6-17-2020 2:00 PM

Essays on Housing Markets

Yuxi Yao, *The University of Western Ontario*

Supervisor: MacGee, James C., *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Economics

Abstract

Housing assets have long been the most important assets in household portfolios. Therefore, understanding the long-run trends of the key indicators in the housing market is important for understanding household housing tenure choice and asset allocation, evaluating household welfare, and analyzing financial stability. My thesis consists of three chapters accounting for the changes in homeownership rates, cross-city distribution of prices and rents, and aggregate housing prices in the United States for the past several decades.

Chapter 2 focuses on the decline in the homeownership rate among young households since 1980. I find that while some of the college graduates merely postpone home purchases, a considerable fraction of non-college graduates have become long-term renters. I show the diverging ownership decisions between college and non-college graduates is driven by changes in the income distribution, due to an increasing population share of college graduates and unbalanced income growth among young and middle-aged college graduates. My findings suggest that low-income, non-college graduates are facing an affordability problem.

Chapter 3 focuses on cross-city variation in the two shelter costs: prices and rents. Since owning and renting are the most prevalent options to obtain housing services, understanding the joint distribution of prices and rents across cities has important implications on the ownership decisions and life quality in each city. I first document three stylized facts about the distributions of prices and rents across cities in the U.S.: (i) prices are more dispersed compared to rents across cities; (ii) the dispersion of house prices has increased more than the dispersion of rents from 1980 to 2010; (iii) prices and rents are highly correlated in both levels and growth rates. As most owners live in detached houses while most renters live in apartments, this chapter examines the implication of the difference in land use between houses and apartments on these observations. I develop a city-level housing tenure choice model where owner-occupied houses take more land to build compared to rental apartments. I calibrate the model to house prices, rents, and the fraction of households living in houses for each of the largest 181 cities in the U.S. in 1980. Feeding in the model population, income, down payment requirement, and residential land supply in 2010, I show the model can account for 82% of the large increase in house price dispersion and 56% of the increase in rent dispersion from 1980 to 2010.

Chapter 4, which is co-authored with Yifan Gong, investigates the contribution of four demographic-related factors i.e. changing fertility, rising life expectancy, urbanization, and international immigration, on the growth of the aggregate housing price since 1970. Conceptually, the total housing demand is determined by the age profile of housing demand aggregated over the age distribution of the population. Among these four factors, declining fertility, rising life expectancy, and international immigration affect the age distribution of the population. In addition, rising life expectancy changes the age profile of housing demand. Specifically, it leads to an increase in the housing demand for senior households. Urbanization that moves people from rural areas with high supply elasticity to urban areas with low supply elasticity further increases house prices. To quantitatively evaluate the importance of these four factors and to make projections on future house prices, we develop a general equilibrium model and find these four factors can account for 41% of the observed housing price growth from 1970 to 2010. Applying the projected changes in these four factors, we predict housing prices will keep growing by about 5% to 25% from 2010 to 2050. The growth rates vary with urbanization rates and the levels of immigration.

Keywords: Home Ownership, College Share, Income Growth, Dispersion of Prices, Dispersion of Rents, Land Intensity, Population Structure, Housing Demand, Urbanization

Summary for Lay Audience

My thesis consists of three chapters that explain the long-run trends in homeownership rates, cross-city variation in prices and rents, and the contribution of demographic-related factors on the aggregate housing prices, in the United States.

Chapter 2 focuses on the decline in the homeownership rate among young households since 1980. I find that while some of the college graduates merely postpone home purchases, a considerable fraction of non-college graduates have become long-term renters. I show that the diverging ownership decisions between college and non-college graduates can be accounted for by the increase in the numbers of high-income college graduates who experienced an increase in income, especially among the middle-aged households. These changes result in a higher fraction of high-income households, which pushes up the housing price. As a result, non-college graduates find owning less affordable, while college graduates who expect higher income in the future choose to delay home purchases.

Chapter 3 focuses on the distribution of the two shelter costs: prices and rents, across cities. I find that prices are more dispersed across cities compared to rents. Moreover, the dispersion of prices has increased more than rents over time. Motivated by the fact that most owners live in detached houses while most renters live in apartments, this chapter examines the implication of the land use difference between houses and apartments on the joint distribution of prices and rents across cities. Land values vary across cities with economic fundamentals, such as income and population. As houses use more land to produce, the cost of building houses consequently varies more compared to apartments.

Chapter 4 (co-authored with Yifan Gong) investigates the contribution of four demographic-related factors, i.e. changing fertility, rising life expectancy, urbanization, and international immigration, on the growth of the aggregate housing price in a general equilibrium framework. Our estimated model shows that changes in these factors can explain 41% of the price growth from 1970 to 2010. We also find that expected changes in these factors predict a sustained housing price growth from 2010 to 2050.

Co-Authorship Statement

This thesis contains co-authored material. Chapters 4 is co-authored with Yifan Gong. Both authors are equally responsible for the work.

Acknowledgments

I am heavily indebted to my advisor James MacGee for his continued support and guidance. He has served as a great supervisor, an excellent mentor, and an ideal role model for a junior researcher. During all these years, I have learned tremendously from Jim about critical thinking, conducting original research, and, most importantly about being open-minded, grateful, and positive. I am extremely grateful to him for the time he spent on helping me improve my drafts and prepare presentations. Without his unwavering encouragement and advice, I would not be able to meet the challenges and complete this thesis.

I would like to offer my deepest gratitude to my committee members, Simona Cociuba and Elizabeth Caucutt, who have provided invaluable comments and suggestions on various aspects of my research and my life. Besides my thesis committee, I have also benefited substantially from the feedbacks provided by the participants of the Brownbag Seminar at Western, including, but certainly not limited to, Tim Conley, Ananth Ramanarayanan, Igor Livshits, and Jacob Short. I would like to thank Todd Stinebrickner and Nirav Mehta for their incisive suggestions on the motivation of the third chapter of my thesis and their help in the job search process.

During my Ph.D. studies, I have been fortunate to interact with many intelligent and generous researchers. I wish to thank Andrii Parkhomenko and Jason Allen for their detailed comments on the early draft of the third chapter of my thesis. I am grateful for the internship opportunity provided by the Bank of Canada. I would like to thank Joel Wagner, Katya Kartashova, Cesaire Meh, Geoffrey Dunbar, Brian Peterson, Nuno Paixao, and Oleksiy Kryvtsov for their insightful comments on my research.

I greatly enjoyed the time at Western with my fellow students and my friends. I would like to thank Wenya Wang, Sha Wang, Qian Liu, Ziyu Zhang, Jin Zhou, Phuong Vu, Zhuang Liu, Matthew Carew, Samantha Black, Youngmin Park, Youjin Choi, and Hyeongsuk Jin for their companion and support. I am also grateful for the service and support I receive from the Department. I would like to thank Sandra Augustine, Karin Feulgen, Sharon Phillips, Gary Kim, Leslie Kostal, Debra Merrifield, and Maureen O'Connell.

I am truly grateful for the constant love and support from my parents and my grandparents. They have tried their best to provide me with the ground to pursue my interest ever since I was born. My last but deepest thanks are to my beloved husband and co-author, Yifan Gong. I could not imagine my life to be half as happy as I am if I were not fortunate enough to meet him. I am grateful for the time that we spend together and excited about the endeavors that we are going to take in the future.

Contents

List of Figures

List of Tables

List of Appendices

Chapter 1

Introduction

Housing assets have long been the most important part of household portfolios. The Federal Reserve's 2016 Survey of Consumer Finance reveals that 70% of the net wealth of the median household is in residential housing. Therefore, understanding the long-run trends of the key indicators in the housing market, such as house prices, rents, and homeownership rates, is important for understanding households housing tenure choice and asset allocation, evaluating household welfare, and analyzing financial stability.

My thesis consists of three chapters accounting for changes in different aspects of the U.S. housing market. The second chapter focuses on the decline in the homeownership rate among young households. In the third chapter, I develop a model that can quantitatively account for the cross-city variation in the two shelter costs: prices and rents, and the changes in these two distributions over time. The fourth chapter looks at the impact of changes in four demographic-related factors on the aggregate housing price and predicts future housing prices.

The second chapter, "Accounting for the Decline in Homeownership Among the Young", studies the decline in the homeownership rate amongst young households since 1980. The previous literature suggests that the drop in young homeownership rate of the young is temporary, i.e., households are postponing the purchase of their first home. Most recently, there are papers and reports that highlight the impact of rising student loans on delaying home purchasing for college graduates. In this chapter, I document that while some college graduates postpone home purchases, a considerable fraction of non-college graduates have become long-term renters. I provide a mechanism that shows the diverging homeownership decisions between college and non-college graduates can be accounted for by a change in the income distribution due to an increasing population share of college graduates and unbalanced income growth among young and middle-aged college graduates. Specifically, middle-aged college graduates see a larger increase in their household income, compared to young households. These changes result in a higher fraction of high-income households, which pushes up the equilibrium housing price. As a result, non-college graduates find owning less attractive, while college graduates who expect higher income in the future choose to delay home purchases. I examine the implications of this mechanism using cross-city variation in house prices, college share and homeownership rates and incomes of young and middle-aged, college and non-college households. I find that changes in the income distribution can account for the majority of the observed changes in young and middle-aged homeownership rates for both college and non-college graduates for all years from 1980 to 2018. My findings suggest that low-income, non-college graduates are having a more severe affordability problem.

The third chapter, ”Land and the Rise in the Dispersion of House Prices and Rents across U.S. Cities”, focuses on cross-city variation in the two shelter costs: prices and rents. Shelter cost is the largest component of household expenditures and there have been large swings in the two shelter costs across cities since 1980, which affect homeownership as well as living standards across cities. In this chapter, I first document three stylized facts about the distributions of prices and rents across cities in the U.S.: (i) while both prices and rents vary across cities, prices are more dispersed compared to rents; (ii) the dispersion of house prices has increased more than the dispersion of rents from 1980 to 2010; (iii) despite the discrepancy between the two distributions, prices and rents are highly correlated both in terms of levels and growth rates. Cities with higher rents usually have even higher prices. Motivated by the fact that most owners live in detached houses while most renters live in apartments, this chapter examines the implication of the land use difference between houses and apartments on these stylized facts. I develop a city-level housing tenure choice model where owner-occupied houses take more land to build compared to rental apartments. When land values vary across cities, house prices vary more compared to rents due to the intensive use of land in the construction of houses. I calibrate the model to house prices, rents, and the fraction of households living in detached/attached homes for each of the largest 181 cities in the U.S. in 1980. Feeding in the model population, income, down payment requirement, and residential land supply in 2010, I show the model can account for 82% of the large increase in house price dispersion and 56% of the moderate increase in rent dispersion. In addition, the model generates an increase in the dispersion of price-rent ratios that matches the data. It suggests the increase in the price-rent ratios in big cities can be explained by changes in economic fundamentals. It challenges the view that the price-rent ratio can be used as an indicator of housing bubbles. In addition, this chapter has implications on inferring the user cost of owner-occupied houses from the observed rent of rental units.

In the fourth chapter, “Demographics and the Housing Market”, my coauthor and I investigate the contribution of four demographic-related factors, i.e. changing fertility, rising life expectancy, urbanization, and international immigration, on the growth of the aggregate housing price since 1970. Conceptually, the total housing demand is determined by the age profile of housing demand aggregated over the age distribution of the population. Among these four factors, changing fertility, rising life expectancy, and international immigration affect the age distribution of the population. In addition, we propose and quantify another channel through which rising life expectancy increases the housing demand for senior households. In other words, rising life expectancy changes the age profile of housing demand. Moreover, urbanization that moves people from rural areas with high supply elasticity to urban areas with low supply elasticity further increases housing prices. To quantitatively evaluate the importance of these four factors and to make projections on future house prices, we develop a general equilibrium model and find these four factors can account for 41% of the observed housing price growth from 1970 to 2010. In addition, we find that among these factors, urbanization makes the largest contribution. Applying the projected changes in these four factors provided by the Census and the United Nations, we use our model to predict housing prices through 2050. We find that the housing price will keep growing in the next 40 years by about 5% to 25%. The growth rates vary with urbanization rates and the levels of immigration.

Chapter 2

Accounting for the Decline in Homeownership among the Young

2.1 Introduction

The past four decades have witnessed a significant decline in the homeownership rate of households with "heads" aged 25-34. While the aggregate homeownership rate has been stable around 68%, young households saw a 10 percentage point drop in their ownership rate. Previous literature suggests that the drop in young homeownership rate is temporary, i.e., households are postponing the purchase of their first home (Fisher and Gervais, 2011; Anagnostopoulos, Atesagaoglu, and Carceles-Poveda, 2013). Some more recent studies highlight the impact of rising student loans on delaying home purchases by college graduates (see e.g. Mezza, Ringo, Sherlund, and Sommer, 2020). I find that while some college graduates postpone buying their first home, a considerable fraction of non-college graduates have become long-term renters.

This chapter proposes a mechanism that can largely account for the diverging ownership decisions between college- and non-college-educated households from 1980 to 2010. The mechanism is motivated by the observed changes in the income distribution that have been caused by (i) unbalanced income growth among college graduates; (ii) a rising share of college graduates. Specifically, college graduates have enjoyed an increase in their household income, especially among those that are middle-aged during the past several decades. Meanwhile, non-college graduates have barely seen any growth in their household income. I show these changes, when examined through the lens of a general equilibrium model, can quantitatively account for the delayed purchasing of college graduates and the switch towards renting by non-college graduates.

An increase in the household income of college graduates drives up the aggregate housing demand for owner-occupied units. As the supply is not perfectly elastic (see e.g. Glaeser, Gyourko, and Saks, 2005), a rise in demand leads to higher equilibrium house prices. As a result, a considerable fraction of households headed by non-college graduates find owning less affordable and become long-term renters. In other words, an increase in the income gap between college and non-college graduates shifts ownership from non-college-educated households to college-educated households. Meanwhile, homeownership has shifted from young to middle-aged college graduates as the middle-aged ones have seen a larger increase in their household income. In the presence of credit constraints, i.e., down payment requirements, young college graduates postpone the purchase of their first home as their income profiles have become steeper. In addition, a growing share of college graduates results in a lower ownership rates for both college and non-college graduates by fueling house price increases, with the downward pressure on ownership for college graduates partially offset by the increase in their household income.

To illustrate the mechanism and to guide the empirical analysis, I develop a stylized three-period tenure choice model which shows that changes in income and college share can affect the ownership decisions of college and non-college graduates in different ways. My model extends the frame-work of Ortalo-Magne and Rady (2006) to allow for two types of households, College and Non-college, to capture the widening gap in their household incomes and the divergence in their ownership decisions. Within each type, households differ in preference towards owning and in the endowment streams which are described by an ability ranking. Conditional on the ability ranking and age, college graduates earn more than non-college graduates. Moreover, the lifetime earning profile for college graduates is steeper than that for non-college graduates. For all households, owning is preferred to renting. The owner-occupied units are in limited supply.

The model yields several testable implications on house prices and homeownership rates of four groups of households (henceforth “the four groups”): young college, young non-college, middle-aged college, and middle-aged non-college. The comparison between young and middle-aged households from the same educational background allows me to distinguish delaying home purchasing from switching to long-term renters. The model implies that, first, house prices are increasing in college share, while homeownership rates of all groups are declining in college share. Second, homeownership rates of non-college graduates are decreasing in the income of college graduates. Third, an increase in the income of middle-aged college graduates lowers the homeownership rate of young college graduates due to the credit constraint, i.e. households cannot borrow against their future income, and therefore are forced to delay home purchases.

To evaluate the empirical relevance of the model and to quantify the contribution of these changes on ownership decisions of college and non-college graduates, I examine cross-city variations in house prices, the share of households headed by college graduates, and homeownership rates and household incomes of the four groups using decennial data from the Integrated Public Use Microdata Series (IPUMS) for 1980 to 2010. I regress local house prices and homeownership rates of the four groups on college share, average household income of the four groups, housing supply elasticity, total number of households and year dummies that are supposed to capture any potential aggregate trends.

Consistent with the model's predictions, I find that: a 1 percentage point increase in households headed by holders of a bachelor degree or above pushes up the average house price by 2.1-2.3 percent. For homeownership rates, a 1 percentage point increase in the share of college-educated households leads to a 0.47-0.70 percentage point drop in the homeownership rate for young non-college graduates, a 0.63-0.84 percentage point drop for young college graduates, a 0.34-0.62 percentage point drop for middle-aged non-college graduates and a 0.31-0.40 percentage point drop for middle-aged college graduates. A 1 percent increase in the average household income of college-educated households is associated with a 0.04-0.14 percentage point drop in the homeownership rate among young non-college graduates, and a 0.03-0.17 percentage point drop among middle-aged non-college graduates. Moreover, a 1 percent increase in the average income of middle-aged college graduates is associated with a 0.18-0.22 percentage point drop in the homeownership rate among young college graduates. The estimated coefficients on the college share and the household income of the other three groups become smaller and less significant after controlling for local house prices, indicating that growing college share and changing household income of other groups affect homeownership rates mostly through their impact on local house prices, as the mechanism suggests.

Concerns with the empirical analysis include endogeneity and reverse causality. For instance, higher house prices could induce less-educated households to move to cities with low house prices, resulting in a higher college share. Therefore, I construct an Instrumental Variable (IV) for the college share, which exploits the cross-industry variations in the labor demand growth for college graduates. I construct the predicted college share by interacting the 1970 city-level industry structure with the labor demand growth for college graduates and the non-college graduates in other cities. This Instrument allows me to isolate the impact of increasing college share on local house prices and homeownership rates from alternative explanations. The results are similar to the OLS ones.

To quantitatively evaluate the impact of the mechanism on ownership decisions for college and non-college gradates, I apply the estimated coefficients on the changes in college share and in the average household income of the four groups to project their impact on homeownership rates of the four groups for all years from 1980 to 2018 for the aggregate economy. The model does a good job in fitting the trends in homeownership rates for the four groups from 1980 to 2018, which implies that the proposed mechanism can largely account for the diverging ownership decisions between college and non-college gradates. My findings suggest that the low-income non-college graduates become long-term renters due to the high house prices caused by the changes in the income distribution, which implies that they are facing a more severe affordability problem.

This chapter offers new insights into the discussion about the drop in the young homeownership rate. Unlike Fisher and Gervais (2011) and Anagnostopoulos, Atesagaoglu, and Carceles-Poveda (2013), who argue that the drop in the young homeownership rate is temporary, I find that the drop in the young homeownership rate is more persistent among less-educated households. Fisher and Gervais (2011) argue that young households postpone home purchases due to the delay in marriage and the increase in the income risk. Anagnostopoulos, Atesagaoglu and Carceles-Povedda (2013) argue that skill-biased technological change towards experience lowers the income-to-house price ratio for the young, but increases it for the old. Consequently, it takes young households longer to save for a down payment. Most recently, Mezza, Ringo, Sherlund, and Sommer (2020) highlight the impact of rising student loans on delaying home purchases for college graduates. The empirical analysis in this chapter controls for potential aggregate trends in marriage, income risks, skill-biased technological change, and the rising student debt by introducing year dummies. My analysis suggests that, in spite of these trends, the change in income distribution caused by the increasing share of college graduates and the widening household income gap between college- and non-college-educated households can account for a large fraction of the observed dynamic in ownership rates for both college and non-college graduates. My results suggest that changes in income distribution have pushed up house prices and resulted in a housing affordability issue among low-income non-college-educated households.

The findings in this chapter are consistent with the previous work showing that educational attainment has an increasing impact on the propensity of owning (Gyourko and Linneman, 1996, Gyourko and Linneman, 1997, and Segal and Sullivan, 1998). This chapter provides a mechanism that rationalizes the growing importance of education attainment on housing tenure choice. In the empirical part, this chapter adopts a more Macro approach. I look into the cross-metropolitan variations in homeownership rates, house prices, population share of college graduates and household incomes. I find that homeownership rates for college graduates are less sensitive to their average household income compared to non-college graduates.

The rest of this chapter is organized as follows. Section 2.2 presents the motivating facts. Section 2.3 outlines a simple OLG model that illustrates the mechanism. Section 2.4 describes the empirical exercise. Section 2.5 concludes.

2.2 Empirical Evidence

This section documents the stylized facts that motivate the mechanism. First, I present changes in homeownership rates by age for college- and non-college-educated households separately. Second, I provide evidence on the change in the income distribution due to the increasing share of households headed by college graduates and the widening gap in household income between college and non-college graduates.

2.2.1 Trends in Homeownership Rates: 1976-2015

While the aggregate homeownership rate has been relatively flat since 1976, the homeownership of households with heads aged 25-34 has decreased from 53% to 40% (Figure 2.1). The young homeownership rate recovered slightly during the 2001-2005 mortgage credit expansion, followed by an even sharper decline after 2006.

Figure 2.1: U.S. Homeownership Rates

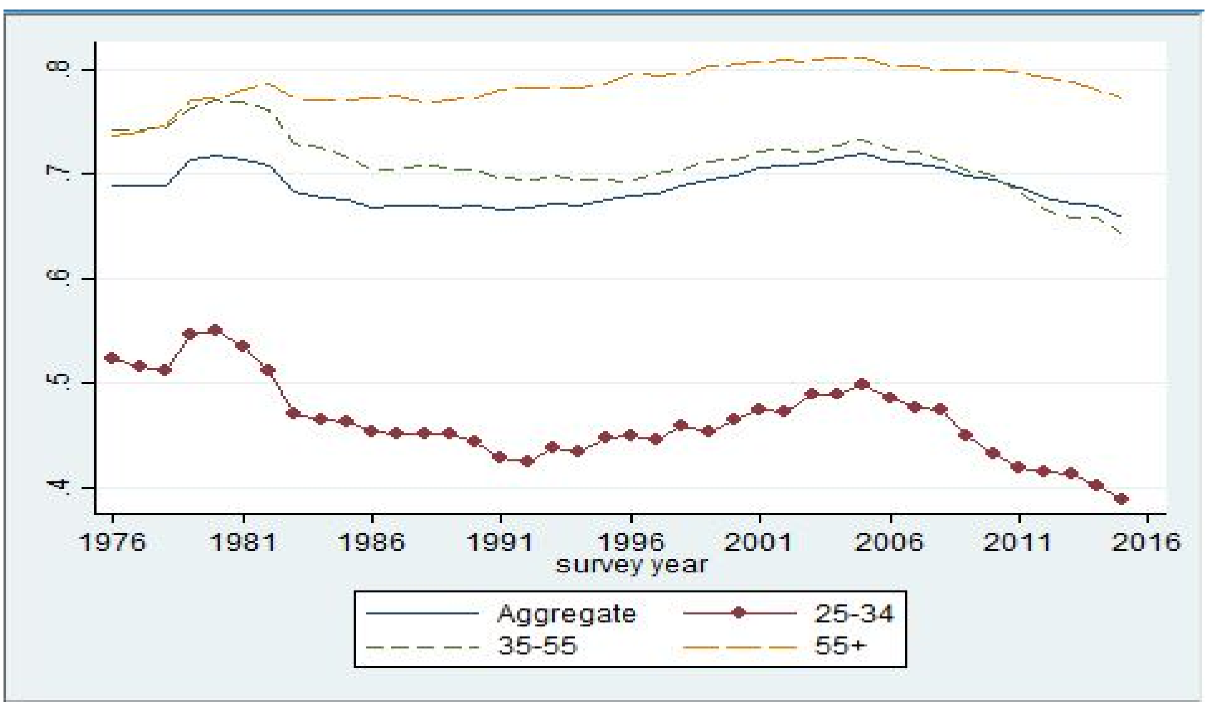


Table 2.1 presents the homeownership rates by age and by education of the household head. Following the Census Bureau and other researchers (e.g. Fisher and Gervais, 2011), homeownership rates are defined as the number of households living in owner-occupied dwellings divided by the total number of households. Households are identified by the age and education

attainment of the household head. College graduates are defined as people who complete four years of college education.

The drop in the young homeownership rate is larger among households headed by non-college graduates. Specifically, households headed by 25-to 29-year-old college graduates see a 7 percentage points decline in their homeownership rate compared to a 15 percentage points drop experienced by households headed by non-college graduates between 1976-2015. The changes in homeownership rates for all age groups from 25-54 are negative. However, there is a noticeable difference in the levels of decline in homeownership rates between college and non-college graduates. For all age groups, the drop in the homeownership rate for non-college graduates is between 1.5 to 3 times as large as the drop for college graduates. Most importantly, 58% (4% out of 7%) of the drop in the homeownership by the young recovers when households reach middle-age (45-54) among college graduates. In contrast, only 33% (5% out of 10%) of the drop in young homeownership recovers when households hit middle-age for non-college graduates. This comparison suggests that while some college graduates postpone home purchases, a considerable fraction of non-college graduates have become long-term renters.

Table 2.1: Change in Homeownership Rates by Age and Education

Age	25-34	35-44	45-54
Households with Non-College-Educated Heads			
1976-1980	0.52	0.70	0.76
2011-2015	0.37	0.53	0.66
Change	0.15	0.17	0.10
Households with College-Educated Heads			
1976-1980	0.55	0.79	0.85
2011-2015	0.48	0.73	0.82
Change	0.07	0.06	0.03

Note: Author's calculation using data from IPUMS-CPS

The cross-sectional analysis above compares homeownership rates of different cohorts. To check the robustness of my main finding, i.e. diverging homeownership decisions between college- and non-college-educated households, I conduct a cross-cohort comparison. Figure 2.2 plots the age profiles of homeownership by the education of the household head calculated using the CPS data extracted from the Integrated Public Use Microdata Series (IPUMS). We see similar patterns. Among the less-educated households, homeownership profiles of newly born cohorts are flatter compared to the older generations. For households headed by college graduates, newly born cohorts catch up with the older generations in terms of ownership around age 40. The sharp contrast in the ownership profiles confirms the diverging ownership decisions between college and non-college graduates, i.e. while a large fraction of non-college-educated households become long-term renters, some college-educated households merely postpone the transition to ownership.

2.2.2 Trends in Income Distribution

This subsection documents the driving forces behind the changing income distribution. The rise in the college premium and the increase in the number of households headed by college graduates has changed the income distribution substantially since 1980 (see e.g. Goldin and Katz, 2001). As Gyourko and Linneman (1996) and Gyourko and Linneman (1997) suggest, the propensity towards owning increases with household income. Therefore, the change in the income distribution can lead to a reallocation of owner-occupied houses among education (income) and age groups through its impact on house prices.

The fraction of households headed by a bachelor or above has more than doubled, climbing from less than 15% to 33% from 1976 to 2015.[1] Meanwhile. the gap in real average household income between college and non-college graduates has risen significantly (see Figure 2.3). Young households (25-34) and middle-aged households (35-54) constitute prime age buyers in housing markets. Since the household income of non-college-educated households has barely changed, the widening income gap between those with and without college education is mainly driven by the rising household income of college graduates. It is worth noting that middle-aged households headed by college graduates have experienced a larger increase in their household income compared to the young college-educated households. In the following analysis, I show that this steeper earning profile is key to understanding the postponement of house purchases by college graduates.

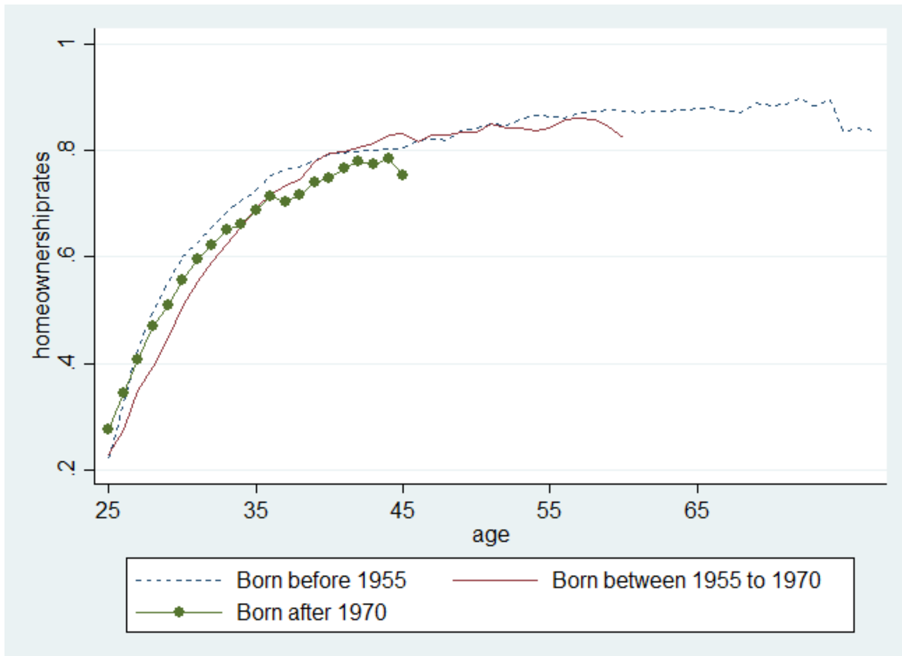
2.3 Model

This section presents a model that I use to qualitatively illustrate the impact of the change in the income distribution documented in Section 2.2.2: (i) an increase in the share of college graduates, (ii) a moderate rise in the household income of young college graduates and (iii) a large increase in the household income of middle-aged college graduates; on house prices and homeownership rates of the four groups.

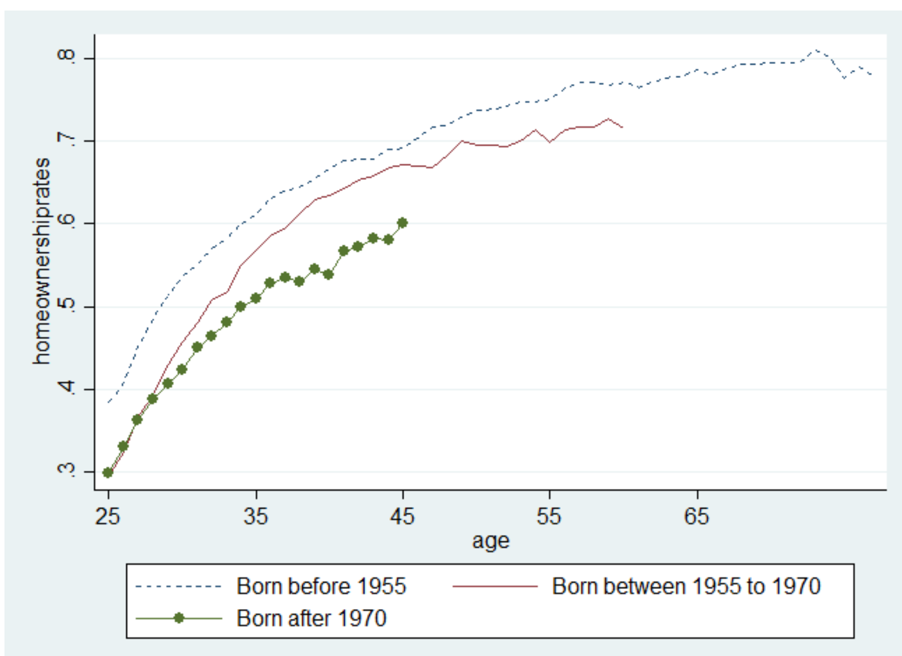
The model generates several testable implications: (i) an increase in the share of college graduates pushes up the equilibrium house prices and lowers homeownership rates for all groups; (ii) a rise in the household income of college graduates lowers homeownership rates

[1] Author's calculation using IPUMS.

