Title
Understanding the Effects of Financial Globalization and Macroeconomic Shocks: Essays on International Finance and Consumption

Author
Shen, Sheng

Publication Date
2019

Peer reviewed|Thesis/dissertation

Understanding the Effects of Financial Globalization and Macroeconomic Shocks: Essays on International Finance and Consumption

by

Sheng Shen

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 Pierre-Olivier Gourinchas, Chair
Professor Yuriy Gorodnichenko
Professor Ulrike Malmendier
Professor Annette Vissing-Jorgensen

Spring 2019

Understanding the Effects of Financial Globalization and Macroeconomic Shocks:
Essays on International Finance and Consumption

Abstract

Understanding the Effects of Financial Globalization and Macroeconomic Shocks: Essays on International Finance and Consumption

by

Sheng Shen

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Pierre-Olivier Gourinchas, Chair

This dissertation consists of three essays that aim to provide a deeper understanding of the consequences of financial globalization and macroeconomic fluctuations—for banks, firms, and households—by analyzing both micro-foundations and macroeconomic implications.

The first chapter studies how globalization of banking systems has affected credit allocation and the macroeconomy. I provide a new theory of credit allocation in financial systems with both global and local banks, and tests it using cross-country loan-level data. I first point out that the traditional theory in banking and corporate finance of firm-bank sorting based on hard versus soft information does not explain the sorting patterns between firms and global versus local banks. In light of this puzzle, I propose a new perspective: global banks have a comparative advantage in extracting *global* information, and local banks have a comparative advantage in extracting *local* information. I formalize this view in a model in which firms have returns dependent on global and local risk factors, and each bank type can observe only one component of the firms' returns. This double information asymmetry creates a segmented credit market with a double adverse selection problem: in equilibrium, each bank lends to the worst type of firms in terms of the unobserved risk factors. Moreover, I show that the adverse selection problem has important macroeconomic implications. When one of the bank types faces a funding shock, the adverse selection affects credit allocation at both the extensive and intensive margins, generating spillover and amplification effects through adverse interest rates. I formally test the model using detailed firm-bank micro data and empirical strategies that tightly map to the model set-up. I find firm-bank sorting patterns, and effects of US and Euro area monetary policy shocks on credit allocation, that support the model predictions. This evidence reveals a novel *adverse selection channel* of international monetary policy transmission.

The second essay studies the long-run implications of financial crises and macroeconomic shocks on consumption behavior. We show that personal lifetime experiences can

"scar" consumption. Consumers who have lived through times of high unemployment have persistent pessimistic beliefs about their future financial situation, though their actual future income is uncorrelated with past experiences. Nevertheless worse lifetime experiences predict significantly reduced consumption spending, controlling for income, wealth, demographics, and time effects. As a result of their experience-induced frugality, scarred consumers also build up more wealth. The results are robust to a battery of variations in the liquid- and illiquid wealth and income controls in the PSID, and replication in the Nielsen Homescan Panel and the CEX. Scarred consumers use more coupons and purchase more sale items and lower-end products. We use the stochastic life-cycle model of Low et al. (2010) to show that the estimated negative relationship between experiences and consumption cannot be generated by financial constraints, income scarring, or unemployment scarring, but is consistent with *experience-based learning*. As predicted by experience-based learning, the estimated effects of a macro shock are stronger for younger cohorts. The results suggest a novel micro-foundation of fluctuations in aggregate demand, and imply long-run effects of macroeconomic shocks.

The third chapter studies the impact of capital inflows on the real economy in the context of a "China shock" in the US real estate market. We document an unprecedented surge in housing purchases by foreign Chinese in the US over the past decade and estimates its effect on US local economies. Using transaction-level data on housing purchases, we find that the share of purchases by foreign Chinese in the California real estate market increased almost twentyfold during the period of 2007-2013. In particular, these purchases have been concentrated in zip codes that are historically populated by ethnic Chinese, making up for more than 10% of the total real estate transactions in these neighborhoods in 2013. We exploit the cross-sectional variation in the concentration of Chinese population settlement across zip codes during the pre-sample period to instrument for housing transactions by foreign Chinese. The results show that the surge in capital inflow from China into the US real estate market significantly increases local housing prices and local employment. We present evidence showing that this effect is primarily driven by a housing net worth channel. The evidence highlights the role of foreign capital inflow on the local real economy, especially in times of economic downturns.

Contents

Contents i

List of Figures iii

List of Tables v

1 Global vs. Local Banking: A Double Adverse Selection Problem 1
- 1.1 Introduction 1
- 1.2 Traditional Theory and New Perspective 9
- 1.3 A Model with Global and Local Banking 12
- 1.4 Comparative Analysis and Implications 28
- 1.5 Mapping Theory to Empirics 33
- 1.6 Empirical Analysis: Firm-Bank Sorting 37
- 1.7 Empirical Analysis: Adverse Selection Channel of Monetary Policy Transmission 43
- 1.8 Conclusion 51

2 Scarred Consumption **52**
- 2.1 Introduction 52
- 2.2 Data and Variable Construction 59
- 2.3 Baseline Results: Consumption and Consumer Optimism, Future Income and Future Wealth 66
- 2.4 Consumption with Experience-based Learning 76
- 2.5 Robustness using PSID, CEX, and Nielsen 85
- 2.6 Further Implications and Discussions of Experience Effects 93
- 2.7 Aggregate Implications and Conclusion 101

3 Capital Flows, Asset Prices, and the Real Economy: A "China Shock" in the US Real Estate Market **104**
- 3.1 Introduction 104
- 3.2 Housing Purchases by Foreign Chinese in the US 110
- 3.3 Foreign Chinese Housing Purchases and Local Economies 115

3.4 Channel Linking Foreign Chinese Housing Purchases to the Real Economy 128
3.5 Simple Model 135
3.6 Conclusion 141

Bibliography **142**

A Appendix for Chapter 1 **156**
A.1 Model Solutions and Proofs 156
A.2 Additional Figures and Tables 162

B Appendix for Chapter 2 **166**
B.1 Empirical Analysis 166
B.2 Model 183

C Appendix for Chapter 3 **192**
C.1 Additional Figures and Tables 192

List of Figures

1.1 **Global Banking Credit to Private Sector** 2
1.2 **Firm-Bank Sorting, by Firm Size and Age Quartile** 11
1.3 **Loan Currency Denominations by Global and Local Banks** 13
1.4 **Model Timeline** 15
1.5 **Benchmark Equilibrium: First-Best and Closed Economy** 18
1.6 **Illustration of Interest Rate Functions and Threshold Functions** . 21
1.7 **Firm-Bank Sorting Firm Space** 23
1.8 **Firm-Bank Sorting and Interest Rates Under Symmetric Equilibrium** 25
1.9 **Firm-Bank Sorting under Asymmetric Equilibrium** 26
1.10 **Effects of Financial Integration** 27
1.11 **Effects of a Positive r^G Shock (r^G lowers)** 29
1.12 **Effects of a Decline in Funding Cost r^G** 32
1.13 **Estimates of Global Factor z^G** 39
1.14 **Firm-Bank Sorting, by z_i^G/z_i^L Quartile (Method 1)** 40

2.1 **Monthly Consumption Expenditure by Age Group** 60
2.2 **Unemployment Experience by Age Group and by Region** 64
2.3 **Examples of Experience Shocks from the Recession (PSID)** 68
2.4 **Wealth Build-up: Estimated Coefficients and Confidence Intervals for Experience Measures** 77
2.5 **Wealth and Income Controls: Coefficients and Confidence Intervals for Experience Measures** 87
2.6 **Example of Unemployment Experience Shock from Recession, Nielsen** 95
2.7 **Aggregate Unemployment Experience and Consumer Spending** . 103

3.1 **All-Cash Housing Purchases by Ethnic Chinese and Anglo-American** 112
3.2 **Housing Purchases by Foreign Chinese and Buyers of Other Ethnicities** 114
3.3 **Foreign Chinese Housing Purchases across Zip Codes based on Historical Ethnic Chinese Population Share** 118

3.4 **Zillow Home Price Index by Historical Ethnic Chinese Population Share** . . . 119
3.5 **Thought Experiment: Counterfactual Housing Prices and Employment Conditions** . . . 129

A.1 **Estimates of Average Productivity Measure** $log\ z_{it}$ **by Country** . . . 163
A.2 **Firm-Bank Sorting, by** z_i^G/z_i^L **Quartile (Method 2)** . . . 164
A.3 **Three-Month Euribor Rates around ECB Announcements** . . . 165

B.1 **Average Life-Cycle Consumption** . . . 189
B.2 **Average Life-Cycle Consumption for Agents with Good Realizations Early in Life** . . . 190
B.3 **Average Life-Cycle Consumption Patterns for Agents with Bad Realizations Early in Life** . . . 191

C.1 **Housing Purchases by Foreign Chinese in the Los Angeles Region** 200

List of Tables

1.1 **Firm-Bank Sorting, by Firm Size and Age Quartile: Statistical Test** 12
1.2 **Banks' Overall Risk Before and After a Decline in r^G** 33
1.3 **Summary Statistics: Loan and Firm Count by Country (Method 1)** 36
1.4 **Summary Statistics: Firm Characteristics by Bank Type** 36
1.5 **Firm-Bank Sorting, by z_i^G/z_i^L Quartile: Statistical Test** 41
1.6 **Firm-Bank Sorting, Traditional Theory and New Perspective** . . . 41
1.7 **Determinants of z_i^G/z_i^L and Global Banking Credit** 42
1.8 **Monetary Policy Shocks and Credit Allocation: Extensive Margin** 49
1.9 **Monetary Policy Shocks and Credit Allocation: Intensive Margin** 50

2.1 **Summary Statistics (PSID)** 65
2.2 **Experience Effects and Annual Consumption (PSID)** 69
2.3 **Experience Effects and Expectations** 73
2.4 **Experience Effects and Future Income** 75
2.5 **Key Simulation Parameters** 81
2.6 **Estimations with Model-Simulated Data** 82
2.7 **Estimations with Model-Simulated Data, Unemployment Scarring** 83
2.8 **Summary Statistics (CEX and Nielsen)** 89
2.9 **Experience Effects and Quarterly Consumption (CEX)** 90
2.10 **Experience Effects and Monthly Consumption (Nielsen)** 94
2.11 **Experience Effects and Monthly Consumption Quality (Nielsen)** . 97
2.12 **Age-Heterogeneity in Reaction to Unemployment Fluctuation** . . 98
2.13 **Experience Effects and Consumption, GMM regressions** 101

3.1 **Summary Statistics** 116
3.2 **Foreign Chinese Housing Demand and Local Home Prices (Zillow)** 121
3.3 **Foreign Chinese Housing Demand and Local Home Prices (Transaction Prices)** 123
3.4 **Foreign Chinese Housing Demand and Total Employment** 124
3.5 **Foreign Chinese Housing Demand and Ex-ante Employment: Placebo Test** 126
3.6 **Testing for Direct Demand Channel: Number of Tax Return Filings** 131

3.7 **Foreign Chinese Housing Demand and Tradable/Non-tradable Sector Employment** . . . 133
3.8 **Foreign Chinese Housing Demand and Non-Tradable Sector Employment, Excluding Construction Sector** . . . 134
3.9 **Foreign Chinese Housing Demand and Housing Net Worth** . . . 135
3.10 **Foreign Chinese Housing Demand and Foreclosure (Share)** . . . 136

A.1 **Summary Statistics: Loan and Firm Count by Country (Method 2)** 162

B.1 **Summary Statistics (PSID), Full Sample** . . . 169
B.2 **Consumption (PSID), Full Sample** . . . 170
B.3 **Consumption (PSID), Alternative Experience Measure: Gap Years** 171
B.4 **Consumption (PSID): Alternative Experience Measure: Spousal Experience** . . . 172
B.5 **Consumption (PSID), Alternative Experience Measure: Lagged Experience** . . . 173
B.6 **Consumption (PSID), Alternative Experience Measure: Different Weights (λ)** . . . 174
B.7 **Consumption (PSID), Alternative Clustering Units** . . . 175
B.8 **Consumption (PSID), Alternative Weights: PSID Weights** . . . 176
B.9 **Consumption (PSID), Additional Wealth Controls** . . . 177
B.10 **Consumption (PSID), Additional Income Controls** . . . 178
B.11 **Consumption (PSID), Additional Liquidity Controls** . . . 179
B.12 **Wealth Accumulation** . . . 180
B.13 **Summary Statistics (Nielsen-CEX Matched Data)** . . . 181
B.14 **Consumption (Nielsen-CEX Matched Sample)** . . . 182
B.15 **Model Parameters Used in Simulations** . . . 186
B.16 **High-Education Estimations with Model-Simulated Data** . . . 188
B.17 **High-Education Estimations with Model-Simulated Data with Unemployment Scarring** . . . 188

C.1 **Foreign Chinese Housing Demand and Housing Prices, Alternative IV** . . . 193
C.2 **Foreign Chinese Housing Demand (Count) and Housing Prices** . 194
C.3 **Foreign Chinese Housing Demand (Count) and Housing Transaction Prices** . . . 195
C.4 **Foreign Chinese Housing Demand (Count) and Total Employment** 196
C.5 **Testing for Migration Channel: Number of Tax Return Filings (Lead)** . . . 197
C.6 **Foreign Chinese Housing Demand (Count) and Tradable/Non-Tradable Sector Employment** . . . 198
C.7 **Foreign Chinese Housing Demand and Foreclosure (Count)** . . . 199

Acknowledgments

I am extremely grateful to my advisers—Pierre-Olivier Gourinchas, Ulrike Malmendier, and Yuriy Gorodnichenko—for their guidance and support over the years. I am indebted to Pierre-Olivier for his advice, critique, and painstaking effort in helping me make decisions at key moments in my career. His depth of knowledge and perspicacity have been instrumental in the creation of this thesis, and are an endless source of inspiration. I am grateful to Ulrike for her mentorship and collaboration. Working with her has been fundamental to my growth as a researcher, thinker, writer, and presenter, and her mentorship has been central to my personal development as an economist. I thank Yuriy—who has been a source of advising and support since my first day in graduate school—for his insightful suggestions, encouragement, and generosity at every stage.

In addition, I have benefited tremendously from discussions with many faculty at UC Berkeley, including my committee member Annette Vissing-Jorgensen, as well as Amir Kermani, Ross Levine, Andrés Rodríguez-Clare, Andrew Rose, and Jim Wilcox. I also have been fortunate to have received continued mentorship and encouragement from economists in the larger academic community, including Linda Goldberg, Paolo Pesenti, Rebecca Hellerstein, Galina Hale, and Joseph Joyce. This thesis has benefited from generous financial support from the Clausen Center, and data support from the Fisher Center for Real Estate & Urban Economics and the Berkeley Data Lab.

Graduate school would not have been as stimulating or enjoyable without the conversations and help of my peers, including Carlos Avenancio, Maria Coelho, Satoshi Fukuda, Chris Jauregui, Casper Nordal Jørgenson, Dmitry Koustas, Zhimin Li, Waldo Ojeda, Vincenzo Pezone, Nick Sander, Ganesh Viswanath Natraj, and Calvin Zhang.

I dedicate this thesis to my family, to whom I owe everything. I thank my parents and grandparents for instilling in me a love for learning and strength to overcome all obstacles. I am deeply grateful to my husband, Jonathan Ragan-Kelley, who constantly uplifts me with his love and support.

Chapter 1

Global vs. Local Banking: A Double Adverse Selection Problem

1.1 Introduction

One of the most striking developments in credit markets across the world over the past two decades has been the increase in global banking credit—loans given by global banks to firms abroad.[1] Global banking credit has more than tripled since the mid-1990s, reaching almost $15 trillion and accounting for around 20% of total domestic private credit on average for a developed or major emerging market economy (Figure 3.2). This implies that there has been both a transformation of the competitive structure of credit markets and an expansion of financing sources for corporations: in a typical financial system today, firms can get credit not only from local banks but also from global banks.

Global banking credit also has important macroeconomic and policy implications. The global financial crisis has revealed that fluctuation in this credit serves as a key channel through which monetary policy and liquidity conditions get transmitted abroad (see, e.g., Cetorelli and Goldberg 2012a, Schnabl (2012), Rey 2016, Bräuning and Ivashina 2017). This, in turn, has prompted debate on optimal bank regulation and macroprudential policies in the presence of globalized credit markets (see e.g., Stein 2014, Fischer 2015, Rajan 2015, and Bernanke 2017).

Despite extensive debates on measures to minimize the risks entailed by global banking credit, it remains an open question how credit is allocated in financial systems with both global and local banks in the first place. Why do some firms get loans from global banks instead of local banks? Is existing theory in banking and corporate finance sufficient for explaining patterns of firm-bank sorting in globalized financial systems?

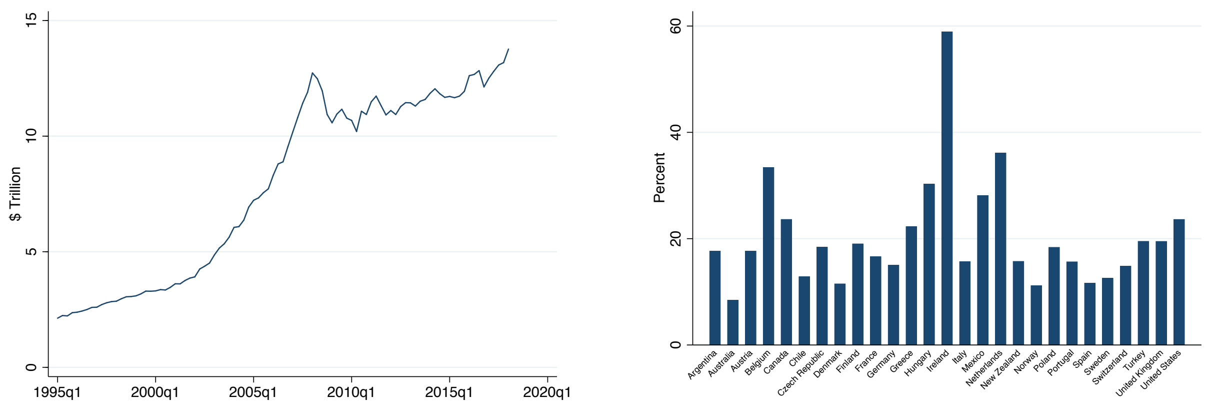
[1] Global banking credit, as described here, can be summarized as cross-border loans. Global banks are defined as banks that make cross-border loans and thereby have sizable foreign positions on their balance sheets.

How do global banks propagate monetary policy and liquidity shocks across borders?

Figure 1.1: **Global Banking Credit to Private Sector**

(a) Total Global Banking Credit, All Countries

(b) Share in Total Private Credit, by Country



Notes. Panel (a) plots a time-series of total cross-border credit to the non-bank private sector across all BIS reporting countries. Source: BIS Locational Banking Statistics. Panel (b) plots the share of cross-border credit in total private credit, averaged over 2005-2016, for a cross-section of developed and major emerging market economies. Source: BIS Locational Banking Statistics and IMF International Financial Statistics.

In this paper, I study these questions theoretically and empirically. I point out that the traditional theory in banking and corporate finance of firm-bank sorting based on hard versus soft information does not explain the sorting patterns between firms and global versus local banks. Instead, I show that bank specialization in *global* versus *local* information constitutes a key mechanism driving firm-bank sorting in financial systems with both global and local banks. Global banks specialize in information on global risk factors, and local banks specialize in information on local risk factors. This micro-foundation reveals a problem of double adverse selection in credit allocation in globalized financial systems, and has important macroeconomic implications. In particular, the double adverse selection constitutes a novel *adverse selection channel* of international transmission.

I start the analysis by testing whether the sorting pattern between firms and global versus local banks follows the prediction from traditional banking and corporate finance theory. The traditional theory posits that banks and firms sort based on hard and soft information[2]: large banks are more likely to lend to firms with more readily available hard information, which tend to be large and established firms, while small banks are

[2] A well-established strand of literature in finance has used the distinction between hard and soft information to explain lending relationships between banks and firms. Section 1.2 provides an overview of the traditional theory.

more likely to establish relationships with firms with more soft information, which tend to be small and young firms. Mapping this theory to the context of firm-bank sorting in globalized financial systems, one would conjecture that global banks are more likely to lend to firms with more hard information, since global banks tend to be larger than local banks. However, using a cross-country firm-bank loan-level dataset, I find that the traditional theory does *not* predict the sorting patterns between firms and global versus local banks: both global and local banks lend to firms across the entire asset size and age distribution. This points to a puzzle in the mechanism driving global banking credit: why do firms of similar size and age borrow from different types of banks?[3]

In light of this puzzle, I raise a new perspective. I argue that global and local banks differ in their specialization in global and local information: global banks have a comparative advantage in extracting information on global risk factors, and local banks have a comparative advantage in extracting information on local risk factors. Each bank type's comparative informational advantage plays a key role in determining firm-bank sorting in financial systems with both bank types. This idea is motivated by the observation that global banks are uniquely positioned to extract information on global factors through global market making activities and research efforts within the banking organizations.[4] At the same time, local banks are more conveniently positioned to extract information on local factors through local lending relationships (Petersen and Rajan 1994, Berger et al. 2005).

To formalize the new perspective and provide guidance for empirical testing, I first develop a model with global and local banks in which each bank type's comparative informational advantage serves as the key ingredient. From this one key ingredient, the model generates a sharp prediction about the equilibrium credit allocation in a two-bank-type economy: firms with higher expected return based on global factors relative to local factors are more likely to borrow from global banks, and vice versa for firms with returns more dependent on local factors. Using cross-country firm-bank loan-level data and empirical strategies that closely map to the model, I find empirical evidence that is consistent with the prediction.

To make this result more concrete, consider two firms: Oil States International, an American multinational corporation that provides services to oil and gas companies, and Zale Corporation, an American jewelry retailer that has a large presence in malls around the US. While both firms are public firms, headquartered in Texas, and of similar size (with total assets around \$1.3 billion in 2017), Oil States International's return is more dependent on global risk factors, since, as a multinational firm in the

[3] Another mechanism we may conjecture driving the sorting may be bank specialization in loans of particular currency denominations. I provide evidence in Section 1.2 showing that, in fact, global and local banks lend in both local and non-local currencies.

[4] For instance, global banks heavily recruit PhD economists to work in their macro research departments. See past and current job listings from global banks such as Citi, JP Morgan, and Goldman Sachs on the American Economic Association's Job Openings for Economists site: https://www.aeaweb.org/joe/listings.

petroleum industry, it is highly exposed to global demand and supply shocks. On the other hand, Zale Corporation's return is more exposed on local risk factors, since its main sources of sales revenue are local customers. The model predicts that on average, Oil States International is more likely to borrow from global banks, while Zale Corporation is more likely to borrow from local banks. The data confirms this prediction: banks that lend to Oil States International are mostly global banks, including Bank of Nova Scotia, Credit Suisse, and Royal Bank of Canada, while mostly local banks such as Bank of Boston, First Republic Bank Dallas, and Rhode Island Hospital Trust National Bank lend to Zale Corporation.

This result of firm-bank sorting based on banks' comparative informational advantage and firms' relative exposure to global and local risk reveals a problem of double adverse selection: both global and local banks are adversely selected against through firm selection, since firms select into borrowing from the bank which observes the more favorable component of their returns. I further demonstrate that this adverse selection problem has important macroeconomic implications. Given a funding shock to one of the banks, the adverse selection affects credit allocation at both the extensive and intensive margins, generating spillover and amplification effects through adverse interest rates. That is, a decrease in the funding cost of one of the bank types induces firm switching, attracting higher-return firms to contract with it (amplification effects) and leaving the other bank type with a riskier pool of firm (spillover effects). This constitutes a new channel through which monetary policy and liquidity shocks from abroad can be transmitted to firms. I test these predictions by analyzing how US and Euro area monetary policy shocks affect credit allocation in the Euro area, using tick-by-tick data on Federal Funds futures and Euribor futures to identify monetary policy shocks. The empirical results support the model predictions.

This adverse selection channel of international transmission is not only new relative to existing views on channels of international transmission through bank credit, but also clarifies the forces underlying the "international risk-taking channel" of monetary policy transmission.[5] It reveals that the empirical results which the existing literature (e.g., Morais et al. 2018) points to as evidence for risk-taking behavior by global banks could be confounded with a force generated by the adverse selection problem, namely, substitution between global banking credit and local banking credit.

The main features of the model are as follows. I consider an economy comprised of global and local banks, and firms that have returns dependent on global and local risk factors. There is perfect competition within each bank type. Each faces a problem of asymmetric information: global banks have the technology to extract information on global factors but not local factors, and vice versa for local banks. This double information asymmetry is common knowledge and thereby incorporated in the loan

[5] The international risk-taking channel of monetary policy transmission is based on the view that low monetary policy rates and QE in developed economies could induce banks to lend to riskier firms abroad (Bruno and Shin 2015a, Coimbra and Rey 2017, and Morais et al. 2018).

contracts offered by the banks. Consequently, each bank prices loans based on the component of firm return it observes, as well as its expectation of the component of return it does not observe for the subset of firms that selects the respective bank. Each bank type holds Nash-type conjectures about the other bank type's loan pricing and plays best response strategies. Firms, in turn, select the best loan contract. Given the setup, I characterize the equilibrium in the economy and then conduct comparative statics analysis to study how the equilibrium changes in response to changes in bank funding cost.

The model generates three sharp predictions. First, in equilibrium, firm-bank sorting and credit allocation are affected by double adverse selection. Both types of banks are adversely selected against through firm selection, since firms with higher expected return based on global factors relative to local factors are more likely to borrow from global banks, and vice versa for firms with higher expected return based on local factors. The intuition is straightforward. Given the information asymmetry, banks can only assign interest rates contingent on the component of firms' risk exposure that they observe (global or local), but not on the unobserved component, for which their rates must be uniform. Since firms select the bank that offers the best loan contract, they select into borrowing from the bank which observes the more favorable component of their return, resulting in adverse selection against banks. Moreover, banks, knowing firms' selection process, assign interest rates based on the expected risk of the firms which will approach them: they directly observe one component of risk, but assume the expected value of the other. As a result, relative to the first-best outcome, firms that are riskier in their unobserved exposure component face more favorable interest rates, and firms with relatively balanced global and local risk exposure face more adverse interest rates.

Second, shocks to bank funding costs affect credit allocation at the extensive margin. Specifically, suppose global banks face a decrease in funding cost due to expansionary monetary policy in the home country of the global banks. The model predicts that firms with relatively balanced global and local risk exposure components are more likely to switch into contracting with global banks. The result is driven by adverse selection: since the firms with relatively balanced global and local risk exposure are more adversely selected against, they are more likely to switch lenders given any changes in the credit market. These marginal firms that switch away from local banks into global banks are less risky than the infra-marginal firms that continue to borrow from either the local banks or the global banks.

Third, shocks to bank funding costs affect credit allocation at the intensive (interest rate) margin, and generate spillover and amplification effects. Continuing with the scenario of a lowering of global banks' funding cost due to expansionary monetary policy, the model predicts that i) the interest rates of the infra-marginal firms that remain with the local banks are expected to increase (i.e., a spillover effect), and ii) the interest rates of the infra-marginal firms that remain with the global banks are expected to decrease by more than the direct effect caused by the funding cost change (i.e., an amplification effect). The spillover effect on the infra-marginal firms that continue to borrow from local banks is solely driven by an exacerbation of the adverse selection problem. Since the marginal firms that switch away from local banks are less risky than these infra-marginal firms, local banks are left with a riskier pool of firms, which induces the banks to increase interest rates, despite no changes to their funding cost. On the other hand, the impact of the funding cost change is positively amplified for infra-marginal firms that continue to borrow from global banks because the marginal firms that switch into global banks are less risky than these infra-marginal firms, which alleviates the adverse selection problem for the global banks.

The model shows that adverse selection resulting from competitive interactions between banks with differing specialization in global versus local information forms a novel channel of international monetary policy transmission. Next, I formally test the three model predictions, using data on global syndicated corporate loans from Dealscan, matched with international firm-level databases including Amadeus, Orbis, Compustat, and Compustat Global. I further categorize the lead bank on each loan into global banks and local banks. The resulting sample includes 115,166 loans, borrowed by 12,979 firms across 24 countries, over the period 2004-2017. This cross-country firm-bank loan-level dataset is uniquely appropriate for this study because it captures a significant portion of cross-border lending that other loan datasets such as credit registry data would not capture.

To test the model prediction on firm-bank sorting, I implement an empirical strategy that tightly maps to the model set-up to construct measures for each firm's global and local risk exposure. I first compute a total exposure measure for each firm that can be interpreted as exposure to both demand and productivity risk, from which I estimate the firm's global and local risk exposure using principal component analysis. The results based on the new measures show a stark pattern of firm-bank sorting: as predicted by the model, global banks lend more to firms with higher exposure to global risk relative to local risk, and vice versa for local banks. I further show that, once I control for bank specialization in global and local information using the new measures, the firm-bank sorting patterns predicted by the traditional banking theory based on hard versus soft information are confirmed.

To test the model predictions of how funding shocks to banks affect credit allocation, I take the Euro area as an empirical laboratory and analyze how US and Euro area monetary policy shocks affect credit allocation across firms in the Euro area, through US and Euro area banks. To identify exogenous shocks to US and Euro area monetary policy, I use high-frequency data on Federal Funds futures and Euribor futures. I find that an expansionary shock to US monetary policy induces firms in the Euro area with relatively balanced global and local risk components to switch their borrowing from Euro area banks to US banks, conditional on Euro area monetary policy. The analogue applies to an expansionary shock to Euro area monetary policy.

Furthermore, I find that, conditional on Euro area (US) monetary policy and given expansionary US (Euro area) monetary policy, the interest rates of the infra-marginal firms that continue to borrow from Euro area (US) banks increase, reflecting a spillover effect. Specifically, a 25-basis-point expansionary US (Euro area) monetary policy shock increases the interest rate spread for the infra-marginal firms that continue to borrow from Euro area (US) banks by 22-27 (25-32) basis points. At the same time, the interest rate spreads of the infra-marginal firms that continue to borrow from US (Euro area) banks decrease, reflecting an amplification effect. A 25-basis-point expansionary shock to the US (Euro area) monetary policy decreases the interest rate spread for the infra-marginal firms that continue to borrow from US (Euro area) banks by 25-32 (34-40) basis points. The results are consistent with the model prediction on the effects of bank funding shocks on credit allocation at both the extensive and intensive margins, revealing an adverse selection channel of monetary policy transmission.

Related Literature The primary contribution of this paper—formalizing and providing empirical evidence of a novel micro-founded theory of credit allocation in globalized banking systems—adds to two broad strands of literature in finance, macroeconomics, and international finance: banking, and the macroeconomic implications of banking.

First, the new perspective I propose builds on the traditional information view of banking from classic papers by Campbel and Kracaw (1980), Diamond (1984), Ramakrishnan et al. (1984), and Boyd and Prescott (1986). They argue that the special role of banks derives from their ability to collect and process information. Through this lens, a subsequent strand of literature in banking and corporate finance, including Petersen and Rajan (1994), Stein (2002), Berger et al. (2005), and Liberti and Petersen (2018), argues that different banks specialize in hard versus soft information, and lend to different types of firms as a result. I provide evidence showing hard versus soft information is insufficient for explaining firm-bank sorting in globalized banking systems, and propose an alternative dimension of bank specialization.[6]

In the context of global banking specifically, this paper is related to the strand of banking literature that studies the effects of foreign bank entry on credit access. The framework developed in this paper builds on the work by Dell'Ariccia and Marquez (2004), Sengupta (2007), Detragiache et al. (2008), and Gormley (2014), which emphasize the importance of (imperfect) information in shaping competition and credit allocation in economies with both local banks and foreign banks. The focus of that line of studies is foreign bank entry into low-income countries, where overall information asymmetries may be large. Local banks are considered to have an informational advantage over the foreign banks, which, as a result, are able to target only the largest or the least informationally opaque firms. In contrast, the focus of this paper is cross-border lending by global banks in developed economies, where the majority of global banking activity occurs. What sets this paper apart is the new perspective on how banks' comparative advantage in different types of information, or global and local in-

[6] Section 1.2 describes the traditional theory and the relevant empirical tests in detail.

formation specifically, can affect credit allocation.[7] While the existing models predict that the smaller, more informationally opaque firms are more likely to borrow from local banks[8], the framework in this paper predicts that some large and informationally transparent firms are still likely to borrow from local banks, as long as their returns are more dependent on local risk factors.

Detragiache et al. (2008), Beck and Peria (2010) and Gormley (2014) also explore the impact of foreign banking on overall credit access, relating it to debates on the benefits and costs of financial openness. They argue that foreign banking entry undermines overall access to credit since it worsens the credit pool left to local banks, gives rise to adverse selection, and thereby lowers overall financial development. While my model also points to the possibility of a decline in aggregate credit due to adverse selection, I show that access to global banking credit actually leads to a more efficient credit allocation in the financial system. This is in line with papers which argue that the benefits of financial openness outweigh the costs, such as Levine (1996), Claessens et al. (2001), Edison et al. (2002), Claessens (2006), and Beck et al. (2007).

Second, this paper also contributes to the literature on the macroeconomic implications of banking. The global financial crisis put the spotlight on the importance of financial intermediaries for macroeconomic stability and monetary policy transmission.[9] In particular, global banks have emerged as a key channel for international transmission of liquidity conditions and monetary policy, sparking both theoretical and empirical research. On the theoretical front, several recent papers have introduced models with global banks for studying international transmission, including Dedola et al. (2013), De Blas and Russ (2013), Niepmann (2015), Bruno and Shin (2015b), and Aoki et al. (2016). While these models solely focus on emergence and implications of one type of bank,[10] this paper argues that the competitive interaction between global and local banks plays an important role for international transmission. On the empirical front, a growing literature uses bank-level and loan-level data to trace out the channels through which global banking affects domestic bank lending, including Cetorelli and Goldberg (2012b), Popov and Udell (2012), Schnabl (2012), De Haas

[7] The key ingredient incorporated in my model to formalize the idea of banks' differing specialization in global versus local information, double asymmetric information, and the ensuing result of double adverse selection, is new to the line of research in contract theory on adverse selection in credit markets, starting with the classic papers such as Stiglitz and Weiss (1981) and De Meza and Webb (1987).

[8] Papers including Berger et al. (2001), Clarke et al. (2005), Mian (2006), and Gormley (2010) provide empirical evidence in support of this prediction, though the empirical settings studied in these papers are all low-income economies.

[9] In the domestic macro literature, an emerging set of papers have introduced macroeconomic models with financial frictions in the form of balance sheet constraints on financial intermediaries to study aggregate economic activities, including Gertler and Kiyotaki (2010) and Gertler and Karadi (2011).

[10] In the framework in Bruno and Shin (2015b), there are both global and local banks. But local banks simply act as a conduit that intermediates funds from global banks to firms, which essentially make only one type of bank active in the economy.

and Lelyveld (2014), Ivashina et al. (2015), and Baskaya et al. (2017). This paper contributes to this line of work by pointing out a new channel of international transmission through global banks—adverse selection.

Furthermore, the adverse selection channel of international transmission raised in this paper is new to the literature on international transmission of monetary policy. Recent papers by Rey (2016) and Miranda-Agrippino and Rey (2018) provide evidence of large spillovers of US monetary policy on credit creation around the world, suggesting global banks as the main source for transmission. Existing work has pointed to currency mismatches on global banks' balance sheets (Ongena et al. 2017, Bräuning and Ivashina 2017, Bräuning and Ivashina 2018) and internal capital markets within global banks (Cetorelli and Goldberg 2012a) as channels of international monetary policy transmission. In addition, low international monetary policy rates and expansive quantitative easing in large developed economies over the past decade have prompted debates on the extent of a bank risk-taking channel of monetary policy transmission, as explained in Borio and Zhu (2012), Bruno and Shin (2015a), and Coimbra and Rey (2017). Morais et al. (2018), using firm-bank loan data, show that low monetary policy rates and QE in developed economies led global banks to increase credit supply to firms in Mexico, especially firms with higher-than-average ex-ante loan rates. They consider this to be evidence of bank risk-taking. Contrary to their explanation, I show that the force driving increased credit supply to riskier firms could be substitution between global banking credit and local banking credit, raising adverse selection as a new channel of international transmission of monetary policy.

The rest of this paper is structured as follows. Section 1.2 reviews the traditional theory and presents a new puzzle on firm-bank sorting in globalized credit markets. Section 1.3 presents the model. Section 1.4 applies the model to analyze the effects and implications of changes to bank funding costs on credit allocation. Section 1.5 outlines the model predictions and describes the data used for empirical testing. Section 1.6 presents the empirical strategy used to test the prediction on firm-bank sorting and discusses the results. Section 1.7 presents the empirical strategy used to test the predictions on credit allocation given bank funding shocks and discusses the results. Section 1.8 concludes. Proofs are relegated to A.1.

1.2 Traditional Theory and New Perspective

In this section, I review the traditional theory on firm-bank sorting and test whether it predicts the patterns of firm-bank sorting in globalized credit markets.

Classic banking theory argues that banks exist because of their unique ability to collect and process information. Based on this view, a long strand of literature in banking and corporate finance has used the distinction between hard and soft information to explain how banks and firms establish relationships (see, e.g., Petersen and Rajan 1994, Stein 2002, Petersen and Rajan 2002, and Liberti and Petersen 2018). Hard information is information that is quantifiable, independent of its collection process, and easy to transmit in impersonal ways. Soft information is information that is not easily quantifiable, dependent on its collection process, and requires context-specific knowledge to fully understand. Theories based on this view conjecture that large banks are more likely to lend to firms with more readily available hard information, while small banks are more likely to establish relationships with firms with more soft information.

As a first step to understand patterns of firm-bank sorting in globalized credit markets, I test whether the sorting patterns between firms and global versus local banks bear out the predictions of the traditional banking theory. Since global banks tend to be larger, I test whether global banks are more likely to lend to firms with more hard information, and local banks are more likely to lend to firms with more soft information, using a firm-bank loan-level dataset that spans across 24 countries and covers the period 2004-2017.[11] For measures of hard and soft information, I follow the empirical literature (e.g., Berger et al. 2005 and Mian 2006), which often uses firm asset size and firm age to proxy for hard information.

I sort firms into quartiles based on the distribution of firm asset size and firm age in each year in each country, and then calculate the proportion of loans given by global banks and local banks in each quartile. Figure 1.2 plots the distribution of lending from global and local banks over the entire sample. The plot shows that both global banks and local banks lend to firms of *all* sizes and ages, revealing that the traditional theory does not predict the pattern of firm-bank sorting in financial systems with both global and local banks.

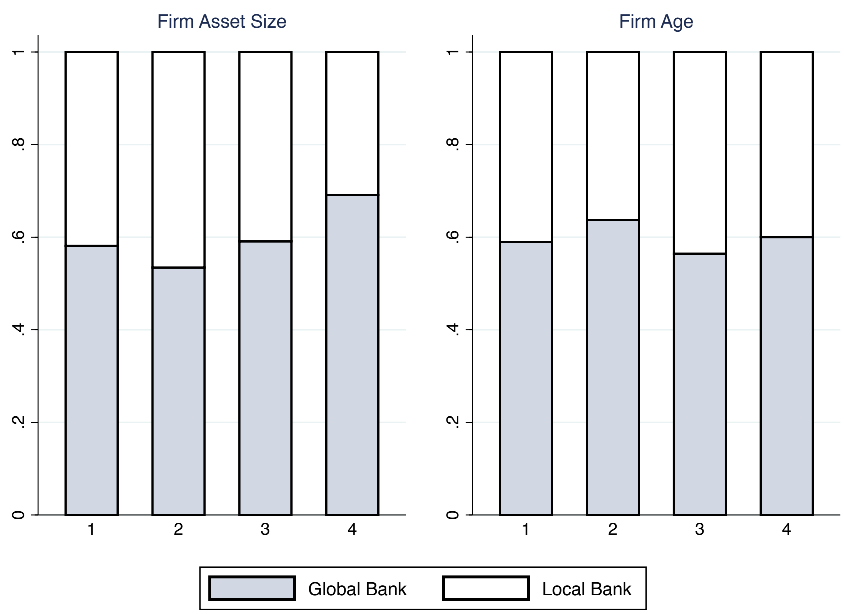
I further test whether the differences between global and local banks illustrated in Figure 1.2 are indeed insignificant in a statistical sense. For each given variable measuring hard information, I test whether the value-weighted mean of that variable for global banks is different from that for local banks. Table 1.1 presents these means and their differences. The results confirm the takeaways from the graphical analysis: the differences in value-weighted means are statistically *insignificant* between global and local banks for firm asset size and firm age.

Another conjecture about the mechanism driving the sorting between firms and global versus local banks may be bank specialization in loans of particular currency denominations. This is particularly motivated by recent papers by Maggiori et al. (2018) and Gopinath and Stein (2018) that highlight the prevalence of Dollar loans, and to a lesser extent Euro loans, in international financial markets. Given these considerations, I test whether global banks specialize in lending in non-local currencies, while local banks specialize in lending in local currency. As shown in Figure 1.3, in fact, global and local banks make loans in both local and non-local currencies. This empirical observation holds even when the US or both the US and Euro area countries are excluded from the sample.

The empirical evidence shows that the traditional banking theory of bank specializa-

[11] See Section 1.5 of the paper for a detailed discussion of the data and data-cleaning procedure.

Figure 1.2: **Firm-Bank Sorting, by Firm Size and Age Quartile**



Notes. The plot shows sorting patterns between firms and global versus local banks, with firms sorted into quartiles by asset size and age. The data sample consists of syndicated loans between global and local banks and firms across 24 countries from 2004-2017. Source: Dealscan, Amadeus, Orbis, Compustat, Compustat Global, and author's calculation.

tion in hard or soft information, as well as the view of bank specialization in particular currency denominations, are insufficient to explain observed sorting patterns between firms and global versus local banks. This points to a puzzle in the mechanism driving firm-bank sorting in globalized credit markets. In light of the puzzle, I propose a new perspective. I argue that global and local banks' differing specialization in global and local information constitutes a key mechanism for firm-bank sorting and credit allocation in financial systems with both types of banks. Global banks have a comparative advantage in extracting information on global risk, and local banks have a comparative advantage in extracting information on local risk.

This new perspective builds on the classic information view of banking. Furthermore, it incorporates global banks' unique position to acquire "global" information through global market-making activities and research efforts they invest in for analyzing global economic and market trends. Next, I proceed to formalize the new perspective by developing a model with global and local banks in which each bank type's comparative informational advantage serves as the key ingredient.

Table 1.1: **Firm-Bank Sorting, by Firm Size and Age Quartile: Statistical Test**

	(1) Size	(2) Age
(1) Mean: Global Bank	3.196***	2.759***
	(0.0299)	(0.0208)
(2) Mean: Local Bank	3.099***	2.726***
	(0.0674)	(0.0367)
(3) Difference	0.0969	0.0330
	(0.0716)	(0.0426)
Observations	115,166	114,323

Notes. The dependent variable in each regression (Y) is one of the hard information variables, firm size (column 1) or firm age (column 2), coded 1-4 based on the quartile number to which each respective firm belongs. Note the firms are sorted every year by country. Row 1 and row 2 show the means for each variable for global banks and local banks, respectively, by running a value-weighted regression of Y on a constant. For differences in means of the two types of banks, the whole data is used in the regression and a dummy for global banks is added (row 3). Standard errors reported in parentheses are clustered at the bank level. Source: Dealscan, Amadeus, Orbis, Compustat, Compustat Global, and author's calculation.

1.3 A Model with Global and Local Banking

In this section I develop a model to study firm-bank sorting and credit allocation in an economy with two types of banks—global banks and local banks—and firms heterogenous in their exposure to global and local risks. Each type of bank can perfectly observe only one component of firms' risk exposure, giving rise to a double information asymmetry. I show that firm-bank sorting and credit allocation in equilibrium reveal a problem of double adverse selection.

Set-up

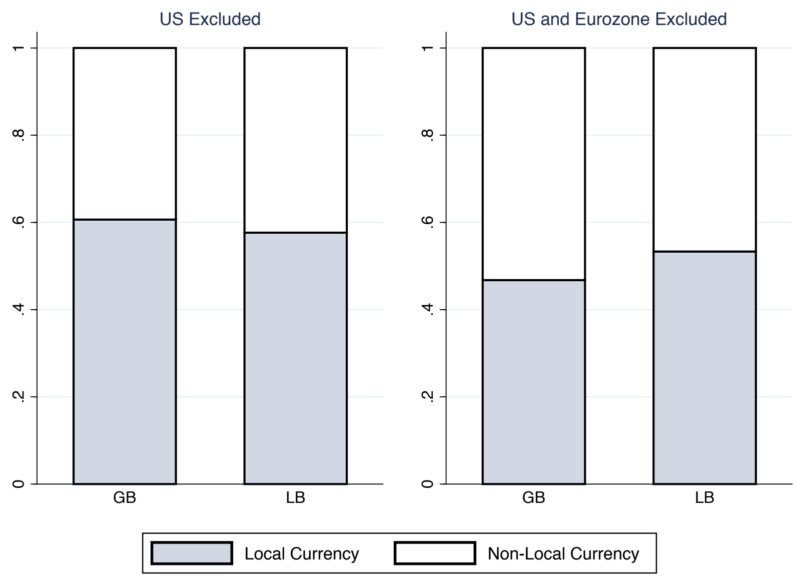
Consider an economy with two periods ($t = 0, 1$), a single good, and two classes of agents: firms and banks. All agents are risk neutral and cannot end with a negative amount of cash due to limited liability.

Firms. There is a continuum of heterogenous firms that have access to a fixed-size project requiring an investment of one. Each firm i's production technology is characterized by the following production function:

$$z_i = z_i^G + z_i^L + u_i \tag{1.1}$$

where z_i^G denotes firm i's component of return due to global risk, z_i^L denotes firm i's component of return due to local risk, and u_i denotes firm i's idiosyncratic risk.

Figure 1.3: **Loan Currency Denominations by Global and Local Banks**



Notes. The plot shows the share of loans in local currency versus non-local currency given by global and local banks. The left panel is based on loans from all countries in the sample except the US. The right panel is based on loans from all countries in the sample except the US and Euro area countries. Source: Dealscan, Amadeus, Orbis, Compustat, Compustat Global, and author's calculation.

Each component is independently and uniformly distributed, with $z_i^G \sim U(0,1)$, $z_i^L \sim U(0,1)$, and $u_i \sim U(0,1)$. More specifically, z_i^G can be considered to encompass two components, $z_i^G = \beta_i^G z^G$, where β_i^G denotes firm i's exposure to global risk and z^G denotes global risk. Similarly, z_i^L can be considered to encompass two components, $z_i^L = \beta_i^L z^L$, where β_i^L denotes firm i's exposure to local risk and z^L denotes local risk.[12]

Firms have full information on their returns due to global and local risk at the time of investment (period 0), while idiosyncratic risks are not realized until after investment (period 1). Firms have no private wealth; to implement the project, they need to raise one unit of external funds from a bank j through a loan contract.

Banks. There are two types of banks, *global banks* (G) and *local banks* (L), denoted as $j \in \{G, L\}$. They can enter the financial market and compete on projects by offering standard debt contracts. There is perfect competition within each bank type in the financial market.

The key feature that differentiates global banks from local banks is their information acquisition technology on global and local information: global banks have the

[12] These considerations become more applicable when mapping the model to empirics, which I describe more in detail in Section 1.6.

technology to evaluate firms' return due to global risk (z_i^G) but are not able to extract information on firms' return due to local risk (z_i^L), while local banks are able to evaluate firms' return due to local risk but are not able to extract information on firms' return due to global risk. This gives rise to an environment with double information asymmetry. The nature of the double information asymmetry problem and the distributions of the firms' return due to global risk and local risk are common knowledge across banks and firms.

Given the information structure, the loan rate offered by the two types of banks can be made contingent on the component of firm return observable to each respective bank type. Each type-contingent interest rate applies uniformly for all firms of the given observable component regardless of their unobserved return component. More specifically, global banks can offer type-contingent gross interest rate $R^G(z_i^G)$ for firms with return component z_i^G, and that rate applies for all firms with a given z_i^G regardless of z_i^L. Similarly, local banks can offer type-contingent interest rate $R^L(z_i^L)$ for firms with return component z_i^L, and that rate applies for all firms with a given z_i^L regardless of z_i^G.

It follows that the interest rates offered by each type of bank can be generated by interest rate functions that map the observable return components to type-contingent interest rates from the respective bank type: global banks offer contracts based on the interest rate function $\mathcal{R}^G : z_i^G \mapsto R^G(z_i^G)$, and local banks offer contracts based on the interest rate function $\mathcal{R}^L : z_i^L \mapsto R^L(z_i^L)$. For both types of banks, each bank's objective is to maximize expected profit across firms of each observable type: global banks maximize expected profit across firms of each given z_i^G, and local banks maximize expected profit across firms of each given z_i^L.

Global banks and local banks face gross funding rate r^G and r^L, respectively, for the funds they intermediate.[13]

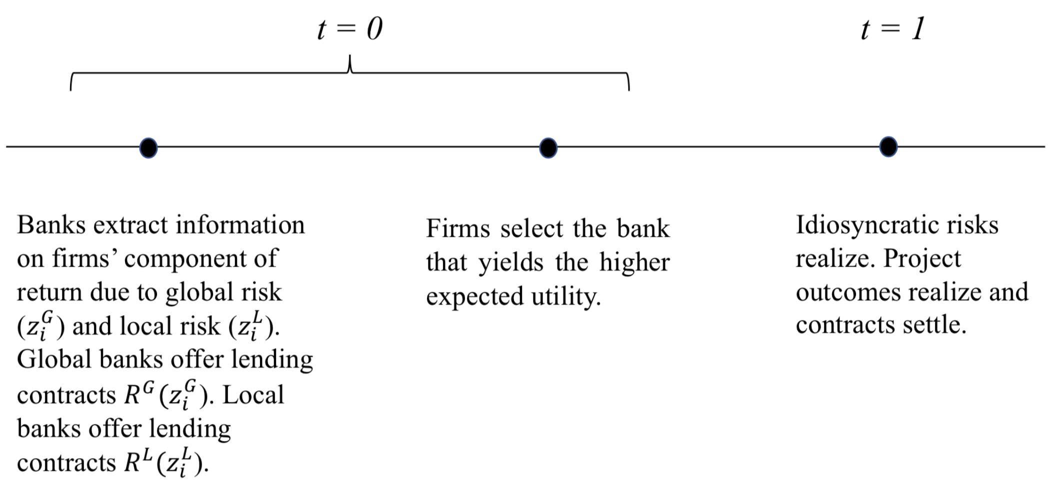
Firm-Bank Sorting. This environment in which each type of bank can perfectly observe only one component of the firms' return, while firms have full information on both return components, gives rise to a sorting process between banks and firms. The timing of the model is presented in Figure 1.4.

Let E_i denote the expectation of firm i conditional on its information set. Between global banks and local banks, each firm i selects the contract offered by bank $j \in \{G, L\}$ that yields the higher expected utility $E_i\big[\max(z_i - R^j(z_i^j), 0)\big]$.[14] Firm selection results in a partition of the set of all firms into two subsets, as each firm i with return component (z_i^G, z_i^L) selects to borrow from either a global bank or a local bank given the interest rate functions of the two bank types. One subset, denoted as S^G, chooses

[13] Since the funding market is not of central importance to this paper, it is not explicitly modeled for analytical convenience. The funding rates r^G and r^L could reflect funding conditions in the interbank market, the deposit market, or other risk premiums. While funding is fully elastic here, the model predictions do not change if r^G and r^L are considered to be decreasing in loan amounts.

[14] Note that the expectation here is taken with respect to idiosyncratic shocks only.

Figure 1.4: **Model Timeline**



to sign a lending contract with a global bank, and the other subset, denoted as S^L, chooses to sign a lending contract with a local bank:

$$S^G = \left\{ (z_i^G, z_i^L) : E_i\big[\max(z_i - \mathcal{R}^G(z_i^G), 0)\big] \geq E_i\big[\max(z_i - \mathcal{R}^L(z_i^L), 0)\big] \right\}; \tag{1.2a}$$

$$S^L = \left\{ (z_i^G, z_i^L) : E_i\big[\max(z_i - \mathcal{R}^L(z_i^L), 0)\big] > E_i\big[\max(z_i - \mathcal{R}^G(z_i^G), 0)\big] \right\}. \tag{1.2b}$$

The following assumptions about firm selection hold throughout the paper.

Assumption 1 *Suppose $\mathcal{R}^G(z_i^G) > z_i^G + z_i^L + 1$ or $\mathcal{R}^L(z_i^L) > z_i^G + z_i^L + 1$. Then $(z_i^G, z_i^L) \in S^G$ if $\mathcal{R}^G(z_i^G) \leq \mathcal{R}^L(z_i^L)$; and $(z_i^G, z_i^L) \in S^L$ otherwise.*

Assumption 1 states that in the region of the parameter space when the firm's expected utility is zero when it borrows from either a global bank or a local bank, it chooses the bank that offers the lower interest rate. This assumption ensures that there is no ambiguity in firm selection across all regions of the parameter space.

Remark 1 *Based on Equations (1.2a) and (1.2b) and Assumption 1, each firm i selects into borrowing from a global bank if and only if $\mathcal{R}^G(z_i^G) \leq \mathcal{R}^L(z_i^L)$, and each firm i selects into borrowing from a local bank if and only if $\mathcal{R}^G(z_i^G) > \mathcal{R}^L(z_i^L)$. In sum, each firm chooses the bank that offers the lowest rate.*

The selection of firms directly affects global and local banks' expected profits. Let E_G denote the expectation of a global bank conditional on its information set and E_L

denote the expectation of a local bank conditional on its information set. The expected profits for a global bank (G) from lending to firms of a given z_i^G and a local bank (L) from lending to firms of a given z_i^L are given by

$$\begin{aligned} \text{G:}\ \ & E_G[\pi_G(z_i^G)] = \int_{G_1} \min\left(z_i^G + z_i^L + u_i, \mathcal{R}^G(z_i^G)\right) dF_{G_1}(z_i^L, u_i) - r^G, \\ & \text{where}\ \ G_1(z_i^G) = \left\{(z_i^L, u_i) \mid z_i^L\colon (z_i^G, z_i^L) \in S^G,\ 0 \le u_i \le 1\right\}; \end{aligned} \tag{1.3a}$$

$$\begin{aligned} \text{L:}\ \ & E_L[\pi_L(z_i^L)] = \int_{L_1} \min\left(z_i^G + z_i^L + u_i, \mathcal{R}^L(z_i^L)\right) dF_{L_1}(z_i^G, u_i) - r^L, \\ & \text{where}\ \ L_1(z_i^L) = \left\{(z_i^G, u_i) \mid z_i^G\colon (z_i^G, z_i^L) \in S^L,\ 0 \le u_i \le 1\right\}. \end{aligned} \tag{1.3b}$$

The first term on the right hand side of Equations (1.3a) and (1.3b) is the expected gross return across loan contracts to firms of a given z_i^G and z_i^L for a global bank and a local bank, respectively. In the global bank's expected profit function, $G_1(z_i^G)$ summarizes the set of firms which select global banks given z_i^G. This includes firms with idiosyncratic risk u_i from any part of the u_i distribution, and z_i^L such that they are in the subset of firms that choose the global bank's contract. Similarly in the local bank's expected profit function, $L_1(z_i^L)$ summarizes the set of firms which select local banks given z_i^L. This includes firms with idiosyncratic risk u_i from any part of the u_i distribution, and z_i^G such that they are in the subset of firms that choose the local bank's contract. The integrand in both equations shows the relationship between bank profit and firm profit in a standard debt contract: for each firm, when its realized return is less than the contractual interest rate, it defaults and gives up any realized project returns to the lending bank; otherwise, the firm is able to repay the loan at the contractual rate and keeps the difference between the project return and rate as profit. $F_{G_1}(.)$ and $F_{L_1}(.)$ denote the cumulative distribution function of the relevant variable conditional on G_1 and L_1, respectively. The last terms in Equations (1.3a) and (1.3b) are the funding costs for the global bank and local bank, respectively.

Strategies and Equilibrium Definition

As shown in Equations (1.3a) and (1.3b), each type of bank's choice of the interest rate function affects the expected profit of the other type of bank since it influences the subset of firms that selects loan contracts from one versus the other. I consider the competitive interplay between a global bank and a local bank as a non-cooperative game.

In the game, the global bank's strategy set U^G consists of the set of possible interest rate functions $\mathcal{R}^G$, and the local bank's strategy set U^L consists of the set of possible interest rate functions $\mathcal{R}^L$. The payoff function for the global bank is the expected profit function $E_G[\pi_G(\mathcal{R}^G, \mathcal{R}^L)]$ across all firms, and that for the local bank is the expected profit function $E_L[\pi_L(\mathcal{R}^G, \mathcal{R}^L)]$.[15] A given strategy $\mathcal{R}^G$ is a best response to the strategy $\mathcal{R}^L$ if $E_G[\pi_G(\mathcal{R}^G, \mathcal{R}^L)] \geq E_G[\pi_G(\mathcal{R}^{G'}, \mathcal{R}^L)]\ \forall\, \mathcal{R}^{G'} \in U^G$, and vice versa for $\mathcal{R}^L$.

In a competitive equilibrium in this credit market, both global and local banks play best responses to each other's strategies. Each operating bank earns an expected profit of zero given perfect competition and free entry, and the selection of firms is consistent with the banks' equilibrium strategies.

Formally, the definition of the competitive equilibrium is as follows:

Definition 1 *For a given set of parameters r^G, r^L, and the distributions of z_i^G, z_i^L, and u_i, a competitive equilibrium with free entry in the credit market is a strategy profile $\{\mathcal{R}^G, \mathcal{R}^L\}$ and sets S^G and S^L such that:*

1. *(No Unilateral Deviation):*
   $E_G[\pi_G(\mathcal{R}^G, \mathcal{R}^L)] \geq E_G[\pi_G(\mathcal{R}^{G'}, \mathcal{R}^L)]\ \ \forall\, \mathcal{R}^{G'} \in U^G$;
   $E_L[\pi_L(\mathcal{R}^G, \mathcal{R}^L)] \geq E_L[\pi_G(\mathcal{R}^G, \mathcal{R}^{L'})]\ \ \forall\, \mathcal{R}^{L'} \in U^L$;

2. *(Zero Profit Condition, Global Bank):*
   $$\int_{G_1} \min\left(z_i^G + z_i^L + u_i, \mathcal{R}^G(z_i^G)\right) dF_{G_1}(z_i^L, u_i) = r^G;$$

3. *(Zero Profit Condition, Local Bank):*
   $$\int_{L_1} \min\left(z_i^G + z_i^L + u_i, \mathcal{R}^L(z_i^L)\right) dF_{L_1}(z_i^G, u_i) = r^L;$$

4. *(Firm Selection):*
   $$S^j = \left\{(z_i^G, z_i^L) : E_i\{\max[z_i - \mathcal{R}^j(z_i^j), 0]\} \geq E_i\{\max[z_i - \mathcal{R}^k(z_i^k), 0], j \neq k \in \{G, L\}\right\}.$$

Part 1 of Definition 1 requires that no unilateral deviation in strategy by any bank is profitable for that bank. Parts 2 and 3 impose zero profit among global banks and local banks, respectively. Part 4 defines the set of firms that select the loan contract with either of the two types of banks in an incentive-compatible fashion. All banks that enter the market hold correct expectations about both banks' pricing choices and the pool of firms that will accept the contract. As a consequence, the allocations of credit across firms are consistent with the banks' equilibrium strategies.

Before turning to characterizing the equilibrium in the credit market of two bank types under double information asymmetry, I describe two useful benchmarks.

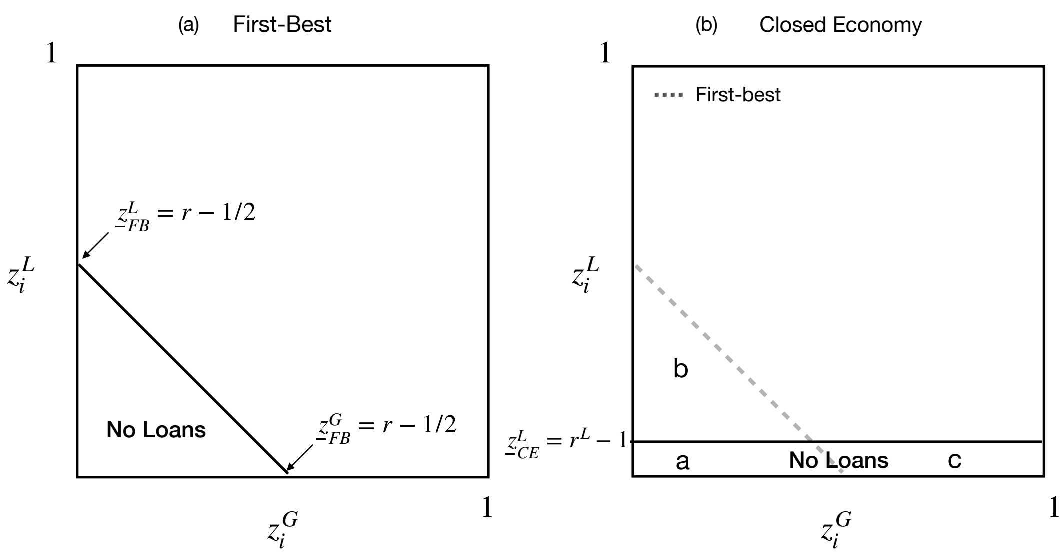
First Best. In an environment where both types of banks observe full information on each firm's return due to global and local risk, the only margin that differentiates

[15] Banks also strictly prefer making a loan with zero expected profit to not making a loan.

the loan rate charged by global banks versus local banks is the funding cost faced by each bank type. As a result, only the bank type with lower funding cost (r) exists in the credit market in equilibrium, and its interest rate function is strictly decreasing in $(z_i^G + z_i^L)$. Panel (a) of Figure 1.5 shows an illustration of the first-best equilibrium in an economy with full information. The diagonal line $\underline{z}_{FB}^L + \underline{z}_{FB}^G + 1/2 = r$ denotes a threshold.[16] The firms in the region below this threshold are not able to receive loans, as their expected profits are too low for the bank to break even in expectation.

Closed Economy. In an environment where there exist only local banks that observe information on each firm's return due to local risk, the interest rate function $\mathcal{R}^L(z_i^L)$ is strictly decreasing in z_i^L and uniform across the entire distribution of z_i^G. Panel (b) of Figure 1.5 shows an illustration of the equilibrium in this economy. Firms with z_i^L below $\underline{z}_{CE}^L = r^L - 1$ (firms in Regions a and c) are not able to receive loans. Relative to the first-best allocation without information asymmetries, the equilibrium in a closed economy overfunds firms whose return due to local risk is high relative to return due to global risk (firms in Region b) and underfunds firms whose return due to local risk is low relative to return due to global risk (firms in Region c).

Figure 1.5: **Benchmark Equilibrium: First-Best and Closed Economy**



Notes. Panel (a) illustrates the first-best equilibrium in an economy with full information. Panel (b) illustrates the equilibrium credit allocation in a closed economy in which there are only local banks.

Equilibrium Characterization

In the following I characterize the equilibrium in a credit market of two bank types under double information asymmetry. I start by establishing the properties of the bank

[16] Note $E[u_i] = 1/2$.

interest rate functions in equilibrium.

Subject to the zero profit conditions from Parts 2 and 3 of Definition 1, Equation (1.3a) determines the global banks' type-contingent interest rate function $\mathcal{R}^G$ given firm selection as specified in Equation (1.2a), and Equation (1.3b) determines the local banks' type-contingent interest rate function $\mathcal{R}^L$ given firm selection as specified in Equation (1.2b). Since firm selection depends on interest rates from both types of banks in equilibrium, Equations (1.3a) and (1.3b) given Equations (1.2a) and (1.2b) simultaneously determine the type-contingent interest rate functions $\mathcal{R}^G$ and $\mathcal{R}^L$ in equilibrium.

Let $E_G[z_i^L \mid (z_i^G, z_i^L) \in S^G, z_i^G]$ denote the global banks' expectation of the average z_i^L for the set of firms with (z_i^G, z_i^L) in S^G conditional on z_i^G, and $E_L[z_i^G \mid (z_i^G, z_i^L) \in S^L, z_i^L]$ denote the local banks' expectation of the average z_i^G for the set of firms with (z_i^G, z_i^L) in S^L, conditional on z_i^L. Proposition 1 characterizes $\mathcal{R}^G$ and $\mathcal{R}^L$.

