controls for the individual's credit conditions, which include outstanding credit card debt, student debt, auto loans and other unsecured borrowing. Zip code level rents include controls for the current zip code level rent and the year-on-year growth of rents from the ACS. Standard Errors are clustered by quarter, age and county. Significance levels: *** $\Rightarrow$ p < 0.01, ** $\Rightarrow$ p < 0.05, * $\Rightarrow$ p < 0.1.

3.4.2 Long Horizon Regressions

We next ask whether local housing-market experiences also predict homeownership at longer horizons. For this exercise, we keep one observation per person from the last quarter that we observe them in the panel - and estimate:

$$Y_{iac} = \alpha + \delta_c + \delta_a + \beta_1^L \text{EXP\_HR}_{it} + \beta_2^L \text{EXP\_SD}_{it} + \mathbf{X}_i' \gamma^L + \epsilon_i \tag{3.13}$$

where the outcome variable equals 1 if, at the end of the sample, the individual has an open mortgage (our proxy for homeownership status). We restrict the sample to ages 30–60, for whom mortgages are less likely to be fully repaid. We include controls for demographics, current zip code level house prices and rents, and credit conditions. We also include county, age, education and occupation fixed effects. We estimate our regressions at the county, zip code and state level. The results from this exercise are given in Table 3.4.

Column (1) (zip code level experiences) yields a positive but imprecisely estimated coefficient in the linear probability model. The weaker zip code level precision in the long-horizon regressions reflects two compounding forces: (i) greater measurement error in the zip code level price index, and (ii) the lack of rich time-varying controls (e.g., county×quarter fixed effects) that, in the panel, soak up much of this noise. By contrast, Columns (2) and (3) (county and state level experiences) deliver positive and statistically significant coefficients. These estimates imply that a 1 p.p. increase in experienced returns is associated with roughly a 2.9–3.7 p.p. higher probability of being a homeowner at the end of the sample.

Turning to volatility, the long-horizon results in Table 3.4 are similar to the short-horizon patterns that we observed in Table 3.3: experienced volatility has a negative, significant effect on the probability of having a mortgage open at the end of the sample only when constructed at the zip code level. In terms of magnitudes, a 1 p.p. increase in experienced volatility lowers the probability of being a homeowner at the end of the sample by about 3.2 p.p. By contrast, the county and state level volatility coefficients in Columns (2) and (3) are small and imprecise. Taken together, we observe that higher experienced volatility of housing returns reduces the likelihood of homeownership only when experiences are constructed at the zip code level.

3.4.3 Effect of Experiences on Leverage at the Time of Purchase

Finally, we study how zip code level experienced returns and volatility affect leverage at the time of mortgage origination. For each individual i, we retain a single observation: the quarter t in which they first become a homeowner in our sample.

Table 3.4: Homeownership Status Regressions, $\lambda = 1$, End of Sample

Dependent Variable:	Homeownership Status, End of Sample		
Model:	(1)	(2)	(3)
Variables			
Experienced Returns, Zip Level ($\lambda = 1$)	1.008		
	(0.9232)		
Experienced Std. Dev., Zip Level ($\lambda = 1$)	-3.179***		
	(0.4179)		
Experienced Returns, County Level ($\lambda = 1$)		2.858***	
		(0.8765)	
Experienced Std. Dev., County Level ($\lambda = 1$)		-0.5860	
		(0.4940)	
Experienced Returns, State Level ($\lambda = 1$)			3.746***
			(1.169)
Experienced Std. Dev., State Level ($\lambda = 1$)			-0.2307
			(0.7389)
Controls			
Demographics	✓	✓	✓
Zip Code Price Level	✓	✓	✓
Credit Vars	✓	✓	✓
Fixed-effects			
Age	✓	✓	✓
County	✓	✓	✓
Education	✓	✓	✓
Quarter	✓	✓	✓
Occupation	✓	✓	✓
Fit statistics			
Observations	61,371	61,371	61,371
R^2	0.20255	0.20161	0.20160
Within R^2	0.09198	0.09090	0.09090

Table 3.4 reports estimates of Equation (3.13) for $\lambda = 1$, including experiences at the county and state level in addition to zip code level experiences. Demographics include controls for gender, marital status, and number of adults and kids in the family. Zip code level prices include controls for the level of house prices at the zip code level. "Credit Vars" includes controls for the individual's credit conditions, which include outstanding credit card debt, student debt, auto loans and other unsecured borrowing. Standard Errors are clustered by quarter, age and county. Significance levels: *** $\Rightarrow$ p $<$ 0.01, ** $\Rightarrow$ p $<$ 0.05, * $\Rightarrow$ p $<$ 0.1.

$$LZV_{it} = \alpha + \delta_{ct} + \delta_a + \beta_1^{\text{LEV}}\,\text{EXP\_HR}_{it} + \beta_2^{\text{LEV}}\,\text{EXP\_SD}_{it} + \mathbf{X}_{it}'\boldsymbol{\gamma}^{\text{LEV}} + \varepsilon_{it}^{\text{LEV}}. \quad (3.14)$$

Here, LZV_{it} is the log of the mortgage origination amount divided by the contemporaneous zip code level house price; δ_{ct} are quarter-by-county fixed effects and δ_a are age fixed effects. The covariate vector $\mathbf{X}_{it}$ is added cumulatively across columns as shown in Table 3.5: demographics are included in all columns; current zip code level prices and rents (and their lagged growth rates) enter starting in column (2); credit covariates enter from column (3); education fixed effects are added in column (4); and occupation fixed effects are added in column (5). EXP_HR_{it} and EXP_SD_{it} denote experienced returns and experienced volatility at the zip code level (constructed with $\lambda = 1$).

Table 3.5: EFFECT OF EXPERIENCES ON LEVERAGE CHOICE AT THE TIME OF MORTGAGE ORIGINATION

Dependent Variable:	Log (Mortgage Origination Amount / Zip Level House Price)				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Experienced Returns, Zip Code ($\lambda = 1$)	3.769***	2.578*	3.270**	4.019***	3.958***
	(1.204)	(1.318)	(1.301)	(1.335)	(1.373)
Experienced Std. Dev., Zip Code ($\lambda = 1$)	0.9892	-1.100	-0.3865	0.6539	0.9157
	(1.433)	(1.248)	(1.241)	(1.249)	(1.229)
Controls					
Demographics	✓	✓	✓	✓	✓
Zip-Level Prices		✓	✓	✓	✓
Zip-Level Rents		✓	✓	✓	✓
Credit Vars			✓	✓	✓
Fixed-effects					
Quarter–County	✓	✓	✓	✓	✓
Age	✓	✓	✓	✓	✓
Education				✓	✓
Occupation					✓
Fit statistics					
Observations	25,126	23,538	23,495	23,495	23,495
R^2	0.73528	0.74670	0.74949	0.75364	0.75641
Within R^2	0.02367	0.05276	0.06284	0.06784	0.06794

Table 3.5 reports estimates of Equation (3.14) for zip code level experiences, with $\lambda = 1$. Demographics include controls for gender, marital status, and number of adults and kids in the family. Zip code level prices include controls for the level of house prices at the zip code level, in addition to lagged zip code level growth rates. zip code level rents control for median rents and the year-on-year growth rate in rents from the ACS. "Credit Vars" includes controls for the individual's credit conditions, which include outstanding credit card debt, student debt, auto loans and other unsecured borrowing. Standard Errors are clustered by quarter, age and county. Significance levels: *** $\Rightarrow$ p $<$ 0.01, ** $\Rightarrow$ p $<$ 0.05, * $\Rightarrow$ p $<$ 0.1.

Across specifications in Table 3.5, the coefficient on EXP_HR_{it} is positive and statistically significant. The estimates imply that a one p.p. increase in experienced returns is associated with a 2.6%−4.0% increase in the mortgage-to-price ratio at origination. This pattern is difficult to reconcile with the view that our earlier homeownership results reflect unobserved income or financial capacity: if higher experienced returns simply picked up greater permanent income or better credit, we would expect lower leverage (larger down payments) at origination. Moreover, the estimates remain stable as we progressively add controls and fixed effects across columns, which proxy for, and partially absorb, differences in income and borrowing capacity. The coefficient on experienced volatility in Table 3.5 is small and statistically indistinguishable from zero.

3.5 Conclusion

This chapter highlights the important role of personal experiences in local housing markets in shaping homeownership choices of households. We show that these experiences predict entry into homeownership in the directions implied by our model: higher experienced returns are associated with a greater likelihood of first-time entry into homeownership, whereas higher experienced volatility is associated with a lower likelihood. These patterns are robust to rich demographics, zip code level prices and rents, borrower credit conditions, and extensive fixed effects (county×quarter, age, education, and occupation). The age profile aligns with our life-cycle model: effects attenuate with age and are most evident for younger households. Geographic aggregation matters: returns predict transitions at the zip code, county, and state levels, but volatility matters only when these experiences are constructed at the zip code level. Consistent qualitative patterns hold in longer-horizon specifications, and conditional on entry into homeownership, higher experienced returns are also associated with greater leverage at origination.

Our study complements work showing that past house prices inform expectations about future prices [Kuchler and Zafar, 2019, Armona et al., 2019, Agarwal et al., 2015, Adelino et al., 2018]. Case et al. [2014] hypothesize that relatively high long-run expectations before the financial crisis were formed due to having risen over decades. Although we do not measure expectations in this chapter, understanding how experiences shape long-run beliefs in the housing market is a promising direction for future research.

Heterogeneity in individuals' experiences—shaped by where they have lived—is essential to our analysis of experienced volatility in shaping homeownership decisions. The same approach could, in principle, be extended to rich data on individual portfolios to examine how cross-sectional differences in experienced volatility of stock returns translate into behavior in the stock market. Our results are primarily concerned with the extensive margin of housing demand: lived experiences that raise expected returns and lower perceived volatility increase the probability of entry into homeownership, while adverse experiences do the opposite.

A natural next step is to discipline a richer quantitative model with down-payment constraints and other realistic housing-market frictions to study how this extensive-margin channel feeds back into prices. Such a model would enable a focused analysis of how the entry of new homebuyers—shaped by optimistic long-run expectations—can propel house prices during booms. It would also provide a relatively understudied mechanism for slow recoveries: negative returns and elevated perceived volatility in house prices depress sentiment, reducing entry into homeownership, muting transactions, and slowing the recovery of house prices and economic activity.

The first two chapters of this dissertation have taken households' information and beliefs about their financial environment as given; the next chapter asks how households come to acquire the financial knowledge that underlies these decisions in the first place.

Chapter 4 Dollars and Sense: Self-Selection in Financial Literacy Education[1]

4.1 Introduction

The first two chapters of this dissertation studied how households make financial decisions given their existing beliefs and knowledge. A natural prior question is how they come to acquire that knowledge in the first place. Standard economic models often assume that individuals make investment decisions with complete information and rational expectations. As a result, models of educational attainment predict that individuals invest until marginal benefit equals marginal cost; namely, that individuals with the greatest potential benefits are most likely to invest in education [Willis and Rosen, 1979, Card, 1999]. Existing literature suggests this is also true with respect to financial literacy education; better educated, higher-income individuals may have the most to gain and therefore may be most likely to invest in acquiring financial knowledge [Lusardi et al., 2017].

However, several frictions may prevent individuals who *would* benefit from financial education from investing optimally. For example, it may be difficult for individuals to correctly value the benefits of financial literacy education or predict the costs of acquiring financial knowledge when their starting level of financial knowledge is low. Individuals with limited financial knowledge may also misjudge their starting level of financial literacy. This could lead to underinvestment in financial education due to overconfidence if they believe they already know everything they need to, or underconfidence if they incorrectly assume they cannot learn effectively. Finally, individuals with low starting levels of financial literacy may also lack access to education resources.