Figure 2.2: Homeownership Profile

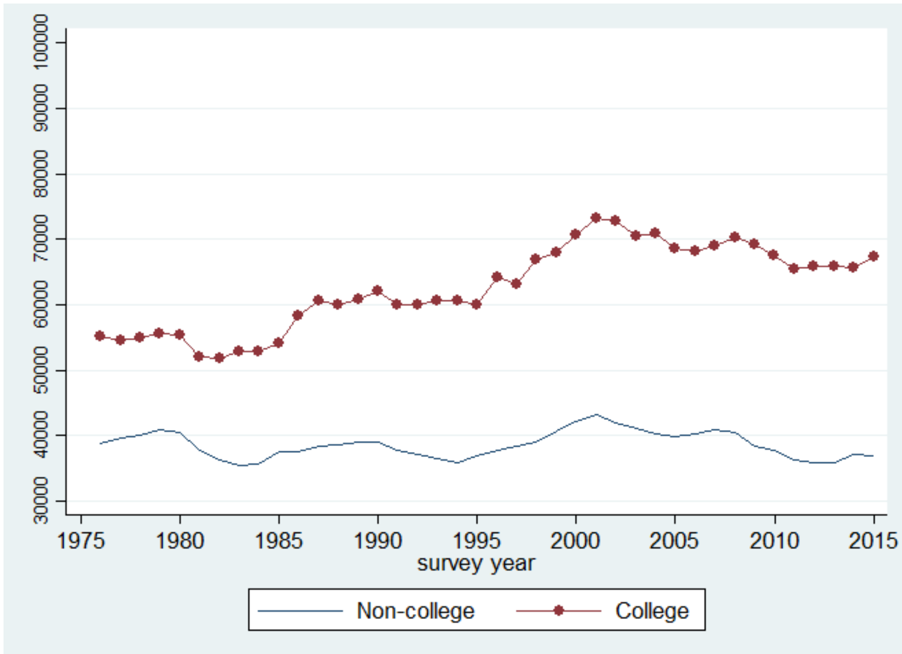


(a) College

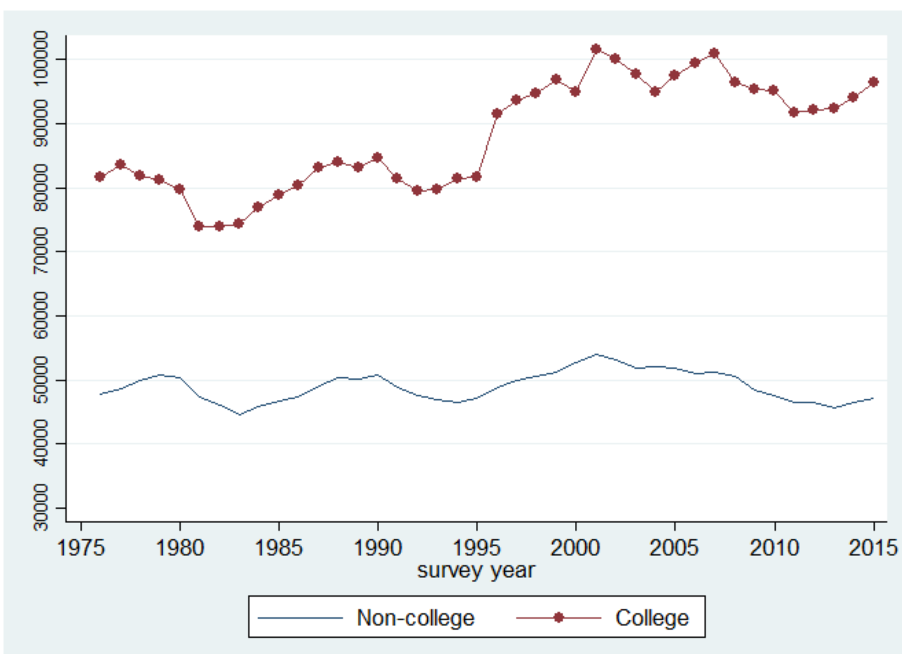


(b) Non-College

Figure 2.3: Evolution of Household Income



(a) Household Income (25-34)



(b) Household Income (35-54)

for non-college graduates; and (iii) an increase in the household income of middle-aged households lowers homeownership rates for the young among college graduates. These implications guide my empirical analysis in Section 2.4.

The three-period OLG model extends the framework in Ortalo-Magne and Rady (2006). The model is modified to consider two types of households, college and non-college, to capture the diverging ownership decisions and changes in household income between these two groups. In addition, it allows me to consider the impact of an increase in the fraction of households headed by college graduates. Households within each type differ in the utility premium they derive from living in an owner-occupied house and in their endowment streams, which are characterized by their ability ranking. Conditional on the ability ranking, college graduates earn more than non-college graduates, which is referred to as the "college premium". Owner-occupied houses are in limited supply.

2.3.1 Population

A measure one of agents is born at the start of each period. A fraction κ are college graduates (C), and the remaining $1-\kappa$ are non-college graduates (N). Each agent lives for 3 periods, so the total population in each period is 3.

Within each type (College or Non-college) of each cohort, agents are uniformly distributed over the unit square. Each agent of type $g \in \{N, C\}$ is identified by the indices $(i, m) \in [0,1] \times [0,1]$ that determine the ability and preference towards owning, respectively. Agents learn their types and indices at the beginning of life. I assume that i is independent of m, such that households of all abilities from all educational backgrounds draw their preference towards owning from the same distribution.[2] College and non-college graduates differ in their endowment stream conditional on their ability ranking.

2.3.2 Commodities

There is a numeraire consumption good and S units of identical owner-occupied houses. Each house can accommodate one household only, who must own it. There are $3-S$ units of identical rental units. Housing choices $h \in \{\emptyset, H\}$, where $\emptyset$ stands for renting, of which cost and utility are normalized to 0.

[2]Relaxing this independence assumption to allow for a positive correlation between ability and preference towards owning will not change my results qualitatively.

2.3.3 Endowment

Agents are born with no initial wealth. At age $j = 1$ and 2, agents with (i, m) of type g receive an endowment of $w_j^g(i)$ units of the numeraire goods where the mapping from ability ranking to endowment, $w_j^g : [0, 1] \rightarrow R_+$, is continuous and monotonically increasing.

2.3.4 Preference

Following Ortalo-Magne and Randy (2006), I assume a linear utility function.

$$\sum_1^3 c_t + U(h_2, m) + U(h_3, m) \tag{2.1}$$

$$U(h, m) = \begin{cases} 0, & if \;\; h = \emptyset \\ m\,\triangle & if \;\; h = H \end{cases} \tag{2.2}$$

$$\begin{aligned} s.t. c_1 + s_1 + \mathbb{1}_{h_2=H} P^* &\leq w_1^g(i) \\ c_2 + s_2 + \mathbb{1}_{h_2=\emptyset, h3=H} P^* &\leq w_2^g(i) + s_1 * r \\ c_3 &\leq s_2 * r + \mathbb{1}_{h3=H} P^* \\ c_t &\geq 0, \quad t \in \{1, 2, 3\} \end{aligned} \tag{2.3}$$

c_t represents the consumption of the numeraire goods and $m\triangle$ is the additional utility derived from living in owner-occupied houses, i.e. the ownership premium. Consumption in each period has to be non-negative.

The linear utility implies that all non-housing consumption will be postponed until the last period of life. This feature keeps the model analytically tractable, particularly with respect to the equilibrium house price and homeownership rates of different groups.

The supply of owner-occupied houses is fixed at S, so that the aggregate homeownership rate is fixed at $S/3$.[3] Households have access to a storage technology for the numeraire good that yields an exogenously given rate of interest $r > 1$.

In each period, there is a competitive market for houses with the equilibrium price P^*. There is no rental market for dwellings and no other asset markets. Households cannot borrow against their future income.

[3]Fixed supply of owner-occupied houses is not critical to my results. As long as the supply is not perfectly elastic, the results hold. It is worth noting the aggregate homeownership rate in the U.S. has been stable at 69%, which supports the fixed supply assumption.

2.3.5 Timing

Within each period, agents first derive utility from housing. Then they receive an endowment of the numeraire good, after which, they trade in the housing market and finally, they consume the numeraire good.

Households with ability i of type g receive $w_1^g(i)$ at age 1 and $w_2^g(i)$ at age 2. At age 3, households have no labor income and consume their savings. Owners at age 3 sell their house and consume. Denote $W^g(i) = rw_1^g(i) + w_2^g(i)$ as household's lifetime income valued at age 2.

2.3.6 Assumptions on Endowment Streams

- Fix the ability rank i, college graduates receive more endowments than non-college graduates $w_j^C(i) \geq w_j^N(i)$ for $j = 1, 2$.

- Households experience wage growth as they age. $w_2^g(i) \geq w_1^g(i)$. Consistent with the observation that college graduates have steeper earning profiles than non-college graduates, I assume $w_2^C(i) - w_1^C(i) \geq w_2^N(i) - w_1^N(i)$.

- Following Ortalo-Magne and Rady (2006), I adopt the following convention. Given a continuous and strictly increasing function $w : [0,1] \rightarrow R_+$, set $w^{-1}(i) = 1 \;\; if \;\; i > w(1), \;\; and \;\; w^{-1}(i) = 0 \;\; if \;\; i < w(0)$.

2.3.7 Equilibrium

The independence between endowment stream and ownership premium implies that a household's preference towards owning determines whether he/she wants to buy, while the household income determines whether he/she can afford to purchase a house.

As households postpone consumption until the last period, the value of buying a house at the end of the first period and holding it until the last period is the total lifetime income valued at the end of life plus the ownership premium for two periods net of the forgone interest on the equilibrium house price for two periods.

$$V^1(i,m,g) = \begin{cases} W^g(i)r + 2m \triangle -(r^2-1)P^*, \;\; if \;\; w_1^g(i) \geq P^* \\ -\infty \;\; otherwise \end{cases} \tag{2.4}$$

Similarly, the value of buying a house at the second period is the total life time income valued at the end of life plus the ownership premium for one period net of the forgone interest on the equilibrium price for the second period.

$$V^2(i,m,g) = \begin{cases} W^g(i)r + m\triangle -(r-1)P^*, & if \;\; w^g(i) \geq P^* \\ -\infty \;\; otherwise \end{cases} \tag{2.5}$$

The value of permanent renters is simply the lifetime income valued at the end of life.

$$V^0(i,m,g) = W^g(i)r \tag{2.6}$$

Households that prefer buying in the first period satisfy $V^1(i,m,g) \geq V^2(i,m,g)$ and $V^1(i,m,g) \geq V^0(i,m,g)$. It implies that at the end of period 1, households with $i \geq w_1^{g^{-1}}(P^*)$ and $m \geq m_1^*$ become owners, where $m_1^* = \frac{r(r-1)P^*}{\triangle}$.

Households that postpone buying a house to the end of period 2 are characterized as $i \geq W^{g^{-1}}(P^*)$ and $m \geq m_2^*$, where $m_2^* = \frac{(r-1)P^*}{\triangle}$.

The lower income cutoff among the second-period buyers $w_1^{g^{-1}}(P^*) > W^{g^{-1}}(P^*)$ implies that some of the households postpone buying due to credit constraints. Households with $m \geq m_1^*$ always prefer to buying a house in the first period. However, households without enough endowment have to save for one more period.

Households that prefer buying a house at the end of the first period also find it optimal to hold it in the second period. As there is no uncertainty on house prices or income, owner-occupied houses are modeled as a consumption good. Thus, the cost of owning per period is simply the forgone interest on the house price. Households that buy a house are those with a high attachment to owning and therefore would prefer to hold it until the last period of life.

The model has two types of households, college and non-college, and three ages for each type. In total, I have 6 groups. Because households buy house at the end of age 1 and age 2, homeownership rates for the four groups are calculated: young college, young non-college, middle-age college, and middle age non-college.

Lemma 2.3.1 *There is a unique steady-state equilibrium. The price of houses P^* solves*

$$\begin{aligned} S = &(1-\kappa)(1-m_1^*)(1-w_1^{N^{-1}}(P^*)) + (1-\kappa)(1-m_2^*)(1-W^{N^{-1}}(P^*)) \\ &\kappa(1-m_1^*)(1-w_1^{C^{-1}}(P^*)) + \kappa(1-m_2^*)(1-W^{C^{-1}}(P^*)) \end{aligned} \tag{2.7}$$

Given the uniform distribution of ability and preference towards owning, steady state home-ownership rates for young non-college, young college, middle-age non-college, and middle-age *college are* $\{n_1^N, n_1^C, n_2^N, n_2^C\}$

$$n_1^g = (1 - m_1^*)(1 - w_1^{g^{-1}}(P^*)), \quad g \in \{N, C\} \tag{2.8}$$

$$n_2^g = (1 - m_2^*)(1 - W^{g^{-1}}(P^*)), \quad g \in \{N, C\} \tag{2.9}$$

The uniqueness of the equilibrium can be proven by showing the right-hand side of equation 2.7 is strictly decreasing in P^* as both m^* and $w^{g^{-1}}$ are increasing in P^*.

Proposition 2.3.2 *Holding the income of each group constant, an increase in the college share (rising* κ*) pulls up the equilibrium housing price and therefore reduces homeownership rates for all groups* $\frac{\partial P^*}{\partial k} > 0$, $\frac{\partial n_i^g}{\partial k} < 0$ *for* $i = \{1, 2\}$ *and* $g = \{N, C\}$.

Rising college share drives up the aggregate demand for owner-occupied houses for a given level of house price, because college graduates with high income are more likely to be able to afford a house. As long as the housing supply is not perfectly elastic, growing aggregate demand reflects itself through the equilibrium housing price.

Proposition 2.3.3 *Holding the share of college graduates fixed, a rise in the income of college (non-college) graduates leads to an increase in the equilibrium housing price, an increase in the homeownership rate of college (non-college) graduates and a decrease in the homeownership rate for both young and middle-aged non-college (college) graduates.*

An increase in the endowment of middle-age college-(non-college-)educated households lowers the homeownership rate for the young among college-(non-college-)graduates.

A rise in the endowment of young college- (non-college-) educated households increases the homeownership rate of themselves. The impact on the homeownership rate for middle-aged college-(non-college-)educated households is uncertain.

A rise in endowment of one type, $g \in \{N, C\}$, leads to an increase in the housing demand of that type, as a larger fraction of households from that type can afford to buy a house. Holding the housing supply unchanged, the housing price has to adjust in the presence of excess demand. As a result, the equilibrium house price increases to clear the housing market. The homeownership rate drops for the other type whose income does not change. As the aggregate homeownership rate is fixed at $S/3$, the homeownership rate for the type that experiences an income growth increases at the expense of a drop in homeownership of the other type.

Within each type, a rise in the endowment of middle-age households pushes up the equilibrium house price. Both m_1^* and $w_1^{g^{-1}}$ increase. Because young households cannot borrow against future income, an increasing fraction of them choose to postpone house purchases.

Within one type, an increase in the endowment of young households leads to an increase in the equilibrium house price. As a result, m_2^* increases, indicating that less middle-aged households find owning attractive. Meanwhile, the life time income $W^g(i)$ increases, suggesting that owning is more affordable. The direction of overall impact depends on these two forces.

The qualitative impact of increasing college share, rising household income of young college households (YC), and an even bigger rise in household income of middle-aged college households (MC) on homeownership rates of the four groups is summarized in Table 2.2.[4] A combination of the three changes predicts a drop in the homeownership rate for both young and middle-aged households among non-college graduates, which indicates that some non-college graduates choose to become long-term renters. Its impact on college graduates is undetermined. When the increase in the household income of middle-aged households is large enough, we should see some college graduates postponing the purchase of their first home.

Table 2.2: Predicted Impact on Homeownership Rates

Homeownership Rate	↑ College Share	↑ HH Income of YC	↑↑ HH Income of MC	Overall
Young College	↓	↑	↓	?
Young Non-College	↓	↓	↓	↓
Middle Age College	↓	?	↑↑	?
Middle Age Non-college	↓	↓	↓	↓

In the next section, I apply the model to data to quantify the impact of the change in the income distribution on house prices and homeownership rates of the four groups.

2.4 Empirical Analysis

I use cross-city variation for the largest 161 cities in the U.S. to verify Proposition 2.3.2 and Proposition 2.3.3 and to quantify the impact of the change in income distribution on the ownership decision for college and non-college graduates.[5] I regress house prices and homeownership rates of the four groups on the share of households headed by college graduates, average household income of the four groups, local number of households, housing supply elasticity and year dummies. To control for the endogeneity issue and reverse causality, I adopt an instrumental variable that predicts share of college graduates using the city-level industry structure in the year 1970, aggregate labor demand growth, and labor demand growth for college graduates in other cities from 1980 to 2010.

[4]The comparative statics mainly focuses on the change in income of college graduates as that is what we see in the data. The propositions apply for non-college graduates in the same way.

[5]Appendix A.1 provides a full list of the cities in OLS and IV estimation.

I find that, consistent with the propositions, cities with more college graduates tend to have higher house prices and lower homeownership rates for the four groups. The homeownership rate of one type (College or Non-college) is increasing in its own household income and decreasing in the household income of the other type. Moreover, an increase in the household income of middle-age households lowers the homeownership rate for the young among college graduates.

I apply the estimates to quantify the impact of the changes in college share and household income of the four groups on aggregate homeownership rates for the four groups. I find that these changes can largely account for the diverging ownership decisions: delayed home purchasing of college graduates and the switch towards renting by non-college graduates.

2.4.1 Data Description

I use data from the 1970, 1980, and 1990 waves of the Census and the 2000 and 2010 waves of the American Community Survey (ACS), taken from IPUMS-USA, aggregated to the metropolitan area level.[6] I obtain information on the mean/median house values of owner-occupied houses, household income, age and education of household heads, and the geographic location of residence. All dollar values are converted into constant 1999 dollars using the Consumer Price Index (CPI). For the analysis across metropolitan areas, I use a sample of the largest 161 cities that have at least 30 observations in each of the four groups for every year for the OLS estimation and 108 cities for the IV estimation.[7]

The average self-reported house value in a city is used as the representation of local house prices in the main analysis. To check the robustness of the results, two other measurements of housing prices are used: the median local house price that is self-reported and the Freddie Mac Conventional Mortgage Home Price Index that is based on repeated sales.

2.4.2 Summary Statistics

Table 2.3 presents the summary statistics of the 161 metropolitan areas for 1980 and 2010. During this period, the (unweighted) average share of college-educated households increased from 20.5% to 31%, with the standard deviation increasing from 0.058 to 0.074. The average house price went up, as did the cross-city variation in mean house prices. Similar to the aggregate economy, we can see a larger and more persistent drop in homeownership rates among

[6]Census 1970 is used to construct the Instrumental Variable.

[7]To calculate the homeownership rate of one group, I need enough observations for each group. So I exclude cities with less than 30 observations in a group for at least one year. I conduct robustness using metropolitan areas with more than 40 (20) households in each group. Despite ending up with fewer (more) areas, I obtain similar results.

the non-college graduates compared to college graduates. While non-college graduates experienced a decline in their average household income, college graduates enjoyed an increase in their average household income, especially among the middle-aged ones.

Figure 2.4 plots the unconditional correlation between average house prices and the share of households headed by college graduates across metropolitan areas in 2010. A 1% point increase in a city's college share in 2010 is associated with a 3.05% increase in the average city house price.

Figure 2.5 displays the cross-city comparison in college share and homeownership rates in 2010. Panel (a) shows that a 1 percentage point increase in a city's college share is associated with a 0.39 percentage point drop in the young homeownership rate among non-college-educated households and a 0.68 percentage point drop in the young homeownership rate among college-educated households, indicating housing tenure choice of young college graduates might be more sensitive to the college share that shifts the local house prices. One potential explanation is that college graduates with steeper earning profiles choose to smooth consumption by postponing home purchases due to credit constraints. Figure 2.5(b) confirms this conjecture. A 1 percentage point increase in a city's college share is associated with a 0.27 percentage point drop in middle-aged homeownership rate among the non-college-educated households and a 0.16 percentage point drop in the middle-aged homeownership rate among college-educated households. It suggests that in the presence of higher housing prices due to higher college share, more than 76 percent of the drop in the young homeownership rate is recovered as the household head ages for college graduates, compared to the 30 percent recovery among non-college graduates. These unconditional correlations support the two propositions. In the next subsection, I proceed to check the impact of changes in both college share and income of the four groups on the ownership decisions for college and non-college graduates.

2.4.3 Verification of Propositions

I estimate the following regressions using the panel of 161 metropolitan areas over four decades from 1980 to 2010.

$$log(P_{j,t}) = \beta_0 + \beta_1 \kappa_{j,t} + \sum_g \beta_2^g log(I_{j,t}^g) + \beta_3 E_j + \beta_4 log(N_{j,t}) + \delta_t + \epsilon_{j,t} \tag{2.10}$$

$$OR_{j,t}^g = \alpha_0^g + \alpha_1^g \kappa_{j,t} + \alpha_2^g log(I_{j,t}^g) + \sum_{g' \neq g} \alpha_{3,g'}^g log(I_{j,t}^{g'}) + \alpha_4 E_j + \alpha_5 log(N_{j,t}) + \delta_t + \epsilon_{j,t} \tag{2.11}$$

Table 2.3: Summary Statistics

		1980		2010	
Variable	Obs	Mean	Std. Dev	Mean	Std. Dev
Fraction of college-educated households	161	0.205	0.058	0.314	0.0739
Homeownership rate YNC	161	0.510	0.081	0.360	0.090
Homeownership rate YC	161	0.584	0.082	0.542	0.118
Homeownership rate MNC	161	0.747	0.061	0.627	0.073
Homeownership rate MC	161	0.842	0.042	0.810	0.061
Ln Average HH income YNC	161	10.56	0.093	10.30	0.137
Ln Average HH income YC	161	10.82	0.098	10.85	0.158
Ln Average HH income MNC	161	10.86	0.10	10.67	0.117
Ln Average HH income MC	161	11.27	0.071	11.34	0.129
Ln Total Number of HHs	161	11.92	0.99	12.46	0.96
Ln Average House price	161	11.76	0.253	12.03	0.370
Saiz's supply elasticity	161	2.13	1.099		

Figure 2.4: College Share and Average Local House Price

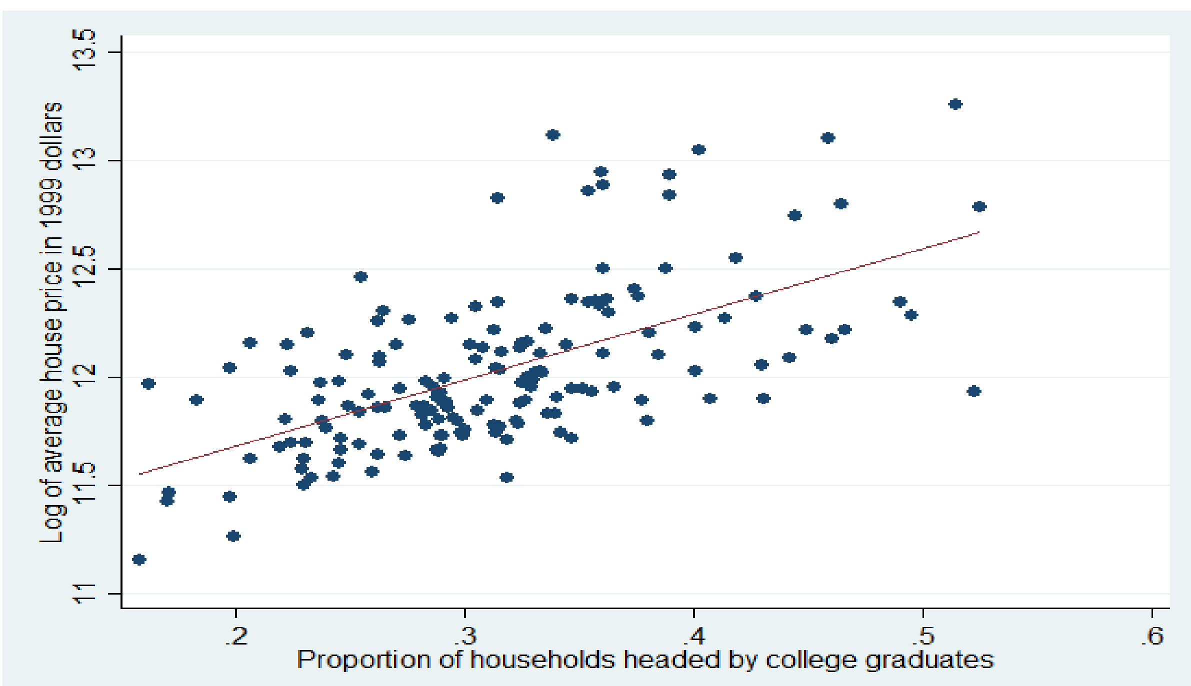
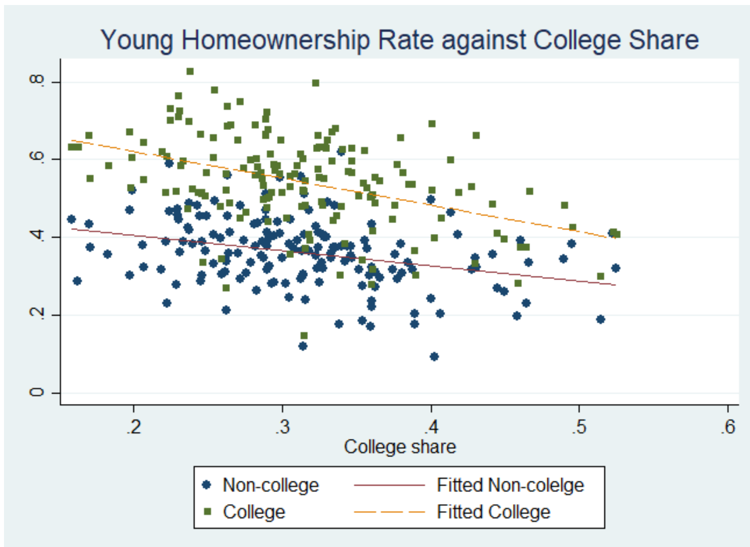
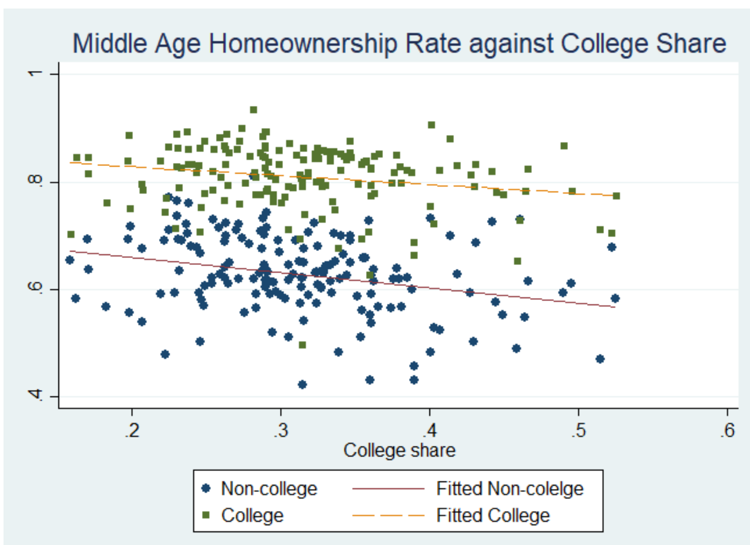


Figure 2.5: College Share and Homeownership Rates



(a) Homeownership Rate (25-34)



(b) Homeownership Rate (35-54)

Equation 2.10 and 2.11 detail the primary specification used in testing Proposition 2.3.2 and Proposition 2.3.3 from the model. Equation 2.10 focuses on the equilibrium house price, and Equation 2.11 is for homeownership rates of the four groups. $P_{j,t}$ is the real average house price in metropolitan j at time t. $OR^{g}_{j,t}$ stands for the homeownership rate of group g in metropolitan j at time t. $\kappa_{j,t}$ is the share of households headed by someone with a college degree or above in metropolitan j at time t. $I^{g}_{j,t}$ is the real average household income of group g in metropolitan j at year t. I examine four groups defined by age and education of the household head: young college-educated household (YC), young non-college-educated household (YNC), middle-aged college-educated household (MC), and middle-aged non-college-educated household (MNC). $N_{j,t}$ is the total number of households in metropolitan area j at time t. δ_t is a year dummy that is supposed to capture any potential aggregate shocks, such as a drop in the marriage rate, an increase in the income volatility, an expansion of mortgage credit, a change in consumer confidence, and/or a trend. E_j is local housing supply elasticity provided by Saiz (2010), which also contains information on amenity and regulations.

OLS

Table 2.4 reports the OLS results. Consistent with Proposition 2.3.2, cities with a higher share of households headed by college graduates tend to have higher house prices and lower homeownership rates for all groups. The homeownership rate of one type is generally positively correlated with the average income of its own group and negatively correlated with the average income of the other type, as Proposition 2.3.3 suggests. Moreover, an increase in the household income of middle-age households lowers the homeownership rate for the young within the same type. In addition, cities with high housing supply elasticity tend to have lower house prices and higher homeownership rates.

The mechanism suggests that holding the household income fixed for non-college graduates, rising college share and increasing household income of college graduates affect homeownership rate of non-college graduates through the general equilibrium effect, i.e. their impact on the equilibrium house prices. To test the importance of the general equilibrium effect, I introduce the city-level average house price into the homeownership rate regressions in columns (6)-(9). The estimates on the share of college-educated households become smaller in magnitude and less significant, as do the estimates on the average household income of other groups. In other words, the rising share of college graduates and the increasing household income of college graduates affect homeownership rates of non-college graduates through their impact on local house prices, as the model suggests.

Some of the year fixed effects are significant, suggesting the existence of time trends. The estimated coefficients on year dummies differ across groups, indicating that the trends vary across groups. For the two time-related factors that are well-discussed in the literature, declining marriage rates and mortgage credit expansion, previous studies suggest that changes in these two factors may differ across groups. For instance, Goldstein and Kenney (2001) argue that marriage is increasingly becoming a province of the most educated women. For the impact of changes in credit conditions, Sufi and Mian (2009) suggest that the mortgage credit expansion is concentrated in subprime ZIP codes with sharply declining relative income growth. Although they do not divide households by the education of the household head, research on the skill-biased technological change (see, Katz et al. 1999, for example) indicates that households with negative income growth are likely to be less educated.

Declining marriage rates and credit expansions have opposite impacts on homeownership rates. Declining marriage rates lower homeownership rates (Fisher and Gervais, 2010).[8] Meanwhile, relaxing credit constraints boosts homeownership. The negative impact dominates through 1980 to 2000. From 2000 to 2010, the positive impact takes over as we see a significant positive coefficient on the 2010 year dummy on the homeownership rate of young college-educated households.

IV Results

One possible concern about the OLS regression is the omitted-variable bias or reverse causality. For instance, higher house prices could induce less educated households to move to other cities, resulting in a higher college share (Gyourko, Mayer and Sinai, 2013). In this case, the OLS estimator could be downward biased in the price equation estimation.

I control for these possibilities using an instrumental variable (IV), in which I use the industry structure in the year 1970, total labor demand growth and the labor demand growth for college graduates in other cities to project the college share in one city for 1980-2010. The instrumental variable exploits the cross-industry variations in the labor demand growth for college graduates. It requires that the industry structure in 1970 is independent of the housing market conditions in the following years.

The instrumental variable is constructed in two stages. In the first stage, I use Equation 2.12 to construct the predicted local labor demand for college graduates $Z_{j,t}$

$$Z_{jt} = \sum_{h}^{41} n_{h,j,1970} \times (n_{h,-j,t}/n_{h,-j,1970}) \tag{2.12}$$

[8]The cross-sectional variation in marriage rates is not big enough to test the impact of changing marriage rates on homeownership rates. Meanwhile, the change in marriage rates over time can be largely captured by the time dummies. When I regress marriage rates on time dummies, I find more than 70% of variation in the marriage rate can be explained.

Table 2.4: OLS Results

	(1) $log(P)$	(2) OR^{YNC}	(3) OR^{YC}	(4) OR^{MNC}	(5) OR^{MC}	(6) OR^{YNC}	(7) OR^{YC}	(8) OR^{MNC}	(9) OR^{MC}
κ	2.053***	-0.474***	-0.632***	-0.335***	-0.311***	-0.00809	-0.353***	0.0661	-0.0733*
	(0.142)	(0.0515)	(0.0608)	(0.0429)	(0.0345)	(0.0510)	(0.0690)	(0.0427)	(0.0396)
$log(I^{YNC})$	0.434***	0.222***	0.0352	-0.0304	-0.0118	0.321***	0.0943*	0.0544	0.0384
	(0.141)	(0.0505)	(0.0528)	(0.0387)	(0.0277)	(0.0438)	(0.0544)	(0.0366)	(0.0316)
$log(I^{YC})$	0.355***	-0.0402	0.330***	-0.0494	-0.0343	0.0403	0.378***	0.0199	0.00679
	(0.0792)	(0.0324)	(0.0408)	(0.0300)	(0.0224)	(0.0268)	(0.0388)	(0.0251)	(0.0206)
$log(I^{MNC})$	0.585***	0.0495	0.126**	0.276***	0.133***	0.182***	0.205***	0.390***	0.201***
	(0.135)	(0.0576)	(0.0621)	(0.0454)	(0.0351)	(0.0512)	(0.0617)	(0.0401)	(0.0342)
$log(I^{MC})$	0.0196	-0.117**	-0.222***	-0.134***	0.0714**	-0.112***	-0.219***	-0.131***	0.0737***
	(0.117)	(0.0502)	(0.0552)	(0.0407)	(0.0311)	(0.0413)	(0.0501)	(0.0326)	(0.0279)
$log(P)$						-0.227***	-0.136***	-0.195***	-0.116***
						(0.0128)	(0.0207)	(0.0115)	(0.0127)
$log(N)$	-0.0144	-0.00824**	-0.00954**	-0.00952**	-0.00598**	-0.0115***	-0.0115***	-0.0123***	-0.00764***
	(0.0120)	(0.00417)	(0.00459)	(0.00371)	(0.00303)	(0.00310)	(0.00404)	(0.00287)	(0.00250)
Elasticity	-0.115***	0.0223***	0.0217***	0.0174***	0.0104***	-0.00382	0.00599*	-0.00505**	-0.00291
	(0.00897)	(0.00304)	(0.00335)	(0.00235)	(0.00210)	(0.00292)	(0.00343)	(0.00225)	(0.00194)
1990.year	-0.0908***	-0.0388***	-0.0801***	-0.0391***	-0.0384***	-0.0594***	-0.0925***	-0.0568***	-0.0489***
	(0.0238)	(0.00928)	(0.00860)	(0.00760)	(0.00519)	(0.00743)	(0.00840)	(0.00607)	(0.00469)
2000.year	-0.0682***	-0.00175	-0.0238**	-0.0120	-0.0119*	-0.0172*	-0.0331***	-0.0253***	-0.0198***
	(0.0249)	(0.0109)	(0.0118)	(0.00862)	(0.00618)	(0.00912)	(0.0112)	(0.00725)	(0.00567)
2010.year	0.251***	-0.0157	0.0687***	-0.0231*	0.0211**	0.0412***	0.103***	0.0259**	0.0501***
	(0.0374)	(0.0163)	(0.0175)	(0.0126)	(0.00932)	(0.0146)	(0.0173)	(0.0120)	(0.00960)
Constant	-3.236***	-0.476	-2.024***	0.267	-0.799***	-1.211***	-2.465***	-0.365	-1.174***
	(1.065)	(0.395)	(0.424)	(0.295)	(0.238)	(0.312)	(0.393)	(0.239)	(0.215)
Observations	644	644	644	644	644	644	644	644	644
R-squared	0.734	0.491	0.487	0.476	0.295	0.669	0.544	0.672	0.452

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Where $n_{h,j,t}$ is the number of college-educated workers in industry h, city j, year t, $n_{h,-j,t}$ is the number of college workers in industry h and year t, excluding city j. The first stage uses the change in the number of college workers in other cities adjusted by local industrial college employment in the base year to predict the local labor demand for college graduates in other years.