Proposition 1 *(Type-Contingent Interest Rate Functions)*

1. *$\mathcal{R}^G$ is strictly decreasing in z_i^G for $z_i^G \in [\underline{z}^G, 1]$, where $\underline{z}^G \equiv r^G - E_G[z_i^L \mid (\underline{z}^G, z_i^L) \in S^G, z_i^G] - 1/2$.*
2. *$\mathcal{R}^L$ is strictly decreasing in z_i^L for $z_i^L \in [\underline{z}^L, 1]$,, where $\underline{z}^L \equiv r^L - E_L[z_i^G \mid (z_i^G, \underline{z}^L) \in S^L, z_i^L] - 1/2$.*

Part 1 of Proposition 1 establishes that the global banks' interest rate function is strictly monotone for $z_i^G \in [\underline{z}^G, 1]$. The lower bound $\underline{z}^G$ pins down a cut-off point on z_i^G below which the expected profits of the pertinent firms are too low for the global banks to break even in expectation. In other words, $\underline{z}^G$ defines the lowest z_i^G firm to which the global banks lend. The lower bound $\underline{z}^G$ is increasing in global bank's funding cost (r^G), decreasing in the average z_i^L of the set of firms that are expected to select the global bank, and decreasing in the expected idiosyncratic shocks for firms. The explanation for local banks' interest rate function $\mathcal{R}^L$ established in Part 2 of Proposition 1 is analogous. Panel (a) of Figure 1.6 illustrates the interest rate functions in a graph with z_i^L on the x-axis. Since global banks cannot observe z_i^L, $\mathcal{R}^G$ is uniform across z_i^L. $\mathcal{R}^L$ is strictly decreasing in z_i^L, as established in Proposition 1.

Using strict monotonicity, I next establish that the competitive interplay between global and local banks generates a unique form of horizontal segmentation in equilibrium, in which there exists a set of marginal firms that are indifferent between taking loans from global banks and local banks. Formally,

Proposition 2 *(Threshold Functions) Let $RG = \{\mathcal{R}^G(z_i^G) \mid z_i^G \in [\underline{z}^G, 1]\}$ and $RL = \{\mathcal{R}^L(z_i^L) \mid z_i^L \in [\underline{z}^L, 1]\}$. In the region $RG \cap RL$, there exist threshold functions $\bar{z}^L(z_i^G)$ and $\bar{z}^G(z_i^L)$ such that:*

1. *$\mathcal{R}^G(z_i^G) = \mathcal{R}^L(\bar{z}^L(z_i^G))$.*
   $\mathcal{R}^L(z_i^L) = \mathcal{R}^G(\bar{z}^G(z_i^L))$.

2. $S^G = \{(z_i^G, z_i^L) : z_i^L \leq \bar{z}^L(z_i^G)\}$, *and* $S^L = \{(z_i^G, z_i^L) : z_i^L > \bar{z}^L(z_i^G)\}$.

   $S^L = \{(z_i^G, z_i^L) : z_i^G < \bar{z}^G(z_i^L)\}$, *and* $S^G = \{(z_i^G, z_i^L) : z_i^G \geq \bar{z}^G(z_i^L)\}$.

Part 1 of Proposition 2 establishes that, for every firm with z_i^G (resp. z_i^L), there exists a threshold on z_i^L (resp. z_i^G), denoted as $\bar{z}^L(z_i^G)$ (resp. $\bar{z}^G(z_i^L)$), at which both the global bank and local bank offer the same interest rate. Panel (b) of Figure 1.6 illustrates the threshold: for a given z_i^G, there exists a threshold $\bar{z}^L(z_i^G)$ at which the interest rate functions of the two banks intersect, $\mathcal{R}^G(z_i^G) = \mathcal{R}^L(\bar{z}^L(z_i^G))$.

Part 2 of Proposition 2 follows from the monotonic property of the type-contingent interest rate. Given $\mathcal{R}^G(z_i^G)$ and $\mathcal{R}^L(z_i^L)$ are strictly decreasing in z_i^G and z_i^L, respectively, firms (z_i^G, z_i^L) with $z_i^L < \bar{z}^L(z_i^G)$ face a lower rate from global banks and therefore select global banks (i.e, the firms are in S^G). Firms with $z_i^L > \bar{z}^L(z_i^G)$ face a lower rate from local banks and thereby select local banks (i.e, they are in S^L). This idea is shown in Panel (b) of Figure 1.6. An analogous explanation applies to firms with $z_i^G < \bar{z}^G(z_i^L)$ and $z_i^G > \bar{z}^G(z_i^L)$.

Parts 1 and 2 of Proposition 2 establish the existence of thresholds that segment the credit market into two parts, with global banks as the lender in one, and local banks as the lender in the other. In equilibrium, the threshold values $\bar{z}^L(z_i^G)$ and $\bar{z}^G(z_i^L)$ are determined by the interaction between the interest rate schedules of the global and local banks, where $\bar{z}^L(z_i^G) = (\mathcal{R}^L)^{-1}(\mathcal{R}^G(z_i^G))$ and $\bar{z}^G(z_i^L) = (\mathcal{R}^G)^{-1}(\mathcal{R}^L(z_i^L))$.

The following corollary characterizes the threshold functions, describing how they change given changes in z_i^G, z_i^L, and the interest rate functions. Let $\tilde{z}^G$ be a cut-off that pins down an upper bound on z_i^G, above which firms with z_i^L from any part of the z_i^L distribution are expected to select the global bank (i.e., $\bar{z}^L(z_i^G) = 1$ for all $z_i^G \geq \tilde{z}^G$), and the analogue applies to $\tilde{z}^L$.

Corollary 1 *(Threshold Functions Characterization) Let* $\tilde{z}^G = min\{z_i^G : \bar{z}^L(z_i^G) = 1\}$ *and* $\tilde{z}^L = min\{z_i^L : \bar{z}^G(z_i^L) = 1\}$.

1. $\bar{z}^L(z_i^G)$ *is increasing in* z_i^G *for* $z_i^G \in [\underline{z}^G, \min(\tilde{z}^G, 1)]$.

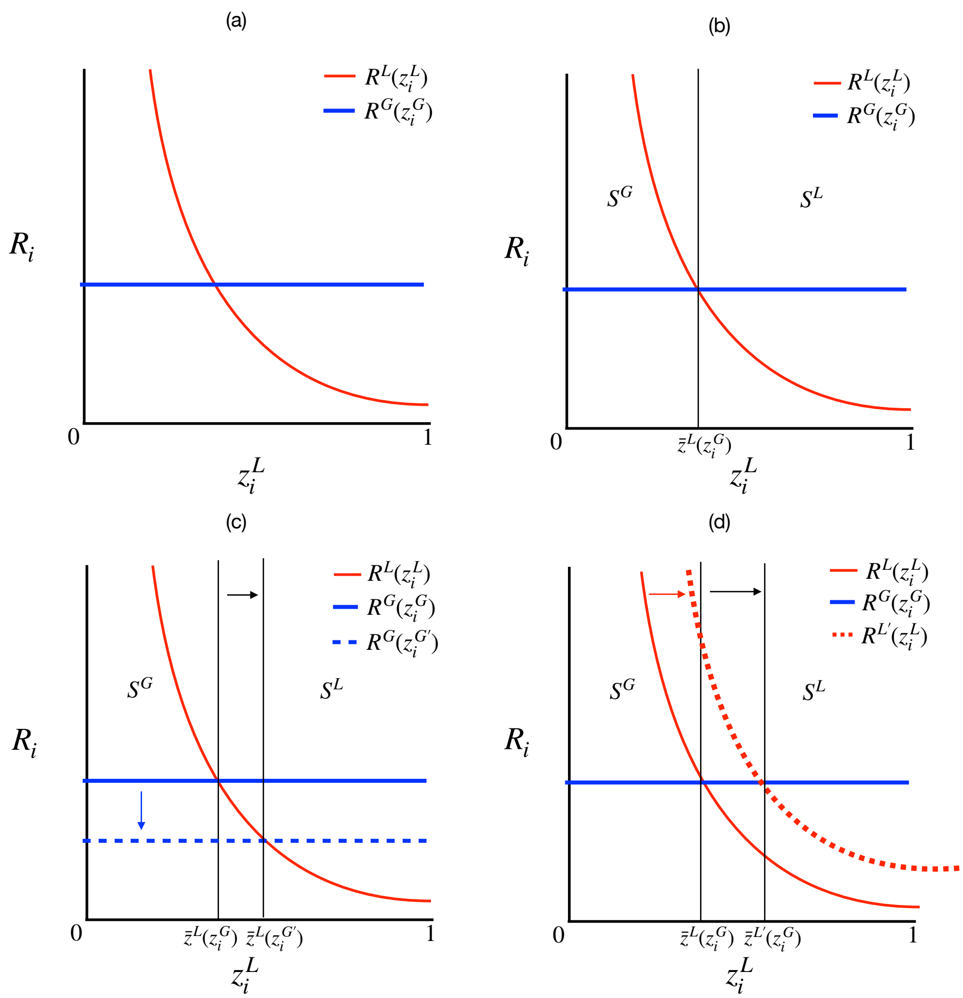
   $\bar{z}^G(z_i^L)$ *is increasing in* z_i^L *for* $z_i^L \in [\underline{z}^L, \min(\tilde{z}^L, 1)]$.

2. $\bar{z}^G(z_i^L)$ *is decreasing in* $\mathcal{R}^L(z_i^L)$ *and* $\bar{z}^L(\bar{z}^G(z_i^L))$ *is increasing in* $\mathcal{R}^L(z_i^L)$*, for* $z_i^G \in [\underline{z}^G, \min(\tilde{z}^G, 1)]$ *and* $z_i^L \in [\underline{z}^L, \min(\tilde{z}^L, 1)]$.

   $\bar{z}^L(z_i^G)$ *is decreasing in* $\mathcal{R}^G(z_i^G)$ *and* $\bar{z}^G(\bar{z}^L(z_i^G))$ *is increasing in* $\mathcal{R}^G(z_i^G)$*, for* $z_i^G \in [\underline{z}^G, \min(\tilde{z}^G, 1)]$ *and* $z_i^L \in [\underline{z}^L, \min(\tilde{z}^L, 1)]$.

The intuition for Part 1 of Corollary 1 is straightforward. Suppose there is an increase in z_i^G from z_i^G to $z_i^{G'}$, or in other words, the global component of firm i's return strengthens. Global banks' expected profit increases, and perfect competition drives down $\mathcal{R}^G(z_i^G)$. At the margin, this attracts firms with higher z_i^L to contract with

Figure 1.6: **Illustration of Interest Rate Functions and Threshold Functions**



Notes. Panel (a) illustrates Proposition 1, showing the monotonically decreasing property of the interest rate functions, given information asymmetry. Panel (b) illustrates Part 1 and 2 of Proposition 2, showing, for a given z_i^G, there exists a threshold $\bar{z}^L(z_i^G)$ at which $\mathcal{R}^G(z_i^G) = \mathcal{R}^L(\bar{z}^L(z_i^G))$. Firms below the threshold borrow from global banks; firms above which borrow from local banks. Panel (c) illustrates Part 3 of Proposition 2, showing an increase in z_i^G lowers $\mathcal{R}^G(z_i^G)$ and increases $\bar{z}^L(z_i^G)$, holding all else constant. Panel (d) illustrates Part 4 of Proposition 2, showing an increase in $\mathcal{R}^L(z_i^L)$ increases $\bar{z}^L(z_i^G)$, holding all else constant.

global banks. Thus, the threshold on z_i^L increases, $\bar{z}^L(z_i^{G'}) > \bar{z}^L(z_i^G)$. This relationship is illustrated in Panel (c) of Figure 1.6.

The intuition for Part 2 of Corollary 1 (shown in Panel (d) of Figure 1.6) is as follows. Suppose the local banks' interest rate function changes such that $\mathcal{R}^L(z_i^L)$ increases for some $z_i^L \in [\underline{z}^L, \min(\tilde{z}^L, 1)]$. A higher interest rate induces a set of marginal firms to switch from contracting with local banks to global banks, holding constant z_i^G and $\mathcal{R}^G(z_i^G)$. In particular, the local component (z_i^L) of the switching firms is greater than that of the firms in global banks' original portfolio, which implies an increase of the threshold $\bar{z}^L(z_i^G)$. At the same time, the global component (z_i^G) of the switching firms is higher than that of the firms that remain with local banks, which implies a decrease of the threshold $\bar{z}^G(\bar{z}^L(z_i^G))$.

Based on the results from Proposition 1 and 2 and Corollary 1, I next characterize the competitive interaction between the two interest rate functions offered by the two types of banks.

Proposition 3 *(Interaction of Rate Functions in Equilibrium) Given z_i^G, for any increase in $\mathcal{R}^L(z_i^L)$ such that $\bar{z}^L(z_i^G)$ increases, $\mathcal{R}^G(z_i^G)$ declines. Given z_i^L, for any increase in $\mathcal{R}^G(z_i^G)$ such that $\bar{z}^G(z_i^L)$ increases, $\mathcal{R}^L(z_i^L)$ declines.*

Proposition 3 points out that each bank's type contingent interest rate function is determined by two inputs: the observed risk component of each firm's return and the threshold value of the unobserved risk component. For a given z_i^G, if there is a change in the local banks' interest rate function $\mathcal{R}^L$ such that the threshold $\bar{z}^L(z_i^G)$ increases, a set of marginal firms with z_i^L greater than all the z_i^L's in global banks' original portfolio switches into borrowing from global banks. As a result, the global banks offer a lower $\mathcal{R}^G(z_i^G)$ for the firms with the given z_i^G. The interaction between the interest rates functions of global and local banks point to a stable equilibrium in which the two banks interact as strategic substitutes.

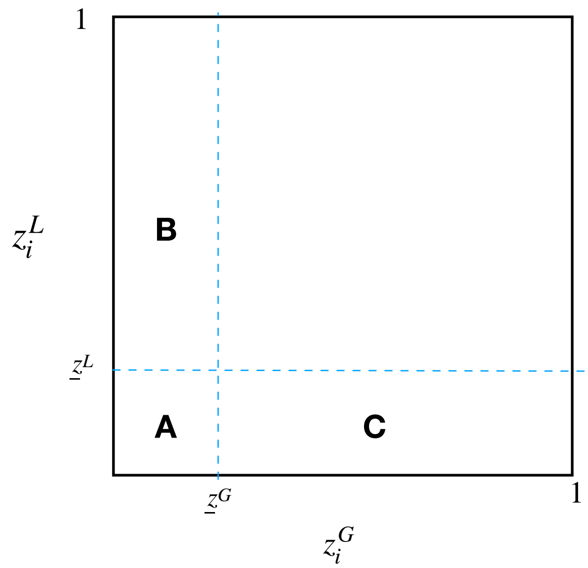
Propositions 1–3 lead to a full characterization of the equilibrium solution on $\mathcal{R}^G$ and $\mathcal{R}^L$. The solutions for the equilibrium interest rates $\mathcal{R}^G(z_i^G)$ and $\mathcal{R}^L(z_i^L)$, and thresholds $\bar{z}_i^L = \bar{z}^L(z_i^G)$ and $\bar{z}_i^G = \bar{z}^G(z_i^L)$, for $z_i^G \in [\underline{z}^G, 1]$ and $z_i^L \in [\underline{z}^L, 1]$ are described in detail in A.1, A.1.

Firm-Bank Sorting Under Double Asymmetric Information

I proceed to study the equilibrium firm-bank sorting in this economy.

Symmetric Equilibrium. To build intuition, I first study firm-banking sorting in the case where global and local banks face the same funding cost, $r^G = r^L = r$. This can be motivated by the idea that both types of banks have access to funds from a global interbank market that provides an elastic supply of funds at the risk-free interest

Figure 1.7: **Firm-Bank Sorting Firm Space**



Notes. The plot summarizes all the firms in the economy. The bounds $\underline{z}^G$ and $\underline{z}^L$ define the cut-offs below which global banks and local banks, respectively, would not make loans. Firms in Region A are not offered loans. Firms in Region B can only receive loans from local banks. Firms in Region C can only receive loans from global banks.

rate r. This case allows me to focus solely on the implications of the double information asymmetry on firm-bank sorting.

Given the assumption $r^G = r^L = r$, the expected profit functions of the global and local banks become completely symmetric. With perfect competition and free entry, the equilibrium thresholds also become symmetric.

Lemma 1 *(Thresholds: Symmetric Case) If $r^G = r^L$, then $\bar{z}^L(z_i^G) = z_i^G$ and $\bar{z}^G(z_i^L) = z_i^L$.*

Given Lemma 1, sorting between firms and global versus local banks immediately follows.

Corollary 2 *(Firm-Bank Sorting: Symmetric Case) Let $r^G = r^L$. A firm selects a global bank if and only if $z_i^G \geq z_i^L$. A firm selects a local bank if and only if $z_i^L > z_i^G$.*

Panel (a) of Figure 1.8 provides a simple illustration of firm-bank sorting for the symmetric case. Global and local banks compete for loans borrowed by firms with $z_i^G \in [\underline{z}^G, 1]$ and $z_i^L \in [\underline{z}^L, 1]$. In equilibrium, the thresholds form a 45 degree line that segments the credit market. Firms in Region L, which have $z_i^L > z_i^G$, select into local banks, and firms in Region G, which have $z_i^G \geq z_i^L$, select into global banks.

Corollary 2 reveals that the information asymmetry problem faced by global and local banks creates a segmented credit market affected by double adverse selection. Both types of banks are adversely selected against, as firms select into borrowing from the bank which observes the more favorable component of their risk exposure. Specifically, firms with a weaker local component (z_i^L) relative to their global component (z_i^G) select into a global bank — the bank that cannot observe the weaker component.

Furthermore, firms are borrowing at higher interest rates relative to the first-best outcomes. This is because banks, given the information asymmetry problem, can only assign interest rates contingent on the component of firms' risk exposure that they observe, but not on the unobserved component, for which their rates must be uniform, as shown by the iso-interest rate curves in Panel (b) of Figure 1.8. Knowing the firm selection process, they assign interest rates based on the expected risk of the firms which will approach them. This gives rise to heterogeneity among firms in the degree to which they are charged higher interest rates relative to the first-best outcomes. The firms that are riskier in their unobserved exposure component face more favorable interest rates, and firms with relatively balanced global and local risk exposure (i.e., closer to the thresholds) face more adverse interest rates. Specifically, consider firms a and b in Panel (a) of Figure 1.8. In this economy, both firms select into borrowing from a global bank in equilibrium, and are offered the same interest rate $\mathcal{R}^G(z_i^G)$ since their z_i^G component is the same. However, the z_i^L component of firm a is much stronger than that of b, which means that firm a faces a worse outcome relative to the first-best outcome.

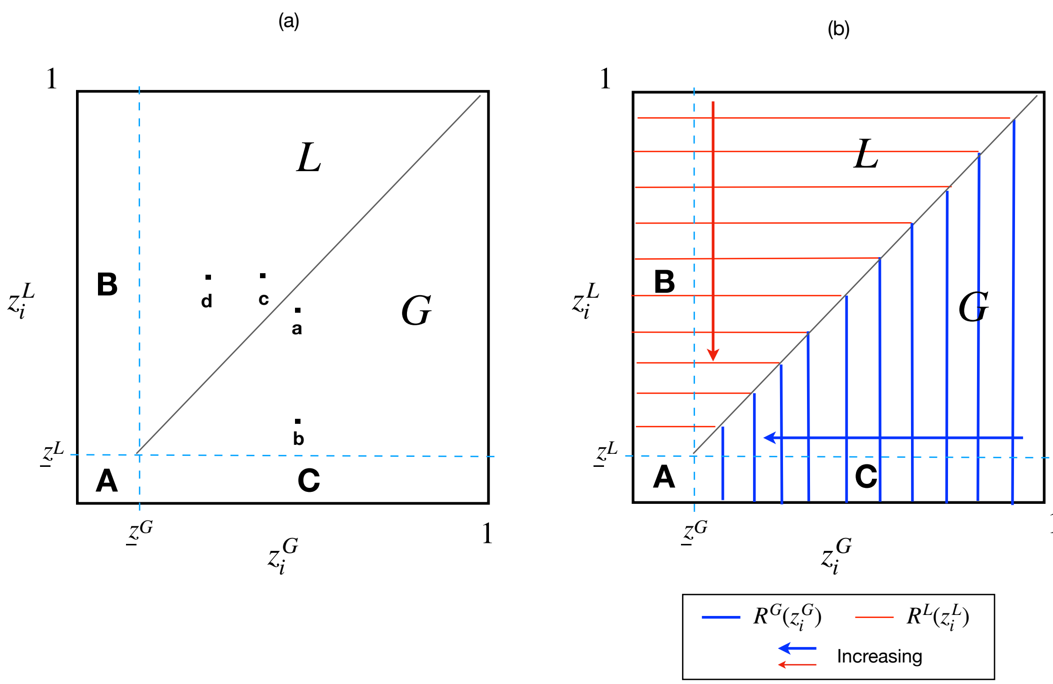
Asymmetric Equilibrium. Next I solve the model numerically to study firm-bank sorting in the general case when there is variation between the funding costs of global and local banks ($r^G \neq r^L$).

Panel (a) of Figure 1.9 provides an illustration of the equilibrium when $r^G < r^L$, where $r^G = 1.00$ and $r^L = 1.01$. Compared to the symmetric case, global banks are able to capture a greater share of the loan market given their funding advantage. In particular, they are able to attract all the firms with $z_i^G > \tilde{z}^G$, and they provide loans to firms with lower z_i^G components than before, since the cut-off $\underline{z}^G$ is increasing in r^G (Equation (A.3)).

Panel (b) of Figure 1.9 illustrates the equilibrium when $r^G > r^L$, where $r^L = 1.00$ and $r^G = 1.01$. The results are analogous.

Closed Economy vs. Financial Integration. An interesting counterfactual to consider is how this financially integrated economy compares with the benchmark closed economy, in terms of firm-bank sorting, aggregate credit, and efficiency. In a closed economy where there are only local banks, firms with $z_i^L < r^L - 1$ are considered too risky to get loans (illustrated in Panel (b) of Figure 1.5). With financial integration, most of those firms, specifically firms with $z_i^G > \underline{z}^G$, would be able to get loans from

Figure 1.8: **Firm-Bank Sorting and Interest Rates Under Symmetric Equilibrium**

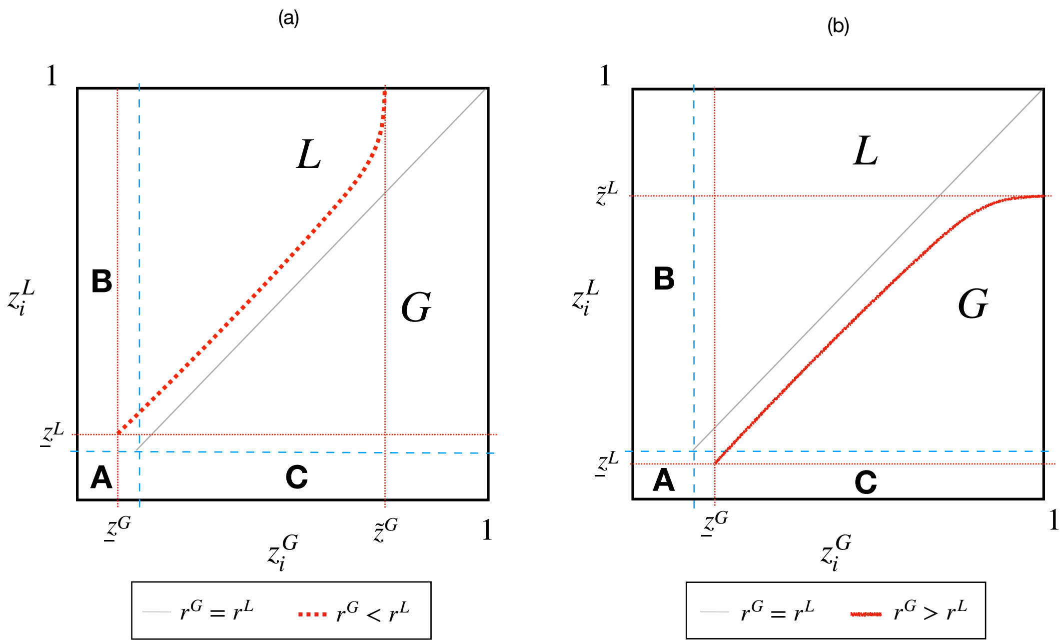


Notes. Panel (a) shows the equilibrium firm-bank sorting when $r^G = r^L$. Panel (b) shows iso-interest rate curves by global banks and local banks. For both plots, Region A depicts the region where no loans are given. Region B depicts the region where only local bank loans are given and no global banks would give loans. Region C depicts the region where only global bank loans are given and no local banks would give loans. Region L depicts the region where both global and local bank compete for loans, and loans are given by local banks in equilibrium. Region G depicts the region where both global and local bank compete for loans, and loans are given by global banks in equilibrium.

global banks (firms in Region n in Panel (a) of Figure 1.10). Furthermore, a set of firms with stronger global components (z_i^G) relative to their local components (z_i^L) would switch into borrowing from global banks (firms in Region G in Panel (a) of Figure 1.10), since they would receive lower interest rates from global banks, as shown in Panel (b) of the figure. Those firms would all benefit from financial integration.

However, the switching of firms leaves local banks with a riskier pool of firms, inducing an increase in interest rate for the infra-marginal firms that remain with local banks (firms in Region L in Panel (a) of Figure 1.10), as shown in Panel (b) of the Figure. This means that financial integration can give rise to an adverse selection problem. Moreover, this adverse selection problem would force a set of firms to exit the credit market (firms in Region e in Panel (a) of Figure 1.10). This result suggests that financial integration can induce a decline in aggregate credit due to adverse selection, which is in line with the arguments raised in Detragiache et al. (2008) and Gormley (2014).

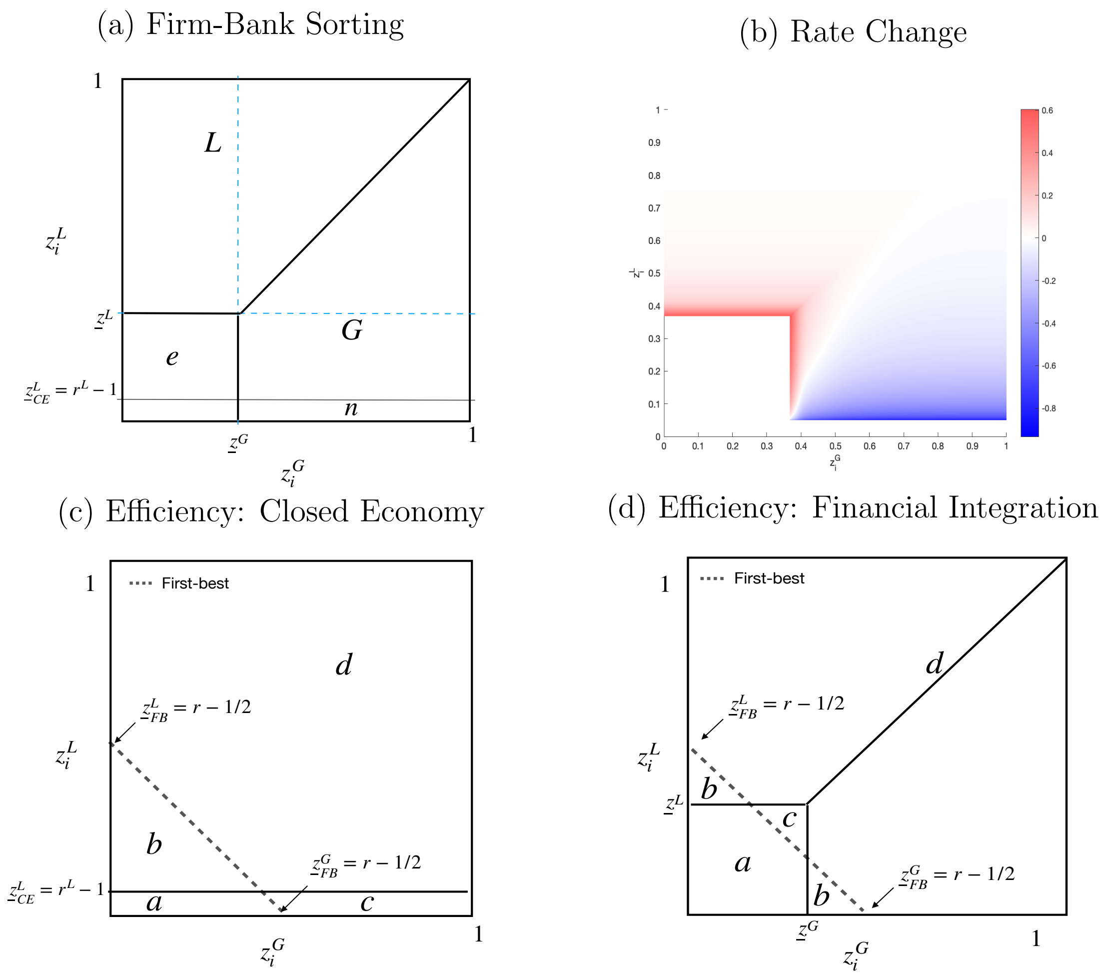
Figure 1.9: **Firm-Bank Sorting under Asymmetric Equilibrium**



Notes. Panel (a) illustrates the firm-bank sorting when $r^G < r^L$, where $r^G = 1.00$ and $r^L = 1.01$. Panel (b) illustrates the firm-bank sorting when $r^G > r^L$, where $r^L = 1.00$ and $r^G = 1.01$. For both plots, Region A depicts the region where no loans are given. Region B depicts the region where only local bank loans are given and no global banks would give loans. Region C depicts the region where only global bank loans are given and no local banks would give loans. Region L depicts the region where both global and local banks compete for loans, and loans are given by local banks in equilibrium. Region G depicts the region where both global and local banks compete for loans, and loans are given by global banks in equilibrium.

Despite the potential decline in aggregate credit, it is important to point out that credit allocation in a fully integrated financial system is more efficient relative to a closed economy. I define efficiency in terms of how closely credit allocation corresponds to that in the benchmark full-information economy. As shown in Panels (c) and (d) in Figure 1.10, in a full information economy, firms in Regions a and b would not get loans, and firms in regions c and d would get loans. In both a closed economy and a financially integrated economy, firms in Region b are overfunded, while firms in Region c are underfunded. Nevertheless, for all reasonable parameters values, Regions b and c in a financially integrated economy are smaller than the corresponding regions in a closed economy. Quantitatively, let efficiency be defined as the share of total credit in the economy relative to the benchmark full-information economy ($Efficiency = 1-(b+c)/(a+b+c+d)$ based on the illustrations Panels (c) and (d) in Figure 1.10). Given parameter values $r^G = 1.05$ and $r^L = 1.05$, the closed economy is 85% efficient, while a financially integrated economy is 95% efficient.

Figure 1.10: **Effects of Financial Integration**



Notes. Panel (a) characterizes the equilibrium characterization after financial integration. Relative to a closed economy, upon financial integration, firms in Region e are no longer able to get loans, and firms in Region n are able to get loans. Panel (b) shows the interest rate change as measured by $\Delta R_i = R_i^{\text{FI}} - R_i^{\text{CE}}$ upon financial integration. The plot is based on simulations using parameter values $r^G = 1.05$ and $r^L = 1.05$. Panel (c) and (d) compares the firm space in a closed economy and a financially integrated economy, respectively, to that in the benchmark full-information economy. According to the first-best outcome, firms in Regions a and b would not get loans, and firms in Regions c and d would get loans.

1.4 Comparative Analysis and Implications

In this section, I explore the macroeconomic implications of the model. I study how the equilibrium credit allocation responds to changes in banks' funding cost (e.g., a change in monetary policy of the home country of one of the banks) at both the extensive (firm selection) and intensive (interest rate) margins. In addition, I apply the model to clarify the forces underlying the international risk-taking channel of monetary policy, and examine the impact of changes to banks' funding conditions on the riskiness of the banks' portfolios.

The following corollary summarizes the effects of a shock to banks' funding cost on the thresholds and the equilibrium interest rates.

Corollary 3 *(Funding Shock) Holding all else constant,*

1. $\bar{z}^L(z_i^G)$ *is decreasing in* r^G *and increasing in* r^L*;* $\bar{z}^G(z_i^L)$ *is decreasing in* r^L *and increasing in* r^G*.*
2. $\mathcal{R}^G(z_i^G)$ *is increasing in* r^G *and decreasing in* r^L*;* $\mathcal{R}^L(z_i^L)$ *is increasing in* r^L *and decreasing in* r^G*.*

To expand on its intuition and implications, I describe the results from Corollary 3 in the context of a decrease in global banks' funding cost, e.g., a decrease in funding rate due to expansionary monetary policy in the home country of the global banks. The effects of a lower funding cost, r^G, are also illustrated in Figure 1.11, which is based on simulation results with parameter values $r^G = 1.015$, $r^{G'} = 1.005$, and $r^L = 1.040$, where $r^{G'}$ denotes the new gross funding rate for global banks.

Extensive Margin Effects. A decrease in global banks' funding costs lowers the equilibrium interest rates offered by global banks for all firms. Based on Part 4 of Proposition 2, $\bar{z}^L(z_i^G)$ increases, and $\bar{z}^G(z_i^L)$ decreases, which implies that a set of marginal firms switch from local banks to global banks. The changes in the thresholds are illustrated in Panel (a) of Figure 1.11. It is interesting to point out that the marginal firms that switch into global banks are less risky than the infra-marginal firms that continue to borrow from either the local banks or the global banks. Moreover, the funding cost change affects $\underline{z}_i^G$ and $\underline{z}_i^L$, the cut-offs on z_i^G and z_i^L below which global and local banks, respectively, would not make loans. A set of risky firms that initially were not able to get loans from either bank can now get loans from global banks (firms in Region G_2'), while a set of firms that initially were getting loans from local banks are no longer able to borrow from either class of bank (firms in Region G_3').

This result shows that a shock to bank funding cost affects credit allocation at the extensive margin. Specifically, the model predicts that firms near the thresholds, which are firms with relatively balanced global and local risk exposure components, are more likely to switch into contracting with global banks. The result is driven by adverse selection: since the firms with relatively balanced global and local risk exposure are more adversely selected against, they are more likely to switch lenders given any changes in the credit market.

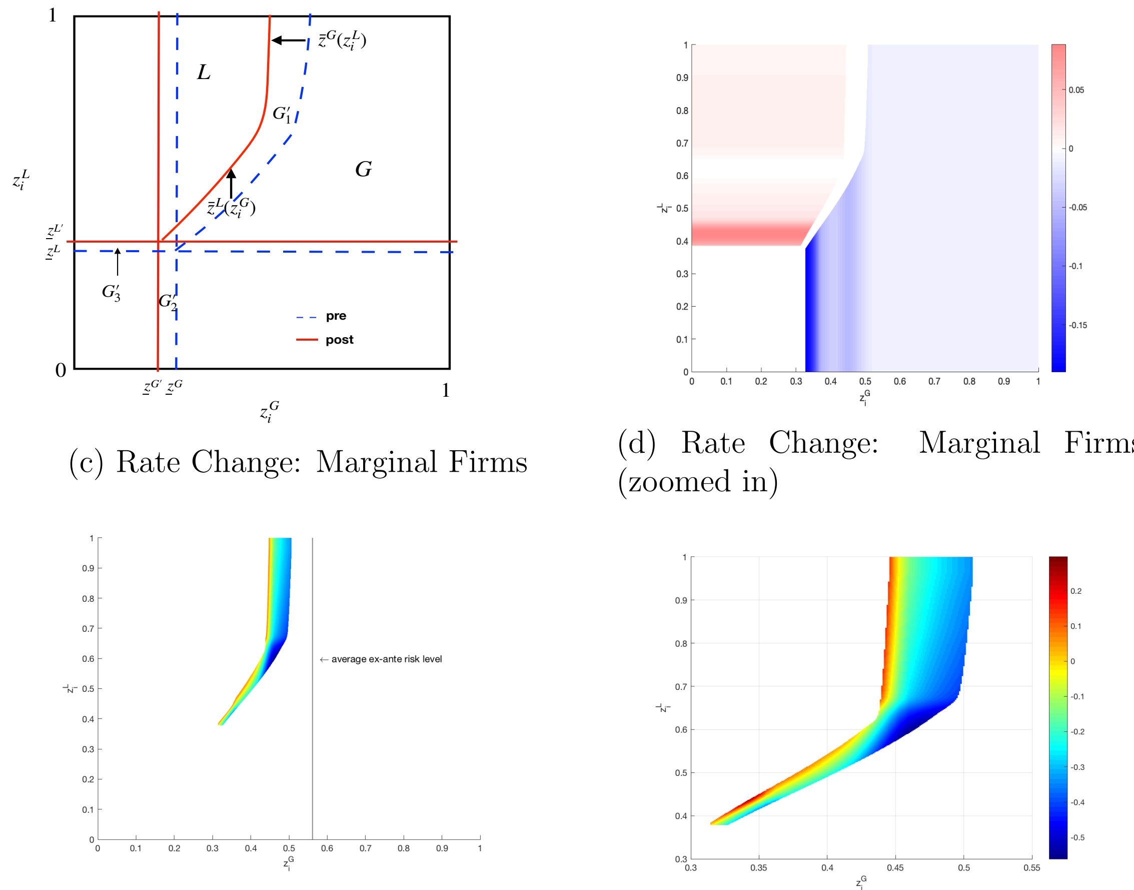
Figure 1.11: **Effects of a Positive r^G Shock (r^G lowers)**

(a) Equilibrium Characterization

(b) Rate Change: Infra-marginal Firms

(c) Rate Change: Marginal Firms

(d) Rate Change: Marginal Firms (zoomed in)



Notes. Simulations based on parameter values $r^G = 1.015$, $r^{G'} = 1.005$, and $r^L = 1.040$. Panel (a) Illustrates the equilibrium characterization before and after a decrease in r^G. Panel (b) shows $\Delta R_i = R_i^{\text{post}} - R_i^{\text{pre}}$ for the infra-marginal firms. Panel (c) shows $\Delta R_i = (R_i^{\text{post}} - R_i^{\text{pre}})/(R_i^{\text{pre}} - 1)$ for the marginal firms. Panel (d) shows a zoomed-in version of part (c) of this figure.

Intensive Margin Effects. Changes in bank funding cost also affect credit allocation at the intensive margin. Given a decline in r^G, for each value of z_i^L, the z_i^G components of the marginal firms that switch away from local banks are higher than those of all the infra-marginal firms that remain with the local banks. Since the local banks are left with a riskier pool of firms, they charge higher interest rates, despite no changes to their funding cost. This points to a spillover effect, one that is solely driven by an exacerbation of the adverse selection problem. Simulation results show that, given a 100 basis point decrease in r^G (specifically a decrease from $r^G = 1.015$ to $r^{G'} = 1.005$), the interest rates that local banks offer to the infra-marginal firms that continue to borrow from them increase by 126 basis points on average, as shown in the red region in Panel (b) of Figure 1.11.

From the global banks' perspective, the z_i^L components of the marginal firms that switch into them are higher than those of all the infra-marginal firms that were getting loans from them in the initial equilibrium, conditional on z_i^G. Since the pool of firms that borrows from global banks is less risky given the funding cost shock, they lower $\mathcal{R}^G(z_i^G)$. In other words, the interest rates of the infra-marginal firms that remain with the global banks are expected to decrease by an amount more than that caused by the decrease in global banks' funding cost, reflecting an amplification effect. The impact of the funding shock is positively amplified for those infra-marginal firms because firm switching alleviates the initial adverse selection problem for the global banks. Simulation results show that a decrease of 100 basis points in r^G translates to a decrease of 180 basis points for an average infra-marginal firm that borrows from global banks, as shown in the blue region in Panel (b) of Figure 1.11.

Panels (c) and (d) of Figure 1.11 illustrate the change in interest rate for the marginal firms that switch banks given the funding cost shock (firms in Region G_1' in panel (a) of the Figure). The effects are heterogenous across the firms: while interest rates decrease for the switching firms that are closer to initial threshold; rates increase for firms closer to new threshold. Nevertheless, those firms would have been worse off if there were frictions to switching that left them with the local banks.

Altogether, this analysis of the effects of a funding cost shock on credit allocation reveals an adverse selection channel of international transmission of funding conditions. It results from the key ingredient in the model: competitive interactions between banks with differing specialization in global versus local information. One of factors that can affect banks' funding cost is monetary policy rate changes. When this happens, the model points to a novel adverse selection channel of international monetary policy transmission through bank lending, one that is distinct from the channels discussed in the existing literature, including currency mismatches on global banks' balance sheets (Bräuning and Ivashina 2017, Ongena et al. 2017, Bräuning and Ivashina 2018) and internal capital markets within global banks (Cetorelli and Goldberg 2012a).

International Monetary Policy Transmission. One channel of international monetary policy transmission that has received much attention in recent years is the risk-taking channel. Papers, including Bruno and Shin (2015a) and Coimbra and Rey (2017), argue that low international monetary policy rates and QE could induce global banks to reach for yield and take on excess risk. In particular, Morais et al. (2018), using loan-level data, show that low monetary policy rates and QE in developed economies led global banks in Mexico to increase credit supply to firms charged higher-than-average ex-ante interest rates (riskier firms). They consider this result as evidence for risk-taking behavior by global banks.

To better understand the forces underlying their result, I implement the empirical exercise in Morais et al. (2018) in my model using numerical simulation and examine whether bank risk-taking is indeed the main driving force. Following their procedure, I first categorize each firm in the model into a high-risk group and a low-risk group based on whether the firm's ex-ante rate is above or below the average interest rate in the credit market in the initial equilibrium. I then examine, given a decline in global banks' funding cost due to expansionary monetary policy in their home country, whether it is the high-risk firms that receive more loans from the global banks.

The specific parameter values I use for the simulation are $r^G = 1.015$, $r^{G'} = 1.005$, and $r^L = 1.040$, where the change in r^G reflects the decline in monetary policy rate in developed economies in the post-global financial crisis period and r^L reflects the average monetary policy rate in Mexico over the period. Panel (c) of Figure 1.11 shows a line pinpointing the firm with the average ex-ante interest rate in that parameter space. As shown, the set of marginal firms that switch into global banks in response to the funding cost change are firms in the high-risk group. Therefore, this model recovers the result that Morais et al. (2018) find in the paper, predicting that an expansionary monetary policy in the home country of the global banks leads to a higher supply of credit to high-risk firms in the local economy. However, in contrast to their explanation, in my model the driving force for the result is substitution between global banking credit and local banking credit.

Overall Riskiness in Bank Portfolios. The prior exercise suggests that credit substitution driven by adverse selection is an important effect of monetary policy transmission. Furthermore, it could potentially confound with bank risk-taking behavior. I investigate this issue further by analyzing how a funding shock affects the overall riskiness of banks' portfolios, and decomposing the overall effect into the changes due to credit substitution and those due to bank risk-taking.

Let the riskiness of the portfolio held by a bank j be expressed in terms of the firms' average output $R_j = \sum_{i=1}^{n}(z_i^G + z_i^L)/n$, where j denotes either a global bank or local bank, and i denotes the firm in the respective bank portfolio. Higher average output R_j implies lower risk.

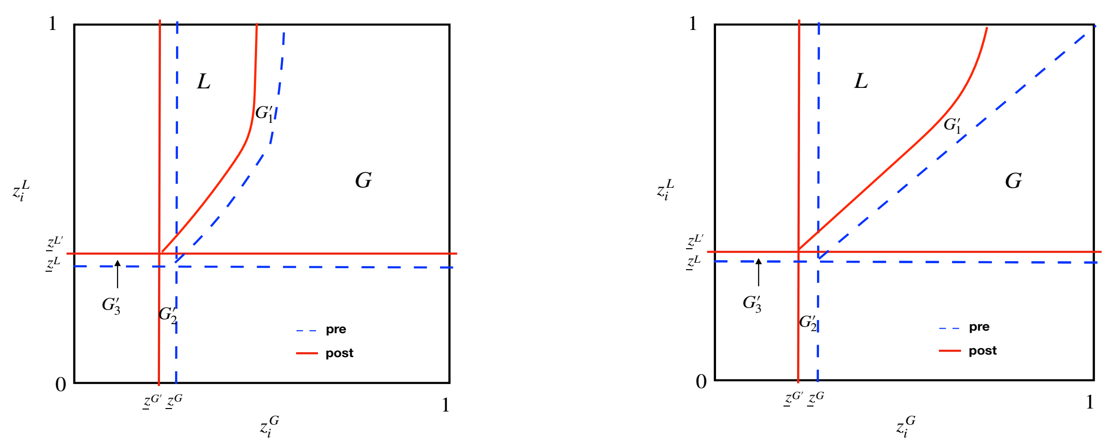
I compute R_j before and after a decline in r^G using numerical simulation, and examine the change in R_j of each bank's portfolios given the change. Specifically, I run the simulation for two sets of parameter values for the initial equilibrium. In scenario 1, $r^G < r^L$ in the initial equilibrium: $r^G = 1.015$, $r^L = 1.050$, and $r^{G'} = 1.005$ ex-post. In scenario 2, $r^G = r^L = 1.015$ in the initial equilibrium, and $r^{G'} = 1.005$. Table 1.2 presents the results. The local bank's portfolios become unambiguously riskier after the funding cost change due to negative spillover effects. On the other hand, the overall riskiness of the global bank's portfolio may increase or decrease depending on the relationship between r^G and r^L in the initial equilibrium.

In scenario 1, the overall riskiness of the global bank's portfolio increases given the decline in funding cost. This is due to the risk profiles of both the marginal firms that switch into the global bank and the newly added firms that were too risky to receive loans before (Region G_1' and Region G_2' in Panel (a) of Figure 1.12, respectively). The average risk of the firms that newly enter the credit market and borrow from the global bank is unambiguously higher than that of the infra-marginal firms that were getting loans from the global bank, driving up the overall riskiness of the global bank's portfolio. This change can be attributed to bank risk-taking. The marginal firms that switch into borrowing from the global bank, despite having higher z_i^L components conditional on z_i^G, have lower z_i^G components on average—and, as a result, higher combined average risk—than those of the infra-marginal firms. This further drives up the overall riskiness of the global bank's portfolio, and the driving force is credit substitution.

In scenario 2, the overall riskiness of the global bank's portfolio lowers. While the riskiness of the firms that newly enter the credit market is still unambiguously higher than that of the infra-marginal firms (Region G_2' in Panel (b) of Figure 1.12), the average riskiness of the switching firms is lower. The average riskiness of both the z_i^L and z_i^G components of the switching firms are lower than the infra-marginal firms that were initially getting loans from global banks. The risk profile of the marginal firms dominate the risk adjustments in global bank's portfolio given the change in r^G. In other words, the effects due to credit substitution dominate the effects due to bank risk-taking in this scenario.

Figure 1.12: **Effects of a Decline in Funding Cost** r^G

(a) Scenario 1, $r^G < r^L$ (b) Scenario 2, $r^G = r^L$



Notes. Panel (a) Illustrates the equilibrium before and after a decline in r^G based on simulations with parameter values $r^G = 1.015$, $r^{G'} = 1.005$, and $r^L = 1.050$. Panel (a) Illustrates the equilibrium before and after a decline in r^G based on simulations with parameter values $r^G = 1.015$, $r^{G'} = 1.005$, and $r^L = 1.015$.

Table 1.2: **Banks' Overall Risk Before and After a Decline in r^G**

		Pre	Post	Switching	New
Scenario 1	G	1.163	1.157	1.155	0.509
	L	0.943	0.917	–	–
Scenario 2	G	1.087	1.155	1.516	0.508
	L	1.155	1.085	–	–

Notes. The table shows the riskiness of the portfolios held by a global bank (G) and a local bank (L) before ("pre") and after ("post") a decline in r^G. The post effect is further decomposed by showing the riskiness of the "switching" firms and "new" firms that select into global banks after the change. Riskiness of bank portfolios is measured as $R_j = \sum_{i=1}^{n}(z_i^G + z_i^L)/n$, where j denotes either a global bank or local bank, i denotes all the firms in the respective bank portfolio. The higher the R_j measure, the lower the risk. In scenario 1, $r^G < r^L$ in the initial equilibrium: $r^G = 1.015$, $r^L = 1.050$, and $r^{G'} = 1.005$ ex-post. In scenario 2, $r^G = r^L = 1.015$ in the initial equilibrium, and $r^{G'} = 1.005$.

1.5 Mapping Theory to Empirics

The model presented in Sections 1.3 and 1.4 delivers three sharp empirical predictions on firm-bank sorting and credit allocation:

Prediction 1: Conditional on funding cost differences between global and local banks, global banks lend more to firms with higher return due to global risk relative to local risk, and local banks lend more to firms with higher return due to local risk relative to global risk.

Prediction 2: A shock to the funding cost of one type of bank induces the segment of firms with relatively balanced global and local risk components (i.e., the marginal firms near the thresholds $\bar{z}^L(z_i^G)$ and $\bar{z}^G(z_i^L)$) to switch to borrowing from the other type of bank.

Prediction 3: Given a decrease in global banks' funding cost, the interest rates of the infra-marginal firms that remain with the local banks are expected to increase (spillover effect). The interest rates of the infra-marginal firms that remain with the global banks are expected to decrease by more than the direct effect due to the decrease in funding cost (amplification effect). The effects on interest rates of the marginal firms that switch banks are ambiguous.

I proceed to test these predictions in the subsequent sections. First, I provide a description of the data used in the empirical analysis.

Data and Summary Statistics. The main data source for the empirical analysis is syndicated corporate loans from Loan Pricing Corporation's Dealscan database. Syndicated loans are extended by a group of banks to a borrower under a single loan contract. Within each group of lenders, the "lead arranger" is the bank that establishes a relationship with the borrowing firm, negotiates terms of the contract, and guarantees a loan amount for a price range. It then turns to "participant" lenders that fund part of the loan.[17] Ivashina and Scharfstein (2010) report that syndicated loan exposures represent about a quarter of total commercial and industrial loan exposures on US banks' balance sheets, and about a third for large US and foreign banks. De Haas and Van Horen (2013) note that syndicated loans are a key source of cross-border funding for firms from both advanced and emerging market countries.

For the purpose of this study, the ideal dataset is one that encompasses the universe of loans to firms that genuinely have access to both global and local banking credit, which are likely to be firms above a certain threshold in size. The global syndicated loans are viewed as a proxy of that universe of loans. Despite potential selection issues, syndicated loans are uniquely appropriate for this study because they capture a significant portion of cross-border lending, which would not be captured by other loan datasets such as credit registry data.

In the Dealscan data, there is detailed information on each loan contract, including terms of the loans at origination (interest rate, whether or not the loan is secured, the maturity of the loan), the type of loan (e.g., line of credit versus term loan), the purpose of the loan, the size of the loan, and the contract activation and ending dates. The dataset also contains information on the name of the borrowers and lenders as well as the country of syndication. Using the names of the borrowers, I hand-match the Dealscan data with international firm-level databases including Orbis, Amedeus, Compustat, and Compustat Global to extract firm balance sheet data.[18] I further implement a series of data-cleaning procedures to correct for basic reporting mistakes, including dropping firm-year observations that have missing information on total assets and operating revenues, dropping firms with negative total assets or employment in any year, and dropping firm-year observations with missing information regarding their industry of activity. Finally, I also exclude firms in financial industries, identified by SIC codes 60 through 64 from the sample.

For the purpose of this empirical analysis, one of the key variables needed is one that identifies whether the lender of each loan is a global bank or a local bank. To this end, I categorize the lead lender(s) of each loan as global or local. The focus is on the lead bank(s) of each loan contract because they are the entities that are responsible for due diligence prior to loan syndication, while the participant banks rely on the information collected by the lead banks (Ivashina and Scharfstein 2010).[19]

The bank categorization is based on the following criteria:

[17] See Sufi (2007) and Ivashina (2009) for more background description of syndicated loans.

[18] The Amadeus and Orbis datasets are mainly used to extract information on European and other non-US firms, including private firms. Compustat is used to extract information on US firms. A well-known problem in the Orbis and Amadeus dataset is that key variables, such as employment and materials, are missing once the data are downloaded. I follow the data collection process described in Kalemli-Ozcan et al. (2015) to maximize the coverage of firms and variables for the sample. Specifically I merge data across historical disks instead of downloading historical data all at once from the WRDS website.

[19] For loans that involve multiple lead banks of which some are global banks and some are local

1. Local banks: a lender is categorized as a local bank if the corresponding loan is not a cross-border loan, i.e., the borrower of the loan operates in the country where the lender resides. This includes local subsidiaries of foreign banks.[20]

2. Global banks:

   - Method 1: a lender is categorized as a global bank if the corresponding loan is a cross-border loan.
   - Method 2: a lender is categorized as a global bank if the corresponding loan is a cross-border loan, or if it is considered a globally systemically important bank (G-SIB).

The resulting sample encompasses 115,166 loans, borrowed by 12,979 firms across 24 countries, in the period 2004-2017. Table 1.3 presents the summary statistics on the loan counts and firm counts for each country in the sample, with the loan counts decomposed into the share given by global banks and that given by local banks, based on Method 1 of the categorization criteria for global banks.[21] The majority of the countries in the sample are developed economies, where most global banking activities take place. For most of the countries, the loans are split relatively evenly between global banking credit and local banking credit.

Table 1.4 presents the summary statistics on a set of firm balance sheet variables. All the variables in the table are in billions of dollars, except for age and employment. Value added, wage bill, total assets, and exporter revenue are deflated with gross output price indices with a base year of 2017. I first calculate the means and standard deviations of each variable across firms in each given year and country without weighting across firms. Entries in the table denote the means and standard deviations averaged across all years and countries. The summary statistics exhibit significant variation in each variable in the sample, which shows that the sample contains firms from a wide distribution of asset size and age. For all variables except exporter revenue, there does not seem to be a significant difference between the firms that borrow from global banks and the ones that borrow from local banks. On the other hand, it seems that firms that borrow from global banks export significantly more than firms that borrow from local banks.

---

banks, I consider a loan is given by global bank if $\geq 50\%$ of the lenders are global banks. These cases make up around 20% of the loans. Based on the model predictions, I conjecture that firms with relatively balanced global and local risk components are more likely to get loans that involve both global and local lead banks. I find empirical evidence that supports this conjecture.

[20] E.g., for firms in Germany, JP Morgan Holding Deutschland is a local bank, while JP Morgan Chase USA is a global bank. Local subsidiaries are considered separate legal entities from their parent bank, incorporated in host countries and supervised by the host regulator.

[21] Table A.1 in A.2 presents summary statistics on the same variables as Table 1.3 but with the banks categorized based on Method 2 of the categorization criteria for global banks.

Table 1.3: **Summary Statistics: Loan and Firm Count by Country (Method 1)**

Country	Loan	GB	LB	Firm	Country	Loan	GB	LB	Firm
Australia	4507	0.70	0.30	701	Japan	21341	0.29	0.71	2865
Austria	387	0.53	0.47	61	Mexico	601	0.70	0.30	137
Belgium	704	0.61	0.39	123	Netherlands	2028	0.28	0.72	406
Canada	6760	0.64	0.36	903	New Zealand	1023	0.70	0.30	127
Czech Republic	197	0.68	0.32	77	Norway	1017	0.66	0.34	253
Denmark	327	0.56	0.44	84	Poland	318	0.54	0.46	87
Finland	587	0.65	0.35	113	Portugal	254	0.65	0.35	64
France	5876	0.43	0.57	996	Spain	4380	0.60	0.40	839
Germany	5987	0.54	0.46	942	Sweden	875	0.62	0.38	190
Greece	309	0.66	0.34	47	Switzerland	790	0.58	0.42	175
Ireland	404	0.63	0.37	107	United Kingdom	6810	0.43	0.57	1528
Italy	2378	0.58	0.42	688	United States	46732	0.40	0.60	1466

Notes. Sample constructed from Dealscan, Amadeus, Orbis, Compustat, Compustat Global, and author's calculation. Sample period covers the year 2004-2017.

Table 1.4: **Summary Statistics: Firm Characteristics by Bank Type**

	Global Bank		**Local Bank**	
	Mean	SD	Mean	SD
Value Added	512.55	1256.45	468.55	895.09
Age	25.29	24.67	25.23	24.98
Employees	1657.34	6073.34	1719.58	5326.32
Wage Bill	209.73	1030.76	163.35	786.41
Working Capital	110.58	1089.56	123.34	1173.32
Fixed Asset	918.03	2465.80	732.76	2987.84
Total Assets	1344.5	4658.56	1134.53	4034.32
Exporter Revenue	587.00	1789.34	113.31	456.68

Notes. Value added is constructed as the difference between operating revenue and materials with negative values dropped. Age of the firm is calculated as the difference between the year of the balance sheet information and the year of firm incorporation plus one. Except for age and employment, all entries in the table are in billions of dollars. Value added, wage bill, total assets, and exporter revenue are deflated with gross output price indices with a base year of 2017. I first calculate the means and standard deviations without weighting across firms for each year in each country. Entries in the table denote the means and standard deviations averaged across all years and countries. Data from Amadeus, Orbis, Compustat, and Compustat Global. Sample period covers the year 2004-2017.

1.6 Empirical Analysis: Firm-Bank Sorting

In this section, I test whether the firm-bank sorting patterns predicted by the model are consistent with the observed patterns in the data (model Prediction 1). To that end, I follow an empirical strategy that tightly maps to the model set-up.

Methodology

In order to test whether global banks lend more to firms with higher return due to global risk (z_i^G) relative to local risk (z_i^L), and vice versa for local banks, I need to construct measures for z_i^G and z_i^L for each firm in the sample. Recall from the model that the production function for each firm is $z_i = z_i^G + z_i^L + u_i$. I take that as a simplified version of a typical Cobb-Douglas production function $Y_i = z_i K_i^{\gamma} L_i^{1-\gamma}$, where there is one unit of K_i and L_i. The parameter z_i, in turn, can be interpreted as a firm revenue productivity measure that captures total exposure to both productivity and demand risk, and z_i^G (z_i^L) can be interpreted as total exposure to global (local) productivity and demand risk.

Estimating z_i. I start by estimating a time-varying revenue productivity measure z_{it} for each firm in each year based on the method of Solow growth accounting.[22] Specifically, I compute the z_{it} based on the following equation:

$$log\, z_{it} = log\,(Y_{it}/L_{it}) - \gamma_t log\,(K_{it}/L_{it}) \tag{1.4}$$

where Y_{it} denotes nominal value added divided by the 2-digit industry-level output price deflator for each country, where value added is constructed as the difference between operating revenue and material costs with negative values dropped, L_{it} denotes the wage bill divided by the same output price deflator, K_{it} denotes fixed assets divided by the aggregate price of investment goods, and the factor share γ_t uses country-specific and industry-specific shares extracted from the National Accounts of each country.

Figure A.1 plots the estimates of the productivity measure, $log\ z_{it}$, averaged across firms and time by country. As expected, average productivity is higher for the relatively more developed economies such as the US and high-income European economies.

Estimating z_i^G and z_i^L. Next I decompose the firm-specific productivity measure, z_i, which captures total exposure to productivity and demand risk, into two components:

[22] Gorodnichenko (2012) shows that this can be used as a robust non-parametric method to estimate productivity. He also points out that a number of existing parametric methods for estimating productivity are misspecified or poorly identified. In particular, inversion/control-function estimators (e.g., Olley and Pakes 1996, Levinsohn and Petrin 2003) can lead to inconsistent estimates because they ignore variation in factor prices. GMM/IV estimators using lags of endogenous variables as instruments (e.g., Blundell and Bond 1998a) can be poorly identified because of economic restrictions on the comovement of inputs and output.

exposure to global risk (z_i^G) and exposure to local risk (z_i^L). Firms' total exposure to global risk can be considered to encompass two components, $z_i^G = \beta_i^G z^G$, where β_i^G denotes firm i's exposure to global risk and z^G denotes global risk. The same applies to firms' total exposure to local risk: $z_i^L = \beta_i^L z^L$, where β_i^L denotes firm i's exposure to local risk and z^L denotes local risk.

I implement a principal component analysis to extract estimates for z^G and z^L, following Stock and Watson (2002). Specifically, I estimate the following equation:

$$z_{ict} = \beta_{ic}^G z_t^G + \beta_{ic}^L z_{ct}^L + u_{ict} \tag{1.5}$$

where z_{ict} is the productivity measure for firm i in country c in year t, z_t^G is the global factor, z_{ct}^L is the local factor in country c, and u_{ict} is a firm-specific component.

The factors can be estimated consistently with a two-step procedure. In the first step, the common global factor is obtained from the principal components of the z_{ict} series across the 24 countries in the sample. The first principal component explains 58% of the total variance, which I take as the global factor, z_t^G. Figure 1.13 plots the global factor.[23] As shown, it declines around 2007-2008, the period of the global financial crisis, and gradually recovers thereafter.

In the second step, I orthogonalize the global component by regressing the productivity measures z_{ict} on the global factor and taking the residuals. I then extract local (country) factors by computing the principal components based on the residualized z_{ict} series for each country. The first principal component from output for each country is taken as the local factor, z_{ct}^L. Finally, I estimate the firm-specific global and local exposure measures using OLS regressions. β_i^G and β_i^L are extracted from the loadings on the global and local factor, respectively.

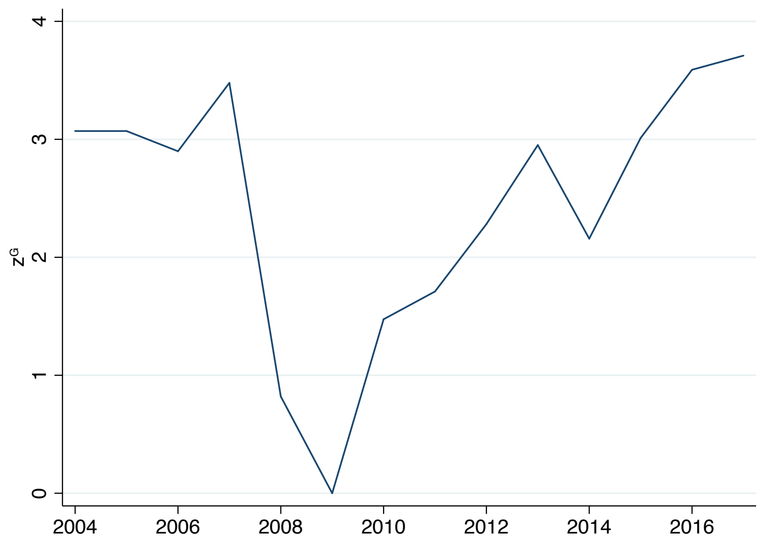
Results

Using the estimated measures for z_i^G and z_i^L, I proceed to test the first model prediction on firm-bank sorting. Similar to the procedure I used to test the traditional theory on firm-bank sorting in Section 1.2 but now using the new measures, I sort firms into quartiles based on the distribution of firm exposure to global versus local risk (z_i^G/z_i^L) in each year by country, and calculate the proportion of loans given by global banks and local banks in each quartile. Figure 1.14 plots the resulting distribution of lending from global and local banks over the entire sample. The plot shows a stark pattern of firm-bank sorting: global banks lend more to firms with higher return given global risk (z_i^G) relative to local risk (z_i^L), and local banks lend more to firms with higher return due to local risk relative to global risk.[24]

[23] To map closely to the model setup where z_i^G and z_i^L only take positive values, the factor values have been adjusted upward by their minimum so that all the values are positive.

[24] Figure A.2 parallels Figure 1.14, with the banks categorized based on Method 2 of the bank categorization criteria for global banks.

Figure 1.13: **Estimates of Global Factor** z^G

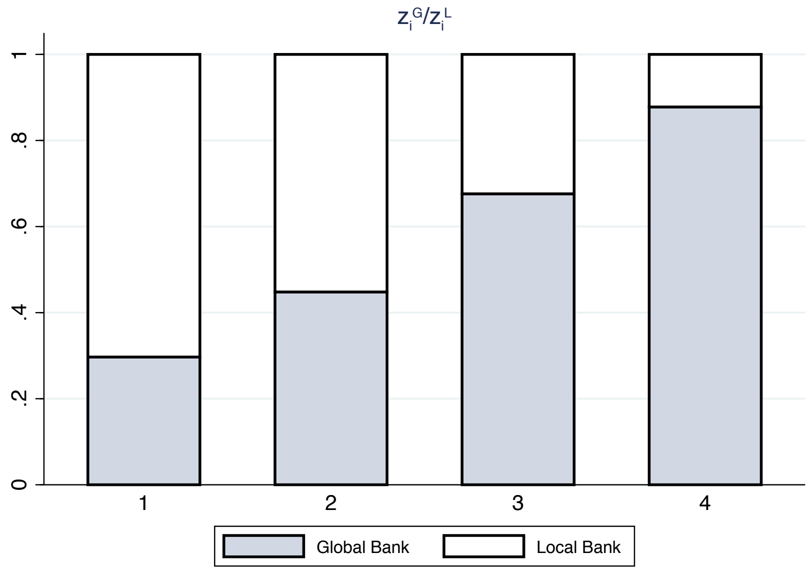


Notes. A plot of the global factor z^G, extracted from the first principal component of the z_{ict} series. The factor values have been adjusted upward by their minimum so that all the values are positive. Source: Dealscan, Amadeus, Orbis, Compustat, Compustat Global, and author's calculation.