This chapter studies self-selection in financial literacy education – specifically, whether individuals correctly value financial education and choose to invest in it. First, we document selection on baseline ability in both observational data and in a lab setting. We show that individuals with higher levels of financial knowledge are more likely to invest in financial literacy education. Next, we use a lab experiment to measure demand for education and incentivize individuals to complete education within the experiment. This allows us to identify whether there is selection *into* education on learning gains or treatment effects. To evaluate whether new technologies can lower barriers to participating in financial literacy

[1]IRB of record: UC Berkeley Protocol #2024-04-17372 (Observational Data), 2024-03-17249 (Lab Experiment).

education, we vary the format and perceived cognitive difficulty of the education material offered. Experimentally varying the perceived cognitive cost of education allows us to better understand the mechanisms driving current selection patterns while exploring policy-relevant ways to change existing selection patterns.

We study self-selection into financial literacy education using two complementary approaches. The first uses rich observational data on UC Berkeley students to document selection into a large, highly rated introductory personal finance class. We combine survey measures of financial literacy with student records from 2010-2025 from the UC Berkeley Registrar's Office. We document selection on observables and on the intensive margin, the latter measured by the time students spend on the course website. This allows us to capture another previously unobserved aspect of selection: selection on attention.

To explore the mechanisms behind self-selection into financial education and to test for selection on treatment effects on a separate, more representative population, we move to a controlled lab setting. We conduct an online experiment with n=$1,047$ US participants on Prolific, a large online survey platform. In the experiment, each participant completes two sets of incentivized financial decision-making exercises. Between the two exercises, participants are given an opportunity to purchase different versions of financial education material and a portion of participants are randomly assigned to complete the education material. This allows us to precisely measure beliefs about ability and expected benefits from education. We then examine whether the true benefits from education are positively correlated with WTP, as standard models of investment in education would predict, or if individuals *mispredict* the benefits of education. Finally, we experimentally vary the perceived difficulty, or cognitive cost, of the education material to test whether new technology can lower barriers to entry and change existing selection patterns.

We document selection on observables into financial literacy education in both settings, among two distinct populations. We find that students who enroll in a large financial literacy class (UGBA 135) tend to be positively selected on family income, quantitative skills, and financial decision-making ability. Similarly, in the US Prolific population, individuals with the highest latent return to financial education have the lowest willingness to pay for education. Using measurements from the lab setting, we find that this is driven by systematic belief errors: low-ability individuals tend to underestimate the benefits from education and overestimate their starting level of knowledge. While technology-backed education solutions that lower the cost of education do increase demand compared to the traditional "classroom style" lecture education material, the AI and video options fail to alter selection patterns, indicating that additional behavioral barriers persist even when cognitive costs are reduced.

This chapter contributes to a large literature on mistakes in financial decision-making and financial literacy education. Given the increasingly complex financial landscape, there is growing concern about the ability of households to make sound financial decisions. Studies have shown that individuals with lower levels of financial literacy have lower rates of retirement savings [Mitchell and Lusardi, 2015], higher costs of borrowing [Huston, 2012, Lusardi and de Bassa Scheresberg, 2013], lower rates of asset accumulation, and lower rates of investment in stocks and other assets with high rates of return [Lusardi and Mitchell, 2023, van Rooij et al., 2011]. These studies also find that financial mistakes are more prevalent among low-income individuals and minorities, suggesting that disparities in financial literacy may exacerbate existing income and wealth gaps.

Despite the seemingly large benefits of financial literacy, several studies have pointed out difficulties encouraging participation in financial education [Bruhn et al., 2013, Lara Ibarra et al., 2017]. Furthermore, the individuals who take up financial literacy education are often *not* those who would ex ante have been expected to have the largest benefits. In fact, existing studies suggest that individuals who choose to invest in financial education are often *positively selected* along some dimension. For example, Meier and Sprenger [2013] find that individuals who choose to participate in a free credit counseling program tend to be more patient (less present-focused) than individuals who do not participate. Similarly, Bhattacharya et al. [2012] find that investors who opt into free investment advice tend to be more financially sophisticated—older, wealthier, and investors who hold more diversified, less home-biased portfolios—than those who do not. This chapter advances our understanding of selection by documenting both selection on baseline ability and selection on treatment effects.

Selection may limit the efficacy of financial literacy education even when education is mandatory. For example, financial literacy education mandates may be less effective if the students who pay attention or exert effort in the class are the ones who have a higher starting level of financial knowledge and therefore recognize the importance of acquiring additional knowledge. These selection patterns may also have distributional consequences, especially as states increasingly adopt high school financial literacy mandates and as financial literacy education becomes more prevalent on college campuses [Luedtke and Urban, 2023, Urban et al., 2020]. If high-income, high-literacy students disproportionately invest time and effort into improving their personal finances then state-mandated publicly funded financial literacy programs may actually contribute to increasing income inequality by teaching wealthy individuals how to grow their wealth.[2]

This chapter proceeds as follows. Section 4.2 describes the observational data used to identify selection on observables and presents preliminary results. Section 4.3 describes the

[2] Over 29 states in the US either currently have or plan to adopt a financial literacy mandate that requires students to take a financial literacy class before graduating from high school.

lab experiment used to decompose mechanisms. We then test for selection on treatment effects and decompose the mechanisms driving selection. Finally, Section 4.4 concludes and discusses areas for future research.

4.2 Observational Data

4.2.1 Data and Empirical Strategy

We use two complementary approaches to understand the mechanisms behind self-selection in financial education. The first combines observational data from two sources to document patterns of selection into a large personal finance class at UC Berkeley: Intro to Personal Finance (UGBA 135). The class is taught by Professor Terrance Odean and currently enrolls approximately 600 students per semester. The class is highly rated, available to any UC Berkeley student, and has no finance or math prerequisites. UGBA 135 has been offered since 2008 (and in its current form since 2013) and is often fully enrolled.

To build the dataset, we combine observational data from two sources: 1) course records from the UC Berkeley Registrar's Office and 2) survey data from a personal finance survey fielded through the Associated Students of the University of California (ASUC).

Data from the Registrar's office includes anonymized individual-level degree, admissions, and course data for all undergraduate entry cohorts from 2007-2025.[3] This includes college application data (parental education level, the applicant's self-reported family annual income, University of California Application Fee Waiver status, and intended major), degree date (major/GPA at graduation, graduation date), and demographic data (gender, race/ethnicity, and age), transfer student status, course enrollment data, and student's GPA for each semester.

We are also interested in whether attention or effort (although endogenously allocated) varies systematically across enrolled students. To control for attention, we include data on the total amount of time (in minutes) that each student spends on the course website as a proxy for attention and effort. Notably, time spent on the course website is an especially relevant measure of attention for this class: the weekly homework assignment for the class involves watching pre-recorded videos and completing online multiple choice quizzes on the course website. The instructors also regularly post recordings of the in-class lectures, lecture notes, and budgeting and other financial resources on the course website.

To measure financial literacy knowledge using similar measures to our lab setting, we partner with the ASUC Financial Literacy Task Force to administer a survey to students at

[3] Transfer student data is from 2016-2025.

UC Berkeley. In the survey we measure financial knowledge and stress and attitudes towards personal finances. To measure financial knowledge, we ask the Big Three financial literacy questions and have students answer three hypothetical decision-making questions. Finally, we ask students to provide their hypothetical willingness to pay or accept for UGBA 135.

The hypothetical decision-making questions ask students to make decisions about investing (index funds), debt, and compound interest in scenarios where there is a financially dominant choice. Students are told to maximize the amount of money they have in the scenario. While unincentivized, these scenarios give us a measure of financial decision-making ability against an objective benchmark and allow us to measure financial mistakes as the amount of money the participants lose by choosing a dominated option. While both the willingness to pay elicitation and the financial decision-making questions are unincentivized, these questions are similar to those used in our experiment, and give us another measure of selection patterns.

4.2.2 Observational Data Results – Selection on Baseline Ability

Standard human capital accumulation theories suggest a negative correlation between baseline ability and demand for education, as those with the lowest starting knowledge typically have the lowest marginal cost (i.e. opportunity cost) and highest marginal benefits from education. However, both our results from the observational data and lab setting contradict this prediction.

We first document selection into the personal finance class at UC Berkeley. Using individual-level data, we compare the demographic and major characteristics of UGBA 135, the large personal finance class at UC Berkeley, against the broader undergraduate population. We focus on demographic variables and proxies for income such as Pell Grant status and family education background.

We find strong evidence of positive selection on observables. As shown in Table 4.1, students who enroll in the financial literacy course are significantly different from non-enrollees across several dimensions. Specifically, enrollees are 12 percentage points more likely to come from quantitative majors (e.g., math, economics, and data science) (78% vs. 66%, $p < 0.01$), tend to have higher cumulative GPAs, and are significantly less likely to be underrepresented minorities (12% vs. 21%, $p < 0.01$). Furthermore, students from higher socioeconomic backgrounds appear more likely to invest in financial education; enrollees are less likely to be low-income or Pell Grant Eligible (proxied by eligibility for UC Fee Waivers) and are less likely to come from families without a college degree when compared to the broader population of Berkeley students.

	All Mean (SE)	Fin Lit Class Mean (SE)	No Fin Lit Class Mean (SE)	Difference Δ	Reg P-value
Male	0.45	0.48	0.45	0.03	0.04
	(0.01)	(0.01)	(0.01)		
White	0.26	0.22	0.27	-0.05	0.00
	(0.01)	(0.01)	(0.01)		
Asian	0.41	0.54	0.38	0.15	0.00
	(0.01)	(0.01)	(0.01)		
Underrepresented Minority	0.20	0.12	0.21	-0.10	0.00
	(0.00)	(0.01)	(0.01)		
Quantitative Majors	0.68	0.78	0.66	0.12	0.00
	(0.01)	(0.01)	(0.01)		
Cumulative GPA	3.54	3.60	3.53	0.07	0.00
	(0.00)	(0.01)	(0.01)		
Parents - No College	0.16	0.13	0.17	-0.03	0.01
	(0.00)	(0.01)	(0.01)		
Parents - No 4Yr College	0.24	0.20	0.25	-0.05	0.00
	(0.01)	(0.01)	(0.01)		
UC Fee Waiver	0.22	0.18	0.23	-0.05	0.00
	(0.01)	(0.01)	(0.01)		
N	6,639	1,234	5,405		

Table 4.1: UC Berkeley Student Characteristics by Financial Literacy Class - Class of 2023

Notes: Table 4.1 shows demographics of students from the 2023 graduating class who did and did not take a large personal finance class at UC Berkeley. Each row shows the results from a one-sample t-test comparing students who take a large personal finance class at UC Berkeley (UGBA 135) to students who do not take the class from the same graduating year. Negative numbers and p-values less than 0.10 are highlighted in red. Students who self-selected into the personal finance class tend to have higher cumulative GPAs and are more likely to have quantitative majors, less likely to be underrepresented minorities, less likely to be first-generation college students, and less likely to qualify for a Pell Grant or UC Fee Waiver (a proxy for low-income status).

Using data from the time spent on the course website, we also find evidence of selection on the intensive margin, or selection on attention. Again, we measure attention using the time (in minutes) spent on the course website. Unlike other courses that do not actively use their website, the majority of homework assignments for this class are completed on the course website. The instructors also regularly post recordings of in-class lectures, lecture notes, and other financial resources on the course website.

We find that even among the positively selected students who choose to enroll in UGBA 135, there are large differences in the amount of time students spend on the course website.

These differences also tend to be correlated with education outcomes – students who are above the median in time spent on the class website tend to earn a higher grade in the class (UGBA 135 GPA) and have higher GPAs in general (Term GPA). These differences highlight selection on attention as another important dimension and may help explain the limited efficacy of financial literacy mandates.