Predicted total local labor demand $L_{j,t}$ is constructed in a similar way in the second stage.

$$L_{jt} = \sum_{h}^{41} l_{h,j,1970} \times (l_{h,-j,t}/l_{h,-j,1970}) \tag{2.13}$$

Where $l_{h,j,t}$ is the number of workers, including both college and non-college graduates, in industry h, city j, year t, $l_{h,-j,t}$ is the number of workers in industry h and year t, excluding city j.

The Bartik Instrument is defined as the predicted college share, i.e. the ratio between predicted local demand for college graduates and predicted aggregate local labor demand.

$$B_{jt} = \frac{Z_{jt}}{L_{jt}} \tag{2.14}$$

I use two-stage estimation. In the first stage, I regress the share of college-educated households on the instrument, the average household income of all groups, housing supply elasticity, the number of households, and year dummies. The first-stage regression result is reported in the last column of Table 2.5. In the second stage, I regress the variables of interest, i.e., local house prices and homeownership rates of the four groups, on the predicted fraction of college-educated households, average household income of the four groups, housing supply elasticity, the number of households, and year dummies.

As the construction of the Instrument variable requires the industry structure in 1970. I use the 1-in-100 national random sample of the population of which the smallest identifiable geographic units are metropolitan areas. In order to construct the employment of different industries for both college and non-college graduates, cities without enough observations in all of the 41 industries for both education groups are dropped. Therefore, I end up with a smaller sample size in the IV estimation. To make the OLS results and IV results more comparable, I use the sample of 107 cities to run the OLS regressions. Appendix A.2 reports the results. The estimates are quantitatively similar to the ones presented in Table 2.4.

The IV results are reported in Table 2.5. The first-stage estimation result is reported in the last column. The F statistic is 713.485, rejecting the weak instrument null hypothesis. For the price equation, the IV estimates are similar to the OLS estimates. A 1 percentage point increase in the share of college-educated households leads to a 2.3% increase in average local house prices. In terms of the ownership estimations, IV estimators are larger in magnitude than the OLS estimators, especially for college-educated households. A 1 percentage point increase in the share of college-educated households leads to a 0.70 percentage point drop in homeownership rate among young non-college-educated households, a 0.84 percentage point drop in homeownership rate for young college-educated households, a 0.62 percentage point drop in homeownership rate for middle-aged non-college-educated households, and a 0.41 percentage point drop in homeownership rate among middle-aged college-educated households.

In the ownership rate regressions, estimated coefficients on the average income of the other type are negative and the estimated coefficient on the household income of its own group is positive, which is consistent with Proposition 2.3.3. For instance, a 1 percent increase in the household income of young college graduates lowers the ownership of young non-college graduates by 0.145 percentage point and middle-aged non-college graduates by 0.177 percentage point. In addition, a 1 percent increase in the household income of middle-aged college graduates lowers the ownership of young college graduates by 0.177 percentage point.

A comparison of the coefficients on income of college and non-college households reveals that in general, homeownership rates of college graduates are less sensitive to the change in their current income, which suggests the impact of household income on ownership rate might not be linear, most likely due to the down payment requirement.

Note that in the price equation estimations, the IV estimate on college share is larger than the OLS estimate, which suggests that the omitted variable is negatively correlated with the college share. One possible explanation is that higher house prices could induce less educated households to move to other cities, resulting in a higher college share (Gyourko, Mayer and Sinai, 2013).

2.4.4 Robustness

In addition to the average house prices, two measurements of housing prices are commonly used: the median house price reported by owners and the Freddie Mac Conventional Mortgage Home Price Index (CMHPI).[9] Compared to the mean house price, the median house price is less likely to be affected by extreme values but it may overlook the increasing demand in the high quality market. According to Goodman and Ittner (1992), the self-reported house value is subject to measurement errors. Therefore, I also run the regressions using CMHPI based on repeated sales to approximate house prices. It controls for quality by holding constant property type and location, but it may overlook the price of newly built houses. To check the robustness of the results and to investigate the impact of changing income distribution across markets with

[9]CMHPI is combined with the median single-family home values from the 2000 Census to make the cross metropolitan areas comparison possible.

Table 2.5: IV Results

	$log(P)$	OR^{YNC}	OR^{YC}	OR^{MNC}	OR^{MC}	κ
κ	2.312***	-0.700***	-0.843***	-0.621***	-0.405***	
	(0.237)	(0.0907)	(0.0966)	(0.0784)	(0.0598)	
Bartik						0.652***
						(0.0255)
I^{YNC}	0.442***	0.277***	0.0908	-0.00849	-0.0266	-0.0383
	(0.164)	(0.0628)	(0.0669)	(0.0543)	(0.0414)	(0.0268)
I^{YC}	0.550***	-0.145***	0.158***	-0.177***	-0.100***	-0.0780***
	(0.126)	(0.0483)	(0.0515)	(0.0418)	(0.0319)	(0.0197)
I^{MNC}	0.604***	0.0872	0.161**	0.335***	0.154***	0.138***
	(0.198)	(0.0759)	(0.0809)	(0.0657)	(0.0501)	(0.0309)
I^{MC}	-0.0374	-0.0744	-0.177***	-0.0215	0.111***	0.114***
	(0.160)	(0.0612)	(0.0652)	(0.0529)	(0.0404)	(0.0247)
N	-0.0297**	-0.00304	0.00159	-0.00457	-0.00318	0.00424**
	(0.0123)	(0.00469)	(0.00500)	(0.00406)	(0.00310)	(0.00197)
Elasticity	-0.135***	0.0301***	0.0335***	0.0228***	0.0173***	0.00100
	(0.00912)	(0.00349)	(0.00372)	(0.00302)	(0.00230)	(0.00151)
1990.year	-0.0903***	-0.0178*	-0.0498***	-0.0223**	-0.0255***	-0.00120
	(0.0282)	(0.0108)	(0.0115)	(0.00934)	(0.00712)	(0.00466)
2000.year	-0.0938***	0.0261**	0.00574	0.00696	-0.000569	0.00239
	(0.0340)	(0.0130)	(0.0139)	(0.0112)	(0.00858)	(0.00557)
2010.year	0.239***	0.0366*	0.119***	0.0246	0.0360***	0.0439***
	(0.0495)	(0.0190)	(0.0202)	(0.0164)	(0.0125)	(0.00727)
Constant	-4.825***	-0.847*	-1.761***	-0.517	-0.627*	-1.545***
	(1.338)	(0.512)	(0.546)	(0.443)	(0.338)	(0.192)
Observations	428	428	428	428	428	428

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

different housing qualities, I run the price regression using the median house price and CMHPI. Table 2.6 reports the OLS and IV results.

The impact of increasing college share on local house prices is robust with respect to different measurements of house prices. It is worth noting that the increasing college share has a similar impact on local median price compared to local average house price. Its impact on the CMHPI is significantly smaller. A 1 percentage point growth in college share increases CMHPI by 1.94%, 17% lower than its impact on the median or mean house price. As CMHPI only measures housing price based on repeated sales, it may overlook the price of newly built houses. The results indicate that college graduates with high lifetime incomes may prefer newly built houses with better quality. In other words, the impact of an increasing college share on housing prices is not uniform across markets with different housing qualities. It has a larger impact on newly built houses of better qualities.

2.4.5 Effects of Changing Income Distributions on Homeownership Rates

Decomposition

I use the IV estimates to project the impact of increasing college share and widening gap in the household income between college- and non-college-educated households that occurred between 1980 to 2010 on the homeownership rates of the four groups. Specifically, I apply the estimated coefficients from the IV estimation to the change in college share and changes in the average household income of the four groups to project their impact on homeownership rates since 1980 on the national level.[10]

Table 2.7 presents the results. The increasing share of households headed by bachelors can account for over half of the observed changes in homeownership rates for all groups. Rising income partially alleviates the downward pressure on homeownership among college graduates. The increasing share of college graduates combined with the changes in households income tend to over-predict the drop in homeownership rates, indicating the possibility that relaxing mortgage credit constraints mitigates the downward pressure on homeownership rates caused by the change in the income distribution.[11]

Projections: 1980 to 2018

In this section, I extend the projection to all years from 1980 to 2018. I use the IV estimates to project the impact of the change in income distribution caused by increasing college share

[10] Coefficients not statistically significant are treated as 0.

[11] Recall that I include time dummies to capture potential time trends in the regression. Time dummies are not applied to the decomposition.

Table 2.6: Robustness Check

	CMHPI-OLS	CMHPI-IV	Median House Price-OLS	Median House Price-IV
κ	1.914***	2.088***	2.070***	2.417***
	(0.202)	(0.272)	(0.196)	(0.261)
I^{YNC}	0.629***	0.649***	0.412**	0.451**
	(0.192)	(0.190)	(0.177)	(0.181)
I^{YC}	0.598***	0.638***	0.525***	0.604***
	(0.136)	(0.146)	(0.125)	(0.139)
I^{MNC}	0.790***	0.732***	0.933***	0.816***
	(0.227)	(0.229)	(0.206)	(0.219)
I^{MC}	-0.109	-0.159	-0.204	-0.304*
	(0.194)	(0.185)	(0.173)	(0.176)
N	-0.0508***	-0.0537***	-0.0357**	-0.0416***
	(0.0170)	(0.0141)	(0.0178)	(0.0135)
Elasticity	-0.145***	-0.145***	-0.144***	-0.143***
	(0.0111)	(0.0106)	(0.0110)	(0.0101)
1990.year	-0.0293	-0.0352	-0.0589*	-0.0706**
	(0.0333)	(0.0326)	(0.0314)	(0.0311)
2000.year	-0.0105	-0.0181	-0.0909***	-0.106***
	(0.0363)	(0.0392)	(0.0320)	(0.0375)
2010.year	0.250***	0.228***	0.257***	0.213***
	(0.0553)	(0.0570)	(0.0464)	(0.0546)
Constant	-8.386***	-7.824***	-5.954***	-4.830***
	(1.421)	(1.540)	(1.323)	(1.476)
Observations	424	424	428	428
R-squared	0.717		0.733	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.7: Effects of Changing Income Distributions on Homeownership Rates

	Observed			Predicted by the model			
	1980	2010	$\triangle$ 2010-1980	$\uparrow$ college share	Own Income Change	Income Change Other	Total
Young Non-College	0.51	0.36	-0.15	-0.084	-0.061	-0.0145	-0.16
Young College	0.55	0.50	-0.061	-0.10	0.016	-0.049	-0.13
Middle Age Non-College	0.73	0.62	-0.11	-0.075	-0.054	-0.018	-0.15
Middle Age College	0.82	0.79	-0.02	-0.05	0.0144	-0.035	-0.07

and changing income of college- and non-college-educated households on the homeownership rates of the four groups for each single year from 1980 to 2018 on the national level. Note that in the empirical analysis, I only use data for four years (1980, 1990, 2000, and 2010). Therefore this exercise includes both in-sample and out-of-sample projections.

The data used in this section comes from the Current Population Survey (CPS). CPS provides household level data on household income, ownership status, and age and education of the household head.[12] For each year, I compute the homeownership rates, the average household income of each of the four groups, and the share of households headed college graduates.

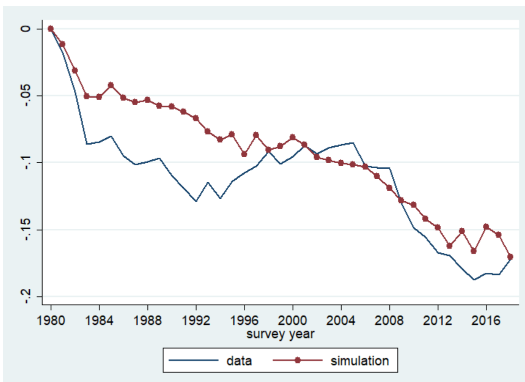
Figure 2.4.5 presents the results. Overall, the model captures the trends in homeownership rates of the four groups both for in-sample test and for the out-of-sample projection, which suggests that changes in the income distribution can account for the observation that while college graduates postpone home purchases, a large fraction of non-college graduates have become long-term renters as owning becomes less affordable. In general, the model over-predicts homeownership rates for all groups between 1980 and 1990 and under-predicts homeownership rates at the beginning of the 2000s. One possible explanation is that the high inflation rate at the beginning of 1980s made owning more attractive (see, e.g. Poterba, 1984). The following adjustments in the inflation rate lowered homeownership rates for all groups. Meanwhile, the mortgage credit expansion in the early 2000s reduced the cost of mortgage for households and led to an increase in homeownership rates for all groups. The large spikes in the observed homeownership rates for young college graduates suggest that their housing tenure choice decisions are more vulnerable to macroeconomic conditions.

2.5 Conclusion

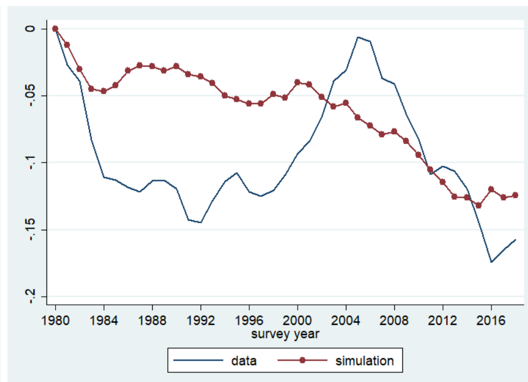
This chapter finds that the drop in young homeownership rate is more persistent among less-educated households. While a considerable fraction of college-educated households are postponing home purchases, many non-college-educated households find owning less attractive and remain as long-term renters. My analysis suggests that the changing income distribution caused by a growing numbers of college graduates and the rising household income of college graduates can account for the diverging ownership decisions between these two groups. The changing income distribution pushes up house prices and lowers homeownership rates for all groups. The increasing college premium partially alleviates the decline in homeownership rate among the college graduates, while non-college graduates without any income growth find owning less affordable.

[12]The sample size in CPS is much smaller than that in the Census and ACS. That is the main reason that I only use CPS to do aggregate analysis.

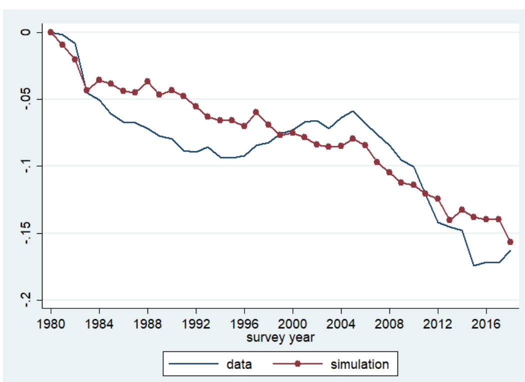
Figure 2.6: Projected Homeownership Rates: 1980-2018



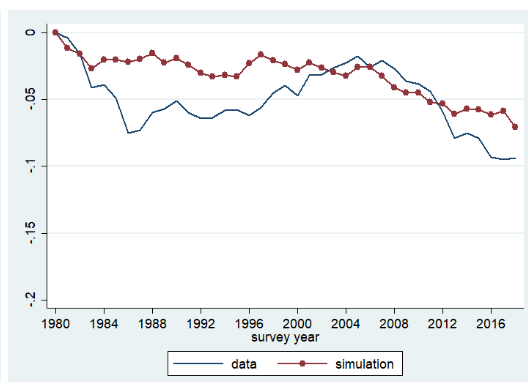
(a) Homeownership Rate: Young Non-college



(b) Homeownership Rate: Young College



(c) Homeownership Rate: Middle Age Non-college



(d) Homeownership Rate: Middle Age College

My results add to the studies which argue the drop in the homeownership rate among young households is temporary. This chapter points out that in the presence of rising income inequality and limited housing supply, owning has become less affordable for the low income, less-educated households. As owner-occupied houses provide a hedge against fluctuations in rents (Sinai and Souleles, 2005) and constitute a major part of the household investment portfolio (Flavin and Yamashita, 2002), reducing access to homeownership among non-college graduates might aggravate wealth and welfare inequality.

In terms of policy implications, my findings suggest that in the presence of rising income inequality, policies that disproportionately favor homeowners could have resulted in an even larger increase in the wealth inequality. One example is the mortgage interest deduction, which allows owners with more mortgage to collect more after-tax savings. As argued in Sommer and Sullivan (2018), eliminating the mortgage interest deduction will lead to a decline in house prices and an increase in homeownership. My findings imply that such a policy modification could be the solution to the affordability problem experienced by the low-income, less educated households.

Bibliography

[1] Alexis Anagnostopoulos, Orhan Erem Atesagaoglu, and Eva Carceles-Poveda. Skill-biased technological change and homeownership. *Journal of Economic Dynamics and Control*, 37(12):3012–3033, 2013.

[2] Jonas DM Fisher and Martin Gervais. Why has home ownership fallen among the young?*. *International Economic Review*, 52(3):883–912, 2011.

[3] Marjorie Flavin and Takashi Yamashita. Owner-occupied housing and the composition of the household portfolio. *American Economic Review*, 92(1):345–362, 2002.

[4] Edward L Glaeser, Joseph Gyourko, and Raven Saks. Why is manhattan so expensive? regulation and the rise in housing prices. *The Journal of Law and Economics*, 48(2):331–369, 2005.

[5] Claudia Goldin and Lawrence F Katz. Decreasing (and then increasing) inequality in america: a tale of two half-centuries. *The causes and consequences of increasing inequality*, pages 37–82, 2001.

[6] John L Goodman Jr and John B Ittner. The accuracy of home owners' estimates of house value. *Journal of Housing Economics*, 2(4):339–357, 1992.

[7] Joseph Gyourko and Peter Linneman. Analysis of the changing influences on traditional households' ownership patterns. *Journal of Urban Economics*, 39(3):318–341, 1996.

[8] Joseph Gyourko and Peter Linneman. The changing influences of education, income, family structure, and race on homeownership by age over time. *Journal of Housing Research*, 8:1–26, 1997.

[9] Joseph Gyourko, Christopher Mayer, and Todd Sinai. Superstar cities. *American Economic Journal: Economic Policy*, 5(4):167–199, 2013.

[10] Lawrence F Katz, David Autor, Orley Ashenfelter, and David Card. Handbook of labor economics. *Handbook of Labor Economics*, 3, 1999.

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

[12] Atif Mian and Amir Sufi. House prices, home equity–based borrowing, and the us household leverage crisis. *American Economic Review*, 101(5):2132–2156, 2011.

[13] Francois Ortalo-Magne and Sven Rady. Housing market dynamics: On the contribution of income shocks and credit constraints. *The Review of Economic Studies*, 73(2):459–485, 2006.

[14] James M Poterba. Tax subsidies to owner-occupied housing: an asset-market approach. *Quarterly Journal of Economics*, 99(4):729–752, 1984.

[15] Albert Saiz. The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125(3):1253–1296, 2010.

[16] Lewis M Segal and Daniel G Sullivan. Trends in homeownership: Race, demographics, and income. *Economic Perspectives-Federal Reserve Bank of Chicago*, 22:53–72, 1998.

[17] Todd Sinai and Nicholas S Souleles. Owner-occupied housing as a hedge against rent risk. *Quarterly Journal of Economics*, 120(2):763–789, 2005.

[18] Kamila Sommer and Paul Sullivan. Implications of us tax policy for house prices, rents, and homeownership. *American Economic Review*, 108(2):241–74, 2018.

Chapter 3

Land and the Distribution of House Prices and Rents across U.S. Cities

3.1 Introduction

Shelter cost is a major part of households expenditures. Therefore, understanding why it differs across cities is important for evaluating the quality of life in different places (see, e.g., Moretti, 2013 and Albouy, 2008). As owning and renting are the most prevalent options to obtain housing services, both prices and rents have been commonly used to measure shelter costs in the literature. In this chapter, I find that although prices and rents are highly correlated, the distribution of house prices across cities differs from rents. Moreover, house prices have become more dispersed across cities during the past 30-50 years compared to rents.[1] These findings suggest that using prices and rents interchangeably to measure shelter costs could be problematic. Moreover, understanding the joint distribution between the two shelter costs has important implications for housing tenure choice and household welfare in different places. In this chapter, I first establish three stylized facts about the distributions of prices and rents across cities in the U.S. and then propose a mechanism that can account for these observations.

Distribution of Prices and Rents across Cities: First, while both prices and rents differ substantially across cities, price dispersion is 80% larger than rent dispersion. For instance, as Table 3.1 shows, the Coefficient of Variation (CV) of house prices is 0.33 while the CV of rents across cities is only 0.18 in 1980. Second, the price dispersion has increased more than rent dispersion between 1980 to 2010. The dispersion of house prices has almost doubled while the dispersion of rents has only increased by 50% (Table 3.1). Third, the correlation between prices and rents is high, both in levels and in growth rates. Cities that have higher (growth rates

[1] Gyourko, Mayer, and Sinai (2013) and Van Nieuwerburgh and Weill (2010) among others document the significant increase in house price dispersion during the past 30-50 years.

of) prices normally have higher (growth rates of) rents (Section 3.2).[2]

The main mechanism that I propose to account for the joint distribution of prices and rents across cities is motivated by the fact that most owners live in detached houses while most renters live in apartments. Houses and apartments differ in the use of land, which is the scarce resource in big cities. To quantitatively evaluate the implication of land use difference between houses and apartments on the joint distribution between prices and rents, I develop a city-level housing tenure choice model where owner-occupied houses take more land to build compared to rental apartments. The model is calibrated to the distribution of house prices and rents across the cities in the year 1980. Feeding into the model factors that affect the housing demand and residential land supply in 2010, I show that the model can account for the majority of changes in the dispersion of prices and rents from 1980 to 2010.[3]

Table 3.1: First and Second Moments of Prices and Rents in the data

Variable [a]	Obs.	1980		2010	
		Mean	CV	Mean	CV
Price	181	120900	0.33	146270	0.62
Rent	181	6038	0.18	6264	0.27

[a]Prices are measured by the trimmed mean value of detached houses. Rents are measured by the trimmed mean rent of 2 bedroom apartment. More details about the data can be found in Section 2.1.

According to the American Housing Survey (2009), around 93% of owner-occupied homes are detached with a median lot size of 14,000 square feet (sqf), while the majority of renters (around 63%) live in multifamily apartments with a median land usage of 367 sqf.[4] In other words, the "land footprint" of owner-occupied homes is much larger than that of rental apartments. Moreover, in cities where land is expensive, apartments economize on land use by building up. However, detached houses cannot have less land than the size of the first floor and can be constrained by the minimum lot size requirement imposed by zoning regulations. As a result, the cost of building houses is disproportionally higher compared to apartments in these cities due to the heavy reliance on land as an input.

To quantitatively evaluate the implication of different land use between houses and apartments on the dispersion of prices and rents as well as their changes over time, I develop a

[2]Appendix B.1 provides the density plot of prices and rents.

[3]I choose the year 2010 in the simulation for two reasons. First, the dataset I am using adopts a consistent definition of metropolitan areas from 1980 to 2010. Second, during the housing boom period, house prices may contain bubbles, i.e. part of the house price growth cannot be rationalized by economic fundamentals. As this chapter focuses on the impact of economic fundamental in the steady states, I choose the year 2010 in the simulation instead of the years around 2007.

[4]Author's calculation. Land use per unit is defined as the ratio between unit size and the number of floors in the multifamily building.

city-level housing tenure choice model with competitive housing supplies. The key feature is that while both owner-occupied houses and rental apartments are constructed using land and material, owner-occupied houses are more land-intensive.

The model allows me to examine how the land values are determined by economic fundamentals (i.e., income and population) and land supply, through calibration in the baseline year 1980. Then I test the performance of the model by feeding in the economic fundamentals and land supply in 2010 to see whether it can account for changes in the distributions of prices and rents from 1980 to 2010.

In the model, each city is an isolated island inhabited by finitely-lived households. Households choose between purchasing a house and renting an apartment, as well as the size to buy/rent in each period. A down payment is required for households who purchase a house. Competitive construction firms build houses and apartments using land and material. A key model feature is that the production function for houses has a higher land share compared to apartments. In addition, houses are subject to a city-specific minimum lot size requirement, which specifies a lower bound on the land input per house. The total amount of land available for residential construction in each city is assumed to be exogenous.

The higher land intensity in building houses allows the model to endogenously generate the feature that prices increase more in land values than rents do, consistent with the empirical evidence. In cities where the price of land is high, both house and apartment developers substitute expensive land with relatively cheap material. However, the reduction in land used for houses is smaller than apartments for two reasons. First, the higher land share in the production function of houses implies that land and material are less substitutable, i.e., the ratio between the marginal productivity of land and material is larger in producing houses compared to apartments for any land-material input ratio. Second, the minimum lot size requirement puts a floor on the land use per house. As a result of the intensive use of land, the total cost of building houses, which equals the house price in a competitive market, grows more than that of apartments, which equals the net present value of rents, when land value rises. From the cross city perspective, when the distribution of land values across cities changes, the dispersion of prices should change more than rents.

The model is first calibrated to house prices, rents, and the fraction of households living in detached or attached houses for each of the largest 181 cities in the U.S. in 1980.[5] My main numerical exercise is to change population, income, residential land supply, and the down payment requirement to the 2010 level to examine whether the model can account for the large increase in price dispersion and the moderate increase in rent dispersion at the same time.

I find that changes in income, population, residential land supply, and the down payment

[5]Appendix B.3 provides a full list of the 181 cities.

requirement between 1980 and 2010 can account for 82% of the increase in price dispersion and 56% of the increase in rent dispersion. The model captures the changing relationship between prices and rents as well. It can also account for 90% of the increase in the dispersion of price-rent ratios across cities from 1980 to 2010, which implies that most of the rise in the price-rent ratios in large cities can be attributed to these fundamentals. This challenges the conventional view that a rise in the ratio of house prices to rents signals a housing bubble.

This chapter provides additional insights on the interaction between owner-occupied and rental markets, as well as the dynamics of prices, rents, and homeownership rates.[6] As households' owning/renting decision depends on both prices and rents, understanding how prices and rents interact is important for estimating the response of aggregate housing demand to policy changes. The previous literature largely adopts one of the two frameworks. In models where rental units are provided by risk-neutral real estate firms (Gervais, 2002; Yang, 2009), rents change at the same rate as prices. In models where the supply of rental units is endogenously provided by the owners (Chambers, Garriga, and Schlagenhauf, 2009; Sommer, Sullivan, and Verbrugge, 2013; Favilukis and Van Nieuwerburgh, 2017), price and rent may change in different directions. For instance, an increase in the credit supply may lead to higher prices and more rental supply, which contributes to lower rents. This chapter examines another channel, the role of different land intensities between houses and apartments, which delivers the positive correlation between prices and rents while allowing the growth rates of prices and rents to differ.

The mechanism proposed in this chapter helps reconcile the documented difference in price and rent elasticity with respect to either demand shocks or supply changes. For instance, Saiz (2007) finds that the impact of changing income or immigration on median prices is 40% to 80% larger than that on median rents. Focusing on the impact of relaxing credit constraints, Greenwald and Guren (2019) find that the elasticity of prices to credit is between 0.30 to 0.38 while the elasticity of rents is between 0.21 to 0.26. In addition, Parkhomenko (2018) finds that the elasticity of prices with respect to land scarcity is 0.051, which is twice as big as the elasticity of rents with respect to land scarcity, 0.024. The mechanism proposed in this chapter suggests that in the long-run equilibrium, demand for houses and apartments are aggregated into demand for land through different production functions. The higher land share in producing houses implies that price responds more than rent when the underlying land value changes due to either changes in total land demand and/or supply.

This chapter builds on a growing literature which examines alternative explanations for the rising dispersion of house prices across U.S. cities. Gyourko, Mayer, and Sinai (2013) discuss

[6]For instance, Greenwald and Guren (2019) argue that perfect segmentation of these two markets implies a large impact of credit supply while complete integration implies no impact of credit supply on price.

the impact of sorting on preference towards specific locations on housing price dispersion across cities. Van Nieuwerburgh and Weill (2010) explore how sorting on ability, i.e., people with high ability moves to more productive places, drives up price dispersion across cities. Noticeably, Van Nieuwerburgh and Weill (2010) over-predict the increase in the rent dispersion as they do not distinguish owner-occupied from rental units. In this chapter, I consider another type of dwelling which is less land-intensive and only available for rent. Such an extension allows me to complement the previous literature by accounting for both the substantial rise in the price dispersion and the simultaneous moderate increase in the rent dispersion.

Closest in spirit to this chapter are Davis and Heathcote (2007), Davis and Palumbo (2008), and Larson et al. (2019), who estimate residential land values using data on home values and costs of housing structures on county-, city-, and country-level.[7] They find that land prices have become more important in determining house prices over the past three decades. This chapter differs in modeling the demand and supply of the land market explicitly. The housing tenure choice model developed in this chapter maps economic fundamentals to the demand of houses and apartments, which is then translated into the demand of land via production functions. The upside of working with this equilibrium model is that it allows me to evaluate the impact of changes in economic fundamentals on the land values which determine house prices and rents.

The mechanism proposed in this chapter allows for a flexible relationship between prices and rents, which is important for explaining the cross-city variation in the price-rent ratios. Previous studies that treat rents as the dividends of housing assets find that the standard asset-pricing approach does not work well to account for the variation in price-rent ratios across cities. Therefore, Glaeser and Gyourko (2007) argue that owner-occupied houses and rental units differ, such that prices and rents are not directly comparable. However, I find that prices and rents are highly correlated both in levels and in growth rates. The mechanism that I propose allows a non-linear relationship between prices and rents. It rationalizes the high correlation (> 0.7) between prices and rents. As both houses and apartments are constructed using the same inputs, land and material, the equilibrium house prices and apartments rents both depend on land values and material values. Meanwhile, it allows price and rent to diverge when land value changes. Due to the higher land share in building houses, the equilibrium house prices change more with land values compared to rents.

My study also complements the literature on the estimation of housing production functions. Previous work focuses on estimating the production function of single-family detached homes (Albouy, Ehrlich, and Shin, 2018; Epple, Gordon, and Sieg, 2010). In addition to

[7]The county-level average price of land used in singe-family housing is available from 2012 to 2018. The city-level price of residential land is available from 1984 to 2014. The country-level price of residential land is available at ten year intervals between 1930 and 2000, and annually from 1975 to 2006.

single-family detached homes, I estimate the production function of multi-family apartments and find that the land share in constructing houses is almost twice as big as the land share in the construction of multi-family buildings.

The remainder of the chapter is organized as follows. Section 3.2 provides empirical evidence. Section 3.3 lays out the model. Section 3.4 describes the calibration strategy. Section 3.5 discusses the calibration results. Section 3.6 presents and discusses the quantitative exercise. Section 3.7 concludes.

3.2 Empirical Evidence

This section documents empirical evidence supporting the main mechanism explored in the chapter. First, I show the two shelter costs, prices and rents, are highly correlated both in levels and in growth rates. Cities that have higher (growth in) prices tend to have higher (growth in) rents. Second, I examine the implications of the main mechanism. Specifically, the higher land share in the production function of owner-occupied houses implies that (1) the ratio between price and rent is increasing in land value and (2) the cost of building houses grows more than the cost of building apartments when land value changes. The data confirms these two implications. I find that the price-rent ratio is increasing in land values. Moreover, for around 30 cities with a sufficient number of condo owners, the price growth for single family detached homes is higher compared to condos from 1980 to 2010. In addition, as land plays an important role in my mechanism, I examine the use of land between houses and apartments across cities. I find that as population density, i.e. an approximation for land scarcity, increases, both houses and apartments use less land: houses have smaller lots while apartments increase the number of stories. On average, apartments reduce land use more compared to houses.

3.2.1 Data

Apartment rents: I construct an estimate of the rent of a standard two-bedroom apartment using data from the 1980 Census and the American Community Survey in 2010. To eliminate the influence of rents of luxury and low-quality apartments, I use the trimmed mean and discard the top and bottom 5 percent.

House prices: To construct an estimate of house prices, I combine the mean value of single-family homes from the 1980 Census with the Freddie Mac Conventional Mortgage Home Price Index (CMHPI), an index based on repeated sales to approximate the house price of constant quality, following Van Nieuwerburgh and Weill (2010).[8] To be consistent with the rent data, I

[8] The CMHPI is a quarterly index. I get the annual index by taking the average of the four quarters.

use the trimmed mean house value in 1980 with the top and bottom 5 percent discarded.[9]

House price and apartment rents are deflated by the national Consumer Price Index (CPI).[10]

Characteristics of houses and apartments: Land use and the unit size for houses and apartments are computed using data from the American Housing Survey (AHS), 2009 National Sample.[11] The AHS provides detailed information on the year of construction, type of the dwelling (e.g., detached, attached, or multi-family buildings), lot size, unit size, number of stories, and housing tenure status.