As before, I further test whether the differences between global and local banks illustrated in Figure 1.14 and A.2 are statistically significant. For the measure on firm exposure to global versus local risk (z_i^G/z_i^L), I test whether the value-weighted mean of that variable for global banks is different from that for local banks. Table 1.5 presents these means and their differences. The results confirm the graphical analysis: the differences in value-weighted means are statistically *significant* between global and local banks for the measure of firm exposure to global versus local risk (z_i^G/z_i^L), supporting the model prediction on firm-bank sorting.

The results show that the new perspective I raise in this paper, bank specialization in global versus local information, plays an important role in determining firm-bank sorting in financial systems with both global and local banks. But does the traditional theory of bank specialization in hard versus soft information still play a role? I investigate this question by studying how the measures that capture global information and the measures that capture hard information jointly predict the likelihood of getting loans from global banks. I run a set of regressions with the dependent variable being a dummy variable that takes the value 1 if the loan is given by a global bank and 0 otherwise. The independent variables in the regressions are firm exposure to global risk relative to local risk (z_i^G/z_i^L), firm asset size, and/or firm age, each coded by the quartile number to which each observation of the respective variable belongs. The results are presented in Table 1.6. Results in column 1 show that between firms

Figure 1.14: **Firm-Bank Sorting, by z_i^G/z_i^L Quartile (Method 1)**



Notes. The plot shows sorting patterns between firms and global versus local banks, with firms sorted into quartiles by their exposure to global versus local risk (z_i^G/z_i^L), with the banks categorized based on Method 1 of the bank categorization criteria for global banks. Data sample consists of syndicated loans between firms global and local banks and firms across 24 countries from 2004-2017. Source: Dealscan, Amadeus, Orbis, Compustat, Compustat Global, and author's calculation.

in two consecutive quartiles based on the measure of exposure to global risk relative to local risk, the firms in the higher quartile group are 33% more likely to get loans from a global bank. Columns 2 and 3 present results from regressions that include firm asset size and firm age, respectively. The results show that, controlling for firm exposure to global risk relative to local risk, firms that are larger and more established are significantly more likely to get loans from global banks, which is consistent with the predictions from the traditional banking theory. The results in column 4 show that each of the three measures still have predictive power on the likelihood of getting loans from global banks, even when the other two measures are also included as regressors. Overall, these results suggest that the firm-bank sorting patterns predicted by the traditional banking theory can be recovered once bank specialization in global and local information are taken into account.

Finally, I explore the characteristics of the firms that borrow from global banks, and the characteristics of the loans are given by global banks. For the former, I study if exporters are more likely to have a higher value of z_i^G/z_i^L and thereby more likely to get loans from global banks. I run a firm-level panel regression with z_i^G/z_i^L as the dependent variable, and a dummy variable that takes the value 1 if the exporting revenue for the respective firm for a given year is nonzero and 0 otherwise as the

Table 1.5: **Firm-Bank Sorting, by z_i^G/z_i^L Quartile: Statistical Test**

	Method 1 z_{iG}/z_{iL}	Method 2 z_{iG}/z_{iL}
Mean: Global Bank	2.905***	3.382***
	(0.046)	(0.040)
Mean: Local Bank	2.107***	2.507***
	(0.113)	(0.097)
Difference	0.798***	0.875***
	(0.122)	(0.105)
Observations	98,345	98,345

Notes. The dependent variable in each regression (Y) is the measure of firm exposure to global versus local risk, (z_i^G/z_i^L), coded 1-4 based on the quartile number to which each respective firm belong. Note the firms are sorted based on the exposure measure every year by country. Row 1 and row 2 show the means for each variable for global banks and local banks, respectively, by running a value-weighted regression of Y on a constant. For differences in means of the two types of banks, the whole data is used in the regression and a dummy for global banks is added (row 3). Standard errors reported in parentheses are clustered at the bank-level. Results in column 1 and column 2 are based on the banks categorized using Method 1 and Method 2, respectively, of the bank categorization criteria for global banks. Source: Dealscan, Amadeus, Orbis, Compustat, Compustat Global, and author's calculation.

Table 1.6: **Firm-Bank Sorting, Traditional Theory and New Perspective**

	(1) 1(GB)	(2) 1(GB)	(3) 1(GB)	(4) 1(GB)
z_i^G/z_i^L	0.329***	0.221***	0.261***	0.198**
	(0.086)	(0.074)	(0.080)	(0.081)
Size		0.268***		0.236***
		(0.081)		(0.073)
Age			0.157**	0.138*
			(0.075)	(0.078)
Industry FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Observations	98,345	98,345	98,345	98,345

Notes. Results from regressions with the dependent variables being a dummy variable that takes the value 1 if the loan is given by a global bank and 0 otherwise. The independent variables are firm exposure to global risk relative to local risk (z_i^G/z_i^L), firm asset size, and/or firm age, each coded by the quartile number to which each observation of the respective variable belongs. Each regression controls for industry and country fixed effects. Standard errors reported in parentheses are clustered at the firm level. Source: Dealscan, Amadeus, Orbis, Compustat, Compustat Global, and author's calculation.

main regressor, controlling for time and country fixed effects. The results, reported in column 1 of Table 1.7, show that exporting firms tend to have significantly higher z_i^G/z_i^L values, or higher exposure to global risk relative to local risk. Combined with the results from the sorting exercises, this empirical evidence suggests that exporters are more likely to get loans from global banks.

In light of these evidence, I further investigate into the loan-level data to see whether loans of specific purposes such as trade finance are more likely to be funded by global banks. I run a loan-level regressions with the main regressors being dummies on specific loan purposes, including project finance, working capital, trade finance, and others[25]. The dependent variable of the regression is a dummy variable that takes the value 1 if the loan is given by a global bank and 0 otherwise. The results (column 2 of Table 1.7) show that it is not the case that global banks mainly finance loans for the purpose of trade finance. A significant portion of the loans they finance are for general project finance and working capital.

Table 1.7: **Determinants of z_i^G/z_i^L and Global Banking Credit**

	(1) z_i^G/z_i^L	(2) 1(GB)
Exporter	0.565***	
	(0.103)	
Project purpose		
Project finance		0.013***
		(0.001)
Working capital		0.020***
		(0.001)
Trade finance		0.004**
		(0.002)
Year FE	Yes	Yes
Industry FE	Yes	Yes
Country FE	Yes	Yes
Observations	129,309	98,345

Notes. Column 1 reports results from a firm-level panel regression with z_i^G/z_i^L as the dependent variable, and a dummy variable that takes the value 1 if the exporting revenue for the respective firm for a given year is nonzero and 0 otherwise as the main regressor. Column 2 reports results from a loan-level regression with a dummy variable that takes the value 1 if the loan is given by a global bank and 0 otherwise as the dependent variable, and dummy variables on loan purpose as the main regressors. Time, industry and country fixed effects are included in both regressions. Standard errors reported in parentheses are clustered at the firm level. Source: Dealscan, Amadeus, Orbis, Computstat, Compustat Global, and author's calculation.

[25] Others include IPO related finance, real estate, stock buyback, etc. They are grouped together in one variable.

1.7 Empirical Analysis: Adverse Selection Channel of Monetary Policy Transmission

In this section, I study how shocks to bank funding cost, specifically monetary policy shocks, affect credit allocation at the extensive and intensive margins, testing model Predictions 2 and 3. I take the Euro area as the empirical laboratory of this study, and analyze how US and Euro area monetary policy, through US and Euro area banks, respectively, affect credit allocation across firms in the Euro area. From the perspective of Euro area firms, US banks are global banks, and Euro area banks are local banks. Given this context, I raise two conjectures based on the model predictions and the results on firm-bank sorting from the last section:

i) Conditional on Euro area monetary policy, an expansionary US monetary policy induces firms in the Euro area with relatively balanced global and local risk components—firms in the second tercile of the z_i^G/z_i^L distribution—to switch their borrowing from Euro area banks to US banks.

ii) Conditional on Euro area monetary policy and given expansionary US monetary policy, the interest rates of the infra-marginal firms that continue to borrow from Euro area banks—firms in the first tercile of the z_i^G/z_i^L distribution (firms with relatively low z_i^G relative to z_i^L)—are expected to increase (spillover effect). The interest rates of the infra-marginal firms that continue to borrow from US banks—firms in the third tercile of the z_i^G/z_i^L distribution (firms with relatively high z_i^G relative to z_i^L)—are expected to decrease by more than the direct effect due to expansionary US monetary policy (amplification effect). The effects on interest rates of the marginal firms that switch banks—firms in the second tercile of the z_i^G/z_i^L distribution—are ambiguous.

To test the conjectures, I perform regressions of the following form, using data on loans borrowed by Euro area firms in the loan-level data and the firm-specific z_i^G/z_i^L measure:

$$\Delta Y_{it} = \sum_{q=1}^{3} \beta^q(\Delta USR_t \text{ x } T^q_{it-1}) + \sum_{q=1}^{3} \delta^q(\Delta EUR_t \text{ x } T^q_{it-1}) + \sum_{q=2}^{3} \gamma^q T^q_{it-1} + \nu_i + \sigma_t + \epsilon_{it} \quad (1.6)$$

where i indexes firm, t indexes the date on which a specific loan is issued, $\Delta(.)$ denotes the difference in the referred variable between the date on which the current loan is issued and the date on which the last loan was issued, Y denotes the applicable dependent variable which I explain below, USR denotes US monetary policy shocks, EUR denotes Euro area monetary policy shocks, q indexes each of the three terciles of the z_i^G/z_i^L distribution, T^q_{it-1} are dummy variables that take the value 1 when firm i's z_i^G/z_i^L measure at the time of the last loan issuance belongs to tercile q and 0 otherwise, ν_i are firm dummies, and σ_t are year dummies. The standard errors are clustered by time, to take into consideration potential correlations across firms in borrowing behavior or borrowing term changes since the monetary policy shocks are aggregate.

For measures of US and Euro area monetary policy shocks, I use intraday data on the Federal Funds 30-day futures contracts and the three-month Euribor futures contracts, respectively, from Gorodnichenko and Weber (2016) and CQG Data Factory.[26] The Federal Funds futures data is based on trading on the Chicago Board of Trade (CBOT) Globex electronic trading platform. It reflects the market expectation of the average effective Federal Funds rate during that month. The Euribor futures rates is based on trading on ICE Futures Europe and reflects the market expectation of the Euribor rate for three-month Euro deposits.[27] Therefore, both series provide a market-based measure of the anticipated path of the monetary policy rates for the respective region.

In order to identify exogenous shocks to US and Euro area monetary policy, the monetary policy shocks are calculated as changes in the futures rates within a time window around the Federal Open Market Committee (FOMC) or European Central Bank (ECB) monetary policy announcements.[28] The identifying assumption is that changes in the interest rate futures within the specified windows around the announcements only reflect market responses to the monetary policy news, not changes in other domestic or foreign economic conditions. For measures of US monetary policy shocks, I consider a window of 60 minutes around the announcements that starts 15 minutes (Δ^-) prior to the event, following Gorodnichenko and Weber (2016) and Nakamura and Steinsson (2018).

As for ECB monetary policy, its key target rate decision since 2001 has been announced at 13:45 CET through a press release, followed by a press conference at 14:30 pm CET. At the press conference, the ECB President and Vice-President discuss the future path of monetary policy and announce any additional non-conventional measures.[29] To give a sense of how the ECB policy rate announcement and the press conference affect the market expectation of the Euribor rate, I illustrate the three-month Euribor futures rate in high frequency on two specific announcement dates in

[26] The US monetary policy shock measure based on intraday data on the Federal Funds futures contracts has been used in a number of papers, including Kuttner (2001), Cochrane and Piazzesi (2002), Rigobon and Sack (2004), Gertler and Karadi (2015), Gorodnichenko and Weber (2016), Nakamura and Steinsson (2018), and Wong (2018). The Euro area monetary policy measure based on the three-month Euribor futures has been used in papers including Bernoth and Hagen (2004), Rosa and Verga (2008), and Ranaldo and Rossi (2010). They show that the three-month Euribor futures rate is an unbiased predictor of Euro area policy rate changes.

[27] To be more specific, the three-month Euribor future is a commitment to engage in a three-month loan or deposit of a face value of 1,000,000 Euros. Futures prices are quoted on a daily basis. There are four delivery dates during a year, namely the third Wednesday of March, June, September and December.

[28] I obtain the dates of the FOMC meetings from the Federal Reserve Board website at http://www.federalreserve.gov/monetarypolicy/fomccalendars.htm, and those of the ECB meetings from the ECB website at https://www.ecb.europa.eu/press/govcdec/mopo. I also verify the exact times of the monetary policy announcements using the first news article about them on Bloomberg.

[29] See Rosa and Verga (2008) for a description of the institutional features unique to ECB monetary policy announcements.

Figure A.3. The upper panel plots the Euribor futures rate from 08:00 to 18:00 CET for April 6, 2006. At 13:45 CET, the ECB announced through a press release that it is keeping the target rate unchanged. Since this decision was expected by the market, the futures rate did not exhibit significant change around the press release time. But it decreased sharply during the press conference window. This is because, contrary to market expectation of an interest rate hike later in the year, Jean-Claude Trichet told the press that "the current suggestions regarding the high probability of an increase of rates in our next meeting do not correspond to the present sentiment of the Governing Council." The decline in Euribor futures rate during the press conference time window thus reflect market's revision of its expectations. The bottom panel of Figure A.3 plots the Euribor futures rate for November 3, 2011, when the ECB unexpectedly cut interest rates by 25bps for the first time in two years. The sharp decline in the Euribor futures rate around the time of the press release reflect the change in market expectation. Given the unique institution features of ECB monetary policy announcements, I apply a window of 120 minutes that starts 10 minutes (Δ^-) prior to the press release and ends 10 minutes (Δ^+) after the press conference to construct measures of ECB monetary policy shocks.

Furthermore, I consider two measures of monetary policy shocks for each region: a current period shock based on current month futures ($mp1$), and a long-term path shock based on three-month-ahead futures ($mp4$). The long-term path shock is aimed at capturing any persistent effects of current period shocks on long-term investment, which can occur when the current period shocks change expectations about the future path of monetary policy rates.

The shock measures take the general form:

$$mp_t = (f^x_{t+\Delta^+} - f^x_{t-\Delta^-}) \tag{1.7}$$

where t is the time when the FOMC or ECB issues an announcement, $f_{t+\Delta^+}$ is the Federal Funds futures or the Euribor futures Δ^+ minutes after t, $f_{t-\Delta^-}$ is the Federal Funds futures or the Euribor futures Δ^- minutes before t, and x denotes either 1 for current month futures or 4 for three-month-ahead futures. For the US current monetary policy shock measure ($mp1$), Equation (1.7) is adjusted by the term $\frac{D}{D-t}$, where D is the number of days in the month. This is because the Federal Funds futures settle on the average effective overnight Federal Funds rate.

I aggregate up the identified shocks to obtain monthly measures of monetary policy shocks, following Cochrane and Piazzesi (2002). I use the monetary policy measures from the month prior to the loan dates (t) when estimating Equation (1.6), to ensure time consistency.

Extensive Margin To analyze how monetary policy shocks affect credit allocation across firms in the Euro area at the extensive margin, I estimate Equation (1.6) with the the dependent variable being the change in a dummy variable that takes the value 1 if the loan is given by a US bank and 0 if the loan is given by a Euro area bank between two consecutive loans for each given firm i (denoted as ΔUSB_{it}). The main coefficients of interest are β^q and δ^q. I conjecture β^2 to be negative, and δ^2 to be positive, since, based on the model prediction, contractionary US monetary policy would induce firms in the second tercile of the z_i^G/z_i^L distribution to switch away from US banks, and contractionary ECB monetary policy would induce firms in the second tercile of the z_i^G/z_i^L distribution to switch into US banks. All the specifications include firm fixed effects to account for potential demand-driven explanations for changes in the trends of firms' borrowing behavior, as well as time fixed effects to control for common shocks.

Table 1.8 reports the regression results. Columns 1 and 3 show the average effects of the US and Euro area monetary policy shocks, based on measures of $mp1$ and $mp4$, respectively, on the firms' switching behavior. Results in Column 1 show that, on average, a 25-basis-point shock to the current US monetary policy rates decreases the probability of firm switching from a Euro area bank into a US bank by 3.4 percentage points, while a 25-basis-point shock to the Euro area monetary policy rates increases the probability of a firm switching from a Euro area bank into a US bank by 4.1 percentage points. The effects are larger and more significant when considering shocks to the path of monetary policy rates. Results in Column 3 show that, on average, a 25-basis-point shock to the path of US monetary policy rates decreases the probability of firm switching into a US bank by 5.2 percentage points, while such shock to the path of Euro area monetary policy rates increases the probability of a firm switching into a US bank by 5.3 percentage points.The coefficients are statistically significant at the 5% level. The findings point to evidence of firm switching in the Euro area in response to monetary policy shocks on average. In particular, firms respond slightly more to domestic monetary policy shocks.

Turning to the coefficients of interest, columns 2 and 4 in Table 1.8 show the estimations of how these effects vary for firms in different terciles of the z_i^G/z_i^L distribution (Equation (1.6)). Across both specifications, the effects of US and Euro area monetary policy shocks on the probability of firm switching are around two times larger in the second tercile of the z_i^G/z_i^L distribution than the other terciles, and highly significant. The point estimates of β^2 imply that a 25-basis-point shock to the current and long-term US monetary policy rate decreases the probability of firm switching into a US bank by 6.0 and 7.6 percentage points, respectively, for firms in the second tercile of the z_i^G/z_i^L distribution. For those firms, the point estimates of δ^2 imply that a 25-basis-point shock to the Euro area monetary policy increases the probability of firm switching into a US bank by 6.6-8.5 percentage points. The effects are again larger when considering shocks to the path of monetary policy rates, suggesting that firm investments respond more to changing expectations about the future path of monetary policy rates. The results for the other two terciles are mostly statistically insignificant.

Overall, the results suggest that most of the firm switching effects are concentrated in the second tercile of the z_i^G/z_i^L distribution, where firms have relatively balanced exposure to global risk relative to local risk. This evidence supports the model prediction on the effects of bank funding shocks on credit allocation across firms at the extensive margin.

Intensive Margin Next, I turn to analyzing how monetary policy shocks affect credit allocation across firms in the Euro area at the intensive (interest rate) margin. I implement Equation (1.6) with the dependent variable being the change in the interest rate spread between two consecutive loans for each given firm i (denoted as ΔR_{it}).[30] The spread describes the amount the borrower pays in basis points over the LIBOR. The main coefficients of interest are again β^q and δ^q. The model predicts that, conditional on Euro area (US) monetary policy and given contractionary US (Euro area) monetary policy, the interest rates of the infra-marginal firms that continue to borrow from Euro area (US) banks decrease, reflecting a (positive) spillover effect. Thus, β^1 (which summarizes the group of firms that are more likely to be borrowing from Euro area banks) and δ^3 (which summarizes the group of firms that are more likely to be borrowing from US banks) are conjectured to be negative. The model also predicts that, under the above scenario, the interest rate spreads of the infra-marginal firms that continue to borrow from US (Euro area) banks increase, reflecting a (negative) amplification effect. Thus, β^3 and δ^1 are conjectured to be positive.

Since these predictions are based on the assumption that there is stronger pass-through from US monetary policy to the interest rates offered by US banks, and similarly Euro area monetary policy to Euro area banks, I first perform a series of regressions to validate these assumptions. Columns 1, 2, 4 and 5 in Table 1.9 report the results from regressions of changes in firm interest rate spreads (ΔR_{it}) on changes in US and Euro area monetary policy shocks (ΔUSR and ΔEUR, respectively), a dummy variable that takes the value 1 for US or Euro area banks and 0 otherwise ($1(USB)$ or $1(EUB)$), and interactions of these two variables: either an interaction between the US monetary policy shock and the US bank dummy variable ($USR * 1(USB)$), or one between the Euro area monetary policy shock and the Euro area bank dummy variable ($\Delta EUR * 1(EUB)$). The results confirm the assumption. Columns 1 and 4 show that a 25-basis-point shock to the current and long-term US monetary policy rate disproportionately increases the interest rate spread charged by US banks, by around 25 and 33 basis points, respectively, on average relative to other banks. Results in columns 2 and 5 show that a 25-basis-point shock to the current and long-term Euro area monetary policy rate disproportionately increases the interest rate spread charged by Euro area banks, by 34 and 37 basis points, respectively, on average relative to other banks.

Turning to the coefficients of interest, columns 3 and 6 in Table 1.9 report the results of how the effects of monetary policy shocks on interest rate spreads vary for firms in different terciles of the z_i^G/z_i^L distribution (Equation (1.6)). As predicted, the coefficients β^1 and δ^3 are negative across all specifications. Specifically, a 25-basis-

[30] To make the interest rate spreads as comparable as possible, the type of loan facilities (e.g., revolving line, bank term loan, and institutional term loan) between two consecutive loans are matched.

point shock to the current US monetary policy rate decreases the interest rate spread for the infra-marginal firms that continue to borrow from Euro area banks by 22 basis points, while such a shock to the Euro area monetary policy rate decreases the interest rate spread for the infra-marginal firms that continue to borrow from US banks by 25 basis points. The effects are larger and more significant when considering shocks to the path of monetary policy rates (column 6). A 25-basis-point shock to the long-term US (Euro area) monetary policy rate decreases the interest rate spread for the infra-marginal firms that continue to borrow from Euro area (US) banks by 27 (32) basis points. These results point to a (positive) spillover effect.

Furthermore, the coefficients β^3 and δ^1 are positive across all specifications, as predicted, and highly statistically significant. Specifically, a 25-basis-point shock to the current US monetary policy rate increases the interest rate spread for the infra-marginal firms that continue to borrow from US banks by 25 basis points. The effect increases to 32 basis points given a 25-basis-point shock to the path of US monetary policy rate. Similarly, a 25-basis-point shock to the current and long-term Euro area monetary policy rate increases the interest rate spread for the infra-marginal firms that continue to borrow from Euro area banks by 34 and 40 basis points, respectively. These results point to a (negative) amplification effect. Furthermore, the effects on interest rates of the firms in the second tercile of the z_i^G/z_i^L distribution, which, based on the results from Table 1.8, is mostly comprised of marginal firms that may switch banks, are ambiguous, as predicted.

Overall, the results in Table 1.9 support the model prediction on the effects of bank funding shocks on credit allocation across firms at the intensive margin. Combined with the results on the extensive margin effects, they point to evidence of a novel adverse selection channel of monetary policy transmission.

Table 1.8: **Monetary Policy Shocks and Credit Allocation: Extensive Margin**

	(1)	(2)	(3)	(4)
	$mp1$	$mp1$	$mp4$	$mp4$
ΔUSR	-0.134*		-0.209**	
	(0.071)		(0.083)	
ΔEUR	0.164**		0.211**	
	(0.074)		(0.089)	
$\Delta USR * T^1$		-0.049		-0.054
		(0.119)		(0.128)
$\Delta USR * T^2$		-0.241**		-0.302**
		(0.120)		(0.131)
$\Delta USR * T^3$		-0.117		-0.163
		(0.118)		(0.127)
$\Delta EUR * T^1$		0.057		0.062
		(0.118)		(0.137)
$\Delta EUR * T^2$		0.264**		0.339***
		(0.118)		(0.135)
$\Delta EUR * T^3$		0.173		0.220*
		(0.116)		(0.127)
Firm FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	11,454	11,454	11,454	11,454
R-squared	0.067	0.068	0.067	0.068

Notes. Regressions with the dependent variable being the change in a dummy variable that takes the value 1 if the loan is given by a US bank and 0 if the loan is given by a Euro area bank between two consecutive loans for each given firm i (denoted as ΔUSB). USR denotes US monetary policy shocks, and EUR denotes Euro area monetary policy shocks. T^q is a dummy variable that takes the value 1 when the firm's z_i^G/z_i^L measure at the time of the last loan issuance belongs to tercile q and 0 otherwise. For the specifications in columns 1 and 2, the monetary policy measures used are current period shocks constructed from current month futures ($mp1$). For the specifications in columns 3 and 4, the monetary policy measures used are long-term path shocks constructed from three-month-ahead futures ($mp4$). Year and firm fixed effects are included in all specifications. Standard errors reported in parentheses are clustered by time. Source: Dealscan, Amadeus, Orbis, Computstat, Compustat Global, and author's calculation. Significance at the 1 percent, 5 percent, and 10 percent levels is indicated by ***, **, and *, respectively.

Table 1.9: **Monetary Policy Shocks and Credit Allocation: Intensive Margin**

	(1) $mp1$	(2) $mp1$	(3) $mp1$	(4) $mp4$	(5) $mp4$	(6) $mp4$
$\Delta USR * 1(USB)$	98.543**			132.458***		
	(43.765)			(47.986)		
$\Delta EUR * 1(EUB)$		136.633***			147.375***	
		(42.543)			(49.864)	
$\Delta USR * T^1$			-89.354*			-108.564*
			(48.542)			(54.875)
$\Delta USR * T^2$			62.796			78.342
			(52.769)			(60.875)
$\Delta USR * T^3$			98.427**			126.653**
			(46.293)			(58.975)
$\Delta EUR * T^1$			136.864**			158.539***
			(56.249)			(57.986)
$\Delta EUR * T^2$			76.563			83.457
			(52.087)			(59.357)
$\Delta EUR * T^3$			-101.876*			-127.978**
			(54.681)			(54.975)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,367	3,367	3,367	3,367	3,367	3,367
R-squared	0.051	0.051	0.052	0.051	0.051	0.052

Notes. Regressions with the dependent variable being the change in the interest rate spread between two consecutive loans for each given firm i (denoted as ΔR). USR denotes US monetary policy shocks, and EUR denotes Euro area monetary policy shocks. $1(USB)$ and $1(EUB)$ are dummy variables that takes the value 1 for US and Euro area banks, respectively, and 0 otherwise. T^q is a dummy variable that takes the value 1 when the firm's z_i^G/z_i^L measure at the time of the last loan issuance belongs to tercile q and 0 otherwise. For the specifications in columns 1-3, the monetary policy measures used are current period shocks constructed from current month futures ($mp1$). For the specifications in columns 4-6, the monetary policy measures used are long-term path shocks constructed from three-month-ahead futures ($mp4$). The specifications in column 1 and 4 include USR, $1(USB)$, and EUR as regressors. The specifications in column 2 and 5 include USR, $1(EUB)$, and EUR as regressors. Year and firm fixed effects are included in all specifications. Standard errors reported in parentheses are clustered by time. Source: Dealscan, Amadeus, Orbis, Computstat, Compustat Global, and author's calculation. Significance at the 1 percent, 5 percent, and 10 percent levels is indicated by ***, **, and *, respectively.

1.8 Conclusion

The rise of global banking has transformed financial systems and corporate financing across the world over the past two decades. This paper provides a new theory on the mechanism driving credit allocation in globalized financial systems, and tests it using cross-country loan-level data. I show that bank specialization in global versus local information—information on global versus local risk factors—plays a key role in determining firm-bank sorting and credit allocation in financial systems with both global and local banks.

I first point out that that the traditional theory of bank specialization in hard or soft information is insufficient to explain observed sorting patterns between firms and global versus local banks, revealing a puzzle in the mechanism driving global banking credit. Given the puzzle, I develop a model of banking in which there are global and local banks, and firms that have return dependent on exposure to global and local risk. Each bank faces a problem of asymmetric information: global banks have the technology to extract information on global risk factors but not local risk factors, and vice versa for local banks. The model shows that this double information asymmetry creates a segmented credit market affected by double adverse selection: banks are adversely selected against by firm selection, as firms select into borrowing from the bank which observes the more favorable component of their risk exposure.

I further apply the model to analyze the macroeconomic implications of the adverse selection problem, studying the impact on credit allocation of funding shocks to banks. The model demonstrates that, given a monetary policy shock, adverse selection affects credit allocation at both the extensive and intensive margins. It induces firms with relatively balanced global and local risk components to switch banks, and generates spillover and amplification effects through adverse interest rates. I test the model using a cross-country firm-bank loan-level dataset matched with firm balance sheet data. I find firm-bank sorting patterns, and evidence of firm switching behavior and interest rate changes given US and Euro area monetary policy shocks, that support the model predictions. The results point to a novel adverse selection channel of international monetary policy transmission.

Overall, the evidence substantiates that bank specialization in global versus local information is a key mechanism driving credit allocation in globalized banking systems. This mechanism has potentially important policy implications. Relative to the traditional view that firms and banks sort based on hard versus soft information, this new mechanism suggests that global banks' balance sheet may be more loaded on global risk than previously thought, since firms with returns more dependent on global risk are more likely to select into borrowing from them. This, in turn, calls for considerations from policy-makers for bank regulations on exposure limits and macroprudential policies.

Chapter 2

Scarred Consumption

> *The crisis has left deep scars, which will affect both supply and demand for many years to come.* — Blanchard (2012)

2.1 Introduction

More than a decade after the Great Recession, consumers have been slow to return to prior consumption levels (Petev et al. (2011), De Nardi et al. (2012)). As the quote above suggests, the crisis appears to have "scarred" consumers. Consumption has remained low not only in absolute levels, but also relative to the growth of income, net worth, and employment—a pattern that challenges standard life-cycle consumption explanations, such as time-varying financial constraints. For the same reason, low employment due to the loss of worker skills or low private investment, as put forward in the literature on "secular stagnation" and "hysteresis," cannot account for the empirical pattern either.[1]

What, then, explains such long-term effects of a macroeconomic crisis on consumption? The hypothesis we put forward starts from the observation in Pistaferri (2016) that long-lasting crisis effects appear to be explained by consumer confidence remaining low for longer periods than standard models imply. We relate this observation to the notion of *experience effects*, and show that consumers' past lifetime experiences of economic conditions have a long-lasting effect on beliefs and consumption expenditures, which is not explained by income, wealth, liquidity, and other life-cycle determinants.

This chapter is based on joint work with Ulrike Malmendier. Permission to reprint this material as a chapter of the present dissertation has been obtained.

[1] The literature on secular stagnation conjectured protracted times of low growth after the Great Depression (Hansen (1939)). Researchers have applied the concept to explain scarring effects of the Great Recession (Delong and Summers (2012), Summers (2014a) Summers (2014a), Summers (2014b)). Blanchard and Summers (1986) introduce the term "hysteresis effects" to characterize the high and rising unemployment in Europe. Cf. Cerra and Saxena (2008), Reinhart and Rogoff (2009), Ball (2014), Haltmaier (2012), and Reifschneider et al. (2015).

Prior literature on experience effects has shown that personally experienced economic outcomes, such as realizations of inflation, stock returns, or interest rates, receive additional weight in individuals' expectations about future realizations of the same variables, even among highly educated and well informed individuals (e. g., central bankers), and significantly influence their behavior.[2] Here, we ask whether a similar mechanism is at work when individuals experience high unemployment rates. We apply the linearly declining weights estimated in prior work to both national and local unemployment rates that individuals have experienced over their lives so far, and to their personal unemployment experiences, to predict consumption and beliefs. We show that past experiences predict both consumer pessimism and consumption scarring as well as several other empirical regularities, including generational differences in consumption patterns, after controlling for wealth, income and other standard determinants.

We start by presenting four baseline findings on the relation between past experiences and consumption, beliefs, future income, and wealth build-up. We first documenting the long-lasting effect of past experiences on consumption. Using the *Panel Study of Income Dynamics* (PSID) from 1999-2013, we find that past macro and personal unemployment experiences have significant predictive power for consumption, controlling for income, wealth, age fixed effects, a broad range of other demographic controls (including current unemployment), as well as state, year, and even household fixed effects. Without household dummies, the identification comes both from cross-household differences in consumption and unemployment histories, and from how these differences vary over time. With household dummies, the estimation solely relies on within-household variation in consumption in response to lifetime experiences.[3] In both cases, the effects are sizable. A one standard deviation increase in the macro-level measure is associated with a 3.3% ($279) decline in annual food consumption, and a 1.6% ($713) decline in total consumption. A one standard deviation increase in personal unemployment experiences is associated with 3.7% ($314) and 2.1% ($937) decreases in annual spending on food and total consumption, respectively. The results are robust to variations in accounting for the spouse's experience, to excluding last year's experience, or using different weights, from equal to steeper-than-linearly declining.[4]

[2] Theoretical papers on the macro effects of learning-from-experience in OLG models include Ehling et al. (2018), Malmendier et al. (2018), Collin-Dufresne et al. (2016), and Schraeder (2015). The empirical literature starts from Kaustia and Knüpfer (2008), Malmendier and Nagel (2011), and the analysis of FOMC members is in Malmendier et al. (2018).

[3] We have also estimated a model with cohort fixed effects. In that case, the identification controls for cohort-specific differences in consumption. The results are very similar to estimations without cohort fixed effects. Note that, differently from most of the prior literature on experience effects (Malmendier and Nagel (2011) Malmendier and Nagel (2011), Malmendier and Nagel (2015)), the experience measure is not absorbed by cohort fixed effects as the consumption data sets contains substantial within-cohort variation in experiences. The unemployment experience measure of a given cohort varies over time depending on where the cohort members have resided over their prior lifetimes.

[4] We also included lagged consumption in the estimation model to capture habit formation but do not find a significant effect, while the experience proxy remains significant.

Second, we document that consumers' past experiences significantly affect beliefs. Using the *Michigan Survey of Consumers* (MSC) from 1953 to 2012, we show that people who have experienced higher unemployment rates over their lifetimes so far have more pessimistic beliefs about their financial situation in the future, and are more likely to believe that it is not a good time to purchase major household items in general. Importantly, these estimations control for income, age, time effects, and a host of demographic and market controls.

Third, we relate the same measure of lifetime unemployment experiences to actual future income, up to three PSID waves (i. e., six years) in the future. Again, we control for current income, wealth, demographics, as well as age, state, year and even household fixed effects. We fail to identify any robust relation. In other words, while there is a strong reaction to prior lifetime experiences in terms of beliefs and actual consumption choices, actual future income does not appear to explain these adjustments.

Our fourth baseline result captures the wealth implications of consumption scarring. If consumers become more frugal in their spending after negative past experiences, even though they do not earn a reduced income, we would expect their savings and ultimately their wealth to go up. Our fourth finding confirms this prediction in the data. Using a 6- to 14-year horizon, we find past lifetime experiences predict liquid and illiquid wealth build up, in particular for past personal unemployment experiences. Unobserved wealth effects, the main alternative hypothesis, do not predict wealth build up, or even predict the opposite.

These four baseline results—strong experience effects on consumption expenditures and on consumer optimism, but lack of an effect on actual future income, plus positive wealth build-up—are consistent with our hypothesis: Consumers over-weigh experiences that have occurred during their lifetimes so far when forming beliefs about future realizations and making consumption choices, as predicted by models of experience-based learning (EBL). Considered jointly, and given the controls included in the econometric models, the results so far already distinguish EBL from several alternative explanations: The inclusion of age controls rules out certain life-cycle effects, such as increasing precautionary motives and risk aversion with age (cf. Caballero (1990), Carroll (1994)), or declining income and liquidity constraints during retirement (cf. Deaton (1991), Gourinchas and Parker (2002)). The controls for labor market status and demographics take into account intertemporal allocation of expenditure as in Blundell et al. (1994) or Attanasio and Browning (1995). The inclusion of time fixed effects controls for common shocks and available information such as the current and past national unemployment rates. The PSID also has the advantage of containing information on wealth, a key variable in consumption models. Moreover, the panel structure of the PSID data allows for the inclusion of household fixed effects and thus to control for time-invariant unobserved heterogeneity.

To further distinguish EBL from other determinants that can be embedded in a life-cycle permanent-income model, we simulate the Low et al. (2010) model of consumption and labor supply. Their model accounts for various types of shocks, including productivity and job arrival, and allows for financial constraints as well as "income scarring"—the notion that job loss may have long-lasting effects on future income because it takes time to obtain an offer of the same job-match quality as before unemployment. We further extend the Low et al. (2010) model to allow for "unemployment scarring"—the notion that job loss itself may induce a negative, permanent wage shock.[5] We contrast these explanations with EBL by simulating the model for both Bayesian and experience-based learners.

First, we show that the main empirical features of experience effects—over-weighing of past lifetime experiences and resulting adjustments of consumption—are not generated when consumers have rational beliefs about the probability of being unemployed next period: There is no negative relation between lifetime experiences of unemployment and consumption, after controlling for income and wealth. This holds both when we allow for financial constraints and income scarring, as in the original Low et al. (2010) model, and when we add unemployment scarring. In fact, given the income control, the simulate-and-estimate exercise often predicts a *positive* relation between unemployment experiences and consumption. Intuitively, a consumer who has the same income as another consumer despite worse unemployment experiences likely has a higher permanent income component, and rationally consumes more.

We then turn to consumers who overweight their own past experiences when forming beliefs. Here, we find the opposite effect: Higher life-time unemployment experiences predict lower consumption among EBL agents, controlling for income and wealth. Thus, the simulate-and-estimate exercise disentangles EBL from potential confounds such as financial constraints, income scarring, and unemployment scarring. There is a robust negative relation between past experiences and consumption under EBL, consistent with the empirical estimates, but not under Bayesian learning.

The model also helps to alleviate concerns about imperfect wealth controls. We conduct both simulate-and-estimate exercises leaving out the wealth control in the estimation. In the case of rational consumers we continue to estimate a positive rather than negative relationship between prior unemployment experiences and consumption; in the case of experience-base learners, we continue to estimate a negative relationship. This holds whether or not we allow "only" for financial constraints and income scarring à la Low et al. (2010), or also for unemployment scarring. Only if we alter the experience-effect proxy to increase the overweighting of experiences far in the past, the (mis-specified) estimation without wealth controls misattributes the omitted wealth effect to EBL even though consumers are simulated to be Bayesian learners. The misattribution is insignificant under financial constraints and income scarring as in the Low et al. (2010) model, and becomes significant when we add unemployment scarring.

Guided by these simulation results, we perform three more empirical steps: (1) a broad range of robustness checks and replications using variations in the wealth, liquid-

[5] We thank the audience at the University of Minnesota macro seminar for this useful suggestion.

ity, and income controls, and using different data sets; (2) a study of the implications of EBL for the quality of consumption and of the heterogeneity in consumption patterns across cohorts, and (3) a discussion of the potential aggregate effects of EBL for consumption and savings.

First, we replicate the PSID results using four variants of wealth controls: third- and fourth-order liquid and illiquid wealth; decile dummies of liquid and illiquid wealth; separate controls for housing and other wealth; and controls for positive wealth and debt. Similarly, we check the robustness to four variants of the income controls: third- and fourth-order income and lagged income; quintile dummies of income and lagged income; decile dummies of income and lagged income; and five separate dummies for two-percentile steps in the bottom and in the top 10% of income and lagged income. All variants are included in addition to first- and second-order liquid and illiquid wealth and first- and second-order income and lagged income, and all estimations are replicated both with and without household fixed effects. We also subsample households with low versus high liquid wealth (relative to the sample median in a given year), and find experience effects in both subsamples.

We replicate the PSID results in two additional data sets, the Consumer Expenditure Survey (CEX) and the Nielsen Homescan Data. The CEX contains a more comprehensive list of product categories, and sheds light on the impact of unemployment experience on durable consumption and total consumption. The Nielsen data is a panel of consumption purchases by representative households in all U.S. markets. It contains detailed data on the products that households purchase at the Universal Product Code (UPC) level for each shopping trip, which allows us to control more finely for time (year-month) effects. The estimated magnitudes in the Nielsen and CEX data are very similar to those in the PSID.[6]

Next, we exploit the rich high-frequency nature of the Nielsen data and show that prior experiences affect consumption also at the qualitative margin. We estimate a significant increase in several measures of frugality: (i) the use of coupons, (ii) the purchase of lower-quality items (as ranked by their unit price, within product module, market, and month), and (iii) the purchases of on-sale products. For example, at the 90$^{\text{th}}$ percentile of unemployment experiences households purchase 9% more sale items annually than at the 10$^{\text{th}}$ percentile.

We then test a unique prediction of EBL: Macroeconomic shocks have particularly strong effects on younger cohorts, who increase their consumption more than older cohorts during economic booms, and lower their spending more during busts. This prediction captures that a given shock makes up a larger fraction of the lifetime-experience of younger people. We confirm the prediction both for aggregate and personal unemployment experiences, and both in the positive and in the negative direction.

[6] We have also explored the Health and Retirement Survey (HRS), which contains information on consumption (from the Consumption and Activities Mail Survey) and wealth on a biennial basis since 2001. However, given that cross-cohort variation is central to our identification, the lack of cohorts below 50 makes the HRS is not suitable for the analysis.

Our results imply that experience effects constitute a novel micro-foundation of fluctuations in aggregate demand and long-run effects of macro shocks. We provide suggestive evidence by correlating aggregate lifetime experiences of past national unemployment among the U.S. population with real personal consumption expenditure (PCE) from the U.S. Bureau of Economic Analysis (BEA) from 1965 to 2013. The resulting plot shows that times of higher aggregate past-unemployment experience in the population coincide with lower aggregate consumer spending. This suggests that changes in aggregate consumption may reflect not only responses to recent labor-market adjustments, but also changes in belief formation due to personal lifetime experiences of economic shocks. Overall, our findings imply that the long-term consequences of macroeconomic fluctuations can be significant, thus calling for more discussion on optimal monetary and fiscal stabilization policy to control unemployment and inflation (Woodford (2003) Woodford (2003), Woodford (2010)).

Related Literature Our work connects several strands of literature. Foremost, the paper contributes to a long, rich literature on consumption. Since the seminal work of Modigliani and Brumberg (1954) and Friedman (1957), the life-cycle permanent-income model has been the workhorse to study consumption behavior. Consumption decisions are an intertemporal allocation problem, and agents smooth marginal utility of wealth across predictable income changes over their life-cycle. Subsequent variants provide more rigorous treatments of the assumptions about uncertainty, time-separability, and the curvature of the utility function (see Deaton (1992) and Attanasio (1999) for overviews). A number of empirical findings, however, remain hard to reconcile with the model predictions. Campbell and Deaton (1989) point out that consumption does not react sufficiently to unanticipated innovation to the permanent component of income (excess smoothness). Instead, consumption responds to anticipated income increases, over and above what is implied by standard models of consumption smoothing (excess sensitivity; cf. West (1989), Flavin (1993)).

The empirical puzzles have given rise to a debate about additional determinants of consumption, ranging from traditional explanations such as liquidity constraints Gourinchas and Parker (2002)[7] to behavioral approaches such as hyperbolic discounting Harris and Laibson (2001), expectations-based reference dependence Pagel (2017); Olafsson and Pagel (2018), and myopia Gabaix and Laibson (2017).[8] Experience-based learning offers a unifying explanation for both puzzles. The lasting impact of lifetime income histories can explain both consumers' lack of response to permanent shocks and their overreaction to anticipated changes.

Overall, our approach is complementary to the existing life-cycle consumption literature: Experience effects describe consumption behavior after taking into account the established features of the life-cycle framework. EBL can explain why two individuals

[7] See also Kaplan et al. (2014); Deaton (1991); Aguiar and Hurst (2015).

[8] See also Dynan (2000) and Fuhrer (2000) on habit formation.

with similar income profiles, demographics, and household compositions still make different consumption choices if they lived through different macroeconomic or personal employment histories.

Our predictions and findings are somewhat reminiscent of consumption models with intertemporal non-separability, such as habit formation models (Meghir and Weber (1996), Dynan (2000), Fuhrer (2000)). In both cases, current consumption predicts long-term effects. However, the channel is distinct. Under habit formation, utility is directly linked to past consumption, and households suffer a loss of utility if they do not attain their habitual consumption level. Under EBL, households adjust consumption patterns based on inferences they draw from their past experiences, without direct implications for utility gains or losses.

Related research provides evidence on the quality margin of consumption. Nevo and Wong (2015) show that U.S. households lowered their expenditure during the Great Recession by increasing coupon usage, shopping at discount stores, and purchasing more goods on sale, larger sizes, and generic brands. While they explain this behavior with the decrease in households' opportunity cost of time, we argue that experience effects are also at work. The key element to identifying this additional, experience-based source of consumption adjustment are the inter-cohort differences and the differences in those differences over time. Relatedly, Coibion et al. (2015) show that consumers store-switch to reallocate expenditures toward lower-end retailers when economic conditions worsen.

The second strand of literature is research on experience effects. A growing literature in macro-finance, labor, and political economy documents that lifetime exposure to macroeconomic, cultural, or political environments strongly affects their economic choices, attitudes, and beliefs. This line of work is motivated by the psychology literature on the availability heuristic and recency bias (Kahneman and Tversky (1974), Tversky and Kahneman (1974)). The availability heuristic refers to peoples' tendency to estimate event likelihoods by the ease with which past occurrences come to mind, with recency bias assigning particular weight to the most recent events. Taking these insights to the data, Malmendier and Nagel (2011) show that lifetime stock-market experiences predict subsequent risk taking in the stock market, and bond-market experiences explain risk taking in the bond market. Malmendier and Nagel (2015) show that lifetime inflation experiences predict subjective inflation expectations. Evidence in line with experience effects is also found in college students who graduate into recessions (Kahn (2010), Oreopoulos et al. (2012)), retail investors and mutual fund managers who experienced the stock market boom of the 1990s (Vissing-Jorgensen (2003), Greenwood and Nagel (2009)), and CEOs who grew up in the Great Depression (Malmendier and Tate (2005), Malmendier et al. (2011)). In the political realm, Alesina and Fuchs-Schündeln (2007), Lichter et al. (2016), Fuchs-Schündeln and Schündeln (2015), and Laudenbach et al. (2018) provide evidence of the long-term consequences of living under communism, its surveillance system, and propaganda on preferences, norms, and financial risk-taking.

Our findings on experience effects in consumption point to the relevance of such effects in a new context and reveal a novel link between consumption, life-cycle, and the state of the economy. A novelty of our empirical analysis, compared to the existing literature, is that the detailed panel data allow us to identify effects using within-household variation, whereas earlier works such as Malmendier and Nagel (2011) (Malmendier and Nagel (2011), Malmendier and Nagel (2015)) rely solely on time variation in cross-sectional differences between cohorts.

In the rest of the paper, we first present the data and variable construction (Section 2.2), followed by the four baseline findings on consumption, beliefs, future income, and wealth build-up (Section 2.3). The stochastic life-cycle consumption model in Section 2.4 illustrates the differences between the consumption of rational and experience-based learners. Guided by the simulate-and-estimate exercise, we present additional wealth and income robustness tests in Section 2.5, and replicate the results in the CEX and Nielsen data. Section 2.6 shows further results on the quality margins of consumption and the cross-cohort heterogeneity in responses to shocks. Section 2.7 discusses the aggregate implications of experience-based learning for consumer spending and concludes.

2.2 Data and Variable Construction

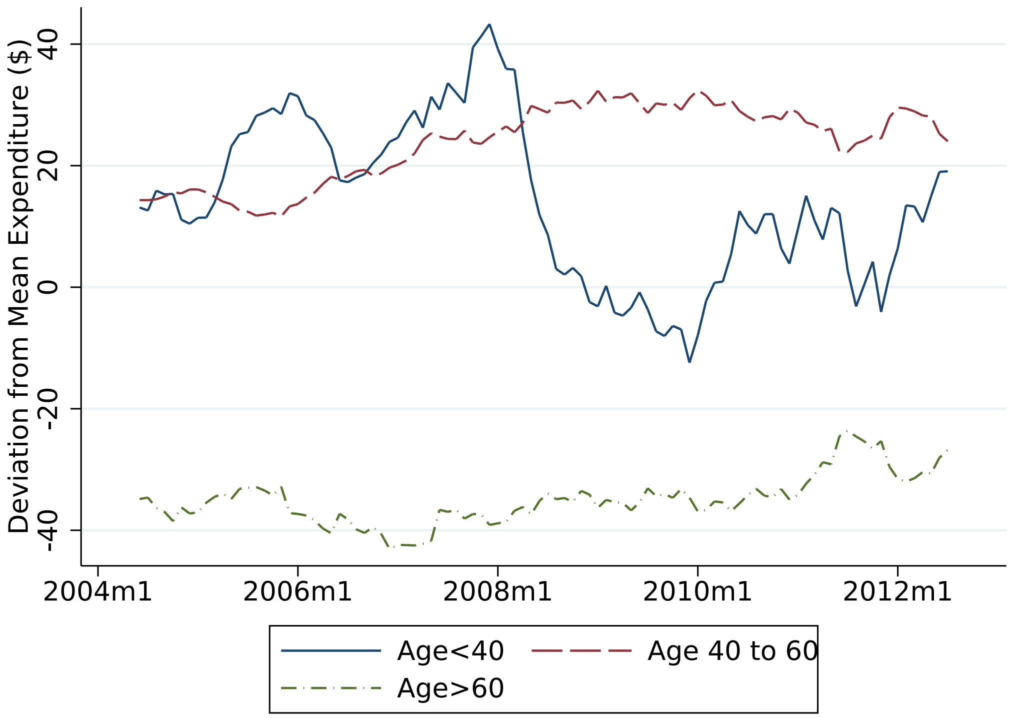
Measure of Experience Effects

The experience-effect hypothesis is based on the idea that individuals overweight realizations that have occurred during their lifetimes. In the context of consumption, the conjecture is that individuals who have lived through difficult economic times have more pessimistic beliefs about future job loss and income, and thus spend less relative to those who have lived through mostly good times, after controlling for wealth and income. The opposite holds for extended exposure to prosperous times. Moreover, the cross-sectional differences vary over time as households accumulate different experiences. Younger cohorts react more strongly to a shock than older cohorts since it makes up a larger fraction of their life histories so far.

The raw time-series of household expenditures (from the Nielsen data) in Figure 2.1 helps to illustrate the hypothesized effects. Expenditures are expressed as deviations from the cross-sectional mean in the respective month. In general, the spending of younger cohorts (below 40) is more volatile than that of older cohorts, consistent with younger cohorts exhibiting greater sensitivity. Zooming in on the Great Recession period, we also see that the spending of younger cohorts was significantly more negatively affected than those of the other age groups. Such patterns are consistent with consumers being scarred by recession experiences, and more so the younger they are.

To formally test the experience-effect hypothesis, we construct measures of past

Figure 2.1: **Monthly Consumption Expenditure by Age Group**



Notes. Six-month moving averages of monthly consumption expenditures of young (below 40), mid-aged (between 40 and 60), and old individuals (above 60) in the Nielsen Homescan Panel, expressed as deviations from the cross-sectional mean expenditure in the respective month, and deflated using the personal consumption expenditure (PCE) price index of the U.S. Bureau of Economic Analysis (BEA). Observations are weighted with Nielsen sample weights.

experiences that apply the weighting function estimated in prior work to the experience of times of high and low unemployment rates. We focus on experiences of unemployment rates following Coibion et al. (2015), who single out unemployment as a the most spending-relevant variable, and construct experience measures both on the macro (national and local) level and on the personal level. The macro measure captures the experience of living through various spells of unemployment rates. The personal measure captures the personal situations experienced so far.

Specifically, unemployment experience is measured as

$$E_t = \sum_{k=0}^{t-1} w\left(\lambda, t, k\right) W_{t-k}, \tag{2.1}$$

where W_{t-k} is the unemployment experience in year $t-k$, and k denotes how many years ago the unemployment was experienced.[9] The weights w are a function of t, k, and λ, where λ is a shape parameter for the weighting function. Following Malmendier and Nagel (2011), we parametrize the weighting function as

$$w(\lambda, t, k) = \frac{(t-k)^{\lambda}}{\sum_{k=0}^{t-1} (t-k)^{\lambda}}. \tag{2.2}$$

The specification of experience weights is parsimonious in that it introduces only one additional parameter to capture different possible weighting schemes for past experiences: If $\lambda > 0$, then past observations receive less weight than more recent realizations, i. e., weights are declining in time lag k. In that case, the weighting scheme emphasizes individuals' recent experiences, letting them carry higher weights, while still allowing for an impact of earlier life histories. For example, consider a 30-year-old in the early 1980s, when the national unemployment rate reached over 10%. While the experience of living through relatively low unemployment in the early 1970s (around 5-6%) as a 20-year-old may still influence her behavior, the influence is likely to be smaller relative to more recent experiences. In the case $\lambda \to \infty$, we have convergence towards the strongest form of recency bias. In our main empirical analyses, we will apply linearly declining weights ($\lambda = 1$), which approximate the weights estimated in Malmendier and Nagel (2011) (Malmendier and Nagel (2011), Malmendier and Nagel (2015)). For robustness, we also conduct the analysis using weight parameter, $\lambda = 0$ and $\lambda = 3$.

Empirically, we construct national, local, and individual measures of unemployment experiences, depending on the data set and individual information available. For the national macro measure, we combine several historical time series on national unemployment rates: a) the unemployment data from Romer (1986) for the period 1890-1930; b) the unemployment data from Coen (1973) for the period 1930-1939; c) the BLS series that counts persons aged 14 and over in the civilian labor force for the period 1940-1946; and d) the BLS series that counts persons aged 16 and over in the civilian labor force for the period 1947-present.[10]

[9] In the empirical implementation, we utilize unemployment information from birth up to year $t-1$ while the theoretical p_t is constructed based on realizations of W_{t-k} for $k = 0, ..., t-1$, i. e., from the moment of birth to the realization at the beginning of the current period. It is somewhat ambiguous what corresponds best to the theoretical set-up, especially as, in practice, only backward looking (macro) information becomes available to every individual. However, since we do control for (macroeconomic and personal) contemporaneous unemployment in all regressions, the inclusion or exclusion of macro or personal unemployment at time t in the experience measure does not make a difference to the estimation results.

[10] An alternative and widely cited source of historical 1890-1940 data is Lebergott (1957) 1957; 1964. Later research has identified multiple issues in Lebergott's calculations and has sought to modify the estimates to better match the modern BLS series. Romer (1986) singles out Lebergott's assumptions that (1) employment and output in some sectors move one-to-one, and (2) the labor force does not vary with the business cycle, as invalid and generating an excessively volatile time series. Coen (1973) finds that both armed forces and cyclical variations in average hours per worker have

For the more local, region-specific measure of unemployment experiences, we combine information on where a family has been living (since the birthyear of the household head) with information about local historical unemployment rates. Ideally, both sets of information would be available since the birthyear of the oldest generation in our data. However, the Bureau of Labor Statistics (BLS) provides state-level unemployment rates only since 1976, and there do not appear to be reliable sources of earlier historical unemployment data for all US states. These data limitations imply that, if we were to work with "all available" data to construct the region-specific experience measure, the values for family units from the later periods would be systematically more precise than those constructed for earlier periods, biasing the estimates. Hence, we have to trade off restricting the sample such that all family units in a given data set have sufficient location and employment rate data, and having sufficient sample to construct a reliable experience measure. We choose to use the five most recent years state-level unemployment rates, $t-5$ to $t-1$, either by themselves or combined with national unemployment rate data from birth to year $t-6$. In the former case, we first weight past experiences as specified in equation (2.2), applied to $k = 1, ..., 5$, and then renormalized the weights to 1. In the latter case, we use weights exactly as delineated in (2.2). As we will see, the estimation results are very similar under all three macro measures, national, regional, and combined. We will show the combined measure in our main regressions whenever geographic information on the individual level is available.

Finally, to construct the personal experience measure, we use the reported employment status of the respondent in the respective data set. We face the same data limitations as in the construction of the state-level macro experience measure regarding the earlier years in the lives of older cohorts. Mirroring our approach in the construction of the macro economic measure, we use the personal-experience dummy variables from year $t-5$ to $t-1$ and national unemployment rates from birth to year $t-6$, with weights calculated as specified in (2.2).

Consumption Data

Our main source of data is the PSID. It contains comprehensive longitudinal data on consumption at the household level and has long time-series coverage, which allows us to construct experience measures for each household. We will later replicate the results in Nielsen and CEX data. Compared to those data, the PSID has the advantage of containing rich information on household wealth, a key variable in consumption models.

The PSID started its original survey in 1968 on a sample of $4,802$ family units. Along with their split-off families, they were repeatedly surveyed each year until 1997, when the PSID became biennial. We focus on data since 1999 when the PSID started to cover more consumption items (in addition to food), as well as information on household

---

been ignored in previous studies, and these variables appear to have significant effects on measures of labor participation.

wealth. The additional consumption variables include spending on childcare, clothing, education, health care, transportation, and housing, and approximately 70% of the items in the CEX survey (Andreski et al. (2014)).

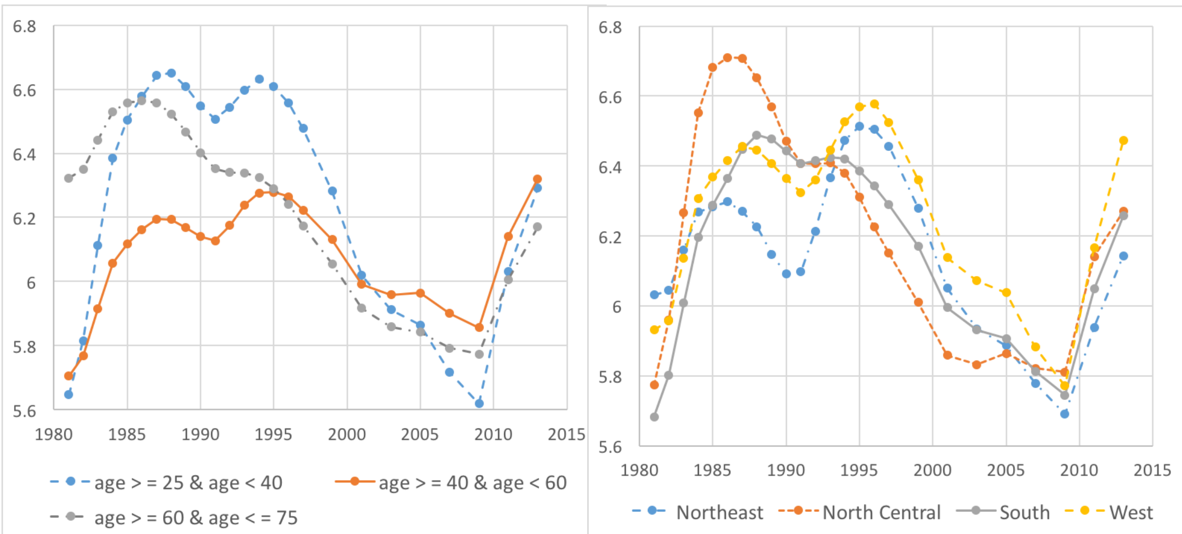
Regarding household wealth, the survey asks about checking and saving balances, home equity, and stock holdings. Those wealth variables allow us to control for consumption responses to wealth shocks, and to tease out the effects of experiences on consumption for different wealth groups. Indeed, compared to the Survey of Consumer Finances (SCF), which is often regarded as the gold standard for survey data on wealth, Pfeffer et al. (2016) assess the quality of the wealth variables in the PSID to be generally quite similar. The exceptions are "business assets" and "other assets," for which the PSID tends to have lower values. We construct separate controls for liquid and illiquid wealth, using the definitions of Kaplan et al. (2014). Liquid wealth includes checking and savings accounts, money market funds, certificates of deposit, savings bonds, treasury bills, stock in public companies, mutual funds, and investment trusts. Illiquid wealth includes the net values of home equity, other real estate, and vehicles, private annuities, IRAs, investments in trusts or estates, bond funds and life insurance policies.

The PSID also collects income information and a range of other household demographics, including years of education (ranging from 0 to 17), age, gender, race (White, African American or Others), marital status, and family size. The information on the head of household is significantly more complete than for other family members. Therefore, while the family is our unit of analysis, we focus on the experiences and demographic variables of the heads of the family in our baseline estimations, including our key explanatory variable measuring unemployment experiences. We then show the robustness to including the spouse's experiences.

The key explanatory variable is the past experience of each household head at each point in time, calculated as the weighted average of past unemployment experiences as defined in (2.1) and (2.2). The PSID allows us to construct measures of both macroeconomic and personal unemployment experiences, and to further use both national and more local (statewide) rates for the macro experience measure. As discussed above, the more local measure of unemployment experiences has to account for several data limitations. The oldest heads of household in the survey waves we employ are born in the 1920s, but the PSID provides information about the region (state) where a family resides only since the start of the PSID in 1968, and the Bureau of Labor Statistics (BLS) provides state-level unemployment rates only since 1976. As specified above, we use the five most recent years state-level unemployment rates, $t-5$ to $t-1$, either by themselves or combined with national unemployment rate data from birth to year $t-6$. The estimation results under all three macro measures, national, regional, and combined are very similar. We will show the combined measure in our main regressions.

To construct the personal experience measure, we first create a set of dummy variables indicating whether the respondent is unemployed at the time of each survey, using the reported employment status.[11] We face the same data limitations as in the construction of the state-level macro experience measure regarding the earlier years of the older cohorts, and employ the same approach.

Figure 2.2: **Unemployment Experience by Age Group and by Region**



Notes. The left graph shows the unweighted means of local unemployment experiences of different age groups. The right graph shows the unweighted means of local unemployment experiences in different regions.

Figure 2.2 illustrates the heterogeneity in lifetime experiences using our (combined) macroeconomic experience measure, both in the cross-section and over time, for the PSID families in our sample. The left panel of Figure 2.2 plots the unweighted mean experiences of young (below 40), middle-aged (between 40 and 60), and old individuals (above 60), while the right panel of Figure 2.2 plots the measures for individuals in the Northeast, North Central, South, and West. The plots highlight the three margins of variation that are central to our identification strategy: At a given point in time, people differ in their lifetime experience depending on their cohort and residential location, and these differences in experiences evolve over time.

Table 2.1 shows the summary statistics for our sample. We focus on household heads from age 25 to 75.[12] In the main analysis, we run the regressions excluding observations

[11] The PSID reports eight categories of employment status: "working now," "only temporarily laid off," "looking for work, unemployed," "retired," "permanently disabled," "housewife; keeping houses," "student," and "other". We treat "other" as missing, and "looking for work, unemployed" as "unemployed." We code all other categories as "not unemployed." One caveat here is that the PSID is biennial during our sample period. For all gap years t, we assume that the families stay in the same state and have the same employment status as in year $t-1$. Alternatively, we average the values of $t-1$ and $t+1$, as discussed in B.1.

[12] With the control for lagged income in our main estimations, the actual minimum age becomes

Table 2.1: **Summary Statistics (PSID)**

Variable	Mean	SD	p10	p50	p90	N
Age	47.61	12.06	32	47	65	33,164
Experience (Macro) [in %]	6.00	0.28	5.67	5.97	6.36	33,164
Experience (Personal) [in %]	4.55	14.27	0.00	0.00	18.92	33,164
Household Size	2.75	1.45	1	2	5	33,164
Household Food Consumption [in $]	8,452	5,153	2,931	7,608	14,999	33,164
Household Total Consumption [in $]	44,692	31,786	16,626	39,608	76,823	33,164
Household Total Income [in $]	80k	51k	22k	70k	155k	33,164
Household Liquid Wealth [in $]	38k	320k	-23k	0k	91k	33,164
Household Illiquid Wealth [in $]	222k	919k	1k	71k	513k	33,164
Household Total Wealth [in $]	260k	1,007k	-3k	72k	636k	33,164

Notes. Summary statistics for the estimation sample, which covers the 1999-2013 PSID waves. We exclude from the sample observations with total family income below the 5^{th} or above the 95^{th} percentile in each wave from 1999 to 2013, as well as the pre-sample 1997 wave (because we control for lagged income). Age, Experience (Macro), and Experience (Personal) are calculated for the heads of households. Household total income includes transfers and taxable income of all household members from the last year. Liquid wealth and illiquid wealth are defined following Kaplan, Violante and Weidner (2014). All values are in 2013 dollars using the PCE. Observations are annual and not weighted.

with total family income below the 5^{th} or above the 95^{th} percentile in each wave. The sample truncation addresses known measurement errors in the income variable.[13] After dropping the individuals for whom we cannot construct the experience measures (due to missing information about location or employment status in any year from t to $t-5$), and observations with missing demographic controls or that only appear once, we have 33,164 observations. The mean of the macroeconomic experience measure is 6.0%, and that of the personal experience measure is 4.6%. The average household food consumption and the average household total consumption in our sample are $8,452 and $44,692, respectively, measured in 2013 dollars.

27. Additionally, we also conduct the analysis on a subsample that excludes retirees (households over the age of 65) since they likely earn a fixed income, which would not be affected by beliefs about future economic fluctuations. The results are similar.

[13] Gouskova and Schoeni (2007) evaluate the quality of the family income variable in the PSID by comparing it to family income reported in the Current Population Survey (CPS), which is used for compiling the government's official estimates of income and poverty. The comparison shows that the income distributions from the two surveys closely match for incomes between the 5^{th} and 95^{th} percentiles. However, there is less consensus in the upper and lower five percentiles of the income distributions. As a robustness check, we also re-estimate the regression model on the full sample. The summary statistics are in Appendix-Table B.1.