	All Mean (SE)	High Attention Mean (SE)	Low Attention Mean (SE)	Difference Δ	Reg P-value
UGBA 135 GPA	3.63	3.72	3.53	0.18	0.00
	(0.02)	(0.02)	(0.03)		
Term GPA	3.46	3.54	3.37	0.17	0.00
	(0.02)	(0.03)	(0.03)		
Male	0.51	0.48	0.55	-0.07	0.02
	(0.02)	(0.02)	(0.02)		
White	0.19	0.16	0.22	-0.06	0.03
	(0.01)	(0.02)	(0.02)		
Asian	0.52	0.56	0.48	0.08	0.01
	(0.02)	(0.02)	(0.02)		
Underrepresented Minority	0.17	0.16	0.18	-0.02	0.38
	(0.01)	(0.02)	(0.02)		
Quantitative Majors	1.00	1.00	1.00	0.00	.
	(0.00)	(0.00)	(0.00)		
Parents - No College	0.19	0.20	0.17	0.03	0.27
	(0.01)	(0.02)	(0.02)		
Parents - No 4Yr College	0.27	0.30	0.25	0.04	0.15
	(0.01)	(0.02)	(0.02)		
UC Fee Waiver	0.22	0.25	0.20	0.04	0.13
	(0.01)	(0.02)	(0.02)		
N	977	489	488		

Table 4.2: Student Characteristics by Attention in UGBA 135 - 2023-2024 Academic Year

Notes: Table 4.2 shows student characteristics by attention for students who took UGBA 135 during the 2023-2024 academic year. Attention is measured as the amount of time (in minutes) the student spent on the UGBA 135 course website, which includes time spent watching videos and completing comprehension check quizzes, watching lecture videos, and reviewing course notes and other materials. Students are in the high attention category if they are above the median in time spent on the course website. Students in the low attention category are below the median. Note that time spent on the course website may be missing for a small number of students. Negative numbers and p-values less than 0.10 are highlighted in red.

We also provide suggestive evidence of selection on financial decision-making ability. Using the ASUC survey data, we examine the relationship between the hypothetical willingness to pay (WTP) for the course and performance on hypothetical financial decision-making questions (e.g., maximizing returns in compound interest questions and debt repayment tasks). Figure 4.1 plots these results. We find that WTP is positively correlated with financial literacy scores: students with the highest demand for the course are arguably students who already possess high levels of financial knowledge. While the decisions in the survey are unincentivized, the correlation raises further concerns that students who stand to benefit the most from financial literacy education – students starting with lower levels of financial knowledge or from less privileged backgrounds – may not be the ones self-selecting into the class.

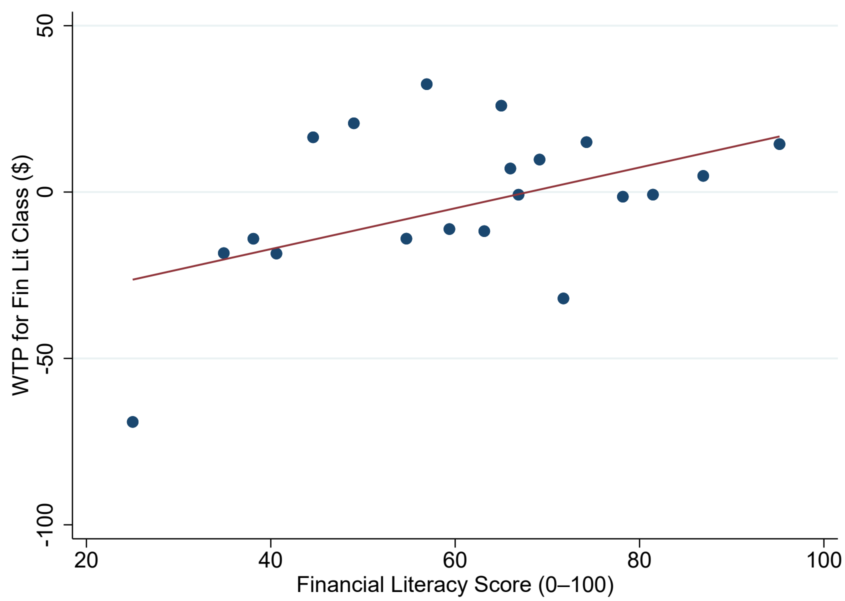


Figure 4.1: WTP for UGBA 135 versus Financial Decision-Making Score

Notes: Figure 4.1 shows the results from the ASUC survey conducted at UC Berkeley. We plot the relationship between the students' score on a set of financial decision-making questions and the hypothetical demand for UGBA 135. Both the decision-making questions and the questions are hypothetical and unincentivized, however the results provide additional suggestive evidence of the positive selection trend we observe in the observational data.

Taken together, the observational evidence suggests strong positive selection into financial education along observable characteristics and financial decision-making ability. However, these patterns are inherently descriptive. In the administrative data, demand for education is jointly determined by expected returns, perceived costs, institutional constraints, peer effects, and information frictions, making it difficult to isolate the underlying mechanism driving selection. In particular, we cannot observe individuals' beliefs about their own baseline ability or the returns to education, nor can we measure how these beliefs relate to actual learning outcomes. Our controlled lab experiment, described below in Section 4.3, helps us address these limitations and isolate the relationship between baseline ability, expected returns, and demand for education.

4.3 Lab Experiment

4.3.1 Data and Empirical Strategy

To demonstrate that the selection patterns we observe are not limited to classroom settings and to better understand mechanisms, we turn to our lab experiment. We start with a sample of $1,159$ Prolific participants who are 18 years or older and located in the United States. After excluding participants with crossover switching in the willingness-to-pay multiple price lists, who are flagged by our invisible text as having used AI to complete their written responses, and participants who have their education material choices implemented, we are left with a final sample of n=$1,047$ observations. Participants in our experiment have an average age of about 43 years old, are 66% white, and have an average income of about $65,000. The key demographic characteristics of our final sample are presented in Table 4.3.

	Obs.	Mean	SD	P(25)	P(75)
Age (yrs)	1,092	42.70	12.54	33.00	51.00
Male	1,092	0.43	0.50	0.00	1.00
income	1,092	5.41	2.30	4.00	7.00
Time (in minutes)	1,092	37.28	17.37	25.60	44.67
Race - White	1,092	0.66	0.47	0.00	1.00
Race - Hispanic	1,092	0.07	0.25	0.00	0.00
Race - Asian	1,092	0.17	0.37	0.00	0.00
Race - Black	1,092	0.06	0.25	0.00	0.00
Income ($ est)	1,092	65650.18	49693.96	30000.00	87500.00
College Degree	1,092	0.69	0.46	0.00	1.00

Table 4.3: Summary Statistics for Experiment Participants

Notes: Table 4.3 presents the summary statistics of the demographics of the participants in the lab experiment. Participants are recruited through Prolific, a large online survey platform commonly used by researchers. All participants must be 18 years or older and located in the United States. We exclude participants with crossover switching in the willingness-to-pay multiple price lists, participants who are flagged by our invisible text as having used AI to complete their written responses, and participants who have their education material choices implemented.

In the experiment, each participant completes two sets of incentivized financial decision-making exercises. In each financial decision-making exercise, participants choose between two financial options: a fixed dollar amount, or an investment in which the starting amount and growth rate are varied. We then elicit demand for each version of the education material. 10%

of participants have their choices implemented and 90% of participants are randomly assigned to one of the four experiment arms: 1) AI education, 2) Video education, 3) University Lecture education, or 4) control. Finally, all participants complete an endline set of financial decisions and are asked multiple choice questions about their experience with financial education and their use of AI in financial decisions. We also collect information on age, gender, race, and income. Figure 4.2 provides an overview of the experiment.

Demand for education, or willingness to pay (WTP) or willingness to accept (WTA), is measured using three Becker, Degroot, and Marschak (BDM) Mechanism [Becker et al., 1964] multiple price lists. Participants are given the choice between the following options:

A) Dollar amounts ranging from -$10 to +$10 + no education material, or

B) Dollar amounts ranging from -$10 to +$10 + complete the education material

If participants receive the education material, they must fully complete the material by passing attention checks and correctly answering three multiple choice comprehension check questions before they are able to move on in the experiment.

We then test whether these selection patterns are in part driven by the perception that traditional versions of financial literacy education may be unengaging or cognitively difficult. We offer three different formats of the education material: 1) a classroom style university lecture format, 2) a video education format, and 3) an interactive AI-powered learning assistant. The video and AI based learning formats allow us to test whether supply-side policies making education more accessible and less difficult can change selection patterns. All versions of the education material teach the exact same material; the Rule of 72, a simple mental shortcut for calculating the value of the investment options. We intentionally try to hold all versions of the education material constant by keeping the definitions, comprehension check questions, and explanations to the multiple choice questions exactly the same. We also emphasize to participants that all of the materials are based on the same content and vary only in presentation format. The university lecture format is framed as the most difficult version while the AI-powered learning assistant is framed as the most beginner friendly, interactive, and easy to understand version. We describe the video education version using similar language and framing as popular online learning platforms like Khan Academy and Coursera.[4]

We also measure each participant's beliefs about their starting or "pre" score (measured as their expected bonus), their expected bonus if they receive each version of the education material (their perceived benefit), and the reasons they prefer each version. Eliciting expectations also allows us to measure how overconfident participants are in their own

[4] https://www.khanacademy.org/college-careers-more/personal-finance/x12915ee2df79104f:welcome-to-personal-finance/x12915ee2df79104f:untitled-171/a/welcome-to-personal-finance-v2

knowledge. Finally, we ask the participants a set of five unincentivized multiple choice questions on compound interest from Ambuehl et al. [2022], as well as questions about their experience with financial literacy education, personal finances, and AI usage.

The treatment condition is assigned based on the actual WTP/WTA for 10% of participants. The remaining 90% are randomly assigned to one of the four treatment arms; 1) AI education material, 2) video education material, 3) university lecture education material, or 4) no education material. Randomly incentivizing participants with both positive and negative WTPs to receive financial literacy education allows us to identify whether the true benefits from education are positively correlated with demand, as standard models of investment in education would predict, or if individuals actually *mispredict* the benefits of education. Payment for the education bundle is implemented only if they receive a bonus for their endline decision. This ensures that the costs or benefits education are realized *only if* the participant receives the bonus. Table 4.4 shows the balance table for the four treatment arms.

	Control		AI Edu		Video Edu		Lecture Edu	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Male	0.45	(0.50)	0.44	(0.50)	0.42	(0.49)	0.39	(0.49)
Age	42.49	(12.48)	42.02	(12.23)	43.67	(12.39)	43.19	(12.73)
Time (mins)	28.37	(15.05)	41.58	(18.83)	39.52	(17.58)	39.81	(14.98)
Race - White	0.69	(0.46)	0.65	(0.48)	0.62	(0.49)	0.70	(0.46)
Race - Black	0.06	(0.23)	0.07	(0.25)	0.08	(0.27)	0.05	(0.22)
Income ($ est)	65694.44	(48295.28)	61785.71	(45480.35)	69755.64	(55056.89)	62659.18	(47097.10)
College	0.67	(0.47)	0.68	(0.47)	0.73	(0.45)	0.66	(0.47)
Observations	252		252		266		267	

Table 4.4: Balance Table by Treatment Arm

Notes: Table 4.4 presents the balance table of the key participant demographics by treatment arm. Treatment status is randomly assigned by the computer at the individual level.