3.2.2 Correlation between Prices and Rents

The correlation between prices and rents is high both in levels and in growth rates. Figure 3.1 (a) displays the correlation between prices of detached houses and rents of apartments across MSAs in the year 1980. The correlation between these two series is 0.71. This positive correlation is persistent over time and rose to 0.90 in 2010. Figure 3.1 (b) plots the growth rates of apartment rents against the growth rates of house prices from 1980 to 2010. The correlation is 0.7. Specifically, a 1% increase in house price is associated with a 0.48% increase in apartment rent.

3.2.3 Price-rent ratios and Land Values

One key model prediction is that price grows more than rent when the cost of land increases as owner-occupied houses are more land-intensive than rental apartments. Figure 3.2 plots the ratios between prices and rents against the transaction-based land values provided by Albouy, Ehrlich, and Shin (2018) in 2010. Each point corresponds to one city. The correlation is positive (around 0.6) and significant, indicating that price grows more than rent with land values, which is consistent with the implication of the mechanism.

3.2.4 Price Growth: Houses and Apartments

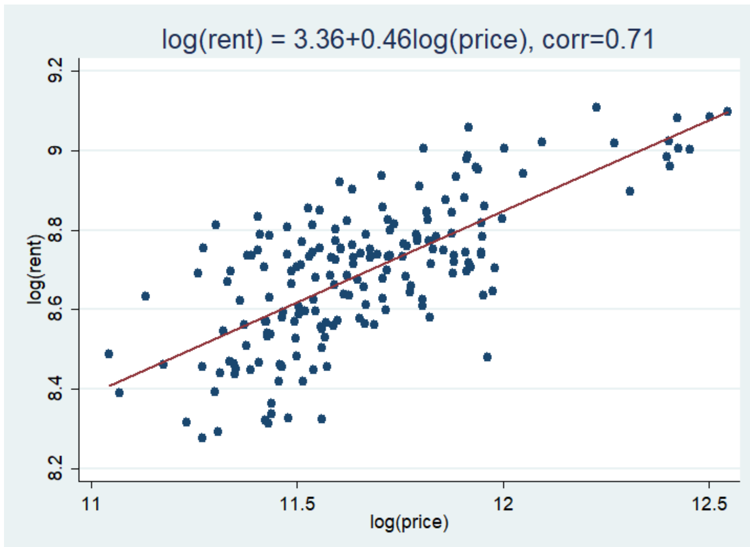
I have been comparing the price of houses with the rent of apartments due to the fact that most owners live houses while most renters live in apartments and we want to understand the

[9] I use trimmed mean house values and apartment rents instead of the commonly used median as house values and contract rents reported in Census were categorical in 1980.

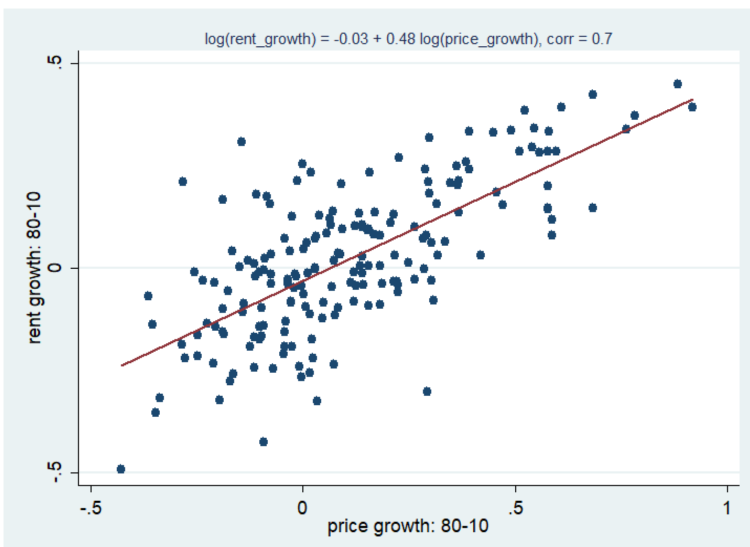
[10] As I focus on the dispersion of prices and rents, which is measured by the Coefficient of Variation, using GDP deflator to adjust would not affect the results.

[11] AHS 2009 is subtracted from National Microdata from the Inter-university Consortium for Political and Social Research (ICPSR).

Figure 3.1: Correlation between House Prices and Apartment Rents

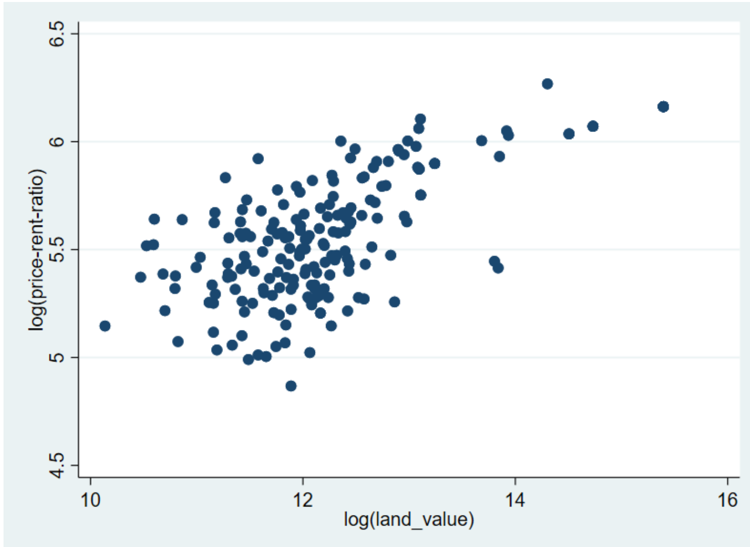


(a) Level in 1980



(b) Growth Rates: 1980-2010

Figure 3.2: Price-rent Ratio and Land Values



relationship between these two shelter costs.[12] A more direct comparison in prices between houses and apartments would allow us to examine the importance of the difference in land intensities in determining the values of these two types of dwelling. Figure 3.3 displays the mean price growth for houses and apartments that were built between 1950 to 1970 for cities where we observe owner-occupied houses and owner-occupied apartments at the same time.[13] A 45-degree line is included to facilitate the comparison. Consistent with the implications of the mechanism, price growth for the land-intensive single-family detached houses is larger than that for multi-family apartments for all cities in the sample from 1980 to 2010.

3.2.5 Land Use in Houses and Apartments

According to the 2009 American Housing Survey, 95% of owners live in a detached or semi-detached house, while 63% of renters live in a multi-family apartment. As land is the key component in my mechanism, I compare the land use of houses and apartments across cities.[14] I find that land input in building apartments is more responsive to land scarcity, which is approximated by the population density, compared to houses.

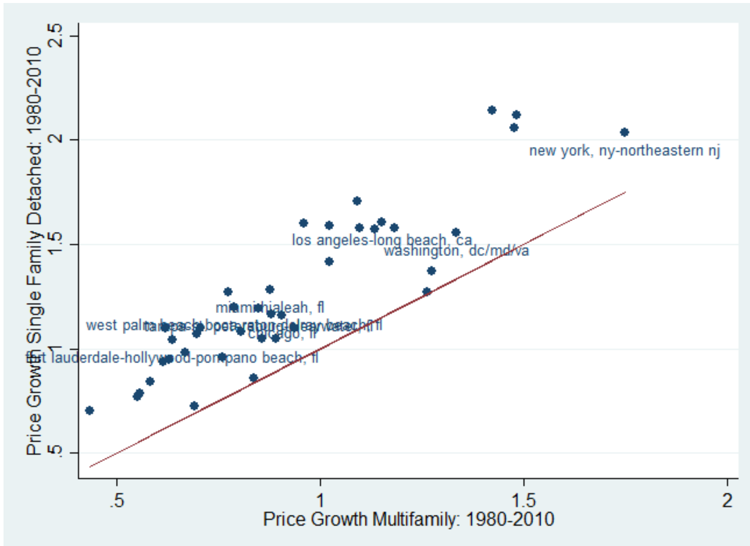
Figure 3.4 plots the land input in building one square foot of living space against local

[12]For instance, even in the biggest city, New York City, 55% of the owners live in houses, and 94% of the renters live in apartments.

[13]Due to the limited stock of owner-occupied condos in the 1980, a small numbers of high-quality expensive apartments built in recent years will significantly affect the mean/median value of owned condos. Therefore, I exclude apartments and houses that are built in recent years.

[14]Appendix B.2 presents more evidence about the land use in different cities and its change over time for houses and apartments.

Figure 3.3: Price Growth: Houses and Apartments



population density for houses in panel (a) and apartments in panel (b) in 2010.[15] A 1% increase in local population density is associated with a 10% decline in the land use for building one square foot living space for apartments. Meanwhile, the land input response against population density is not very significant in house constructions.

3.3 Model

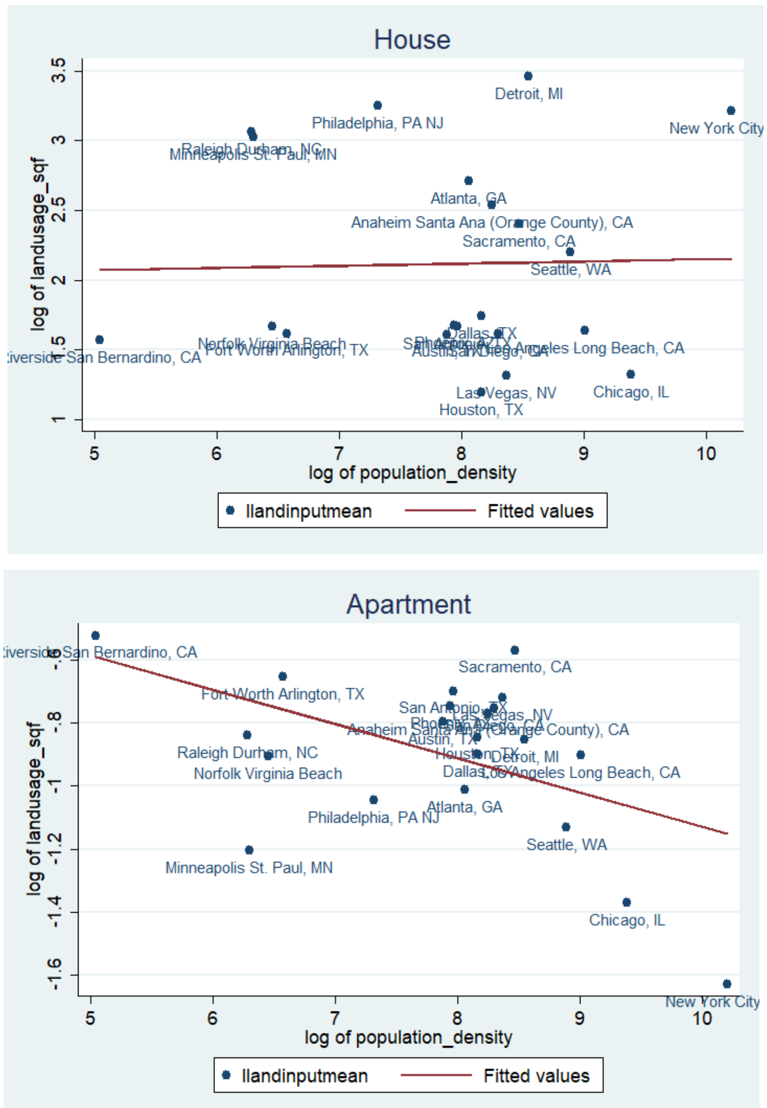
The mechanism suggests that prices and rents differ across cities as underlying land values vary. In addition, the distributions of prices and rents across cities differ due to the differences in land intensities. To quantify the difference in land intensities between houses and apartments and to investigate how land values are determined, I develop a city level housing tenure choice model. I use this model to answer the following questions. First, what factors determine the land values? Second, can change in these factors across cities predict the change in land values that can account for the rise in the dispersion of prices and rents at the same time?

3.3.1 Model Overview

Each city is an isolated island. In other words, households do not move across islands. Cities differ in total population, income, land supply, material cost, the ownership premium of the local residents, and the minimum lot zoning regulation.

[15]The land input is measured by the one divided by the number of stories in the building for apartments and lot size divided by the unit size for houses.

Figure 3.4: Land Input against Population Density: Houses vs Apartments



On each island, households live up to J periods. Households are ex-ante identical. They maximize their expected discounted lifetime utility from the consumption of a non-durable good and housing services. At the beginning of each period, households receive income that depends on their age, the city they live in, and an idiosyncratic shock.

To obtain housing services, households choose between buying a house and renting an apartment. Owners get additional utility from living in owned houses compared to renters. Only owners can borrow against (a certain fraction of) the value of their housing asset.

Houses and apartments are produced using land and material but via different production technologies. Houses are more land-intensive. In addition, the construction of houses is subject to a city-specific minimum lot size requirement, which places a lower bound on the amount of land input.[16] Following Albouy and Ehrlich (2018) and Epple, Gordon, and Sieg (2010), I assume that the markets of houses and apartments exhibit perfect competition. As a result, the equilibrium price of a house equals to the total construction cost. Similarly, the net present value of rents equals to the total construction cost of the apartment.

3.3.2 Households

Preference

Each household has preferences defined over a non-durable good and housing service represented by:

$$\sum_{j=1}^{J} \beta^{j-1} (\prod_{l=1}^{j-1} \psi_l) u(c_j, s_j) \tag{3.1}$$

where β is the discount factor, ψ_l is the probability that a household survives from age j to age $j+1$, c_j is consumption, and s_j is housing service received at age j. The $u(.)$ is a C^2, increasing, and concave function.

A household can obtain housing service by buying a house or renting an apartment.

$$s = \begin{cases} h & \text{if Rent} \\ \theta_k h \mathbb{1}_{j \geq j_o} \zeta & \text{if Own} \end{cases}$$

Owners on island k derive additional utility θ_k compared to renters. θ_k is supposed to capture the quality difference between a standard owner-occupied house and a standard rental apartment and how residents in that city evaluate this quality difference. Similar to Fisher and Gervais (2011), I assume that old households, i.e., over age j_o, discount ownership premium by a factor ζ. Senior households may not enjoy big houses as much as young households due to smaller

[16]This captures zoning laws, see e.g., Isakson, 2004; Bucovetsky, 1984.

family size, and the fact that houses with yards involve heavy housework such as cleaning the snow or maintaining the lawn.

Household Income

Households from island k receive exogenous household income that depends on their age, location, and an idiosyncratic shock. The household income of a household i of age j on island k is represented by

$$lny^i_{j,k} = ln(w_j) + ln(\bar{w_k}) + log(\epsilon^i_j) \tag{3.2}$$

where w_j is the average household income of an age j household, $\bar{w_k}$ is the average income of households on island k relative to the average income of the whole economy, and ϵ^i_j is the idiosyncratic shock that follows an AR(1) process.

$$ln(\epsilon^i_j) = \nu ln(\epsilon^i_{j-1}) + \xi_{i,j}, \xi_{i,j} \sim N(0, \Sigma^2_y) \tag{3.3}$$

The AR(1) process is the same for households of all ages and on all islands.

Asset Arrangement

Following the literature, I assume that only collateralized credit is available.[17] The borrowing interest (i.e., the interest on mortgages) equals the deposit interest r plus a spread r_m. The net asset position is denoted by a. To buy a house, a household must satisfy a minimum down payment requirement of γ. That is, a household's financial asset always satisfies

$$a \geq -(1-\gamma)P(h) \tag{3.4}$$

where $P(h)$ represents the price of a size h house.[18]

A newborn household has no housing asset and draws his/her initial wealth from a probability distribution Π_w defined on $\mathbb{R}_+$.

Costs Related to Owner-Occupied Houses

It is costly to buy or sell a house. These costs include the opportunity cost of time associated with the market search, brokerage, and moving, as well as legal fees. Following the literature (see e.g., Sommer, Sullivan, and Verbrugge, 2013), I assume that the transaction costs are proportional to the value of the house. Specifically, a buyer incurs a total transaction cost of

[17] See e.g. Yang (2009), Sommer, Sullivan, and Verbrugge (2013).

[18] Equation 3.13 shows that due to the minimum lot size requirement, the price of a house is not linear in its size for the whole support.

$k_bP(h)$, and a seller incurs a total transaction cost of $k_sP(h)$. In addition, houses depreciate at a rate δ, each year.

Household's Recursive Problem

At the beginning of each period, the state variable of a household in one city is given by (h, a, ϵ, j), which correspond the current housing stock, financial stock, income shock, and age, respectively.

Owner's Problem An owner enjoys his/her current housing stock h, and chooses consumption c, future asset a', and future housing stock h' to maximize his/her expected value

$$V(h, a, \epsilon, j) = max_{c,a',h'} u(c, h) + \beta \sum_{\epsilon'} \pi(\epsilon'|\epsilon)V(h', a', \epsilon', j + 1) \tag{3.5}$$

subject to

$$\begin{aligned}
&(k_bP(h') + k_sP(h))\mathbb{1}_{h\neq h'} + c + P(h') + a' + \tau P(h) = (1 - \tau_w)w_j\bar{w}_k\epsilon + \\
&\mathbb{1}_{a\geq 0}(1 + r)a + \mathbb{1}_{a\leq 0}(1 + r + r_m)a + (1 - \delta)P(h) \\
&a' \geq -(1 - \gamma)P(h') \\
&c \geq 0
\end{aligned} \tag{3.6}$$

where k_b and k_s are the transaction cost on buyers and sellers, respectively. τ is the property tax on owners, r is the real interest rate, and $r + r_m$ is the mortgage interest. τ_w is the income tax rate, and δ is the depreciation rate on houses.

Renter's Problem A renter with current housing stock $h = 0$, chooses consumption c, future housing stock h', future asset a', and the size of apartment to rent in the current period h^r to maximize his/her value

$$V(0, a, \epsilon, j) = max_{c,a',h',h^r} u(c, h^r) + \beta \sum_{\epsilon'} \pi(\epsilon'|\epsilon)V(h', a', \epsilon', j + 1) \tag{3.7}$$

subject to

$$\begin{aligned}
&c + (1 + k_b)P(h') + R(h^r) + a' = (1 - \tau_w)w_j\bar{w}_k\epsilon + (1 + r)a \\
&a' \geq -(1 - \gamma)P(h') \\
&c \geq 0
\end{aligned} \tag{3.8}$$

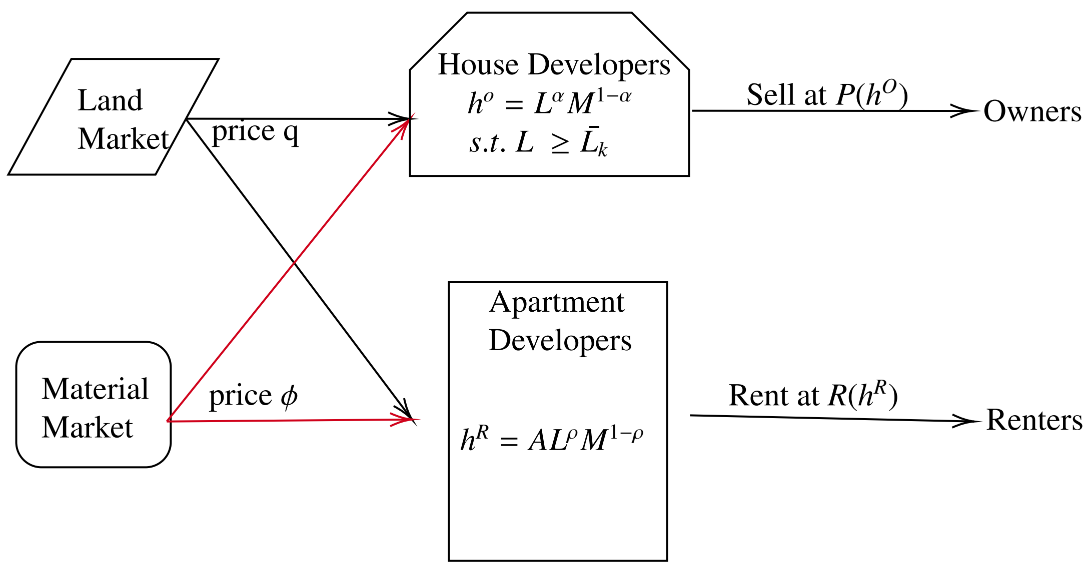
where $R(h^r)$ is the total rent which is a function of the apartment size.

3.3.3 Housing Supply

Owner-occupied houses and rental apartments are produced through Cobb-Douglas production functions that differ in land shares.[19] Moreover, the construction of houses is subject to a city-specific minimum lot size requirement, which may prevent developers from reducing land inputs when land prices are high. The land is owned by absentee landlords who consume the profit from selling land.

Figure 3.5 displays the market structure. Each island has a continuum of competitive developers that purchase material and land to construct houses and apartments. Developers are price-takers. I assume that developers on island k can purchase material at an island-specific constant price ϕ_k. The supply of land is given exogenously with the price of land q_k adjust to clear the land market.

Figure 3.5: Housing Market Structure



Houses are constructed through a production function with a land share α.

$$h^O = L^\alpha M^{1-\alpha} s.t. \quad L \geq \bar{L_k} \tag{3.9}$$

To produce a house of size h^O, a developer needs L units of land input and M units of material input. The land input L (lot size of a house) cannot fall below the minimum lot size $\bar{L_k}$ in city k.

[19]The substitutability between land and other inputs (material) is a common assumption in the literature, see e.g., Albouy and Ehrlich (2018), and Epple, Gordon, and Sieg (2010). Moreover, Larson et al. (2019) find that land prices tend to rise faster than house prices, which supports the functional form used in this chapter.

Apartments are built using a production function with land share, ρ.

$$h^R = AL^\rho M^{1-\rho} \tag{3.10}$$

where A is the relative productivity of apartments compared to houses. A captures the fact that a standard rental apartment uses less material and less land compared to a standard house due to the physical difference between these two types of dwellings.[20]

To build a house of size h^O, developers take the price function $P_k(.)$, land price q_k, and material price ϕ_k as given, and choose land input and material input to minimize the cost of each house they build. Upon finishing, developers sell the house to owners.

$$\begin{aligned} &\min_{L^o, M^o} q_k L^o + \phi_k M^o \\ \text{s.t. } h^o &= (L^o)^\alpha (M^o)^{1-\alpha} \\ L^o &\geq \bar{L}_k \end{aligned} \tag{3.11}$$

Similarly, apartment developers choose land and material input to minimize the cost of each apartment they produce. Upon finishing, developers rent the apartments to renters and collect rents every period.

$$\begin{aligned} &\min_{L^r, M^r} q_k L^r + \phi_k M^r \\ \text{s.t. } h^r &= A(L^r)^\rho (M^r)^{1-\rho} \end{aligned} \tag{3.12}$$

3.3.4 Characterization of Stationary Competitive Equilibrium

As cities are isolated (i.e. households cannot move across cities), each city is a closed economy that is described by a unique competitive equilibrium. Appendix B.4 provides the definition for the stationary competitive equilibrium. This section focuses on characterizing the equilibrium.

The competitive market assumption implies zero profit for developers. The price of a house equals the total cost of building it. Solving the house developers' maximization problem (Equation 3.11) delivers the following price function of house size.

$$\begin{aligned} P_k(h^o) &= f(\alpha) q_k^\alpha \phi_k^{1-\alpha} h^o \;\; if \;\; {\frac{\alpha\phi_k}{(1-\alpha)q_k}}^{1-\alpha} h^O \geq \bar{L}_k \\ &= q_k \bar{L}_k + {\frac{h^o}{\bar{L}_k^{\,\alpha}}}^{\frac{1}{1-\alpha}} \phi_k \;\; otherwise \end{aligned} \tag{3.13}$$

[20] $A > 1$ implies that using the same amount of land and material, developers can produce more apartments than houses due to the fact that apartments are smaller. A captures the structural difference between a standard owner-occupied house and a rental apartment, while θ_k in the utility function captures how people treat this difference.

where $f(\alpha) = \alpha^{-\alpha}(1-\alpha)^{-1+\alpha}$. The lot size of a house or the land input depends on the relative price of material and land, as well as the house size. When the value of land is high compared to material, developers substitute land with materials, until the minimum lot size binds. The minimum lot size may also bind if the house size is too small such that not much land is required. When the minimum lot size does not bind, as the production function is constant return to scale, the price is linearly increasing in house size. When the minimum lot size binds, the price function is convex in house size. The minimum lot size puts a lower bound on the price function, which equals the cost of buying the minimum lot.

The corresponding optimal land use for a house of size h^o on island k is

$$\begin{aligned} L_j^o(h^o) &= \frac{\alpha\phi_k}{(1-\alpha)q_k}^{1-\alpha} h^o \text{ if } \frac{\alpha\phi_k}{(1-\alpha)q_k}^{1-\alpha} h^o \geq \bar{L}_k \\ &= \bar{L}_k \text{ otherwise} \end{aligned} \tag{3.14}$$

Assuming apartment developers are risk neutral, the zero profit condition for rental units implies that the net present value of future rent flows equals the construction cost of that unit.

$$\frac{(1-\delta_r-\tau)R(h^R)}{r} = \frac{f(\rho)q_k^{\rho}\phi_k^{1-\rho}h^r}{A} \tag{3.15}$$

where δ_r is the management cost of rental units and τ is the property tax. The management cost includes salaries, insurance, utilities, management fee, administrative, marketing, contract services, and repair/maintenance (Goodman, 2004).

The corresponding optimal land use of an apartment of size h^r is

$$L_k^r(h^r) = \frac{\rho\phi_k}{(1-\rho)q_k}^{1-\rho} h^r \tag{3.16}$$

The literature on housing tenure choice typically assumes that both price and rent increase linearly in house/apartment size. To prevent households from buying a very small home and to match homeownership rates, a common assumption is that there is a minimum house size for owner-occupied dwellings (see e.g., Chambers, Garriga, and Schlagenhauf, 2009; Sommer, Sullivan, and Verbrugge, 2013). This chapter assumes a minimum lot size constraint, which is qualitatively equivalent to a minimum house size constraint.[21] Such a modification allows prices to grow at an increasing rate with land values when land values are high.

Proposition 3.3.1 *A 1% increase in land price leads to an increase in price bounded from below by α%, and a ρ% growth in rent.*

[21]Equation 3.14 presents the linear relationship between land usage and house size. Conditional on land price and material price, minimum house size binding is equivalent to the minimum lot size binding.

Proof According to Equation 3.15, $\frac{\partial lnP_k(h^o)}{\partial lnq_k} = \alpha$ if $\frac{\alpha\phi_k}{(1-\alpha)q_k}^{1-\alpha} h^o \geq \bar{L}_k$, that is, if minimum lot size does not bind, a 1% increase in the land price q_k leads to an α% increase in the price of a fixed-quality house.

When it binds, $\frac{\partial lnP_k(h^o)}{\partial lnq_k} = \frac{q_k L_k^{1/(1-\alpha)}}{q_k L_k^{1/(1-\alpha)} + \phi_k (h^o)^{1/(1-\alpha)}} > \alpha$ as $\frac{\alpha\phi_k}{(1-\alpha)q_k}^{1-\alpha} h^o < \bar{L}_k$, the price grows at a rate greater than α.

On the other hand, if $\frac{\partial lnR_k(h^r)}{\partial lnq_k} = \rho$, a 1% increase in land price leads to a ρ% increase in the rent of a fixed-quality apartment.

Proposition 3.3.1 implies that as long as $\alpha \geq \rho$, when land price changes, the price of a fixed-quality house changes more than the rent of a fixed-quality apartment. Proposition 3.3.1 formalizes the idea that prices may diverge from rents when land value changes due to the intensive use of land in building houses.

The total land used for constructing houses and apartments equals to the exogenous supply of land, LS_k.

$$\int L^o(h^o)H^o d(h^o) + \int L^r(h^r)H^r d(h^r) = LS_k \tag{3.17}$$

where $L^o(h^o)$ is the land used for constructing a house of size h^o. H^o is the number of households who desire owning a house of size h^o given the equilibrium price function $P^*(h^o)$. Similarly, $L^r(h^r)$ is the land used for constructing an apartment of size h^r. H^r is the number of renters who desire renting an apartment of size h^r given the equilibrium rent function $R^*(h^r)$. The land market clearing condition pins down the equilibrium land price q_k^*.

3.4 Calibration Strategy

The model is calibrated in three stages. In the first stage, values are assigned to parameters that can be determined directly from the data or from the literature. In the second stage, parameters in the production functions are estimated through an Instrumental Variable (IV) approach. In the last stage, city-specific parameters are calibrated to homeownership rates of different age groups. Due to the model's assumption that owners live in houses and renters live in apartments, the homeownership rates in the model are calibrated to the fraction of households living in detached/attached homes in the data.[22] Given this chapter has a special focus on the land market and the high correlation between homeownership rates and the fractions of households living in detached/attached homes, this assumption highlights the land use channel and simpli-

[22]As the majority of owners live in detached/attached houses while most of the renters live in apartments, the correlation between homeownership rate and the fraction of households living in detached/attached houses is around 0.80 across cities in 1980. The fraction of households who are owners and living in apartments is 6% and the fraction of households who are renters and living in detached/attached homes is 9.6% for the 181 cities in my sample.

fies the analysis.

3.4.1 Pre-determined Parameters

Due to the large number of cities with each being described by a general equilibrium model and the city-specific parameters to estimate, it is computationally burdensome to estimate the common parameters, i.e., parameters that apply to all cities, through solving the model. Therefore, I take these common parameters directly from the literature. Table 3.2 summarizes these parameters and their sources.

Preference

Following Chambers, Garriga, and Schlagenhauf (2009) and Fisher and Gervais (2011), the per-period utility takes the following form

$$u(c, s) = ln(c) + \frac{s^{1-\sigma}}{1-\sigma} \tag{3.18}$$

This utility function treats housing as a necessity so that the expenditure share on housing rises with the costs of housing services.

Following the literature on housing tenure choice (see, for example, Sommer, Sullivan, and Verbrugge, 2013; and Yang, 2009), the risk aversion parameter, σ is 2.5, and the discount factor β is 0.95 per year. The discount factor on the ownership premium for old households, $\zeta = 0.9116$, is taken from Fisher and Gervais (2011).

Market Arrangements

Transaction costs for buyers and sellers are set to, $k_b = 0.07$ and $k_s = 0.025$, based on Gruber et al. (2004). Following Sommer, Sullivan, and Verbrugge (2013), the depreciation rate of owner-occupied houses $\delta = 0.025$. The management cost of rental unit $\delta_r = 0.33$, comes from Goodman (2004). Consistent with the literature (see e.g. Sommer, Sullivan, and Verbrugge 2013; Chambers, Garriga, and Schlagenhauf, 2009), down payment requirement, $\gamma = 20\%$, in the baseline calibration. It is lowered to 10% in the simulation of 2010 to capture the mortgage credit expansion from 1980 to 2010.

The risk-free interest rate is, $r = 4\%$, per year and the spread on mortgage, $r_m = 1.5\%$, which are consistent with the literature (see e.g. Amior and Halket, 2014).

Demography and Income

Each period in the model is set to 5 years. At age 20, households enter the model. At the beginning of each period, households receive exogenous income, which depends on location, age, and an idiosyncratic shock. The persistence of income residuals and the standard deviation of error of income residuals, ν and Σ, are set to be 0.75 and 0.45, respectively, following Fernandez and Wong (2014) and Chang and Kim (2006). Specifically, I follow Tauchen (1986) to approximate the continuous process with a discrete number of seven states.

Initial Wealth of Young Households

At the beginning of their lives, households receive initial non-housing wealth. I calibrate the wealth distribution of newborns using the distribution of wealth among 21-25-year-olds in the 2016 Survey of Consumer Finances (SCF). Households with negative wealth or no income are dropped from the sample. I parameterize the initial wealth distributions with a log-normal distribution with the mean and standard deviation calibrated to the data.

I translate the initial wealth distribution in the data to that in the model by scaling by the ratio of average household income among 21-25-year-olds in the model to the average household income of the same age group in different cities. In other words, I assume that the initial wealth distribution in city k follows a log-normal distribution with a city-specific mean $\mu_w \bar{w}_k w_0$ and a city-specific standard deviation $\sigma_w \bar{w}_k w_0$, where μ_w and σ_w are the mean and standard deviation of the wealth distribution of 21-25-year-old households adjusted by, $\bar{w}_k w_0$, the average household income of 21-25-year-old households in city k.

Taxes

The income tax is set to be $\tau_w = 0.2$, following Piketty and Saez (2007) and Amior and Halket (2014). The property tax is chosen to be $\tau = 0.01$, which is standard in the literature.

3.4.2 Land Shares in Production Functions

The two land shares and the relative productivity $\{\alpha, \rho, A\}$ in the production functions are estimated through regressions. Assuming that the minimum lot size does not bind for the standard owner-occupied houses, I regress the house prices of a standard single-family detached house and the rents of a standard two-bedroom apartment on the cross-sectional transaction-based land values provided by Albouy, Ehrlich, and Shin (2018) for 2010. As I allow for the variation in unobserved material prices, it is possible that material prices are correlated with land

Table 3.2: Predetermined Parameters

Parameter	Meaning	Value	Target or Source
σ	Risk aversion	2.5	Sommer, Sullivan, and Verbrugge (2013)
β	Discount factor	0.96	Sommer, Sullivan, and Verbrugge (2013)
ζ	Discount factor on ownership premium for seniors	0.9116	Fisher and Gervais (2011)
δ	Depreciation rate of owner-occupied units	2.5%	Sommer, Sullivan, and Verbrugge (2013)
δ_R	Management cost of rental units	33%	Goodman (2004)
k_b	Buying cost	2.5%	Yang (2009)
k_s	Selling cost	7%	Yang (2009)
r	Risk-free interest	0.04	Sommer, Sullivan, and Verbrugge (2013)
r_m	Mortgage interest	0.015	Amior and Halket (2014)
τ_w	Income Tax	0.2	Piketty and Saez (2007)
τ	Tax on residential properties	0.01	Sommer, Sullivan, and Verbrugge (2013)
ν	AR(1) Coefficient of income	0.75	Fernandez and Wong (2014)
Σ	Innovation of income process	0.45	Chang and Kim (2006)
μ_w	Mean of initial wealth distribution adj by income	3.4	Survey of Consumer Finance 2016
σ_w	std of initial wealth distribution adj by income	28.76	Survey of Consumer Finance 2016

prices. Therefore, I use the fraction of undevelopable land provided by Saiz (2010) as an instrumental variable for the land value. After backing out the land shares, the relative productivity A is estimated by combining Equation 3.13 and Equation 3.15.