2.3 Baseline Results: Consumption and Consumer Optimism, Future Income and Future Wealth

Our analysis starts from the observation that macro shocks appear to have a long-lasting impact on consumer behavior, and that the puzzling persistence of reduced consumer expenditures appears to correlate with consumer confidence remaining low for longer than standard models would suggest (Pistaferri (2016)). Building on a growing literature on experience effects, we ask whether we can better predict consumer confidence and consumer behavior if we allow for a role of consumers' prior lifetime experiences of economic conditions. Measures of prior lifetime experiences have been found to have longlasting effects on individual beliefs and decision-making in the realms of stock returns, bond returns, inflation, and mortgage choices. Here, we ask whether a similar mechanism might help to explain patterns in consumption expenditures. Specifically, we measure spending-relevant macro conditions in terms of higher or lower unemployment rates as in Coibion et al. (2015), both on the aggregate level (unemployment rates) and on the personal level. We then show that past experiences of unemployment have a measurable, lasting effect on individual beliefs and consumption expenditures, but fail to predict (lower) future income or future wealth.

Past Experiences and Consumption

Do lifetime experiences of unemployment predict consumption spending in the long run? We relate expenditures to prior experiences of economic conditions by estimating:

$$C_{it} = \alpha + \beta UE_{it} + \psi UEP_{it} + \gamma' x_{it} + \eta_t + \varsigma_s + \upsilon_i + \varepsilon_{it}, \tag{2.3}$$

where C_{it} is consumption, UE_{it} is i's macroeconomic unemployment experience over her prior life so far, UEP_{it} is her personal unemployment experience over her prior life so far, and x_{it} a vector of control variables including wealth (first and second order of the logarithm of liquid and illiquid wealth), income (first and second order of the logarithm of income and lagged income), age dummies, and household characteristics (unemployment dummy indicating if the household head is currently unemployed, family size, gender, years of education (ranging from 0 to 17), marital status, and race (White, African American, and other)). Finally, η_t are time (year) dummies, ς_s are state dummies, and υ_i are household dummies. Standard errors are clustered at the cohort level.[14] We conduct our empirical analysis both with food consumption as the dependent variable, following the earlier consumption literature, and with total

[14] We have also estimated the model including region∗year fixed effects, and the results remain very similar. Note that we do not include state∗year fixed effects in the model since one of the key margins of variation in our main regressor of interest, macroeconomic unemployment experience (UE_{it}), is at the state∗year level.

consumption.[15] The standard errors are clustered at the cohort level. All the regression results are quantitatively and qualitatively similar when clustered by household, household-time, and cohort-time, and two-way clustered at the cohort and time level.[16]

Our main coefficients of interest are β and ψ. The rational null hypothesis is that both coefficients are zero. The alternative hypothesis, based on the idea of experience effects, is that consumers who have experienced higher unemployment spend less on average, and hence that both coefficients are negative.

Identification. We estimate the model both with and without household dummies. In the former case, we identify experience effects in consumption solely from time variation in the within-household co-movement of consumption and unemployment histories. In the latter case, identification also comes from time variation in cross-sectional differences in consumption and unemployment histories between households.

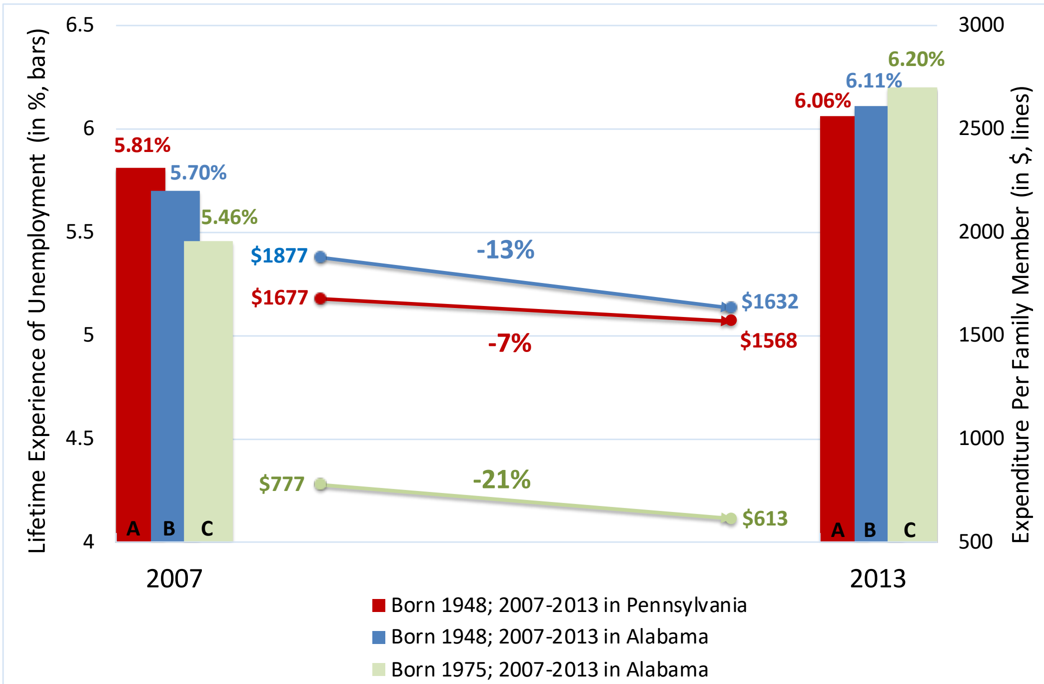
We illustrate the sources of identification with a simple example of the unemployment experiences and household consumption of three individuals in our PSID data over the course of the Great Recession. Consider two individuals (A and B) who have the same age (born in 1948) but live in different states during the 2007-2013 period (Pennsylvania and Alabama), and a third (C) who lives in the same state as B (Alabama) but differs in age (born in 1975).

The two sets of bars in Figure 2.3 illustrate their lifetime experiences of unemployment at the beginning and at the end of the 2007-2013 period, based on the weighting scheme in (2.2) and their states of residence. Person A enters the crisis period with a higher macroeconomic unemployment experience than Person B (5.81% versus 5.70%), but her lifetime experience worsens less over the course of the financial crises and becomes relatively more favorable by 2013 (6.06% versus 6.11%) because unemployment rates were lower in Pennsylvania than in Alabama during the crisis period. Person C has even lower macroeconomic unemployment experiences before the crisis period than person B (5.46%); but being the younger person, C is more affected by the crisis, leading to a reversal of the lifetime unemployment experience between old and young by the end of the crisis (6.11% versus 6.20%). Figure 2.3 relates these differences-in-differences of lifetime experience over the crisis period to consumption behavior. The increase in unemployment experiences of Person A, B, and C by 0.25%, 0.41%, and 0.74%, respectively, were accompanied by decreases in consumption in the same relative ordering, by 7%, 13%, and 21%, respectively.

[15] Food consumption has been most widely used in the consumption literature largely because food spending used to be the only available consumption variable in the PSID before the 1999 survey wave. We are separating out the results on food consumption post-1999 partly for comparison, but also in case the data is more accurate as some researchers have argued. Food consumption and total consumption come directly from the PSID Consumption Data Package 1999-2013.

[16] The results are robust to clustering at different levels and applying different weights, as we discuss below.

Figure 2.3: **Examples of Experience Shocks from the Recession (PSID)**



Notes. The red (dark) bars depict the 2007 and 2013 unemployment experiences of person A, and the red (dark) line the corresponding change of total consumption per family member in person A's family. Similarly, the blue (medium dark) bars and line show person B's unemployment experiences and consumption, and the green (light) bars and line person C's unemployment experiences and consumption. All consumption expenditures are measured in 2013 dollars, adjusted using PCE. Person A's ID in the PSID is 45249; person B's ID in the PSID is 53472; person C's ID in the PSID is 54014.

Results Table 2.2 shows the estimation results from model (2.3) with (log) food consumption as the dependent variable in the upper panel, and with (log) total consumption in the lower panel. All regressions control for first- and second-order (logs of) income, lag income, liquid wealth, illiquid wealth, for all other control variables listed above, as well as the fixed effects indicated at the bottom of the table. Columns (1)-(3) show results without household fixed effects, and columns (4)-(6) with household fixed effects. All estimated coefficients on the control variables (not shown) have the expected sign, consistent with prior literature.

The estimated negative coefficients indicate that both macroeconomic and personal unemployment experiences predict reduced consumption expenditures in the long-run. In the estimations predicting food consumption, shown in the upper half of the table, we find a significantly negative effect of both macroeconomic and personal experiences,

Table 2.2: **Experience Effects and Annual Consumption (PSID)**

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Food Consumption						
Experience (Macro)	-0.097**		-0.091*	-0.120**		-0.117**
	(0.047)		(0.047)	(0.054)		(0.055)
Experience (Personal)		-0.322***	-0.320***		-0.263**	-0.260**
		(0.097)	(0.097)		(0.119)	(0.119)
R-squared	0.192	0.193	0.193	0.541	0.542	0.542
Dependent Variable: Total Consumption						
Experience (Macro)	-0.022		-0.018	-0.059***		-0.057***
	(0.019)		(0.019)	(0.021)		(0.021)
Experience (Personal)		-0.178***	-0.177***		-0.148***	-0.147***
		(0.030)	(0.030)		(0.031)	(0.031)
R-squared	0.573	0.574	0.574	0.788	0.788	0.788
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Income controls	Yes	Yes	Yes	Yes	Yes	Yes
Wealth controls	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	No	No	No	Yes	Yes	Yes
Observations	33,164	33,164	33,164	33,164	33,164	33,164

Notes. The consumption variables come from the 1999-2013 PSID Consumption Expenditure Data package. We take the logarithm of consumption, income, and wealth; non-positive values are adjusted by adding the absolute value of the minimum plus 0.1 before being logarithmized. "Experience (Macro)" is the macroeconomic experience measure of unemployment, and "Experience (Personal)" is the personal experience measure, as defined in the text. Demographic controls include family size, heads' gender, race, marital status, education level, and a dummy variable indicating whether the respondent is unemployed at the time of the survey. Income controls include the first and second order of the logarithm of income and lagged income. Wealth controls include the first and second order of the logarithm of liquid and illiquid wealth. We exclude from the sample observations with total family income below the 5^{th} or above the 95^{th} percentile in each wave Robust standard errors (in parentheses) are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

controlling for the current unemployment status, when we do not control for household fixed effects. The economic magnitudes remain the same whether we include the two types of experience measures separately (in columns 1 and 2) or jointly (in column 3), though the statistical significance of the macro measure diminishes somewhat when we include both measures jointly in the specification without household fixed effects.

When we introduce household fixed effects (in columns 4 to 6), the estimated coefficient on macro experience becomes larger and more precisely estimated. Based on the column (6) estimates, a one standard-deviation increase in macroeconomic unemployment experience leads to a 3.3% decrease in food consumption, which translates to \$279 less annual spending. Hence, the economic magnitude of the macro experience effect alone is large, particularly considering that the estimates reflect behavioral change due to fluctuation in the macro-economy, not direct income shocks.

As expected, the estimated personal experience effects become slightly smaller when we include household fixed effects. The decrease reflects that experience effects (also) predict cross-sectional differences in consumption between households with "mostly good" versus "mostly bad" lifetime experiences, and this component of experience effects is now differenced out. Nevertheless, the effect of personal experience is more than two times larger than macroeconomic experience in absolute value. The estimated effect of a one standard-deviation increase in personal unemployment experiences is similar to a one standard-deviation increase in macro experiences and predicts a 3.7% decrease in food consumption, which is approximately \$314 in annual spending.

When we use total consumption as the dependent variable, in the lower half of Table 2.2, the economic magnitude of the macro-economic experience effect decreases in the specification without household fixed effects (columns 1-3) but is again as precise as in the case of food consumption when we include household fixed effects (columns 4-6). In terms of economic magnitude, a one standard-deviation increase in macroeconomic experience lowers total consumption by 1.6%, or \$713 in annual spending, based on the estimated coefficient from column (6). A one standard-deviation increase in personal lifetime unemployment experience lowers total consumption by 2.1%, or \$937 annually.

We also re-estimate the results on the entire sample, without excluding observations in the top and bottom 5 percentiles of income. As shown in Appendix-Table B.2, the coefficients on macroeconomic and personal unemployment experiences become both larger (in absolute value) and more statistically significant.

The results are also robust to several variations in the construction of the key explanatory variable. First, as discussed above, our baseline specification fills the gap years of the (biennial) PSID assuming that families stay in the same state and have the same employment status as in the prior year. Alternatively, we average the values of the prior and the subsequent year, $t-1$ and $t+1$. This variation affects both the experience proxy and several control variables. As shown in Appendix-Table B.3, the results are robust. Second, our results are robust to including both the head of the household and the spouse in the construction of the experience measure (Appendix-Table B.4), to excluding the experience of year $t-1$ from the measure (Appendix-Table B.5), and to using different weighting λ (Appendix-Table B.6). In terms of alternative approaches to calculating standard errors, we estimate regressions with standard errors clustered at different levels in Appendix-Table B.7. We also vary the weighting of observations by applying the PSID family weights, shown in Appendix-Table B.8. (We do not use PSID family weights in the main regression due to efficiency concerns.)

Overall, the results robustly show that consumers with more adverse macroeconomic and personal unemployment experience tend to spend less, both on food and in total, controlling for wealth, income, employment, family structures, and demographics.

Past Experiences and Beliefs

Given the robust findings of a negative and significant relationship between people's lifetime experiences of economic conditions and their consumption behavior, we turn to explore the channels through which lifetime experiences affect consumption choices. To what extent do personal lifetime experiences color beliefs about future outcomes?

We relate past lifetime experiences of economic fluctuations to current beliefs about future economic prospects, using microdata on expectations from the Reuters/Michigan Survey of Consumers (MSC). The MSC has been conducted by the Survey Research Center at the University of Michigan since the early 1950s, quarterly until Winter 1977, and monthly since 1978. The dataset is in repeated cross-section format and includes a total of 213,177 observation. On average, 630 individuals are surveyed each month (or quarter). Our sample period runs from 1953 to 2012.

Among the multitude of belief elicitations, we identify two questions that capture expectations about economic conditions and consumption. The first question elicits beliefs about one's future financial situation: "Now looking ahead – do you think that a year from now you will be better off financially, or worse off, or just about the same as now?" The second question is about expenditures for (durable) consumption items and individuals' current attitudes towards buying such items: "About the big things people buy for their homes – such as furniture, refrigerator, stove, television, and things like that. Generally speaking, do you think now is a good or bad time for people to buy major household items?" For the empirical analysis, we construct two binary dependent variables. The first indicator takes the value of 1 if the respondent expects better or the same personal financial conditions over the next 12 months, and 0 otherwise. The second indicator is 1 if the respondent assesses times to be good or the same for durable consumption purchases, and 0 otherwise.

We also extract income and all other available demographic variables, including education, marital status, gender, race, and age of the respondent. The explanatory variable of interest is again our measure of lifetime unemployment experiences. Since the MSC does not reveal the geographic location of survey respondents, we apply (2.1) to the national unemployment rates to construct the "Experience (Macro)" variable for each of individual i from birth until year t, and apply (2.2) to calculate the weighted average of past unemployment experiences. We construct the measure for each respondent at each point in time during the sample. As discussed, this construction of lifetime experiences emphasizes individuals' recent experiences, letting them carry higher weights, while still allowing for an impact of earlier life histories.

We regress the indicators of a positive assessment of one's future financial situation and of a positive buying attitude on past unemployment experiences, controlling for all available other predictors of consumer optimism – current unemployment, income, demographics, age fixed effects and year fixed effects. Year fixed effects, in particular, absorb all current macroeconomic conditions as well as all historical information available at the given time.

Table 2.3 shows the results of the corresponding linear least-squares estimations. In columns (1) to (3), we present the estimates of the relation between of lifetime experiences of national unemployment rates and respondents' forecasts of their own future situation. We find that people who have experienced times of greater unemployment during their lives so far expect significantly worse future financial conditions. The statistical and economic significance of the estimated experience effect is robust to variations in the set of controls: Whether we include only the fixed effects (age and time dummies), or add a control for income, or include all demographic variables, we always estimate a highly significant coefficient between -0.017 and -0.014 on past lifetime unemployment experiences.

The robustness of the estimates to the income control is interesting since the MSC data provides only limited controls for respondents' financial situation. This renders the estimates in columns (1) to (3) open to alternative explanations, especially in terms of unobserved wealth effects. When we include income in columns (2) and (3), it has the expected positive coefficient, and the same holds for demographics that might proxy for unobserved wealth (e. g., education) in column (3). Still, the coefficient of past experiences of national unemployment rates remains highly significant and negative.

In terms of the economic magnitude, consider the inter-decile range of lifetime experiences: Respondents who have experienced unemployment rates at the 90th percentile of the sample are around 2.5 percent more likely to say financial conditions will be worse in the next 12 months than respondents in the 10th percentile.

The estimations based on the second question, shown in columns (4) to (6), generate very similar results. We use the same econometric model and same variations in control variables, but substitute the dependent variable with our indicator for "buying attitude." We estimate a significantly negative effect of lifetime experiences of unemployment. The coefficient is again fairly stable, ranging from -0.059 to -0.046. Respondents who have experienced unemployment rates at the 90th percentile of the sample are 8.5 percent more likely to say now is a bad time to buy major household items than respondents in the 10th percentile. This analysis also addresses concerns about alternative explanations based on unobserved, which should not affect the respondent's assessment of general economic conditions.

This second baseline analysis suggests that the economic conditions individuals have experienced in the past have a lingering effect on their beliefs about the future. Differ-

Table 2.3: **Experience Effects and Expectations**

	Expected financial condition coming year (1 = Better or Same, 0 = Worse)			Good/bad time to buy major household items (1 = Good or Same, 0 = Bad)		
	(1)	(2)	(3)	(4)	(5)	(6)
Experience (Macro)	-0.017***	-0.015***	-0.014***	-0.059***	-0.050***	-0.046***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Unemployment rate	-0.015***	-0.015***	-0.015***	-0.044***	-0.044***	-0.043***
	(0.004)	(0.005)	(0.005)	(0.001)	(0.001)	(0.002)
Income		0.017***	0.021***		0.051***	0.039***
		(0.001)	(0.001)		(0.001)	(0.002)
Demographic controls	No	No	Yes	No	No	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	209,786	197,129	193,483	203,492	191,420	187,909
R-squared	0.047	0.048	0.048	0.057	0.065	0.069

Notes. All variables are from the Michigan Survey of Consumers (MSC). The dependent variable in columns (1)-(3) is the response to the question "Now looking ahead – do you think that a year from now you will be better off financially, or worse off, or just about the same as now?" (1 = Better off or about the same, 0 = Worse off) reported by individual respondents in the Michigan Survey of Consumers. Dependent variable in columns (4)-(6) is response to the question "About the big things people buy for their homes – such as furniture, refrigerator, stove, television, and things like that. Generally speaking, do you think now is a good or bad time for people to buy major household items?" (1 = Good (or Same), 0 = Bad) reported by individual respondents. Estimation is done with least squares, weighted with sample weights. "Experience (Macro)" is the macroeconomic experience measure of unemployment, constructed as a lifetime linearly-declining weighted national unemployment rate experienced by households. Demographic controls include education, marital status, gender, and race. Age controls are dummy variables for each age. The sample period runs from 1953 to 2012. Standard errors, shown in parentheses, are robust to heteroskedasticity. *, **, *** denote 10%, 5%, and 1% significance, respectively.

ently from the discussion about consumption expenditures and concurrent consumer confidence during the post-Great Recession period, the experience measure captures the predictive power of prior experiences on current consumer confidence. Individuals who have lived through worse experiences in the past consider their own financial future to be less rosy, and times to be generally bad for spending on durables, after controlling for all historical data, current unemployment, and other macro conditions. This evidence on the beliefs channel is consistent with prior literature on experience effects, including Malmendier and Nagel (2011) (Malmendier and Nagel (2011), Malmendier and Nagel (2015)). While it is possible that the responses to the first question reflect unobserved wealth effects and other unobserved financial constraints, the second question suggests that unobserved wealth effects are of limited explanatory power also in the context of beliefs. Here, respondents are asked about "times in general" and seem to strongly rely on their personal experiences to draw conclusions about economic conditions more broadly.

Past Experiences and Future Income

Next we ask whether the long-term reduction in consumption after past unemployment experiences, as well as the ensuing consumer pessimism, might be the (rational) response to lower employment and earnings prospects. Might the consumer pessimism be explained by (unobserved) determinants of households' future income that are correlated with past unemployment experiences? As we will see, the answer is no.

To test whether past unemployment experiences are correlated with (unobserved) determinants of households' future income, we re-estimate our baseline model from equation (2.3) with the dependent variables changed to future income either one or two or three survey waves in the future, i. e., two, four, and six years ahead.

The estimation results are in Table 2.4. They suggest that unemployment experiences do not play a significant role in explaining future household income. After controlling for income, wealth, employment status, the other demographics, and fixed effects, the estimated coefficients of personal unemployment experiences are all positive, small, and insignificant. For macroeconomic experiences, we estimate small negative coefficients, which are also insignificant with the exception of the estimation predicting income four years ahead, where it is marginally significant. In summary, our results imply that past lifetime experiences do not predict future earnings prospects.

Relatedly, one may be concerned whether past unemployment experiences affect the volatility (rather than the mean) of future income. To test for such a relationship, we change the dependent variable in our baseline model (2.3) to income volatility. Following Meghir and Pistaferri (2004) and Jensen and Shore (2015), we construct two measures of income volatility, one reflecting the variance of permanent income and one reflecting the variance of transitory income. The permanent-variance measure is the product of two-year and six-year changes in excess log income (from year $t-2$ to t and $t-4$ to $t+2$, respectively). The transitory-variance measure is the squared two-year

Table 2.4: **Experience Effects and Future Income**

	Income_{t+2}	Income_{t+4}	Income_{t+6}
Experience (Macro)	-0.030	-0.044*	-0.050
	(0.020)	(0.023)	(0.030)
Experience (Personal)	0.010	0.021	0.017
	(0.013)	(0.013)	(0.021)
Income controls	Yes	Yes	Yes
Wealth controls	Yes	Yes	Yes
Demographic controls	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes
Observations	15,710	11,258	7,641
R-squared	0.865	0.884	0.903

Notes. The dependent variables are future income in two, four, and six years, respectively. "Experience (Macro)" is the macroeconomic experience measure of unemployment, and "Experience (Personal)" is the personal experience measure. Demographic controls include family size, heads' gender, race, marital status, education level, and a dummy variable indicating whether the respondent is unemployed at the time of the survey. Income controls include the first and second order of the logarithm of income and lagged income. Wealth controls include the first and second order of the logarithm of liquid and illiquid wealth. We exclude from the sample observations with total family income below the 5^{th} or above the 95^{th} percentile in each wave from 1999 to 2013, as well as the pre-sample 1997 wave (because we control for lagged income). We take the logarithm of income, and wealth; non-positive values are adjusted by adding the absolute value of the minimum plus 0.1 before being logarithmized. Robust standard errors (in parentheses) are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

change in excess log income, where excess log income is defined as the residual from an OLS regression of log income on our full slate of control variables. Re-estimating the regressions with either measure as the dependent variable, we do not find any significant correlations between unemployment experiences and income volatility.

Past Experience and Wealth Build-up

The significant effect of lifetime unemployment experiences on consumption, and the lack of a relation with future income, imply that household experiences could even affect the build-up of wealth. In the case of negative lifetime experiences, consumers appear to restrain from consumption expenditures more than rationally "required" by their income and wealth position. This experience-induced frugality, in turn, predicts more future wealth. Vice versa, consumers who have lived through mostly good times are predicted to be spenders, and should thus end up with less wealth.

In order to test whether experience effects are detectable in long-run wealth accumulation, we relate households' lifetime experiences to their future wealth, using up to seven survey waves (14 years) into the future. We consider both liquid wealth and total wealth. Note that this analysis also ameliorates potential concerns about the quality of the consumption data and alternative life-cycle interpretations of our findings.

Figure 2.4 summarizes the coefficients of interest graphically for 10 regressions, namely, the cases of wealth at $t+6$, $t+8$, $t+10$, $t+12$, and $t+14$. The upper part shows the coefficient estimates when studying the impact on liquid wealth, and the lower part shows the estimates for total wealth. All coefficient estimates are positive, though the impact of macro experiences is smaller and (marginally) significant only in a few cases, namely, the more recent years for total wealth and the years further in the future for liquid wealth. The estimates of the role of personal lifetime experiences are also all positive, much larger, and typically significant, with coefficients ranging from 0.02 to 0.03 for liquid wealth and from 0.08 to 0.10 for total wealth. These estimates imply that a one-standard deviation increase of personal lifetime experiences of unemployment will lead to additional precautionary savings and resulting wealth build-up of about 1.3% or $4,500 10 years later. In other word, households who have experienced high unemployment tend to accumulate more wealth down the road. Appendix-Table B.12 provides the details on the coefficient estimates of both experience measures.

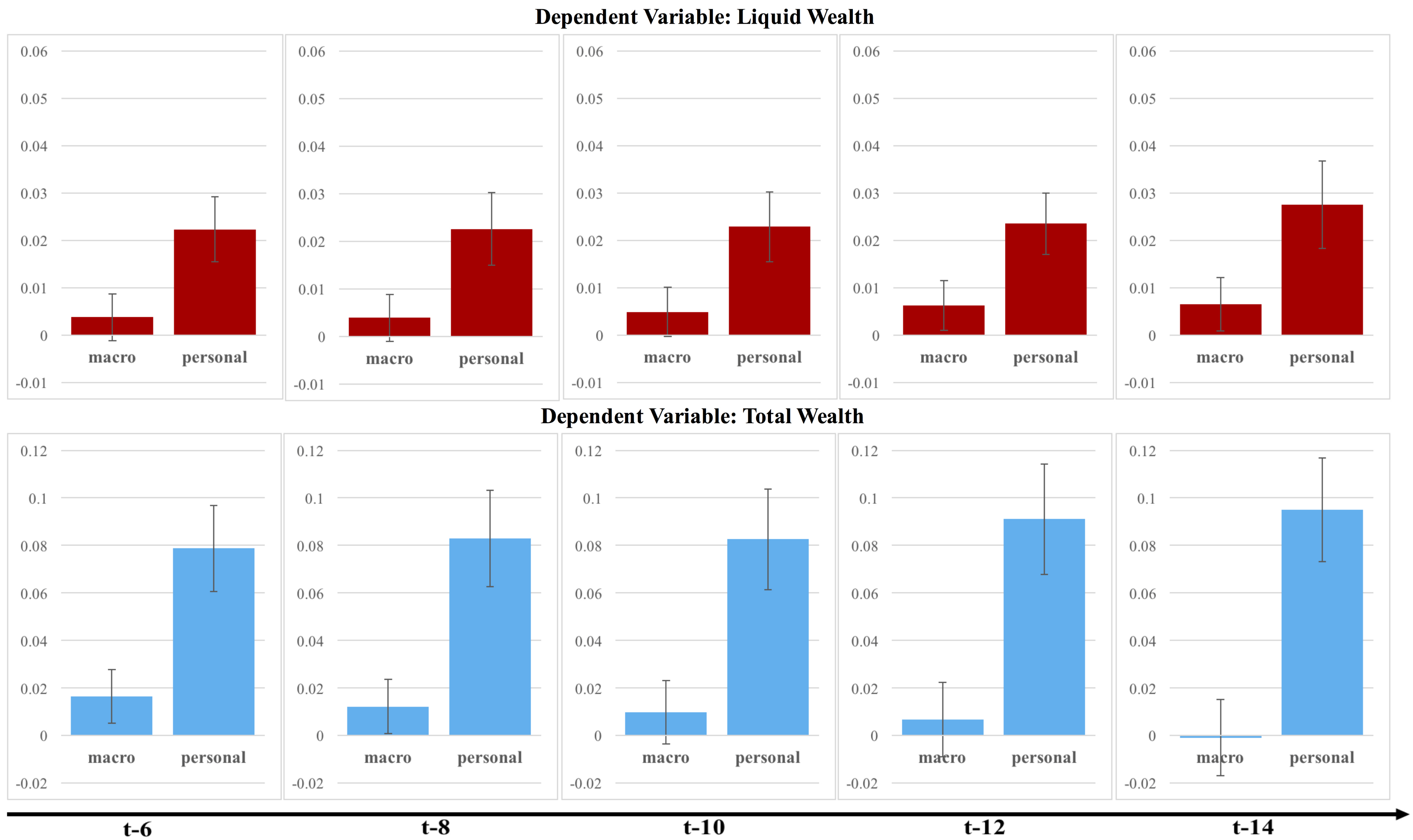
In summary, individuals' lifetime experiences strongly predict consumption expenditure, and beliefs about future economic conditions appear play a role in explaining this result. However, such beliefs do not seem to be consistent with actual income and wealth changes. In fact, we see evidence of a positive relationship between past experience and future wealth build-up.

2.4 Consumption with Experience-based Learning

Our four baseline results suggest that past experiences can "scar" consumption choices. We utilize the Low et al. (2010) model to explore a rich set of possible explanations for the observed relationship between past experiences and consumption.

The Low et al. (2010) framework captures standard life-cycle consumption factors, including financial constraints, social-insurance programs, and "income scarring," i. e., the notion that job loss reduces income flows because of lower match quality in future jobs. The focus of Low et al. (2010) is on the interaction of different types of risk (productivity shocks, employment risk) with social insurance (unemployment insurance, food stamps, and disability insurance). The social-insurance programs add richness to our analysis, and ensure that the experience-effect estimates are not confounded by the funding they provide. However, neither social insurance nor taxation are the focus

Figure 2.4: **Wealth Build-up: Estimated Coefficients and Confidence Intervals for Experience Measures**



Notes. The upper five graphs (red bars) present the estimates when we re-estimate our empirical model using liquid wealth as the dependent variable. The lower five graphs (blue bars) show the estimates when we use total wealth as the dependent variable. The five graphs in horizontal order show the estimated coefficients when we use 6-year lagged, 8-year lagged, 10-year lagged, 12-year lagged and 14-year lagged experience measures respectively. All the confidence intervals are at 90% confidence level.

of this paper. Instead, we utilize the Low et al. (2010) model to provide guidance in distinguishing experience effects from a rich set of alternative explanatory factors. Moreover, we extend the Low et al. (2010) to include "unemployment scarring," i. e., the notion that unemployment, once experienced, makes individuals inherently less employable. We distinguish both mechanisms as well as other life-cycle features from experience effects.

Towards that end, we introduce two classes of consumers into the model: standard rational agents and experience-based learners. Rational consumers use all available historical data to update their beliefs about the probability of being unemployed next period. Experience-based consumers overweight their own experiences when forming beliefs. We simulate intertemporal consumption and labor decisions for both types of consumers, and the estimate the relation between experiences and consumption in both setting, i. e., also for rational consumer. The simulate-and-estimate exercise illustrates the basic mechanism of experience-based learning, and distinguishes it from features of the standard consumption model, such as wealth or liquidity constraints. It thus provides guidance towards empirical robustness checks and additional tests.

Low, Meghir, and Pistaferri (2010) Model Setup. Consumers can work for 40 years, until age 62 (starting at age 23), then have mandatory 10 years of retirement where they receive social-security benefits, and die at the end of retirement. Periods are quarters, amounting to $L = 200$ periods of consumption and labor decisions in total. Their utility function is

$$U(c, P) = \frac{\left(c \times e^{\eta P}\right)^{1-\gamma}}{1-\gamma}, \tag{2.4}$$

where c is consumption, and P an indicator equal to 1 if a person works. In each t, consumer i chooses consumption $c_{i,t}$ and, when applicable, labor supply $P_{i,t}$ to maximize lifetime expected utility

$$\max_{\substack{c_{i,t} \\ P_{i,t}}} V_{i,t} = U(c_{i,t}, P_{i,t}) + \mathrm{E}_t \left[\sum_{s=t+1}^{L} \beta^{s-t} U(c_{i,s}, P_{i,s}) \right]. \tag{2.5}$$

We impose $c_{i,t} < A_t$, which rules out borrowing. As we will see below, by maximizing the financial constraints of consumers, we are able to derive the sharpest distinction between the role of experience effects and financial constraints.[17] We assume that

[17] The reason is that (unobserved) financial constraints, especially of younger cohorts, are a potential confound in the interpretation of the empirical relation between lifetime experiences and consumption: Younger cohort are predicted to react more strongly to a given shock than older cohorts under the experience-effect hypothesis, and they also tend to be more constrained in their borrowing ability. By eliminating borrowing altogether from the simulation, we maximize the impact of financial constraints.

flow utility takes a near CRRA form which induces a precautionary savings motive. (A more detailed description of the intertemporal budget constraint and the various social-insurance programs that affect it, is in B.2.)

Income Process The wage in this model is determined by the following formula

$$\ln w_{i,t} = d_t + x'_{i,t}\psi + u_{i,t} + a_{i,j,t_0}, \tag{2.6}$$

where d_t is the log-price of human capital at time t, $x'_{i,t}\psi$ is the component determined by i's age at time t, $u_{i,t}$ is the stochastic component, and $a_{i,,j,t_0}$ is the job-fit component of i's wage at firm j for a job offered (and accepted) in period t_0. Gross quarterly income is $w_{i,t}h$, where h is the number of hours worked in a quarter. The three social-insurance programs Low et al. (2010) include in their model are detailed in B.2.

Agents have the ability to make decisions about whether or not to work. For example, agents need not work if an offer is too low. They can also retire early. Note that this implies that experienced-based learners may make different labor supply choices depending on their concern about future employment and desire to save.

The Deterministic Component of Wage. We denote $d_t + x'_{it}\psi$ as the deterministic component of the wage as it is the same for all individuals of a given age at time t. The size of this component is estimated via regression in Low et al. (2010) and is of the form [18]

$$d_t + x'_{i,t}\psi = \alpha + \beta_1 \cdot \text{age} + \beta_2 \cdot \text{age}^2. \tag{2.7}$$

The Permanent Component of Wage. The stochastic component of the wage, $u_{i,t}$, is determined by a random walk. Consumers receive a shock to this component on average once a year. If consumer i has an income shock in period t, then $u_{i,t}$ is

$$u_{i,t} = u_{i,t-1} + \zeta_{i,t}, \tag{2.8}$$

where $\zeta_{i,t}$ is i. i. d. normal with mean 0 and variance σ_ζ^2.

The Job-Match Component of Wage. A key element of the Low et al. (2010) model is its job-match process. The consumer-firm job-match component, a_{i,j,t_0}, is drawn from a normal distribution with mean 0 and variance σ_a^2. It is indexed by the period t_0 in which the consumer joined firm j, and not by t, since it is constant throughout the duration of the consumer-firm interaction.

Job Arrival. In each period, the probability of job destruction is δ, the probability that an employed worker receives a job offer is $(1-\delta)\lambda^e$, and for an unemployed worker it is λ^n. Agents receive job offers with varying job matches. By construction, they accept all offers with a higher job match and reject all offers with a lower job match.

The job match component, in combination with the processes of job destruction and job generation, is at the core of the "income scarring" result of Low et al. (2010). While

[18] While $x'_{i,t}$ includes a larger set of control variables in the empirical portion of Low et al. (2010), only age and age squared are used to fit a general lifetime income profile to the model.

employed, people successively trade up for jobs that have a better match. They thus gain higher incomes over their life-cycle. In turn, if they experience job destruction, they lose their job match and must (re-)start getting better and better job offers. Hence, agents typically earn a lower income after an unemployment spell. Low et al. (2010) refer to this effect as "income scarring" as job destruction leads to a long-lasting reduction in earnings. Note that, by accounting for income scarring, we impose a high bar on our hypothesis as we have to demonstrate that experienced-based learners reduce their consumption beyond the reduction due to "income scarring."

Belief Formation. We consider two types of consumers, standard rational agents and experience-based learners. Both types know the model but differ in their belief about the probability of job destruction, δ. We denote the believed probability of job destruction for consumer i in time t as $\delta^b_{i,t}$.

Rational consumers hold a constant belief about δ during their lifetime. They can be viewed as Bayesian learners who have used all available data on unemployment rates to update their belief. If they have lived long enough, they know (or closely approximate) the true value of δ. Thus for rational consumers $\delta^b_{i,t} = \delta \; \forall t$.

Experience-based learners, instead, form their belief $\delta^b_{i,t}$ at time t based on the history of realizations in their lives prior to time t. Applying the specification of experienced-learning in (2.1), with weighting scheme (2.2), we obtain

$$\delta^b_{i,t} = \sum_{k=1}^{t-1} w\left(\lambda, t, k\right) P_{i,t} D_{i,t-k}, \tag{2.9}$$

where $D_{i,t}$ is an indicator of i experiencing job destruction in t, and

$$w(\lambda, t, k) = \frac{\left(t-k\right)^{\lambda}}{\sum_{k=0}^{t-1} P_{i,t}\left(t-k\right)^{\lambda}}. \tag{2.10}$$

is the weight assigned to realizations D at k periods before period t.

Model Estimates on Experience Effects in Consumption. We show the simulated consumption-saving decisions for both rational and behavioral consumers. Table 2.5 reports the key parameter values we use to simulate the model.[19] We choose values identical to those in Low et al. (2010) whenever possible. As in Low et al. (2010), we distinguish between high- and low-education individuals by varying the corresponding parameters. Since we aim to distinguish the role of financial constraints from other factors affecting consumption, we focus on the low-education group in the main text. (The result for the high-education subgroup are in B.2.)

We show several plots of the resulting consumption paths for both rational and experience-based learners in B.2. In particular, we separate consumers who were

[19] The full list of parameters is in Appendix-Table B.15.

Table 2.5: **Key Simulation Parameters**

Parameter		Benchmark value(s)	
Preference parameters			
Relative risk aversion coefficient	ρ	1.5	
Interest rate	r	1.5%	
Discount factor	β	$1/(1+r)$	
Lifetime parameters			
Working years		40	
Retirement years		10	
Income process		High education	Low education
Standard deviation of job matches	σ_a	0.226	0.229
Standard deviation of permanent shocks	σ_ζ	0.095	0.106

"lucky" and "unlucky" early in life, in terms of their earnings in Figures B.2 and B.3. The graphs illustrate the corresponding over- and underconsumption of experience-based learners during their early lifetime, relative to rational consumer behavior, as well as the need to then curtail consumption later in the first case (good experiences). In the second case (bad experiences), the graph shows excess wealth build-up among experience-based learners with early lifetime unemployment experiences—the empirical relationship we found in Section 2.3.

Using the simulated values, we estimate the relationship between consumers' unemployment experience and consumption behavior, controlling for income and wealth. The corresponding OLS regressions are in Table 2.6, where columns (1) and (2) for rational consumers and columns (3) and (4) for experience-based learners.[20] In both cases, we include a measure of prior lifetime unemployment experiences. In the case of the rational agent, prior experiences do not actually enter the belief formation, and the purpose of including the experience measure is to guide our intuition about possible confounds affecting the significantly negative effect we have seen in the PSID data. Specifically, as we are concerned that it might capture unobserved wealth effects, we estimate one model where we do not include wealth as a control (column 1) and one where we include wealth (column 2), in both cases in addition to the experience-effect proxy.

"Income scarring" with $\lambda = 1$. We first conduct the simulation using linearly declining weights $\lambda = 1$ for prior experiences. As shown in the top panel of Table 2.6, we find that income has the expected positive sign and significance level, as does wealth when it is included. However, higher unemployment experiences predict *higher* consumption. This is the opposite of what we find empirically, and is a first step towards ameliorating concerns about confounds. However, the positive sign also seems to contradict the basic intuition of "income scarring," namely, that an unexpected job

[20] We present the result for the high education subgroup in Appendix-Table B.16.

Table 2.6: **Estimations with Model-Simulated Data**

	(1) Rational	(2) Rational	(3) EBL	(4) EBL
$\lambda = 1$:				
Income	0.674	0.488	0.691	0.514
	(276.11)	(64.00)	(446.39)	(54.36)
Wealth		0.0194		0.021
		(46.92)		(43.26)
Unemployment Experience	601.5	1126.1	-1582.0	-2212.8
	(3.27)	(3.40)	(-9.31)	(-8.80)
$\lambda = 0$:				
Income	0.674	0.485	0.688	0.511
	(280.26)	(63.24)	(336.73)	(56.39)
Wealth		0.0194		0.0188
		(45.92)		(40.33)
Unemployment Experience	-176.6	965.8	-1951.4	-2768.8
	(-1.46)	(3.16)	(-8.18)	(-7.88)

Notes. Estimations with the simulated consumption values as the dependent variable and the simulated same period income and same period wealth values as the regressors for rational consumers in columns (1) and (2), and experienced-based learning (EBL) consumers in columns (3) and (4). Estimations are for the low education subgroup with $\lambda = 1$ in the top panel and the same subgroup with $\lambda = 0$ in the bottom panel. Rational consumers hold a constant belief about the probability of being employed next period, and EBL consumers form beliefs based on their employment history in their lifetime as specified in equations (2.1) and (2.2). All estimations include period fixed effects and use period clustered standard errors. Simulations are based on the working periods of 10,000 simulated consumers and thus 1,600,000 observations. t statistics in parentheses.

destruction should lead to lower lifetime income and thus lower consumption. To understand this result, consider two consumers, A and B, with the same income. Consumer A has not had any unexpected job destruction, while consumer B has experienced job loss in the past. All else held equal, "income scarring" would predict that B has a lower income. However, A and B have the same income, suggesting that B's wage is driven by his permanent-income component rather than his job-match component. As a result, person B is less worried about unexpected job destruction and thus rationally consumes more. In other words, if one introduces a proxy for experience effects into a world with rational agents, it can act as a proxy for wages being driven by the permanent-income component and generate the opposite sign. Under this scenario, then, there is little worry about confounding experience effects with traditional determinants of lower consumption, including (unobserved) wealth effects and income scarring, as long as we control for current income.

"Income scarring" with $\lambda = 0$. To recover the intuition of "income scarring" and generate a negative relationship between between the experience measure and consumption, we need to look at a case where unemployment experience is more backward looking. In our model, this is amounts to lowering λ. In the bottom panel of Table 2.6, we repeat the estimation setting $\lambda = 0$, so all prior experiences get equal weight. In this case, the specification without wealth control (column 1) shows a negative correlation between unemployment experience and consumption. Here we see “income scarring” at work: If two people have the same income today, but one person got fired more in the past, then that person likely has earned less in the past, thus has lower assets today, and consumes less today. To support this interpretation, we also estimate the regression with the wealth control in column (2) and estimate a positive coefficient on unemployment experiences, with all coefficients being similar to the ones estimated in Table 2.6 for the rational case. In other words, under this specification of experience effects the potential wealth confound materializes: If we do not control for wealth, the experience-effect proxy might pick up those effects.

We then alter the belief-formation process in the simulation to experience-based learning, and re-estimate the relationship between unemployment experience and consumption, again both without and with wealth control. The results are in columns (3) and (4) of Table 2.6. Under experienced-based learning, the coefficient estimates on the experience variable are negative and highly significant in all cases. That is, lifetime experiences appear to strongly affect the consumption behavior of experience-based learners, even after taking into account their income and wealth.

Table 2.7: **Estimations with Model-Simulated Data, Unemployment Scarring**

	(1) $\lambda = 1$ Rational	(2) $\lambda = 1$ Rational	(3) $\lambda = 0$ Rational	(4) $\lambda = 0$ Rational
Income	0.611	0.399	0.609	0.393
	(66.28)	(157.20)	(66.07)	(134.03)
Wealth		0.021		0.022
		(66.08)		(62.11)
Unemployment Experience	324.1	936.2	-909.4	676.6
	(3.47)	(5.50)	(-3.79)	(4.54)

Notes. t statistics in parentheses. Estimations with the simulated consumption values as the dependent variable and the simulated same period income and same period wealth values as the regressors for rational consumers. Experience calculated with $\lambda = 1$ in columns (1) and (2) and $\lambda = 0$ in columns (3) and (4). Estimations are for the low education subgroup with unemployment scarring. Rational consumers hold a constant belief about the probability of being employed next period, and EBL consumers form beliefs based on their employment history in their lifetime as specified in equations (2.1) and (2.2). All estimations include period fixed effects and use period clustered standard errors. Simulations are based on the working periods of 10,000 simulated consumers and thus 1,600,000 observations.

“Employment scarring”. So far our simulate-and-estimate exercise alleviates several concerns about alternative interpretations of our empirical experience-effect coefficient estimates: When consumers are Bayesian rather than experience-based learners, financial constraints and income scarring both fail to generate a negative relation between the past unemployment experiences and consumption, and under the standard proxy for experience effects, this is true even without the inclusion of proper wealth controls. As a last step, we consider an even higher hurdle in terms of alternative explanations for long-term scarring, and introduce an additional negative correlation between unemployment and future income in the model. Our motivation for introducing such "unemployment scarring" comes from a growing literature on the persistent negative effect of being unemployed on future income, especially during a recession (Davis and Von Wachter (2011), Huckfeldt (2016), Jarosch (2015)). While those findings might in fact be evidence for experience effects, the existing literature proposes alternative, more traditional interpretations. The model of Low et al. (2010) already shares many parallels with this literature. For example, their "job-match component" of wages resembles the job-security component in the model of "unemployment scars" in Jarosch (2015), [21] albeit with the difference that the wage gains lost due to "income scarring" can be regained by working for an extended period. We now introduce additional "unemployment scarring" that increases the the negative effect of job destruction on income in a permanent manner: Every time a consumer experiences job destruction, their permanent component of the wage decreases by σ_ζ, the average size of a permanent income shock.

We re-simulate the model with the additional unemployment scarring effect, and then re-estimate columns (1) and (2) of Tables 2.6, i. e., analyze again which effects the experience proxy might pick up in an estimation using data simulated with rational agents.[22] Table 2.7 shows the results. We find that the signs of all coefficients remain the same, while the size of the coefficients becomes (mechanically) lower. In the specification of Table 2.6 with $\lambda = 1$, the coefficient on the experience-effect proxy remains significantly positive. Intuitively, it still acts as an indirect proxy for a high permanent component. Observing two people A and B with the same income today, even though only B has experienced unemployment, still suggests that B has a higher permanent component since B has lost the previous job-match (income scarring). However, the distribution of the permanent component, conditional on having unemployment experience, will be shifted down by one standard deviation (unemployment scarring). In other words, in the baseline model without unemployment scarring, the comparison person B, who is earning the same income as A, would have earned lower income prior to the job loss. In other words, while unemployment experience still acts as a proxy for the permanent component it now does so for a subgroup with a systematically lower permanent component than in the baseline model which, mechanically produces

[21] See the θ_y component of the firm-type vector described in Section 2.1 of Jarosch (2015).

[22] We replicate our baseline result from Table 2.6 for the high education subgroup with unemployment scarring in Appendix-Table B.17.

a lower coefficient on the unemployment experience variable. In the specification mirroring Table 2.6 with $\lambda = 0$, shown in columns (3) and (4), we also recover previous estimates in terms of sign and also observe both numbers being lowered.

Overall, these results provide evidence that our predictions are robust even when future income and unemployment and strongly negatively correlated by both "income scarring" and "unemployment scarring." Financial constraints, unobserved wealth factors, income scarring, and unemployment scarring fail to generate a negative relation between the standard proxy for past unemployment experiences and consumption. We have also noted that the same holds even when varying the unemployment-experience proxy as long as we control for wealth effects. However, if we fail to appropriately control for wealth effects the confound might materialize. A second conclusion, then, is that it will be important to conduct exhaustive robustness checks with a variety of alternative wealth specifications – including varying proxies for liquid versus illiquid wealth, higher-order terms, decline dummies; separate dummies for housing wealth or for positive wealth versus debt, and for completeness a similar battery of variations of the income controls.

We will use the guidance these estimates help to further disentangle the role of experience-based learning from two potentially confounding factors, wealth and income, and generate additional predictions of the experience-effect model that are not generated by possible alternative interpretations.

2.5 Robustness using PSID, CEX, and Nielsen

Guided by the model-based results, we conduct a broad range of robustness checks and replications in this section, starting from reestimating the consumption model using a battery of alternative and additional wealth, income, and liquidity controls using the PSID data, and then turning to the CEX and the Nielsen data.

PSID: Wealth, Income, and Liquidity

We start from the remaining concerns about unobserved wealth, income, or liquidity components: Could imperfect measurement of individual wealth affect our estimates of experience effects? Our simulate-and-estimate exercise in Section 2.4 has indicated that, even in the presence of such mismeasurement, a confound might not be likely given our empirical specification of experience effects and given the controls for unemployment status and income. Nevertheless, we use a battery of alternative wealth measures, which we include in addition to the first- and second-order liquid- and illiquid-wealth controls that are already included in Table 2.2: (1) third and fourth order controls of (log) illiquid and illiquid wealth, (2) wealth decile dummies, separately for liquid and illiquid wealth, (3) log home equity value (home price minus mortgage) and log non-housing wealth, and (4) log total debt and log positive wealth separately. The left panel of Figure 2.5 plots the key coefficients on the macro and person experience measures, and all the detailed results are shown in Appendix-Table B.9. We find that under each of these alternative specifications all coefficients of interest remain very similar, both in terms of economic magnitude and in terms of statistical significance.

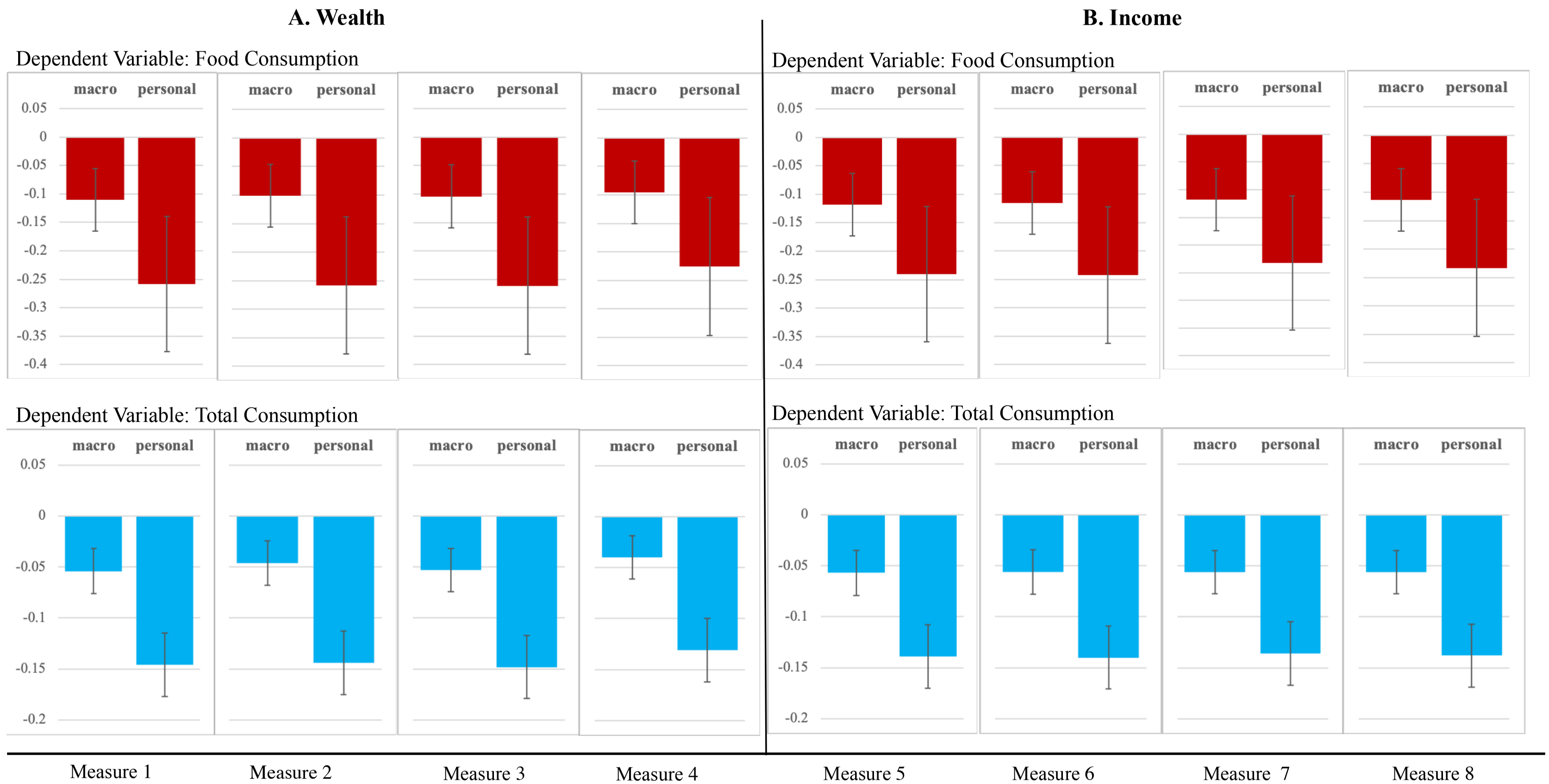
Another related concern is measurement error in the PSID income variable. As with wealth, we re-estimate our empirical model using varying constructs of income measures: (1) third and fourth order of (log) income and lagged income, (2) quintile dummies of income and lagged income, (3) decile dummies of income and lagged income, and (4) controls the bottom 2, $2^{nd}-4^{th}$, $4^{th}-6^{th}$, $6^{th}-8^{th}$, $8^{th}-10^{th}$, $90^{th}-92^{nd}$, $92^{nd}-94^{th}$, $94^{th}-96^{th}$, $96^{th}-98^{th}$, and top 2 percentile dummies of income and lagged income. The estimated coefficients of interest, shown in the right panel of Figure 2.5, are also similar in terms of both magnitude and significance. All results are shown in Appendix-Table B.10.

A more specific concern regards the role of liquidity. Even if our results are robust to various wealth measures, might the result on the impact of household unemployment experience on consumption still be confounded with the presence of (unmeasured) liquidity constraints? Our separate controls for liquid and illiquid wealth, both in the baseline estimations in Table 2.2 and in columns (2) and (6) of Appendix-Table B.9, ameliorate these concerns. As a further step, we test whether the consumption of households that are disproportionately likely to be liquidity constrained, as proxied by their low liquid-assets position, are more affected by their unemployment experience. Closely following the practice in the consumption literature, such as Johnson et al. (2006) and Parker et al. (2013), we sort households year by year into two groups based on whether their liquid wealth lies above or below the median liquid-wealth level in the sample. We then construct an indicator variable that takes the value 1 if a household's wealth position falls into the below-median group. Expanding equation (2.3), we interact the low-liquidity indicator and the experience variables. As shown in Appendix-Table B.11, households in the bottom half of the liquid-wealth group tend to spend less relative to households in the top half on average. However, their consumption expenditure does not exhibit a significantly stronger reaction to unemployment experience. All coefficient estimates are either insignificant or point in the opposite direction. This suggests that the negative effect of households' unemployment experiences on consumption is not explained by liquidity constraints.

CEX

Next, we turn to a second source of consumption data, the Consumer Expenditure Survey (CEX). So far, we have estimated strong experience effects both on food and total consumption in the PSID data. We now enlarge the set of consumption items further to include durable consumption as well as the CEX measure of total consumption, which has been widely used in the literature and which encompasses further categories

Figure 2.5: **Wealth and Income Controls: Coefficients and Confidence Intervals for Experience Measures**



Notes. The left upper four graphs (red bars) present the estimates when we re-estimate our empirical model with food consumption as the dependent variable and four alternative wealth controls. Column (1) controls for third- and fourth-order liquid and illiquid wealth. Column (2) includes decile dummies of liquid wealth and illiquid wealth. Column (3) controls for housing wealth and other wealth (total wealth minus housing wealth). Column (4) controls for positive wealth and debt. All wealth controls are in addition to the controls of first and second order of liquid and illiquid wealth. The lower four graphs (blue bars) show the estimates when we use total consumption as the dependent variable and the four wealth controls. The right upper four graphs (red bars) present the estimates when we re-estimate our empirical model with food consumption as the dependent variable and four alternative income controls. Column (1) controls for third and fourth order of income and lagged income. Column (2) includes quintile dummies of income and lagged income. Column (3) includes decile dummies of income and lagged income. Column (4) includes separately for the bottom 2, $2^{nd}-4^{th}$, $4^{th}-6^{th}$, $6^{th}-8^{th}$, $8^{th}-10^{th}$, $90^{th}-92^{nd}$, $92^{nd}-94^{th}$, $94^{th}-96^{th}$, $96^{th}-98^{th}$, and top 2 percentile dummies of income and lagged income. All income controls are in addition to the controls of first and second order of income and lagged income. The lower four graphs (blue bars) show the estimates when we use total consumption as the dependent variable and the four income controls.All regressions include household fixed effects. All the confidence intervals are at 90% confidence level.

of expenditures, in addition to durables and non-durable items, including healthcare and education expenses. [23]

The CEX is a repeated cross-sectional survey that contains household spending data across a comprehensive list of product categories at the quarterly frequency and is considered to be the benchmark dataset in the consumption literature. Compared to the PSID, its two main disadvantages are the lack of panel structure as the ability to study experience effects within households, i. e., after including household fixed effects, is one of the advances in this paper over prior studies of experience effects in different contexts, and the lack of wealth information.

As in the analysis of the PSID data, we link the measures of consumption in the CEX data to measures of households' lifetime unemployment experiences. As before we construct lifetime experiences as the weighted average of experienced unemployment outcomes since birth, using linearly declining weights. Note that we are not able to construct the same type of macro and personal unemployment experience measures as in the PSID because CEX provides neither information on where households resided prior to the sample period nor on their prior employment status. The data limitations necessitate that we re-construct the macro-level experience measure based on national unemployment rates (rather than state-level unemployment rates for the more recent years) at the quarterly frequency.

The top panel of Table 2.8 provides summary statistics for the CEX sample, including age, income profile, characteristics of the households, and consumption expenditure. The average income of the sample, \$47k, is in line with the average income at the national level. The sample period runs from 1980 to 2012. Note durable consumption and non-durable consumption do not add up to total consumption because total consumption encompasses categories of expenditure that are not considered durable or non-durable, including healthcare and education expenses. The average non-durable and durable consumption spending amount to 67.9% and 20.0% of the mean total consumption expenditures, respectively. Non-durable spending and durable spending are weakly positively correlated, with durable spending being much more volatile than non-durable spending.

We re-estimate the sensitivity of consumption to experienced unemployment conditions in the CEX data, using an estimation model that closely mirrors the PSID model from equation (2.3) but accounts for the limitations of the CEX data. Table 2.9 shows results. In columns (1) - (3) we use total, durable, and non-durable consumption as the outcome variable, respectively.

The results strongly confirm our prior findings and reveal new quantitative implications for the different components of total consumption. All experience effect coef-

[23] Note that estimations involving durable consumption may be partly affected by the timing of household durable purchases. Prior research such as Bar-Ilan and Blinder (1992) and Berger and Vavra (2015) shows that durable purchases tend to be discontinuous and go down during recessions. However, these concerns do not apply to our estimates of experience effects on food and other non-durable consumption items.

Table 2.8: **Summary Statistics (CEX and Nielsen)**

Variable	Mean	SD	p10	p50	p90	N
CEX (Quarterly)						
Age of male head of HH	51	17	29	49	75	417,607
Income	47,220	48,925	8,634	33,728	100,000	417,607
Household size	2.7	1.5	1	2	5	417,607
Total expenditure	6,116	6,145	1,902	4,490	11,479	417,607
Non-durable expenditure	4,152	3,189	1,537	3,452	7,364	417,607
Durable expenditure	1,226	4,082	0	170	2,085	417,607
Experience (Macro)	6.1	0.3	5.8	6.0	6.5	417,607
Nielsen (Monthly)						
Age of male head of HH	50	12	33	49	67	3,171,833
Income	\$50-\$60k		\$20-\$25k	\$50-\$60k	\$100k+	3,171,833
Household size	2.8	1.5	1	2	5	3,171,833
Total expenditure	714	537	205	586	1,366	3,171,833
Coupon use	0.03	0.05	0	0.01	0.09	3,171,833
Product ranking	0.47	0.11	0.34	0.47	0.61	3,171,833
Purchase of sale items	0.24	0.24	0	0.17	0.62	3,171,833
Experience (Macro)	6.0	0.2	5.8	5.9	6.3	3,171,833

Notes. The top panel consists of summary statistics on data from the CEX sample. The sample period runs quarterly from 1980 to 2012. Observations are quarterly and not weighted. The bottom panel consists of summary statistics on data from the Nielsen sample. Coupon use is the value of coupons used divided by total expenditures. Product ranking ranges from 0 to 1 based on the unit price of a good within its product module and market in the given month, where a lower value represents a lower-priced good. Purchase of sale items is the number of sale items divided by the total number of items bought. Experience (Macro) is the household's lifetime experience of national unemployment rates. Nielsen reports income in 13 brackets. The sample period runs monthly from 2004 to 2013.

ficients are negative and typically highly significant. In other words, households who have experienced worse unemployment conditions during their lifetime spend significantly less in total (on all goods), and also specifically on durable and on non-durable items. The coefficients are stable across specifications, and the economic magnitudes are large: a one standard deviation increase in lifetime unemployment experience is associated with a \$432 decline in annual non-durable consumption and \$564 decline in annual total consumption (using the estimates of columns 3 and 1 respectively). The estimate on non-durable consumption is larger than the estimate from the PSID as the earlier set of results shows that a one standard deviation increase in lifetime experience is associated with a \$276 decline in annual food consumption, while the estimate on total consumption is smaller than the one from the PSID (\$912 decline in annual total consumption). This may be attributed to the fact that non-durable consumption in the CEX encompasses leisure expenses, which tend to be elastic, while total consumption in the CEX encompasses healthcare and education expenses, which tend to be more

Table 2.9: **Experience Effects and Quarterly Consumption (CEX)**

	Total	Durables	Non-durable
Experience (Macro)	-0.077***	-0.085***	-0.086***
	(0.010)	(0.027)	(0.005)
Income control	Yes	Yes	Yes
Household controls	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes
Observations	417,607	417,607	417,607
R-squared	0.390	0.126	0.409

Notes. Pooled regressions with (log) total consumption expenditure, durable consumption, and non-durable consumption as the dependent variables. "Experience (Macro)" is the macroeconomic experience measure of unemployment, constructed as a lifetime linearly-declining weighted national unemployment rate experienced by households. Household characteristics include unemployment status, household size, education, and race. Time fixed effects include year-quarter fixed effects. Region fixed effects include dummies for the Northeast, Midwest, South, and West region. Regressions are weighted by household sampling weights from CEX. The sample period runs from 1980 to 2012. Robust standard errors (in parentheses) are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

inelastic. The new estimate for durable consumption is also highly significant. A one standard deviation increase in lifetime unemployment experience is associated with a $120 decline in annual durable consumption.

Nielsen

As a final source of data on consumption choices, we turn to the Nielsen Homescan Dataset to test the experience-effect hypothesis. The Nielsen Homescan Dataset contains information on product purchases made by a panel of more than 100,000 U.S. households from 54 geographically dispersed markets, each roughly corresponding to a Metropolitan Statistical Area (MSA), over the period 2004-2013. The households in the sample provide detailed information about the products they purchase. For each product, the data reports price, quantity, date of purchase, identifier on the store from which the purchase was made, as well as product characteristics, including brand, size and packaging, at the Universal Product Code (UPC) level. The households record whether the purchase involves coupon use or sale items. When coupons were used, the households record the dollar value of the coupons. An item is defined as being on sale if the household recorded that the item purchased involved a deal from the retailer.

The products encompass categories of food and non-food grocery, health and beauty aids, and general merchandise items, summing up to approximately 3.2 million unique UPCs covering 125 general product categories. [24]

Households also report information on their demographics, including age, sex, race, education, occupation, employment status, family composition, household income, and location of residency up to the zip code level. Note that the geographic information is more precise than the state-level identification in the PSID, as it allows us to control for the local (county-level) unemployment rate U_{mt}. The information is updated annually, and the demographics of the households are representative of the population demographics at the national level. For our analysis, we drop households with heads below the age of 25 or above 75, as in the PSID sample.[25]

Our data sample consists of 3,171,833 observations of 105,061 households. The bottom panel of Table 2.8 provides the summary statistics. We note that the average consumption expenditure from Nielsen approximately corresponds to the food consumption expenditures in the PSID, which cross-validates the quality of the data sets as the Nielsen data cover mostly food products. We also conduct a robustness analysis that keeps the advantages of panel structure of Nielsen but also exploit the comprehensiveness of the CEX by match the two datasets. Specifically, we create a synthetic panel using the Nielsen and CEX data and study the relation between past experiences and consumption. See Appendix-Section B.1 for more details.

The high-frequency nature of the Nielsen data allows us to construct more precise experience measures than the PSID, which vary at monthly frequency. However, we are not able to construct the same type of macro and personal unemployment experience measures as in the PSID because, as the CEX, Nielsen provides neither information on where households resided prior to the sample period nor on their prior employment status. We thus re-construct the macro-level experience measure based on national unemployment rates. For the personal experience measure, we can, at best, construct a variable that accounts for unemployment experiences since the beginning of the Nielsen data set. Such a measure is necessarily biased, as it is less precise at the beginning of the sample and for shorter household spells. We therefore choose to report the estimations employing only the macro-experience measure in the main text.[26]

The Nielsen data lack information about consumers' wealth, which is an important component of consumption analyses. Our prior estimations using the PSID data allow

[24] Several studies have examined the quality of the data. For example, Einav et al. (2010) compare the self-reported data in the Nielsen Homescan data with data from cash registers. They conclude that the reporting error is of similar magnitude to that found in commonly used economic data sets.