One challenge the financial education literature has often faced is the question of how to measure financial literacy in objective, welfare-relevant terms. This stems from the fact that financial literacy encompasses a wide variety of skills, including numeracy, statistics and probability, attention to detail, knowledge of financial terms and definitions, and the ability to read and process complex financial contracts. Additionally, in many cases these shortcomings can be overcome through exerting effort (for example, looking up the concepts online, or following the advice of knowledgeable friends or experts). This raises the question of whether teaching financial literacy is simply increasing motivation or teaching individuals to exert effort. Finally, there is the question of how knowledge of financial concepts are translated into real-world behavior that improves welfare, as measured against an objective

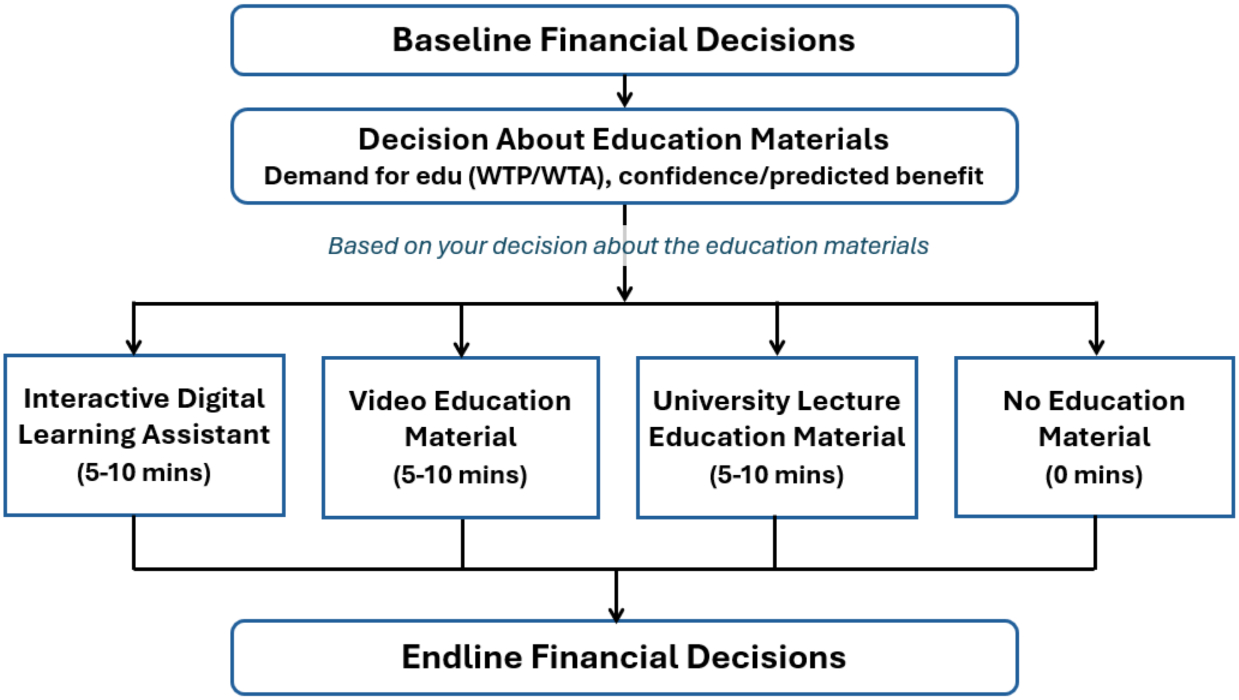


Figure 4.2: Lab Experiment Overview

Notes: The diagram in Figure 4.2 provides an overview of the experimental design for the lab experiment. Participants were first asked to answer a set of 10 baseline financial decision-making questions that test the participant's knowledge of the compound interest. We then ask participants to predict how they performed on the 10 baseline questions, elicit a willingness to pay for each version of the education material, and ask participants to predict how they would do on the endline questions *if they were to receive the specified version of the education material.* 90% of participants are then randomized into one of four experimental arms: the AI education material, video education material, lecture education material, or control (no education material). Finally, all participants answer a set of 10 endline financial decision-making questions, which test the same concepts and are very similar in structure and difficulty to the baseline questions.

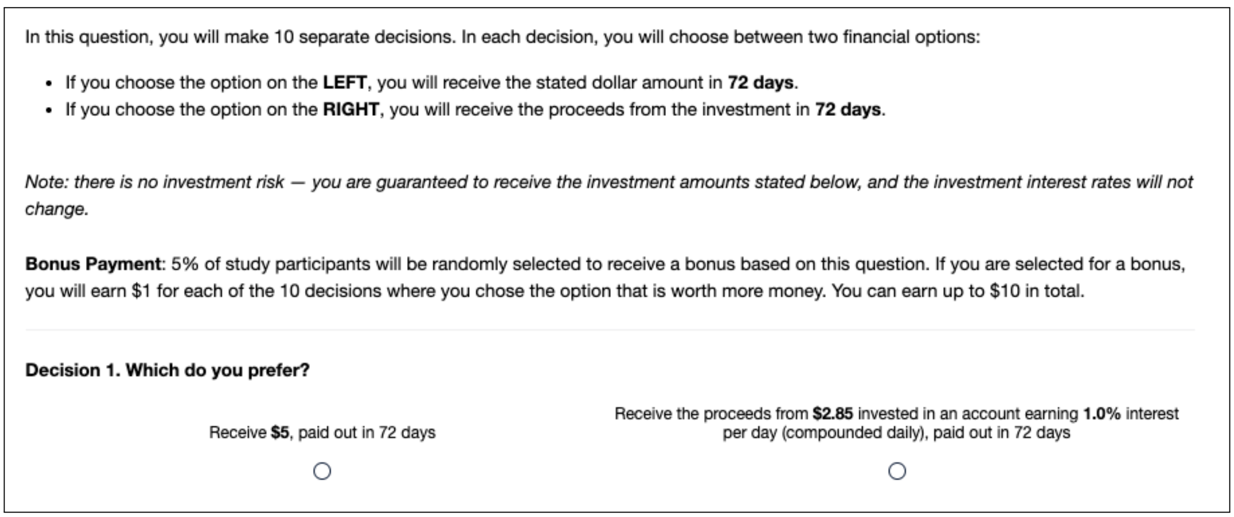
benchmark.[5] In fact, there is evidence that being able to *recall* financial concepts is not the same as being able to *apply* these concepts and translate them into improvements in financial decision-making [Ambuehl et al., 2022].

Acknowledging this difficulty, we choose to focus on what we believe is one of the core components of financial literacy: numeracy. Specifically, our questions test for an understanding of interest rates and compound interest. We focus on these concepts for two main reasons. The first is that we believe that quantitative principles are a core component of making good financial decisions. In fact, "The Big Three," the widely-used, standard measures of financial literacy developed by Lusardi and Mitchell [2011], are at their core

[5]Early papers focus on outcomes such as saving which may not be universally welfare improving. See Bernheim and Taubinsky [2018] for a discussion.

questions about numeracy.[6] The second benefit is that focusing on numeracy allows us to measure financial knowledge in a way that directly maps to monetary gains and losses. This allows us to measure and quantify the benefits of our financial education intervention using an objective, welfare-relevant metric.

Our question is based on the methodology developed in Ambuehl et al. [2022]. We ask participants to choose between two financial instruments: one that is a complexly-framed dollar amount and one that is a simply-framed dollar amount. Participants are told that there is a chance $1 will be added to their bonus for each decision where they choose the financial instrument that is worth more money in 72 days. Since there are ten questions in total, the potential bonus ranges from $0 to $10. We also tell participants that there is no interest rate or investment risk and no uncertainty. Figure 4.3 shows an example of the question.


In this question, you will make 10 separate decisions. In each decision, you will choose between two financial options:

- If you choose the option on the **LEFT**, you will receive the stated dollar amount in **72 days**.
- If you choose the option on the **RIGHT**, you will receive the proceeds from the investment in **72 days**.

Note: there is no investment risk — you are guaranteed to receive the investment amounts stated below, and the investment interest rates will not change.

Bonus Payment: 5% of study participants will be randomly selected to receive a bonus based on this question. If you are selected for a bonus, you will earn $1 for each of the 10 decisions where you chose the option that is worth more money. You can earn up to $10 in total.

Decision 1. Which do you prefer?

Receive **$5**, paid out in 72 days	Receive the proceeds from **$2.85** invested in an account earning **1.0%** interest per day (compounded daily), paid out in 72 days
○	○

Figure 4.3: Financial Decision-Making Question

Notes: Figure 4.3 shows the financial decision-making question used in the lab experiment. The question is based on a core concept in personal finance; understanding compound interest and exponential growth and is based on a similar question in Ambuehl et al. [2022].

Of the two financial instruments, there is one option that maximizes the amount of money the participant will receive at the end of the study. Since one option is framed as a compound interest calculations, participants should calculate the dollar value of the complexly framed option, thus testing their knowledge of compound interest. As in Ambuehl et al. [2022], the complexly framed amount can be calculated quickly using the Rule of 72, a mental shortcut

[6]Two of the Big Three questions test for an understanding of interest rates (one directly and one indirectly by asking about inflation). The third question asks about diversification, a statistical concept that is closely related to numeracy.

that allows you to quickly estimate how long it would take a dollar amount to double at a fixed interest rate. We then define the change in score (and educational benefit) as the endline score (0-10) minus the baseline score (0-10).

4.3.2 Results – Lab Experiment

Replicating Selection on Baseline Ability in the Lab

We first seek to replicate the selection patterns we see in the observational data in our lab setting. One potential concern is that a controlled setting may attenuate the selection mechanisms we observe in the real world. Beyond the ability to correctly predict the costs and benefits of education, real-world selection patterns may encompass factors such as present focus, uncertainty about the future, and concerns about the decay of knowledge that we control for in the lab. For example, we remove any uncertainty about the potential benefits of education by telling participants they will use the material immediately after (in the endline questions). We may also reduce uncertainty about the quality of the education material by providing it as a part of the experiment, essentially signaling its usefulness. Finally, we remove any concerns about knowledge decays over time, thereby increasing the potential benefits of the education material, by asking the endline questions immediately after the education material. Thus, it would be unsurprising if any selection patterns are more muted, or even non-existent, in a lab setting.

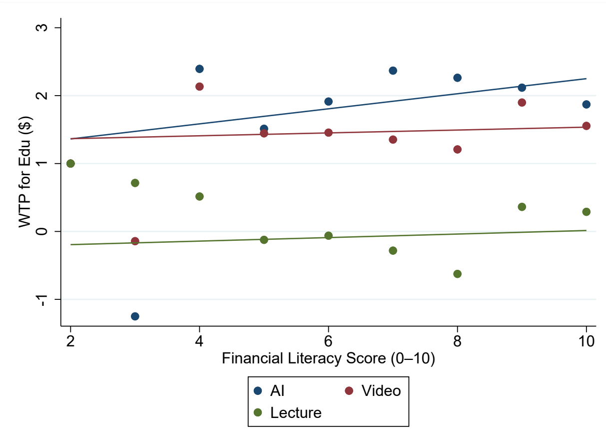
Despite these limitations, we still find evidence that demand for education is positively correlated with baseline financial knowledge in our lab setting. As illustrated in Figures 4.4(a) and 4.4(b), participants with higher scores on the baseline "Rule of 72" decision-making questions and five financial literacy multiple choice questions about compound interest exhibit higher WTPs for the education modules. Conversely, individuals with the lowest baseline performance—who theoretically stand to gain the most—report the lowest WTP.

To estimate the treatment effect of the education material, we run the following specification, which corresponds to Column (1) of Table 4.5:

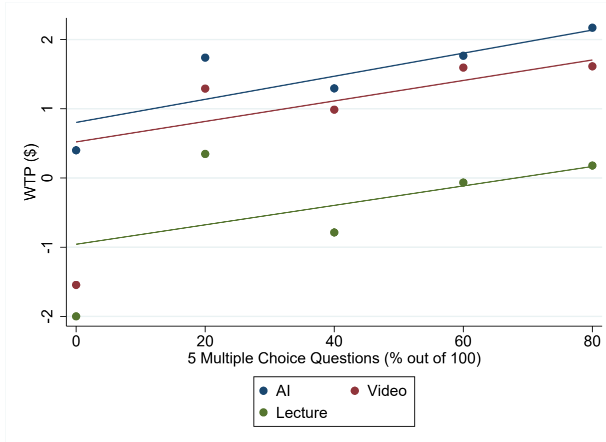
$$Y_{i,1} = \alpha + \beta_1 T_i + \beta_2 Y_{i,0} + \mathbf{X_i} + \varepsilon_i \tag{4.1}$$

where $Y_{i,1}$ is endline score, $Y_{i,0}$ is baseline score, T_i is an indicator for treatment status, and $\mathbf{X_i}$ is a vector of baseline demographic controls for participant i. Baseline demographic controls include college, gender, race, income, and age. Column (2) of Table 4.5 runs the same specification with an indicator for each education material type.