Table 3.3 presents the regression results for the two land shares using data in 2010.[23] The first two columns present the results for OLS estimates, and the last two columns present the results using the fraction of undevelopable land as Instrumental Variable. The land share in the house production function is almost twice as high as the land share in the apartment production function. Consistent with the prediction of Proposition 3.3.1, when the underlying land price changes, house price grows more than apartment rent due to the intensive use of land.

3.4.3 Local Specific Parameters

For each metropolitan area k, the minimum lot size $\bar{L}_k$ and ownership premium θ_k are calibrated to the fraction of households living in houses of three age groups: young (20-40-year-olds), middle-aged (41-65-year-olds), and old (65-90-year-olds).

The basic idea of the identification is that when land and material price are fixed, an increase in the ownership premium θ_k lifts the homeownership rates of all age groups as households derive more utility from owned houses. On the contrary, an increase in minimum lot size disproportionally affects households with low income. Figure 3.6 illustrates the impact of an increase in minimum lot size on the price function of house size. A rise in the minimum lot size increases the house price for small homes while having no impact on the large ones, as smaller

[23] The cross-sectional indices of transaction-based land values provided by Albouy, Ehrlich, and Shin (2018) are only available between 2005 and 2010.

Table 3.3: Estimates of Land Shares

VARIABLES	OLS $log(P)$	OLS $log(R)$	IV $log(P)$	IV $log(R)$
$log(Land\_Value\_Albouy)$	0.377***	0.202***	0.539***	0.280***
	(0.0265)	(0.0158)	(0.0503)	(0.0309)
Constant	7.436***	3.759***	5.479***	2.813***
	(0.319)	(0.192)	(0.608)	(0.373)
Observations	182	182	182	182
R-squared	0.625	0.532	0.510	0.452

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

homes do not need that much land. As a result, an increase in minimum lot size will crowd out low-income households who prefer buying a small home to renting. Given the hump-shaped income profile, young households are more likely to quit owning when minimum lot size goes up. In other words, a rise in minimum lot size has an unbalanced impact on homeownership rates across age groups as it lowers the homeownership rate for young more than the middle-aged households.

More details about estimation can be found in Appendix B.5.

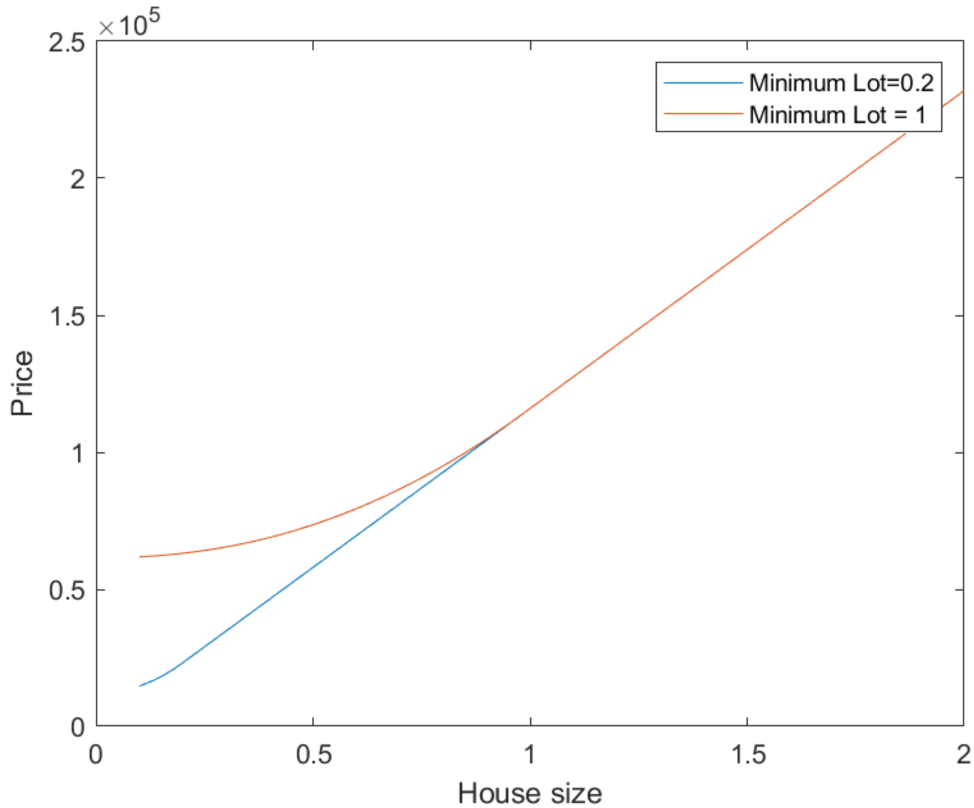
3.4.4 Data

This chapter uses the 5 percent sample of the 1980 U.S. Census and the American Community Survey (ACS) 2010 from the Integrated Public Use Microdata Series (IPUMS). These two datasets provide household-level data on household income, age of the household head, geographic location of residence, housing tenure choice (own or rent), self-valued house price for owner-occupied units, contract rent for rental units, and types of the dwelling (single-family detached house, attached house and multi-family buildings). The geographical unit of analysis is the metropolitan statistical areas (MSA) of residence. The dataset adopts a consistent definition of metropolitan areas from 1980 to 2010.

Young households are defined as households headed by 20-40-year-olds. Middle-age households are households headed by 40-60-year-olds, while old households are households with a 60-90-year-old head.

The age-profile of household income is constructed using the national-level data. Specifically, I group households according to the age of the household head and then put them into five-year bins from 20 to 85. Then for each bin, I calculate the average household income. Finally, for each metropolitan area, I calculate the number of households and average household

Figure 3.6: Price Function



income.[24] Data on survival probabilities comes from the life tables provided by the National Center for Health Statistics. For each age group, I calculate the probability of surviving for another 5 years.

Residential land supply is computed from the Land-Use and Land-Cover Data Sets of U.S. Geological Survey for 1982 and 2012.[25] The dataset classifies the conterminous land area into 19 categories. For each metropolitan area, I calculate the land used for residential constructions for 1982 and 2012. Then I use the growth rate of observed residential land from 1982 to 2012 to approximate the growth of residential land supply from 1980 to 2010.[26]

I start by comparing the growth of the number of households with the growth of residential land for the 181 cities from 1980 to 2010. For most of the cities, land growth does not catch up with population growth, as most of the points fall below the 45-degree line in Figure 3.7, which implies that land has become scarcer for most of the cities. Land values increase which leads to an increase in prices and rents in most of the cities. This is consistent with the observation that the average price and rent for the cities in my sample have both increased.

Figure 3.7: Population Growth and Residential Land Growth

[24]Data on house prices and apartment rents has been discussed in Section 3.2.1.
[25]https://pubs.er.usgs.gov/publication/ds240
[26]The Geological Survey is not available for 1980 or 2010.

3.5 Calibration Results

Table 3.4 shows the estimates of the relative productively A, ownership premium θ_k, and minimum lot size $\bar{L}_k$. The model closely matches the average and the standard deviation of the fraction of households living in houses among young, middle-aged, and old households across cities.

Table 3.4: Parameters and Moments

Estimated Parameters	Mean	Std		
Relative Productivity A	1.3			
Ownership Premium θ_k	1.7	0.43		
Minimum Lot Size $\bar{L}_k$	0.96	0.1		
Moments	Mean (data)	Std (data)	Mean (model)	Std (model)
Frac in Houses young	0.65	0.097	0.68	0.093
Frac in Houses middle-aged	0.83	0.075	0.86	0.085
Frac in Houses old	0.72	0.113	0.69	0.114

As an additional check of the calibration, I plot the model generated fractions of households living in houses (owners) against their data counterparts for the three age groups in Figure 3.8 - Figure 3.10. Each bubble represents one city and the bubble size corresponds to the city population in 1980. A 45-degree line is included to facilitate the comparison. The model does a great job in matching these moments for each individual city. The correlation between model-generated moments and the data exceeds 0.9 for all the age groups.

3.6 Quantitative Exercises

3.6.1 Baseline Results

In this section, I use the calibrated model to quantitatively investigate the impact of changing population, income, land supply, and credit constraint from 1980 to 2010 on the dispersion of house prices and rents. Note that by construction, the model can perfectly match the distribution of prices and rents in 1980. Therefore, I examine the performance of the model with its prediction of prices and rents for each individual city in my sample in 2010.

Specifically, I change population, income, and residential land supply to the 2010 level for each city in my sample. The survival probabilities and the age income profiles are also changed to the 2010 level. In addition, I lower the down payment requirement from 20% to 10% to capture the impact of mortgage innovation on the housing demand during the past

Figure 3.8: Fraction of Young Households living in Detached/Attached Houses 1980: Model and Data

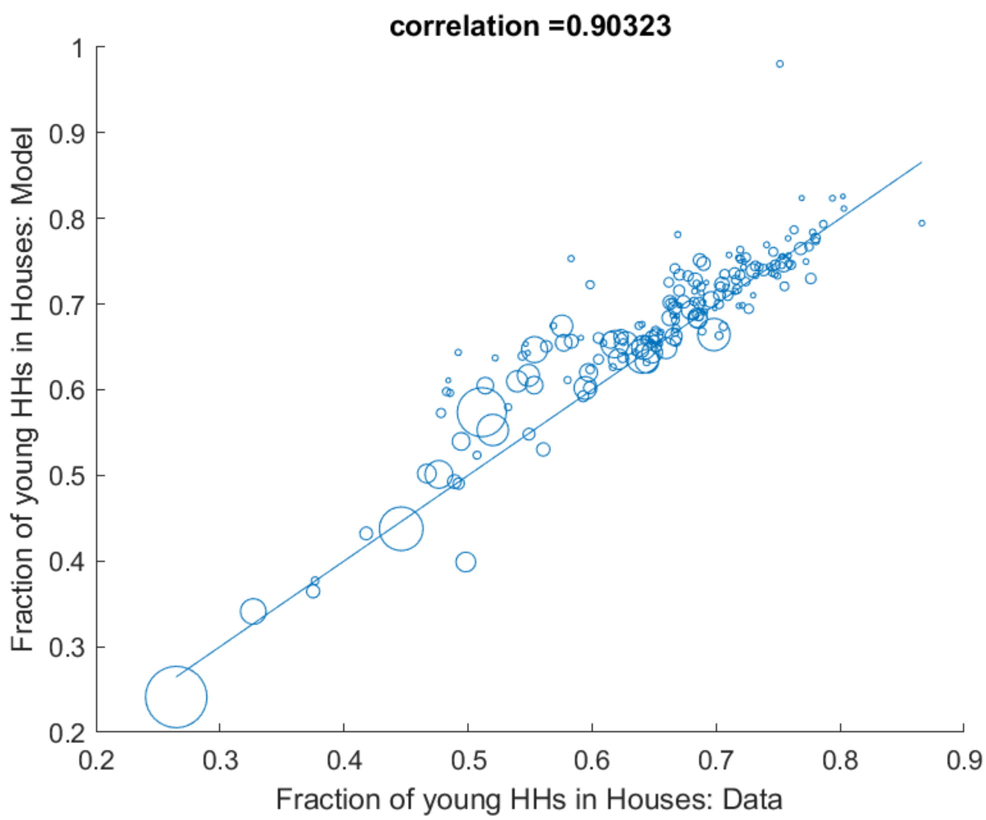


Figure 3.9: Fraction of Middle-aged Households living in Detached/Attached Houses 1980: Model and Data

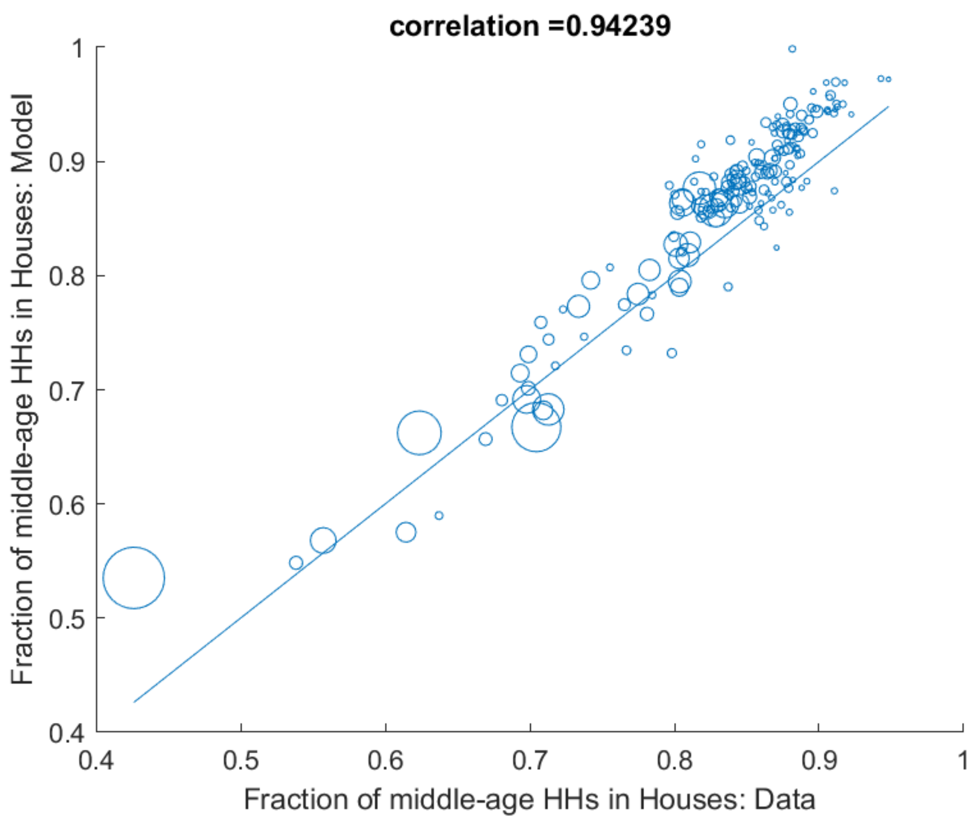
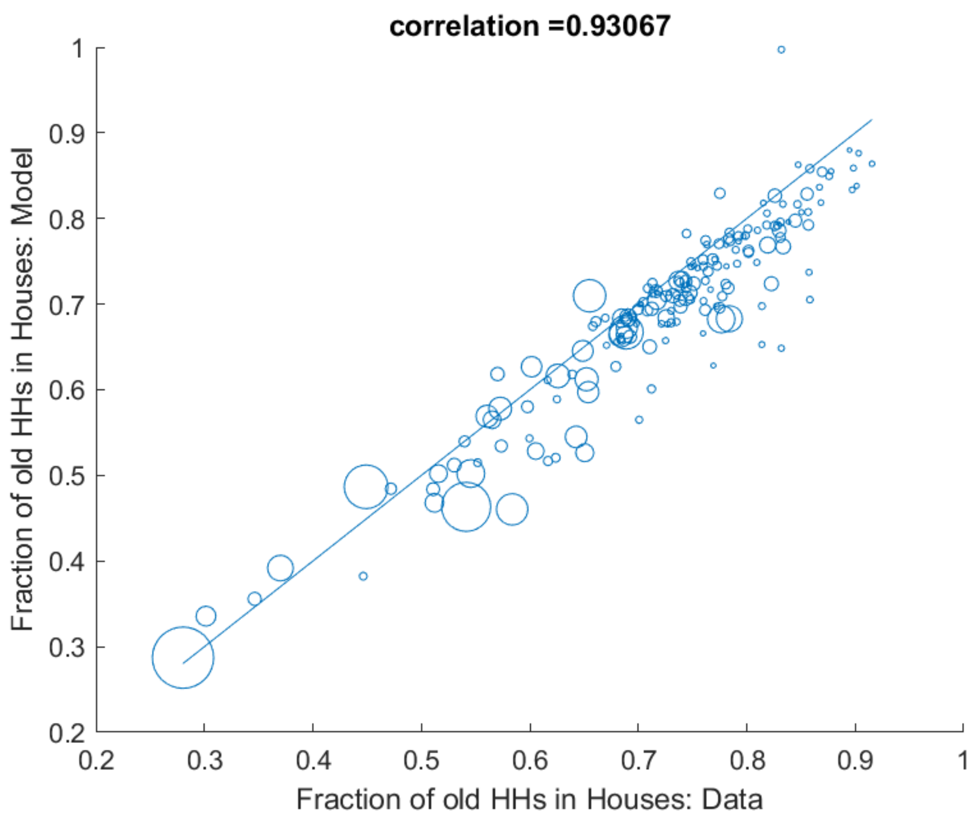


Figure 3.10: Fraction of Old Households living in Detached/Attached Houses 1980: Model and Data



several decades.[27] I simulate the model to calculate the equilibrium land price in 2010 for each individual city. Then I combine the land price with material price to calculate prices and rents for all cities. Note that in the simulation, material prices are fixed at the 1980 level.[28]

The results are summarized in Table 3.5. The model successfully generates a large increase in the price dispersion and a relatively moderate increase in the rent dispersion that match the data. Changes in population, income, land supply, and the down payment requirement can account for 82% of the increase in CV of house prices and 56% of the increase in CV of rents in the data. In addition, the model successfully generates an increase in the dispersion of price-rent ratios. As the last two columns in Table 3.5 show, the model can account for 90% of the increase in CV of price-rent ratios. These results suggest that most of the rise in the price-rent ratios in large cities can be attributed to fundamentals. This challenges the conventional view that the rise in the ratio of house prices to rents signals a housing bubble. In addition, it shows that the land use difference is important for understanding the relationship between these two shelter costs.

Table 3.5: Baseline Results

	mean(P)	CV(P)	mean(R)	CV(R)	mean(P/R)	CV(P/R)
1980	120900	0.33	6038	0.18	19.89	0.21
2010	146270	0.62	6264	0.27	22.15	0.31
Simulated 2010 10% down payment	129266	0.57	6085	0.23	20.47	0.30
Simulated 2010: 20% down payment	126095	0.52	6046	0.22	20.26	0.28

In addition to the aggregate moments, the model does a good job in predicting prices and rents for each individual city. Figure 3.11 and 3.12 plot the simulated prices and rents against the observed prices and rents in 2010, respectively. Each bubble represents a city, with the size of the bubble determined by the population size in 2010. A 45-degree line is included to facilitate the comparison. The model-generated and observed house prices line up well, with a correlation of 0.9. Similarly, the correlation between the model generated rent and the data is 0.78. The model slightly under-predicts the average level of prices and rents in 2010. Potential explanations are provided in the discussion.

As a robustness check, I also compare the model generated ratios between prices and rents with the data. Figure 3.13 demonstrates the comparison. The correlation between model-generated price-rent ratios and data is as high as 0.81. In other words, the model successfully captures the relationship between these two shelter costs across cities.

[27]Duca, Muellbauer, and Murphy (2012) show that the loan-to-value ratio for 1st-time home-buyers has increased to 90% in 2010.

[28]Material prices in 1980 is calculated in Section 3.4.3.

Figure 3.11: House Prices in 2010: Model vs Data

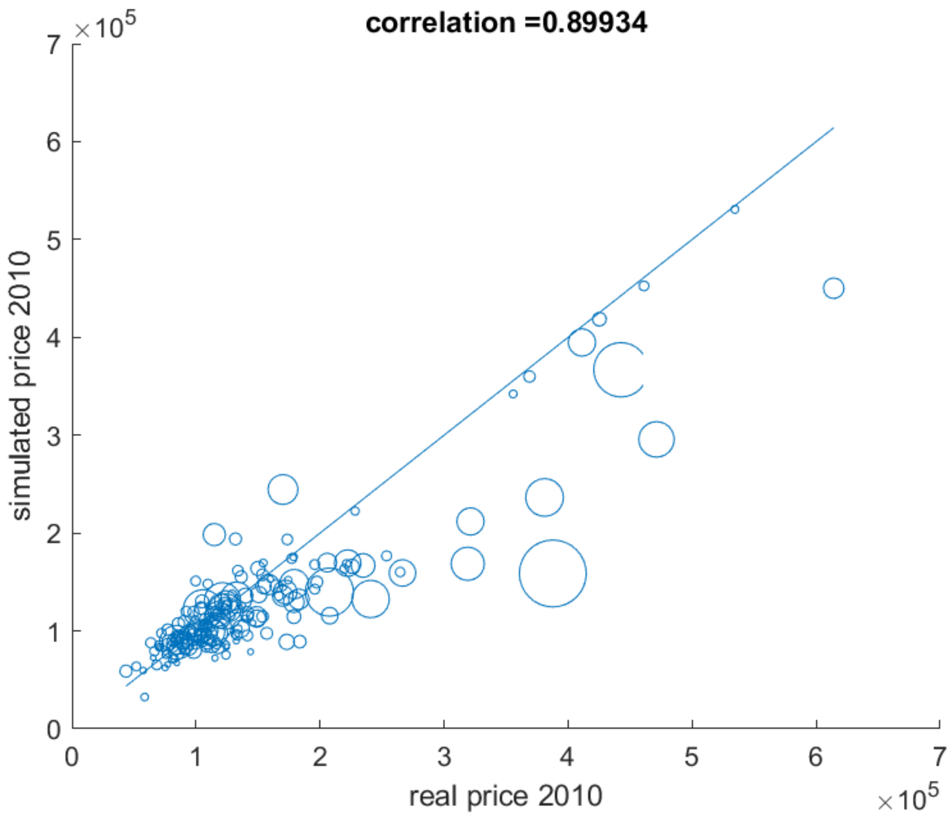


Figure 3.12: Rents in 2010: Model vs Data

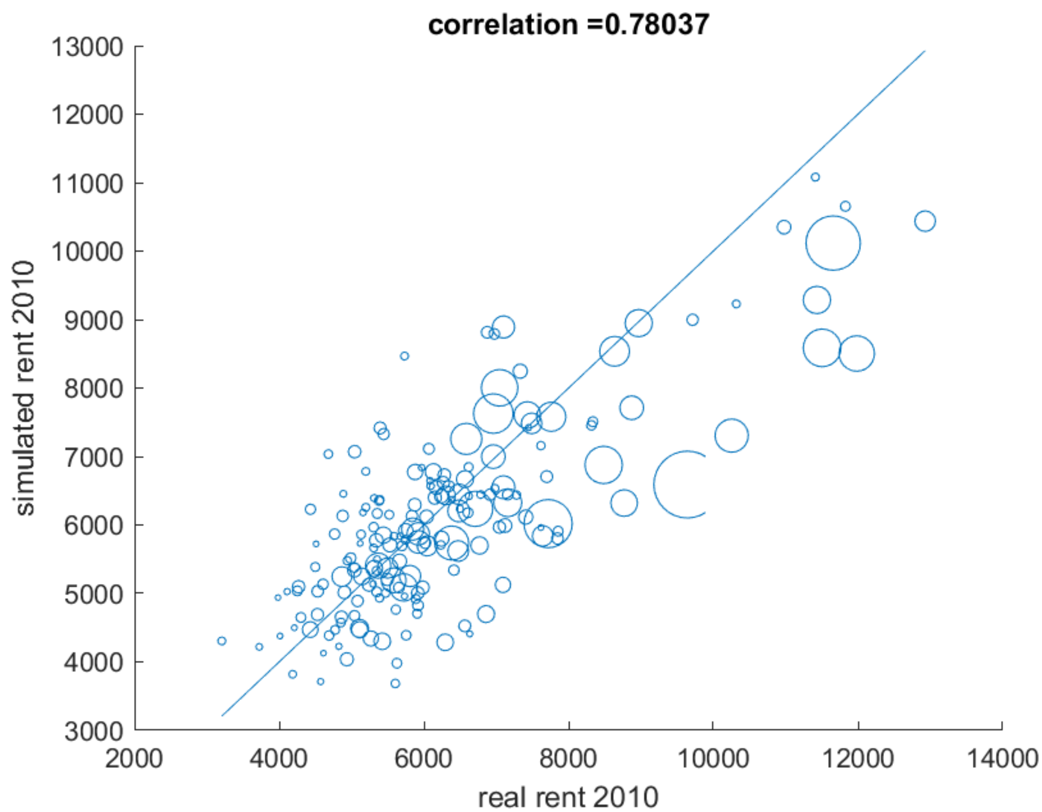
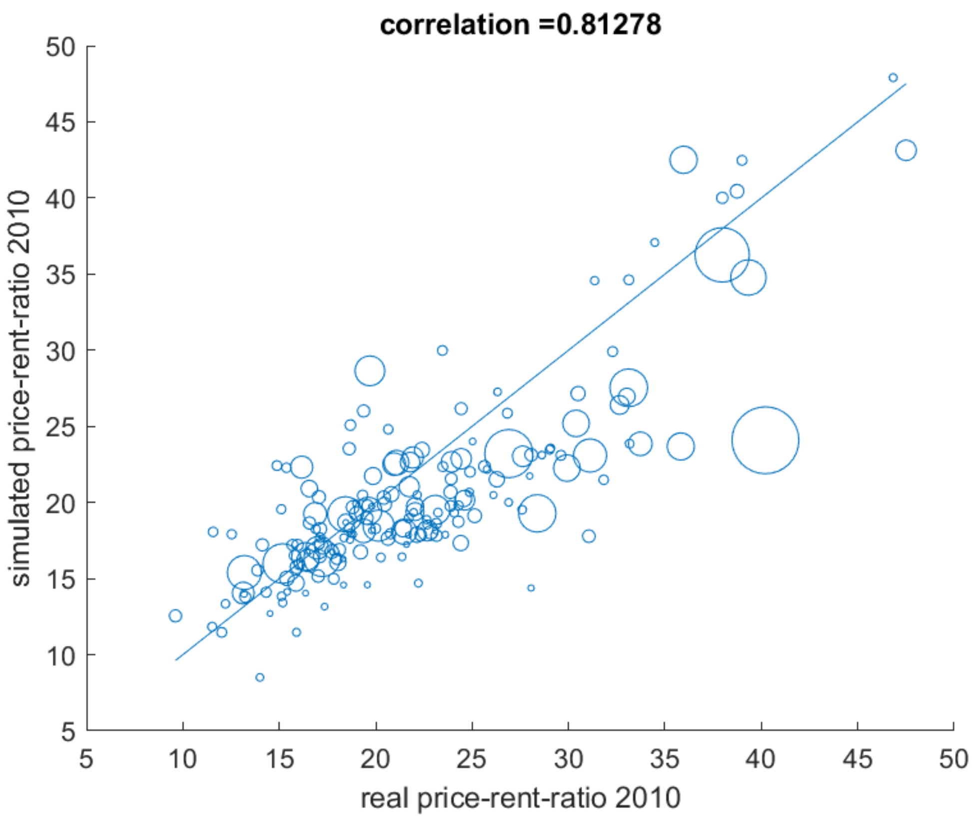


Figure 3.13: Price-rent ratios in 2010: Model vs Data



To sum up, the quantitative exercise shows that while changes in land values lead to changes in prices and rents, the difference in land intensities is key to understanding the divergence between these two shelter costs.

3.6.2 The Contribution of Lowering Down Payment Requirement

Motivated by the recent literature on the heterogeneous effect of mortgage-related policies, such as interest rate cuts across regions (See e.g. Hurst et al., 2016), I investigate the impact of lowering down payment requirement on house prices, rents, and housing price-rent ratios in different cities by changing population, income, and residential land supply to 2010 while holding the down payment requirement at its initial level, 20%. The results are presented at the bottom line in Table 3.5. Consistent with the previous literature, I find that the impact of lowering the down payment requirement on the underlying land prices is not identical across cities, as evidenced by the changes in the dispersion of prices, rents, and price-rent ratios. The difference between the last two rows in Table 3.5 reveals that the lowering down payment ratio accounts for 17% of the increase in house price dispersion, 11% of the increase in the rent dispersion, and 22% of the increase in the dispersion of housing price-rent ratios.

3.6.3 Discussion

Unchanged material price: In the quantitative exercise, I use material prices in 1980, which are estimated using data on prices and rents in the calibration. One concern is that increasing housing demand may push up wages of construction workers, which leads to an increase in the material cost, i.e., total construction cost net of land cost. In the Appendix B.6, I exploit cross-states variation and show that the relative wages of construction workers to other workers are not affected by the number of building permits, which is an approximation for housing demand. This is consistent with Gyourko and Saiz (2006), who document the significant differences in construction costs across U.S. housing markets and find that the construction costs do not change much in regards to housing permits.[29]

Underpredicted land price for big cities: The model noticeably under-predicts house prices, rents, and housing price-rent ratios in some big cities, which suggests the land prices in these cities are under-predicted. One potential explanation is the lack of redevelopment costs in my model, i.e., costs associated with changing the use of existing land. In reality, it is costly both in terms of time and money to tear down a house and build apartments on the same piece of land, and vice versa. Moreover, there are regulations, such as single-family zoning, that

[29] I use cross-state variation rather than cross-city variation due the data availability. For most of the cities, I do not enough observations for construction workers.

prevent the construction of multifamily buildings in certain areas. In other words, some land in expanding cities is not efficiently used, which may put upward pressure on the land price through the overuse of land.[30]

Figure 3.14 plots the ratio between the growth of land used for apartment constructions generated by the model and the growth of the amount of land used for high-density construction in the data from 1980 to 2010 against the ratio between model-generated price-rent ratio and price-rent ratio in data in 2010.[31] The negative slope suggests that the model underestimates price-rent-ratios in 2010, i.e., the underlying land prices, in some cities as it over-projects the growth of land used for high-rise buildings. For example, in the case of New York, the model under-predicts the housing price-rent ratio by 40% as it over-predicts the land growth for high-density construction by 200%. The living costs, prices and rents, in these big cities could thus be much lower if land use were flexibly adjusted.[32]

Figure 3.14: Price-Rent Ratio and Land Growth for Apartments: Model V.S. Data

[30]New York Times: Cities Start to Question an American Ideal: A House With a Yard on Every lot https://www.nytimes.com/interactive/2019/06/18/upshot/cities-across-america-question-single-family-zoning.html?mtrref=www.google.com&assetType=REGIWALL.

[31]I use price-rent ratio to approximate land values as it is strictly increasing in land values.

[32]I approximate the land use growth for apartments in the data by the land used for high-density construction.

Allowing for migration across cities For simplicity, in the quantitative exercise, I assume the city specific ownership premium is fixed at the 1980 level. One concern is that ownership premium in a city may change if internal migration is allowed. Intuitively, allowing people to move across cities ("sorting") will affect the population composition. Specifically, people who value owning more will move to cities with relatively low house prices, resulting in a decline in ownership premium in cities where house price is high and an increase in cities where house price is low from 1980 to 2010. In other words, when I apply the ownership premium estimated using data in 1980 to the 2010 simulation, I should over-predict the fraction of people living in houses in big cities with high house prices and under-predict it in small cities with low house prices, as people with high ownership premium move to these small cities. In other words, the prediction for demand of houses could be biased.

I apply the estimated ownership premium θ_k to simulate the demand of houses in 2010 and compare it to the data. Figure 3.15 - Figure 3.16 display the comparison for young and middle-aged households. Overall, the model's prediction of demand for houses lines up well with the data. Noticeably, instead of over-predicting demand for houses, the model slightly under-estimates the fraction of households living in houses for young and middle-age groups especially among the big cities.[33] In other words, allowing for migration across cities would not have a significant impact on my results.

One potential explanation is that the estimated ownership premium of a specific city combines two pieces of information: the structure difference between a standard owner-occupied house and a standard rental apartment in that city as well as how local residents treat this difference, i.e., the taste of local residents.[34] Changing population composition only affects the taste of local residents on the difference between owned-houses and rental apartments. If the variation of the estimated ownership premium θ_k mainly comes from the structure differences across cities, i.e., the relative quality of houses compared to apartments differ substantially across cities instead of local residents taste, allowing people to move across cities based on their taste towards owning would not have a significant impact on the results.

[33]The comparison for old households (65-90-year-olds) is ruled out from the comparison due to another dynamic aspect that is missing from the model. Assuming stationary equilibrium in 1980 and 2010 alleviates the burden from computing the transitional dynamics. Due to the high transaction cost of housing assets, households may not adjust the size of their houses frequently. Specifically, a large fraction of house owners aged from 65 to 90 bought their house in 1980 when the house prices were lower than today.

[34]The average structure difference is captured by the relative productivity A, while the local residents' taste is captured by the ownership premium θ. As these two terms are not separable in the model, they are identified using different sets of moments. The average relative productivity A is estimated using the relationship between prices and rents while the ownership premium is calibrated to the ownership rates.