[25] As in the PSID data, we also conduct the analysis on a subsample that excludes households over the age of 65 (retirees) whose expectation of their future income should be immune to beliefs about future economic fluctuations. The results from both sets of regressions are similar.

[26] We have re-estimated our model using such a proxy for personal unemployment experience, constructed as a binary variable that takes the value 1 at time t if the head of household has ever been unemployed since the beginning of the sample period up to time $t-1$, and 0 otherwise. The results on our main coefficient of interest remain similar.

us to gauge potential biases (and alleviate such concern) to some extent, given the comparable consumption outcome variables across the two data sets. To further address the issue of the missing wealth control, we follow recent advancements in the literature, such as Stroebel and Vavra (2017) and Dube et al. (2018), and use ZIP-code level house prices as a measure of housing wealth. According to these studies, consumption responds strongly to house price movements, suggesting an important role for housing wealth in consumption dynamics (see, e. g., Mian et al. (2013a), Stroebel and Vavra (2017), and Berger and Vavra (2015)). Empirical analyses can exploit this insight since better measures of housing prices have now become available. Specifically, we extract Zillow's Home Value Index at the local ZIP code level, [27], and merge the data with the Nielsen Homescan sample. The match rate lies around 75%, and the resulting data set contains almost 3.2 million observations. We include this proxy for local housing prices, as well as an indicator variable for being a homeowner and its interaction with the Home Value Index in all of our estimations to partially address the concern about the lack of direct controls for total wealth.[28]

To re-estimate the sensitivity of consumption to experienced unemployment conditions in the Nielsen data, we use an estimation model that again closely mirrors the PSID model from equation (2.3), but accounts for the additional details as well as the limitations of the Nielsen data.[29]

Table 2.10 present results from regression specification (2.11). Columns (1) and (2) show estimates from pooled OLS regressions, and columns (3) and (4) report the estimates from regressions with household fixed effects, thus controlling for time-invariant

[27] Zillow Inc. collects detailed data on individual housing values across the U.S. and constructs ZIP-code level indices on a monthly bases, using the median value for a ZIP code. The calculations use Zillow's estimates of housing values ("Zestimates"), which aims to provide a realistic market value given the size, rooms, and other known attributes of the house, recent appraisals, geographic location, and general market conditions. (The exact formula is proprietary.) More details about the data and the quality of Zillow coverage across the U.S. are provided in Dube et al. (2018).

[28] We also conduct the analysis without including the set of wealth controls in the regressions, and the coefficient on unemployment experience remains significant and of very similar magnitude.

[29] Specially, we estimate the following

$$C_{it} = \alpha + \beta UE_{it} + \kappa U_{mt} + \gamma^{'} x_{it} + \eta_t + \varsigma_m + \upsilon_i + \varepsilon_{it}. \tag{2.11}$$

C_{it} represents the measures of consumption and UE_{it} denotes the lifetime (macro) experience of unemployment rates. U_{mt} is the current county-level unemployment rates; x_{it} is a vector of control variables including income controls, wealth controls, household characteristics (unemployment status, household size, education, race, and a dummy variable indicating whether the respondent is unemployed at the time of the survey), and age dummies; η_t are time (year-month) dummies; ς_m are local-market dummies. (Local markets denote the Nielsen designated market areas (DMAs). They are slightly bigger than county but smaller than MSA. We control for location at the local market level instead of county level because people may travel outside of counties to purchase goods. The results are similar if we use county fixed effects instead.); and υ_i are household dummies. The standard errors are clustered at the cohort level.All regression results are quantitatively and qualitatively similar when clustered by household, household-time, cohort-time, or two-way clustered at the cohort and time level.

unobserved heterogeneity at the household level. We find that, exactly as in the PSID data, households who have experienced higher unemployment conditions during their lifetimes so far spend significantly less, controlling for contemporaneous macro conditions, local market conditions, and a range of household controls including income, age, and employment status. The economic magnitude is significant: A one standard deviation increase in lifetime experience of unemployment is associated with a \$708 decline in annual consumption of non-durable goods, which amounts to around 8% of average spending for the households in our sample. When we introduce household fixed effects, the estimated experience effects become smaller, as expected given the differencing out of the cross-sectional differences in consumption between households with "mostly good" versus "mostly bad" lifetime experiences. With household fixed effects, a one standard deviation increase in lifetime experience of unemployment is associated with a \$300 decline in annual consumption of non-durable goods, comparable to the estimates from regressions using the PSID.

In Figure 2.6, we illustrate the economic magnitude of the estimates in the context of unemployment conditions during the Great Recession, which falls in the Nielsen sample period. The average monthly unemployment rate from 2008-2012 was 8.1%, with the maximum during the period being 10%. Comparing these numbers with historical averages, the average unemployment rate during the 60 years prior to 2008, from 1947-2007, was 5.6%. Now consider two individuals, a 25-year-old and a 60-year-old as of December 2007. Their lifetime unemployment experience, based on our experience weighting scheme, was 5.3% and 5.8%, respectively, when they entered the crisis in 2008. By the end of 2012, their lifetime unemployment experience was 6.3% vs. 6.1%, respectively. In other words, the unemployment experience for the 25-year-old increased by 1%, whereas that for the 60-year-old increased by 0.3%. Relating these experiences to consumption behavior, our model estimates imply that the monthly consumption expenditure of the 25-year-old decreased by approximately 18% while that of the 60-year-old decreased by approximately 5%.

2.6 Further Implications and Discussions of Experience Effects

Building on the robust results on the relation between past experiences and consumption, we further study two implications of experience effects and discuss an alternative channel, besides belief, that may be driving the effects.

Consumption Quality

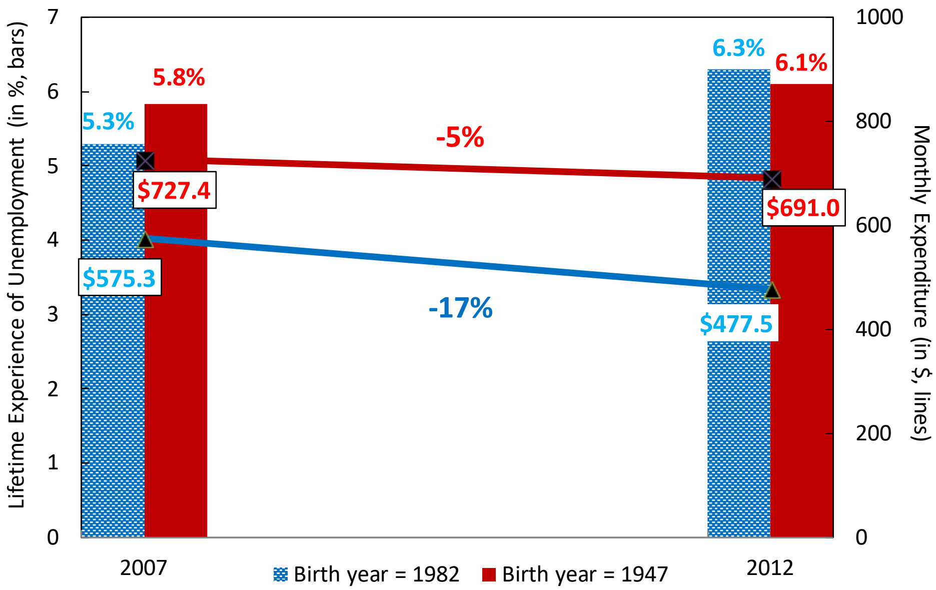
Motivated by the robust results on the link between past experiences and consumption expenditure, we further ask whether people's lifetime unemployment experiences affect the quality margins of their consumption. To that end, we make use of the rich

Table 2.10: **Experience Effects and Monthly Consumption (Nielsen)**

	(1)	(2)	(3)	(4)
Experience (Macro)	-0.415***	-0.415***	-0.178***	-0.177***
	(0.044)	(0.044)	(0.034)	(0.034)
Unemployment rate (county)		-0.002		-0.005***
		(0.003)		(0.001)
Income control	Yes	Yes	Yes	Yes
Wealth control	Yes	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes
Market-area fixed effects	Yes	Yes	Yes	Yes
Household fixed effects	No	No	Yes	Yes
Observations	3,171,833	3,171,833	3,171,833	3,171,833
R-squared	0.116	0.116	0.526	0.526

Notes. Pooled OLS and fixed effects regression with (log) total consumption expenditure as the dependent variable. Experience (Macro) is the macroeconomic experience measure of unemployment, constructed as a lifetime linearly-declining weighted national unemployment rate experienced by households. Wealth controls include the ZIP-code level house-price index from Zillow, an indicator variable for households that own at least one house, and an interaction term between the house price index and the homeowner dummy. Household characteristics include unemployment status, household size, education, race, and a dummy variable indicating whether the respondent is unemployed at the time of the survey. Time fixed effects are year-month fixed effects. Regressions are weighted using the household sampling weights from Nielsen. The sample period runs from 2004 to 2013. Robust standard errors (in parentheses) are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

Figure 2.6: **Example of Unemployment Experience Shock from Recession, Nielsen**



Notes. Example of the impact of the Great Recession on weighted lifetime experiences of unemployment rates and monthly consumption expenditure of a 25-year-old vs. a 60-year-old (as of 2007) from December 2007 to December 2012. The bars show the weighted lifetime experiences of unemployment rates based on a linearly-declining weighting scheme. The lines show the monthly expenditures: the values for 2007 are from actual data, and the values for 2012 are calculated based on model estimates.

high-frequency micro-level information on purchases and products in the Nielsen data, which capture the qualitative margins of household consumption. Specifically, we construct three monthly measures of consumption quality: (1) coupon use, normalized by total expenditures, (2) the ranking of purchased products, constructed based on their unit price within each given product module, market, and month, and normalized between 0 and 1, where lower value represents lower-priced goods, and (3) number of on-sale products purchased, normalized by the total number of products purchased. The summary statistics are shown in Table 2.8.

The estimation model we use to study the effect of lifetime unemployment experience on coupon use, the purchase of lower-end products (within a product category), and the purchase of sale items is exactly the same as delineated in equation (2.11), other than the switch in outcome variable. The estimates are shown in Table 2.11. We display only the main coefficients of interest but include the same battery of controls as in Table 2.10. We find that households who have lived through periods of worse employment conditions are more likely to use coupons, purchase lower-end products, and allocate more expenditures toward sale items. For example, considering the inter-decile range of unemployment experiences, our estimates suggest that households who have experienced unemployment rates at the 90th percentile of the sample experiences use \$13 more in coupon and purchase 8% more sale items monthly than respondents at the 10th percentile. This set of results show that people who have lived through large fluctuations in unemployment adjust the quality margins of their consumption accordingly. This suggests a thorough study on the long-term impact of macroeconomics shocks on consumption calls for analysis not only based on aggregate spending figures but also evidence on product substitution and consumption reallocation—margins that entail important welfare implications.

Heterogeneity Across Cohorts

The analyses of consumption decisions in the PSID, Nielsen, and CEX data indicate that people overweight their lifetime experiences, which naturally gives rise to heterogeneity in consumption behavior across cohorts. In particular, we see that consumers overweight more recent experiences, and the experience-effect hypothesis implies that younger cohorts do so more strongly than older cohorts. One implication of our findings, then, is that a given unemployment shock should have a stronger effect on cohorts with shorter lifetime histories so far. In other words, we predict that the young lower their consumption expenditure to a greater degree than older cohorts during economic busts and, vice-versa, increase their spending significantly more than older cohorts during booms.

We test this implication directly, regressing the log change in consumption in the Nielsen data on the interaction of age with the log change in unemployment conditions from month t to $t-1$, controlling for the same battery of controls as in Table 2.10. We do so separately for positive and negative changes (in absolute value) in unemployment rates in order to identify possible asymmetries in the reaction to improving versus tightening economic conditions. Moreover, as we focus on the most recent unemployment experience and know where a household resides during that period, we can use either changes in the national unemployment rate or changes in the local (county-level) unemployment rate as our proxy for the experienced unemployment shock, controlling for the respective other rate change.[30]

[30] Note that it would be more difficult to estimate the relationship between changes in consumption and recent changes in unemployment experience in the PSID. The low (biannual rather than monthly) frequency of survey waves makes it harder to define the "most recent" experience in a uniform way, and also drastically reduces statistical power as we have only eight waves.

Table 2.11: **Experience Effects and Monthly Consumption Quality (Nielsen)**

	(1)	(2)	(3)	(4)
A: Coupons				
Experience (macro)	0.036***	0.035***	0.005*	0.005*
	(0.005)	(0.005)	(0.003)	(0.003)
Unemployment rate (county)		0.001***		0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.040	0.041	0.690	0.690
B: Product Ranking				
Experience (macro)	-0.104***	-0.104***	0.004**	0.004**
	(0.0338)	(0.0338)	(0.002)	(0.002)
Unemployment rate (county)		-0.001**		-0.009***
		(0.001)		(0.002)
R-squared	0.083	0.083	0.680	0.680
C: On-sale Items				
Experience (macro)	0.159***	0.156***	0.009**	0.009*
	(0.018)	(0.018)	(0.004)	(0.004)
Unemployment rate (county)		0.003***		0.005***
		(0.000)		(0.001)
R-squared	0.073	0.074	0.830	0.830
Income control	Yes	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes
Market area fixed effects	Yes	Yes	Yes	Yes
Household fixed effects	No	No	Yes	Yes
Observations	3,171,833	3,171,833	3,171,833	3,171,833

Notes. OLS regressions with the ratio of coupons used over total expenditure as the dependent variable in Panel A; the (transformed) ranking of goods, based on their unit price in their specific product modules, markets, and months in Panel B (where we use the logit transformation ln(y/(1-y)) to map the original ranking, which ranges from 0 to 1, to the real line); and with the ratio of on-sale items purchased over the total number of items purchased as the dependent variable in Panel C. Experience (Macro) is the macroeconomic experience measure of unemployment, constructed as a lifetime linearly-declining weighted national unemployment rate experienced by households. Other controls are as in Table 2.10. The sample period runs from 2004 to 2013. Robust standard errors (in parentheses) are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

Table 2.12: **Age-Heterogeneity in Reaction to Unemployment Fluctuation**

	(1)	(2)	(3)	(4)	(5)
	Δln(C)	Δln(C)	Δln(C)	Δln(C)	Δln(C)
Age * Δln(National unemp-down)	-0.023***	-0.023***			-0.021***
	(0.005)	(0.005)			(0.005)
Age * Δln(National unemp-up)	-0.006***	-0.007***			-0.000
	(0.002)	(0.002)			(0.003)
Age * Δln(Local unemp-down)			-0.002*	-0.003**	-0.002
			0.00121)	(0.00135)	(0.00138)
Age * Δln(Local unemp-up)			-0.008***	-0.008***	-0.009***
			(0.001)	(0.001)	(0.001)
Local unemployment control	Yes	Yes	Yes	Yes	Yes
Income control	Yes	Yes	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Market area fixed effects	Yes	Yes	Yes	Yes	Yes
Household fixed effects	No	Yes	No	Yes	Yes
Observations	3,171,833	3,171,833	3,171,833	3,171,833	3,171,833
R-squared	0.005	0.005	0.005	0.005	0.005

Notes. OLS regression with dependent variable being the log change in monthly total consumption expenditure and the main regressors being the interaction term between age and the log change in national or local unemployment rate separated into two variables depending on whether the change is positive or negative, both from time t to $t-1$. Local unemployment controls are the log change in local unemployment rate separated into two variables depending on whether the change is positive or negative. Household characteristics include household size, education, and race. Time fixed effects include year-month fixed effects. The sample period runs monthly from 2004 to 2012. Regressions are weighted by Nielsen household weights. Robust standard errors in parentheses are clustered by cohort and time. *, **, *** denote 10%, 5%, and 1% significance, respectively.

The results are shown in Table 2.12. We interact age with the national-rate shock in columns (1) and (2), and with the local (county-level) rate shock in columns (3) and (4). We include all interactions in column (5). Note the log changes in the national unemployment rate are absorbed by the time (year-month) fixed effects, and we include the positive and negative log changes in the local unemployment rate across all specifications.

We find that unemployment shocks, whether positive or negative, have a smaller effect on expenditures as age increases, as shown by the estimated effects of the age-unemployment interaction. Both when we consider the most recent change in national unemployment rates (columns 1 and 2) and local unemployment rates (columns 3 and 4), the coefficients on the interaction between age and the most recent change in unemployment are significant and negative. The effects are a bit stronger for increases in national unemployment and for decreases in local unemployment. When we include all four interaction effects, their coefficient sizes remain similar, with the exception of the interaction of age with lower national employment, where the estimated coefficient becomes smaller and insignificant. Overall, the results support our prediction of a significantly stronger response to recent experiences among the young than among the old.

This finding also helps further distinguish the experience-effect hypothesis from alternative theories in the existing consumption literature such as liquidity constraints of the young (e.g. Zeldes (1989), Gourinchas and Parker (2002)). Models with liquidity constraints predict that the young react more strongly to negative unemployment shocks than the old, as they are more likely to hit liquidity constraints; but they do not easily predict a more positive reaction to positive shocks. To generate the latter prediction, these models need to rely on the argument that the young were previously constrained, and a positive shock allows them to adjust to their permanent-income optimum. However, our identification also exploits the differences in consumption of the young at better and worse economic times. Here, an adjustment to the PIH optimum would predict the opposite outcome relative to the experience effect hypothesis: the young with more negative prior experiences would exhibit a stronger reaction to recent good outcomes according to the PIH.[31] Thus, our findings highlight experience effects as a distinct force in affecting people's consumption behavior.

[31] To that end, we ran a set of regressions that augments the specifications from Table 2.12 with a triple interaction regressor involving age, positive and negative national or local unemployment shocks, and a dummy variable for negative experience that takes the value 1 if the respondent's unemployment experience is above the median unemployment experience for her age. The results show that for a given age, positive national and local unemployment shocks have weaker effects on the consumption of respondents with worse unemployment experiences, as predicted by experience-based learning but not by a standard PIH framework.

Preference Channel

In Section 2.3, we showed that individuals' lifetime experiences significantly predict beliefs about one's own financial situation in the future. Nevertheless, it is possible that an alternative channel, the preference channel, may also play a role in driving experience effects. It is difficult to further distinguish the relative importance of experience-based learning (beliefs channel) and the hypothesis of experience-based taste changes (preference channel). There are many possible specifications of the preference-based interpretation, and it is thus impossible to conclusively reject the instable-preferences explanation. As in the case of the beliefs-based channel, we can at best aim to provide evidence in favor of specific formalizations. Specifically, we explore one preference specification that has garnered significant support in prior empirical literature: We study whether our findings on the significant relationship between consumption and lifetime experience may be correlated with habit persistence in consumption. To that end, we estimate an alternative version of the empirical model in equation (2.11) that includes a lagged consumption measure on the right hand side.

This dynamic specification, with the lagged dependent variable included, requires a correction for the correlation between the lagged dependent variable and the fixed effects in the error term, which gives rise to "dynamic panel bias" (Nickell (1981)). To obtain unbiased and consistent coefficients, we estimate the specification using a dynamic GMM panel estimator, following Holtz-Eakin et al. (1988), Arellano and Bond (1991), Arellano and Bover (1995), and Blundell and Bond (1998b). Accordingly, both level and differenced equations are used, and the lagged dependent variable is instrumented using lagged differences for the level equation and lagged levels for the differenced equation.[32] The goodness of fit statistics for the system GMM estimators are calculated as the square of the correlation coefficients between the actual and the fitted values of the dependent variable.

In Table 2.13, we present results on whether our findings on the significant relationship between consumption and lifetime experience may be correlated with habit persistence in consumption. The estimation results indicate that they do not operate through the channel of habit formation. The results show that the effects of unemployment experience on consumption remain highly significant after taking into account consumption habit. The estimation results both confirm the robustness of experience effects and indicate that they do not operate through the channel of habit formation.

[32] Note that we test for first- and second-order autocorrelation in the first-differenced errors and find that they are first-order serially correlated, but not second-order serially correlated. This supports the validity of the moment conditions used by the system GMM estimators.

Table 2.13: **Experience Effects and Consumption, GMM regressions**

	PSID	Nielsen	CEX
Experience (macro)	-0.181***	-0.266***	-0.045***
	(0.063)	(0.051)	(0.006)
Experience (personal)	-0.635**	—	—
	(0.120)		
Income control	Yes	Yes	Yes
Wealth control	Yes	Yes	No
Household characteristics	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes
Location fixed effects	Yes	Yes	Yes
Observations	29,813	3,016,952	235,834
R-squared	0.45	0.41	0.64

Notes. System GMM regressions with food consumption (in logarithm) as the dependent variable. "Experience (Macro)" is the macroeconomic experience measure, "Experience (Personal)" is the personal experience measure, specified as described above for the respective datasets. Time fixed effects include year fixed effects for the PSID sample, year and month fixed effects for the Nielsen sample, and year and quarter fixed effects for the CEX sample. Location fixed effects include state fixed effects for the PSID sample, market area fixed effects for the Nielsen sample, and region fixed effects for the CEX sample. The sample period runs from 1999-2013 for the PSID, 2004 to 2013 for the Nielsen sample, and 1980 to 2012 for the CEX sample. Robust standard errors in parentheses are clustered on cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

2.7 Aggregate Implications and Conclusion

While it has been a decade since the start of the Great Recession, effects of the crisis still linger, and a better understanding of the long-term effects of economic shocks has proven to be of utmost importance for both academics and policy-makers. In this paper, we have put forward the idea that experiences of macroeconomic and personal unemployment shocks play a significant role in affecting household consumption and thereby serve as an important force in determining the long-term consequences of macroeconomic shocks.

Estimation results from detailed household panel data and three different data sources confirm this conclusion. Households who have experienced times of higher local and national unemployment rates and more personal unemployment spend significantly less, after controlling for income, wealth and demographics, and they tend to choose lower-quality items. We further show that beliefs about one's own financial situation in the future play a strong role in explaining the results on consumption, but such belief does not seem to be consistent with actual income and wealth changes. In fact, we see evidence of a positive relationship between past experience and future wealth build-up.

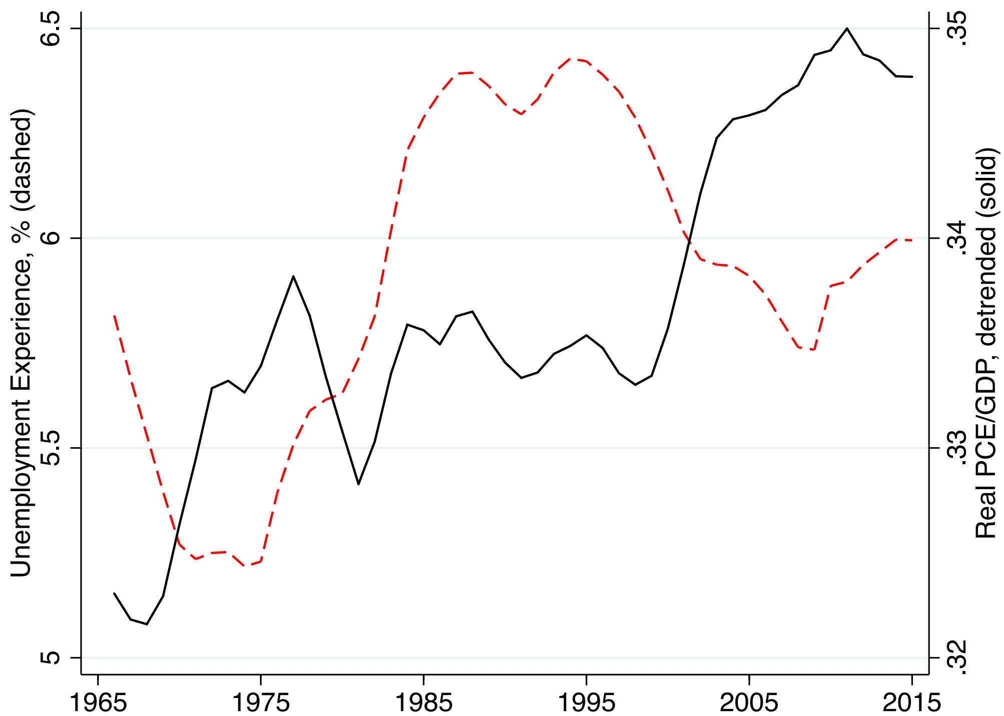
Our results on the lasting effects of past experiences on consumption suggest that experience effects could constitute a novel micro-foundation underlying fluctuations in aggregate demand and long-run effects of macroeconomic shocks. While a thorough investigation of the macroeconomic implications of experience effects is beyond the scope of this paper, we provide some suggestive evidence on the aggregate level to point to experience effects as a factor of macroeconomic significance.

Specifically, we relate an aggregate measure of lifetime experiences in the U.S. population to a measure of aggregate consumption expenditure in the U.S. from 1965 to 2013. For the former measure, we take a weighted average of national unemployment experience, as defined in Equation (2.1), using data on U.S. population broken down by age (age 25 to 75) from the Census as weights. For aggregate consumer spending, we use data on real personal consumption expenditure (PCE) from the U.S. Bureau of Economic Analysis (BEA) normalized by real gross domestic product (GDP). As shown in Figure 2.7, there exists a negative relationship between the two measures: times of higher aggregate unemployment experience coincide with times of lower aggregate consumer spending. The strong negative correlation pattern not only adds credibility to our micro-level estimates but also suggests the possibility that personally experienced labor market conditions may be a significant granular source of aggregate fluctuations.

The evidence on experience effects in consumption has potentially important policy implications. They appear to significantly dampen macroeconomic fluctuations, which in turn calls for considerations from policy-makers on optimal stabilization policy, monetary or fiscal.

For future research, our empirical methodology could be applied to a larger cross-section of countries, particularly countries that have undergone more drastic and volatile macroeconomic events such as the emerging market countries and some European countries. Such exercises would help to determine the extent to which personal experiences affect household consumption—the key ingredient in all macro and macro-finance frameworks.

Figure 2.7: **Aggregate Unemployment Experience and Consumer Spending**



Notes. Aggregate unemployment experience calculated as a weighted average of national unemployment experience, as defined in Equation 2.1, with the weights being U.S. population by age (restricted to age 25 to 75) from the Census. Aggregate consumer spending is measured as real personal consumption expenditure (PCE) from the U.S. Bureau of Economic Analysis (BEA) normalized by real gross domestic product (GDP), detrended by removing a linear time trend from the series.

Chapter 3

Capital Flows, Asset Prices, and the Real Economy: A "China Shock" in the US Real Estate Market

3.1 Introduction

With greater international financial integration, unprecedented volume and varied types of international capital have flowed across national borders over the past decades.[1] The economic consequences of these flows are at the heart of a passionate debate among researchers and policymakers: Do capital inflows stimulate the real economy? There is a wide range of views about this question. On the one hand, capital inflows are thought to augment local saving, enhance capital allocation efficiency and productivity, and thus growth in the recipient countries (Fischer 1997, Summers 2000, Harrison et al. 2004, Tong and Wei 2010). On the other hand, capital inflows can fuel booms in asset prices, which imply potential downside risks to financial stability and economic performance (Rodrik 1998, Stiglitz 2002, Aizenman and Jinjarak 2009; Gourinchas and Obstfeld 2012). The wide range of views arises in part from the difficulty of evaluating the casual impact of capital inflows on the real economy, as the timing of capital inflows is rarely exogenous, and unobserved factors may simultaneously affect capital inflows and the real economy. In this paper, we use a local economy analysis to address this empirical challenge and to provide quantitative evidence on the effects of foreign

This chapter is based on joint work with Zhimin Li and Calvin Zhang. Permission to reprint this material as a chapter of the present dissertation has been obtained.

[1] Based on data from the IMF International Financial Statistics (IFS) database, gross capital inflows increased from \$0.4 trillion in 1980 and approximately \$3.7 trillion in 2000 to \$11 trillion in 2007 — an amount equivalent to 19 percent of world GDP. Of this amount, almost 90 percent of gross capital inflows went to developed economies.

residential real estate capital inflows on local economies in the US.

The specific context of our study is related to one of the most notable economic phenomena over the past two decades—the rise of China in the global economy, which a growing literature has dubbed the "China shock". While existing literature has mostly focused on a "China shock" on the goods side[2], China's intention to play a more active role in global finance in recent years has prompted a new question: as China becomes more financially integrated into the global economy, what are the impacts of a "China shock" on the finance side for the rest of the world? Our paper is the first academic work that studies this increasingly relevant question. In this paper, we call attention to a "China shock" in the US real estate market, documenting an unprecedented surge in residential housing purchases by foreign Chinese[3] in the US since 2007. We estimate the effects of these capital inflows on the US local housing and labor markets, and shed light on the mechanism underlying the effects. We show that the surge in capital from foreign Chinese into US real estate has a large positive causal effect on local housing prices and local employment, and this effect appears to be transmitted through a housing net worth channel.

The surge in housing purchases by foreign Chinese in the US over the past decade has grabbed many headlines in the press. According to the National Association of Realtors, foreign Chinese have taken the lead among all foreign buyers of US real estate by a wide margin, as measured by both value and quantity.[4] Moreover, they tend to concentrate the purchases in specific regions, with California being the most popular destination[5]. Even though these purchases have been widely reported by the media, to the best of our knowledge, no academic study has provided a formal quantification of the phenomenon and explored its effects on the US real economy.

Motivated by the numerous reports from the mass media, we first examine whether they can be corroborated by actual housing transaction data. Using data covering all real estate transactions in the three largest core-based statistical areas (CBSA) in California from DataQuick, we find two striking patterns on purchasing behavior by foreign Chinese in the US real estate market. First, house purchases by foreign Chinese increased almost twentyfold over the 2007-2013 period, while there was no significant change in the purchasing behaviors by buyers of other ethnicities over the period. Second, the increase in house purchases by foreign Chinese has been concentrated in zip codes that are historically populated by ethnic Chinese.

---

[2] In particular, a growing strand of literature in international trade, starting with Autor et al. (2013), studies the effects of China's rising import competitions on local economies in the US.

[3] In this paper, "foreign Chinese" is used to denote Chinese who do not regularly reside in the US, and "ethnic Chinese" denotes Chinese who regularly residents the US.

[4] Foreign Chinese buyers spent $28.6 billion on residential property in the US in 2014, which is a 30% increase from the previous year and more than two and a half times the amount spent by Canadians, the next biggest group of foreign buyers of real estate in the US.

[5] A survey published by the California Association of Realtors found that Chinese bought 32% of homes sold to foreigners in California in 2014, and a recent RealtyTrac report found that 80% of new construction homes in the city of Irvine were sold to Chinese buyers that year.

Given these two new stylized facts, we proceed to study the effects of the surge in capital flows from China in the US real estate market on US local economies. Empirically establishing the causal impact from foreign Chinese capital inflow on the real economy is challenging due to an issue of endogeneity: it is difficult to distinguish whether increasing purchases by foreign Chinese are affecting the local economy or whether foreign Chinese select into neighborhoods that are more prosperous. To overcome with this issue, we devise an instrumental variable (IV) strategy which makes use the stylized fact that foreign Chinese tend to purchase homes in areas that historically have a higher share of ethnic Chinese population. We exploit the historical cross-market variation in the concentration of Chinese population across zip codes and instrument foreign Chinese housing transaction value by the aggregate housing transaction value in California weighted by the historical share of ethnic Chinese population across zip codes. This strategy is numerically equivalent to a Bartik instrument, since the underlying identifying assumption of our instrument also is in terms of cross-sectional local shares. Our empirical model also controls for county-year fixed effects and zip code level characteristics that may systematically affect local house markets and labor markets, including population, education, and pre-sample period trends in housing prices and employment.

We find that the surge in real estate purchases by foreign Chinese since 2007 have a positive and significant effect on local housing prices. Since the year 2007 was start of the Great Recession, thus our analysis also delves into the question of whether capital inflows from China into the US real estate market played a stabilizing role during economic downturns in the US.[6] Our results show that a 1% increase in housing demand by foreign Chinese, as measured by transaction value, increases home prices by 0.074% during the housing market crash period (2007-2011), which corresponds to an increase of $433 per home. To put it another way, zip codes that experienced more real estate purchases by foreign Chinese exhibit a lower decline in housing prices during the housing market crash of 2007-2011, which suggests that foreign cash inflow can have stabilizing effects during economic downturns. During the housing market recovery period (2012-2013), a 1% increase in housing demand by foreign Chinese induces a 0.102% increase increase in home prices, which corresponds to $597 per home.

This first set of results provides a quantification of the impact of foreign real estate capital inflows on asset prices. Do such inflows have an effect on the real economy? We next assess their impact on local labor markets using employment data from the Census Bureau as the dependent variable. Our results show that foreign Chinese housing purchases significantly increases local total employment, controlling for county-year

[6] Besides the housing market crash in the US, the 2007/2008 period was also the time that the real estate market in China began to boom significantly, the Yuan was increasing in value against the Dollar, and the Chinese government increased the limit on how much Chinese citizens can exchange yuan to other currencies annually (up from $20,000 to $50,000). All of these factors likely played a role in inducing the surge of housing purchases by foreign Chinese in the US. The main focus of this paper is to understand the implications of these purchases on the US economy.

fixed effects and zip code level population, education, and pre-trend in employment. A 1% increase in foreign Chinese housing transaction value induces a 0.102% increase in a zip code's total employment levels during the housing market crash years and a 0.149% increase during the recovery years.

In addition to the main empirical model, we also run a set of placebo test and regressions with alternative specifications to check the robustness of our baseline results. First, we run a placebo test to address the concern that neighborhoods which historically attracted ethnic Chinese settlement could have characteristics that systemically differentiate their local economic conditions from other neighborhoods. To that end, we test whether changes in housing purchases by foreign Chinese after 2007 are related to local employment prior to 2007. We do not find a significant relationship. Second, we use an alternative IV in the regression specification: China's gross domestic product weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. This IV is motivated by idea that the surge in foreign Chinese housing demand was demand-driven, due to changes in China's capital flow regulation, currency appreciation, and domestic real estate market boom. We find positive and significant results on the effects of foreign Chinese housing purchases on housing prices and employment from regressions using both alternative IVs.

Given the robust findings on the impact of foreign Chinese housing purchases on the local economies, we next turn to explore the mechanism underlying the results. We explore two mechanisms potentially driving the results: (i) a migration channel, which entails that foreign Chinese housing prices induce a migration inflow into the corresponding neighborhoods, and consequently push up demand for local goods and local employment; ii) a housing net worth channel, which posits that housing net worth affects employment by changing consumer demand through either a direct wealth effect or less binding borrowing constraints driven by the rise in the collateral value.

We test the migration channel by relating foreign Chinese housing transaction, CHTV, to the number of income tax returns at the zip code level, which is often considered a measure of the number of households. We do not find a positive relationship between CHTV and migration inflow. In fact, results show that foreign Chinese house purchases crowd out residents on net.

We test the housing net worth channel by estimating the effects of foreign Chinese capital inflows on local non-tradable and tradable sector employment. This is motivated by the consideration that the impact of consumption changes in an area due to housing net worth fluctuations on local employment should show up foremost in the employment size of non-tradable sectors in that area, since non-tradable sector employment depends primarily on local demand, while the tradable sector is more diversified in its geographic origins of demand. We find that real estate purchases by foreign Chinese since 2007 significantly increase employment in the local non-tradable sectors, while the effects on the tradable sections are negligible and insignificant. This result supports the housing net worth channel conjecture. The result holds even if we exclude the construction sector from non-tradable industries when constructing the non-tradable sector employment variable.

Finally, we build a simple model that incorporates the housing net worth channel to rationalize the empirical results. The model clarifies how a foreign housing demand shock affects local employment through the housing wealth channel. The key intuition from the model is that a positive housing wealth shock due to housing purchases by foreign Chinese increases the local demand for non-tradable goods and hence local non-tradable sector employment because demand for non-tradable goods are centralized in local economies. At the same time, an increase in the demand for tradable goods can be supplied by the production elsewhere, diffusing the effect on local employment in the tradable sector. Our results support this prediction as we find that foreign Chinese purchases significantly impact employment in the non-tradable sectors but not the tradable sectors.

Related Literature. This paper is related to several strands of literature and contains important policy implications. First, our paper is related to a growing literature that studies the effects of housing purchases by foreigners on local housing markets. Badarinza and Ramadorai (2015) examines the effects of housing demand by foreigners on domestic housing prices in London. Using political shocks in a source country as an exogenous instrument, they estimate the effects of foreign buyers on house prices in London neighborhoods with a large pre-existing share of residents born in that source country and find substantial price effects in such areas. Sa (2016) also studies the effect of foreign investment on UK house prices and home ownership rates, using a different data set. Cvijanovic and Spaenjers (2015) find that house purchases by non-resident foreigners push up house prices and crowd out residents in highly desirable neighborhoods of Paris. They also show empirically that relatively few properties bought by non-residents are rented out, which corresponds to reports on foreign Chinese and validates an important assumption we make in our model. Favilukis and Van Nieuwerburgh (2017) develop a spatial equilibrium model of a city with heterogeneity among residents to study the welfare implications out-of-town buyers of local housing markets. Our paper contributes to this literature by going beyond the price effects of foreign housing purchases and examines the consequences on local employment as well as the underlying mechanisms. We aim to bridge the literature in the macro-finance and urban economics by presenting empirical evidence on how a foreign shock on the finance side via real estate investments affects the real economy.

To that end, our paper is related to the line of research that explores the effects of housing investments on the real economy. Green (1997) and Parker (2000) are among the earlier works that point out a significant link between real estate investment and the macro-economy. Recent papers by Mian et al. (2013b) and Mian and Sufi (2014) argue that deterioration in household balance sheets, the housing net worth channel, played a significant role in the sharp decline in US spending and employment during the 2007-2009 financial crisis. Our paper presents results that support the housing net worth channel in the context of a positive housing net worth shock driven by foreign Chinese demand.

More broadly, our paper is related to papers that study the impact of foreign investments on domestic local economy, including papers that look into the effects of foreign direct investment on domestic economic growth (e.g.,Borensztein et al. 1998). Our analysis quantifies the effect of foreign housing investment, a specific form of capital inflow that has not been emphasized in the international finance literature, on the local economy and draws a link between international capital inflow to the housing sector and domestic economy. Moreover, given the surge in housing purchases by foreign Chinese coincided with the housing market crash in the US, our results show that investments by foreigners can play a stabilizing role in times of economic downturns. Our work also is related to papers that estimate the effects of stabilization policies such as fiscal stimulus on local economies during economic downturns, including Ramey (2011), Nakamura and Steinsson (2014), and Chodorow-Reich et al. (2012).

Our paper also contributes to the literature that aims to better understand the impacts and implications of China's increasing integration into the global economy on the rest of the world. A growing literature explores the effects of China's rapid growth in trade activities on US local economies, starting with the paper by Autor et al. (2013) who study the effects of rising Chinese import competition on US local labor markets and find that such competition explains one-quarter of the aggregate drop in US manufacturing employment. A number of subsequent papers find that Chinese import competition significantly affect innovation (Autor et al. (2017)), electoral consequences (Autor et al. (2017)), and marriage market outcomes (Autor et al. (2018)) in the US. While China's integration into the global economy indeed has been most acutely manifested in its trade activities over the past two decades, China has been seeking to open up its capital markets, which has prompted growing interests in the academic, policy and business community to better understand the implications of a "China shock" on the finance side for the rest of the world. This paper is one of the first academic papers on that front. We focus on a specific source of "China shock" on the finance side—the surge of cash inflows from China for residential real estate purchases in the US, and analyze its economic impacts on the US local economies.

The remainder of the paper is organized as follows. Section 2 describes our measure of foreign Chinese housing transaction and the data. Section 3 discusses the identification strategy and baseline results on the effects of foreign Chinese housing transaction on local housing markets and labor markets. Section 4 presents evidence on the mechanism driving the effects. Section 5 presents a model used for interpreting the empirical results. Section 6 concludes.

3.2 Housing Purchases by Foreign Chinese in the US

Measure of Chinese Housing Transaction

Data. Our main data source is the housing transaction data from DataQuick, which contains the universe of housing transaction records collected from county register of deeds and assessor offices throughout the US. For each home sale, the data include the sales price, the closing date, the precise address of the home, home characteristics, information on home financing, and names of the buyers and sellers. In our analysis, we focus on housing transactions in the three largest core-based statistical areas (CBSAs) in California (Los Angeles-Long Beach-Riverside, San Jose-San Francisco-Oakland, and San Diego-Carlsbad-San Marcos), which covers 17 counties and 912 zip codes. We restrict our sample to single family residential homes as the focus of this study is residential real estate purchases not commercial real estate purchases. The final dataset contains 1,796,669 residential housing transactions over the period 2001-2013.

Measure Construction. Given our objective is to study the impact of housing demand by foreign Chinese on local economies, we need a measure to capture housing purchases by these Chinese in the US. The key challenge in generating the measure is that county offices do not collect information on the country of origin of the buyers. To overcome this difficulty, we develop a three-step method to construct a measure of Chinese housing transaction value ($CHTV$).

First, we identify the ethnicity of the housing buyers in our sample using the ethnic name-matching technique from Kerr (2008a), Kerr (2008b), and Kerr and Lincoln (2010).[7] The technique both applies the ethnic-name database from the Melissa Data Corporation, which specializes in the identification of Asian (especially Chinese, Indian/Hindi, Japanese, Korean, Russian, and Vietnamese) ethnicities, and manually codes the unmatched names.[8] Exploiting the fact that certain names are unique or more common to one ethnicity, this technique assigns each housing buyer, based on their first and last names, a probability of belonging to a specific ethnicity, with the probabilities summed up to one. If a name is unique to one ethnicity, the buyer with that name will be assigned to the respective ethnicity with a probability of one. For names that are common among multiple ethnicities, the technique uses the demographic breakdown in the MSA in which the corresponding buyers reside for assigning

[7] We thank William Kerr for running the buyer names from our sample through his ethnic name-matching algorithms.

[8] See Kerr (2008b) and Kerr (2008a) for more comprehensive details on the names matching process and descriptive statistics from their matching exercises. Kerr (2008b) originally created the technique to identify the ethnicity of inventors who were granted patents by the US Patent and Trademark Office. Kerr and Lincoln (2010) use it to investigate the impact of H-1B Visa reforms on Chinese and Indian inventors and patents.

probabilities. For example, a person with the surname Chen would be assigned to the Chinese ethnicity group with a probability of one, while someone with the surname Lee, which could be of Chinese, Korean or American ethnicity, would be assigned to each of the three ethnic groups with probabilities based on the proportion of Chinese, Koreans, and Americans in the MSA in which the person resides. In total, nine ethnicities are distinguished by the name-matching technique: Chinese, Anglo-Saxon/English, European, Hispanic/Filipino, Indian/Hindi, Japanese, Korean, Russian, and Vietnamese. The match rate on the names from the housing transaction data sample was 97%. As the first step to constructing the Chinese housing demand measure, we keep only transactions made by ethnic Chinese buyers. In particular, we consider a housing transaction to be made by an ethnic Chinese buyer only if the name-matching technique assigns that buyer as an ethnic Chinese with a probability of one.

Second, we keep only housing transactions that are made entirely in cash by ethnic Chinese buyers for the Chinese housing transaction value measure. This step is motivated by the fact that foreigners have limited access to the US mortgage market: non-US citizens without lawful residency in the US are not eligible for Fannie Mae, Freddie Mac or FHA home loans, and it is difficult for these borrowers to finance homes through private lenders.[9] Furthermore, the National Association of Realtors reports that most non-resident foreign buyers make all-cash purchases, while a much smaller fraction of resident foreign buyers paid all-cash.[10] While this filtering step may exclude some foreign Chinese who purchased houses in the US using mortgages from US private lenders, it ensures that the $CHTV$ measure we construct is as conservative as possible and can be viewed as a lower bound on foreign Chinese housing purchases.

Third, we recognize that restricting the sample to all-cash purchases by ethnic Chinese is a necessary but not sufficient criterion for identifying foreign Chinese housing purchases since ethnic Chinese who reside in the US also can purchase homes with all-cash. To address this concern, we make use of an interesting observation from the data: ethnic Chinese who reside in the US behave similarly as Anglo-Americans in terms of making all-cash real estate purchases. More specifically, the percentage of housing transactions that are made in all-cash by ethnic Chinese and Anglo-Americans share a similar trend before 2007, as shown in Figure 3.1.[11] Given this observation, the third step we take to construct the Chinese housing demand measure is adjusting the ethnic Chinese all-cash transaction value by the share oethnicf all-cash transactions by Anglo-

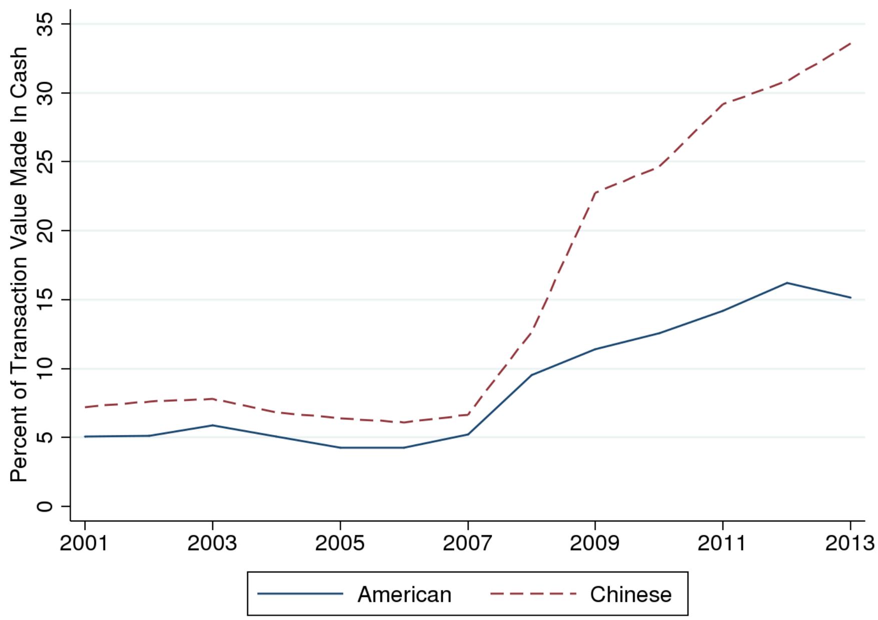
[9] Mortgages to foreigners through private lenders carry high interest rates and require borrowers make large down payments, in the range of 30 to 50 percent.

[10] According to the "Profile of International Activity in US Residential Real Estate" published by the National Association of Realtors, 72 percent of non-resident foreign buyers made all-cash purchase, while 35 percent of resident foreign buyers paid all-cash.

[11] We take note of the observation that, after 2007, the share of housing transactions made in all-cash increased much faster for ethnic Chinese relative to Americans. Given reports from the National Association of Realtors, we conjecture that this is driven by increases in all-cash purchases by foreign Chinese.

Americans for each zip code and year, as an effort to exclude purchases made by US-based ethnic Chinese from the measure. Specifically, we construct the measure $CHTV$ as follows: $CHTV_{zt}$=(ethnic Chinese all-cash transaction value)$_{zt}$*(1-Anglo-American all-cash transaction value/Anglo-American transaction value)$_{zt}$, where z denotes zip code and t denotes year.

Figure 3.1: **All-Cash Housing Purchases by Ethnic Chinese and Anglo-American**



Notes: This figure plots the percentage of all-cash housing purchases (by transaction value) by Anglo-American and ethnic Chinese between 2001 and 2013. The ethnicity assignments are made based on Kerr's ethnic name-matching technique. A buyer is considered to be an Anglo-American or ethnic Chinese if the technique assigns the corresponding ethnicity for the buyer with a probability of one. Data source: DataQuick and authors' calculations.

Two Stylized Facts

We next use our measure of Chinese housing transaction value ($CHTV$) to study housing purchasing behavior by foreign Chinese in the US. In particular, we document two new stylized facts about Chinese housing purchases in aggregate and across zip codes.

We begin by examining the share of housing transaction value by foreign Chinese and buyers of other ethnicities in the CA real estate market over the 2001-2013 period

(Figure 3.2a). The figure reveals a striking increase in house purchases by foreign Chinese over the 2007-2013 period. While the percentage of all housing transactions made by foreign Chinese was small (around 0.3%) and comparable to those of other ethnic groups over the 2001-2006 period, it began to increase sharply in 2007 and reached more than 5% of total housing transaction value in California by 2013, overtaking all other ethnic groups as the lead group of buyers in the market. That is almost a twentyfold increase in transaction values. This finding is consistent with reports from the National Association of Realtors that foreign Chinese have become the lead group of foreign buyers in the US real estate market.

Building on this observation about foreign Chinese housing purchases in aggregate, we then turn to explore the spatial distribution of these purchases. Do some zip codes attract more foreign Chinese buyers than others? We find that the the surge of housing purchases by foreign Chinese tend to be concentrated in areas that have been historically populated by ethnic Chinese. Figure 3.2b dissects Figure 3.2a by zooming in on housing purchases by foreigners in zip codes in the top quartile of the ethnic Chinese population based on the 2000 Census Bureau data. In those neighborhoods, they make up more than 10% of the total real estate transaction value by 2013. We also observe this pattern of spatial distribution when we illustrate foreign Chinese housing purchases on maps. Appendix-Figure C.1 shows that these purchases tend to be clustered in zip codes that have been predominately ethnic Chinese.[12]

Motivated by these two stylized facts, we next proceed to study the effects of housing purchases by foreign Chinese on US local economies, namely local housing market and labor markets. In particular, we exploit the second stylized fact in our empirical strategy to assess the causal relationship.

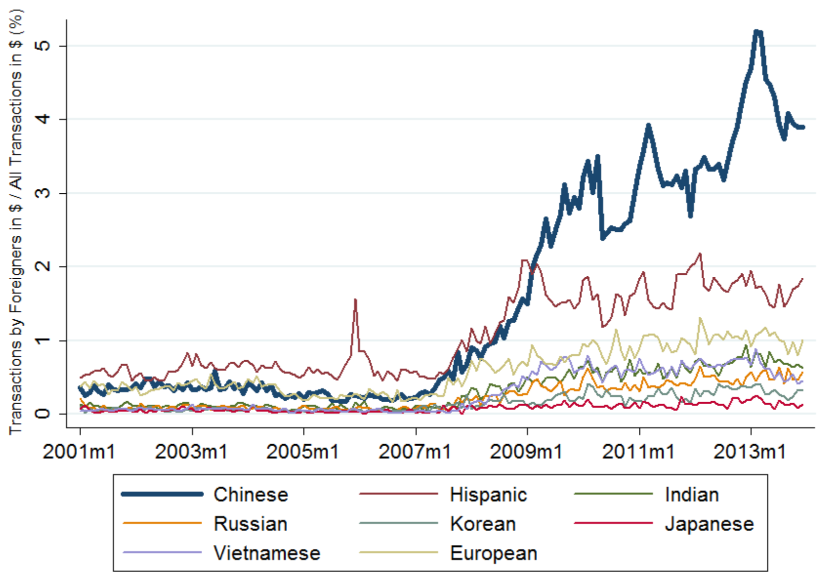
Additional Variables and Summary Statistics

To study the effects of housing purchases by foreign Chinese on local housing and labor markets, the two main sets of outcome variables are zip code level housing prices and employment. For the former, we use the Zillow Home Value Index for single family homes, which is a smoothed, seasonally adjusted measure of the median estimated home value across a given zip code. Employment data are from Zip Codes Business Patterns (ZBP) collected by the Census Bureau. It provides annual statistics for businesses with paid employees within the US at the zip code level for two- through six-digit NAICS code level. We decompose the employment measure into two categories, tradable and non-tradable, using the four-digit industry classification code following the classification scheme in Mian et al. (2013b). A industry is defined as tradable if it has imports plus exports equal to at least \$10,000 per worker, or if total exports plus imports for the NAICS 4-digit industry exceeds \$500M, while non-tradable industries include the

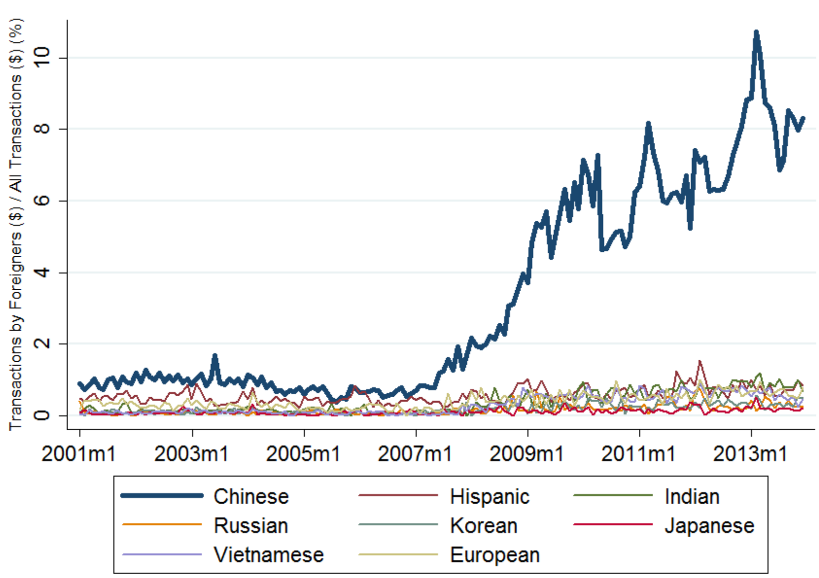
[12] A similar pattern hold when we construct Figure 3.2a and 3.2b using foreign Chinese housing transaction count, instead of value. This suggests that foreign Chinese have been purchasing residential real estate across a full spectrum of types, not only high-end real estate.

Figure 3.2: **Housing Purchases by Foreign Chinese and Buyers of Other Ethnicities**

(a) All Zip Codes



(b) Zip Codes in the Top Quartile of Ethnic Chinese Population Share



Notes: Figure (a) plots the share of housing transactions (by $ value) made by foreign Chinese and buyers of other ethnicities in the three largest CBSAs in California. Figure (b) plots the share of housing transactions (by $ value) made by foreign Chinese and buyers of other ethnicities in zip codes in the top quartile of ethnic Chinese population, based on the 2000 Census data, in the three largest CBSAs in California. The sample period runs monthly from 2001 to 2013. Data source: DataQuick and authors' calculations.

retail sector, restaurants, and industries related to construction, real estate, or land development.

In addition to the outcome variables, we collect information on zip code level population and education, measured as share of population with bachelor degrees, from the Census Bureau, to use as control variables in the empirical analysis.

Table 3.1 presents the summary statistics of our dataset. To give a sense of the aggregate figures, Column 2 shows the dollar value and count of the total housing transactions and foreign Chinese housing transactions across all zip codes in the sample over 2001-2013. Over that period in the three largest CBSAs in CA, there were almost 1.8 million residential housing transactions, amounting to approximately 909 trillion dollars. The house purchases by foreign Chinese make up 1.2% and 1.1% of the total transaction value and count, respectively. Columns 3-8 of Table 3.1 presents the means and standard deviations of the variables at the level of zip code-year. In total, the dataset contains 9,986 zip code-year observations over the sample period. Motivated by the stylized facts we found, we break down the sample into three sub-periods, the housing market boom period (2001-2006), the housing market crash period (2007-2011) and the housing market recovery period (2012-2013), and examine the relevant summary statistics. As shown in the top panel of the table, there was a dramatic increase in housing transactions by foreign Chinese during the post-2007 period relative to earlier years. On average, while each zip code witnessed 0.8 housing transactions by foreign Chinese for an average value of \$0.45 million per year between 2001-2006, these figures jumped to 11 transactions and \$5 million by 2013, respectively. The share of Chinese transaction out of all housing transactions in California in terms of count and value also increased from 0.28% to 4.5% and from 0.28% to 4.36% respectively. The bottom four rows of Table 3.1 present summary statistics on the variables capturing local economic conditions in the dataset. Interestingly, they show that the average home prices, tradable and non-tradable sector employment, and income were similar across the three sub-periods.

3.3 Foreign Chinese Housing Purchases and Local Economies

Identifying the Impact of Foreign Chinese Housing Purchases

Our objective is to estimate the impact of the surge in foreign Chinese housing transactions since 2007 on US local economies. To that end, we estimate the following equation using the foreign Chinese housing transaction value ($CHTV$) measure and data on local economic conditions:

$$ln(Y_{zt}) = \alpha + \theta ln(CHTV_{zt}) + \beta ln(CHTV_{zt}) \times \mathbb{I}\{t \geq 2007\} + \gamma X_{z,0} + \eta_{ct} + \varepsilon_{zt} \quad (3.1)$$

Table 3.1: **Summary Statistics**

	Aggregate	**Zip Code/Year (N=9,986)**					
	2001-2013	2001-2006		2007-2011		2012-2013	
	Total	Mean	SD	Mean	SD	Mean	SD
Total Housing Transaction							
Value ($)	908,614B	160.71M	173.33M	103.25M	132.29M	114.68M	158.54M
Counts	1,796,669	285.71	302.74	241.41	315.13	243.33	303.96
Foreign Chinese Housing Transaction							
Value ($)	10,601B	0.45M	1.30M	3.18M	5.49M	5.00M	8.91M
Counts	18,942	0.80	2.21	7.87	13.33	10.95	16.90
Zillow Single Family Home Price Index	-	0.54M	0.36M	0.54M	0.36M	0.54M	0.40M
Log of Non-Tradable Employment	-	7.34	1.26	7.63	0.97	7.61	1.00
Log of Tradable Employment	-	5.88	1.99	5.89	1.92	5.80	1.93
Household Income	11.94B	68,562.23	57,776.78	76,097.31	62,394.53	85,152.74	82,349.33

Notes: This table presents summary statistics of the key variables (listed in Column 1) in the dataset. Column 2 shows the variables summed across all zip codes in the sample over the period 2001-2013. Columns 3-8 show the means and standard deviations of the variables by zip code and year, broken down into the housing market boom period (2001-2006), the housing market crash period (2007-2011) and the housing market recovery period (2012-2013). Data source: DataQuick, Zillow, Census Bureau, IRS, and authors' calculations.

where Y_{zt} denotes either the Zillow Home Value Index or employment size for zip code z in year t, $CHTV_{zt}$ denotes the foreign Chinese housing transactions value measure, $\mathbb{I}\{t \geq 2007\}$ is an indicator variable that takes the value one if the year is 2007 or later and 0 otherwise, $X_{z,0}$ are zip code level controls including population and education (measured as the population share with bachelor degrees) from the pre-sample year 2000, and pre-sample period trends for the respective dependent variable calculated as the change in either Zillow Home Value Index or employment size between 1996 and 2000, and η_{ct} are county-year fixed effects.

Equation 3.1 takes the form of a difference-in-difference regression framework. Our coefficient of interest is β, which measures whether zip codes with more foreign Chinese housing purchases experienced a greater increase in housing prices and employment after 2007, the year when the sharp increase in foreign Chinese transactions began, controlling for county-year fixed effects and zip code level population, education, and pre-sample period trends in housing prices and employment. In other words, we assess the change in housing prices and employment between zip codes that experienced more foreign Chinese housing purchases and the rest within the same county and year, controlling for zip code level characteristics that may systematically affect local house markets and labor markets.

Equation 3.1 will consistently estimate the coefficient of interest if $Cov(CHTV_{zt}, \varepsilon_{zt}) = 0$. This condition is unlikely to hold, despite the inclusion of county-year fixed effects and zip code level controls, due to an issue of endogeneity: it is difficult to distinguish whether the increase in foreign Chinese purchases affected local home prices and employment or whether foreign Chinese sought to buy homes in neighborhoods that are more likely to experience higher rates of home price appreciation or more employment. To address this concern, we devise an instrumental variables (IV) strategy.

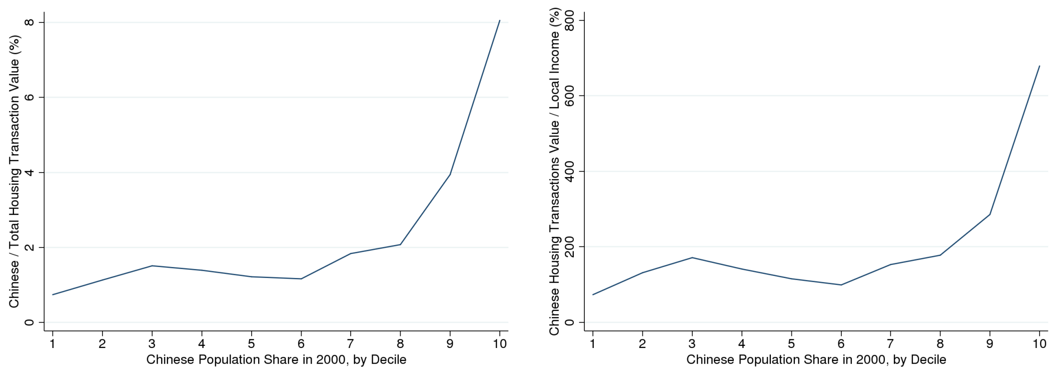
Our IV strategy exploits the stylized fact that foreign Chinese tend to purchase homes in areas that historically have a higher share of ethnic Chinese population. This observation helps to identify quasi-random variation in the spatial distribution of foreign Chinese housing purchases. Specifically, we instrument foreign Chinese housing transaction value ($CHTV$) by the aggregate housing transaction value in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period: $CHNShare_{z,0} \times TTV_t$, where $CHNShare_{z,0}$ is the ethnic Chinese population share in zip code z from the pre-sample year 2000, and TTV_t is a time-varying measure of total housing transaction values in California.

When the instrumental variable is applied to Equation 3.1, the time-varying component of the IV, total housing transaction values in California (TTV_t), is fully absorbed by the county-time fixed effects. Therefore, our IV strategy fundamentally uses cross-sectional variation in local ethnic Chinese population share to identify foreign Chinese housing demand. The identification assumption is that the pre-sample period ethnic Chinese population share is independent from factors that affect local housing prices and employment after 2007. This strategy is numerically equivalent to a Bartik instrument, whose underlying identifying assumption is in terms of cross-sectional local industry shares, as shown in Goldsmith-Pinkham et al. (2018). A similar instrument has been used to study the impact of immigrants on the labor markets by Card (2001).

To illustrate the predictive power of the instrumental variable $CHNShare_{z,0}$, we plot the foreign Chinese housing purchases measure, $CHTV$, normalized by either total housing transaction value or total income, across the distribution of historical ethnic Chinese population share. As shown in Figure 3.3, zip codes that had a higher concentration of ethnic Chinese population historically witnessed significantly more housing purchases by foreign Chinese from 2007-2013. In particular, foreign Chinese houses purchases in the 9th and 10th deciles of the historical ethnic Chinese population share distribution are two and four times higher than the 8th decile, respectively. This suggests that our instrumental variable has strong predictive power for the foreign Chinese housing purchases measure. In all subsequent results tables, we report the first-stage F-statistic to show the statistical significance of the instrument.