To examine treatment effect heterogeneity, we augment Equation 4.1 and estimate the



(a) Rule of 72 Questions



(b) Compound Interest Multiple Choice Questions

Figure 4.4: Selection on Baseline Ability

Notes: Figures 4.4(a) and 4.4(b) show selection on baseline financial decision-making ability. Figure 4.4(a) shows the relationship between WTP for each version of the education material and baseline financial decision-making ability, measured as the baseline or pre score on the Rule of 72 questions. The Financial Literacy Scores can range from zero to ten (one point for each question correct). Figure 4.4(b) shows the relationship between demand for education and the participant's score on five multiple choice questions about compound interest. These five questions are taken from Ambuehl et al. [2022] and are unincentivized. We elicit demand for each of the 3 different versions of the education material: 1) the AI version, 2) the video version, and 3) the university lecture version.

following specification, which corresponds to Column (3) of Table 4.5:

$$Y_{i,1} = \alpha + \beta_1 T_i + \beta_2 Y_{i,0} + \beta_3 H_i + \beta_4 (T_i \times H_i) + \varepsilon_i \tag{4.2}$$

where $Y_{i,1}$ is endline score, $Y_{i,0}$ is baseline score, T_i is an indicator for treatment status, H_i is the dimension of heterogeneity we are interested in, and $\mathbf{X_i}$ is a vector of baseline demographic controls for participant i. We look at treatment effect heterogeneity by education material and by baseline ability.

Table 4.5 reports treatment effects on endline test performance using an ANCOVA specification that controls for baseline ability. Across all specifications, assignment to AI, video, or lecture education leads to large and statistically significant improvements in test scores. Pooling across formats yields an average treatment effect of approximately 1.3 points (s.e. $= 0.117$). The coefficient on baseline performance indicates substantial persistence in test scores, while remaining well below one. Columns (5) and (6) show significant heterogeneity in treatment effects: the interaction between treatment assignment and baseline score is negative, implying that educational interventions generate larger gains for participants with lower initial ability. These results establish that the interventions are effective and that their benefits are concentrated among participants who start from a lower baseline.

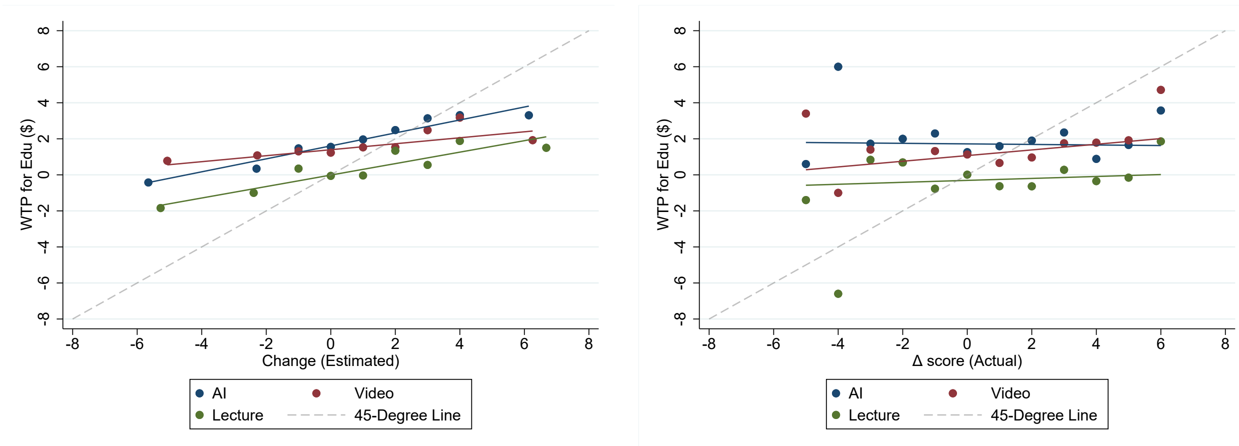
Mechanisms: Belief Errors and Misprediction

Another central prediction of human capital theory is that individuals have rational expectations and continue to invest in education while the marginal benefits (or the gains from education) equal the expected marginal costs. To test these patterns, we turn to a laboratory setting where our experimental design allows us to test the accuracy of these predictions by measuring key variables that we cannot observe in administrative datasets: financial ability, beliefs, and treatment effects.

	(1) Post Score	(2) Post Score	(3) Post Score
Pre Score	0.431***	0.429***	0.619***
	(0.028)	(0.028)	(0.050)
Any treatment = 1	1.311***		2.944***
	(0.117)		(0.400)
AI treatment = 1		1.454***	
		(0.150)	
Video treatment = 1		1.318***	
		(0.145)	
Lecture treatment = 1		1.171***	
		(0.153)	
Pre Score × Any treatment			-0.251***
			(0.059)
Constant	3.932***	3.939***	2.718***
	(0.301)	(0.301)	(0.396)
Observations	1047.000	1047.000	1047.000
R^2	0.237	0.239	0.246
Pre score control	X	X	X
Demographics controls	X	X	X

Table 4.5: Treatment Effects of Education Material

Notes: This table reports treatment effects on endline test performance using an ANCOVA specification. All columns regress the post-treatment test score on indicators for assignment to educational interventions, with the control group omitted. Column (1) pools all interventions into a single "Any treatment" indicator. Column (2) includes separate indicators for AI, video, and lecture treatments. Column (3) includes the specification in Column (1) but additionally allows treatment effects to vary with baseline test performance by interacting treatment assignment with the pre-treatment score. Baseline test scores are included in all specifications as "Pre Score." All columns additionally control for demographics (college education, gender, race, income, and age). Robust standard errors are reported in parentheses.



(a) WTP versus Predicted Change

(b) WTP versus Actual Change

Figure 4.5: WTP versus Predicted and Actual Change

Notes: Figures 4.5(a) and 4.5(b) show the relationship between participant demand (WTP) for the different types of education material versus their predicted and actual change in score. Figure 4.5(a) shows that demand is increasing in estimated or predicted benefit, confirming that the WTP elicitation is positively correlated with the perceived benefits from education, however Figure 4.5(b) shows that WTP is much less correlated with the *actual* change in score. This suggests that individuals have difficulty predicting the *true* benefits from education.

We first compare participants' willingness to pay (WTP) against both their predicted score improvement (perceived return) and their actual score improvement (realized return). Figure 4.5 shows the relationship between demand for financial education and the predicted benefits of education. The contrast between these two relationships is striking. As shown in Figure 4.5(a), demand for financial education is strongly and positively correlated with *predicted* benefits; participants who believe they will learn more are willing to pay more.

However, Figure 4.5(b) highlights the disconnect between beliefs and reality. While there is a positive relationship between expected benefit and demand, there is essentially zero correlation between actual gains from education and WTP for education. This implies that individuals are essentially "optimizing on noise"; given incorrect beliefs, individuals are not able to direct resources in a way that lines up with true benefits.

To test whether prediction errors vary by baseline ability, Figure 4.6 plots the perceived change in score from the education material and actual change in score versus baseline ability. The figure highlights that the prediction error, or the gap between actual change in score and expected change in score, is larger for individuals with lower starting scores and decreases as baseline score increases. Although this pattern is partly driven by the natural cap on the scores at 10 points (given the 0-10 range), this is consistent with the idea that individuals with higher levels of financial knowledge have a better understanding of their abilities.

	(1) T-C Change (AI)	(2) T-C Change (Video)	(3) T-C Change (Lecture)
Baseline number correct (pre)	-0.333***	-0.225***	-0.157***
	(0.060)	(0.061)	(0.060)
WTP for This Material	0.032	0.032	-0.030
	(0.031)	(0.030)	(0.025)
Est. Change (rel. to control)	0.015	-0.029	0.026
	(0.041)	(0.049)	(0.055)
Constant	3.638***	2.809***	2.187***
	(0.437)	(0.438)	(0.451)
R^2	0.108	0.052	0.027
Observations	249	260	263

Table 4.6: Selection on Treatment Effects

Notes: Table 4.6 examines selection on treatment effects by relating actual learning gains to willingness to pay (WTP) and beliefs. The dependent variable in each column is actual learning, defined as the participant's endline score minus the average endline score of control-group participants with the same baseline score (i.e., baseline-score–adjusted treatment–control differences). Each column restricts the sample to participants who received the indicated education format (AI, video, or lecture). "Baseline number correct (pre)" measures initial performance on the financial decision-making task. "WTP for This Material" is the participant's elicited willingness to pay for the specific education format received. "Est. Change (rel. to control)" is the participant's expected learning gain relative to control, measured prior to treatment. Standard errors are reported in parentheses.

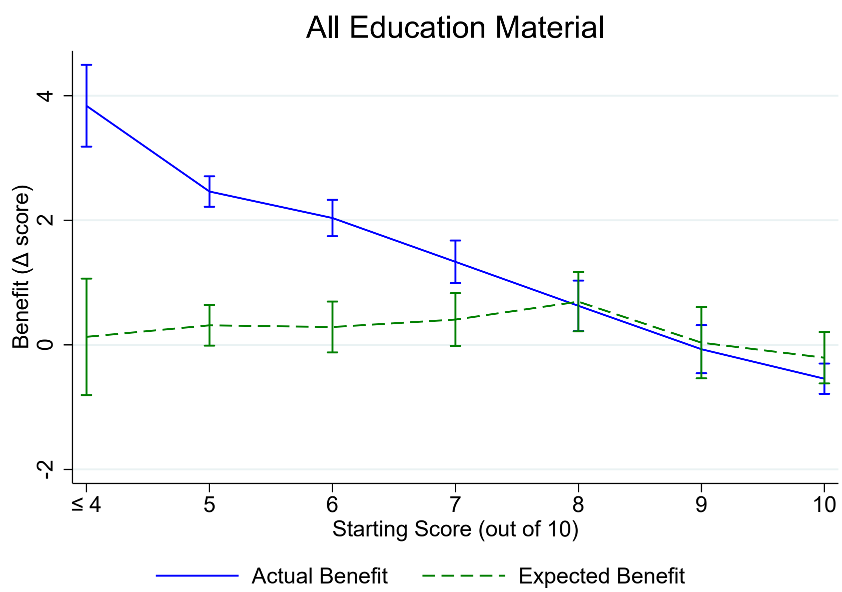


Figure 4.6: Actual versus Expected Benefit by Ability

Notes: Figure 4.6 plots expected or perceived benefits from all versions of the education material (in change in score) against the actual benefits from education by baseline ability. The figure shows that forecast errors, or the gap between the actual and perceived benefit, are smaller for higher ability individuals. This is consistent with the idea that individuals with higher levels of financial knowledge have a better understanding of their abilities. This figure pools all the versions of the education material together and only shows the treated individuals. Error bars show 95% confidence intervals.

	(1) WTP for AI	(2) WTP for Video	(3) WTP for Lecture	(4) Average WTP
Baseline number correct (pre)	0.072	0.035	-0.066	0.003
	(0.095)	(0.086)	(0.092)	(0.080)
Estimated pre score	0.231***	0.182***	0.249***	0.200***
	(0.067)	(0.064)	(0.069)	(0.061)
Change in number correct	0.044	0.111	0.014	0.043
	(0.085)	(0.080)	(0.084)	(0.074)
Estimated change	0.385***	0.244***	0.399***	0.291***
	(0.076)	(0.076)	(0.073)	(0.076)
Perceived difficulty (0 to 10)	-0.083	-0.134**	-0.272***	-0.068
	(0.055)	(0.056)	(0.062)	(0.066)
Estimated change	0.000	0.000	0.000	
	(.)	(.)	(.)	
AI Belief Index (z)	1.825***	0.288	0.356	0.879***
	(0.240)	(0.199)	(0.217)	(0.187)
Constant	-0.530	0.039	1.927*	0.267
	(0.890)	(0.868)	(0.982)	(0.812)
R^2	0.125	0.033	0.078	0.051
Observations	1016	1021	1028	967

Table 4.7: Determinants of Willingness to Pay for Educational Materials (OLS)

Notes: Table 4.7 reports OLS regressions of willingness to pay (WTP) for each educational format (AI, video, lecture) and the average WTP across all formats per participant. Columns (1)–(3) correspond to WTP for AI, video, and lecture, respectively; Column (4) reports results for the average WTP across all three materials. *Baseline number correct (pre)* is the number of questions the participant got right on the first set of decisions. *Change in number correct* is the post–pre difference in actual performance. *Perceived difficulty* and *Estimated change* capture participants' beliefs about how hard the material is and how much they expected to improve. The *AI Belief Index* is the standardized average of participants' attitude and trust in AI. All regressions include demographic controls (college, gender, race, age, and income), which are omitted from the table for brevity. Standard errors are reported in parentheses.