Figure 3.15: Fraction of Young Households living in Detached/Attached Houses 2010: Model and Data

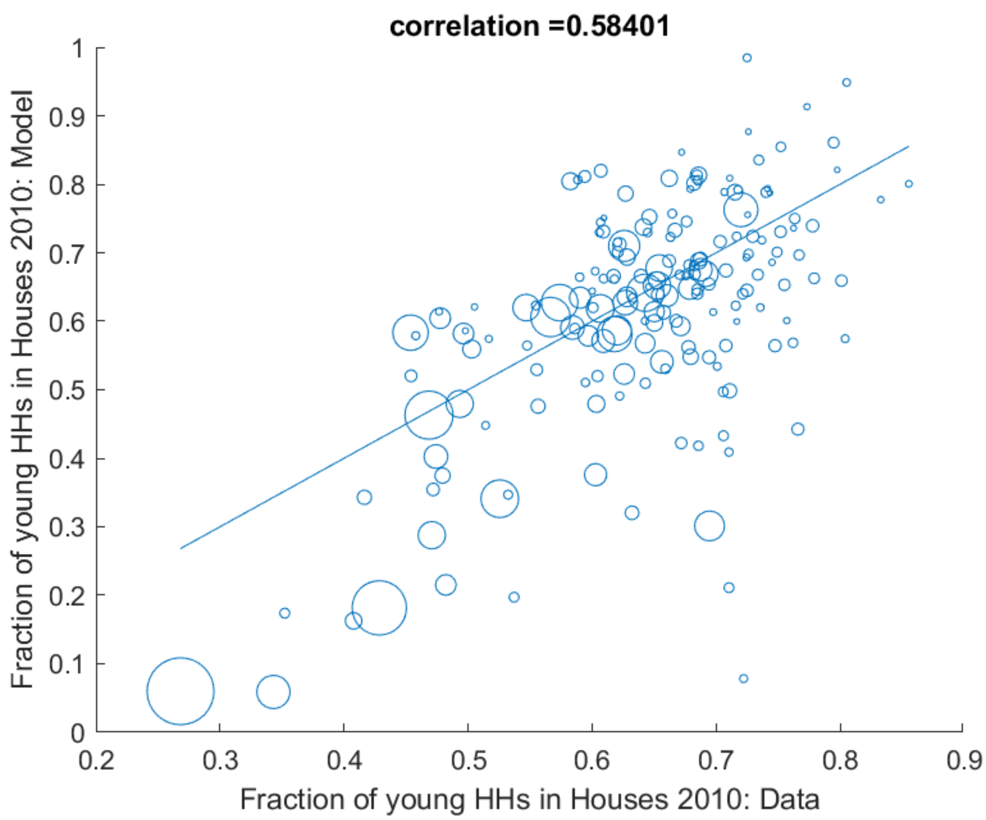
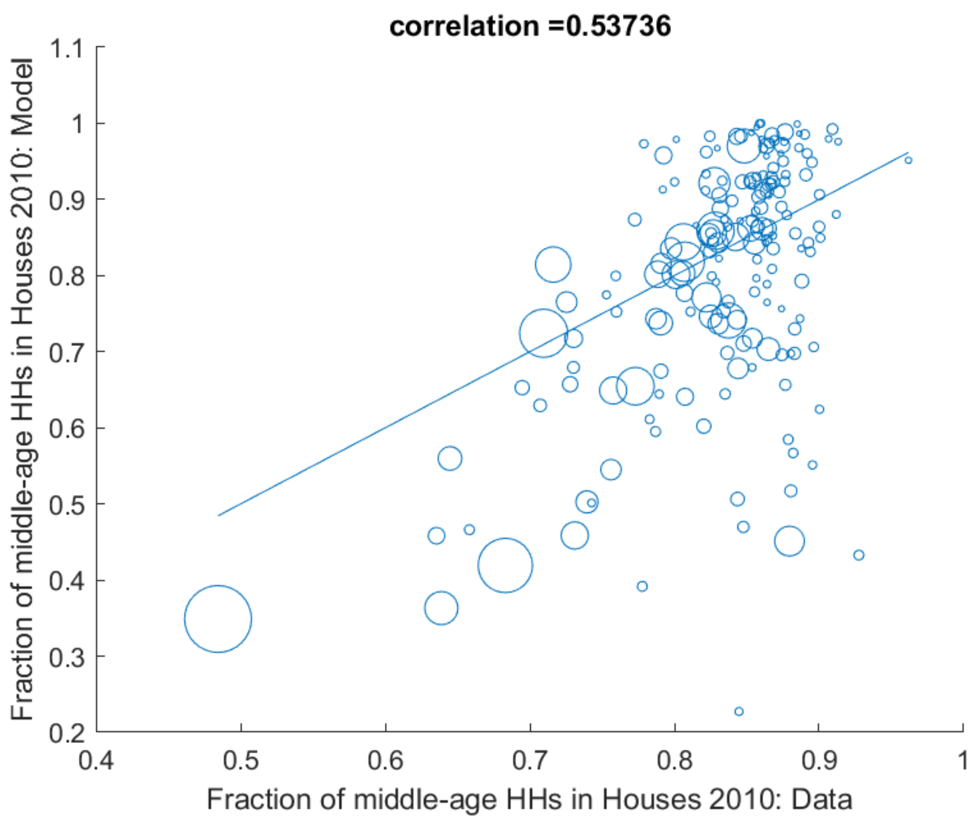


Figure 3.16: Fraction of Middle-aged Households living in Detached/Attached Houses 2010: Model and Data



3.7 Conclusion

This chapter provides new insights into the joint distributions of the two shelter costs, prices and rents, across metropolitan areas. It extends the housing tenure choice model to account for the difference in land use between owner-occupied houses and rental apartments. Houses are more land-intensive compared to apartments. When increasing land demand drives up land prices in expanding cities, the cost of building houses grows faster than that of apartments, leading to a larger increase in price than rent. The quantitative analysis shows the difference in land intensity is crucial in understanding the divergence between the distribution of these two shelter costs across cities.

Throughout my analysis, I maintain the assumption that both the apartment market and the house market are competitive such that the price (net present value of rents) equals the construction cost of houses (apartments), which simplifies the problem and allows me to highlight the mechanism where the house prices and apartment rents grow at different rates when land price changes due to different land intensities. Admittedly, perfect competition is a strong assumption and developers may have monopolistic powers. Nevertheless, even for monopolistic developers, the marginal cost of constructing houses and apartments is one of the most important factors in pricing their products.

This chapter makes the first attempt to study the relationship between house prices and apartment rents by exploring the difference in land intensities. The framework developed in this chapter rationalizes the different supply elasticity between owner-occupied and rental units especially in the long-run when land is reusable.

This chapter provides new insights on measuring the user cost of owner-occupied dwellings. The commonly-used measurements for shelter costs in the Consumer Price Index (CPI) are the Rent of Primary Residence and the Owners Equivalent Rent of primary residence (OER). This chapter illustrates that when land price changes, the construction costs of houses may deviate from that of apartments. As construction cost is one of the crucial components that determine user costs, using the change in rent of apartments, i.e. primary residence, to approximate the change in the user cost of owner-occupied houses could under-estimate (over-estimate) the growth of shelter cost for residents in areas where land price increases (declines).

The Owner Equivalent Rent is based on a hypothetical question that asks the owners about the rent they will receive if their dwellings were rented out. If the owners' answer is based on the rent of regular rental units, the OER could also be biased. Understandably, house prices are not perfect measurements of owners' user costs due to the high volatility and vulnerability to changes in macroeconomic conditions. Nevertheless, policymakers should be aware of the divergence between housing prices and rents that is sustained and can be accounted for by changes in land values and land use difference when evaluating living expenses for owners.

Bibliography

[1] David Albouy. Are big cities bad places to live? estimating quality of life across metropolitan areas. Technical report, National Bureau of Economic Research, 2008.

[2] David Albouy and Gabriel Ehrlich. Housing productivity and the social cost of land-use restrictions. *Journal of Urban Economics*, 107:101–120, 2018.

[3] David Albouy, Gabriel Ehrlich, and Minchul Shin. Metropolitan land values. *Review of Economics and Statistics*, 100(3):454–466, 2018.

[4] Michael Amior and Jonathan Halket. Do households use home-ownership to insure themselves? evidence across us cities. *Quantitative Economics*, 5(3):631–674, 2014.

[5] Sam Bucovetsky. On the effects of minimum-lot-size zoning. *Economics Letters*, 15(1-2):189–193, 1984.

[6] Matthew Chambers, Carlos Garriga, and Don E Schlagenhauf. Accounting for changes in the homeownership rate. *International Economic Review*, 50(3):677–726, 2009.

[7] Yongsung Chang and Sun-Bin Kim. From individual to aggregate labor supply: A quantitative analysis based on a heterogeneous agent macroeconomy. *International Economic Review*, 47(1):1–27, 2006.

[8] Morris A Davis and Jonathan Heathcote. The price and quantity of residential land in the united states. *Journal of Monetary Economics*, 54(8):2595–2620, 2007.

[9] Morris A Davis and Michael G Palumbo. The price of residential land in large us cities. *Journal of urban Economics*, 63(1):352–384, 2008.

[10] John V Duca, John Muellbauer, and Anthony Murphy. Shifting credit standards and the boom and bust in us house prices: Time series evidence from the past three decades. *Federal Reserve Bank of Dallas, unpublished manuscript*, 2012.

[11] Dennis Epple, Brett Gordon, and Holger Sieg. A new approach to estimating the production function for housing. *American Economic Review*, 100(3):905–24, 2010.

[12] Jack Y Favilukis and Stijn Van Nieuwerburgh. Out-of-town home buyers and city welfare. 2017.

[13] Raquel Fernández and Joyce Cheng Wong. Divorce risk, wages and working wives: A quantitative life-cycle analysis of female labour force participation. *The Economic Journal*, 124(576):319–358, 2014.

[14] Jonas DM Fisher and Martin Gervais. Why has home ownership fallen among the young? *International Economic Review*, 52(3):883–912, 2011.

[15] Martin Gervais. Housing taxation and capital accumulation. *Journal of Monetary Economics*, 49(7):1461–1489, 2002.

[16] Edward L Glaeser and Joseph Gyourko. Arbitrage in housing markets. Technical report, National Bureau of Economic Research, 2007.

[17] Jack Goodman. Determinants of operating costs of multifamily rental housing. *Journal of Housing Economics*, 13(3):226–244, 2004.

[18] Daniel L Greenwald and Adam Guren. Do credit conditions move house prices? 2019.

[19] Robert F Martin Joseph W Gruber et al. Does housing wealth make us less equal? the role of durable goods in the distribution of wealth. In *Econometric Society 2004 North American Summer Meetings*, number 15. Econometric Society, 2004.

[20] Joseph Gyourko, Christopher Mayer, and Todd Sinai. Superstar cities. *American Economic Journal: Economic Policy*, 5(4):167–99, 2013.

[21] Joseph Gyourko and Albert Saiz. Construction costs and the supply of housing structure. *Journal of regional Science*, 46(4):661–680, 2006.

[22] Joseph Gyourko, Albert Saiz, and Anita Summers. A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index. *Urban Studies*, 45(3):693–729, 2008.

[23] Erik Hurst, Benjamin J Keys, Amit Seru, and Joseph Vavra. Regional redistribution through the us mortgage market. *American Economic Review*, 106(10):2982–3028, 2016.

[24] Hans Isakson. Analysis of the effects of large lot zoning. *Journal of Real Estate Research*, 26(4):397–416, 2004.

[25] Will Larson, Jessica Shui, Stephen D Oliner, Morris A Davis, et al. The price of residential land for counties, zip codes, and census tracts in the united states. Technical report, American Enterprise Institute, 2019.

[26] Enrico Moretti. Real wage inequality. *American Economic Journal: Applied Economics*, 5(1):65–103, 2013.

[27] Andrii Parkhomenko. The rise of housing supply regulation in the us: Local causes and aggregate implications. In *Technical Report*. mimeo, 2018.

[28] Thomas Piketty and Emmanuel Saez. How progressive is the us federal tax system? a historical and international perspective. *Journal of Economic perspectives*, 21(1):3–24, 2007.

[29] Albert Saiz. Immigration and housing rents in american cities. *Journal of urban Economics*, 61(2):345–371, 2007.

[30] Albert Saiz. The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125(3):1253–1296, 2010.

[31] Kamila Sommer, Paul Sullivan, and Randal Verbrugge. The equilibrium effect of fundamentals on house prices and rents. *Journal of Monetary Economics*, 60(7):854–870, 2013.

[32] George Tauchen. Finite state markov-chain approximations to univariate and vector autoregressions. *Economics letters*, 20(2):177–181, 1986.

[33] Stijn Van Nieuwerburgh and Pierre-Olivier Weill. Why has house price dispersion gone up? *The Review of Economic Studies*, 77(4):1567–1606, 2010.

[34] Fang Yang. Consumption over the life cycle: How different is housing? *Review of Economic Dynamics*, 12(3):423–443, 2009.

Chapter 4

Demographics and the Housing Market

4.1 Introduction

The past several decades have seen significant growth in the real housing price in the United States. For example, the Freddie-Mac house price index has almost doubled while the median sale price of houses has more than doubled from 1975 to 2017.[1] Recent work has primarily focused on the role of mortgage credit expansions, including easy access to mortgage markets and lower mortgage interests, on the housing price growth (see e.g., Chambers, Garriga, and Schlagenhauf, 2009; Sommer, Sullivan, and Verbrugge, 2013; Greenwald, Guren, et al., 2019). In contrast, despite the substantial changes in demographics, the implications of these changes are less studied.

In this chapter, we contribute to the literature by examining to what extent the observed long-run growth of house prices can be accounted for by changes in fertility, life expectancy, urbanization, and international immigration. We focus on these four factors as they are more sustained and predictable, compared to innovations in the financial markets. Therefore, understanding their impact on the housing market is important for predicting future housing prices.

While there is a sizable literature investigating the effect of recent changes in the age composition of the population caused by declining fertility and rising life expectancy on housing prices (see e.g. Hiller and Lerbs, 2016), there is less discussion about how changes in other aspects, such as the population structure, caused by international immigration and the geographic distribution of households caused by urbanization. In this chapter, we develop a unified framework that allows us to evaluate the relative importance of these changes and to predict future housing prices.

We model the total housing demand by aggregating the age profile of housing demand

[1]Real numbers deflated by the Consumer Price Index (CPI).

over the age distribution of households. Three of the four factors of interest: declining fertility rate, international immigration, and rising life expectancy; affect the population structure, which refers to the population by age. Meanwhile, as documented in Section 4.2, rising life expectancy affects the age profile of housing demand as well. Specifically, rising life expectancy leads to a sizable increase in the housing demand among old households, as senior people keep their houses longer, knowing that they will live longer. Urbanization contributes to higher housing prices by relocating people across areas with different housing supply elasticity. From 1970 to 2010, the urbanization rate in the United States has increased from 73.6% to 80.61%. Urbanization moves people from rural areas, where the housing supply elasticity is high, to urban areas, where the housing supply elasticity is low. As a result, the increase in housing prices in the urban area is higher than the decrease in the rural area. In addition, urbanization results in a higher weight on the high housing price in urban areas in the calculation of the aggregate housing price. These two channels suggest that urbanization has a net positive effect on the aggregate housing price.

In Section 4.3, we develop a general equilibrium model of housing tenure choice to quantitatively study the contribution of these four factors. We model each area (urban or rural) in each state as an isolated economy/market. In each market, the aggregate housing demand is the product of the housing demand of different age groups and the age structure of the population, which is determined exogenously. Meanwhile, the aggregate population structure depends on the fertility rates, motility rates, and international immigration. We take advantage of the multiple-market setting and estimate the model using both cross-state and cross-time variations. We use the decennial data from the Census and the American Community Survey and focus on 32 states for which we have a balanced panel. The estimation spans the period from 1970 through 2010, decennially.

Using the estimated model, Section 4.4 quantitatively evaluates the contribution of these factors on the growth of housing prices from 1970 to 2010. The results confirm the importance of these factors discussed in this chapter. We find that declining fertility, rising life expectancy, urbanization, and international immigration account for roughly 41% of the increase in the aggregate housing price from 1970 to 2010. Among these factors, urbanization makes the largest contribution, accounting for 31-40% of the model generated housing price growth. International immigrants and their future generations since 1970 contribute to 24-29% of the model generated housing price growth. Improvement in survival probability accounts for 15-18% of the housing price growth that can be explained by the model. Specifically, rising life expectancy has a significant impact both on the age profile of housing demand and on the population structure.

In Section 4.4.3, we apply the projected population structure, which combines predictions on fertility rates, survival probabilities, and international immigration, survival probabilities, and urbanization rates provided by the Census and the United Nations to predict aggregate housing prices in the next 40 years. We find that housing prices will grow by 5-25% from 2010 to 2050. The growth rates vary with urbanization rates and the levels of immigration.

This chapter is related to the literature that focuses on the impact of changes in the age composition of the population on the real estate market. The seminal work by Mankiw and Weil (1989) focuses on one channel through which population aging caused by declining fertility and rising life expectancy can have a negative impact on the aggregate housing price. The intuition is straightforward. As the middle-aged households have the highest housing consumption, compared to young and elderly households, i.e. households have a hump-shaped housing consumption profile, a decline in the share of middle-aged households would lead to a substantial decline, about 47%, in the housing price in the next 20 years. The following studies by Hamilton (1991) and Holland (1991) question the magnitude of the prediction but agree on the trend. Thirty years after the publication of Mankiw and Weil (1989), house prices have risen rather than fallen. This chapter proposes several competing factors that may mask the negative impact of the declining share of prime-age buyers on the aggregate housing price and makes long-run predictions in housing prices. Specifically, we highlight three factors, i.e. rising life expectancy, international immigration, and urbanization, that have contributed to the housing price growth since 1970. We find that rising life expectancy, which shifts up the housing demand for senior households, combined with the growing number of seniors have increased aggregate housing demand significantly. In addition, international immigrants and their future generations since 1970 make up 18% of the total population, which contributes to the steady growth in the aggregate housing demand. Meanwhile, urbanization that relocates people to cities with lower housing supply elasticity has further boosted the aggregate housing price.

Our finding that housing prices will keep growing in the next few decades is consistent with Green and Lee (2016) who argue that, with rising education and income levels driving up the housing demand per household, the recent massive demographic shift caused by the retirement of the baby boomer generations will not result in a housing crisis. Instead of focusing on the role played by household income, this chapter highlights the contribution of survival probability on household's willingness to pay, and examine the effects of urbanization, and international immigration on aggregate housing demand and housing supply.

This chapter simplifies the general equilibrium framework, which implies that rising life expectancy may lead to delaying home selling (Chambers, Garriga, and Schlagenhauf, 2009, and Anagnostopoulos, Atesagaoglu, and Carceles-Poveda, 2013). We quantitatively evaluate this channel and consider more factors related to changes in demographics that affect the aggregate housing market. The empirical facts documented in Section 4.2 confirm the importance of factors that we propose. The model allows us to quantitatively evaluate the contribution of each factor on the aggregate housing price growth and predict future housing prices.

Also, this chapter contributes to the studies on urbanization (see e.g. Glaeser, Gyourko, and Saks, 2006). While most existing literature studies the impact of urbanization on housing prices in developing countries, such as China (see e.g. Chen, Guo, and Wu, 2011), we focus on the housing markets in the United States. Specifically, we estimate the housing supply elasticity in urban and rural areas separately. We find that housing supply is less responsive to prices in urban areas compared to rural areas. The supply elasticity in rural areas is about three times as high as urban areas.

4.2 Empirical Evidence

In this section, we present empirical evidence showing the importance of factors that may have positive impacts on the housing market: increasing survival probabilities, international immigration, and urbanization.

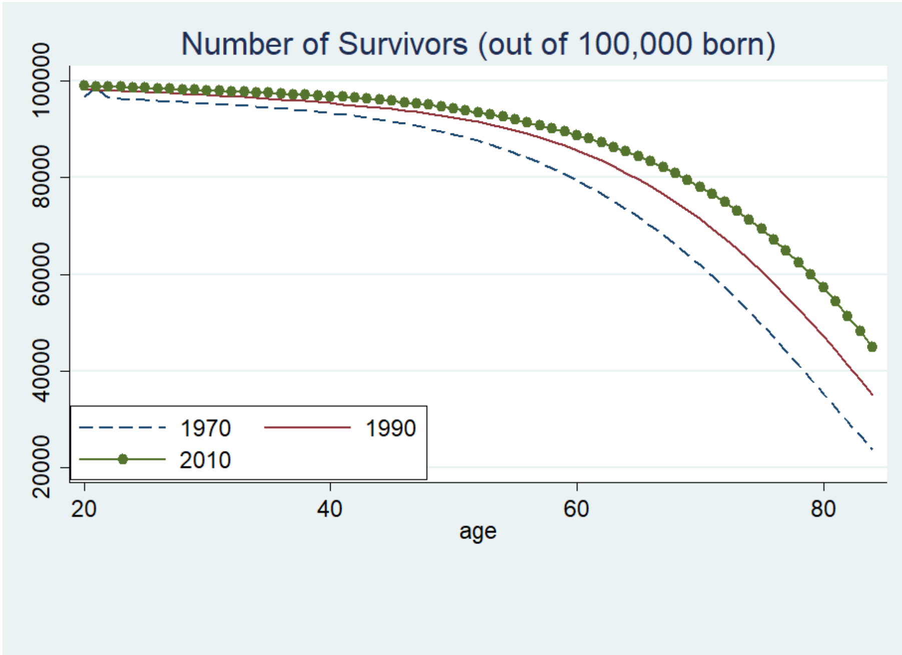
4.2.1 Rising Life Expectancy

While previous studies focus on the contribution of rising life expectancy on the housing demand through its impact on the population structure, we find that it has a direct impact on age profile of housing consumption. We start by documenting the rising life expectancy in the past several decades accompanied by an increasing homeownership rate for senior households in the United States. Panel (a) and (b) in Figure 4.1 plot the survival rate by age, and the homeownership profile, respectively. From 1970 to 2010, we see a significant improvement in the survival probability, especially for the senior households.[2] For example, the probability of living for more than 70 years has increased from 60% to 80%. Meanwhile, the ownership rate for senior households has increased substantially, indicating that senior households keep their homes longer. For instance, homeownership starts to decline at age 55 in 1970, when the average life expectancy is around 71 years. The turning point in ownership profiles comes after age 75 in 2010 as the life expectancy increases to 79 years.[3]

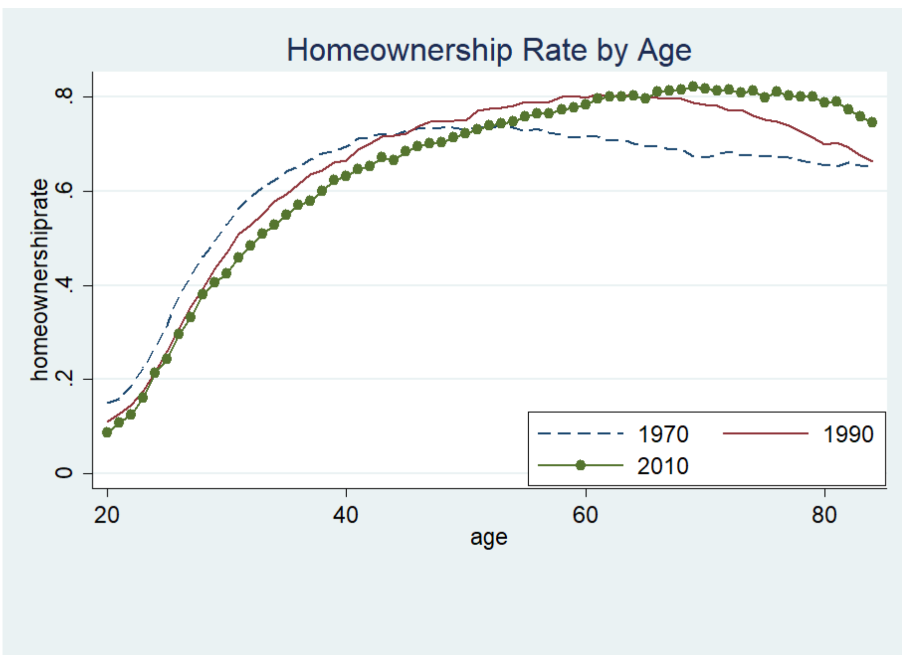
[2]To distinguish the time effect from the cohort effect, we compare the ownership profile for different cohorts in Appendix C.1. Similarly, we find that households born between 1940-1950 have higher ownership rate in their 70s compared to households born between 1920 and 1930.

[3]We approximate housing demand with ownership rate as the price has increased substantially since 1970, while the rent has barely changed (see e.g., Sommer, Sullivan, and Verbrugge, 2013). In addition, there are studies arguing that the owner-occupied market and the rental market are somewhat segmented (see e.g., Greenwald, Guren, et al., 2019)

Figure 4.1: Survival Rates and Homeownership: 1970-2010



(a) Survival Probability



(b) Onwership Profile

Source: Survival rates come from the life table from U.S. Social Security Administration.
Homeownership rates come from authors calculation using data from Census and American Community Survey.

4.2.2 International Immigration

In order to isolate the contribution of international immigration and their future generations on the population, we simulate the population structure in a hypothetical scenario where there were no international immigrants in the United States since 1970. Specifically, we compute the population by age from 1971 to 2010 in this hypothetical scenario by applying the realized fertility rates and mortality rates to the population structure in 1970.[4] Taking the difference between the actual population from the Census and the population structure obtained through our simulation, we find that immigrants and their future generations since 1970 account for 18% of the total adult population in 2010.

4.2.3 Urbanization

In the Census and American Community Survey, the variable that captures the urbanization status of households is only available before 1990. So we use a different variable that captures the metropolitan area status of households. Households living in metropolitan areas are defined as urban households.[5]

From 1970 to 2010, the fraction of households living in metropolitan areas increased from 63% to 76%. Figure 4.2 shows that, while the total population in metropolitan areas has increased steadily, the population size in non-metropolitan areas has been fairly stable. For non-metropolitan areas, the negative impact of urbanization seems to be compensated by the growth of the aggregate population.

Figure 4.3 plots the average housing price in metropolitan and non-metropolitan areas separately. It reveals that, on average, housing prices grew faster in metropolitan areas than in non-metropolitan areas. From 1970 to 2010, on average, the housing price in metropolitan areas increased by about 120%, while non-metropolitan areas have only seen an 80% growth. In addition, the price growth for rural areas was mild before the credit expansions started from the late 1990s, when the price growth for urban areas was still significant.

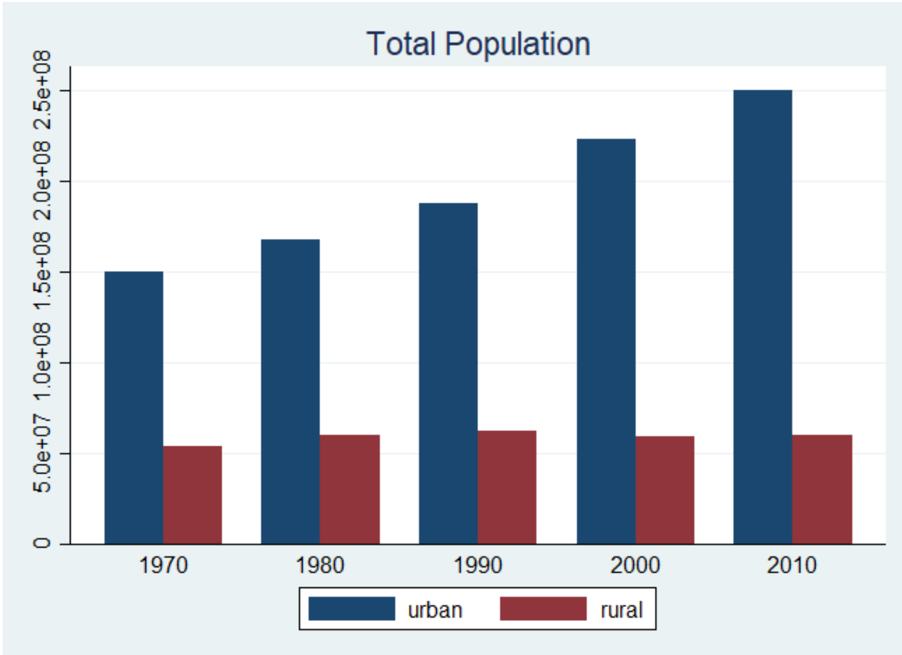
4.3 Model and Estimation

This section presents a parsimonious general equilibrium model to characterize both the demand and supply of housing. Despite its simplicity, our model is rich enough to capture the impact of the factors of interest on housing prices. We specify the model in Section 4.3.1 and describe the estimation strategy in Section 4.3.2.

[4] Appendix C.2 shows the detailed steps.

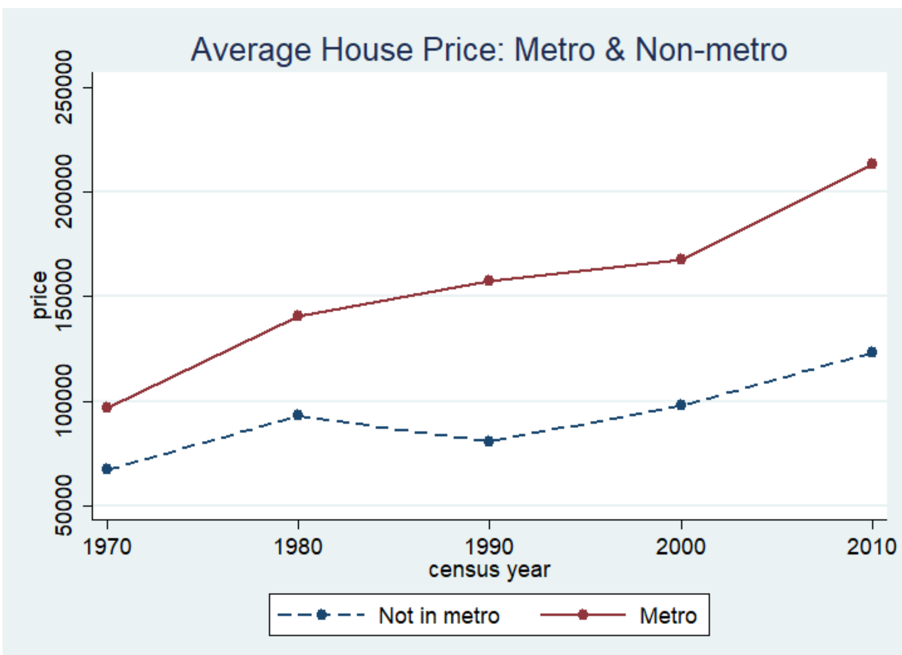
[5] We use metropolitan area and urban area interchangeably.

Figure 4.2: Urbanization



Source: Authors calculation using data from Census and American Community Survey.

Figure 4.3: Housing Price: Urban and Rural Areas



Average housing price in 1999 U.S. dollars.
Source: authors' calculation using data from Census and American Community Survey.

4.3.1 The Model

Overview

In our model, a market is defined by which state ($j \in J$) it is in, whether it is in urban or rural areas (u), and which year it is in (t). Therefore, for each year we have $2J$ segregated markets.[6] There is no interaction across markets. The supply side of the housing markets is fairly standard. For each market, the housing supply is a deterministic function of the housing price. The aggregate housing demand in each market can be computed by combining the housing demand of different age groups and the age structure of the population. While Mankiw and Weil (1989) focus on the changes in the age structure by keeping the housing demand for each age group fixed, our framework captures the possibility that longer life expectancy can lead to an increase in the housing demand for senior households. In other words, when people live longer, they keep their houses longer.

To capture the impact of life expectancy, or equivalently survival probabilities, on the housing demand of individual households, we consider a model where the average housing demand for households from age group a, which we denote F_a, depends not only on the household income w_a and the housing price P, but also on the survival probability S_a:

$$F_a = F_a(w_a, S_a, P). \tag{4.1}$$

Then, the aggregate demand of all households is described as follows:

$$D = \sum_a F_a N_a, \tag{4.2}$$

where N_a is the total number of households in age group a.

Demand

Consider the housing demand for households in state j and area u (urban area: $u = 1$; rural area: $u = 0$) at time t. For this market, we assume that households from the same age group a have identical income $w_{a,j,u,t}$ and survival probability $S_{a,t}$.[7] Households choose whether to buy a house. Houses are identical, and each household can only buy one house.[8] For a representative household of age a, the probability of owning a housing unit, $\lambda_{a,j,u,t}$, is given by:

[6] This setting allows us to estimate the model through cross-state and cross-time variation.

[7] The assumption that the survival probability $S_{a,t}$ does not vary by state and area is made because of data availability issues.

[8] This assumption is well justified by the data. According to the American Community Survey, the fraction of owners who have a second residence in 2005 is only 3%.

$$\lambda_{a,j,u,t} = \frac{exp(\beta_0 + \beta_1 log(w_{a,j,u,t}) + \beta_2 S_{a,t} - \beta_3 log(P_{j,u,t}) + \gamma_a + E_t + \eta_{a,j,u,t})}{1 + exp(\beta_0 + \beta_1 log(w_{a,j,u,t}) + \beta_2 S_{a,t} - \beta_3 log(P_{j,u,t}) + \gamma_a + E_t + \eta_{a,j,u,t})}. \tag{4.3}$$

In this equation, $P_{j,u,t}$ is the housing price in state j, area u at time t. γ_a is the age fixed effect. E_t is the time fixed effect, which may capture the macroeconomic environment that affects the housing demand, such as mortgage interests and access to mortgages. $\eta_{a,j,u,t}$ is an idiosyncratic shock.