There are a couple possible threats to our identification strategy. One is that the neighborhoods which historically attracted ethnic Chinese settlement could have characteristics that systemically differentiate their local housing prices and employment from other neighborhoods. The inclusion of pre-sample period zip code level controls in the empirical model is the first step we take to address this concern. In addition, we examine trends in the outcome variables between neighborhoods with a higher share of ethnic population historically and the rest. Figure 3.4 depicts the quarterly Zillow Home Value Index for zip codes in the top two deciles of the historical ethnic Chinese population share distribution in 2000, which we denote as the "treated" group,

Figure 3.3: **Foreign Chinese Housing Purchases across Zip Codes based on Historical Ethnic Chinese Population Share**



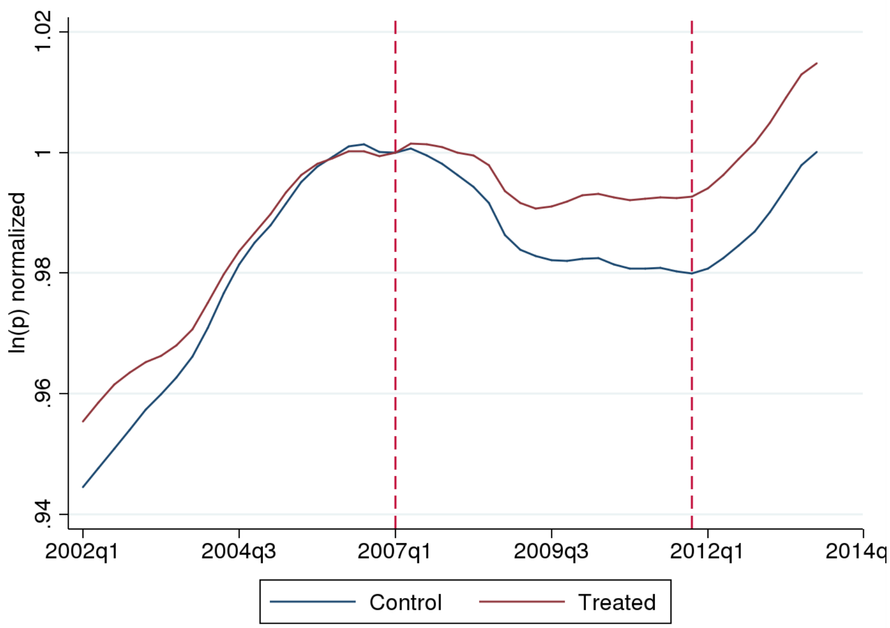
Notes: The left figure shows foreign Chinese housing purchases normalized by total housing transaction value between 2007 and 2013 across zip codes, with the zip codes distributed based on the share of ethnic Chinese population from the 2000 Census Survey. The right figure shows foreign Chinese housing purchases normalized by total income between 2007 and 2013 across zip codes, with the zip codes distributed based on the share of ethnic Chinese population from the 2000 Census Survey.

and zip codes in the bottom eight deciles of the distribution, which we denote as the "control" group. The home price trends between the two groups appear quite similar prior to 2007. Only after 2007, the year when the surge in foreign Chinese housing demand began, does a gap in price trends between the two groups emerge. The gap continues to widen thereafter, with home prices of the treated group at an increasingly higher price level than the control group. To further confirm that the neighborhoods which historically attracted ethnic Chinese settlement are not systemically different from other neighborhoods, we conduct a placebo analysis to test whether changes in housing purchases by foreign Chinese after 2007 are related to local employment prior to 2007. The results show that there is no significant link, which suggests that the neighborhoods which witnessed large foreign Chinese housing demand after 2007 do not exhibit significantly different employment conditions, controlling for local population and education level.[13]

Another concern about our identification strategy is that the surge in foreign Chinese housing purchases in more ethnic-Chinese-concentrated neighborhoods in 2007 may not be exogenous. If this is the case, we would also need to instrument for the 2007 indictor variable in our empirical model. But a number of aggregate factors likely contributed to the increase in foreign Chinese housing demand around that period, including the relaxation of capital flow restrictions in China, the appreciation of the Yuan against the Dollar, and arbitrage opportunities between the US and Chinese real es-

[13] Section 3.4 describes more details about the placebo test and the relevant results.

Figure 3.4: **Zillow Home Price Index by Historical Ethnic Chinese Population Share**



Notes: This figure shows the average quarterly Zillow Home Value index for zip codes in the top two deciles of ethnic Chinese population share based on data from the 2000 Census (the "Treated" group) and those in the bottom eight deciles of ethnic Chinese population share (the "Control" group). For both groups, the index is normalized by its respective value in the first quarter of 2007.

tate market as a result of a housing market boom in China and a housing market crash in the US, not heterogenous policies across US neighborhoods that aimed to attract foreign cash into the US real estate markets. Furthermore, we devise an alternative IV, constructed based on the idea that the surge in foreign Chinese housing demand was demand-driven, in our regressions: China's gross domestic product weighted by the share of ethnic Chinese population across zip codes from the pre-sample period.[14]

Foreign Chinese Housing Purchases and Local Housing Prices

Table 3.2 shows the estimation results from our main regression model based on Equation 3.1 with (log) Zillow Single Family Home Value Index as the dependent variable. All regressions control for zip code level population, education, and a pre-sample period trend of housing prices. Columns 1-3 present results from three distinct specifications over the sample period 2007-2013. Column 1 shows results from a specification with an

[14] Section 3.4 describes more details about the robustness checks and the relevant results.

indicator variable for the post-2007 period but does not contain time fixed effects; column 2 shows results from a specification with county and time fixed effects separately included. Column 3 are estimates from our preferred specification, which includes county-time fixed effects. Across all three specifications, the results indicate that real estate purchases by foreign Chinese induce a significant increase in local housing prices, with the coefficient of interest significant at the 1% level. Based on the estimate from our preferred specification, a 1% increase in housing demand by foreign Chinese as measured by transaction value induces a 0.085% increase in local home prices, which translates to an increase of $497 per home based on the mean Zillow Home Value Index over the sample period ($584,553). The first stage regression F-statistics are highly significant, which indicate that our instrumental variable has strong predictive power for foreign Chinese housing purchases.

We further conduct a sub-period analysis to study whether the effects of foreign Chinese housing purchases on housing prices differ between the housing market crash period of 2007-2011 (column 4) and the housing market recovery period of 2012-2013 (column 5). For the 2007-2011 period, the results show that a 1% increase in housing demand by foreign Chinese, as measured by transaction value, increases local home prices by 0.074% , which corresponds to an increase of $433 per home. To further give a sense of the economic magnitude, we apply this empirical estimate to describe the heterogeneous effects of foreign Chinese housing purchases on home prices across zip codes with varying exposure to ethnic Chinese population settlement historically. Between 2007 and 2011, a zip code in the 90th percentile of the pre-sample period ethnic Chinese share distribution witnessed an increase of 139% in foreign Chinese housing purchase as measured by transaction value ($CHTV$) per year, which is 92 percentage points higher than the increase in the median zip code. Our regression estimate implies that the resulting increase in housing prices of the 90th percentile zip codes is 6.8% higher than the median zip code. Given the context of the housing market crash period during which housing prices declined across all zip code, this result implies that the zip codes which received more real estate capital from China experienced a smaller decline in housing prices. In other words, cash inflow from foreign Chinese played a stabilizing role for US local economies over the housing market crash period.

Compared to the housing market crash period, our results show that effects of foreign Chinese housing purchases on home prices are even greater in magnitude during the housing market recovery years of 2012 and 2013. As shown in column 5 of Table 3.2, a 1% increase in $CHTV$ increases home prices by 0.102% during the recovery years, which corresponds to $596 per house. During that period, foreign Chinese housing purchases increased by 174% for zip codes in the 90th percentile of the historical ethnic Chinese share distribution, while the corresponding increase was 36% for the median zip code. Based on our estimate, this difference leads to a difference of 14.1% in home prices between the two zip codes.

In addition to using the Zillow Home Price Index as a measure for housing prices, we estimate Equation 3.1 using housing transaction prices from DataQuick as the

Table 3.2: **Foreign Chinese Housing Demand and Local Home Prices (Zillow)**

	(1)	(2)	(3)	(4)	(5)
ln(CHTV) $\times \mathbb{I}\{t \geq 2007\}$	0.127***	0.105***	0.085***	0.074***	0.102***
	(0.017)	(0.012)	(0.014)	(0.014)	(0.015)
ln(CHTV)	-0.039*	-0.032	-0.009	0.001	-0.004
	(0.023)	(0.021)	(0.021)	(0.021)	(0.023)
$\mathbb{I}\{t \geq 2007\}$	0.198***				
	(0.020)				
ln(Population)	-0.036**	-0.033**	-0.041**	-0.043***	-0.031*
	(0.016)	(0.016)	(0.017)	(0.017)	(0.016)
Δ ln(HNW), 00-96	1.275***	1.274***	1.268***	1.232***	1.304***
	(0.216)	(0.213)	(0.214)	(0.204)	(0.242)
Education	4.432***	4.398***	4.331***	4.134***	4.250***
	(0.268)	(0.261)	(0.264)	(0.251)	(0.303)
County FE	X	X			
Year FE		X			
County Year FE			X	X	X
Post Period	2007-2013	2007-2013	2007-2013	2007-2011	2012-2013
Model Statistics:					
First Stage F-stat	112.45	116.68	108.67	98.95	85.53
Observations	4588	4588	4571	3474	2470

Notes: The dependent variable is log Zillow Single Family Home Value Index. $CHTV$ denotes foreign Chinese housing transaction values instrumented by the aggregate housing transaction value in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. $\mathbb{I}\{t \geq 2007\}$ is an indicator variable that takes the value 1 if year is post-2007 and 0 otherwise. Education is measured as the population share with bachelor degrees. Additional control variable includes a pre-sample trend variable for the dependent variable calculated as the change in Zillow Home Value Index between 1996 and 2000. Columns 4 shows the results for the housing crash period (2007-2011); Column 5 for the recovery period (2012-2013). The sample period runs from 2001-2013. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

dependent variable. This allows us to study the effects of Chinese capital inflow on local housing prices through another lens: do housing purchases by foreign Chinese push up the prices of newly transacted houses on average? For these regressions, we control for home characteristics including the number of bathrooms, the square footage, and age of the home, in addition to the set of controls specified in Equation 3.1. The results, shown in Table 3.3, are comparable to the estimates in Table 3.2: a 1% increase in the housing demand by foreign Chinese as measured by transaction value increases local home transaction prices by 0.117% on average, which corresponds to an increase of $527 per newly-transacted house. This effect is similar for both the housing market crash period of 2007-2011 (column 4) and the housing market recovery period of 2012-2013 (column 5).

Foreign Chinese Housing Purchases and Local Employment Effects

The first set of results provides a quantification of the impact of real estate capital inflows from China on local asset prices. Do such inflows have an effect on the real economy? We next proceed to study whether foreign Chinese real estate purchases significantly affect local labor markets. To that end, we estimate Equation 3.1 with zip code level employment size as the outcome variable. Our key coefficient of interest, β, provides an estimate of the impact of the surge in foreign Chinese housing purchases on local employment, conditional on county-time fixed effects and zip code level characteristics that may systematically affect labor markets.

The results are shown in Table 3.4. As in Table 3.2, we present results from three distinct specifications over the sample period 2007-2013: i) including post-2007 period indicator variable (column 1); ii) including county and time fixed effects separately (column 2); and iii) including county-time fixed effects (column 3, preferred specification). Across all three specifications, the first-stage F-statistics are again highly significant. Our coefficient of interest is fairly stable. It shows that real estate capital inflows from China have a positive and significant effect on local labor market. Based on estimates from our preferred specification, a 1% increase in housing demand by foreign Chinese, as measured by transaction value, increases local employment by 0.12%.

We further conduct a sub-period analysis to study whether the impact of foreign Chinese housing purchases on local labor market differs between the housing market crash period of 2007-2011 and the recovery period of 2012-2013. Our results show that a 1% increase in foreign Chinese housing purchases increases local employment by 0.10% during the housing market crash period and 0.15% during the recovery period. To highlight the economic magnitude, we compare the employment effects between zip codes from the 90th percentile and 50th percentile of the historical ethnic Chinese share distribution, as we did for the housing price effects. As we noted above, foreign Chinese housing purchases in the 90th percentile zip code was 92% higher than those in the

Table 3.3: **Foreign Chinese Housing Demand and Local Home Prices (Transaction Prices)**

	(1)	(2)	(3)	(4)	(5)
ln(CHTV) $\times \mathbb{I}\{t \geq 2007\}$	0.154***	0.127***	0.117***	0.118***	0.114***
	(0.017)	(0.013)	(0.014)	(0.015)	(0.014)
ln(CHTV)	-0.068***	-0.060***	-0.045**	-0.040**	-0.033*
	(0.022)	(0.019)	(0.019)	(0.019)	(0.020)
$\mathbb{I}\{t \geq 2007\}$	0.015				
	(0.019)				
ln(Population)	-0.009	-0.006	-0.014	-0.022	-0.002
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Δ ln(HTV), 00-96	0.184*	0.158	0.154	0.124	0.113
	(0.102)	(0.102)	(0.102)	(0.103)	(0.105)
Education	5.366***	5.334***	5.279***	5.163***	5.030***
	(0.157)	(0.151)	(0.152)	(0.156)	(0.170)
County FE	X	X			
Year FE		X			
County Year FE			X	X	X
Post Period	2007-2013	2007-2013	2007-2013	2007-2011	2012-2013
Model Statistics:					
First Stage F-stat	123.43	127.84	118.11	107.08	90.52
Observations	4883	4883	4866	3699	2633

Notes: The dependent variable is log housing transaction values, averaged by zip code, from DataQuick. $CHTV$ denotes foreign Chinese housing transaction values instrumented by the aggregate housing transaction value in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. $\mathbb{I}\{t \geq 2007\}$ is an indicator variable that takes the value 1 if year is post-2007 and 0 otherwise. Education is measured as the population share with bachelor degrees. Additional control variable includes a pre-sample trend variable for the dependent variable, calculated as the change in housing transaction values between 1996 and 2000, and variables on home characteristics including the number of bathrooms, the square footage, and age of the home. Columns 4 shows the results for the housing crash period (2007-2011); Column 5 for the recovery period (2012-2013). The sample period runs from 2001-2013. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

median zip code during the 2007-2011 period. Based on our empirical estimate (column 4), this difference translates to a difference of 9.38% in total employment between the two zip codes. Given the context of the housing market crash period, our result implies that foreign Chinese real estate capital can help to alleviate negative employment shocks. During the recovery period between 2012 and 2013, foreign Chinese housing purchases in the 90th percentile zip code was 138 percentage point higher than the median zip code. Based on our estimate (column 5), the resulting increase in total employment in the 90th percentile zip codes is 20.56% higher than the median zip code.

Table 3.4: **Foreign Chinese Housing Demand and Total Employment**

	(1)	(2)	(3)	(4)	(5)
ln(CHTV) $\times \mathbb{I}\{t \geq 2007\}$	0.094***	0.081**	0.120***	0.102**	0.149***
	(0.035)	(0.035)	(0.045)	(0.044)	(0.051)
ln(CHTV)	0.054	0.055	0.023	0.028	0.018
	(0.072)	(0.071)	(0.078)	(0.079)	(0.082)
$\mathbb{I}\{t \geq 2007\}$	-0.105				
	(0.065)				
ln(Population)	0.744***	0.746***	0.737***	0.752***	0.741***
	(0.077)	(0.077)	(0.078)	(0.079)	(0.090)
Δ ln(Emp), 00-96	0.315	0.306	0.311	0.380*	0.332
	(0.194)	(0.193)	(0.195)	(0.197)	(0.210)
Education	2.271***	2.295***	2.292***	2.246***	2.402***
	(0.634)	(0.627)	(0.640)	(0.680)	(0.714)
County FE	X	X			
Year FE		X			
County Year FE			X	X	X
Post Period	2007-2013	2007-2013	2007-2013	2007-2011	2012-2013
Model Statistics:					
First Stage F-stat	131.37	134.33	122.90	110.97	93.24
Observations	4900	4900	4883	3712	2643

Notes: The dependent variable is log total employment size. *CHTV* denotes foreign Chinese housing transaction values instrumented by the aggregate housing transaction value in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. $\mathbb{I}\{t \geq 2007\}$ is an indicator variable that takes the value 1 if year is post-2007 and 0 otherwise. Education is measured as the population share with bachelor degrees. Additional control variable includes a pre-sample trend variable for the dependent variable, calculated as the change in total employment between 1996 and 2000. Columns 4 shows the results for the housing crash period (2007-2011); Column 5 for the recovery period (2012-2013). The sample period runs from 2001-2013. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

Placebo and Robustness Tests

Our two baseline results show that foreign Chinese housing purchases in the US have a strong and positive effect on local house prices and local labor markets. To check the validity of our identification strategy and robustness of the results, we next run a range of placebo test and regressions with alternative specifications.

First, one possible concern about our identification strategy is whether neighborhoods which historically attracted ethnic Chinese settlement have unobserved characteristics that systemically differentiate their local economic conditions from other neighborhoods. To address the concern, we conduct a placebo test. We test whether changes in housing purchases by foreign Chinese after 2007 are correlated with local employment conditions prior to 2007, using the following specification:

$$\ln(Emp, 01-06)_z = \alpha_a + \beta_a \ln(\sum_{2007}^{2013} CHTV)_z + \gamma_a X_z + \varepsilon_{zt}$$

where $(Emp, 01-06)_z$ denotes the change in employment size between 2001 and 2006 for zip code z, $(\sum_{2007}^{2013} CHTV_z)$ denotes the total transaction value of Chinese purchases between 2007 and 2013 for zip code z, and X_z denotes a set of zip code level control variables including population and education. The coefficient of interest is β_a. If the zip codes which attracted more foreign Chinese housing purchases have systematically better employment conditions, β_a is expected to be positive and significant.

Table 3.5 shows the regression results. We find that ex-post foreign Chinese housing purchases do *not* predict ex-ante local employment conditions, controlling for local population and education. In fact, the coefficient β_a is zero and insignificant. This suggests that the zip codes which attracted more ethnic Chinese settlements are not neighborhoods that have more employment opportunities on average. To put it another way, it does not appear that foreign Chinese have been targeting neighborhoods that are systematically different in terms of employment conditions.

Second, our next robustness check addresses the concern that the surge in foreign Chinese housing purchases in more ethnic-Chinese-concentrated neighborhoods in 2007 may not be exogenous. Based on media reports and anecdotal evidence, the surge is likely driven by factors related to changes in China's capital flow regulation, currency appreciation, and domestic real estate market boom around 2007, or, in other words, factors exogenous to US neighborhoods. To test the plausibility of the argument, we re-estimate Equation 3.1 using an alternative instrumental variable that is constructed based on the idea that the surge in foreign Chinese housing demand is demand-driven: China's gross domestic product weighted by the share of ethnic Chinese population across zip codes from the pre-sample period.

The results are shown in the left panel of Appendix-Table C.1. Columns 1-3 present the estimates with Zillow Housing Price Index, housing transaction value, and total employment as the dependent variable, respectively. In these regressions, we include an indicator variable which takes the value one for periods after the surge of home purchases by foreign Chinese (i.e., 2007 or after) and 0 otherwise. County-time dummies are not included here because it would fully absorb the time varying components of the IV. The results show a strong positive impact of foreign Chinese housing purchases on local housing prices and employment. Compared to results using the baseline IV, the coefficient of interest from regressions using the alternative IV is slightly smaller in magnitude across all three specifications: a 1% increase in housing demand by foreign Chinese, as measured by transaction value, increases local home prices by 0.06%, local home transaction prices by 0.09%, and local employment by 0.12%.

Table 3.5: **Foreign Chinese Housing Demand and Ex-ante Employment: Placebo Test**

	(1)	(2)
ln(CHTV, 07-13)	-0.000	0.001
	(0.009)	(0.010)
ln(Population)	-0.074***	-0.077***
	(0.028)	(0.030)
Education		-0.049
		(0.154)
First Stage F-statistic	413.02	328.32
Observations	717	717

Notes: The dependent variable is log change in total employment size between 2001 and 2006. $CHTV, 07-13$ denotes the log change in foreign Chinese housing transaction value between 2007 and 2013 instrumented by the aggregate housing transaction value in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. Education is measured as the population share with bachelor degrees. The sample period runs from 2001-2013. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

Third, we show that the results are robust to using Chinese housing transaction count ($CHTC$), instead of foreign Chinese housing transaction value, as the main regressor. We re-estimate Equation 3.1 with $CHTC$ (instrumented by the aggregate housing transaction count in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period, as before) to check whether our baseline results are driven by purchases of high-end real estate by foreign Chinese. The results are shown in Appendix Tables C.2-C.4. They indicate that higher foreign Chinese housing demand, as measured by transaction count, strongly increases local home prices and local employment. In fact, the coefficient of interest here is more than double the corresponding coefficient in regressions that use $CHTV$. This result suggests that foreign Chinese have been purchasing residential real estate across a full spectrum of types, not only high-end houses.

Quantitative Implications

Given the significant effect of real estate capital inflows from China on local housing prices and employment, we illustrate the quantitative implications of our findings using a counterfactual thought experiment. We estimate how housing prices and employment conditions in zip codes in deciles 1-9 of the historical ethnic Chinese share distribution might have evolved during the housing market crash period and recovery period if each received the same amount of foreign Chinese real estate capital as zip codes in the 10th decile of the distribution. Specifically, we multiply our estimates for β from columns 4 and 5 of Table 3.2 and Table 3.4 by the hypothetical additional amount of foreign Chinese real estate capital for deciles 1-9 of the historical ethnic Chinese share distribution.[15] This is by no means a structural estimate, only a rough quantification.

Figure 3.5 illustrates the predicted outcomes. The left figure plots the predicted house prices and employment size for deciles 1-9 of the historical ethnic Chinese share distribution for the housing market crash period. If zip codes in decile 1 received the same amount of real estate capital from China as decile 10 over the period, its housing prices in 2011 would have been 11.2% higher than the actual price level, and its employment size would have been 15.4% higher than the actual amount. The actual house prices for zip codes in decile 1 was $204,629 in 2011 on average, which is a 50% decline from the average 2007 price level ($406,966). Based on our predicted estimate, real estate capital from China would have alleviated the housing price decline in zip codes in decile 1 by 6 percentage point during the period, increasing the average price to $227,547 in 2011. The actual employment size for zip codes in decile 1 declined by 12% during the 2007-2001 period. Our estimates predict that real estate capital from China would have prevented the employment decline if zip codes in decile 1 received the same amount of real estate capital from China as decile 10 of ethnic Chinese share distribution.

The right plot illustrates the predicted house prices and employment size for deciles 1-9 of the historical ethnic Chinese share distribution for the housing market recovery

[15] The exact formulas we use for the counterfactual calculation are as follow: for the housing market crash period of 2007-2011, $\frac{(CHTV/Total^{2011} - CHTV/Total^{2007})^{decile\ 10} - (CHTV/Total^{2011} - CHTV/Total^{2007})^{decile\ x}}{(CHTV/Total^{2011})^{decile\ x}} * \hat{\beta}_{2007-2011}/4$, where $Total$ denotes the total housing transaction value in California, x denotes deciles 1-9, and $\hat{\beta}_{2007-2011}$ denotes the estimate for β from column 4 from Table 3.2 and Table 3.4 for predictions for housing prices and employment, respectively; for the housing market recovery period of 2012-2013, $\frac{(CHTV/Total^{2013} - CHTV/Total^{2012})^{decile\ 10} - (CHTV/Total^{2013} - CHTV/Total^{2012})^{decile\ x}}{(CHTV/Total^{2013})^{decile\ x}} * \hat{\beta}_{2012-2013}/2$, where $Total$ denotes the total housing transaction value in California, x denotes deciles 1-9, and $\hat{\beta}_{2012-2013}$ denotes the estimate for β from column 5 from Table 3.2 and Table 3.4 for predictions for housing prices and employment, respectively.

period. If zip codes in decile 1 received the same amount of real estate capital from China as decile 10 over the period, its housing prices and employment in 2013 would have been 11.3% and 16.6% higher than the actual amount, respectively. The actual average house prices for zip codes in decile 1 was $252,432 in 2013, which is 19% higher than the average 2012 price level. Based on our predicted estimate, real estate capital from China would have increased the housing price in these zip codes by an additional 17 percentage point during the period, raising the average price to $287,520 in 2013. The actual employment size for zip codes in decile 1 increased by 3% during the 2011-2013 period. Our estimates predict that the increase would have been 17 percentage points larger if zip codes in decile 1 received the same amount of real estate capital from China as decile 10 of ethnic Chinese share distribution.

Overall, our predicted estimates show that real estate capital from China played a stabilization role during a period of economic downturn in the US and helped to accelerate economic recovery during the post-financial crisis period. The effects are heterogenous across zip codes depending on the amount of real estate capital inflow from China, which is correlated with local historical ethnic Chinese share.

3.4 Channel Linking Foreign Chinese Housing Purchases to the Real Economy

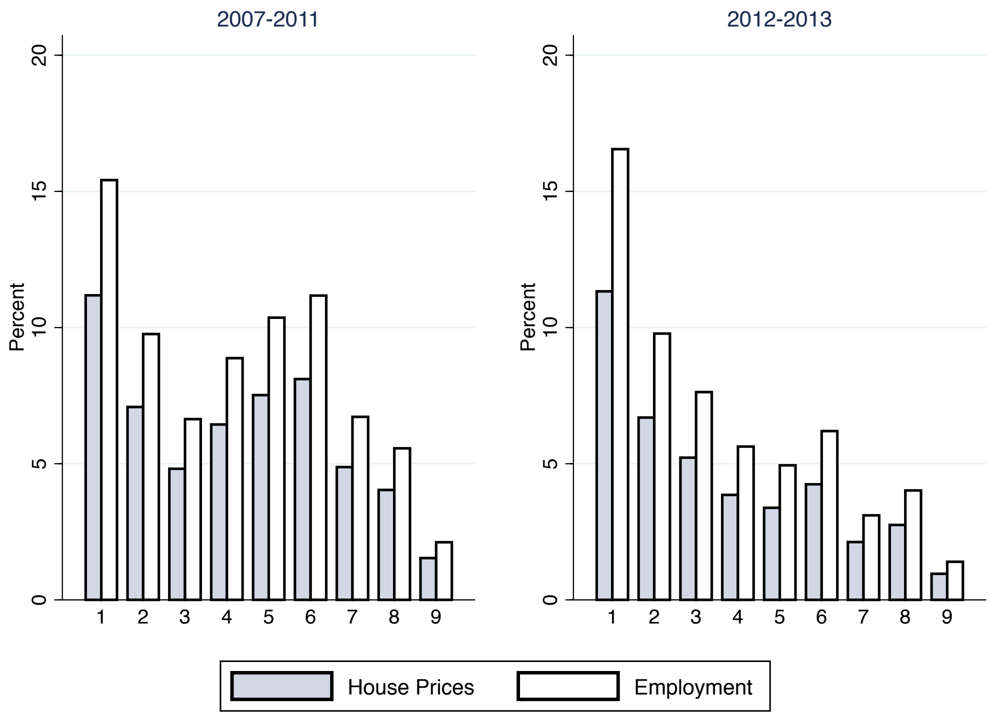
So far, our results robustly show a quantitively large and significant effect of real estate capital inflows from China on local asset prices and local employment. In this section, we explore the channels that link foreign Chinese real estate capital to the real economy.

Migration Channel

One potential mechanism underlying the effects is a migration channel. If foreign Chinese are moving into the houses they purchased in the US or renting them out, that would entail a migration inflow into the corresponding neighborhoods, which could push up demand for local goods and thereby local employment. To test whether this is the case, we relate foreign Chinese housing transaction, $CHTV$, to the number of income tax returns at the zip code level, which is often considered a measure of the number of households, as in Greenland et al. (2019).[16] The income tax returns data are from administrative records of individual income tax returns (Forms 1040) from the Internal Revenue Service (IRS). If the inflow of real estate capital from China induces migration inflow, we expect a positive and significant relationship between $CHTV$ and the number of tax filings.

[16] A more direct dependent variable for this test could be migration inflow and outflow count. However, to the best of our knowledge, migration data are not collected at the zip code.

Figure 3.5: **Thought Experiment: Counterfactual Housing Prices and Employment Conditions**



Notes: This figure illustrates the predicted housing prices and employment size in zip codes in deciles 1-9 of the historical ethnic Chinese share distribution if each received the same amount of foreign Chinese real estate capital as the zip codes in 10th decile of the distribution. The left panel shows the predicted housing prices and employment size for the housing market crash period of 2007-2011. The exact formula for the counterfactual calculation is as follow: $\frac{(CHTV/Total^{2011}-CHTV/Total^{2007})^{decile\ 10}-(CHTV/Total^{2011}-CHTV/Total^{2007})^{decile\ x}}{(CHTV/Total^{2011})^{decile\ x}} * \hat{\beta}_{2007-2011}/4$, where $Total$ denotes the total housing transaction value in California, x denotes deciles 1-9, and $\hat{\beta}_{2007-2011}$ denotes the estimate for β from column 4 from Table 3.2 and Table 3.4 for predictions for housing prices and employment, respectively. The right panel shows the predicted housing prices and employment size for the housing market recovery period of 2012-2013. The exact formula for the counterfactual calculation is as follow: $\frac{(CHTV/Total^{2013}-CHTV/Total^{2012})^{decile\ 10}-(CHTV/Total^{2013}-CHTV/Total^{2012})^{decile\ x}}{(CHTV/Total^{2013})^{decile\ x}} * \hat{\beta}_{2012-2013}/2$, where $Total$ denotes the total housing transaction value in California, x denotes deciles 1-9, and $\hat{\beta}_{2012-2013}$ denotes the estimate for β from column 5 from Table 3.2 and Table 3.4 for predictions for housing prices and employment, respectively.

Table 3.6 reports the results. As in the prior tables, column 1 shows results from a specification with an indicator variable for the post-2007 period but does not contain time fixed effects; column 2 shows results from a specification with county and time fixed effects separately included. Column 3 are estimates from our preferred specification, which includes county-time fixed effects. Columns 4 and 5 show estimates from a sub-period analysis that aims to compare the effects between the housing market crash period of 2007-2011 and the recovery period of 2012-2013. Overall, we do *not* find a positive relationship between $CHTV$ and the number of tax filings. In fact, results show that foreign Chinese house purchases lower the number of filings on average during the sample period. This suggests that real estate capital inflow from China actually drives households out of the respective neighborhoods on net.[17]

The observation that foreign Chinese housing purchases are not accompanied by an inflow of migrants could be reconciled with anecdotal evidence that foreign Chinese tend to leave their house purchases abroad vacant. Studies by Rosen et al. (2017) and Simons et al. (2016) find that foreign Chinese real estate buyers tend to neither use the purchased properties as primary residences nor rent it out. They show that housing purchases by foreign Chinese in the US are positively related to the number of Chinese investors in the EB-5 Immigrant Investor Visa Program, who are primarily interested in obtaining a green card for their children instead of actual returns to their real estate investments. The tendency of foreign Chinese real estate buyers to leave housing properties vacant may not be surprising in light of a similar practice in China: Glaeser et al. (2017) show that housing vacancy rates in China are much higher than in the US, reaching more than 20% in major Chinese cities in 2012.

We conclude from Table 3.6 that a channel of migration inflow, which could raise local demand and employment, is not driving the relationship between foreign Chinese housing purchases and local labor markets.

Housing Net Worth Channel

Another potential channel linking foreign Chinese housing purchases to local employment condition is a housing net worth channel: higher housing wealth could affect employment by changing consumer demand through either a direct wealth effect or less binding borrowing constraints driven by the rise in collateral value, as argued in papers by Mian et al. (2013b) and Mian and Sufi (2014). One of the key predictions of the housing net worth channel hypothesis is that the impact of demand changes due to housing net worth fluctuations on local employment should show up foremost in the non-tradable sector employment, since non-tradable sector employment depends primarily on local demand while the tradable sector is more diversified in its geographic

[17] To address possible concern that there is a time lag in when migration will be reflected in the number of tax returns, we also study the relation between foreign Chinese housing transaction in year t and the number of income tax returns in year $t+1$. The results, reported in Appendix-Table C.5, are quantitively and qualitatively similar to the ones from regressions using contemporaneous measures.

Table 3.6: **Testing for Direct Demand Channel: Number of Tax Return Filings**

	(1)	(2)	(3)	(4)	(5)
ln(CHTV) $\times\mathbb{I}\{year \geq 2007\}$	-0.052***	-0.052***	-0.036***	-0.035***	-0.039***
	(0.013)	(0.013)	(0.010)	(0.010)	(0.010)
ln(CHTV)	0.034**	0.034**	0.020	0.017	0.017
	(0.015)	(0.015)	(0.013)	(0.013)	(0.013)
$\mathbb{I}\{t \geq 2007\}$	0.078***				
	(0.010)				
ln(Population)	0.906***	0.907***	0.907***	0.909***	0.924***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.017)
Education	1.030***	1.026***	1.039***	1.100***	1.123***
	(0.111)	(0.111)	(0.111)	(0.111)	(0.105)
Δ ln(Returns), 01-98	0.701***	0.703***	0.703***	0.674***	0.701***
	(0.100)	(0.100)	(0.101)	(0.093)	(0.092)
County Fixed Effects	X	X			
Year Fixed Effects		X			
County Year Fixed Effects			X	X	X
Post Period	2007-2013	2007-2013	2007-2013	2007-2011	2012-2013
Model Statistics:					
First Stage F-statistic	119.06	124.11	117.32	106.22	93.17
Observations	4582	4582	4566	3409	2343

Notes: The dependent variable is log income tax returns. $CHTV$ denotes foreign Chinese housing transaction values instrumented by the aggregate housing transaction value in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. Education is measured as the population share with bachelor degrees. Δ ln(Returns), 01-98 is a pre-sample trend variable for the dependent variable, calculated as the change in the numbers of returns between 1998 and 2001. Columns 4 shows the results for the housing crash period (2007-2011); Column 5 for the recovery period (2012-2013). The sample period runs from 2001-2013. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

origins of demand. In the context of linking foreign Chinese housing purchases to local employment, the housing net worth channel hypothesis predicts that the surge in real estate capital from China raises employment in non-tradable sectors but not the tradable sectors.

To test for this effect, we decompose total employment into tradable sector employment and non-tradable sector employment, and re-estimate Equation 3.1 with tradable and non-tradable sector employment as the dependent variables. The results are reported in Table 3.7. We find that housing purchases by foreign Chinese have a positive and significant effect on local non-tradable sector employment. A 1% increase in $CHTV$ increases zip code level non-tradable sector employment by 0.122% and 0.137% during the housing market crash period and housing market recovery period, respectively (columns 3 and 5). On the other hand, increases in housing purchases by foreign Chinese seem to have no statistically significant impact on local tradable sector employment during the sample period, based on estimates in columns 2, 4 and 6.[18]

One may be concerned that the strong relation between foreign Chinese housing purchases and non-tradable sector employment is driven by an increase in employment in the construction sections in particular, since it is plausible that higher housing prices due to higher foreign Chinese demand induces higher housing supply, which could increase demand for construction. We test for such mechanism by studying whether the strong relationship between foreign Chinese housing prices and non-tradable sector employment holds when we exclude the construction sector from the definition of the non-tradable sector when constructing the non-tradable sector employment variable. The results are reported in Table 3.8. We continue to find a positive and significant link between $CHTV$ and non-tradable sector employment.

These results support the housing net worth channel hypothesis, as the effects are concentrated in the non-tradable non-construction sector employment. To further check the robustness of the result, we directly test whether the housing net worth channel played a role by adapting the housing net worth tests from Mian et al. (2013b) and Mian and Sufi (2014) to the context of housing price surge due to foreign Chinese demand. Specifically, we study the effects of local home prices, instrumented by $CHTV$, on local tradable and non-tradable sector employment, by estimating following regression:

$$\ln(Emp)_{zt} \;=\; \alpha_0 + \beta_1 \ln(H\tilde{N}W)_{zt} + \beta_2 \ln(H\tilde{N}W)_{zt} \times \mathbb{I}\{t \geq 2007\} + \gamma X_z + \eta_{ct} + \varepsilon_{zt}$$

where $\ln(Emp)_{zt}$ is the log of either tradable or non-tradable sector employment and $\ln(H\tilde{N}W)_{zt}$ is the log of the Zillow Home Value Index instrumented by the share of

[18] We find similar results when using the alternative IV of China's gross domestic product weighted by the share of ethnic Chinese population across zip codes from the pre-sample period or using foreign Chinese housing demand as measured by transaction count, $CHTC$ (instead of $CHTV$), as the main regressor.. The results are reported in the right panel of Appendix-Table C.1 and Appendix-Table C.6.

Table 3.7: **Foreign Chinese Housing Demand and Tradable/Non-tradable Sector Employment**

	(1) ln(NT Emp)	(2) ln(T Emp)	(3) ln(NT Emp)	(4) ln(T Emp)	(5) ln(NT Emp)	(6) ln(T Emp)
ln(CHTV) $\times \mathbb{I}\{t \geq 2007\}$	0.137***	0.082	0.122***	0.046	0.137***	0.144
	(0.044)	(0.101)	(0.043)	(0.099)	(0.044)	(0.116)
ln(CHTV)	-0.060	0.244	-0.057	0.246	-0.060	0.259
	(0.076)	(0.171)	(0.078)	(0.175)	(0.076)	(0.181)
ln(Population)	0.887***	0.833***	0.894***	0.889***	0.887***	0.822***
	(0.070)	(0.142)	(0.071)	(0.146)	(0.070)	(0.158)
Δ ln(Emp$_{NT,T}$), 00-96	-0.103	-0.162	-0.074	-0.153	-0.103	-0.113
	(0.129)	(0.113)	(0.136)	(0.120)	(0.129)	(0.121)
Education	2.570***	-4.817***	2.524***	-4.738***	2.570***	-5.102***
	(0.611)	(1.266)	(0.655)	(1.352)	(0.611)	(1.453)
County Year FE	X	X	X	X	X	X
Post Period	2007-2013	2007-2013	2007-2011	2007-2011	2012-2013	2012-2013
Model Statistics:						
First Stage F-stat	122.57	118.15	111.49	107.49	122.57	90.85
Observations	4876	4811	3708	3668	4876	2607

Notes: The dependent variable is log non-tradable or tradable sector employment. $CHTV$ denotes foreign Chinese housing transaction values instrumented by the aggregate housing transaction value in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. $\mathbb{I}\{t \geq 2007\}$ is an indicator variable that takes the value 1 if year is post-2007 and 0 otherwise. Education is measured as the population share with bachelor degrees. Δ ln(Emp$_{NT,T}$), 00-96 is a pre-sample trend variable for the dependent variable, calculated as the change in non-tradable or tradable sector employment between 1996 and 2000. Columns 4 shows the results for the housing crash period (2007-2011); Column 5 for the recovery period (2012-2013). The sample period runs from 2001-2013. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

ethnic Chinese population in 2000, $\mathbb{I}\{t \geq 2007\}$ is an indicator variable that takes the value 1 if year is post-2007 and 0 otherwise, X_z are time-invariant zip code level controls including population, education measured as the percentage of the population with a bachelor degree, and a pre-sample trend variable for the corresponding dependent variables calculated as the difference between tradable/non-tradable sector employment in 2001 and 2006, and η_{ct} are county-year fixed effects.

The results, reported in Table 3.9, corroborate the earlier findings that zip codes with more foreign Chinese home purchases experience a higher increase in both local home prices and local employment. Furthermore, results in columns 1 and 3 show that a 1% increase in home prices increases non-tradable sector employment by 0.66% - 0.69%, and the effects are similar for both the housing market crash period and the housing market recovery periods. On the other hand, foreign Chinese housing

Table 3.8: **Foreign Chinese Housing Demand and Non-Tradable Sector Employment, Excluding Construction Sector**

	(1) ln(NT Emp)	(2) ln(T Emp)	(3) ln(NT Emp)	(4) ln(T Emp)	(5) ln(NT Emp)	(6) ln(T Emp)
ln(CHTV) $\times\mathbb{I}\{year \geq 2007\}$	0.101**	0.082	0.096**	0.046	0.101**	0.144
	(0.044)	(0.101)	(0.042)	(0.099)	(0.044)	(0.116)
ln(CHTV)	-0.024	0.244	-0.023	0.246	-0.024	0.259
	(0.078)	(0.171)	(0.079)	(0.175)	(0.078)	(0.181)
ln(Population)	1.010***	0.833***	1.012***	0.889***	1.010***	0.822***
	(0.071)	(0.142)	(0.074)	(0.146)	(0.071)	(0.158)
Education	3.219***	-4.817***	3.194***	-4.738***	3.219***	-5.102***
	(0.595)	(1.266)	(0.625)	(1.352)	(0.595)	(1.453)
Δ ln(NT/T Emp), 00-96	-0.165*	-0.162	-0.141*	-0.153	-0.165*	-0.113
	(0.084)	(0.113)	(0.084)	(0.120)	(0.084)	(0.121)
County Year Fixed Effects	X	X	X	X	X	X
Post Period	2007-2013	2007-2013	2007-2011	2007-2011	2012-2013	2012-2013
Model Statistics:						
First Stage F-statistic	121.63	118.15	110.71	107.49	121.63	90.85
Observations	4876	4811	3708	3668	4876	2607

Notes: The dependent variable is log non-tradable (excluding the construction sector) or tradable sector employment. $CHTV$ denotes foreign Chinese housing transaction values instrumented by the aggregate housing transaction value in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. $\mathbb{I}\{t \geq 2007\}$ is an indicator variable that takes the value 1 if year is post-2007 and 0 otherwise. Education is measured as the population share with bachelor degrees. Δ ln($\text{Emp}_{NT,T}$), 00-96 is a pre-sample trend variable for the dependent variable, calculated as the change in non-tradable (excluding the construction sector) or tradable sector employment between 1996 and 2000. Columns 4 shows the results for the housing crash period (2007-2011); Column 5 for the recovery period (2012-2013). The sample period runs from 2001-2013. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

demand-instrumented home prices do not have a significant impact on tradable sector employment. This set of results further suggests that a housing net worth channel drives the relationship between higher housing prices induced by higher foreign Chinese purchase and local employment.

To delve deeper into this relationship, we next explore a potential factor that links foreign Chinese housing demand to housing net worth, and thereby local employment through the housing net worth channel. Papers including Campbell et al. (2011) and Mian et al. (2015) show that higher local foreclosure leads to lower local housing prices. Given such, we test whether foreign Chinese housing demand predicts lower local foreclosure share, which could explain the cross-sectional differences in local housing prices across zip codes. Table 3.10 reports the results. We find that the surge in foreign Chinese housing demand has a negative and significant relationship with local foreclosure share. In other words, zip codes that received more real estate capital from China have

Table 3.9: **Foreign Chinese Housing Demand and Housing Net Worth**

	(1) ln(NT Emp)	(2) ln(T Emp)	(3) ln(NT Emp)	(4) ln(T Emp)
ln(HNW)	0.690***	0.560	0.664***	0.406
	(0.167)	(0.400)	(0.179)	(0.426)
ln(HNW) $\times \mathbb{I}\{year > 2007\}$			0.067	0.418**
			(0.079)	(0.211)
Δ ln(NT/T Emp), 06-01	0.342***	0.202*	0.341***	0.202*
	(0.125)	(0.122)	(0.125)	(0.122)
ln(Population)	0.925***	1.178***	0.924***	1.174***
	(0.058)	(0.138)	(0.058)	(0.139)
County Year Fixed Effects	X	X	X	X
Model Statistics:				
First Stage F-statistic	144.25	129.94	70.38	64.44
Observations	3406	3345	3406	3345

Notes: The dependent variable is log non-tradable or tradable sector employment. HNW denotes Zillow Home Value Index instrumented by the aggregate housing transaction value in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. $\mathbb{I}\{t \geq 2007\}$ is an indicator variable that takes the value 1 if year is post-2007 and 0 otherwise. Δ ln($\text{Emp}_{NT,T}$), 06-01 is a pre-sample trend variable for the dependent variable, calculated as the change in non-tradable or tradable sector employment between 2001 and 2006. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

relatively lower foreclosure rates relative to other zip codes.[19] While it is possible that the higher foreclosure in zip codes which received less foreign Chinese real estate capital entails a migration outflow from these neighborhoods into ones that received more foreign Chinese real estate capital, the earlier results on migration suggest that any forces that might induce migration outflow, if they exist, is dominated by crowding-out effects. Overall, this set of results sheds light on a factor that explains how foreign Chinese housing demand could affect housing net worth, and thereby local employment through the housing net worth channel.

3.5 Simple Model

In this section, we develop a simple partial equilibrium framework to rationalize the empirical results on the impact of housing purchases by foreign Chinese on house prices

[19] The results are similar when we use zip code level foreclosure count as the dependent variable, as shown in Appendix-Table C.7.

Table 3.10: **Foreign Chinese Housing Demand and Foreclosure (Share)**

	(1)	(2)	(3)	(4)	(5)
ln(CHTV) $\times \mathbb{I}\{year \geq 2007\}$	-0.107***	-0.101***	-0.074***	-0.092***	-0.040***
	(0.007)	(0.007)	(0.007)	(0.009)	(0.004)
ln(CHTV)	0.054***	0.052***	0.029***	0.033***	0.006**
	(0.006)	(0.006)	(0.005)	(0.006)	(0.003)
$\mathbb{I}\{t \geq 2007\}$	0.165***				
	(0.006)				
ln(Population)	0.008*	0.004	0.007	0.011**	-0.002
	(0.005)	(0.004)	(0.004)	(0.005)	(0.002)
Education	-0.810***	-0.791***	-0.759***	-0.819***	-0.300***
	(0.040)	(0.038)	(0.037)	(0.044)	(0.023)
County Fixed Effects	X	X			
Year Fixed Effects		X			
County Year Fixed Effects			X	X	X
Post Period	2007-2013	2007-2013	2007-2013	2007-2011	2012-2013
Model Statistics:					
First Stage F-statistic	125.99	129.84	120.23	108.98	92.92
Observations	4900	4900	4883	3712	2643

Notes: The dependent variable is foreclosure count weighted by total residential home count. $CHTV$ denotes foreign Chinese housing transaction values instrumented by the aggregate housing transaction value in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. $\mathbb{I}\{t \geq 2007\}$ is an indicator variable that takes the value 1 if year is post-2007 and 0 otherwise. Education is measured as the population share with bachelor degrees. Additional control includes a pre-sample trend variable for the dependent variable, calculated as the change in foreclosure share between 1996 and 2000. Columns 4 shows the results for the housing crash period (2007-2011); Column 5 for the recovery period (2012-2013). The sample period runs from 2001-2013. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

and employment in the local economy. In particular, we formalize the housing net worth channel and highlight how it drives the effects. We then discuss the mapping between the model predictions and our empirical findings.

Baseline

Consider an economy consisted of Z equally-sized regions indexed by z. Each region produces two types of goods, tradable (indexed by T) and non-tradable (indexed by N). The tradable good is nationally traded and serves as a numerarie good with $P^T = 1$. There is a fixed stock of housing in each region (indexed by H). Regions can freely trade the tradable good, but must consume the non-tradable good produced locally.

Given our empirical finding that migration does not appear to play a significant role in linking foreign Chinese housing demand to the local economy, we impose the restriction that labor cannot move across islands but can move freely between the tradable and non-tradable sectors within an island, for simplicity. Let D_z denote the nominal income in each region, which consists of wages and rental income (rebated to local workers).

Preference Workers in region z have Cobb-Douglas preferences over tradable and non-tradable goods as well as housing (C_z^N, C_z^T, and C_z^H) with prices P_z^N, P^T, and P_z^H, and they spend income shares α, β, and $1-\alpha-\beta$ on the three goods.

Budget Constraint The budget constraint of workers is $P_z^N C_z^N + C_z^T + P_z^H C_z^H = D_z$. From the Cobb-Douglas preference specification, $P_z^N C_z^N = \alpha D_z$, $C_z^T = \beta D_z$, and $P_z^H C_z^H = (1-\alpha-\beta) D_z$ on the non-tradable, tradable, and housing consumption, respectively.

Output All regions face the same tradable good price, while the non-tradable good price may be region-specific since non-tradable good are produced locally. Production is governed by a constant returns technology for tradable and non-tradable goods with employed labor, e as the only factor input and produces output according to $y_z^T = be_z^T$, and $y_z^N = ae_z^N$, respectively, where b and a are productivity parameters. The housing supply is fixed at H_z.

Employment Total employment in each region is normalized to one with $e_z^T + e_z^N = 1$. Wages in the non-tradable and tradable sectors are given by $w_z^N = aP_z^N$ and $w_z^T = bP^T = b$. Free mobility of labor across sectors equates the two wages, which implies $w_z = w = b$ and $P_z^N = \frac{b}{a}$.

Equilibrium In equilibrium, the goods markets clear. For non-tradable goods, $y_z^N = C_z^N$ in each region. For tradable goods, the total demand equate to the total production across all regions: $\sum_{z=1}^{Z} y_z^Z = \sum_{z=1}^{Z} C_z^T$. We solve the model with a symmetry assumption: in the initial state, all regions have the same housing stock $H_z = H_0$, and the economy achieves full employment. Housing demand is equal to supply in equilibrium: $C_z^H = H_0$. Since the nominal income is $D_z = w + P_z^H H_0$, we obtain equilibrium house prices and the nominal income. The equilibrium variables in this simple framework are solved as follows:

$$\text{Prices}: \qquad P_z^{*N} = \frac{b}{a}; \qquad P^{*T} = 1; \qquad P_z^{*H} = \frac{1-\alpha-\beta}{\alpha+\beta}\frac{b}{H_0} \equiv P_0;$$

$$\text{Employment}: \qquad e_z^{*N} = \frac{\alpha}{\alpha+\beta} \equiv e_0^N; \qquad e_z^{*T} = \frac{\beta}{\alpha+\beta} \equiv e_0^T;$$

$$\text{Wages}: \qquad w_z^{*N} = w_z^{*T} = b \equiv w;$$

$$\text{Nominal income}: \qquad D_z^* = w + P_0 H_0 = b + \frac{1-\alpha-\beta}{\alpha+\beta} b \equiv D_0.$$

Effects of House Demand by Foreigners

Suppose now that there is heterogeneous housing demand by foreign Chinese across regions:

$$H_0 = C_z^H + C_{\text{chn},z}^H,$$

where C_z^H is the housing demand by local workers, and C_{chn}^H is the (exogenous) housing demand by foreign Chinese. Since $C_z^H = (1-\alpha-\beta)\frac{D_z}{P_z^H}$ and $D_z = b + P_z^H H_0$, we obtain housing prices:

$$P_z^H = \frac{(1-\alpha-\beta)b}{(\alpha+\beta)H_0 - C_{\text{chn},z}^H},$$

which shows that regions with more housing purchases from foreign Chinese have higher house prices (a housing boom).

Foreign Chinese Housing Demand and Local Housing Prices. Consider two regions, one with a house demand from foreign Chinese (treated region, $C_{\text{chn},z}^H > 0$) and one without (control region, $C_{\text{chn},z}^H = 0$). The cross-sectional difference in house prices between the two regions is

$$P_{z,\text{treated}}^H - P_{z,\text{control}}^H = \frac{(1-\alpha-\beta)b}{(\alpha+\beta)H_0 - C_{\text{chn},z}^H} - P_0,$$

which is an increasing function of $C_{\text{chn},z}^H$ with

$$\frac{\partial(P_{z,\text{treated}}^H - P_{z,\text{control}}^H)}{\partial C_{\text{chn},z}^H} = \frac{(1-\alpha-\beta)b}{[(\alpha+\beta)H_0 - C_{\text{chn},z}^H]^2} > 0. \tag{3.2}$$

Our model thus predicts that higher housing demand by foreign Chinese in the treated region increases the cross-sectional difference in house prices between the two regions. This prediction from Equation 3.2 is supported by our empirical results.

The nominal income now becomes

$$D_z = b + P_z^H H_0 = b + \underbrace{\frac{(1-\alpha-\beta)bH_0}{(\alpha+\beta)H_0 - C_{\text{chn},z}^H}}_{\text{housing net worth channel}}.$$

This equation shows that house demand by foreign Chinese raises the nominal demand via a housing net worth channel.

Foreign Chinese Housing Demand and Local Employment. The non-tradable sector employment becomes

$$e_z^N = \frac{\alpha}{b} D_z = \alpha + \frac{\alpha(1-\alpha-\beta)H_0}{(\alpha+\beta)H_0 - C_{\text{chn},z}^H},$$

which shows that the non-tradable sector expands in regions with higher house demand from foreigners. The cross-sectional difference in local employment in the non-tradable sector between the treated and control regions is an increasing function of $C_{\text{chn},z}^H$:

$$\frac{\partial(e_{z,\text{treated}}^N - e_{z,\text{control}}^N)}{\partial C_{\text{chn},z}^H} = \frac{\alpha(1-\alpha-\beta)}{[(\alpha+\beta)H_0 - C_{\text{chn},z}^H]^2} > 0. \quad (3.3)$$

The prediction on the employment effect in the non-tradable sectors, shown in Equation 3.3, is also supported by our empirical results.

Based on the full employment condition, however, output and employment in the tradable sector, will shrink in treated regions:

$$e_z^T = 1 - e_z^N = 1 - \alpha - \frac{\alpha(1-\alpha-\beta)H_0}{(\alpha+\beta)H_0 - C_{\text{chn},z}^H},$$

so the cross-sectional difference in local employment in the tradable sector between the treated and control regions is a decreasing function of $C_{\text{chn},z}^H$:

$$\frac{\partial(e_{z,\text{treated}}^T - e_{z,\text{control}}^T)}{\partial C_{\text{chn},z}^H} < 0. \quad (3.4)$$

In our partial equilibrium setup, the increased demand for tradable good is met by imports from other regions outside the local economy since we assume that the tradable good is nationally traded with a fixed price $P^T = 1$. So the inflow of Chinese real estate capital acts as financial transfers to recipient regions, allowing them run trade deficits. To put it more concretely, we derive the deficit as the difference between consumption and output in the traded sector:

$$\begin{aligned}
\text{Deficit}_z &= C_z^T - b e_z^T \\
&= \beta D_z - b\left(1 - \frac{\alpha}{b} D_z\right) \\
&= (\alpha+\beta)D_z - b \\
&= (\alpha+\beta)D_z - D_z + P_z^H H_0 \\
&= P_z^H H_0 - (1-\alpha-\beta)D_z \\
&= P_z^H H_0 - P_z^H C_z^H \\
&= P_z^H C_{\text{chn},z}^H.
\end{aligned}$$

In the aggregate, the production and employment in the tradable sector decrease due to the inflow of real estate purchases by foreign Chinese.[20] The prediction in Equation 3.4 is not confirmed by our empirical analysis as we find statistically insignificant effect of Chinese real estate purchases on the employment in the tradable sector.

Overall, our simple model predicts that housing purchases by foreign Chinese raise local house prices, increase local employment in the non-tradable sector via a housing net worth channel, and lower local employment in the tradable sector. The first two are consistent with our empirical results, but we do not find a strong relationship between foreign Chinese real estate capital and local tradable sector employment. It does not seem obvious how to generate a null effect on the tradable sector employment in the model without introducing some friction such as employment slack or migration. In that regard, the static and partial equilibrium nature of our model is not completely satisfactory because it fails to capture the effect on the tradable sector. Nonetheless, it formalizes the economic mechanism underlying the effects of foreign housing purchases on house prices and non-tradable sector employment in the local economy.

Discussion

Our simple model assumes no migration or commuting across regions, which should be relaxed in a more general framework. In such a setting, competing forces will arise in driving the effects of an increase in housing demand by foreigners on local employment across sectors. Suppose that prices and wages are flexible, workers can move and commute across regions, and there are both homeowners versus renters in local economies. We consider the differential effects of a positive shock in foreign housing demand on local homeowners versus renters. A positive housing price shock in the local economy due to housing purchases by foreign Chinese will increase the wealth of local homeowners, which would push up spending on non-tradable goods and raise local employment in the non-tradable sector. At the same time, rental prices would also increase, which would push out renters from the treated regions into cheaper areas (control regions). This would lower consumption demand and employment for non-tradable goods in the treated region. If the positive force from housing net worth channel exceeds the negative force from the outflow of renters (subject to migration and commuting costs), then non-tradable sector employment in treated regions will still increase. Our empirical results support this prediction.

On the other hand, the effect of foreign housing demand on the employment in the tradable sector is less clear. Homeowners' increased demand for tradable goods in treated regions can be supplied by the production elsewhere, so the employment effect of a rising real estate demand from foreign Chinese in the tradable sector is more diffused than the effect on the non-tradable sector. In order for a model with a more general setup to predict a null effect on tradable sector employment—a result from our

[20] This negative impact on the tradable sector can be viewed a form of a "Dutch disease."

empirical analysis, we would need to ensure that the tradable sector in both treated and the control regions expand to the same extent, despite that fact that the living costs in treated regions have risen. This prediction would be difficult to generate in a general equilibrium framework with flexible prices and full employment. For that to happen, we may need to introduce agglomeration spillovers in the treated region. Otherwise, the model would predict a reduction in the tradable sector employment in treated regions, which contradicts our empirical findings, just as our simple model.

3.6 Conclusion

In this paper, we document an unprecedented surge in housing purchases by foreign Chinese in the US over the past decade and analyzes the effects of these real estate capital inflows on US local economies. Using detailed transaction-level housing purchase data, we utilize an instrumental variable strategy that exploits cross-zip code variation in the concentration of Chinese population stemming from pre-sample period differences in Chinese population settlement to instrument for the housing transactions by foreign Chinese. We find that housing purchases by foreign Chinese significantly induces higher local housing prices and higher local employment. In particular, the effects on employment are concentrated in the non-tradable sectors, suggesting that housing net worth played a role in driving the effects. We develop a simple model that helps to illustrate how housing purchases by foreign Chinese affect the local employment through the housing wealth channel.

This paper is the first academic paper that studies the effects of a "China shock" on the finance side on US local economies. Given China has been seeking to open up its capital markets, a better understanding of the implications of a "China shock" on the finance side for the rest of the world is of utmost importance. Our results point to potential welfare gains and losses that come with China's opening up for the rest of the world. Moreover, our evidence highlights the role of capital inflow on domestic real economy, especially in times of economic downturns. During the housing market crash period between 2007-2011, the improvement in household balance sheet resulting from foreign real estate capital inflows played a mitigating role for the US local economies.

Bibliography

Abadie, A. and G. Imbens (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business and Economic Statistics 29*(1), 1–11.

Aguiar, M. and E. Hurst (2015). Consumption vs expenditure. *Journal of Political Economy 113*(5), 919–948.

Aizenman, J. and Y. Jinjarak (2009). Current account patterns and national real estate markets. *Journal of Urban Economics 66*(2), 75–89.

Alesina, A. and N. Fuchs-Schündeln (2007). Goodbye Lenin (or not?): The effect of communism on people's preferences. *American Economic Review 97*(4), 1507–1528.

Andreski, P., G. Li, M. Z. Samancioglu, and R. Schoeni (2014). Estimates of annual consumption expenditures and its major components in the psid in comparison to the ce. *American Economic Review 104*(5), 132–135.

Aoki, K., G. Benigno, and N. Kiyotaki (2016). Monetary and financial policies in emerging markets. Working Paper.

Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies* (58), 277–297.

Arellano, M. and O. Bover (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics* (68), 29–51.

Attanasio, O. and M. Browning (1995). Consumption over the life-cycle and over the business-cycle. *American Economic Review 85*(5), 1118–1137.

Attanasio, O. P. (1999). Consumption (Chapter 11). Volume 1, Part B of *Handbook of Macroeconomics*, pp. 741 – 812. Elsevier.

Attanasio, O. P. and G. Weber (1995). Is consumption growth consistent with intertemporal optimization? Evidence from the consumer expenditure survey. *Journal of political Economy 103*(6), 1121–1157.

Autor, D., D. Dorn, G. Hanson, and K. Majlesi (2017). Importing political polarization? The electoral consequences of rising trade exposure. *Working Paper*.

Autor, D., D. Dorn, G. Hanson, G. Pisano, and P. Shu (2017). Foreign competition and domestic innovation: Evidence from U.S. patents. *Working Paper*.

Autor, D. H., D. Dorn, and G. H. Hanson (2013). The China syndrome: Local labor market effects of import competition in the United States. *The American Economic Review 103*(6), 2121–2168.

Autor, D. H., D. Dorn, and G. H. Hanson (2018). When work disappears: Manufacturing decline and the falling marriage-market value of young men. *Working Paper*.

Badarinza, C. and T. Ramadorai (2015). Home away from home? Foreign demand and london house prices. *Working Paper*.

Ball, L. (2014). Long-term damage from the great recession in OECD countries. NBER Working Paper 20185.

Bar-Ilan, A. and A. Blinder (1992). Consumer durables: Evidence on the optimality of usually doing nothing. *Journal of Money, Credit and Banking 24*(2), 258–272.

Baskaya, Y., J. Giovanni, S. Kalemli-Ozcan, and M. Ulu (2017). International spillovers and local credit cycles. Working Paper.

Beck, T., A. Demirguc-Kunt, and M. S. M. Peria (2007). Reaching out: Access to and use of banking services across countries. *Journal of Financial Economics 85*, 234–266.

Beck, T. and M. S. M. Peria (2010). Foreign bank participation and outreach: Evidence from Mexico. *Journal of Financial Intermediation 19*(1), 52–73.

Berger, A., N. Miller, M. Petersen, R. Rajan, and J. Stein (2005). Does function follow organizational form? Evidence from the lending practices of large and small banks. *Journal of Financial economics 76*(2), 237–269.

Berger, A. N., L. F. Klapper, and G. F. Udell (2001). The ability of banks to lend to informationally opaque small businesses. *Journal of Banking and Finance 25*, 2127–2167.

Berger, D. and J. Vavra (2015). Consumption dynamics during recessions. *Econometrica 83*(1), 101–154.

Bernanke, B. S. (2017). Federal reserve policy in an international context. *IMF Economic Review 65*(1), 5–36.

Bernoth, K. and J. v. Hagen (2004). The Euribor futures market: Efficiency and the impact of ecb policy announcements. *International finance 7*(1), 1–24.

Blanchard, O. (2012). Sustaining a global recovery. *Finance and Development 46*(3), 9–12.

Blanchard, O. and L. Summers (1986). Hysteresis and the European unemployment problem. *NBER Macroeconomic Annual 1*, 15–90.

Blundell, R. and S. Bond (1998a). Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics 87*(1), 115–143.

Blundell, R. and S. Bond (1998b). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* (87), 115–143.

Blundell, R., M. Browning, and C. Meghir (1994). Consumer demand and the life-cycle allocation of household expenditures. *The Review of Economic Studies 61*(1), 57–80.

Borensztein, E., J. De Gregorio, and J.-W. Lee (1998). How does foreign direct investment affect economic growth? *Journal of International Economics 45*(1), 115–135.

Borio, C. and H. Zhu (2012). Capital regulation, risk-taking and monetary policy: A missing link in the transmission mechanism? *Journal of Financial stability 8*(4), 236–251.

Boyd, J. H. and E. C. Prescott (1986). Financial intermediary-coalitions. *Journal of Economic theory 38*(2), 211–232.

Bräuning, F. and V. Ivashina (2017). Monetary policy and global banking. Working Paper.

Bräuning, F. and V. Ivashina (2018). US monetary policy and emerging market credit cycles. Working Paper.

Bruno, V. and H. S. Shin (2015a). Capital flows and the risk-taking channel of monetary policy. *Journal of Monetary Economics 71*, 119–132.

Bruno, V. and H. S. Shin (2015b). Cross-border banking and global liquidity. *Review of Economic Studies 82*(2), 535–564.

Caballero, R. J. (1990). Consumption puzzles and precautionary savings. *Journal of Monetary Economics 25*(1), 113–136.

Campbel, T. S. and W. A. Kracaw (1980). Information production, market signalling and the theory of financial intermediation. *The Journal of Finance 35*(4), 863–882.

Campbell, J. and A. Deaton (1989). Why is consumption so smooth. *Review of Economic Studies 56*, 357–374.

Campbell, J. Y., S. Giglio, and P. Pathak (2011). Forced sales and house prices. *American Economic Review 101*(5), 2108–31.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics 19*(1), 22–64.

Carroll, C. D. (1994). How does future income affect current consumption? *The Quarterly Journal of Economics*, 111–147.

Cerra, V. and S. C. Saxena (2008). Growth dynamics: The myth of economic recovery. *American Economic Review 98*(1), 439–457.

Cetorelli, N. and L. Goldberg (2012a). Banking globalization and monetary transmission. *Journal of Finance 67*(5), 1811–1843.

Cetorelli, N. and L. S. Goldberg (2012b). Liquidity management of US global banks: Internal capital markets in the great recession. *Journal of International Economics 88*(2), 299–311.

Chodorow-Reich, G., L. Feiveson, Z. Liscow, and W. G. Woolston (2012). Does state fiscal relief during recessions increase employment? Evidence from the American Recovery and Reinvestment Act. *American Economic Journal: Economic Policy 4*(3), 118–145.

Claessens, S. (2006). Competitive implications of cross-border banking. In G. Caprio, D. D. Evanoff, and G. G. Kaufman (Eds.), *Cross-Border Banking: Regulatory Challenges*, pp. 151–182. Hackensack, NJ: World Scientific Publishing.

Claessens, S., A. Demirgüç-Kunt, and H. Huizinga (2001). How does foreign entry affect domestic banking markets? *Journal of Banking and Finance 25*(5), 891–911.

Clarke, G. R., R. J. Cull, M. S. M. Peria, and S. M. Sanchez (2005). Bank lending to small businesses in Latin America: Does bank origin matter? *Journal of Money, Credit and Banking 37*(1), 83–118.

Cochrane, J. H. and M. Piazzesi (2002). The Fed and interest rates—A high-frequency identification. *American Economic Review 92*(2), 90–95.

Coen, R. (1973). Labor force and unemployment in the 1920's and 1930's: A reexamination based on postwar experience. *Review of Economics and Statistics 55*(1), 46–55.

Coibion, O., Y. Gorodnichenko, and G. H. Hong (2015). The cyclicality of sales, regular and effective prices: Business cycle and policy implications. *American Economic Review 105*(3), 993–1029.

Coimbra, N. and H. Rey (2017). Financial cycles with heterogeneous intermediaries. Working Paper.

Collin-Dufresne, P., M. Johannes, and L. A. Lochstoer (2016). Asset pricing when 'this time is different'. *Swiss Finance Institute Research Paper* (13-73), 14–8.