Table 4.6 examines how actual learning gains (improvements in score between the pre and post tasks, subtracting out the average score of control participants with the same starting score) vary with baseline ability, willingness to pay, and expected benefits across the three educational formats. Across all materials, learning gains are highest among participants with the lowest baseline scores. In contrast, willingness to pay for a given material is only weakly related to realized learning gains. The estimated coefficient on willingness to pay is small and statistically insignificant for AI and lecture, and only marginally positive for video. Similarly, expected learning gains relative to the control group are not predictive of actual learning outcomes. Together, these results show that treatment effects are strongly heterogeneous and concentrated among lower-ability participants, while willingness to pay is only weakly related to realized gains, consistent with demand being driven by perceived rather than actual returns.

	(1) Estimated pre score	(2) Overestimate pre (est-actual)	(3) Estimated change	(4) Underestimate change (actual-est)
Pre score	0.321***	-0.679***	0.040	-0.040
	(0.045)	(0.045)	(0.057)	(0.057)
Actual change			0.131	0.869***
			(0.175)	(0.175)
Actual change × Pre score			0.005	-0.005
			(0.027)	(0.027)
Constant	3.302***	3.302***	-0.962*	0.962*
	(0.459)	(0.459)	(0.517)	(0.517)
R^2	0.054	0.178	0.016	0.301
Observations	1047	1047	1047	1047

Table 4.8: Actual and Perceived Ability and Learning Gains

Notes: Table 4.8 examines the relationship between actual and perceived ability and learning gains. Column (1) regresses participants' estimated baseline quiz score on their actual baseline score. Column (2) regresses the signed estimation error in baseline ability—defined as estimated pre score minus actual pre score—on actual baseline ability; positive values indicate overestimation. Column (3) regresses participants' estimated learning gains on their actual learning gains (change in number correct from pre to post). Column (4) regresses the signed estimation error in learning gains—defined as actual change minus estimated change—on actual learning gains; positive values indicate underestimation of learning gains. All regressions include controls for college education, gender, race, income, and age. Standard errors are reported in parentheses.

We further dissect this relationship in Table 4.7, where we regress actual and estimated baseline scores and benefits on demand (WTP) for different versions of the education material. Across all three formats, perceived benefit is the clearest and most consistent predictor of WTP. Participants who believed they would learn more were willing to pay more, with especially large effects for the AI-based lesson and smaller but still meaningful effects for video and lecture. Taking the average across all versions of the education material, a one-point increase in expected learning gains is associated with an increase in average willingness to pay of 29.2 cents (s.e. = 0.076). This validates our willingness to pay measure and suggests that demand for education is "rational" conditional on beliefs.

Table 4.7 also highlights that perceived difficulty moves in the opposite direction. Participants who thought the education material would be more difficult had lower demand, particularly for the video and lecture formats. The estimate for AI is negative as well, although smaller and less precisely estimated. These patterns indicate that the cognitive costs of learning do influence demand for education, but to a lesser extent than the predicted benefits of education given the smaller coefficients.

Finally, Table 4.7 highlights that objective measures of learning (or the change in score) show much weaker relationships with WTP. Pre score and actual improvement have small and imprecise coefficients, which is consistent with participants basing their WTP primarily on their beliefs about how much they would benefit, even when these beliefs did not reflect actual performance gains.

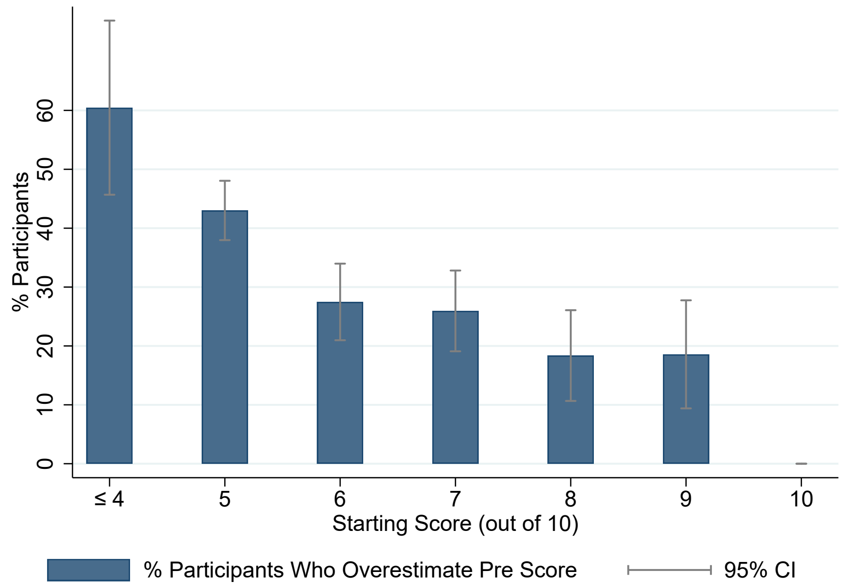


Figure 4.7: Overconfidence by Pre Score

Notes: Figure 4.7 shows the portion of participants who overestimate their starting score by pre score. The figure highlights that participants with lower pre scores are more likely to overestimate their starting score than participants with higher pre scores. Error bars show 95% confidence intervals.

Table 4.8 helps us understand what is driving the prediction errors. Column 1 of Table 4.8 shows the extent to which participants mispredict their starting score; for each additional point (or question) answered correctly, participants increase their estimate of their baseline score by only 0.32 points. Column 2 highlights an interesting heterogeneity: lower-ability participants systematically *overestimate* their baseline knowledge. We refer to this as overconfidence, following Definition One in Moore and Healy [2008]. Individuals with lower starting ability are more likely to be overconfident – each additional point on the baseline quiz is correlated with a 0.68 point reduction in overestimation of the baseline score. The percent of overconfident individuals by baseline score is also shown in Figure 4.7.

Column 3 of Table 4.8 shows that for 1 additional point of actual improvement, predicted improvement only increases by 0.13 points. Comparing the coefficients in Column 1 and 3 suggests that participants have more accurate estimates of their starting score than of their change in score with education. For every 1 point increase in baseline score, participants only increase their estimated baseline score by 0.321, but for every 1 point increase in their actual change in score, participants only increase their estimated change by 0.164 points. Finally, Column 4 shows that the participants who actually see the most improvement in score are more likely to underestimate their improvement: each additional point of actual improvement is correlated with an underestimation of 0.836 points.

Taken together, these results imply a misalignment between demand and returns: individuals who would benefit the most from education are not the ones with the highest willingness to pay. In particular, lower-ability participants tend to overestimate their starting level of knowledge and underestimate the potential gains from education. Combined, this leads to underinvestment in education. Higher-ability participants, whose beliefs are more accurately calibrated, express higher willingness to pay despite experiencing smaller realized gains.

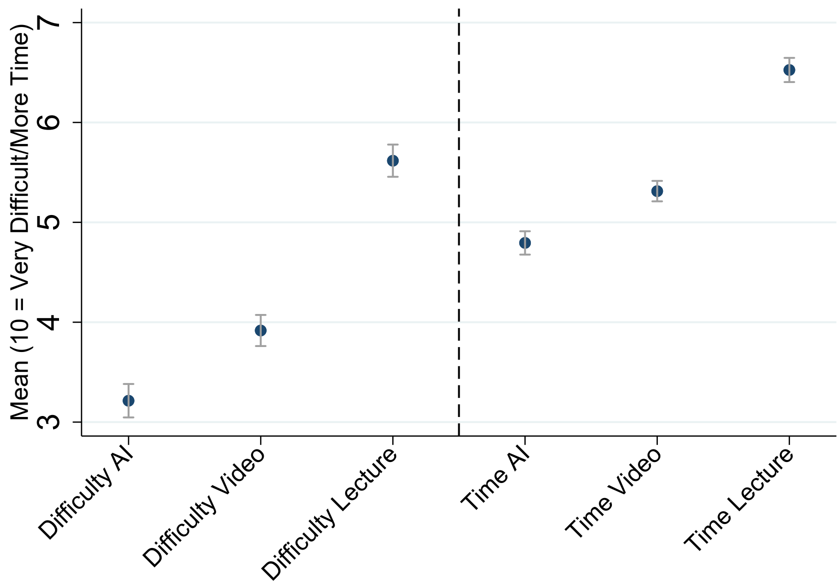


Figure 4.8: Perceived Cost of Education Material

Notes: Figure 4.8 shows the perceived difficulty, or cognitive cost, of the different versions of the education material. Participants perceive the AI education material to be the least cognitively difficult, the video education material to be slightly more difficult, and the lecture education material to be the most cognitively difficult. This is in line with the way we intentionally described the material.

Reducing the Costs of Education

We test whether lowering the perceived cognitive cost of education can correct these selection distortions by offering participants different versions of the education material. We then randomized participants into one of the three educational formats: a standard text-based university lecture (high cognitive cost), a video-based module, and an interactive AI-based learning assistant (low cognitive cost).

Figure 4.8 shows that we are successfully able to vary perceived difficulty of the education material. We elicit perceived difficulty, or perceived cognitive cost, using two measures: 1) how difficult participants perceive the education material to be, and 2) how much time they expect it will take them to complete and fully understand that version of the education material, both before completing the education material). Participants all find the AI material to be the least difficult version, the video education material to be the second most difficult, and the university lecture version to be the most difficult. Participant free-response answers about why they preferred specific materials shows that the primary reason for choosing the AI and video materials was because they were easier to understand and clearer.

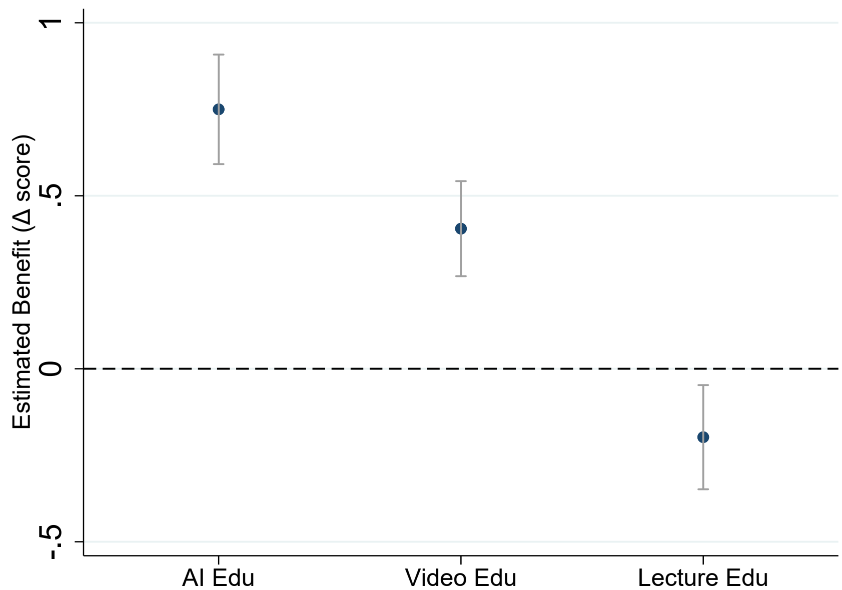


Figure 4.9: Estimated Benefit of Education Material

Notes: Figure 4.9 shows the estimated benefits from education for each version of the education material. Participants expect to see the largest changes in score from the AI education material, followed by the video education material and the lecture education material. This mirrors the perceived cognitive difficulty of the material. Notably, participants expect a positive change in score from the AI and video education materials and zero (or slightly negative) change in score from the lecture education material.