Equation 4.3 shows the extensive margin of the housing demand, which is consistent with a standard logit model (see, e.g. Gyourko and Linneman, 1996).[9] As households are ex ante identical, $\lambda_{a,j,u,t}$ also corresponds to the homeownership rate of age group a.

Then, the aggregate housing demand in state j and area u at time t is given by:

$$\begin{aligned} D_{j,u,t} &= \sum_a F_{a,j,u,t} N_{a,j,u,t} \\ &= \sum_a \lambda_{a,j,u,t} N_{a,j,u,t}, \end{aligned} \tag{4.4}$$

where $N_{a,j,u,t}$ is the number of households of age group a living in state j and area u at time t.

Our specification of the demand side can be considered as a reduced-form approximation of fully specified life-cycle models used in the literature (see e.g. Chambers, Garriga, and Schlagenhauf, 2009). We choose this parsimonious approach mainly for two reasons. First, the economy in our model consists of many segregated markets which allows us to explore both the cross-sectional variation and cross-time variation to estimate the model. Hence, a relatively simpler demand side helps keep the computation and estimation of our model more computationally tractable. Second, this parsimonious specification allows for a more transparent and direct investigation of the impact of key demographic variables on housing demand, and consequently equilibrium housing prices. For example, to examine the importance of survival probability $S_{a,t}$ on the housing demand profile, we can estimate Equation 4.3 using cross-state and cross-time variations in $\lambda_{a,j,u,t}$ and examine the economic and statistical significance of the relevant coefficients β_2.

[9] We abstract away from the intensive margin of the housing demand by assuming that all owner-occupied units are identical. This assumption is made due to a lack of appropriate measures of housing quantity and quality in the Census and the American Community Survey.

Supply

For state j and area u at time t, housing supply is determined through the following fixed effect model.

$$log(P_{j,u,t}) = \alpha_{u,0} + \alpha_{u,1} log(D_{j,u,t}) + \theta_{j,u} + \epsilon_{j,u,t}, \tag{4.5}$$

where $\alpha_{u,1}$ is the price elasticity of supply that may differ between rural and urban areas. $\theta_{j,u}$ is an area fixed effect which may capture permanent housing market conditions, such as regulation or construction costs, for state j and area u. $\epsilon_{j,u,t}$ is an idiosyncratic shock.

4.3.2 Data and Estimation

The primitives of our model are (1) exogenous variables including household income $w_{a,j,u,t}$, survival probability $S_{a,t}$, and number of household $N_{a,j,u,t}$ and (2) the parameters in the demand equations (Equation 4.3) and the supply equation (Equation 4.5). The exogenous variables can be directly computed from the data. To estimate the demand and supply parameters, we first compute the dependent and independent variables in the corresponding equations, then estimate the parameters using both cross-state and cross-time variations.

Data

We use data from the 1970, 1980, and 1990 waves of the Census and the 2000 and 2010 waves of the American Community Survey (ACS). We extract these data from the Integrated Public Use Microdata Series (IPUMS). In this chapter, we focus on the 32 states for which we can construct a balanced panel. These 32 states consist of more than 85% of the total population in the country.[10]

We extract information on the households' geographic location (state of residence and metropolitan status), the age of household head, household income, housing tenure choice, and the value of houses for owners. All dollar values are converted into 1999 dollars using the Consumer Price Index (CPI).

For each state j and area u at time t, we construct the housing price $P_{j,u,t}$ by aggregating the house value reported by the owners in this specific market. We group households based on the age of the household head and put them into five-year bins. For example, the first bin consists of households of which the age of the household head is between 20 and 24. The last bin consists of households of which the age of the household head is between 80 and 84. For

[10]The complete list of these 32 states is provided in Appendix C.3. The Census and ACS adopt either 1% or 5% sample. Therefore some states do not have observations for all age groups in all years for both urban and rural areas.

each age group, we calculate the average household income $w_{a,j,u,t}$, as well as the ownership rate $\lambda_{a,j,u,t}$.

The survival probabilities $S_{a,t}$ come from the life table provided by the U.S. Social Security Administration. For each year, they provide the number of survivors out of 100,000 born by age. We use it to calculate the probability of surviving for another five years and use this probability to measure $S_{a,t}$.

Population Structure across Markets

The age structure of households in each market is calculated through the following steps. First, we compute the total number of households in each age group a at time t for the whole country, $\bar{N}_{a,t}$. Second, for each age group, we calculate the fraction of households living in metropolitan areas, $\theta_{a,t}$. Then for each age group, we compute the share of the population in each state j for metropolitan and non-metropolitan areas separately, $Z_{a,u,t}$. So the number of households from age group a, residing in state j, area u, at time t is given by:

$$N_{a,j,u,t} = \bar{N}_{a,t}((1-u)(1-\theta_{a,t}) + u\theta_{a,t})Z_{a,u,t} \tag{4.6}$$

This specification allows us to conduct counterfactual exercises to isolate the contribution of rising life expectancy, international immigration, and urbanization.

Estimation

The estimation of housing demand is based on Equation 4.3. Specifically, we regress the homeownership rate of different age groups on the average household income, the probability of survival for another five years, the average local house price, and age and time dummies. Homeownership rates are defined as the number of households living in owner-occupied dwellings divided by the total number of households. Households are identified by the age of the household head.

The supply estimation is based on Equation 4.5. The standard OLS regression could be subject to the endogeneity concern. For instance, unobserved amenities that affect housing demand could also be capitalized in the housing price. In other words, households in areas with better amenities are more likely to buy a house as they would like to settle down. We adopt an instrumental variable approach in the spirit of the Bartik demand shifter. Specifically, we extract exogenous variation in housing demand in a state by interacting the housing demand of different age groups in other states with the age structure in the state for rural and urban areas separately.

Formally, we define the Bartik Instrument $ID_{j,t}$ as:

$$\overline{X_{a,j,u,t}} = \frac{\sum_{k \neq j} \lambda_{a,k,u,t}}{N-1}$$
$$ID_{j,u,t} = \sum_{a} \overline{X_{a,j,u,t}} * N_{a,j,u,t}, \tag{4.7}$$

where $\overline{X_{a,j,t}}$ is the average housing demand of households of age a living in area u outside of state j. Aggregating the predicted housing demand by the number of households across different age groups, we obtain the Bartik Instrument Variable that is assumed to be exogenous to local housing market conditions, such as unobserved factors related to the local housing price.

The exclusion restriction assumes the age structure in each market is independent of factors that affect the housing demand of individual households. The intuition of this instrument is straightforward. Assuming that middle-aged households have higher housing demand compared to young households, a market with more young households should have lower aggregate housing demand.

Estimation Results

The estimation results for the demand equations are summarized in Table 4.1. Consistent with the previous literature, the coefficient on prices is negative, and the coefficient on the household income is positive, which suggests that ownership rates of all age groups decline with local house prices and increase with household income. The corresponding coefficients are all significant at a 1% level. Most importantly, we see that survival probability has statistically significant (at a 1% level) positive effects on the extensive margin of housing demand. The model accounts for 89% of the cross-state and cross-time variation in homeownership rates.

For comparison, the second column presents the estimation results without the year dummies. As expected, part of time trends is absorbed by the coefficient on the survival probabilities as the survival probabilities keep growing over time for all age groups. Even without year dummies, the model accounts for 87% of the cross-state and cross-time variation in homeownership rates. The sign and the magnitude of the year dummies are consistent with findings in the previous literature. For instance, year dummies in the ownership regression are positive for all post-1970 years, suggesting the changes in macroeconomic conditions shift up housing demand since 1970. Possible explanations include the high inflation rate that lowers the user cost of owners during the period around 1980 (Poterba, 1984) and the mortgage expansions since 2000 (see e.g. Mian and Sufi, 2009). Note that the time dummy in 1990 is the lowest compared to other years, which suggests that after the end of the high inflation period in 1983 and before the credit expansions in the early 2000s, owning was less desirable.

Table 4.1: Demand Estimation

	Ownership	Ownership
log Price	-0.761***	-0.603***
	(0.0242)	(0.0177)
log Household Income	0.518***	0.341***
	(0.0444)	(0.0351)
Conditional Survival Prob	6.054***	8.528 ***
	(0.553)	(0.384)
1980.year	0.351***	
	(0.0200)	
1990.year	0.146***	
	(0.0192)	
2000.year	0.239***	
	(0.0208)	
2010.year	0.274***	
	(0.0286)	
Age Dummies	Yes	Yes
Constant	-3.923***	-6.240*******
	(0.484)	(0.350)
Observations	4,160	4,160
R-squared	0.880	0.867

Robust standard errors in parentheses
***p<0.01, ** p<0.05, * p<0.1

The estimation results for supply elasticity are summarized in Table 4.2. As we expected, the supply elasticity for rural areas is three times as large as that for urban areas. A one percentage increase in housing price leads to a 2.1% (1/0.48) increase in supply in urban areas and a 10% (1/0.1) increase in supply in rural areas. To compare our estimates with previous studies, we compute the aggregate housing supply elasticity using the urbanization rate in 1990 as weights. We find that the aggregate supply elasticity is around 2.6, which is close to Topel and Rosen (1988), who suggest that the long-run housing supply elasticity is around 3.

4.4 Model Fit and Counterfactual Analysis

4.4.1 Model Fit

We use the estimated model to answer the key question: How much of the housing price growth can be explained by changes in the four factors, i.e. lower fertility, long life expectancy, urban-

Table 4.2: Supply Estimation

	Price Elasticity IV	Price Elasticity OLS
Metro	0.480***	0.488***
	(0.032)	(0.0542)
Non_Metro	0.106*	0.098
	(0.0604)	(0.078)
State Dummies	Yes	Yes

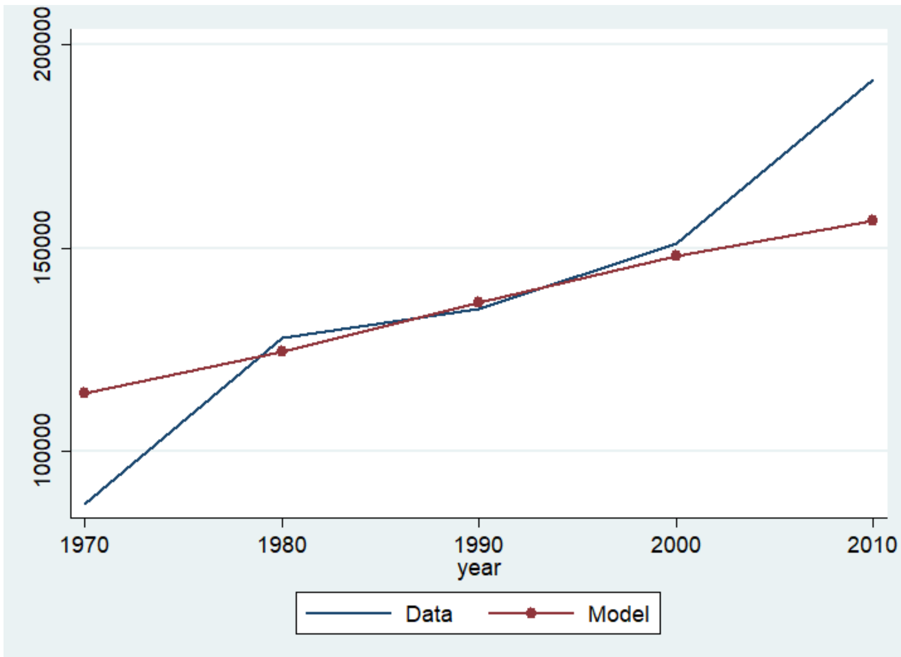
ization, and international immigration, from 1970 to 2010? To answer this question, we feed into the model the survival probabilities, urbanization rates, and the realized age structure of population, which combines changes in fertility rates, survival probabilities, and international immigration from 1970 to 2010. As we focus on the contribution of demographic-related factors instead of changes in the macroeconomic environment, we consider the case where the macroeconomic environment has not changed since 1970 by applying the 1970 time dummy variable (normalized to 0) in the demand equations. Meanwhile, we also adjust the income profile, but its impact is tiny as the income profile has not changed much. Appendix C.4 shows the contribution of the change in income profile.

A market in our model is defined at the state and area level. In each year, we have 64 markets. In order to obtain the national average housing price, we start by separately computing the average housing prices for urban and rural areas. For example, for urban areas we compute a weighted average of state-level urban housing prices using the number of owners in urban areas for each state as weights. Then we further aggregate these two house prices by the number of owners in rural and urban areas. In the simulation, we also adjust the state share as this is another form of reallocating households across areas with different house prices.

Figure 4.4 shows the results. On the aggregate level, rising life expectancy, urbanization, and changes in the population structure, which combines changes in fertility, survival probability, and international immigration, can account for price growth of 42,378 in the dollar value of 1999 from 1970 to 2010. As a comparison, the actual price growth during the same period is 104,001 in 1999 dollars. Hence, the four factors can account for about 41% of the increase in housing prices from 1970 to 2010. Specifically, the predicted housing price growth from 1980 to 2000 almost perfectly lines up with the data. However, it underpredicts the price growth for two periods: 1970-1980 and 2000-2010. One possible explanation is that the high inflation rate from 1970 to 1980 and the mortgage expansions after the 2000s have been the main driving forces for the significant rise in the housing price growth during these periods. In the estimation, these changes in the macroeconomic environment are captured by the year dummies, which are not incorporated in the simulation as they are not the focus of this chapter.

We also examine the housing price for rural and urban areas separately. Figure 4.5 shows the results. The model can account for 48% of the housing price growth in urban areas. Meanwhile, it does not predict much of the increase in the housing price in rural areas. This result is not surprising since Figure 4.2 shows that the total population in rural areas has barely increased. In addition, the supply estimation suggests that housing prices in rural areas are not as responsive to changes in demand as housing prices in urban areas.

Figure 4.4: Model Fit



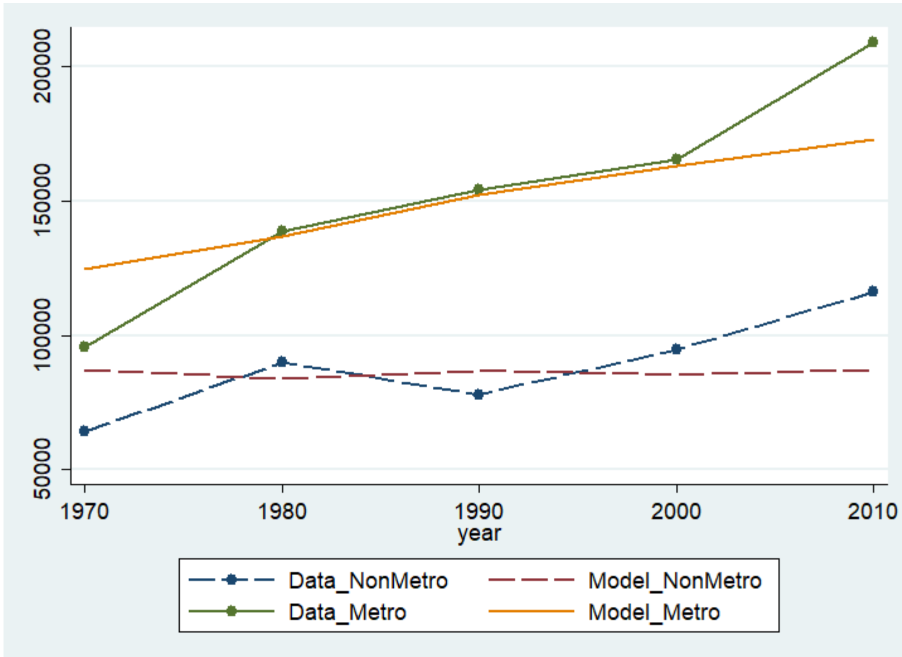
4.4.2 Decomposition

This subsection isolates the contribution of three factors that contribute to the growth of housing prices: rising life expectancy, international immigration, and urbanization. We leave out the changes in fertility rates because even if we fix fertility rates and survival probabilities in the 1970 level, the population structure will still change due to the baby boom generation, which makes the interpretation of changing fertility rates unclear.

We remove these three factors one by one to examine their impact on the aggregate housing price.

Immigration To isolate the impact of immigrants and their future generations on the population structure, we apply the historical fertility and mortality rates to the population structure in 1970 to compute the aggregate population structure ($\bar{N}_{a,t}$ in Equation 4.6) in the years between 1971 and 2010 under the hypothetical scenario where there were no new immigrants entering the United States after 1970. Specifically, for each of these following years, we predict the

Figure 4.5: Model Fit by Area



number of newborns based on the fertility rates and the number of death of different ages based on the mortality rates and then compute the aggregate population structure.[11] Our calculation suggests that immigrants and their future generations since 1970 account for around 18% of the total population in 2010. As housing demand is a household-level decision, we convert the age structure of population into the age structure of households by adjusting the population of different ages with the probability of becoming a household head. Specifically, the probability of becoming a household head for an individual from age group a at time t, $q_{a,t}$ is defined as:

$$q_{a,t} = \frac{\bar{N}_{a,t}}{M_{a,t}}, \tag{4.8}$$

where $N_{a,t}$ is the number of household head of age a and $M_{a,t}$ is the number of individuals of age a at time t.

Urbanization To isolate the impact of urbanization, we fix the fraction of households living in metropolitan areas for each age group, i.e. $\theta_{a,t}$ in Equation 4.6, at its 1970 level.

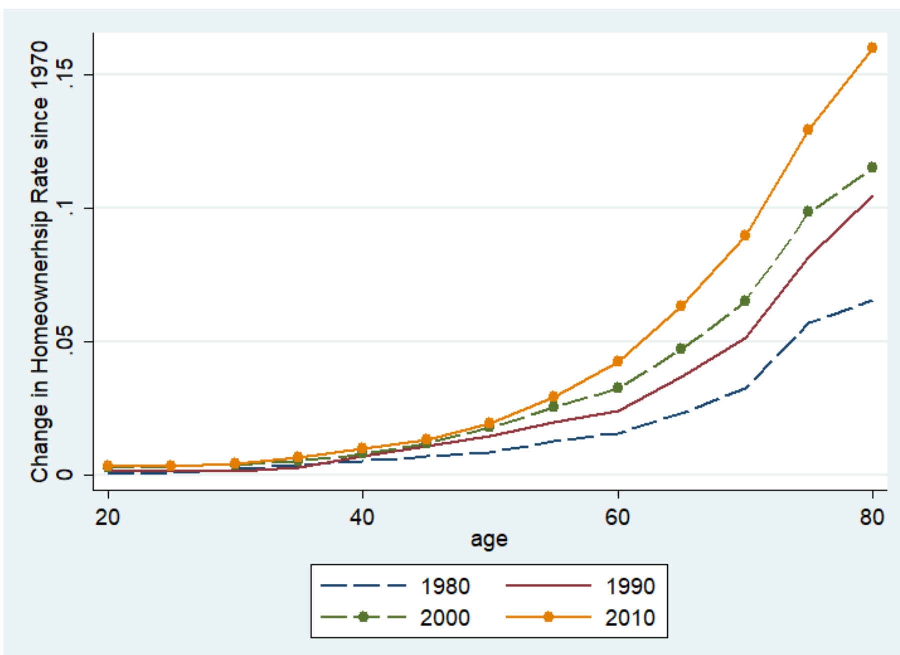
Rising Life Expectancy Rising survival probability affects housing demand through two channels. First, it changes the housing demand for individual households through Equation 4.3. Specifically, it leads to a higher housing demand, especially for senior households, as they have enjoyed the biggest increase in the survival probabilities. Second, it changes the age structure, which leads to a higher fraction of senior households and an increase in the total

[11]More details can be found in Appendix C.2.

population. To isolate the impact of long life expectancy, we follow a two-step procedure. First, we fix the survival probabilities at the 1970 level when computing housing demand using Equation 4.3. Second, we simulate the population structure using the survival probabilities in 1970, following the same procedure that we describe above to compute the population structure in the no-immigrants scenario. This procedure produces the population structure in the scenario where there were neither new international immigrants nor improvement in survival probabilities since 1970.

In addition, we check the marginal impact of increasing survival probability on the housing consumption profile, i.e. the homeownership rate by age. Specifically, we keep all other model primitives at their 1970 levels, change the survival probabilities to other years, and then calculate the changes in homeownership rate of different age groups for each state, urban and rural areas separately.[12] As the marginal impact varies across states and areas, we take the average changes in homeownership rates by age across the 64 markets and plot Figure 4.6. Rising survival probability increases the housing demand, especially for senior households. For households in their 70s, increasing survival probability generates a 10 percentage point increase in their homeownership rate from 1970 to 2010.

Figure 4.6: Marginal Impact of Rising Survival Probability on Homeownership Rate by Age



Due to the interaction among different factors, the order of removing each factor may affect the decomposition results. Therefore, for robustness, we conduct two counterfactual exper-

[12]The logit functional form implies the marginal impact of survival probability varies with other factors such as income and housing price.

iments. In the first experiment (Panel (a) in Figure 4.7), we start by removing international immigration, and then we fix the survival probability at the 1970 level. Finally, we fix the urbanization rate at its 1970 level. In the second experiment, we first fix the urbanization rate at its 1970 level, followed by removing international immigration and applying the survival probability in 1970. The results are shown in Panel (b) in Figure 4.7.

Our substantive results are robust with respect to the counterfactual experiments. Among the three factors we focus on, urbanization makes the largest contribution. Depending on the experiment, reallocating people from rural areas where the price elasticity of supply is high to urban areas where the price elasticity of supply is low contributes to 31-40% of the price growth that the model can account for. International immigrants and their future generations account for 24-29%, followed by the rising survival probability, which contributes to 15-18% of the model generated price growth.

The yellow dashed line in Figure 4.7 shows that, without these three factors or changing macroeconomic environment such as the mortgage expansions, we should have started to see a decline in the aggregate housing price after 2000. Roughly speaking, in this scenario, the only change in the economy is a decreasing fertility rate. Therefore, our finding suggests that an increasing share of elderly people accompanied by declining population size can indeed lead to decreasing housing prices, which echos the main findings in Mankiw and Weil (1989) and Hiller and Lerbs (2016).

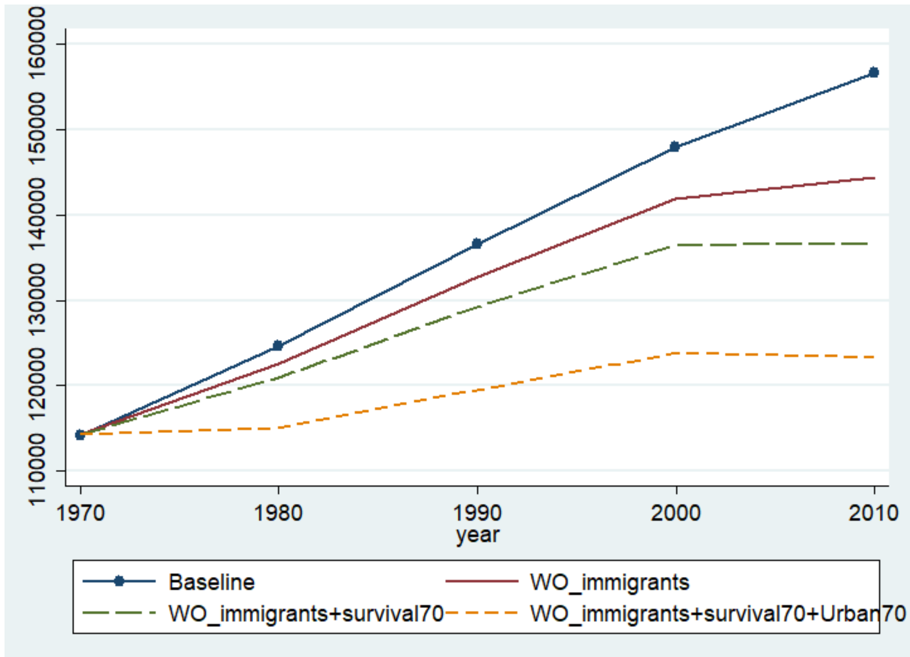
4.4.3 Housing Price Prediction: 2010 to 2050

One reason that we focus on these four factors is that they are sustained and predictable. Our model allows us to evaluate the impact of future trends in these factors on the housing prices, both at the aggregate level and at the area (Urban and Rural) level, in the next several decades. This analysis is of particular importance for understanding whether changes in demographics can lead to a sustained growth of housing prices in the future.

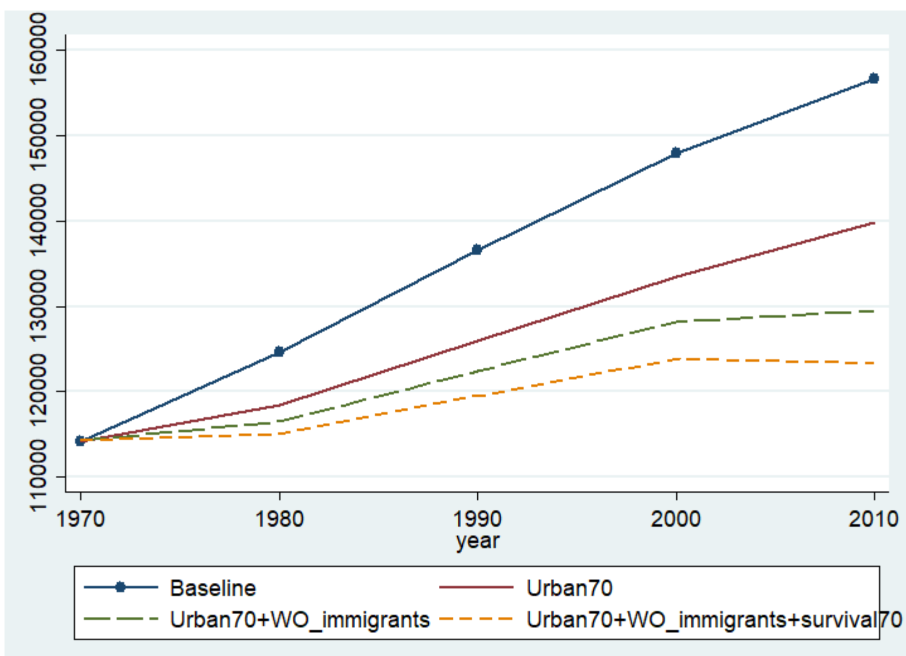
Baseline Predictions

In this section, we provide decennial predictions on housing prices from 2020 through 2050. Our approach requires projections on future demographic trends. We take advantage of the 2017 National Population Projections Tables provided by the Census, which provides predicted population structure every ten years. The population projection is based on a cohort-component method. It takes future trends in births, death, and net international migration into consider-

Figure 4.7: Decomposition



(a) Experiment 1

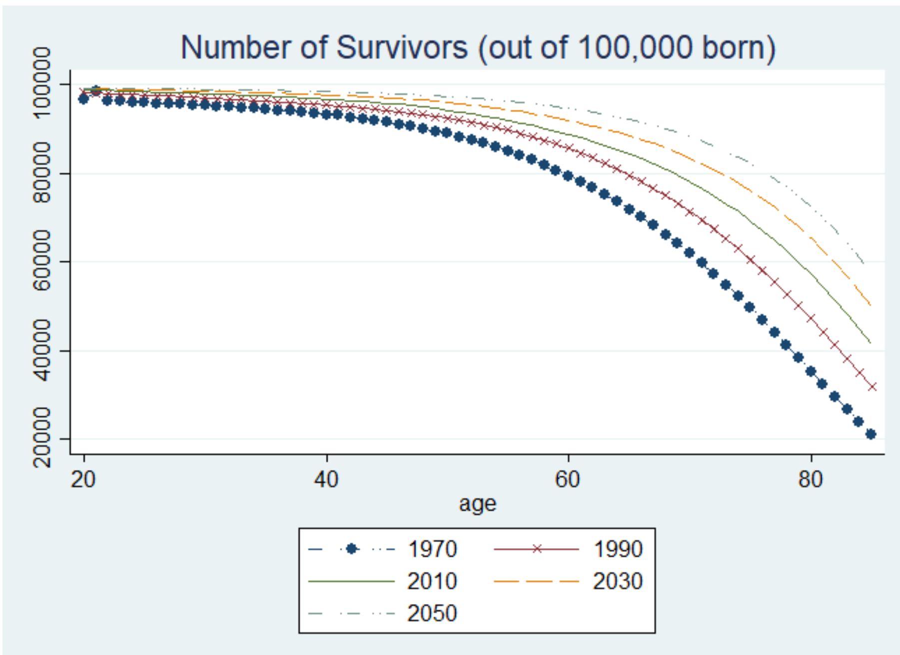


(b) Experiment 2

ation.[13] We apply the household formation rate $q_{a,2010}$ specified in Equation 4.8 to the age structure of the population to compute the age structure of households as housing consumption is a household level decision. In addition, projections of future survival probability can also be obtained from the 2017 Census prediction. Income profile is fixed at the 2010 level.

Figure 4.8 presents the projections on the survival probabilities. We expect to see further improvements in the survival probabilities. Similar to the past decades, senior people will enjoy a larger increase in their survival rates.

Figure 4.8: Predictions on Survival Probabilities



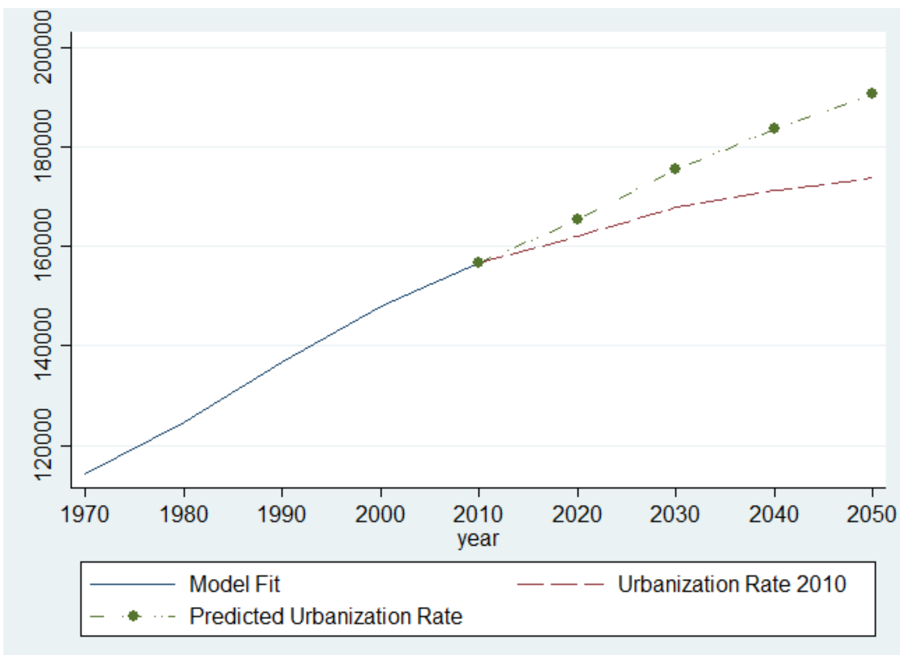
To obtain projections of future urbanization rates, we utilize the data provided by the United Nations and Social Affairs (2018). It predicts that the urbanization rate in the United States will increase from 80.8% in 2010 to 89.2% in 2050. We note that the definition of urbanization adopted in the United Nations projections is somewhat different from the one in this chapter. For example, the urbanization rate (fraction of households living in metropolitan areas) in 2010 in our sample is 76.3%, which is 4.5 percentage points lower than the one reported by the United Nations. This is partly because metropolitan is a stricter definition compared to urban areas. Specifically, to be qualified for a metropolitan area, an area needs to have a higher population density compared to urban areas. To account for this discrepancy, we downward adjust the projections provided by the United Nations by 4.5 percentage points. As a result, our projection of the urbanization rate in 2050 is 84.7%. Roughly speaking, the urbanization

[13] More details can be found at https://www2.census.gov/programs-surveys/popproj/technical-documentation/methodology/methodstatement17.pdf.

rate will keep growing at a rate of 2 percentage points per decade.

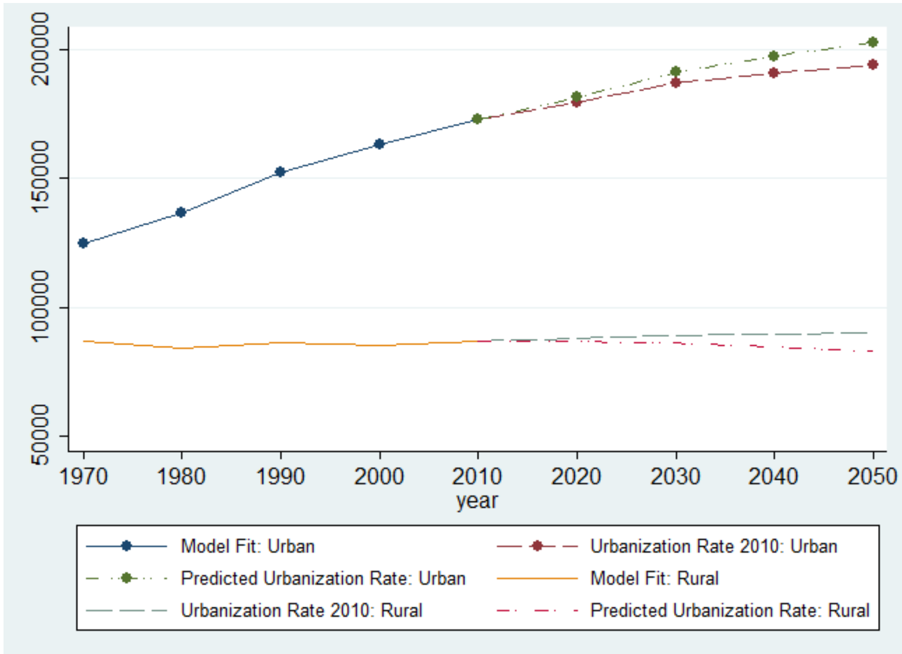
Figure 4.9 depicts the result. We find that housing prices will keep growing in the next 50 years but at a slower rate. The green dashed line shows that stable growth in survival probabilities, sustained urbanization, and expected changes in population structure will generate a 21.7% growth in the aggregate housing price from 2010 to 2050. As a comparison, the solid blue line shows that these factors account for a 41% growth in the aggregate housing price from 1970 to 2010. Consistent with our previous finding that, among the factors considered in this chapter, urbanization makes the largest contribution to the house price growth in the past several decades, we continue to find that the predicted increase in urbanization will have a large impact on the future growth of house prices. The red dashed line shows that, if the urbanization rate is fixed at its 2010 level, we will only see a 10.9% increase in the aggregate housing price until 2050.