Cvijanovic, D. and C. Spaenjers (2015). Real estate as a luxury good: Non-resident demand and property prices in Paris.

Davis, S. J. and T. M. Von Wachter (2011). Recessions and the cost of job loss. Technical report, National Bureau of Economic Research.

De Blas, B. and K. N. Russ (2013). All banks great, small and global: Loan pricing and foreign competition. *International Review of Economics & Finance 26*, 4–24.

De Haas, R. and I. V. Lelyveld (2014). Multinational banks and the global financial crisis: Weathering the perfect storm? *Journal of Money, Credit and Banking 46*(1), 333–364.

De Haas, R. and N. Van Horen (2013). Running for the exit? International bank lending during a financial crisis. *The Review of Financial Studies 26*(1), 244–285.

De Meza, D. and D. C. Webb (1987). Too much investment: A problem of asymmetric information. *Quarterly Journal of Economics 102*(2), 281–292.

De Nardi, M., E. French, and D. Benson (2012). Consumption and the great recession. *Federal Reserve Bank of Chicago Economic Perspectives* (1), 1–16.

Deaton, A. (1991). Saving and liquidity constraints. *Econometrica 59*(5), 1221–1248.

Deaton, A. (1992). *Understanding Consumption*. Oxford: Clarendon Press.

Dedola, L., P. Karadi, and G. Lombardo (2013). Global implications of national unconventional policies. *Journal of Monetary Economics 60*, 66–85.

Dell'Ariccia, G. and R. Marquez (2004). Information and bank credit allocation. *Journal of Financial Economics 72*(1), 185–214.

Delong, B. and L. Summers (2012). Fiscal policy in a depressed economy. *Brookings Papers on Economic Activity 44*(1), 233–297.

Detragiache, E., T. Tressel, and P. Gupta (2008). Foreign banks in poor countries: Theory and evidence. *Journal of Finance 63*(5), 2123–2160.

Diamond, D. (1984). Financial intermediation and delegated monitoring. *Review of Economic Studies 51*(3), 393–414.

Dube, J.-P., G. J. Hitsch, and P. E. Rossi (2018). Income and wealth effects on private-label demand: Evidence from the great recession. *Marketing Science*.

Dynan, K. E. (2000). Habit formation in consumer preferences: Evidence from panel data. *American Economic Review 90*(3), 391–406.

Edison, H. J., R. Levine, L. Ricci, and T. Sløk (2002). International financial integration and economic growth,. *Journal of International Money and Finance 21*, 749–776.

Ehling, P., A. Graniero, and C. Heyerdahl-Larsen (2018, July). Asset prices and portfolio choice with learning from experience. *Review of Economic Studies 85*(1), 1752–1780.

Einav, L., E. Leibtag, and A. Nevo (2010). Recording discrepancies in nielsen homescan data: Are they present and do they matter? *Quantitative Marketing and Economics 2*(8), 207–239.

Favilukis, J. and S. Van Nieuwerburgh (2017). Out-of-town home buyers and city welfare. *Working Paper*.

Fischer, S. (1997). Capital account liberalization and the role of the IMF. In *Should the IMF Pursue Capital-Account Convertibility?*, Volume 19.

Fischer, S. (2015). The Federal Reserve and the global economy. *IMF Economic Review 63*(1), 8–21.

Flavin, M. (1993). The excess smoothness of consumption: Identification and interpretation. *Review of Economic Studies 60*(3), 651–666.

Friedman, M. (1957). *Theory of the Consumption Function*. Princeton University Press.

Fuchs-Schündeln, N. and M. Schündeln (2015). On the endogeneity of political preferences: Evidence from individual experience with democracy. *Science 347*(6226), 1145–1148.

Fuhrer, J. (2000). Habit formation in consumption and its implications for monetary-policy models. *American Economic Review 90*, 367–390.

Gabaix, X. and D. Laibson (2017). Myopia and discounting. Working paper.

Gertler, M. and P. Karadi (2011). A model of unconventional monetary policy. *Journal of Monetary Economics 58*(1), 17–34.

Gertler, M. and P. Karadi (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics 7*(1), 44–76.

Gertler, M. and N. Kiyotaki (2010). Financial intermediation and credit policy in business cycle analysis. In B. M. F. a. M. Woodford (Ed.), *Handbook of Monetary Economics*, Volume 3, pp. 547–599. Elsevier.

Glaeser, E., W. Huang, Y. Ma, and A. Shleifer (2017). A real estate boom with Chinese characteristics. *The Journal of Economic Perspectives 31*(1), 93–116.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2018). Bartik instruments: What, when, why, and how. *Working Paper*.

Gopinath, G. and J. C. Stein (2018). Banking, trade and the making of a dominant currency. Working Paper.

Gormley, T. A. (2010). The impact of foreign bank entry in emerging markets: Evidence from India. *Journal of Financial Intermediation 19*(1), 26–51.

Gormley, T. A. (2014). Costly information, entry and credit access. *Journal of Economic Theory 154*(633-667).

Gorodnichenko, Y. (2012). Using firm optimization to evaluate and estimate productivity and returns to scale. Working Paper.

Gorodnichenko, Y. and M. Weber (2016). Are sticky prices costly? Evidence from the stock market. *American Economic Review 106*(1), 165–99.

Gourinchas, P.-O. and M. Obstfeld (2012). Stories of the twentieth century for the twenty-first. *American Economic Journal: Macroeconomics 4*(1), 226–65.

Gourinchas, P.-O. and J. Parker (2002). Consumption over the life cycle. *Econometrica 70*(1), 787–817.

Gouskova, E. and R. Schoeni (2007). Comparing estimates of family income in the psid and the march current population survey, 1968–2005. *Panel Study of Income Dynamics Technical Paper Series. Michigan: Ann Arbor*.

Green, R. (1997). Follow the leader: How changes in residential and non-residential investment predict changes in GDP. *Real Estate Economics 25*(2), 253–70.

Greenland, A., J. Lopresti, and P. McHenry (2019). Import competition and internal migration. *Review of Economics and Statistics 101*(1), 44–59.

Greenwood, R. and S. Nagel (2009). Inexperienced investors and bubbles. *Journal of Financial Economics 93*, 239–258.

Haltmaier, J. (2012). Do recessions affect potential output? *Board of Governors of the Federal Reserve System Finance and Economics Discussion Series 1066*.

Hansen, A. H. (1939). Economic progress and declining population growth. *The American Economic Review 29*(1), 1–15.

Harris, C. and D. Laibson (2001). Dynamic choices of hyperbolic consumers. *Econometrica 64*(4), 935–957.

Harrison, A. E., I. Love, and M. S. McMillan (2004). Global capital flows and financing constraints. *Journal of development Economics 75*(1), 269–301.

Holtz-Eakin, D., W. Newey, and H. Rosen (1988). Estimating vector autoregressions with panel data. *Econometrica* (56), 1371–1395.

Huckfeldt, C. (2016). Understanding the scarring effect of recessions. Working paper.

Ivashina, V. (2009). Asymmetric information effects on loan spreads. *Journal of financial Economics 92*(2), 300–319.

Ivashina, V. and D. Scharfstein (2010). Bank lending during the financial crisis of 2008. *Journal of Financial Economics 97*(3), 319–338.

Ivashina, V., D. S. Scharfstein, and J. C. Stein (2015). Dollar funding and the lending behavior of global banks. *Quarterly Journal of Economics 130*(3), 1241–1281.

Jarosch, G. (2015). Searching for job security and the consequences of job loss. Working paper.

Jensen, S. T. and S. H. Shore (2015). Changes in the distribution of earnings volatility. *Journal of Human Resources 50*(3), 811–836.

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

Kahn, L. (2010). The long-term labor market consequences of graduating from college in a bad economy. *Labor Economics* (17), 303–316.

Kahneman, D. and A. Tversky (1974). Subjective probability: A judgment of representativeness. In *The Concept of Probability in Psychological Experiments*, pp. 25–48. Springer.

Kalemli-Ozcan, S., B. Sorensen, C. Villegas-Sanchez, V. Volosovych, and S. Yesiltas (2015). How to construct nationally representative firm level data from the ORBIS global database. Working Paper.

Kaplan, G., G. L. Violante, and J. Weidner (2014). The wealthy hand-to-mouth. Technical report, Brookings Papers on Economic Activity.

Kaustia, M. and S. Knüpfer (2008). Do investors overweight personal experience? Evidence from IPO subscriptions. *Journal of Finance 63*, 2679–2702.

Kerr, W. R. (2008a). The ethnic composition of US inventors. *Working Paper*.

Kerr, W. R. (2008b). Ethnic scientific communities and international technology diffusion. *The Review of Economics and Statistics 90*(3), 518–537.

Kerr, W. R. and W. F. Lincoln (2010). The supply side of innovation: H-1B visa reforms and US ethnic invention. *Journal of Labor Economics 28*(3), 473–508.

Kuttner, K. N. (2001). Monetary policy surprises and interest rates: Evidence from the Fed funds futures market. *Journal of monetary economics 47*(3), 523–544.

Laudenbach, C., U. Malmendier, and A. Niessen-Ruenzi (2018). The long-lasting effects of experiencing communism on financial risk-taking. Working Paper.

Lebergott, S. (1957). Annual estimates of unemployment in the united states, 1900-1954. In NBER (Ed.), *The Measurement and Behavior of Unemployment*. Princeton, NJ: Princeton University Press.

Lebergott, S. (1964). *Manpower in Economic Growth: The American Record since 1800*. New York: McGraw-Hill.

Levine, R. (1996). Foreign banks, financial development and economic growth. In C. E. Barfield (Ed.), *International Financial Markets: Harmonization versus Competition*, pp. 224–254. Washington, DC: AEI Press.

Levinsohn, J. and A. Petrin (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies 70*(2), 317–342.

Liberti, J. M. and M. A. Petersen (2018). Information: Hard and soft. Working Paper.

Lichter, A., M. Löffler, and S. Siegloch (2016). The long-term costs of government surveillance: Insights from stasi spying in East Germany. Working Paper.

Low, H., C. Meghir, and L. Pistaferri (2010). Wage risk and employment risk over the life cycle. *American Economic Review 100*(4), 1432–67.

Maggiori, M., B. Neiman, and J. Schreger (2018). International currencies and capital allocation. Working Paper.

Malmendier, U. and S. Nagel (2011). Depression babies: Do macroeconomic experiences affect risk-taking? *Quarterly Journal of Economics 126*, 373–416.

Malmendier, U. and S. Nagel (2015). Learning from inflation experiences. *Quarterly Journal of Economics 131*(1), 53–87.

Malmendier, U., S. Nagel, and Z. Yan (2018). The making of hawks and doves. *Working Paper*.

Malmendier, U., D. Pouzo, and V. Vanasco (2018). Investor experiences and financial market dynamics. Working paper.

Malmendier, U. and G. Tate (2005). CEO overconfidence and corporate investment. *Journal of Finance 60*(6), 2661–2700.

Malmendier, U., G. Tate, and J. Yan (2011). Overconfidence and early-life experiences: The effect of managerial traits on corporate financial policies. *Journal of Finance 66*(5), 1687–1733.

Meghir, C. and L. Pistaferri (2004). Income variance dynamics and heterogeneity. *Econometrica 72*(1), 1–32.

Meghir, C. and G. Weber (1996). Intertemporal nonseparability or borrowing restrictions? a disaggregate analysis using a u.s. consumption panel. *Econometrica 64*(5), pp. 1151–1181.

Mian, A. (2006). Distance constraints: The limits of foreign lending in poor economies. *Journal of Finance 61*(3), 1465–1505.

Mian, A., K. Rao, and A. Sufi (2013a). Household balance sheets, consumption, and the economic slump. *Quarterly Journal of Economics 128*(4), 1687–1726.

Mian, A., K. Rao, and A. Sufi (2013b). Household balance sheets, consumption, and the economic slump. *The Quarterly Journal of Economics 128*(4), 1687–1726.

Mian, A. and A. Sufi (2014). What explains the 2007–2009 drop in employment? *Econometrica 82*(6), 2197–2223.

Mian, A., A. Sufi, and F. Trebbi (2015). Foreclosures, house prices, and the real economy. *The Journal of Finance 70*(6), 2587–2634.

Miranda-Agrippino, S. and H. Rey (2018). US monetary policy and the global financial cycle. Working Paper.

Modigliani, F. and R. Brumberg (1954). Utility analysis and the consumption function: An interpretation of cross-section data. In K. Kufihara (Ed.), *Post-Keynesian Economics*, pp. 128–197. Rutgers University Press.

Morais, B., J.-L. Peydró, J. Roldan, and C. Ruiz (2018). The international bank lending channel of monetary policy rates and quantitative easing: Credit supply, reach-for-yield and real effects. *Journal of Finance*, Forthcoming.

Nakamura, E. and J. Steinsson (2014). Fiscal stimulus in a monetary union: Evidence from us regions. *The American Economic Review 104*(3), 753–792.

Nakamura, E. and J. Steinsson (2018). High frequency identification of monetary non-neutrality: The information effect. *Quarterly Journal of Economics 133*(3), 1283–1330.

Nevo, A. and A. Wong (2015). The elasticity of substitution between time and market goods: Evidence from the great recession. Working paper.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, 1417–1426.

Niepmann, F. (2015). Banking across borders. *Journal of International Economics 96*(2), 244–265.

Olafsson, A. and M. Pagel (2018). The retirement-consumption puzzle: New evidence on individual spending and financial structure. Working paper.

Olley, S. and A. Pakes (1996). Dynamics of productivity in the telecommunications equipment. *Econometrica 64*(6), 1263–1297.

Ongena, S., I. Schindele, and D. Vonnák. (2017). In lands of foreign currency credit, bank lending channels run through? Working Paper.

Oreopoulos, P., T. von Wachter, and A. Heisz (2012). The short- and long-term career effects of graduating in a recession: Hysteresis and heterogeneity in the market for college graduates. *American Economic Journal: Applied Economics* (4), 1–29.

Oyer, P. (2008). The making of an investment banker: Stock market shocks, career choice, and lifetime income. *The Journal of Finance 63*(6), 2601–2628.

Pagel, M. (2017). Expectations-based reference-dependent life-cycle consumption. *Review of Economic Studies 84*(2), 885–934.

Parker, J. (2000). Real estate and the macroeconomy: Discussions. *Brookings Papers on Economic Activity 2*, 150–162.

Parker, J. A., N. S. Souleles, D. S. Johnson, and R. McClelland (2013). Consumer spending and the economic stimulus payments of 2008. *American Economic Review 103*(6), 2530–2553.

Petersen, M. and R. G. Rajan (1994). The benefits of lending relationships: Evidence from small business data. *Journal of Finance 49*(1), 3–37.

Petersen, M. A. and R. G. Rajan (2002). Does distance still matter? The information revolution in small business lending. *The journal of Finance 57*(6), 2533–2570.

Petev, I., L. Pistaferri, and I. Saporta-Eksten (2011). An analysis of trends, perceptions, and distributional effects in consumption. In B. W. David B. Grusky and C. Wimer (Eds.), *The Great Recession*, Handbook of the Economics of Finance, pp. 161–195. Russell Sage Foundation.

Pfeffer, F. T., R. F. Schoeni, A. Kennickell, and P. Andreski (2016). Measuring wealth and wealth inequality: Comparing two us surveys. *Journal of economic and social measurement 41*(2), 103–120.

Pistaferri, L. (2016). Why has consumption grown only moderately after the Great Recession? Working paper.

Popov, A. and G. F. Udell (2012). Cross-border banking, credit access and the financial crisis. *Journal of international economics 87*(1), 147–161.

Rajan, R. (2015). Competitive monetary easing: Is it yesterday once more? *Macroeconomics and Finance in Emerging Market Economies 8*(1-2), 5–16.

Ramakrishnan, R. TS, and A. V. Thakor (1984). Information reliability and a theory of financial intermediation. *Review of Economic Studies 51*(3), 415–432.

Ramey, V. A. (2011). Can government purchases stimulate the economy? *Journal of Economic Literature 49*(3), 673–685.

Ranaldo, A. and E. Rossi (2010). The reaction of asset markets to swiss national bank communication. *Journal of International Money and Finance 29*(3), 486–503.

Reifschneider, D., W. L. Wascher, and D. Wilcox (2015). Aggregate supply in the united states: Recent developments and implications for the conduct of monetary policy. *IMF Economic Review 63*(1), 71–109.

Reinhart, C. and K. Rogoff (2009). The aftermath of financial crises. *American Economic Review 99*(2), 466–472.

Rey, H. (2016). International channels of transmission of monetary policy and the Mundellian Trilemma. *IMF Economic Review 64*(1), 6–35.

Rigobon, R. and B. Sack (2004). The impact of monetary policy on asset prices. *Journal of Monetary Economics 51*(8), 1553–1575.

Rodrik, D. (1998). Who needs capital-account convertibility? *Essays in international finance*, 55–65.

Romer, C. (1986). Spurious volatility in historical unemployment data. *Journal of Political Economy 94*(1), 1–37.

Rosa, C. and G. Verga (2008). The impact of central bank announcements on asset prices in real time. *International Journal of Central Banking 4*(2), 175–217.

Rosen, K., A. Margon, R. Sakamoto, and J. Taylor (2017). Breaking Ground: Chinese Investment in US Real Estate. Fisher Center Working Papers.

Rosenbaum, P. and D. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Journal of Business and Economic Statistics 70*, 41–55.

Sa, F. (2016). The effect of foreign investors on local housing markets: Evidence from the UK. Working Paper.

Schnabl, P. (2012). The international transmission of bank liquidity shocks: Evidence from an emerging market. *Journal of Finance 67*(3), 897–932.

Schraeder, S. (2015). Information processing and non-bayesian learning in financial markets. *Review of Finance*, 1–31.

Sengupta, R. (2007). Foreign entry and bank competition. *Journal of Financial Economics 84*(2), 502–528.

Simons, R. A., J. Wu, J. Xu, and Y. Fei (2016). Chinese Investment in US Real Estate Markets Using the EB-5 Program. *Economic Development Quarterly 30*(1), 75–87.

Stein, J. (2002). Information production and capital allocation: Decentralized versus hierarchical firms. *Journal of Finance 57*(5), 1891–1921.

Stein, J. (2014). Incorporating financial stability considerations into a monetary policy framework: A speech at the international research forum on monetary policy.

Stiglitz, J. E. (2002). *Globalization and its Discontents*, Volume 500. New York Norton.

Stiglitz, J. E. and A. Weiss (1981). Credit rationing in markets with imperfect information. *American Economics Review 71*(3), 393–410.

Stock, J. H. and M. W. Watson (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics 20*(2), 147–162.

Stroebel, J. and J. Vavra (2017). House prices, local demand, and retail prices. *Journal of Political Economy*. Forthcoming.

Sufi, A. (2007). Information asymmetry and financing arrangements: Evidence from syndicated loans. *The Journal of Finance 62*(2), 629–668.

Summers, L. H. (2000). International financial crises: causes, prevention, and cures. *American Economic Review 90*(2), 1–16.

Summers, L. H. (2014a). Reflections on the "new secular stagnation hypothesis". In C. Teulings and R. Baldwin (Eds.), *Secular Stagnation: Facts, Causes and Cures*, pp. 27–40. A VoxEU.org eBook.

Summers, L. H. (2014b). US economic prospects: Secular stagnation, hysteresis, and the zero lower bound. *Business Economics 49*(2), 65–73.

Tong, H. and S.-J. Wei (2010). The composition matters: capital inflows and liquidity crunch during a global economic crisis. *The Review of Financial Studies 24*(6), 2023–2052.

Tversky, A. and D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases. *Science 185*(4157), 1124–1131.

Vissing-Jorgensen, A. (2003). Perspectives on behavioral finance: Does "irrationality" disappear with wealth? evidence from expectations and actions. In *NBER Macroeconomics Annual*.

West, K. D. (1989). The insensitivity of consumption to news about income. *Journal of Monetary Economics 21*, 17–23.

Wong, A. (2018). Transmission of monetary policy to consumption and population aging. Working Paper.

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton: Princeton University Press.

Woodford, M. (2010). Optimal monetary stabilization policy. Volume 3 of *Handbook of Monetary Economics*, pp. 723 – 828. Elsevier.

Zeldes, S. P. (1989). Consumption and liquidity constraints: An empirical investigation. *Journal of Political Economy 97*(2), 305–346.

Appendix A

Appendix for Chapter 1

A.1 Model Solutions and Proofs

Equilibrium Solution

Propositions 1–3 lead to a full characterization of the equilibrium solution on $\mathcal{R}^G$ and $\mathcal{R}^L$. Based on these characterizations, I solve for the equilibrium interest rates $\mathcal{R}^G(z_i^G)$ and $\mathcal{R}^L(z_i^L)$, and thresholds $\bar{z}_i^L = \bar{z}^L(z_i^G)$ and $\bar{z}_i^G = \bar{z}^G(z_i^L)$, for $z_i^G \in [\underline{z}^G, 1]$ and $z_i^L \in [\underline{z}^L, 1]$ as follows.

First, let $\mathbb{R}^G(z_i^G, \bar{z}_i^L)$ and $\mathbb{R}^L(z_i^L, \bar{z}_i^G)$ be the implicit functions which give the rate at which each banks' expected profit (Equation (1.3a) and (1.3b)) would be zero for a given observed component combined with a given threshold on the unobserved component[1]:

$$\mathbb{R}^G(z_i^G, \bar{z}_i^L) = \mathcal{R}^G(z_i^G) \text{ s.t. } E_G[\pi_G(z_i^G, \bar{z}_i^L, \mathcal{R}^G(z_i^G))] = 0;$$
$$\mathbb{R}^L(z_i^L, \bar{z}_i^G) = \mathcal{R}^L(z_i^L) \text{ s.t. } E_L[\pi_L(z_i^L, \bar{z}_i^G, \mathcal{R}^L(z_i^L))] = 0.$$

Based on Proposition 2, for each given z_i^G, the corresponding threshold $\bar{z}_i^L$ is the z_i^L for the firm (z_i^G, z_i^L) for which $\mathcal{R}^L(z_i^L)$=$\mathcal{R}^G(z_i^G)$. By symmetry, $\bar{z}^G(\bar{z}_i^L) = z_i^G$. Therefore, the equilibrium rate $\mathcal{R}^G(z_i^G)$ and threshold $\bar{z}_i^L$ are the solutions to the system of equations:

$$\begin{aligned} \mathbb{R}^G(z_i^G, \bar{z}_i^L) &= \mathbb{R}^L(\bar{z}_i^L, z_i^G). \\ \mathcal{R}^G(z_i^G) &= \mathbb{R}^G(z_i^G, \bar{z}_i^L). \end{aligned} \tag{A.1}$$

Similarly, for each given z_i^L, the equilibrium rate $\mathcal{R}^L(z_i^L)$ and threshold $\bar{z}_i^G$ are the solutions to the system of equations:

$$\begin{aligned} \mathbb{R}^L(z_i^L, \bar{z}_i^G) &= \mathbb{R}^G(\bar{z}_i^G, z_i^L). \\ \mathcal{R}^L(z_i^L) &= \mathbb{R}^L(z_i^L, \bar{z}_i^G). \end{aligned} \tag{A.2}$$

[1] The implicit equations are fully written out in the appendix as Equations (A.5a) and (A.5b).

Furthermore, I apply Proposition 2 to solve for $\underline{z}^G$ and $\underline{z}^L$, the cut-offs below which the expected profits of the firms are too low for the global bank and local bank to break even in expectation, regardless of the rate charged. At these cut-off points, the maximum expected profits of the banks are zero, all firms default given the equilibrium interest rates. The next lemma establishes that the cut-offs $\underline{z}^G$ and $\underline{z}^L$ are thresholds to each other.

Lemma 2 $\underline{z}^G = \bar{z}^G(\underline{z}^L)$*, and* $\underline{z}^L = \bar{z}^L(\underline{z}^G)$.

Given Lemma 2, $\underline{z}^G$ and $\underline{z}^G$ are the solutions to the system of equations:

$$\int_0^{\underline{z}^L}\int_0^1 (\underline{z}_i^G + z_i^L + u_i)\ dF_{\underline{z}^L}(u_i, z_i^L) = r^G;$$
$$\int_0^{\underline{z}^G}\int_0^1 (z_i^G + \underline{z}_i^L + u_i)\ dF_{\underline{z}^G}(u_i, z_i^G) = r^L.$$

where $F_{\underline{z}^L}(.)$ and $F_{\underline{z}^G}(.)$ denote the cumulative distribution function of the relevant variable conditional on $z_i^L \leq \underline{z}^L$ and $z_i^G \leq \underline{z}^G$, respectively. The solutions to this system is:

$$\underline{z}^G = \frac{1}{3}(4r^G - 2r^L - 1) \qquad \text{and} \qquad \underline{z}^L = \frac{1}{3}(4r^L - 2r^G - 1). \tag{A.3}$$

The bounds $\underline{z}^G$ and $\underline{z}^L$ define the cut-offs on z_i^G and z_i^L, respectively, below which global banks and local banks would not make loans. They are increasing in the banks' own funding cost and decreasing in the funding cost faced by the other bank type. In other words, facing higher funding cost induces the respective banks to be more restrictive on the riskiest firm to which they lend, while higher funding cost faced by the other bank type induces them to lend to riskier firms. Interestingly, each banks' own funding cost has a stronger effect on the respective lower bound than the other banks' funding cost. Figure 1.7 illustrates the cut-offs $\underline{z}^G$ and $\underline{z}^L$ in a space that summarizes all the firms in the economy. Given the cut-offs, firms in Region A are not offered loans. Firms in Region B can only receive loans from local banks, and firms in Region C can only receive loans from global banks.

Proofs

Proof of Proposition 1. Based on Equations (1.2a) and (1.3a), $\mathcal{R}^G(z_i^G)$ is given implicitly by the global bank's expected profit function:

$$\begin{aligned} E_G[\pi_G(z_i^G)] =& \left[\int_{G_c} \left(\int_{G_a} (z_i^G + z_i^L + u_i)\ dF(u_i) + \int_{G_b} \mathcal{R}^G(z_i^G)\ dF(u_i) \right) dF(z_i^L) \right] - r^G = 0 \\ \text{where}\quad G_a =& \left\{ u_i \mid 0 \leq u_i < \min(\max(0, \mathcal{R}^G(z_i^G) - z_i^G - z_i^L), 1) \right\} \\ G_b =& \left\{ u_i \mid \ \min(\max(0, \mathcal{R}^G(z_i^G) - z_i^G - z_i^L), 1) \leq u_i \leq 1 \right\} \\ G_c =& \left\{ z_i^L \mid z_i^L\colon (z_i^G, z_i^L) \in S_G \right\} \end{aligned} \tag{A.4}$$

Equation (A.4) can be decomposed into two regions over z_i^G:

1. No loans: z_i^G such that $z_i^G + E_G[z_i^L \mid (z_i^G, z_i^L) \in S^G] + 1/2 < r^G$.
2. Loans: z_i^G such that $z_i^G + E_G[z_i^L \mid (z_i^G, z_i^L) \in S^G] + 1/2 \geq r^G$.

Equilibrium rates $\mathcal{R}^G(z_i^G)$ are defined in region 2.

Analyzing $\frac{\partial E_G[\pi_G(z_i^G)]}{\partial z_i^G}$: An increase in z_i^G lowers the probability of default and increases the bank's expected return. Thus $\frac{\partial E_G[\pi_G(z_i^G)]}{\partial z_i^G} > 0\ \forall z_i^G$.

Given that, I first prove that $\mathcal{R}^G$ is weakly decreasing in z_i^G. Assume otherwise: there exists $z_j^G > z_i^G$ such that $\mathcal{R}^G(z_j^G) > \mathcal{R}^G(z_i^G)$. Given perfect competition with free entry, $E[\pi_G(z_i^G)] = 0$ for $\mathcal{R}^G(z_i^G)$. Because $\frac{\partial E_G[\pi_G(z_i^G)]}{\partial z_i^G} > 0$, another global bank could offer at most the same $\mathcal{R}^G(z_i^G)$ for z_j^G and at least break even. Therefore, it could offer $\mathcal{R}^G(z_j^G) \leq \mathcal{R}^G(z_i^G)$, which is a contradiction. $\mathcal{R}^L$ is similarly weakly decreasing in z_i^L.

Analyzing $\frac{\partial E_G[z_i^L | (z_i^G, z_i^L) \in S^G]}{\partial \mathcal{R}^G(z_i^G)}$: An increase in the rate $\mathcal{R}^G(z_i^G)$ may cause some marginal values of z_i^L to switch from selecting the global to the local bank. Since both $\mathcal{R}^G(z_i^G)$ and $\mathcal{R}^L(z_i^L)$ are non-increasing, those that do will be those with the lowest $\mathcal{R}^L(z_i^L)$ and therefore the highest z_i^L, lowering the expected value of z_i^L over firms which select the global bank. Therefore, $\frac{\partial E_G[z_i^L | (z_i^G, z_i^L) \in S^G]}{\partial \mathcal{R}^G(z_i^G)} \leq 0$.

Analyzing $\frac{\partial E_G[\pi_G(z_i^G)]}{\partial \mathcal{R}^G(z_i^G)}$: An increase in $\mathcal{R}^G(z_i^G)$ drives the expected return to the global bank through two effects:

1. It increases the return in all outcomes where previously there was no default.
2. It decreases the expected value of z_i^L for firms which will select the global bank, which decreases the expected return in case of default.

Absent other constraints, at any point, $\frac{\partial E_G[\pi_G(z_i^G)]}{\partial \mathcal{R}^G(z_i^G)}$ could be dominated by either term and be positive, negative, or zero.

Now I prove that $\mathcal{R}^G$ is strictly decreasing in z_i^G (where loans are made, in region 2). Assume otherwise: there exists $z_j^G > z_i^G$ such that $\mathcal{R}^G(z_j^G) \geq \mathcal{R}^G(z_i^G)$. Consider again the perfect competition and free entry among global banks. $E_G[\pi_G(z_i^G)] = 0$ for $\mathcal{R}^G(z_i^G)$. Because $\frac{\partial E_G[\pi_G(z_i^G)]}{\partial z_i^G} > 0$, if $\mathcal{R}^G(z_j^G) = \mathcal{R}^G(z_i^G)$ there would be excess profit: $E_G[\pi_G(z_j^G)] > 0$. Regardless of the sign of $\frac{\partial E_G[\pi_G(z_i^G)]}{\partial \mathcal{R}^G(z_i^G)}$, another bank could charge a lower rate $\mathcal{R}^G(z_j^G)$ without losing money in expectation:

- If $\frac{\partial E_G[\pi_G(z_i^G)]}{\partial \mathcal{R}^G(z_i^G)} \leq 0$, decreasing the rate would leave profit unchanged or *increased* and clearly be possible.
- If $\frac{\partial E_G[\pi_G(z_i^G)]}{\partial \mathcal{R}^G(z_i^G)} > 0$, a competing global bank could trade the excess profit to offer a lower rate and capture the market while still at least breaking even.

Therefore $\mathcal{R}^G(z_j^G) < \mathcal{R}^G(z_i^G)$, which is a contradiction.

The proof that $\mathcal{R}^L$ is strictly decreasing in z_i^L is entirely analogous.

Further analysis. Consider the two effects which drive $\frac{\partial E_G[\pi_G(z_i^G)]}{\partial \mathcal{R}^G(z_i^G)}$. The first is trivially continuous. The second is continuous because $\mathcal{R}^L$ being strictly decreasing means that differential changes in $\mathcal{R}^G(z_i^G)$ cannot have discontinuous effects on selection S^G.

Consider also the implicit function of $\mathcal{R}^G(z_i^G)$ where the the bank profit is zero: $E_G[\pi_G(z_i^G)] = 0$. By the implicit function theorem, $\frac{d\mathcal{R}^G(z_i^G)}{dz_i^G} = -\frac{\partial E_G[\pi_G]}{\partial z_i^G} / \frac{\partial E_G[\pi_G(z_i^G)]}{\partial \mathcal{R}^G(z_i^G)}$. We know that $\frac{d\mathcal{R}^G(z_i^G)}{dz_i^G} < 0$ ($\mathcal{R}^G$ is strictly decreasing) and $\frac{\partial E_G[\pi_G]}{\partial z_i^G} > 0$. Therefore, $\frac{\partial E_G[\pi_G(z_i^G)]}{\partial \mathcal{R}^G(z_i^G)} > 0$, and the positive profit effect of increasing $\mathcal{R}^G(z_i^G)$ dominates the negative selection effect.

Finally, considering the regions over z_i^G, the boundary between the two regions occurs when $z_i^G + E_G[z_i^L \mid (z_i^G, z_i^L) \in S^G] + 1/2 = r^G$. Since $\frac{\partial E_G[z_i^L \mid (z_i^G, z_i^L) \in S^G]}{\partial \mathcal{R}^G(z_i^G)} < 0$ and $\frac{d\mathcal{R}^G(z_i^G)}{dz_i^G} < 0$, $E_G[z_i^L \mid (z_i^G, z_i^L) \in S^G]$ is increasing in z_i^G. Therefore there is a unique $\underline{z}_i^G = r^G - E_G[z_i^L \mid (\underline{z}_i^G, z_i^L) \in S^G] - 1/2$. Equilibrium rates $\mathcal{R}^G(z_i^G)$ are defined for all $\underline{z}_i^G \leq z_i^G \leq 1$.

All analyses apply to the analogous terms for local banks.

Proof of Proposition 2.

1) In an equilibrium market configuration that supports both types of banks, there must exist a set of marginal firms that are indifferent between the contracts by global banks and local banks, which occur when $\mathcal{R}^G(z_i^G) = \mathcal{R}^L(z_i^L)$. Let $f(z_i^G, z_i^L) = \mathcal{R}^G(z_i^G) - \mathcal{R}^L(z_i^L) = 0$. By Proposition 1, $\frac{\partial f(z_i^G, z_i^L)}{\partial z_i^L} = -\frac{\partial \mathcal{R}^L(z_i^L)}{\partial z_i^L} > 0$ for $z_i^L \in [\underline{z}_i^L, 1]$. By the

implicit function theorem, for each $z_i^G \in [\underline{z}_i^G, 1]$, there exists a threshold function $\bar{z}^L$: $z_i^G \mapsto \bar{z}_i^L$, such that $\mathcal{R}^G(z_i^G) = \mathcal{R}^L(\bar{z}_i^L)$.

The proof on the existence of a threshold function $\bar{z}^G$: $z_i^L \mapsto \bar{z}_i^G$ such that $\mathcal{R}^L(z_i^L) = \mathcal{R}^G(\bar{z}_i^G)$ is analogous.

2) Consider a marginal firm that faces $\mathcal{R}^G(z_i^G) = \mathcal{R}^L(z_i^L)$. As z_i^L decreases, $\mathcal{R}^L(z_i^L)$ increases by Proposition 1, while $\mathcal{R}^G(z_i^G)$ remains constant. Since now $\mathcal{R}^L(z_i^L) > \mathcal{R}^G(z_i^G)$, those firms would select a global bank. Therefore, firms with $z_i^L < \bar{z}_i^L \in S^G$. Conversely, as z_i^L increases, $\mathcal{R}^L(z_i^L)$ decreases by Proposition 1, while $\mathcal{R}^G(z_i^G)$ remains constant. Since $\mathcal{R}^L(z_i^L) < \mathcal{R}^G(z_i^G)$, those firms would select a local bank. Therefore, $S^G = \{(z_i^G, z_i^L) : z_i^L \leq \bar{z}^L(z_i^G)\}$, and $S^L = \{(z_i^G, z_i^L) : z_i^L > \bar{z}^L(z_i^G)\}$

The proof that $S^L = \{(z_i^G, z_i^L) : z_i^G < \bar{z}^G(z_i^L)\}$ and $S^G = \{(z_i^G, z_i^L) : z_i^G \geq \bar{z}^G(z_i^L)\}$ is analogous.

Proof of Proposition 3. The equilibrium interest rate functions are solution to the bank expected profits equations subject to zero profits conditions and firm selection:

$$\begin{aligned} E_G[\pi_G(z_i^G)] =& \left[\int_{G_c} \left(\int_{G_a} (z_i^G + z_i^L + u_i)\ dF(u_i) + \int_{G_b} R^G(z_i^G)\ dF(u_i) \right) dF(z_i^L) \right] - r^G = 0, \\ \text{where}\ \ G_a =& \left\{ u_i \ \middle|\ 0 \leq u_i < \min(\max(0, R^G(z_i^G) - z_i^G - z_i^L), 1) \right\}, \\ G_b =& \left\{ u_i \ \middle|\ \min(\max(0, R^G(z_i^G) - z_i^G - z_i^L), 1) \leq u_i \leq 1 \right\}, \\ G_c =& \left\{ z_i^L \ \middle|\ 0 < z_i^L \leq \bar{z}^L(z_i^G)) \right\}; \end{aligned} \tag{A.5a}$$

$$\begin{aligned} E_L[\pi_L(z_i^L)] =& \left[\int_{L_c} \left(\int_{L_a} (z_i^G + z_i^L + u_i)\ dF(u_i) + \int_{L_b} R^L(z_i^L)\ dF(u_i) \right) dF(z_i^G) \right] - r^L = 0, \\ \text{where}\ \ L_a =& \left\{ u_i \ \middle|\ 0 \leq u_i < \min(\max(0, R^L(z_i^L) - z_i^G - z_i^L), 1) \right\}, \\ L_b =& \left\{ u_i \ \middle|\ \min(\max(0, R^L(z_i^L) - z_i^G - z_i^L), 1) \leq u_i \leq 1 \right\}, \\ L_c =& \left\{ z_i^G \ \middle|\ 0 < z_i^G \leq \bar{z}^G(z_i^L)) \right\}. \end{aligned} \tag{A.5b}$$

Analyzing $\frac{\partial E_G[\pi_G(z_i^G)]}{\partial \mathcal{R}^L(\bar{z}^L(z_i^G))}$: An increase in $\mathcal{R}^L(\bar{z}^L(z_i^G))$ shifts marginal firms from the local to global bank at $(z_i^G, \bar{z}^L(z_i^G))$. This increases the threshold value $\bar{z}^L(z_i^G)$ at z_i^G. As a result, the expected profit of the global bank increases, all else held constant, so $\frac{\partial E_G[\pi_G(z_i^G)]}{\partial \mathcal{R}^L(\bar{z}^L(z_i^G))} > 0$.

The analysis that $\frac{\partial E_G[\pi_G(z_i^G)]}{\partial \mathcal{R}^G(z_i^G)} > 0$ is outlined in the proof for Proposition 1.

By the implicit function theorem, $\frac{d\mathcal{R}^G(z_i^G)}{d\mathcal{R}^L(\bar{z}^L(z_i^G))} = -\frac{\partial E_G[\pi_G(z_i^G)]}{\partial \mathcal{R}^L(\bar{z}^L(z_i^G))} / \frac{\partial E_G[\pi_G(z_i^G)]}{\partial \mathcal{R}^G(z_i^G)} < 0$.

Proof of Lemma 2. At $\underline{z}^G$, the equilibrium rate $\mathcal{R}^G(\underline{z}^G)$ is such that all firms which approach global banks default: $\mathcal{R}^G(\underline{z}^G) = \underline{z}^G + \bar{z}^L(\underline{z}^G) + 1$. Similarly at $\underline{z}^L$, $\mathcal{R}^L(\underline{z}^L) = \bar{z}^G(\underline{z}^L) + \underline{z}^L + 1$. It is clear that at least one entry $\underline{z}^j$ must be the threshold for the other $\underline{z}^k$: $\bar{z}^j(\underline{z}^k) = \underline{z}^j$.

Without loss of generality, let $j = G$ and $k = L$: $\bar{z}^G(\underline{z}^L) = \underline{z}^G$. Assume otherwise, $\bar{z}^L(\underline{z}^G) > \underline{z}^L$. Given $\bar{z}^G(\underline{z}^L) = \underline{z}^G$, $\mathcal{R}^L(\underline{z}^L) = \underline{z}^G + \underline{z}^L + 1$. It follows $\mathcal{R}^G(\underline{z}^G) = \underline{z}^G + \bar{z}^L(\underline{z}^G) + 1 > \underline{z}^G + \underline{z}^L + 1 = \mathcal{R}^L(\underline{z}^L)$. This implies $\mathcal{R}^L(\bar{z}^L(\underline{z}^G)) > \mathcal{R}^L(\underline{z}^L)$, which contradicts the strict monotonicity of $\mathcal{R}^L$. At the same time, $\bar{z}^L(\underline{z}^G) < \underline{z}^L$ is a contradiction, since local banks make no loans to firms with $z_i^L < \underline{z}^L$ by definition. Therefore, $\underline{z}^L = \bar{z}^L(\underline{z}^G)$.

The proof that $\underline{z}^G = \bar{z}^G(\underline{z}^L)$ is analogous.

Proof of Lemma 1. Let $r^G = r^L$. The expected profit equations for global banks and local banks subject to the break even conditions and firm selection, given by Equations (A.5a) and (A.5b), respectively, are symmetric. The result that $\bar{z}^L(z_i^G) = z_i^G$ and $\bar{z}^G(z_i^L) = z_i^L$ follows.

Proof of Corollary 2. Let $r^G = r^L$. Assume firm i selects into borrowing from a global bank. Based on firm selection criteria from Equations (1.2a) and 1.2b and Assumption 1, $\mathcal{R}^G(z_i^G) \leq \mathcal{R}^L(z_i^L)$, which implies $z_i^G \geq z_i^L$ by Proposition 1 and Lemma 1. Now assume $z_i^G \geq z_i^L$. Based on Equations (A.5a) and (A.5b), $R_{Gi}(z_i^G) \leq R_{Li}(z_i^L)$, which implies firm i selects into borrowing from a global bank.

The proof that a firm selects a local bank if and only if $z_i^L > z_i^G$ is analogous.

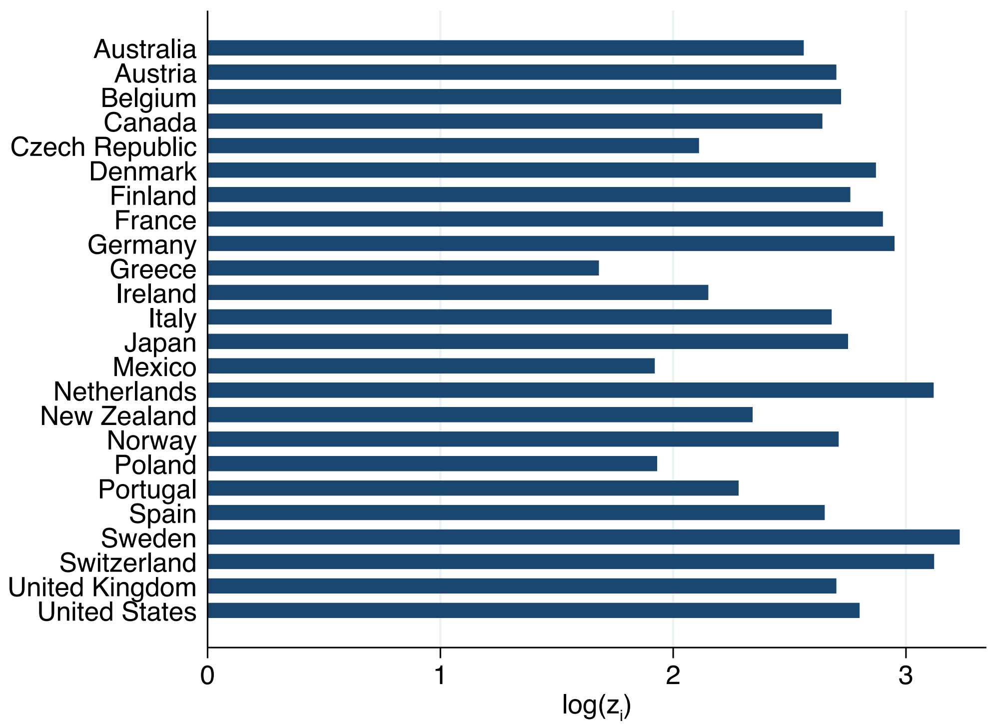
A.2 Additional Figures and Tables

Table A.1: **Summary Statistics: Loan and Firm Count by Country (Method 2)**

Country	Loan	GB	LB	Firm	Country	Loan	GB	LB	Firm
Australia	4507	0.70	0.30	701	Japan	21341	0.45	0.55	2865
Austria	387	0.53	0.47	61	Mexico	601	0.70	0.30	137
Belgium	704	0.69	0.31	123	Netherlands	2028	0.54	0.46	406
Canada	6760	0.64	0.36	903	New Zealand	1023	0.70	0.30	127
Czech Republic	197	0.68	0.32	77	Norway	1017	0.66	0.34	253
Denmark	327	0.56	0.44	84	Poland	318	0.54	0.46	87
Finland	587	0.65	0.35	113	Portugal	254	0.65	0.35	64
France	5876	0.67	0.33	996	Spain	4380	0.68	0.32	839
Germany	5987	0.68	0.32	942	Sweden	875	0.66	0.34	190
Greece	309	0.66	0.34	47	Switzerland	790	0.69	0.31	175
Ireland	404	0.70	0.30	107	UK	6810	0.69	0.31	1528
Italy	2378	0.67	0.33	688	US	46732	0.70	0.30	1466

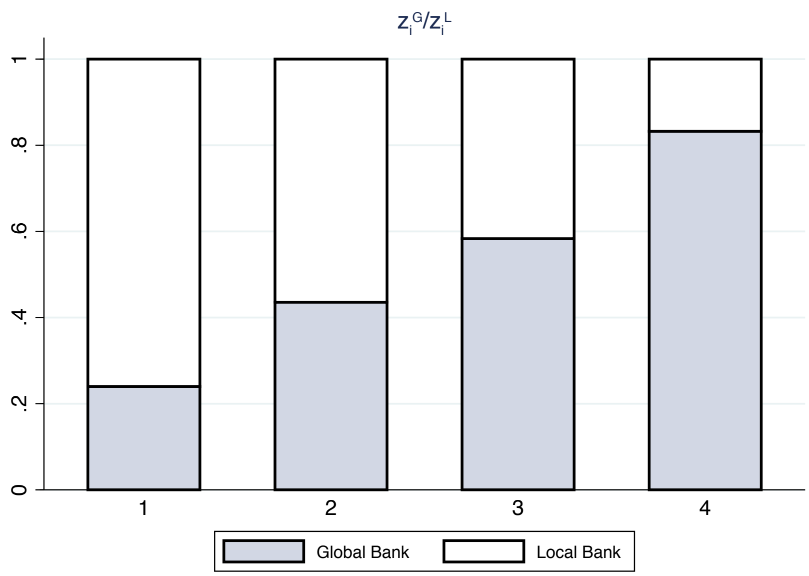
Notes. Sample constructed from Dealscan, Amadeus, Orbis, Compustat, Compustat Global, and author's calculation. Sample period covers the year 2004-2017.

Figure A.1: **Estimates of Average Productivity Measure** $log\ z_{it}$ **by Country**



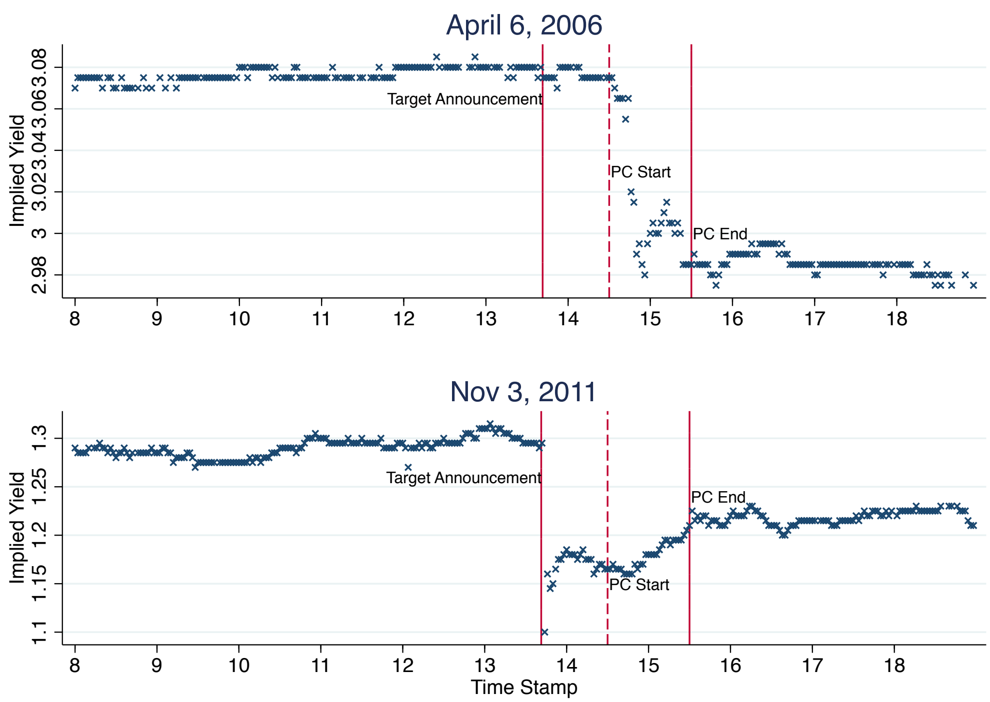
Notes. Estimates of the productivity measure $log\ z_{it}$ averaged across firms and years by country, calculated based on Equation (1.4). Source: Dealscan, Amadeus, Orbis, Compustat, Compustat Global, and author's calculation.

Figure A.2: **Firm-Bank Sorting, by z_i^G/z_i^L Quartile (Method 2)**



Notes. The plot shows sorting patterns between firms and global versus local banks, with firms sorted into quartiles by their exposure to global versus local risk (z_i^G/z_i^L), uses variables that are constructed based on Method 2 of the bank categorization criteria for global banks. Data sample consists of syndicated loans between firms global and local banks and firms across 24 countries from 2004-2017. Source: Dealscan, Amadeus, Orbis, Compustat, Compustat Global, and author's calculation.

Figure A.3: **Three-Month Euribor Rates around ECB Announcements**



Notes. The figure plots the three-month Euribor rates on April 6, 2006 (upper panel) and November 3, 2011 between 08:00 and 18:00. Vertical lines represent the target policy rate announcement (13:45), the start of the press conference (14:30), and the end of the press conference (15:30). All times are in CET. Source: CQG Data Factory.

Appendix B

Appendix for Chapter 2

B.1 Empirical Analysis

Robustness using PSID Data

We present a series of robustness tests of the estimations relating unemployment experiences to consumption, as well as the regression analysis of the wealth build-up hypothesis.

The first twelve tables present statistics and estimations using the PSID data. - In Appendix-Table B.1, we present the summary statistics of the full sample, i. e., including observations with total family income below the 5^{th} or above the 95^{th} percentile in each wave. Otherwise, we apply the same restrictions as in the construction of the main sample, namely, drop individuals for whom we cannot construct the experience measures (due to missing information about location or employment status in any year from t to $t-5$), and observations with missing demographic controls or that only appear once. The resulting sample has 37,156 observations, compared to 33,164 in the main sample. The sample statistics are very similar, with a mean macroeconomic experience measure of 6.0%, mean personal experience of 5.4%, average food consumption of \$8,559, and average total consumption of \$46,256 (both measured in 2013 dollars). In Appendix-Table B.2, we re-estimate the regression model of Table 2.2 on the full sample. The results become even stronger. The estimated macroeconomic experience and personal experience effects are both greater in magnitude and more significant compared to the results in Table 2.2.

In Appendix-Table B.3, we construct the experience measures for the gap years (between the PSID biennial surveys) in an alternative way. For the macroeconomic experience measure in the main text, we fill in the unemployment rate in a gap year t by assuming that the family lived in the same state as in year $t-1$. Here, we assume that respondents spend half of year t in the state in which they lived in year $t-1$ and the other half in the state in which they lived in year $t+1$. (This alternate construction does not change the value if respondents live in the same state in $t-1$ and $t+1$.) Similarly for the personal experience measure, we reconstruct respondents' employment status in year t as the average of their status in years $t-1$ and $t+1$, rather than applying the value from year $t-1$. For example, if a person is unemployed in $t-1$ and is employed in $t+1$, the personal experience in t will be denoted as 0.5. We then re-estimate the model in (2.3) using these alternative constructs of experience. The results are very similar to those in Table 2.2 in the main text.

In Appendix-Table B.4, we present an alternative experience measure that incorporates the experiences of both the head of the household and the spouse. The experience measure for married households is constructed using an average of the household's head and the spouse. We include a couple indicator among the demographic controls, which is equal to 1 for households who are married. All variables other than the couple indicator and the experience measures are defined as in Table 2.2. The coefficients of interest remain very stable, with some of the personal experience effect estimates increasing in (absolute) magnitude.

Appendix-Table B.5 presents yet another alternative experience measure, which excludes unemployment experiences from year $t-1$ to further rule out concurrent factors. All other variables are defined as in Table 2.2. The coefficients of interest remain stable without households fixed effects. When including households fixed effects, the estimates are slightly smaller in magnitude but remain significant.

In Appendix-Table B.6, we use weighting parameters $\lambda = 0$ and $\lambda = 3$ instead of $\lambda = 1$ to construct experience measures, and re-estimate the fixed-effect models of Table 2.2. Higher λ means individuals put more emphasis on their more recent experiences. When $\lambda = 0$, individuals are weighing all their past experiences equally. Note that experience-based learners with $\lambda = 0$ differ from Bayesian learners even though both assign equal weights to past realizations. Bayesian learners use all information to update their beliefs, while experience-based learners focus on information that occurred during their lifetime. As shown in Table B.6, the results remain similar. Hence, the significant relation between experience and consumption appears to be robust to the choice of the weighting parameter.

Appendix-Table B.7 shows the results when using different clustering units. Instead of clustering the standard errors by cohort as in Table 2.2, we cluster the standard errors by cohort$*$year, household, household$*$year, and we two-way cluster by cohort and year. The pooled regressions in Appendix-Table B.7 correspond to the specification in column (3) in Table 2.2, and the specifications with household fixed-effects correspond to column (6) in Table 2.2. As shown, the statistical significance of our results are not affected in most cases. Once we included household fixed effects, both experience variables are significant in predicting total consumption.

In Appendix-Table B.8, we apply the PSID longitudinal family weights. Note that some families are given zero weight and are thus dropped from the estimation, which explains the lower number of observations in the weighted regressions. As before the results remain very similar in the specifications with household fixed effects.

Appendix-Tables B.9, B.10, and B.11 address concerns about unobserved wealth, liquidity, or income components.

Appendix-Table B.9 presents results from estimations using alternative wealth controls. Column (1) controls for third- and fourth-order liquid and illiquid wealth. Column (2) controls decile dummies of liquid wealth and illiquid wealth. Column (3) controls for housing wealth and other wealth (total wealth minus housing wealth). Column (4) controls for positive wealth and debt. All wealth controls are in addition to the measures of liquid and illiquid wealth in Table 2.2. Columns (5) – (8) have the same wealth controls as columns (1) – (4) respectively but include household fixed effects. The coefficients of interest remain stable and (at least marginally) statistically significant.

Appendix-Table B.10 presents results from estimations using alternative income controls. Column (1) controls for third and fourth order of income and lagged income. Column (2) controls the quintile dummies of income and lagged income. Column (3) controls the decile dummies of income and lagged income. Column (4) controls the bottom 2, $2^{nd} - 4^{th}$, $4^{th} - 6^{th}$, $6^{th} - 8^{th}$, $8^{th} - 10^{th}$, $90^{th} - 92^{nd}$, $92^{nd} - 94^{th}$, $94^{th} - 96^{th}$, $96^{th} - 98^{th}$, and top 2 percentile dummies of income and lagged income. All income controls are in adddition to the controls of first and second order of income and lagged income. Columns (5) – (8) have the same income controls as columns (1) – (4) respectively but including household fixed effects. As with the alternative wealth controls, the coefficients of interest remain stable. All of the estimates that were significantly negative before are still significant.

In B.11, we test for whether households that are more liquidity constrained are more affected by their unemployment experience. Closely following the practice in the consumption literature such as Johnson et al. (2006) and Parker et al. (2013), we sort households into two groups based on whether their liquid wealth is above or below the sample median in the respective year. We then add an indicator for below-median liquid wealth as well as its interactions with the experience variables to the estimating equation in (2.3). As shown in Appendix-Table B.11, households in the bottom half of liquid wealth tend to spend less, but do not exhibit a stronger reaction to unemployment experience. This suggests households' experience significantly affect consumption above and beyond potential liquidity constraints.

In Appendix-Table B.12, we study the effects of lifetime experiences on household wealth accumulation. This analysis tests whether, given the significant impact of unemployment experiences on consumption, we can also detect experience effects in the build-up of wealth. The dependent variables are either liquid wealth or total wealth, and the main regressors are lagged experience measures. We lag the experience measures by six, eight, ten, twelve, and 14 years, instead of using the contemporary experience measures, recognizing that the effects of experience on wealth may take time to realize. We include the same set of control variables as in our main analyses, including controls for income in years $t-1$ and $t-2$, and add a control for the average family income between year $t-2$ and the year in which the lagged experience measures are based on (six, eight, ten, twelve, and 14 years ago, respectively). For example, when six-year lagged experience is the main regressor, we control for the average income between $t-2$ and $t-6$. This average-income control addresses the concern that previous experiences of economic boom or crisis may have implications for future income (Oyer (2008); Kahn (2010); Oreopoulos et al. (2012)).[1] In Appendix-Figure A.1, we plot the estimated coefficients on the lagged experience measures. In Appendix-Table B.12, we show the estimates of the coefficients on the 10-year, 12-year, and 14-year lagged experience measures. We find a significant role of past experiences for the build-up of wealth and liquid wealth, especially in the context of personal experiences.

Table B.1: **Summary Statistics (PSID), Full Sample**

Variable	Mean	SD	p10	p50	p90	N
Age	47.65	12.03	32	47	65	37,156
Experience (Macro) [in %]	6.00	0.28	5.67	5.97	6.37	37,156
Experience (Personal) [in %]	5.77	16.57	0.00	0.00	20.00	37,156
Household Size	2.73	1.45	1	2	5	37,156
Household Food Consumption [in $]	8,559	5,630	2,600	7,608	15,451	37,156
Household Total Consumption [in $]	46,256	36,497	14,733	39,559	82,765	37,156
Household Total Income [in $]	93k	133k	17k	69k	178k	37,156
Household Liquid Wealth [in $]	65k	718k	-22k	0k	117k	37,156
Household Illiquid Wealth [in $]	282k	1,268k	0k	72k	606k	37,156
Household Total Wealth [in $]	346k	1,545k	-3k	73k	762k	37,156

Notes. Summary statistics for the estimation sample, which covers the 1999-2013 PSID waves. Age, Experience (Macro), and Experience (Personal) are calculated for the heads of households. Household total income includes transfers and taxable income of all household members from the last year. Liquid wealth and illiquid wealth are defined following Kaplan, Violante and Weidner (2014). All values are in 2013 dollars using the PCE. Observations are annual and not weighted.

[1] The results are similar if, instead of having an average-income control, we include the incomes for all years between year $t-2$ and the year in which the lagged experience measures are based on.