Figure 4.9 shows the estimated benefits from education for each version of the material and Figure 4.10 shows how demand for education varies with each version. We find that the video and AI-versions of the education material (the technology-backed versions of the education material designed to reduce cognitive difficulty) increase the perceived benefit of education and increase demand across all ability levels. This suggests that reducing the cognitive cost of education does increase aggregate demand, however, lowering the cost of education fails to correct the selection pattern. While demand increased across the board, the positive correlation between baseline ability and WTP persisted in the AI and video treatment arms. Low-ability individuals remained less willing to pay than high-ability individuals. The persistence of the selection pattern raises concerns that high ability individuals with higher demand will continue to crowd-out lower ability individuals, even in settings where financial literacy education is taught in an easy, beginner-friendly way. Our results suggest that cognitive cost is not the primary barrier preventing low-ability individuals from investing in education. Instead, the binding constraint appears to be the belief frictions identified above – specifically, mispredicting the benefits of education and an overconfidence in one's baseline ability.

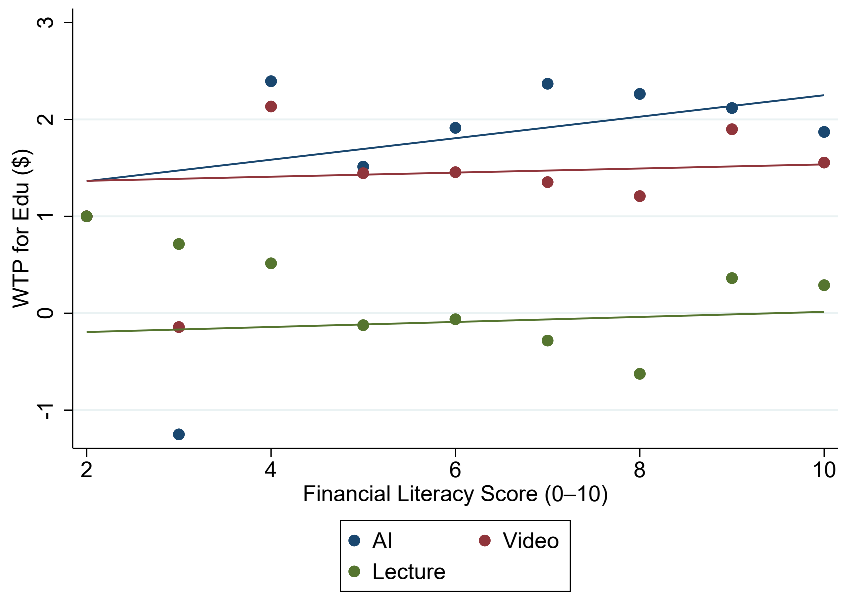


Figure 4.10: WTP by Baseline Ability for each version of the Education Material

4.4 Conclusion

This chapter investigates a critical friction in the accumulation of financial knowledge: the self-selection mechanisms that deter high-return individuals from enrolling in financial literacy education. Standard human capital theory predicts that individuals will invest in education when the expected marginal returns exceed the marginal costs. Under this framework, individuals with the lowest baseline level of knowledge – and thus the highest potential returns to learning – should be the most eager to invest in education. Our findings challenge this assumption in the context of financial literacy education, documenting a troubling pattern where the individuals who stand to benefit the most from education are the least likely to demand it.

Using administrative data from a large university, we first show that enrollment in financial education is characterized by strong positive selection. Students from higher socioeconomic backgrounds and those with stronger quantitative skills effectively select into the course, while underrepresented minorities and those from less privileged backgrounds select out. Our laboratory experiment isolates the mechanism driving this inefficiency. We find that low demand among individuals with lower starting levels of knowledge is not driven by rational cost-benefit analysis, but by systematic belief errors. Low-ability individuals are more likely to underestimate their returns to education. This is in part driven by an overconfidence in their baseline level of knowledge, which leads low-ability individuals to underestimate their potential for improvement.

Additionally, we show that supply-side interventions designed to reduce cognitive costs are insufficient to correct this market failure. While lowering barriers via AI and video-based learning tools does increase aggregate demand compared to traditional lecture-based learning, it fails to alter the selection patterns. The composition of the market remains skewed toward higher ability individuals, suggesting that the binding constraint is not on the difficulty of investing in education, but the misperception of its returns.

These results have important policy implications. Governments and institutions often rely on individuals to "opt-in" to financial education programs to improve decision-making and reduce inequality. Furthermore, even in cases where financial education is required, individuals have to "opt-in" on the intensive margin and choose to allocate attention and effort. Our findings suggest that self-selection patterns may render these education resources less effective, and may even exacerbate inequality by subsidizing the human capital accumulation of individuals who are already financially sophisticated. Because the friction is driven by systematic belief errors about potential benefits rather than access, standard subsidies or even technologies that lower the real and cognitive costs of participation are unlikely to solve the problem.

Ultimately, this chapter highlights a critical friction in human capital accumulation. When returns to education are difficult to assess and beliefs are miscalibrated, individuals who are most in need of education may have the most difficulty evaluating the returns to education. While this chapter focuses on financial literacy education, the implications are likely to be true for other domains such as mental health take-up, preventative health, or job retraining. Acknowledging those most in need of resources often have the most difficulty evaluating the benefits of said resources is an important step towards effective policy design and opens important areas for further research. Together with the preceding chapters on debt management and housing experiences, these results make the case that behavioral frictions are not peripheral to household financial decision-making but central to it, and that policies designed without them in mind are likely to fall short of their intended goals.

Chapter 5 Conclusion

5.1 Summary of Findings

This dissertation has examined three different domains of household financial decision-making—debt management, homeownership, and the acquisition of financial knowledge—and has documented a behavioral mechanism in each that the standard frictionless benchmark cannot accommodate. The three chapters share a common architecture: a clean theoretical benchmark, a behavioral alternative, and an empirical test of which one fits the data, with policy implications drawn out of the result.

Chapter 2 studied the federal student loan repayment pause from 2020 through 2023, a uniquely well-defined episode of temporary debt relief. A neoclassical model of consumption smoothing predicts that relieved households should raise consumption during the pause, lower delinquency mechanically, and revert cleanly to pre-pause behavior once payments resume. Using the University of California Consumer Credit Panel and an instrumental variables strategy that exploits cross-borrower variation in eligibility, we showed that the early-period response was consistent with this benchmark, but that consumption rose only gradually and that delinquency increased during the later quarters of the pause and rose sharply once payments resumed. A life-cycle model with adaptive reference-dependent consumption reproduces these dynamics: temporary liquidity raises spending, which raises a habit-driven reference point, which makes the post-pause cutbacks costly enough that some borrowers default rather than absorb them. Counterfactual simulations show that ending the pause earlier—before habits had time to solidify—would have substantially reduced post-pause delinquency without sacrificing the bulk of the welfare gains from the relief itself.

Chapter 3 examined how households' lived experiences of local housing markets shape their decisions to enter homeownership. We developed a life-cycle model in which households form beliefs about both the mean and the volatility of future house-price returns from the price histories they have personally observed since age 18, and we tested its predictions on a panel that links each individual's residential zip code to local price data from Zillow. The empirical results confirmed both predictions: a one percentage point increase in experienced returns raises the probability of transitioning into homeownership by roughly seven percent relative to the sample's quarterly transition rate, while a one percentage point increase in experienced volatility lowers it by roughly ten percent. The volatility effect attenuates with age, consistent with the life-cycle structure of the model, and is detectable only at the zip-code level—underscoring how granular, local experiences are the operative input to belief formation in this market.

Chapter 4 turned to the question of how households decide whether to invest in financial knowledge in the first place. Using observational data on enrollment in a large personal finance course at UC Berkeley and a controlled online experiment with U.S. participants on Prolific, we documented strong positive selection on baseline ability, family income, and quantitative preparation: the participants with the most to learn are the least willing to pay to learn it. The experiment isolated the mechanism: low-ability participants overestimate their starting knowledge and underestimate their potential learning gains, so demand tracks perceived returns rather than actual returns. Varying the cognitive cost of the educational material—through AI, video, and lecture formats—raises aggregate demand but does not change the selection pattern. The binding constraint is belief miscalibration, not access or difficulty.

5.2 Synthesis

Read together, the three chapters tell a coherent story about household financial decision-making and about the limits of policy designed under the standard frictionless benchmark.

The first lesson is that behavioral frictions in household finance are systematic. They are not noise around a fully informed, forward-looking equilibrium. In each chapter, the deviation from the standard benchmark moves in a predictable direction and admits a tractable structural explanation. Habits form. Experiences become beliefs. Beliefs about one's own ability shape investment in knowledge. None of this is surprising in isolation; what the dissertation shows is that each of these mechanisms is large enough, and persistent enough, to govern first-order outcomes for the households it touches.

The second lesson is that these frictions interact with policy in ways that the standard benchmark misses. In Chapter 2, the welfare consequences of the student loan pause depend not only on its size and duration but on the habits it allows households to form: a shorter pause would have produced essentially the same short-run relief at substantially lower long-run cost. In Chapter 3, household entry into homeownership responds not to the average national housing market but to the local market each household happens to have lived in: housing policy that is calibrated to aggregate conditions can systematically miss the populations it is designed to reach. In Chapter 4, financial literacy mandates and supply-side subsidies that lower the cost of education will raise aggregate take-up but not redirect education toward the households who would benefit most: the households with the highest potential returns will continue to mispredict those returns until the underlying belief errors are addressed.

The third lesson is methodological. In each chapter, identifying the behavioral mechanism required combining the right data with the right model. Administrative panels provided the granularity, the time horizon, and the sample size needed to detect gradual dynamics, geographic heterogeneity, and adverse selection on belief errors. Structural models provided the discipline to map those patterns to the underlying mechanisms and, in two of the three chapters, to evaluate counterfactual policy designs. Neither piece, on its own, would have been sufficient.

5.3 Future Directions

Each chapter closes with directions specific to its own substantive question. Three broader directions cut across the dissertation.

The first is the long-run welfare consequences of debt-relief policies. Chapter 2 documents elevated post-pause delinquency in the quarters immediately following the resumption of payments. Whether this elevation translates into longer-run defaults, foreclosures, or bankruptcies, and how it interacts with subsequent labor-market and credit shocks, are open questions that the UC-CCP and similar data sources are well-positioned to answer as more post-pause years accumulate.

The second is the feedback from experience-based beliefs to asset prices themselves. Chapter 3 treated the local housing market as exogenous to individual decisions. In equilibrium, however, the entry decisions of new homeowners shaped by past returns will themselves influence future returns. Disciplining a richer quantitative model with realistic housing-market frictions and using it to study the feedback from extensive-margin homeownership decisions to prices would shed light on a relatively understudied mechanism for housing booms and slow recoveries.

The third concerns how to design educational interventions that reach the households who would benefit most. Chapter 4 shows that lowering the cost of financial literacy education is, on its own, insufficient: the binding constraint is belief miscalibration, and the supply-side interventions tested in our experiment failed to alter the selection pattern. An open question is whether interventions that directly target beliefs—for example, by giving low-ability individuals personalized feedback on their own performance before eliciting demand for education—can correct the selection pattern that purely cost-side interventions cannot. The same logic likely extends well beyond financial literacy, to other settings where take-up of welfare-improving services is constrained by beliefs about one's own returns to participation.

5.4 Closing

Standard models of household finance have been an indispensable starting point for the questions this dissertation has asked. They identify the natural benchmarks, generate sharp predictions, and discipline the empirical work that follows. What this dissertation has shown is that, in three distinct domains of household decision-making, the data require us to go beyond those benchmarks. Households form habits over consumption that outlast the relief that produced them. They form beliefs about housing markets from the markets they happen to have lived in. They invest in financial knowledge based on what they think they will gain rather than what they will actually gain. Each of these patterns is identifiable, structurally tractable, and consequential for the design of policy. Taking them seriously is, the dissertation argues, central to the next generation of work on household financial decision-making.

Bibliography

Manuel Adelino, Antoinette Schoar, and Felipe Severino. Perception of house price risk and homeownership. Technical report, National Bureau of Economic Research, 2018.

Sumit Agarwal, Chunlin Liu, and Nicholas S Souleles. The reaction of consumer spending and debt to tax rebates—evidence from consumer credit data. *Journal of Political Economy*, 115(6):986–1019, 2007.

Sumit Agarwal, Luojia Hu, and Xing Huang. Rushing into the american dream? house prices growth and the timing of homeownership. *Review of Finance*, 20(6):2183–2218, 12 2015. ISSN 1572-3097. doi: 10.1093/rof/rfv063. URL https://doi.org/10.1093/rof/rfv063.

Sandro Ambuehl, B. Douglas Bernheim, and Annamaria Lusardi. Evaluating Deliberative Competence: A Simple Method with an Application to Financial Choice. *American Economic Review*, 112(11):3584–3626, November 2022. ISSN 0002-8282. doi: 10.1257/aer.20210290. URL https://pubs.aeaweb.org/doi/10.1257/aer.20210290.

Luis Armona, Andreas Fuster, and Basit Zafar. Home price expectations and behaviour: Evidence from a randomized information experiment. *The Review of Economic Studies*, 86 (4):1371–1410, 2019.

Michael Bailey, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel. The economic effects of social networks: Evidence from the housing market. *Journal of Political Economy*, 126 (6):2224 – 2276, 2018. URL https://EconPapers.repec.org/RePEc:ucp:jpolec:doi:10.1086/700073.

Gordon M. Becker, Morris H. Degroot, and Jacob Marschak. Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3):226–232, 1964. ISSN 1099-1743. doi: 10.1002/bs.3830090304. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/bs.3830090304. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/bs.3830090304.

B. Douglas Bernheim and Dmitry Taubinsky. Behavioral Public Economics. In *Handbook of Behavioral Economics: Applications and Foundations 1*, volume 1, pages 381–516. Elsevier, 2018. ISBN 978-0-444-63374-3. doi: 10.1016/bs.hesbe.2018.07.002. URL https://linkinghub.elsevier.com/retrieve/pii/S2352239918300022.

Utpal Bhattacharya, Andreas Hackethal, Simon Kaesler, Benjamin Loos, and Steffen Meyer. Is Unbiased Financial Advice to Retail Investors Sufficient? Answers from a Large Field

Study. *The Review of Financial Studies*, 25(4):975–1032, April 2012. ISSN 0893-9454. doi: 10.1093/rfs/hhr127. URL `https://doi.org/10.1093/rfs/hhr127`.

Matthew J Botsch and Ulrike Malmendier. The long shadows of the great inflation: Evidence from residential mortgages. *Available at SSRN 3888762*, 2023.

Michael Boutros, Nuno Clara, and Francisco Gomes. Borrow now, pay even later: A quantitative analysis of student debt payment plans. *Journal of Financial Economics*, 159: 103898, 2024.

Miriam Bruhn, Gabriel Lara Ibarra, and David McKenzie. Why is Voluntary Financial Education so Unpopular? Experimental Evidence from Mexico. May 2013. doi: 10.1596/1813-9450-6439. URL `https://hdl.handle.net/10986/15580`.

David Card. The Causal Effect of Education on Earnings. In *Handbook of Labor Economics*, volume 3, pages 1801–1863. Elsevier, 1999. ISBN 978-0-444-50187-5. doi: 10.1016/S1573-4463(99)03011-4. URL `https://linkinghub.elsevier.com/retrieve/pii/S1573446399030114`.

Karl E Case and Robert J Shiller. Is there a bubble in the housing market? *Brookings papers on economic activity*, 2003(2):299–362, 2003.

Karl E Case, Robert J Shiller, and Anne Thompson. What have they been thinking? homebuyer behavior in hot and cold markets—a 2014 update. 2014.

Jiafeng Chen and Jonathan Roth. Logs with zeros? some problems and solutions. *The Quarterly Journal of Economics*, 139(2):891–936, 2024.

Raj Chetty and Adam Szeidl. Consumption commitments and habit formation. *Econometrica*, 84(2):855–890, 2016.

Wendy Craig, Lyndall Schumann, Kelly Petrunka, Shahriar Khan, and Ray Peters. Government costs associated with delinquent trajectories. *International Journal of Child, Youth and Family Studies*, 2(2.1):263–293, 2011.

Michael Dinerstein, Constantine Yannelis, and Ching-Tse Chen. Debt moratoria: Evidence from student loan forbearance. *American Economic Review: Insights*, 6(2):196–213, 2024.

Will Dobbie and Jae Song. Debt relief and debtor outcomes: Measuring the effects of consumer bankruptcy protection. *American economic review*, 105(3):1272–1311, 2015.

Will Dobbie, A Liberman, Daniel Paravisini, and Vikram Pathania. Measuring bias in consumer lending. *Review of Economic Studies*, 84(1):126–157, 2017.

Darren Duxbury and Barbara Summers. On perceptions of financial volatility in price sequences. *The European Journal of Finance*, 24(7-8):521–543, 2018.

Peter Ganong and Pascal Noel. Liquidity versus wealth in household debt obligations: Evidence from housing policy in the great recession. *American Economic Review*, 110(10): 3100–3138, 2020.

Mark A Greenwald, Sherry L Jackson, and Michael T Baglivio. The costs of delinquency. *Criminology & Pub. Pol'y*, 13:61, 2014.

Sandra J. Huston. Financial literacy and the cost of borrowing. *International Journal of Consumer Studies*, 36(5):566–572, 2012. ISSN 1470-6431. doi: 10.1111/j.1470-6431.2012.01122.x. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1470-6431.2012.01122.x`.

David S Johnson, Jonathan A Parker, and Nicholas S Souleles. Household expenditure and the income tax rebates of 2001. *American Economic Review*, 96(5):1589–1610, 2006.

Markku Kaustia and Samuli Knüpfer. Do investors overweight personal experience? evidence from ipo subscriptions. *The Journal of Finance*, 63(6):2679–2702, 2008.

THERESA Kuchler and BASIT Zafar. Personal experiences and expectations about aggregate outcomes. *The Journal of Finance*, 74(5):2491–2542, 2019. doi: https://doi.org/10.1111/jofi.12819. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12819`.

Gabriel Lara Ibarra, David McKenzie, and Claudia Ruiz Ortega. Learning the Impact of Financial Education When Take-Up Is Low, November 2017.

Remy Levin and Daniela Vidart. Risk-taking adaptation to macroeconomic experiences. Technical report, 2023.

Igor Livshits, James MacGee, and Michèle Tertilt. Consumer bankruptcy: A fresh start. *American Economic Review*, 97(1):402–418, 2007.

Allison Oldham Luedtke and Carly Urban. High school financial education courses in the United States.: What is the importance of setting state policies? *Journal of Financial Literacy and Wellbeing*, 1(3):431–449, 2023. doi: 10.1017/flw.2024.1. URL `https://www.cambridge.org/core/product/ABD577136362C0289884C41676341271`. Edition: 2024/02/15.

Annamaria Lusardi and Carlo de Bassa Scheresberg. Financial literacy and high-cost borrowing in the United States. Technical report, National Bureau of Economic Research, 2013.

Annamaria Lusardi and Olivia S. Mitchell. Financial Literacy and Planning: Implications for Retirement Wellbeing, May 2011. URL `https://www.nber.org/papers/w17078`.

Annamaria Lusardi and Olivia S. Mitchell. The Importance of Financial Literacy: Opening a New Field. *Journal of Economic Perspectives*, 37(4):137–154, November 2023. ISSN 0895-3309. doi: 10.1257/jep.37.4.137. URL `https://pubs.aeaweb.org/doi/10.1257/jep.37.4.137`.

Annamaria Lusardi, Pierre-Carl Michaud, and Olivia S. Mitchell. Optimal Financial Knowledge and Wealth Inequality. *Journal of Political Economy*, 125(2):431–477, April 2017. ISSN 0022-3808. doi: 10.1086/690950. URL `https://www.journals.uchicago.edu/doi/full/10.1086/690950`.

Ulrike Malmendier. Experience effects in finance: Foundations, applications, and future directions. *Review of Finance*, 25(5):1339–1363, 2021.

Ulrike Malmendier and Stefan Nagel. Depression babies: do macroeconomic experiences affect risk taking? *The quarterly journal of economics*, 126(1):373–416, 2011.

Ulrike Malmendier and Alexandra Steiny Wellsjo. Rent or buy? inflation experiences and homeownership within and across countries. *The Journal of Finance*, 79(3):1977–2023, 2024.

Ulrike Malmendier, Demian Pouzo, and Victoria Vanasco. Investor experiences and financial market dynamics. *Journal of Financial Economics*, 136(3):597–622, 2020.

Stephan Meier and Charles D. Sprenger. Discounting financial literacy: Time preferences and participation in financial education programs. *Journal of Economic Behavior & Organization*, 95:159–174, November 2013. ISSN 0167-2681. doi: 10.1016/j.jebo.2012.02.024. URL `https://www.sciencedirect.com/science/article/pii/S0167268112000625`.

Alvaro Mezza, Daniel Ringo, Shane Sherlund, and Kamila Sommer. Student loans and homeownership. *Journal of Labor Economics*, 38(1):215–260, 2020.

Olivia S. Mitchell and Annamaria Lusardi. Financial Literacy and Economic Outcomes: Evidence and Policy Implications. *The Journal of Retirement*, 3(1):107–114, June 2015. ISSN

2326-6899, 2326-6902. doi: 10.3905/jor.2015.3.1.107. URL https://jor.pm-research.com/content/3/1/107.

Don A. Moore and Paul J. Healy. The trouble with overconfidence. *Psychological Review*, 115 (2):502–517, April 2008. ISSN 1939-1471, 0033-295X. doi: 10.1037/0033-295X.115.2.502. URL https://doi.apa.org/doi/10.1037/0033-295X.115.2.502.

National Association of Realtors. The characteristic and challenges of first time home buyers. Technical report, National Association of Realtors, 2024. URL https://narfocus.com/billdatabase/clientfiles/172/1/4847.pdf.

Jonathan A Parker, Nicholas S Souleles, David S Johnson, and Robert McClelland. Consumer spending and the economic stimulus payments of 2008. *American Economic Review*, 103 (6):2530–53, 2013.

Robert J. Shiller and Anne K. Thompson. What have they been thinking? home buyer behavior in hot and cold markets: A ten-year retrospect. *Brookings Papers on Economic Activity*, 53(1 (Spring):307–366, None 2022. doi: None. URL https://ideas.repec.org/a/bin/bpeajo/v53y2022i2022-01p307-366.html.

Paige Marta Skiba and Jeremy Tobacman. Do payday loans cause bankruptcy? *Journal of Law and Economics*, 57(1):109–145, 2014.

Carly Urban, Maximilian Schmeiser, J. Michael Collins, and Alexandra Brown. The effects of high school personal financial education policies on financial behavior. *Economics of Education Review*, 78:101786, October 2020. ISSN 0272-7757. doi: 10.1016/j.econedurev.2018.03.006. URL https://www.sciencedirect.com/science/article/pii/S0272775718301699.

Maarten C. J. van Rooij, Annamaria Lusardi, and Rob J. M. Alessie. Financial literacy and retirement planning in the Netherlands. *Journal of Economic Psychology*, 32(4): 593–608, August 2011. ISSN 0167-4870. doi: 10.1016/j.joep.2011.02.004. URL https://www.sciencedirect.com/science/article/pii/S0167487011000195.

Brandon C Welsh. Economic costs and benefits of primary prevention of delinquency and later offending: A review of the research. *Early prevention of adult antisocial behaviour*, pages 318–355, 2003.

Robert J. Willis and Sherwin Rosen. Education and Self-Selection. *Journal of Political Economy*, 87(5, Part 2):S7–S36, October 1979. ISSN 0022-3808. doi: 10.1086/260821. URL https://www.journals.uchicago.edu/doi/abs/10.1086/260821.

Stephen P Zeldes. Consumption and liquidity constraints: an empirical investigation. *Journal of political economy*, 97(2):305–346, 1989.