Table B.2: **Consumption (PSID), Full Sample**

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Food Consumption						
Experience (Macro)	-0.181***		-0.165***	-0.171**		-0.163**
	(0.051)		(0.050)	(0.069)		(0.069)
Experience (Personal)		-0.756***	-0.752***		-0.426***	-0.422***
		(0.114)	(0.114)		(0.137)	(0.137)
R-squared	0.199	0.204	0.204	0.542	0.543	0.543
Dependent Variable: Total Consumption						
Experience (Macro)	-0.059*		-0.046	-0.079**		-0.073**
	(0.031)		(0.028)	(0.031)		(0.031)
Experience (Personal)		-0.603***	-0.602***		-0.328***	-0.326***
		(0.073)	(0.073)		(0.082)	(0.082)
R-squared	0.496	0.507	0.507	0.755	0.757	0.757
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Income controls	Yes	Yes	Yes	Yes	Yes	Yes
Wealth controls	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	No	No	No	Yes	Yes	Yes
Observations	37,156	37,156	37,156	37,156	37,156	37,156

Notes. We include all observations i.e., also observations with total family income below the 5^{th} or above the 95^{th} percentile in each wave from 1999 to 2013, as well as the pre-sample 1997 wave (because we control for lagged income). All variables are defined as in Table 2.2. Robust standard errors (in parentheses) are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

Table B.3: **Consumption (PSID), Alternative Experience Measure: Gap Years**

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Food Consumption						
Experience (Macro)	-0.099**		-0.093*	-0.124**		-0.120**
	(0.047)		(0.048)	(0.055)		(0.055)
Experience (Personal)		-0.337***	-0.335***		-0.267**	-0.264**
		(0.104)	(0.104)		(0.127)	(0.128)
R-squared	0.192	0.193	0.193	0.541	0.542	0.542
Dependent Variable: Total Consumption						
Experience (Macro)	-0.022		-0.018	-0.061***		-0.059***
	(0.020)		(0.019)	(0.022)		(0.022)
Experience (Personal)		-0.182***	-0.181***		-0.152***	-0.151***
		(0.031)	(0.031)		(0.033)	(0.033)
R-squared	0.573	0.574	0.574	0.788	0.788	0.788
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Income controls	Yes	Yes	Yes	Yes	Yes	Yes
Wealth controls	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	No	No	No	Yes	Yes	Yes
Observations	33,164	33,164	33,164	33,164	33,164	33,164

Notes. All variables other than the experience measures are defined as in Table 2.2. The construction of the experience measures differs as follows: For any gap year t (between PSID survey waves in $t-1$ and $t+1$), the baseline experience measures in the main text assume that families reside in the same state as in year $t-1$. The alternative construction used in this Appendix-Table assumes that families reside half of year t in their (t-1)-state of residence, and half of the year in their (t+1)-state of residence. (The different assumption does not matter when a family does not move between surveys.) Hence, the macro experience measure in this Appendix-Table uses the average of the year t unemployment rates of the (t-1)-state of residence and the (t+1)-state residence as gap year t's unemployment rate. Similarly, for the personal experience measure, we fill in the employment status of a household head in a gap year with the average of the years before and after. For example, if a person is unemployed in $t-1$ and is employed in $t+1$, then his personal experience in year t is denoted as 0.5. We exclude from the sample observations with total family income below the 5^{th} or above the 95^{th} percentile in each wave from 1999 to 2013, as well as the pre-sample 1997 wave (because we control for lagged income). Robust standard errors (in parentheses) are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

Table B.4: **Consumption (PSID): Alternative Experience Measure: Spousal Experience**

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Food Consumption						
Experience (Macro)	-0.079*		-0.071	-0.111**		-0.106*
	(0.047)		(0.047)	(0.054)		(0.055)
Experience (Personal)		-0.402***	-0.400***		-0.313**	-0.309**
		(0.111)	(0.111)		(0.130)	(0.130)
R-squared	0.192	0.193	0.193	0.541	0.542	0.542
Dependent Variable: Total Consumption						
Experience (Macro)	-0.021		-0.017	-0.059***		-0.056***
	(0.019)		(0.019)	(0.021)		(0.021)
Experience (Personal)		-0.213***	-0.212***		-0.161***	-0.159***
		(0.035)	(0.034)		(0.033)	(0.033)
R-squared	0.573	0.574	0.574	0.788	0.788	0.788
Income controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Income controls	Yes	Yes	Yes	Yes	Yes	Yes
Wealth controls	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33,164	33,164	33,164	33,164	33,164	33,164

Notes. All variables other than the couple indicator, and experience measures are defined as in Table 2.2. Couple is an indicator equal to 1 for households who are married, and is now included as a demographic controle. The experience measures for the married households are constructed using an average of the household's head and the spouse. We exclude from the sample observations with total family income below the 5^{th} or above the 95^{th} percentile in each wave from 1999 to 2013, as well as the pre-sample 1997 wave (because we control for lagged income). Robust standard errors (in parentheses) are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

Table B.5: **Consumption (PSID), Alternative Experience Measure: Lagged Experience**

Dependent Varariable:	Food Consumption			Total Consumption		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Food Consumption						
Experience (Macro)	-0.109*		-0.103*	-0.094**		-0.093**
	(0.056)		(0.057)	(0.046)		(0.046)
Experience (Personal)		-0.320**	-0.318**		-0.066	-0.064
		(0.134)	(0.134)		(0.136)	(0.136)
R-squared	0.204	0.205	0.205	0.587	0.587	0.587
Dependent Variable: Total Consumption						
Experience (Macro)	-0.013		-0.010	-0.047*		-0.046*
	(0.024)		(0.024)	(0.026)		(0.025)
Experience (Personal)		-0.179***	-0.178***		-0.120***	-0.119***
		(0.038)	(0.038)		(0.038)	(0.038)
R-squared	0.572	0.573	0.573	0.806	0.807	0.807
Income controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Income controls	Yes	Yes	Yes	Yes	Yes	Yes
Wealth controls	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20,163	20,163	20,163	20,163	20,163	20,163

Notes. The experience measures (both macro and personal) does not contain unemployment experience from year $t-1$. All other variables are defined as in Table 2.2. We exclude from the sample observations with total family income below the 5^{th} or above the 95^{th} percentile in each wave from 1999 to 2013, as well as the pre-sample 1997 wave (because we control for lagged income). Robust standard errors (in parentheses) are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

Table B.6: **Consumption (PSID), Alternative Experience Measure: Different Weights (λ)**

Dependent Varariable:	Food Consumption			Total Consumption		
	(1)	(2)	(3)	(4)	(5)	(6)
Weighting Parameter $\lambda = 0$						
Experience (Macro)	-0.026		-0.024	-0.108**		-0.104**
	(0.027)		(0.027)	(0.041)		(0.041)
Experience (Personal)		-0.322***	-0.321***		-0.147***	-0.146***
		(0.097)	(0.097)		(0.031)	(0.031)
R-squared	0.192	0.193	0.193	0.788	0.788	0.788
Weighting Parameter $\lambda = 3$						
Experience (Macro)	-0.021**		-0.019*	-0.033***		-0.032**
	(0.010)		(0.011)	(0.012)		(0.012)
Experience (Personal)		-0.175***	-0.174***		-0.150***	-0.149***
		(0.029)	(0.029)		(0.031)	(0.031)
R-squared	0.573	0.574	0.574	0.788	0.788	0.788
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Income controls	Yes	Yes	Yes	Yes	Yes	Yes
Wealth controls	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33,164	33,164	33,164	33,164	33,164	33,164

Notes. All variables other than the experience measures are defined as in Table 2.2. The experience measures are constructed using $\lambda = 0$ in the upper part of the table, and $\lambda = 3$ in the lower part. We exclude from the sample observations with total family income below the 5^{th} or above the 95^{th} percentile in each wave from 1999 to 2013, as well as the pre-sample 1997 wave (because we control for lagged income). Robust standard errors (in parentheses) are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

Table B.7: **Consumption (PSID), Alternative Clustering Units**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable: Food Consumption								
Experience (Macro)	-0.091**	-0.091**	-0.091**	-0.091**	-0.117**	-0.117	-0.117**	-0.117**
	(0.041)	(0.033)	(0.046)	(0.042)	(0.053)	(0.062)	(0.050)	(0.049)
Experience (Personal)	-0.320***	-0.320**	-0.320***	-0.320***	-0.260**	-0.260*	-0.260***	-0.260**
	(0.086)	(0.113)	(0.095)	(0.085)	(0.109)	(0.132)	(0.099)	(0.101)
R-squared	0.193	0.193	0.193	0.193	0.542	0.542	0.542	0.542
Dependent Variable: Total Consumption								
Experience (Macro)	-0.018	-0.018	-0.018	-0.018	-0.057***	-0.057**	-0.057***	-0.057***
	(0.015)	(0.017)	(0.017)	(0.015)	(0.017)	(0.020)	(0.019)	(0.017)
Experience (Personal)	-0.177***	-0.177***	-0.177***	-0.177***	-0.147***	-0.147**	-0.147***	-0.147***
	(0.024)	(0.046)	(0.026)	(0.023)	(0.028)	(0.043)	(0.030)	(0.027)
R-squared	0.574	0.574	0.574	0.574	0.788	0.788	0.788	0.788
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wealth controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	No	No	No	No	Yes	Yes	Yes	Yes
Observations	33,164	33,164	33,164	33,164	33,164	33,164	33,164	33,164

Notes. All variables are defined as in Table 2.2. Standard errors in columns (1) to (4) are clustered by cohort$*$year, cohort and year (two-way clustering), household and household$*$year, respectively, and the same for columns (5) to (8). We exclude from the sample observations with total family income below the 5^{th} or above the 95^{th} percentile in each wave from 1999 to 2013, as well as the pre-sample 1997 wave (because we control for lagged income). Robust standard errors (in parentheses) are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

Table B.8: **Consumption (PSID), Alternative Weights: PSID Weights**

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Food Consumption						
Experience (Macro)	0.016		-0.093*	-0.119**		-0.115**
	(0.056)		(0.047)	(0.054)		(0.054)
Experience (Personal)		-0.302***	-0.324***		-0.262**	-0.260**
		(0.112)	(0.098)		(0.120)	(0.120)
R-squared	0.225	0.226	0.193	0.541	0.541	0.541
Dependent Variable: Total Consumption						
Experience (Macro)	-0.003		-0.021	-0.058***		-0.056**
	(0.023)		(0.019)	(0.021)		(0.022)
Experience (Personal)		-0.162***	-0.176***		-0.150***	-0.149***
		(0.042)	(0.030)		(0.032)	(0.032)
R-squared	0.576	0.576	0.574	0.787	0.788	0.788
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Income controls	Yes	Yes	Yes	Yes	Yes	Yes
Wealth controls	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	No	No	No	Yes	Yes	Yes
Observations	32,834	32,834	32,834	32,834	32,834	32,834

Notes. All variables are defined as in Table 2.2, but observations are now weighted by the PSID family weights. The family with zero weights are dropped. Robust standard errors (in parentheses) are clustered by cohort. We exclude from the sample observations with total family income below the 5^{th} or above the 95^{th} percentile in each wave from 1999 to 2013, as well as the pre-sample 1997 wave (because we control for lagged income). Robust standard errors (in parentheses) are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

Table B.9: **Consumption (PSID), Additional Wealth Controls**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable: Food Consumption								
Experience (Macro)	-0.088*	-0.079	-0.082*	-0.082	-0.110**	-0.101*	-0.104*	-0.096*
	(0.047)	(0.047)	(0.048)	(0.051)	(0.054)	(0.054)	(0.054)	(0.055)
Experience (Personal)	-0.318***	-0.272***	-0.321***	-0.231**	-0.258**	-0.258**	-0.261**	-0.226*
	(0.097)	(0.099)	(0.098)	(0.098)	(0.119)	(0.120)	(0.119)	(0.120)
R-squared	0.194	0.199	0.194	0.203	0.542	0.543	0.542	0.551
Dependent Variable: Total Consumption								
Experience (Macro)	-0.015	-0.005	-0.013	-0.003	-0.054**	-0.046**	-0.053**	-0.040*
	(0.019)	(0.019)	(0.019)	(0.020)	(0.021)	(0.021)	(0.022)	(0.020)
Experience (Personal)	-0.175***	-0.131***	-0.178***	-0.069***	-0.146***	-0.144***	-0.148***	-0.130***
	(0.029)	(0.026)	(0.029)	(0.024)	(0.031)	(0.030)	(0.031)	(0.030)
R-squared	0.577	0.596	0.575	0.633	0.788	0.791	0.788	0.805
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wealth controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	No	No	No	No	Yes	Yes	Yes	Yes
Observations	33,164	33,164	33,164	31,187	33,164	33,164	33,164	31,187

Notes. The pooled regressions and the regressions with household fixed effects here are only different from the regressions in Table 2.2 in terms of the wealth controls. Column (1) controls for third- and fourth-order liquid and illiquid wealth. Column (2) includes decile dummies of liquid wealth and illiquid wealth. Column (3) controls for housing wealth and other wealth (total wealth minus housing wealth). Column (4) controls for positive wealth and debt. All wealth controls are in addition to the controls of first and second order of liquid and illiquid wealth. Columns (5) – (8) have the same wealth controls as columns (1) – (4) respectively. Robust standard errors are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

Table B.10: **Consumption (PSID), Additional Income Controls**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable: Food Consumption								
Experience (Macro)	-0.097**	-0.095**	-0.098**	-0.089*	-0.118**	-0.116**	-0.118**	-0.113**
	(0.047)	(0.047)	(0.048)	(0.048)	(0.055)	(0.055)	(0.056)	(0.055)
Experience (Personal)	-0.267***	-0.274***	-0.238**	-0.227**	-0.240**	-0.243**	-0.232*	-0.233*
	(0.099)	(0.098)	(0.101)	(0.105)	(0.119)	(0.120)	(0.121)	(0.121)
R-squared	0.200	0.200	0.204	0.206	0.543	0.543	0.544	0.544
Dependent Variable: Total Consumption								
Experience (Macro)	-0.020	-0.018	-0.017	-0.017	-0.057**	-0.056**	-0.056**	-0.056**
	(0.019)	(0.019)	(0.019)	(0.019)	(0.022)	(0.022)	(0.021)	(0.021)
Experience (Personal)	-0.153***	-0.154***	-0.142***	-0.141***	-0.139***	-0.140***	-0.136***	-0.138***
	(0.029)	(0.029)	(0.029)	(0.029)	(0.031)	(0.031)	(0.031)	(0.031)
R-squared	0.579	0.579	0.581	0.581	0.789	0.788	0.789	0.789
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wealth controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	No	No	No	No	Yes	Yes	Yes	Yes
Observations	33,164	33,164	33,164	33,164	33,164	33,164	33,164	33,164

Notes. The pooled regressions and the regressions with household fixed effects here are only different from the regressions in Table 2.2 in terms of the income controls. Column (1) controls for third and fourth order of income and lagged income. Column (2) includes quintile dummies of income and lagged income. Column (3) includes decile dummies of income and lagged income. Column (4) includes separately for the bottom 2, $2^{nd} - 4^{th}$, $4^{th} - 6^{th}$, $6^{th} - 8^{th}$, $8^{th} - 10^{th}$, $90^{th} - 92^{nd}$, $92^{nd} - 94^{th}$, $94^{th} - 96^{th}$, $96^{th} - 98^{th}$, and top 2 percentile dummies of income and lagged income. All income controls are in addition to the controls of first and second order of income and lagged income. Columns (5) – (8) have the same income controls as columns (1) – (4) respectively. Robust standard errors are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

Table B.11: **Consumption (PSID), Additional Liquidity Controls**

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Food Consumption						
Experience (Macro)	-0.147**		-0.144**	-0.143**		-0.143**
	(0.058)		(0.058)	(0.063)		(0.063)
Experience (Macro) * LLW	0.097		0.103*	0.048		0.055
	(0.060)		(0.059)	(0.056)		(0.055)
Low Liquid Wealth	-0.572	0.013	-0.602*	-0.316	-0.023	-0.352
	(0.358)	(0.013)	(0.355)	(0.335)	(0.014)	(0.331)
Experience (Personal)		-0.302**	-0.292**		-0.241	-0.236
		(0.142)	(0.142)		(0.149)	(0.149)
Experience (Personal) * LLW		-0.037	-0.053		-0.038	-0.046
		(0.177)	(0.176)		(0.156)	(0.155)
R-squared	0.192	0.193	0.193	0.542	0.542	0.542
Dependent Variable: Total Consumption						
Experience (Macro)	-0.021		-0.023	-0.054**		-0.055**
	(0.022)		(0.022)	(0.024)		(0.023)
Experience (Macro) * LLW	0.001		0.009	-0.012		-0.006
	(0.015)		(0.016)	(0.015)		(0.016)
Low Liquid Wealth	0.034	0.047***	-0.006	0.080	0.013***	0.046
	(0.093)	(0.006)	(0.094)	(0.093)	(0.004)	(0.095)
Experience (Personal)		-0.087**	-0.086**		-0.083**	-0.082**
		(0.036)	(0.036)		(0.033)	(0.033)
Experience (Personal) * LLW		-0.166***	-0.168***		-0.118***	-0.118***
		(0.046)	(0.046)		(0.044)	(0.044)
R-squared	0.573	0.575	0.575	0.788	0.788	0.788
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Income controls	Yes	Yes	Yes	Yes	Yes	Yes
Wealth controls	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	No	No	No	Yes	Yes	Yes
Observations	33,164	33,164	33,164	33,164	33,164	33,164

Notes. Low Liquid Wealth (LLW) is an indicator variable equal to 1 for households whose liquid wealth falls below the sample median of liquid wealth each year. All other variables, excluding the interaction of LLW with Experience (Macro), and Experience (Personal), are defined as in Table 2.2. We exclude from the sample observations with total family income below the 5^{th} or above the 95^{th} percentile in each wave from 1999 to 2013, as well as the pre-sample 1997 wave (because we control for lagged income). Robust standard errors (in parentheses) are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

Table B.12: **Wealth Accumulation**

Dependent Var.:	Liquid Wealth at time t						Total Wealth at time t					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Exp. (Macro)$_{t-10}$	0.006*		0.005	0.003		0.003	0.012		0.010	0.018***		0.019***
	(0.003)		(0.003)	(0.003)		(0.003)	(0.008)		(0.008)	(0.006)		(0.006)
Exp. (Personal)$_{t-10}$		0.023***	0.023***		-0.000	-0.001		0.083***	0.083***	-0.003	-0.005	
		(0.004)	(0.004)		(0.002)	(0.002)		(0.013)	(0.013)		(0.014)	(0.014)
R-squared	0.048	0.048	0.048	0.332	0.332	0.332	0.292	0.294	0.294	0.714	0.714	0.714
Observations	21,691	21,691	21,691	21,691	21,691	21,691	21,691	21,691	21,691	21,691	21,691	21,691
Exp. (Macro)$_{t-12}$	0.007**		0.006*	0.008**		0.007**	0.010		0.007	0.020***		0.020***
	(0.003)		(0.003)	(0.003)		(0.003)	(0.009)		(0.009)	(0.007)		(0.007)
Exp. (Personal)$_{t-12}$		0.026***	0.026***		0.002	0.001		0.092***	0.091***		0.003	0.001
		(0.005)	(0.005)		(0.002)	(0.003)		(0.014)	(0.014)		(0.014)	(0.014)
R-squared	0.049	0.050	0.050	0.333	0.333	0.333	0.294	0.296	0.296	0.730	0.730	0.730
Observations	19,427	19,427	19,427	19,427	19,427	19,427	19,427	19,427	19,427	19,427	19,427	19,427
Exp. (Macro)$_{t-14}$	0.008**		0.007*	0.008**		0.008**	0.002		-0.001	0.011*		0.010
	(0.003)		(0.003)	(0.003)		(0.003)	(0.009)		(0.010)	(0.006)		(0.006)
Exp. (Personal)$_{t-14}$		0.028***	0.028***		0.003	0.002		0.095***	0.095***		0.010	0.009
		(0.005)	(0.005)		(0.003)	(0.003)		(0.013)	(0.013)		(0.009)	(0.009)
R-squared	0.052	0.052	0.052	0.331	0.331	0.331	0.378	0.380	0.380	0.827	0.827	0.827
Observations	17,151	17,151	17,151	17,151	17,151	17,151	17,151	17,151	17,151	17,151	17,151	17,151
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	No	No	No	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes

Notes. "Exp. (Macro)" is the macroeconomic experience measure, and "Exp. (Personal)" is the personal experience measure. Liquid wealth and total wealth are defined in the way as in the main draft. We separately use the $t-10$, $t-12$ experience measures, and $t-14$ experience measures. Income controls include the $t-1$ family total income and the average family total income between $t-2$ and the year we use the experience measures. For gap years (between PSID survey waves), we use the assumption from baseline analysis and use prior-year income. Demographic controls include family size, the household heads' gender, race, marital status, education level, and employment status. We take the logarithm of all income and wealth variables. Robust standard errors (in parentheses) are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

Robustness using CEX-Nielsen Synthetic Panel Data

In order to keep the advantages of panel analysis but also exploit the comprehensiveness of the CEX, we match the two datasets and create a synthetic panel. Specifically, we match a household i from the CEX data with a household j from Nielsen on a set of common covariates (characteristics) $x_i = (x_{i,1}, x_{i,2}, ..., x_{i,p})$ and $x_j = (x_{j,1}, x_{j,2}, ..., x_{j,p})$, which include age, income, marital status, household size, education, race, region of residency, employment status, as well as their consumption of non-durable items, using the nearest-neighbor matching estimator from Rosenbaum and Rubin (1983) and Abadie and Imbens (2011). The distance between x_i and x_j is parameterized by the vector norm $\| x_i - x_j \|_S = ((x_i - x_j)' S^{-1} (x_i - x_j))^{1/2}$, where S is a given symmetric, positive-definite matrix. We find that the set of nearest-neighbor indices for observation i from the CEX in Nielsen as $\Omega_i = (j | t_j = 1 - t_i, \| x_i - x_j \|_S < \| x_i - x_l \|_S, t_l = 1 - t_i, l \neq j)$. In words, the nearest-neighbor propensity-score matching chooses for each observation in the CEX an observation in Nielsen that has the closest estimated propensity score.

Table B.13 provides summary statistics on the matched sample. In the matched dataset, the distributions on total and durable consumption are comparable to those from the underlying CEX data, which is indicative of successful matching. For an average household, its share of durable consumption makes up 10% of total spending, while non-durable consumption amounts to 69% of total spending.

Table B.13: **Summary Statistics (Nielsen-CEX Matched Data)**

Variable	Mean	SD	p10	p50	p90	N
Total consumption expenditure	4,508	4,919	1,838	3,371	7,111	866,819
Durable consumption	1,078	4,466	0	117	1,460	866,819
Non-durable consumption	2,612	1,178	1,423	2,400	4,025	866,819
Non-durable consumption (Nielsen)	2,139	1,602	618	1,757	4,083	3,171,833
Experience (Macro)	5.9	0.2	5.8	5.9	6.2	866,819

Notes. The sample period runs quarterly from 2004 to 2012. Observations are quarterly and not weighted.

Table B.14 shows results from re-estimating specification (2.11) using the matched CEX-Nielsen sample. In columns (1) and (4) we use total expenditures as the outcome variable, in columns (2) and (5), we focus on durable consumption spending, and in columns (3) and (6) we focus on non-durables. As before we show the results both without household fixed effects (columns 1 to 3) and with fixed effects (columns 4 to 6).

For all outcome variables – durable, non-durable, and total consumption – we continue to estimate highly significant negative experience effects. Households who have experienced worse unemployment conditions during their lifetime spend significantly less in total (on all goods), and also specifically on durable and on non-durable items.

Table B.14: **Consumption (Nielsen-CEX Matched Sample)**

	Total	Durables	Non-durable	Total	Durables	Non-durable
Experience (Macro)	-0.358***	-0.797***	-0.220**	-0.266***	-0.796***	-0.033
	(0.038)	(0.122)	(0.019)	(0.051)	(0.145)	(0.028)
Income control	Yes	Yes	Yes	Yes	Yes	Yes
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Market-area fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	No	No	No	Yes	Yes	Yes
Observations	866,819	866,819	866,819	866,819	866,819	866,819
R-squared	0.183	0.053	0.257	0.020	0.008	0.069

Notes. Regressions with (log) total consumption expenditure, durable consumption, and non-durable consumption as the dependent variables. Experience (Macro) is the macroeconomic experience measure of unemployment (household's lifetime experience of national unemployment rates). Household characteristics include unemployment status, household size, education, and race. Time fixed effects include year-quarter fixed effects. Regressions are weighted by household sampling weights from Nielsen. The sample period runs from 2004 to 2012. Robust standard errors (in parentheses) are clustered by cohort. *, **, *** denote 10%, 5%, and 1% significance, respectively.

One exception are non-durables in the case where we identify only within household; here the coefficient becomes small and insignificant. Otherwise, the coefficients are stable across specifications, and the economic magnitudes are large: a one standard deviation increase in lifetime unemployment experience is associated with a \$38 decline in monthly non-durable consumption and \$108 decline in monthly total consumption (using the estimates of columns 3 and 1 respectively). The new estimate for durable consumption is large and highly significant across specifications. A one standard deviation increase in lifetime unemployment experience is associated with a \$57 decline in monthly durable consumption.

B.2 Model

We implement the empirical model of Low et al. (2010) with a few minor adjustments to our setting. All key equations are retained and, when possible, all parameters are set to the same values. As in Low et al. (2010), some parameters are set separately for high- and low-education groups, including the probability of job destruction and job offers.

Parameters governing the income process and utility maximization

The utility function and lifetime expected utility are defined in equations (2.4) and (2.5) in Section 2.4 as $U(c,P) = \frac{\left(c \times e^{\eta P}\right)^{1-\gamma}}{1-\gamma}$ and $U(c_{i,t}, P_{i,t}) + \mathrm{E}_t\left[\sum_{s=t+1}^{L} \beta^{s-t} U(c_{i,s}, P_{i,s})\right]$, respectively. In the simulations, we follow Low et al. (2010) and take risk aversion parameter $\gamma = 1.5$ from Attanasio and Weber (1995), use the estimates for η from their Table 2, and set the discount factor $\beta = 1/R$ in the value function,

Turning to the gross quarterly income $w_{i,t}h$, we follow Low et al. (2010) in setting the number of hours worked per quarter $h = 500$. In the wage process $\ln w_{i,t} = d_t + x'_{i,t}\psi + u_{i,t} + a_{i,j,t_0}$, we recover the parameters α, β_1, and β_2 governing the deterministic component, $d_t + x'_{i,t}\psi = \alpha + \beta_1 \cdot \mathrm{age} + \beta_2 \cdot \mathrm{age}^2$, from the parameters in the Fortran code published alongside Low et al. (2010). In the permanent component $u_{i,t} = u_{i,t-1} + \zeta_{i,t}$, $\zeta_{i,t}$ is i. i. d. normal with mean 0 and variance σ_ζ^2, and we use the value of σ_ζ given in Table 1 of Low et al. (2010). The consumer-firm job-match component, a_{i,j,t_0}, is drawn from a normal distribution with mean 0 and variance σ_a^2, and we use the value of σ_a given in Table 1 of Low et al. (2010).

We obtain the values for the probabilities of job destruction δ, of a job offers $(1-\delta)\lambda^e$ (when employed) and λ^n (when unemployed) from Table 2 in Low et al. (2010). Note that, while the probability of job destruction is constant across time for a given household, the probability of receiving a job offer varies depending on whether or not an agent is employed.

Budget constraint

The intertemporal budget constraint for a working individual i in period t is given by

$$\begin{aligned} A_{i,t+1} = R\left[A_{i,t} - c_{i,t}\right] + \left(w_{i,t}h(1-\tau_w) - F_{i,t}\right)P_{i,t} \\ + \left(B_{i,t}I_{i,t}^{UI}(1-I_{i,t}^{DI}) + D_{i,t}I_{i,t}^{DI}\right)(1-P_{i,t}) + T_{i,t}I_{i,t}^{T} \end{aligned} \tag{B.1}$$

where $A_{i,t}$ is the beginning of period assets in period t, R is the interest factor, τ_w is a tax, F is the fixed cost of working, P an indicator for whether an individual is working, B are unemployment benefits, D are disability benefits, T are food stamp benefits, c is consumption, and the I variables are indicators of receiving the associated social insurance programs.

As in Low et al. (2010), we assume that individuals cannot borrow and thus $A_{i,t} \geq 0 \quad \forall t$. Also as in Low et al. (2010), we set $r = .15$ and define $R = 1 + r$, and we use the estimates for F from their Table 2. In Low et al. (2010) τ_w is a variable of interest and solved for, albeit as fixed percentage (not allowing for progressive or regressive taxation). As we do not focus on the value of social insurance programs, including the tax revenues to be raised to fund them, nor their relation with consumption, we normalize $\tau_w = 0$.

During retirement individuals receive social security equal to the value of disability, so the equation simplifies to

$$A_{i,t+1} = R\left[A_{i,t} + D_{i,t} - c_{i,t}\right].$$

Social Insurance programs

As in Low et al. (2010), we implement three social insurance programs, unemployment insurance, food stamps, and disability insurance.

Unemployment Insurance. Unemployment Insurance is paid only during the quarter following job destruction, and we assume it is perfectly monitored. The value of unemployment benefits is given by

$$B_{i,t} = \begin{cases} bw_{i,t-1}h & \text{if } bw_{i,t-1}h < B_{\max}, \\ B_{\max} & \text{if } bw_{i,t-1}h \geq B_{\max}. \end{cases}$$

where b is the replacement ratio, and $B_{\max}$ is the cap on unemployment benefits. We set $b = .75$ as in Low et al. (2010) and $B_{\max}$ to the value used in the associated code.

Food Stamps (Means-Tested Social Insurance). This program uses the following definition of gross income

$$y_{i,t}^{\text{gross}} = w_{i,t}hP_{i,t} + \left(B_{i,t}I_{i,t}^{UI}(1-I_{i,t}^{DI}) + D_{i,t}I_{i,t}^{DI}\right)(1-P_{i,t}).$$

Then net income is defined as

$$y = (1 - \tau_w) y^{gross} - d.$$

Finally, the amount of food stamps allocated to agent i in period t is

$$T_{i,t} = \begin{cases} \overline{T} - .3 \times y_{i,t} & \text{if } y_{i,t} \leq \underline{y} \\ 0 & \text{otherwise,} \end{cases}$$

where $\overline{T}$ is a maximum payment and $\underline{y}$ is a poverty line. One important implication of this definition is that there is no disincentive to hold assets.

Adjusting to get quarterly values, we set $\overline{T}$ to the actual maximum food stamp allotment for a couple in the US in 1993, we set $\underline{y}$ as the max food stamp allotment for the US in 1993, and we set d to the actual standard deduction for a couple in the US in 1993.

Disability. As in Low et al. (2010), individuals above 50 can apply for disability when they are unemployed and are accepted with a fixed probability of .5. If an application is successful, disability becomes an absorbing state for the remainder of the person's working life. If a person is not accepted, they can only reapply in a future bout of unemployment, after having worked again for at least one year. As a disincentive to applying, the individual must be unemployed in both the period they apply and the period after. We also impose that individuals must have a sufficiently low u and not be working or have a job offer at the time of application. The formula for disability benefits is

$$D_{i,t} = \begin{cases} .9 \times \overline{w}_i & \text{if } \overline{w}_i \leq a_1 \\ .9 \times a_1 + .32 \times (\overline{w}_i - a_1) & \text{if } a_1 < \overline{w}_i \leq a_2 \\ .9 \times a_1 + .32 \times (a_2 - a_1) + .15 \times (\overline{w}_i - a_2) & \text{if } a_2 < \overline{w}_i \leq a_3 \\ .9 \times a_1 + .32 \times (a_2 - a_1) + .15 \times (a_3 - a_2) & \text{if } \overline{w}_i > a_3 \end{cases}$$

where a_1, a_1, and a_3 are fixed thresholds from legislation and $\overline{w}_i$ is the mean earnings prior to application. Similar to Low et al. (2010), we assume $\overline{w}_i$ can be approximated using the agent's value of $u_{i,t}$ at the time of application.

Implementation

Appendix-Table B.15 details all parameters referenced above and where the value was taken from. As discussed, most values are obtained from Low et al. (2010) directly, and some are retrieved from examining the associated Fortran 90 code published with the paper. In the case that we were unable to find certain values in either the paper or the code, as is the case for several welfare values, we used actual values from 1993. The SIPP survey used in Low et al. (2010), from which hourly wage data is sourced,

Table B.15: **Model Parameters Used in Simulations**

Parameter	Low Education	High Education	Source
γ	1.5	1.5	Text
σ_a	0.226	0.229	Table 1
σ_ζ	0.095	0.106	Table 1
$P(\zeta)$	.25	.25	Text
δ	.049	.028	Table 2
λ^e	.67	.72	Table 2
λ^n	.76	.82	Table 2
b	.75	.75	Text
r (yearly)	.015	.015	Text
β	$1/(1+r)$	$1/(1+r)$	Text
F	1088	1213	Table 2
η	-.55	-.62	Table 2
h	500	500	Text
b	.75	.75	Text
UI Cap	3178	3178	Code
P(Disability Acceptance)	.5	.5	Text
a_1	1203	1203	Code
a_2	7260	7260	Code
a_3	16638	16638	Code
α	1.0583	.642	Code
β_1	.0486	.0829	Code
β_2	-0.0004816	-.0007768	Code
Parameter	Low Education	High Education	Source
d	6200/4		Standard couple deduction in 1993[2]
$\underline{y}$	(6970+2460)/4		Actual poverty line in 1993 for couple[3]
$\overline{T}$	203×3		Actual max food stamp allotment for US 1993[4]

begins in 1993. This is also the closest year in the SIPP survey to the PSID data and the values are consistent with the model values.

Like Low et al. (2010), we solve the model numerically. In the last period, all agents consume the entirety of their assets. We then iteratively solve backwards for consumption and other relevant decisions that maximize the agents' value functions. Further details of the model solution can be found in Low et al. (2010).

Figure B.1 depicts the resulting average consumption trends of rational and experience-based learners during their working years, which are the years used in the regressions. The graph hints at a pattern that, early in life, experience-based learners underestimate the probability of job destruction, spend more, and must then save more towards the end of their working life.

Figure B.2 provides an amplified illustration of the differences. In this figure, we only consider the subset of experience-based learners in the simulation who, at period 30, have a believed delta of 0.025 or less and, in the rational case, the subset of agents who would have a believed delta of 0.025 or less at period 30 if they were experience-based learners. Since the true probability of job destruction is 0.049, these agents were "lucky" early in life. For these consumers, the trend of over-consumption among experienced based learners in the early periods is much more pronounced.

Figure B.3 illustrates the opposite scenario. Here, we only consider the subset of experience-based learners in the simulation who, at period 30, have a believed delta of 0.1 or greater, as well as the corresponding rational agents. In light of the true probability of job destruction of 0.049, these agents have had bad luck early in life. This "unlucky" group of experience-based learners has a markedly different savings pattern. They consistently consume less than their rational counterparts for almost there entire lives. Moreover, the illustration hints at an additional prediction, wealth build-up due to excess frugality .

High Education Regressions

For completeness, we replicate our baseline result, presented in Table 2.6, for the high-education subgroup, both with and without the addition of unemployment scarring.

In Appendix-Table B.16, we re-estimate the regression specifications from Table 2.6 using the high-education parameters, partly described in Table 2.5. Some of the other different parameters are the cost of working, the probability of job offers, and the probability of job destruction. As noted in the main text, all parameters are sourced from Low et al. (2010) when possible. Importantly, the signs on the unemployment experience coefficients remains the same. In addition, the coefficient on unemployment experience is greater than and the statistical significance is comparable to the results in Table 2.6.

In Appendix-Table B.17, we re-estimate the regression specifications from Table 2.6 using high education parameters and including unemployment scarring. Again, we find consistent results in terms of sign and statistical significance.

Table B.16: **High-Education Estimations with Model-Simulated Data**

	(1) Rational	(2) Rational	(3) EBL	(4) EBL
Income	0.792	0.638	0.798	0.661
	(189.80)	(127.15)	(230.94)	(91.36)
Wealth		0.0183		0.0170
		(45.06)		(31.60)
Unemployment Experience	3316.0	3112.5	-2173.9	-3724.9
	(5.00)	(4.38)	(-7.31)	(-7.36)

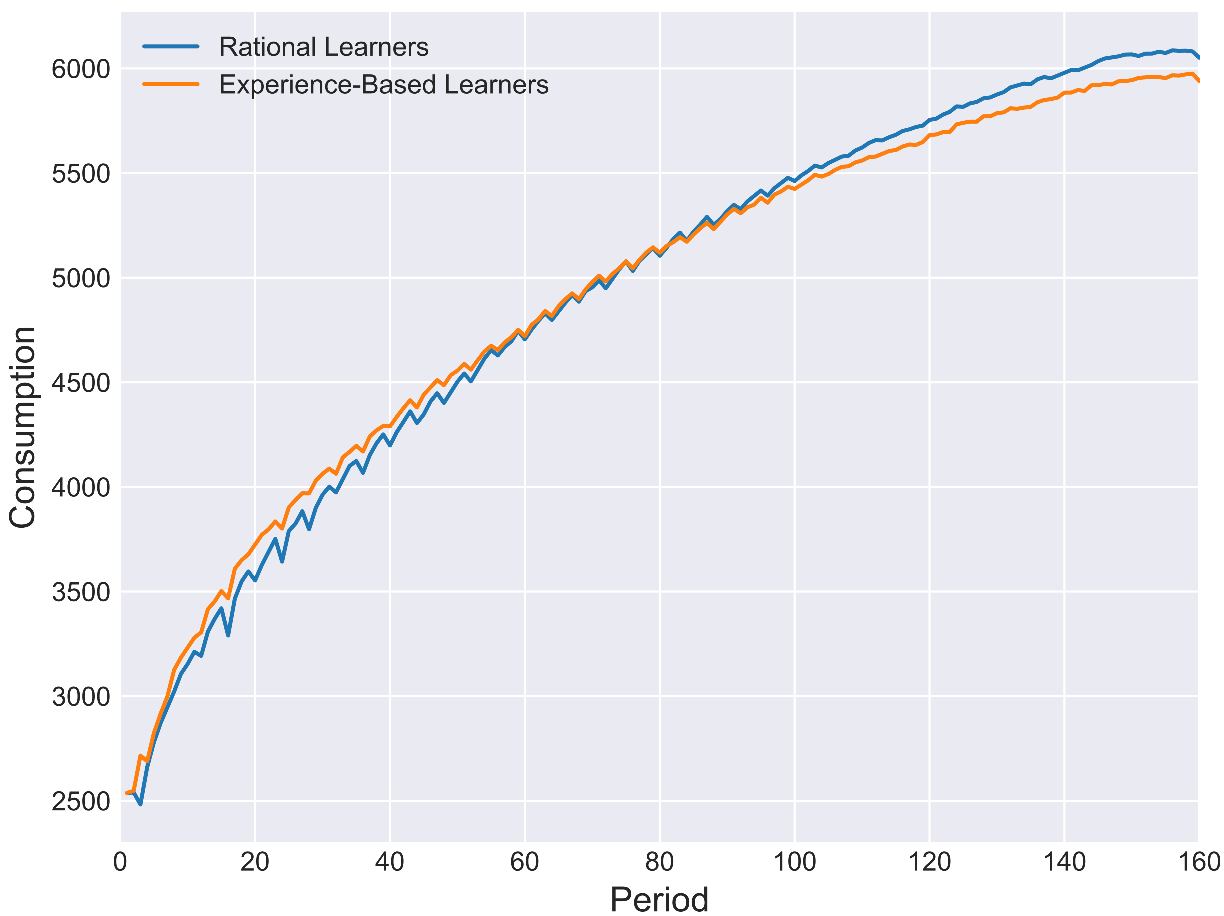
Notes. Estimations with the simulated consumption values as the dependent variable and the simulated same-period income and wealth values as the regressors for rational consumers in columns (1) and (2), and experienced-based learning (EBL) consumers in columns (3) and (4). Estimations are for the high-education subgroup with $\lambda = 1$. Rational consumers hold a constant belief about the probability of being employed next period, and EBL consumers form beliefs based on their prior employment history as specified in equations (2.1) and (2.2). All estimations include period fixed effects and use period-clustered standard errors. t statistics in parentheses. Simulations are based on the working periods of 10,000 simulated consumers and thus 1,600,000 observations.

Table B.17: **High-Education Estimations with Model-Simulated Data with Unemployment Scarring**

	(1) Rational	(2) Rational	(3) EBL	(4) EBL
Income	0.782	0.630	0.790	0.709
	(123.24)	(211.43)	(264.02)	(157.42)
Wealth		0.0168		0.0110
		(314.58)		(24.79)
Unemployment Experience	3398.3	3978.6	-2874.0	-4081.8
	(5.16)	(4.49)	(-9.54)	(-10.64)

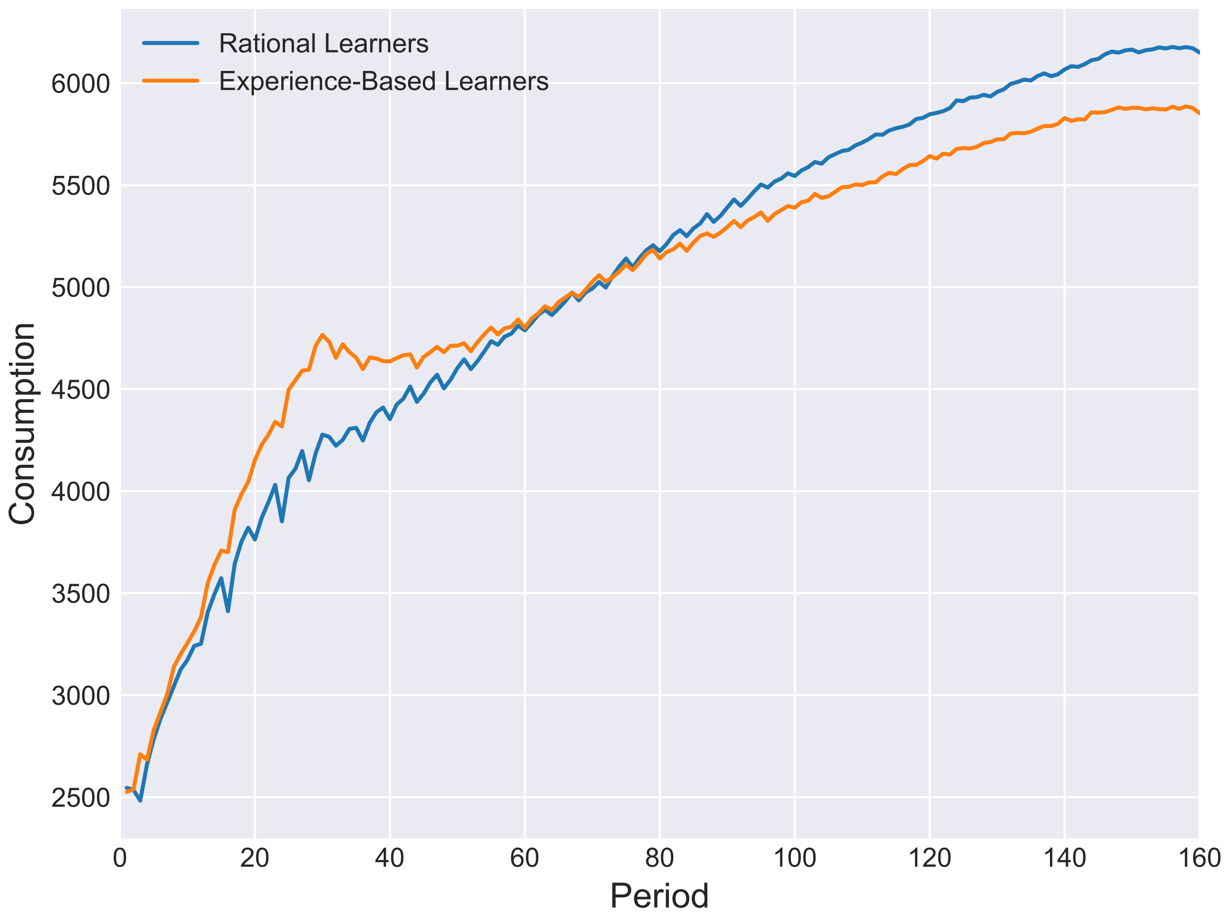
Notes. Estimations with the simulated consumption values as the dependent variable and the simulated same-period income and wealth values as the regressors for rational consumers in columns (1) and (2), and experienced-based learning (EBL) consumers in columns (3) and (4). Estimations are for the high-education subgroup with $\lambda = 1$. Rational consumers hold a constant belief about the probability of being employed next period, and EBL consumers form beliefs based on their prior employment history as specified in equations (2.1) and (2.2). All estimations include period fixed effects and use period-clustered standard errors. t statistics in parentheses. Simulations are based on the working periods of 10,000 simulated consumers and thus 1,600,000 observations.

Figure B.1: **Average Life-Cycle Consumption**



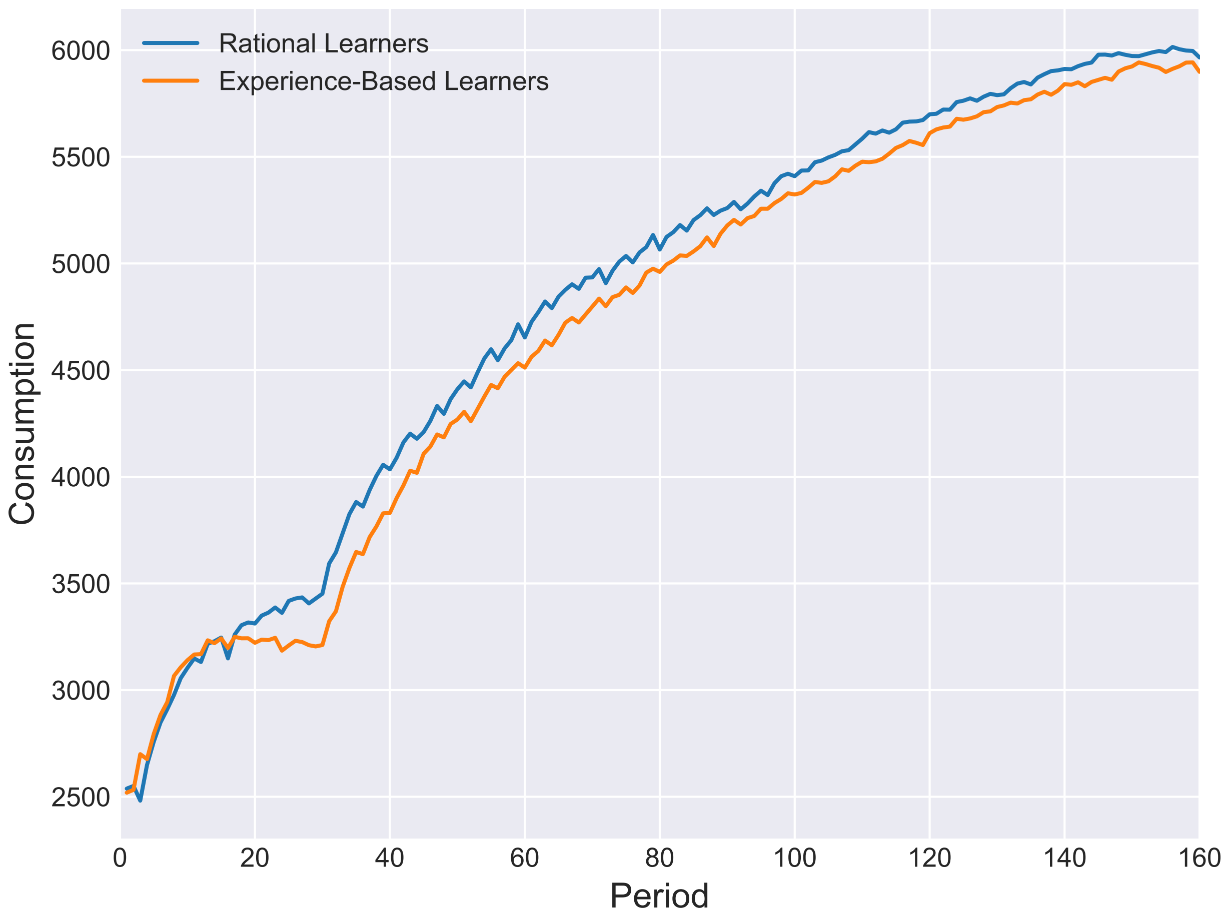
Notes. Average consumption for rational learners and experience-based learners (with $\lambda = 1$) in the low-education group, based on 10,000 lifetime simulations for each type.

Figure B.2: **Average Life-Cycle Consumption for Agents with Good Realizations Early in Life**



Notes. Average consumption for rational learners and experience-based learners (with $\lambda = 1$) in the low-education group, based on 10,000 lifetime simulations for each type and then restricted to those simulations where agents have, or in the rational case would have, a believed delta of 0.025 or less at period 30.

Figure B.3: **Average Life-Cycle Consumption Patterns for Agents with Bad Realizations Early in Life**



Notes. Average consumption for rational learners and experience-based learners (with $\lambda = 1$) in the low-education group, based on 10,000 lifetime simulations for each type and then restricted to those simulations where agents have, or in the rational case would have, a believed delta of 0.1 or greater at period 30.

Appendix C

Appendix for Chapter 3

C.1 Additional Figures and Tables

Table C.1: **Foreign Chinese Housing Demand and Housing Prices, Alternative IV**

	(1) ln(Zillow)	(2) ln(TV)	(3) ln(Emp)	(4) ln(NT Emp)	(5) ln(T Emp)
ln(CHTV) $\times \mathbb{I}\{year \geq 2007\}$	0.055***	0.087***	0.081**	0.093***	0.095
	(0.015)	(0.018)	(0.034)	(0.032)	(0.080)
ln(CHTV)	0.037*	-0.004	0.045	-0.030	0.193
	(0.021)	(0.024)	(0.065)	(0.064)	(0.144)
$\mathbb{I}\{t \geq 2007\}$	0.157***	-0.017	-0.094	-0.045	-0.462***
	(0.018)	(0.017)	(0.059)	(0.057)	(0.128)
ln(Population)	-0.042***	-0.011	0.754***	0.900***	0.862***
	(0.016)	(0.015)	(0.075)	(0.067)	(0.136)
Δ ln(Y), 00-96	1.340***	0.209**	0.302	-0.115	-0.164
	(0.212)	(0.099)	(0.192)	(0.127)	(0.112)
Education	4.203***	5.246***	2.381***	2.635***	-4.483***
	(0.253)	(0.146)	(0.606)	(0.605)	(1.203)
County Fixed Effects	X	X	X	X	X
Post Period	2007-2013	2007-2013	2007-2013	2007-2013	2007-2013
Model Statistics:					
First Stage F-statistic	143.78	154.60	162.55	164.07	155.58
Observations	4588	4883	4900	4893	4828

Notes: The dependent variables are log Zillow Housing Price Index (column 1), log housing transaction value (column 2), log total employment (column 3), log non-tradable sector employment (column 4), and log tradable sector employment (column 5) . $CHTV$ denotes foreign Chinese housing transaction values instrumented by China's GDP weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. $\mathbb{I}\{t \geq 2007\}$ is an indicator variable that takes the value 1 if year is post-2007 and 0 otherwise. Education is measured as the population share with bachelor degrees. Additional control includes a pre-sample trend variable for the dependent variable (Y), calculated as the change in the respective housing price variable between 1996 and 2000. The sample period runs from 2001-2013. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

Table C.2: **Foreign Chinese Housing Demand (Count) and Housing Prices**

	(1)	(2)	(3)	(4)	(5)
ln(CHTC) $\times \mathbb{I}\{year \geq 2007\}$	0.225***	0.180***	0.133***	0.119***	0.154***
	(0.032)	(0.023)	(0.023)	(0.023)	(0.026)
ln(CHTC)	-0.102***	-0.083***	-0.037	-0.021	-0.025
	(0.031)	(0.026)	(0.026)	(0.025)	(0.027)
$\mathbb{I}\{t \geq 2007\}$	0.224***				
	(0.025)				
ln(Population)	-0.042**	-0.037**	-0.046**	-0.049***	-0.036**
	(0.019)	(0.018)	(0.019)	(0.019)	(0.018)
Δ ln(HNW), 00-96	1.357***	1.340***	1.343***	1.291***	1.389***
	(0.238)	(0.232)	(0.234)	(0.221)	(0.266)
Education	4.797***	4.691***	4.653***	4.438***	4.502***
	(0.235)	(0.229)	(0.232)	(0.218)	(0.264)
County Fixed Effects	X	X			
Year Fixed Effects		X			
County Year Fixed Effects			X	X	X
Post Period	2007-2013	2007-2013	2007-2013	2007-2011	2012-2013
Model Statistics:					
First Stage F-statistic	103.06	115.15	90.04	74.56	86.50
Observations	4588	4588	4571	3474	2470

Notes: The dependent variable is log Zillow Single Family Home Value Index. $CHTC$ denotes foreign Chinese housing transaction count instrumented by the aggregate housing transaction count in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. $\mathbb{I}\{t \geq 2007\}$ is an indicator variable that takes the value 1 if year is post-2007 and 0 otherwise. Education is measured as the population share with bachelor degrees. Additional control variable includes a pre-sample trend variable for the dependent variable calculated as the change in Zillow Home Value Index between 1996 and 2000. Columns 4 shows the results for the housing crash period (2007-2011); Column 5 for the recovery period (2012-2013). The sample period runs from 2001-2013. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

Table C.3: **Foreign Chinese Housing Demand (Count) and Housing Transaction Prices**

	(1)	(2)	(3)	(4)	(5)
ln(CHTC) $\times\mathbb{I}\{year \geq 2007\}$	0.268***	0.215***	0.174***	0.179***	0.167***
	(0.033)	(0.025)	(0.025)	(0.027)	(0.024)
ln(CHTC)	-0.144***	-0.121***	-0.081***	-0.074***	-0.056**
	(0.030)	(0.025)	(0.024)	(0.023)	(0.022)
$\mathbb{I}\{t \geq 2007\}$	0.053**				
	(0.024)				
ln(Population)	-0.018	-0.012	-0.021	-0.030*	-0.007
	(0.018)	(0.017)	(0.017)	(0.018)	(0.017)
Δ ln(HTV), 00-96	0.214*	0.181*	0.176	0.145	0.138
	(0.111)	(0.109)	(0.110)	(0.110)	(0.110)
Education	5.719***	5.592***	5.568***	5.432***	5.195***
	(0.151)	(0.144)	(0.144)	(0.146)	(0.149)
County Fixed Effects	X	X			
Year Fixed Effects		X			
County Year Fixed Effects			X	X	X
Post Period	2007-2013	2007-2013	2007-2013	2007-2011	2012-2013
Model Statistics:					
First Stage F-statistic	113.27	126.49	98.17	81.25	88.76
Observations	4883	4883	4866	3699	2633

Notes: The dependent variable is log housing transaction values, averaged by zip code, from DataQuick. $CHTC$ denotes foreign Chinese housing transaction count instrumented by the aggregate housing transaction count in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. $\mathbb{I}\{t \geq 2007\}$ is an indicator variable that takes the value 1 if year is post-2007 and 0 otherwise. Education is measured as the population share with bachelor degrees. Additional control variable includes a pre-sample trend variable for the dependent variable, calculated as the change in housing transaction values between 1996 and 2000, and variables on home characteristics including the number of bathrooms, the square footage, and age of the home. Columns 4 shows the results for the housing crash period (2007-2011); Column 5 for the recovery period (2012-2013). The sample period runs from 2001-2013. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

Table C.4: **Foreign Chinese Housing Demand (Count) and Total Employment**

	(1)	(2)	(3)	(4)	(5)
ln(CHTC) $\times\mathbb{I}\{year \geq 2007\}$	0.195***	0.162***	0.192***	0.169***	0.229***
	(0.058)	(0.055)	(0.063)	(0.062)	(0.072)
ln(CHTC)	-0.005	0.006	-0.018	-0.004	-0.013
	(0.076)	(0.077)	(0.086)	(0.086)	(0.090)
$\mathbb{I}\{t \geq 2007\}$	-0.081				
	(0.069)				
ln(Population)	0.734***	0.739***	0.728***	0.741***	0.732***
	(0.079)	(0.078)	(0.080)	(0.081)	(0.090)
Δ ln(Emp), 00-96	0.336*	0.322*	0.329*	0.393**	0.355*
	(0.194)	(0.193)	(0.195)	(0.196)	(0.212)
Education	3.101***	3.042***	3.018***	2.865***	2.971***
	(0.511)	(0.511)	(0.517)	(0.536)	(0.554)
County Fixed Effects	X	X			
Year Fixed Effects		X			
County Year Fixed Effects			X	X	X
Post Period	2007-2013	2007-2013	2007-2013	2007-2011	2012-2013
Model Statistics:					
First Stage F-statistic	117.30	129.67	99.99	82.20	90.03
Observations	4900	4900	4883	3712	2643

Notes: The dependent variable is log total employment size. *CHTC* denotes foreign Chinese housing transaction count instrumented by the aggregate housing transaction count in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. $\mathbb{I}\{t \geq 2007\}$ is an indicator variable that takes the value 1 if year is post-2007 and 0 otherwise. Education is measured as the population share with bachelor degrees. Additional control variable includes a pre-sample trend variable for the dependent variable, calculated as the change in housing transaction values between 1996 and 2000, and variables on home characteristics including the number of bathrooms, the square footage, and age of the home. Columns 4 shows the results for the housing crash period (2007-2011); Column 5 for the recovery period (2012-2013). The sample period runs from 2001-2013. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

Table C.5: **Testing for Migration Channel: Number of Tax Return Filings (Lead)**

	(1)	(2)	(3)	(4)	(5)
ln(CHTV) $\times\mathbb{I}\{year \geq 2007\}$	-0.047***	-0.042***	-0.035***	-0.036***	-0.032***
	(0.011)	(0.011)	(0.009)	(0.009)	(0.010)
ln(CHTV)	0.027*	0.024*	0.017	0.015	0.013
	(0.015)	(0.014)	(0.013)	(0.013)	(0.014)
$\mathbb{I}\{t \geq 2007\}$	0.060***				
	(0.009)				
ln(Population)	0.904***	0.905***	0.905***	0.905***	0.924***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.018)
Education	1.003***	0.997***	1.009***	1.046***	1.137***
	(0.114)	(0.113)	(0.114)	(0.115)	(0.109)
Δ ln(Returns), 01-98	0.699***	0.699***	0.699***	0.684***	0.677***
	(0.104)	(0.103)	(0.104)	(0.101)	(0.094)
County Fixed Effects	X	X			
Year Fixed Effects		X			
County Year Fixed Effects			X	X	X
Post Period	2007-2013	2007-2013	2007-2013	2007-2011	2012-2013
Model Statistics:					
First Stage F-statistic	111.60	116.15	106.70	100.90	73.77
Observations	4026	4026	4012	3431	1791

Notes: The dependent variable is log income tax returns from year $t+1$. $CHTV$ denotes foreign Chinese housing transaction values instrumented by the aggregate housing transaction value in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. Education is measured as the population share with bachelor degrees. Δ ln(Returns), 01-98 is a pre-sample trend variable for the dependent variable, calculated as the change in the numbers of returns between 1998 and 2001. Columns 4 shows the results for the housing crash period (2007-2011); Column 5 for the recovery period (2012-2013). The sample period runs from 2001-2013. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

Table C.6: **Foreign Chinese Housing Demand (Count) and Tradable/Non-Tradable Sector Employment**

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(NT Emp)	ln(T Emp)	ln(NT Emp)	ln(T Emp)	ln(NT Emp)	ln(T Emp)
ln(CHTC) $\times \mathbb{I}\{year \geq 2007\}$	0.203***	0.173	0.182***	0.123	0.203***	0.253
	(0.061)	(0.136)	(0.060)	(0.135)	(0.061)	(0.157)
ln(CHTC)	-0.103	0.207	-0.092	0.223	-0.103	0.227
	(0.084)	(0.190)	(0.085)	(0.189)	(0.084)	(0.199)
ln(Population)	0.880***	0.821***	0.886***	0.876***	0.880***	0.813***
	(0.071)	(0.145)	(0.073)	(0.149)	(0.071)	(0.158)
Δ ln(NT/T Emp), 00-96	-0.093		-0.065		-0.093	
	(0.129)		(0.136)		(0.129)	
Education	2.868***	-2.922***	2.710***	-3.045***	2.868***	-3.205***
	(0.477)	(1.070)	(0.496)	(1.117)	(0.477)	(1.161)
Δ ln(NT/T Emp), 00-96		-0.164		-0.158		-0.112
		(0.115)		(0.122)		(0.122)
County Year Fixed Effects	X	X	X	X	X	X
Post Period	2007-2013	2007-2013	2007-2011	2007-2011	2012-2013	2012-2013
Model Statistics:						
First Stage F-statistic	101.92	96.23	84.94	79.88	101.92	85.89
Observations	4876	4811	3708	3668	4876	2607

Notes: The dependent variable is log non-tradable or tradable sector employment. $CHTC$ denotes foreign Chinese housing transaction count instrumented by the aggregate housing transaction count in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. $\mathbb{I}\{t \geq 2007\}$ is an indicator variable that takes the value 1 if year is post-2007 and 0 otherwise. Education is measured as the population share with bachelor degrees. Additional control variable includes a pre-sample trend variable for the dependent variable, calculated as the change in housing transaction values between 1996 and 2000, and variables on home characteristics including the number of bathrooms, the square footage, and age of the home. Columns 4 shows the results for the housing crash period (2007-2011); Column 5 for the recovery period (2012-2013). The sample period runs from 2001-2013. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

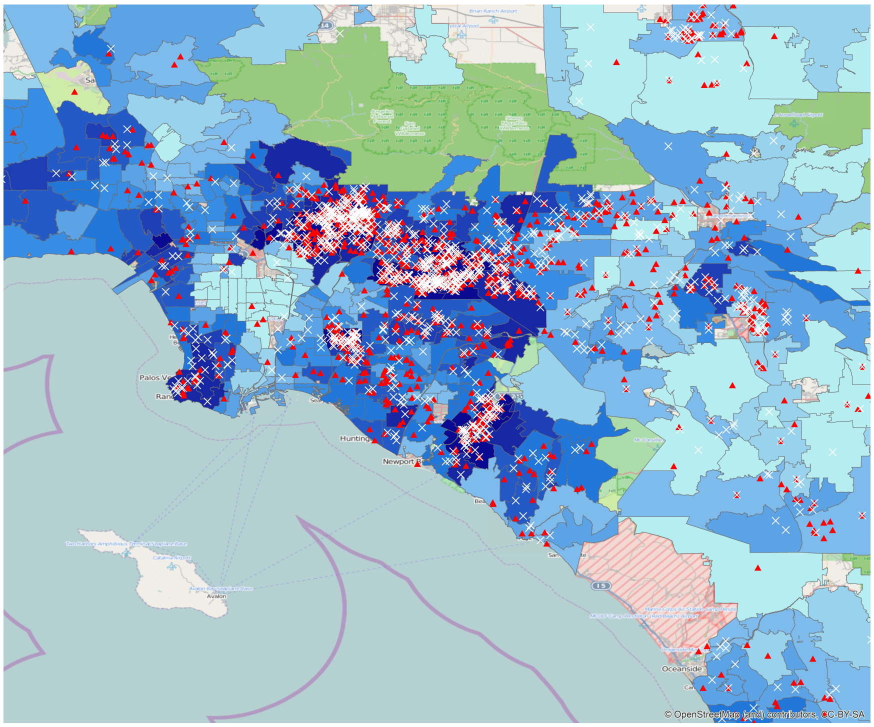
Table C.7: **Foreign Chinese Housing Demand and Foreclosure (Count)**

	(1)	(2)	(3)	(4)	(5)
ln(CHTV) $\times\mathbb{I}\{year \geq 2007\}$	-0.503***	-0.439***	-0.250***	-0.290***	-0.157***
	(0.065)	(0.060)	(0.059)	(0.065)	(0.056)
ln(CHTV)	0.138**	0.133**	-0.038	-0.040	-0.159***
	(0.067)	(0.064)	(0.062)	(0.063)	(0.061)
$\mathbb{I}\{t \geq 2007\}$	2.060***				
	(0.060)				
ln(Population)	0.619***	0.597***	0.609***	0.630***	0.525***
	(0.070)	(0.068)	(0.070)	(0.069)	(0.066)
Education	-4.772***	-4.703***	-4.421***	-4.321***	-2.056***
	(0.466)	(0.460)	(0.472)	(0.479)	(0.486)
County Fixed Effects	X	X			
Year Fixed Effects		X			
County Year Fixed Effects			X	X	X
Post Period	2007-2013	2007-2013	2007-2013	2007-2011	2012-2013
Model Statistics:					
First Stage F-statistic	125.99	129.84	120.23	108.98	92.92
Observations	4900	4900	4883	3712	2643

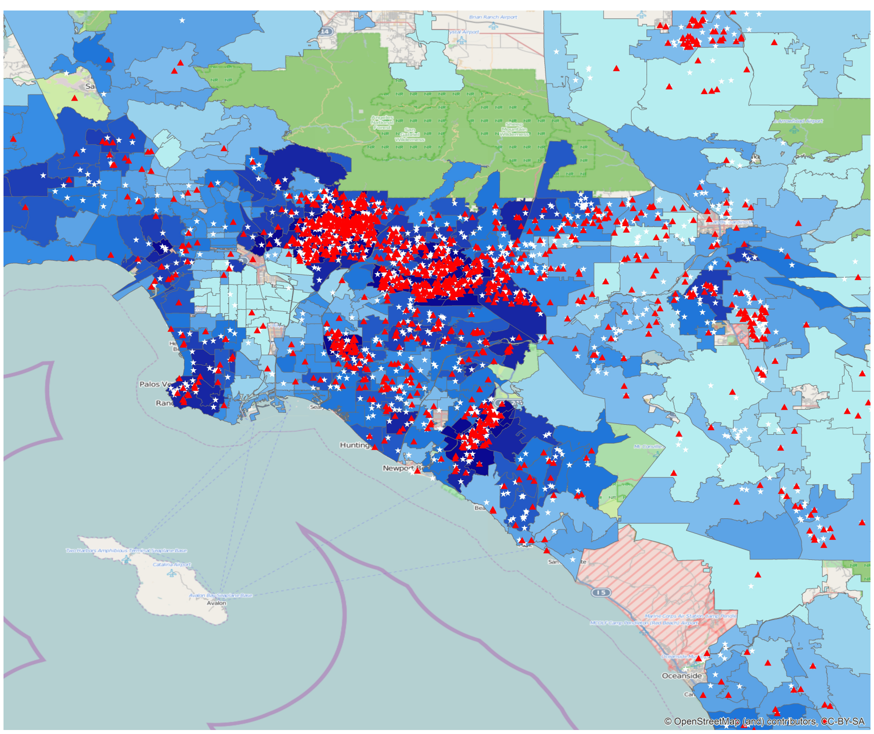
Notes: The dependent variable is log foreclosure count. $CHTV$ denotes foreign Chinese housing transaction values instrumented by the aggregate housing transaction value in California weighted by the share of ethnic Chinese population across zip codes from the pre-sample period. $\mathbb{I}\{t \geq 2007\}$ is an indicator variable that takes the value 1 if year is post-2007 and 0 otherwise. Education is measured as the population share with bachelor degrees. Additional control includes a pre-sample trend variable for the dependent variable, calculated as the change in foreclosure count between 1996 and 2000. Columns 4 shows the results for the housing crash period (2007-2011); Column 5 for the recovery period (2012-2013). The sample period runs from 2001-2013. Standard errors are clustered at the zip code level. *, **, *** denote 10%, 5% and 1% significance respectively.

Figure C.1: **Housing Purchases by Foreign Chinese in the Los Angeles Region**

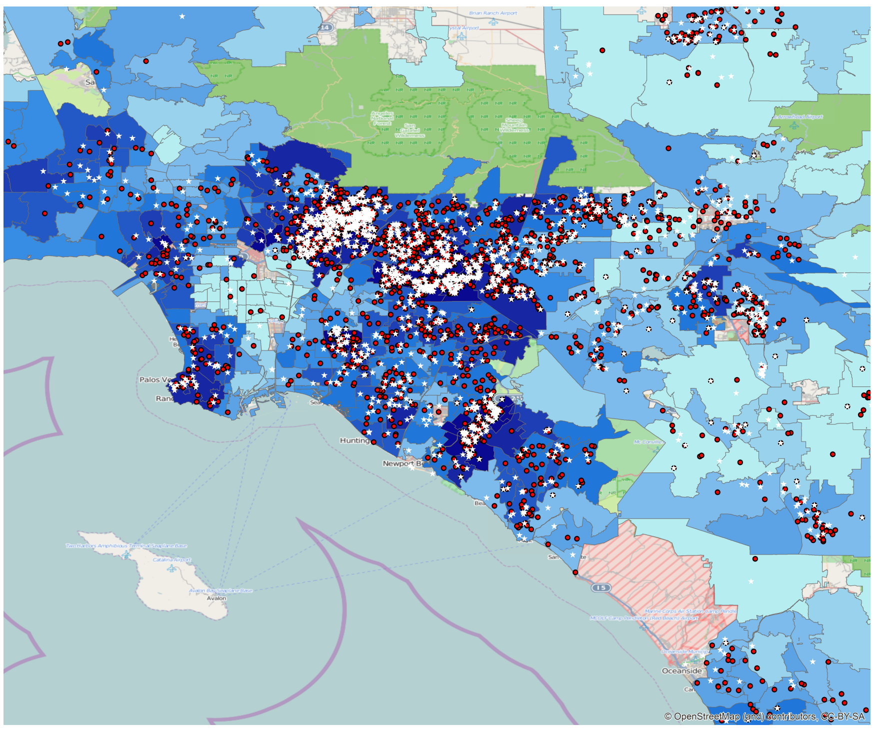
(a) 2010-2011



(b) 2012-2013



(c) 2012-2013



Notes: This figure illustrates housing purchases by foreign Chinese in the Los Angeles Region from 2010 to 2013. The blue shades in the background divides the region based on zip codes, where darker shades represent zip codes with a higher share of ethnic Chinese population in 2000, based on the Census data. Figure (a) illustrates housing purchases by Foreign Chinese in 2010 and 2011, where the white "X"s denote purchases in 2010 and the red triangles denote purchases in 2011. Figure (b) illustrates housing purchases by Foreign Chinese in 2011 and 2012, where the red triangles denote purchases in 2011 and the white stars denote purchases in 2012. Figure (c) illustrates housing purchases by Foreign Chinese in 2012 and 2013, where the white stars denote purchases 2012 and the red circles denote purchases in 2013.