Figure 4.9: Projection with Various Urbanization Rate



Our findings in Section 4.2.3 and Section 4.3.2 suggest that housing prices in urban and rural areas might have diverging trends due to the process of urbanization. To examine whether it will be the case in our context, we also predict the housing price for urban and rural areas separately. Figure 4.10 shows the results. If we fix the urbanization rate at the 2010 level while we make the prediction, we see a steady housing price growth in urban areas and a relatively stable housing price in rural areas. When we apply the predicted urbanization rate provided by the United Nations, we see an even stronger housing price growth in urban areas and a moderate decline in housing prices in rural areas (about 5%).

Figure 4.10: Housing Price Projection: Urban and Rural Areas



This finding is broadly consistent with what has been found in many other developed countries. For example, Japan has suffered negative population growth due to population aging since 2010. Consequently, the aggregate housing price in Japan has fallen (dashed line in Figure 4.11). Nevertheless, when we look at the land price index at different areas, we find that land prices in the 6 biggest metropolitan areas have risen while other places have seen a steady decline in land price (see Figure 4.11).[14]

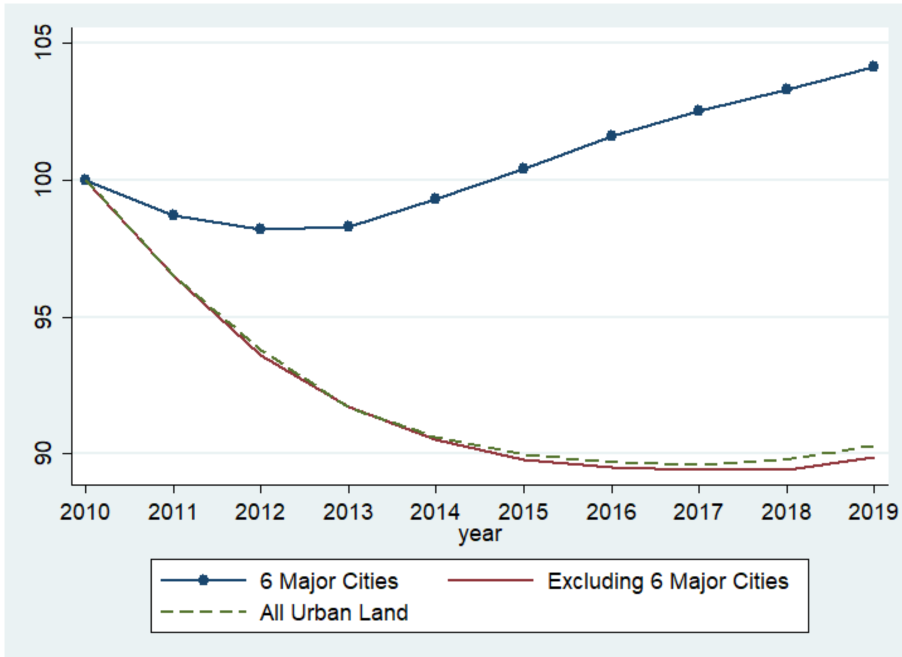
Housing Price Projections with Different Levels of Immigration

The 2017 National Population Projections provide the population structure projections in high, low, and zero immigration scenarios through 2050, based on the official estimates of the resident population in 2016. Taking advantage of these projections, we predict future housing prices with different levels of immigration.

We start by characterizing the population structure under the high, low, and zero immigration scenarios. Figure 4.12 shows that total population projections vary substantially under different scenarios. In the high immigration scenario, the total population will reach 420 million by 2050, which is 28% higher compared to the scenario where no immigration takes place between 2016 and 2050. We note that, if international immigration is completed halted after 2016, the U.S. will see a decline in the total population starting from 2030.

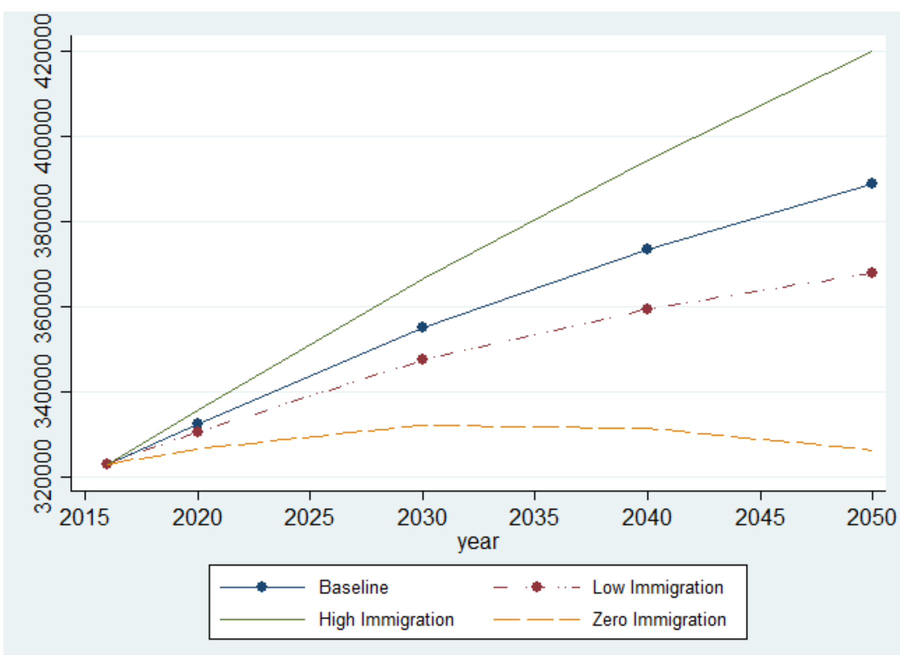
[14]The six biggest metropolitan areas refer to ku-area of Tokyo, Yokohama, Nagoya, Kyoto, Osaka, and Kobe.

Figure 4.11: Residential Urban Land Price Index: Japan



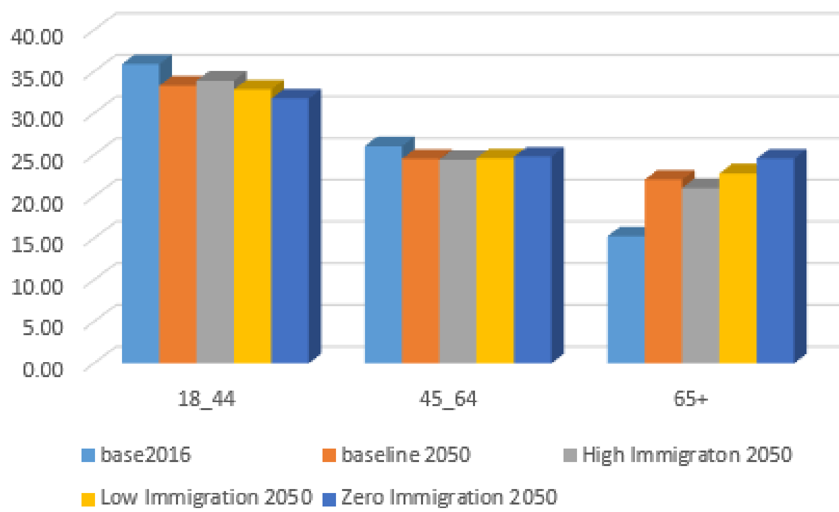
Source: The Land Institute of Japan

Figure 4.12: Total Population with Different Levels of Immigration



In addition, international immigration also affects the age distribution of the population. Figure 4.13 plots the age distribution of population in 2016 and the age distributions in 2050 under different levels of immigration. The population share of young (18-44-year-olds) and middle-aged (45-64-year-olds) declines, accompanied by a rising population share of the old (65+) from 2016 to 2050. As expanding the labor force is one of the main goals of the immigration policies, it is not surprising to see that immigration partially alleviates the declining share of working-age population problem. For instance, the share of working age population (18-64-year-olds) in the zero immigration scenario in 2050 is 56.57 percent, which is 1.8 percentage point lower compared to the high immigration scenario.

Figure 4.13: Age distribution of Population with Different Levels of Immigration

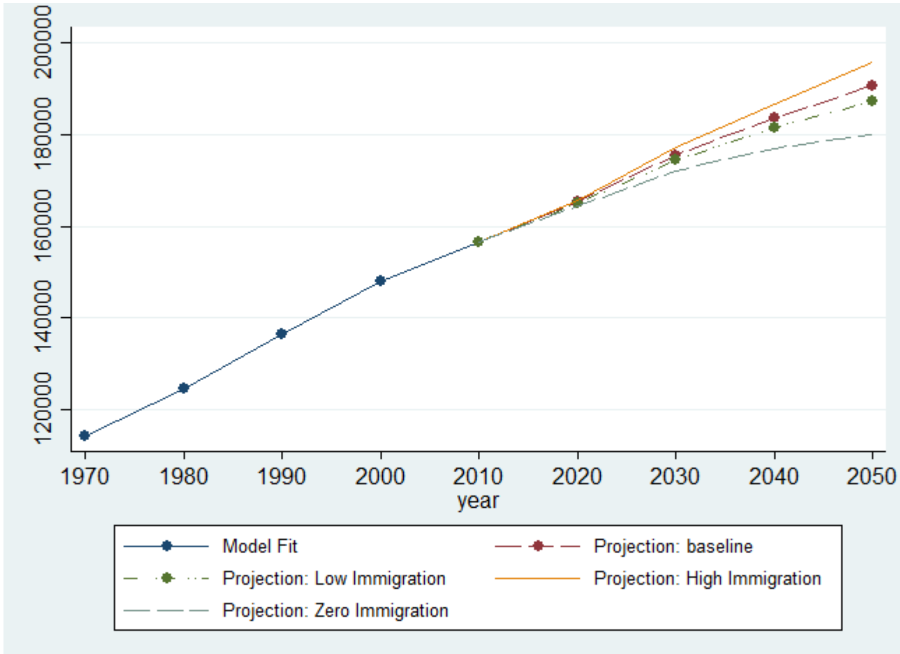


We apply the population structure predictions under different immigration levels to predict the future housing price.[15] Figure 4.14 shows the results. The predicted housing price growth varies with the levels of immigration. In the high immigration scenario, the housing price will grow by 25%, while in the zero immigration scenario the housing price will only increase by 15% from 2010 to 2050.

It seems that the negative population growth after 2030 is not leading to a decline in the aggregate housing prices under the zero immigration scenario. To investigate the underlying driving force for the housing price growth under negative population growth, we examine the role that urbanization plays. Figure 4.15 displays the results. The red connected line and the blue dashed line show that, when urbanization is taking place, we predict an increase in urban housing prices and a decrease in rural housing prices. Even under the zero immigration scenario, the increase in urban housing prices outweighs the decrease in rural housing prices,

[15] We apply the predicted urbanization rates in this exercise.

Figure 4.14: Predicted Housing Prices with Different Levels of Immigration

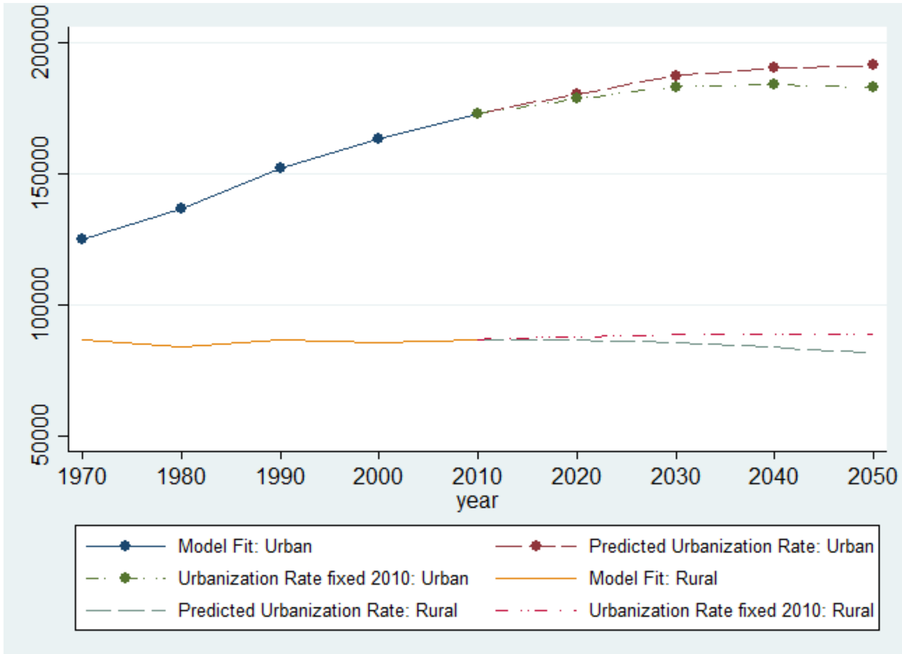


leading to a growth in the aggregate housing price. On the other hand, the green connected line and the red dashed line show that, when the urbanization rate is fixed at the 2010 level, negative population growth after 2030 will lead to a mild decline in the housing price both in urban and rural areas, which suggests that rising life expectancy mitigates the impact of negative population growth on housing prices in this case. The comparison of predicted housing price under different urbanization rates suggest that urbanization alleviates the downward pressure of the negative population growth on house prices, which is consistent with the observation for Japan. Under the zero immigration scenario with the urbanization rate fixed at 2010 level, we are expected to see a 5% growth in the aggregate housing price from 2010 to 2050.

4.5 Conclusion

Housing prices in the United States have enjoyed significant growth in the past several decades. The existing literature often focuses on explanations related to the financial market. In this chapter, we complement the literature by providing a comprehensive quantitative study of the effects of four demographic-related factors on housing prices. Adopting a general equilibrium framework, we find that changing fertility, rising life expectancy, urbanization, and international immigration can account for 41% of the observed growth in housing prices from 1970 to 2010. Leveraging projections of population structure, survival probabilities, and urbanization rates, we use our model to predict the aggregate housing prices through 2050 under different

Figure 4.15: Predicted Housing Prices by Area under Zero Immigration Scenario



urbanization rates and different levels of immigration. We find that the aggregate housing price will keep growing for the next several decades, albeit at a rate slower than before. Among these factors, urbanization is the major contributor, both in terms of explaining past housing price growth and predicting future growth.

We stress that our finding is not in conflict with the scholarly consensus that the rapid increase in housing prices that we have seen can be largely attributed to financial side explanations such as mortgage expansions. Indeed, more than half (60%) of the increase cannot be explained by the factors that we propose. Instead, we think our study sends the message that factors related to economic fundamentals, such as demographics, should also be treated carefully when analyzing the housing market.

Bibliography

[1] Alexis Anagnostopoulos, Orhan Erem Atesagaoglu, and Eva Carceles-Poveda. Skill-biased technological change and homeownership. *Journal of Economic Dynamics and Control*, 37(12):3012–3033, 2013.

[2] Matthew Chambers, Carlos Garriga, and Don E Schlagenhauf. Accounting for changes in the homeownership rate. *International Economic Review*, 50(3):677–726, 2009.

[3] Junhua Chen, Fei Guo, and Ying Wu. One decade of urban housing reform in china: Urban housing price dynamics and the role of migration and urbanization, 1995–2005. *Habitat International*, 35(1):1–8, 2011.

[4] Edward L Glaeser, Joseph Gyourko, and Raven E Saks. Urban growth and housing supply. *Journal of economic geography*, 6(1):71–89, 2006.

[5] Richard K Green and Hyojung Lee. Age, demographics, and the demand for housing, revisited. *Regional Science and Urban Economics*, 61:86–98, 2016.

[6] Daniel Greenwald, Adam Guren, et al. Do credit conditions move house prices? *Unpublished manuscript, MIT*, 2019.

[7] Joseph Gyourko and Peter Linneman. Analysis of the changing influences on traditional households' ownership patterns. *Journal of Urban Economics*, 39(3):318–341, 1996.

[8] Bruce W Hamilton. The baby boom, the baby bust, and the housing market a second look. *Regional Science and Urban Economics*, 21(4):547–552, 1991.

[9] Norbert Hiller and Oliver W Lerbs. Aging and urban house prices. *Regional Science and Urban Economics*, 60:276–291, 2016.

[10] A Steven Holland. The baby boom and the housing market: Another look at the evidence. *Regional Science and Urban Economics*, 21(4):565–571, 1991.

[11] N Gregory Mankiw and David N Weil. The baby boom, the baby bust, and the housing market. *Regional science and urban economics*, 19(2):235–258, 1989.

[12] Atif Mian and Amir Sufi. The consequences of mortgage credit expansion: Evidence from the us mortgage default crisis. *The Quarterly Journal of Economics*, 124(4):1449–1496, 2009.

[13] James M Poterba. Tax subsidies to owner-occupied housing: an asset-market approach. *The quarterly journal of economics*, 99(4):729–752, 1984.

[14] Kamila Sommer, Paul Sullivan, and Randal Verbrugge. The equilibrium effect of fundamentals on house prices and rents. *Journal of Monetary Economics*, 60(7):854–870, 2013.

[15] Robert Topel and Sherwin Rosen. Housing investment in the united states. *Journal of political Economy*, 96(4):718–740, 1988.

[16] Department of Economic United Nations and Population Division Social Affairs. World urbanization prospects: The 2018 revision, online edition. 2018.

[17] Population Division U.S. Census Bureau. Projected population for the united states by age group and alternative immigration scenario: 2020 to 2060 (np2017-d). 2019.

Appendix A

Appendices for Chapter 2

A.1 List of Sample Cities

This section lists the 161 Metropolian areas used in the OLS estimation in Chapter 2 alphabetically. The one followed by (IV) are also used in the IV estimation.

A.2 OLS Regression with 107 Cities

This section presents the OLS estimations in Chapter 2 using the same sample as the IV estimations.

MSA	MSA
Akron, OH (IV)	Charlotte-Gastonia-Rock Hill, NC/SC (IV)
Albany-Schenectady-Troy, NY (IV)	Chattanooga, TN/GA (IV)
Albuquerque, NM (IV)	Chicago, IL (IV)
Allentown-Bethlehem-Easton, PA/NJ (IV)	Chico, CA
Amarillo, TX	Cincinnati-Hamilton, OH/KY/IN (IV)
Ann Arbor, MI	Cleveland, OH (IV)
Appleton-Oshkosh-Neenah, WI (IV)	Colorado Springs, CO
Atlanta, GA (IV)	Columbia, MO
Atlantic City, NJ	Columbia, SC (IV)
Austin, TX (IV)	Columbus, OH (IV)
Bakersfield, CA (IV)	Corpus Christi, TX (IV)
Baltimore, MD (IV)	Dallas-Fort Worth, TX (IV)
Baton Rouge, LA (IV)	Davenport, IA - Rock Island-Moline, IL (IV)
Beaumont-Port Arthur-Orange, TX (IV)	Dayton-Springfield, OH (IV)
Billings, MT	Daytona Beach, FL
Biloxi-Gulfport, MS	Denver-Boulder, CO (IV)
Binghamton, NY (IV)	Des Moines, IA (IV)
Birmingham, AL (IV)	Detroit, MI (IV)
Bloomington-Normal, IL	Duluth-Superior, MN/WI (IV)
Boise City, ID	Erie, PA (IV)
Boston, MA/NH (IV)	Eugene-Springfield, OR
Brownsville-Harlingen-San Benito, TX	Fayetteville, NC
Buffalo-Niagara Falls, NY (IV)	Fayetteville-Springdale, AR
Canton, OH (IV)	Fort Collins-Loveland, CO
Cedar Rapids, IA	Fort Lauderdale-Hollywood-Pompano Beach, FL (IV)
Champaign-Urbana-Rantoul, IL	Fort Myers-Cape Coral, FL
Charleston-N. Charleston, SC (IV)	Fort Wayne, IN (IV)

MSA	MSA
Fresno, CA (IV)	Lincoln, NE
Gainesville, FL	Little Rock-N. Little Rock, AR (IV)
Galveston-Texas City, TX	Los Angeles-Long Beach, CA (IV)
Grand Rapids, MI (IV)	Louisville, KY/IN (IV)
Greeley, CO	Lubbock, TX
Green Bay, WI	Macon-Warner Robins, GA
Greensboro-Winston Salem-High Point, NC (IV)	Madison, WI (IV)
Greenville-Spartenburg-Anderson, SC (IV)	McAllen-Edinburg-Pharr-Mission, TX
Hamilton-Middleton, OH	Melbourne-Titusville-Cocoa-Palm Bay, FL
Harrisburg-Lebanon–Carlisle, PA (IV)	Memphis, TN/AR/MS (IV)
Hartford-Bristol-Middleton- New Britain, CT (IV)	Miami-Hialeah, FL (IV)
Houston-Brazoria, TX (IV)	Milwaukee, WI (IV)
Indianapolis, IN (IV)	Minneapolis-St. Paul, MN (IV)
Jackson, MS (IV)	Mobile, AL (IV)
Johnson City-Kingsport–Bristol, TN/VA	Modesto, CA
Johnstown, PA (IV)	Montgomery, AL
Kalamazoo-Portage, MI	Nashville, TN (IV)
Kansas City, MO/KS (IV)	New Haven-Meriden, CT (IV)
Kileen-Temple, TX	New Orleans, LA (IV)
Knoxville, TN (IV)	New York, NY-Northeastern NJ (IV)
Lafayette, LA	Norfolk-VA Beach–Newport News, VA (IV)
Lafayette-W. Lafayette, IN	Oklahoma City, OK (IV)
Lakeland-Winterhaven, FL	Olympia, WA
Lancaster, PA (IV)	Omaha, NE/IA (IV)
Lansing-E. Lansing, MI (IV)	Orlando, FL (IV)
Las Vegas, NV (IV)	Pensacola, FL
Lexington-Fayette, KY	Peoria, IL (IV)

MSA	MSA
Philadelphia, PA/NJ (IV)	Sarasota, FL
Phoenix, AZ (IV)	Savannah, GA
Pittsburgh, PA (IV)	Scranton-Wilkes-Barre, PA (IV)
Portland, OR/WA (IV)	Seattle-Everett, WA (IV)
Providence-Fall River-Pawtucket, MA/RI (IV)	South Bend-Mishawaka, IN (IV)
Provo-Orem, UT	Spokane, WA (IV)
Raleigh-Durham, NC	Springfield, MO
Reading, PA (IV)	Springfield-Holyoke-Chicopee, MA (IV)
Reno, NV	State College, PA
Richland-Kennewick-Pasco, WA	Stockton, CA (IV)
Richmond-Petersburg, VA (IV)	Syracuse, NY (IV)
Riverside-San Bernardino, CA (IV)	Tacoma, WA (IV)
Roanoke, VA	Tampa-St. Petersburg-Clearwater, FL (IV)
Rochester, NY (IV)	Toledo, OH/MI (IV)
Rockford, IL (IV)	Trenton, NJ (IV)
Saginaw-Bay City-Midland, MI	Tucson, AZ (IV)
St. Louis, MO/IL (IV)	Tulsa, OK (IV)
Salem, OR	Utica-Rome, NY (IV)
Salinas-Sea Side-Monterey, CA (IV)	Ventura-Oxnard-Simi Valley, CA (IV)
Salt Lake City-Ogden, UT (IV)	Visalia-Tulare-Porterville, CA
San Antonio, TX (IV)	Washington, DC/MD/VA (IV)
San Diego, CA (IV)	West Palm Beach-Boca Raton-Delray Beach, FL (IV)
San Francisco-Oakland-Vallejo, CA (IV)	Wichita, KS (IV)
San Jose, CA (IV)	Wilmington, DE/NJ/MD (IV)
Santa Barbara-Santa Maria-Lompoc, CA (IV)	York, PA (IV)
Santa Cruz, CA	Youngstown-Warren, OH/PA (IV)
Santa Rosa-Petaluma, CA	

Table A.1: OLS Results with 107 Cities

	(1) $log(P)$	(2) OR^{YNC}	(3) OR^{YC}	(4) OR^{MNC}	(5) OR^{MC}	(6) OR^{YNC}	(7) OR^{YC}	(8) OR^{MNC}	(9) OR^{MC}
κ	2.092***	-0.575***	-0.718***	-0.434***	-0.334***	-0.0983	-0.429***	0.00541	-0.0588
	(0.171)	(0.0691)	(0.0721)	(0.0563)	(0.0416)	(0.0671)	(0.0796)	(0.0559)	(0.0454)
I^{YNC}	0.417**	0.291***	0.105	0.0127	-0.0186	0.386***	0.163**	0.100*	0.0363
	(0.167)	(0.0634)	(0.0749)	(0.0548)	(0.0417)	(0.0624)	(0.0803)	(0.0511)	(0.0437)
I^{YC}	0.499***	-0.116**	0.186***	-0.134***	-0.0840**	-0.00278	0.255***	-0.0290	-0.0183
	(0.115)	(0.0466)	(0.0644)	(0.0484)	(0.0345)	(0.0365)	(0.0602)	(0.0393)	(0.0305)
I^{MNC}	0.678***	0.0452	0.119	0.272***	0.130**	0.200***	0.213**	0.414***	0.219***
	(0.205)	(0.0772)	(0.0917)	(0.0668)	(0.0523)	(0.0721)	(0.0924)	(0.0572)	(0.0455)
I^{MC}	0.0263	-0.111*	-0.213***	-0.0757	0.0903**	-0.105**	-0.209***	-0.0702	0.0938**
	(0.158)	(0.0647)	(0.0723)	(0.0596)	(0.0424)	(0.0475)	(0.0658)	(0.0447)	(0.0379)
$log(P)$						-0.228***	-0.139***	-0.210***	-0.132***
						(0.0168)	(0.0271)	(0.0145)	(0.0163)
N	-0.0259	-0.00518	-0.000540	-0.00777	-0.00438	-0.0111***	-0.00413	-0.0132***	-0.00779**
	(0.0158)	(0.00569)	(0.00656)	(0.00528)	(0.00458)	(0.00398)	(0.00542)	(0.00357)	(0.00338)
Elasticity	-0.136***	0.0304***	0.0338***	0.0232***	0.0175***	-0.000556	0.0150***	-0.00527*	-0.000391
	(0.0105)	(0.00373)	(0.00404)	(0.00326)	(0.00270)	(0.00358)	(0.00421)	(0.00280)	(0.00219)
1990.year	-0.0828***	-0.0220**	-0.0540***	-0.0287***	-0.0279***	-0.0409***	-0.0655***	-0.0460***	-0.0388***
	(0.0282)	(0.0109)	(0.0102)	(0.00961)	(0.00652)	(0.00866)	(0.00976)	(0.00749)	(0.00576)
2000.year	-0.0841***	0.0206	0.000221	-0.00135	-0.00369	0.00147	-0.0114	-0.0190**	-0.0148**
	(0.0298)	(0.0126)	(0.0132)	(0.0110)	(0.00774)	(0.0106)	(0.0126)	(0.00915)	(0.00726)
2010.year	0.266***	0.0210	0.103***	0.00120	0.0272**	0.0816***	0.140***	0.0571***	0.0622***
	(0.0445)	(0.0189)	(0.0210)	(0.0163)	(0.0120)	(0.0176)	(0.0215)	(0.0146)	(0.0117)
Constant	-5.537***	-0.443	-1.358***	0.0905	-0.399	-1.703***	-2.124***	-1.071***	-1.129***
	(1.170)	(0.467)	(0.483)	(0.376)	(0.282)	(0.393)	(0.469)	(0.304)	(0.261)
Observations	428	428	428	428	428	428	428	428	428
R-squared	0.774	0.572	0.494	0.539	0.351	0.724	0.553	0.727	0.528

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix B

Appendices for Chapter 3

B.1 Distribution of Prices and Rents

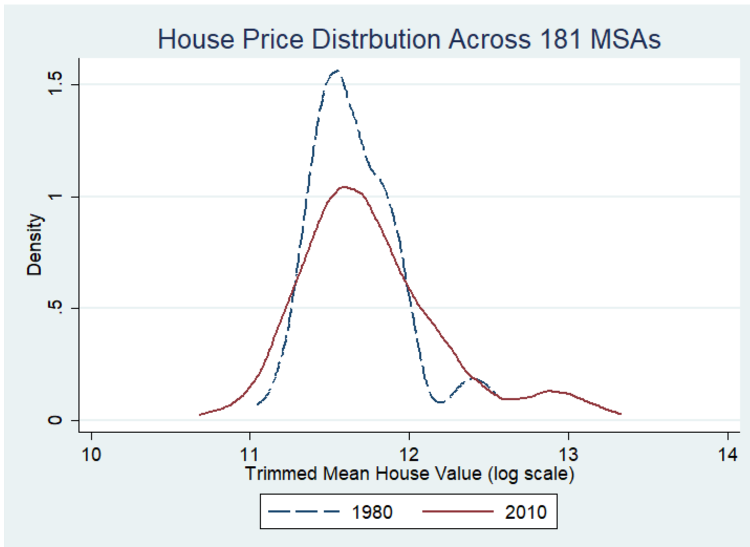
Figure B.1 plots the density distribution of prices and rents across the largest 181 cities for 1980 and 2010. Both prices and rents become more dispersed with fatter tails on both sides over time. Moreover, the dispersion of prices has increased more than rents.

B.2 Land Use

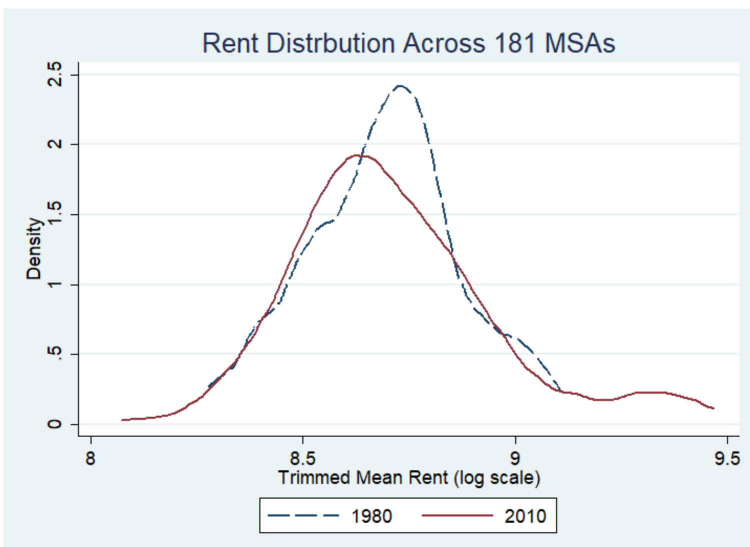
In this section, I present more evidence about the land use in different cities and its changes over time for houses and apartments separately. Houses use more land on average compared to apartments. Table 2 summarizes the population density, the homeownership rate, the fraction of owners that live in houses with a lot, i.e. detached or attached house, the median lot size of owned houses (i.e. land use per house), and the median land use per rental unit for three cities: New York City, Houston, and Memphis. New York City is a representative large city with strict land regulations, while Houston is a large city known for its permissive zoning regulations. Memphis is an example of a mid-sized city. The land use of houses is represented by the lot size, while the land use of apartments is measured by the unit size divided by the number of stories in the building. Population density from the Census 2010 is included as an approximation of the intensity of land demand.

Comparing New York City with Memphis suggests that as land becomes scarcer, i.e., population density increases, land use for both houses and apartment declines. However, apartments are more efficient in economizing land use by building up. For instance, the median lot size in Memphis is only 60% higher than New York, while the median land use for apartments in Memphis is three times that of New York. In New York City, 88% of rental units are in high-

Figure B.1: Distribution of Prices of Rents



(a) Prices



(b) Rents

Table B.1: Characteristics of Owner Occupied Dwellings and Rental Units

		New York City	Houston	Memphis
Owners	Population Density 2000 (per sq miles)	8158.7	705	377.7
	Ownership Rate	37	54	58
	Fraction in Detached/ Attached House	55	92	100
	Median Unit (sqf)	1900	1800	1500
	Median Lot (sqf)	5500	5500	9000
	Lot Size Distributions for Owner-Occupied Houses			
	<=1/2 acre	41.4	60.4	62.8
	1/2-1 acre	4.2	4.4	13.5
	>=1 acre	54.4	35.2	23.7
Renters	Fraction in Detached/ Attached House	6	25	29
	Median Unit (sqf)	700	800	900
	Median Land Use (unit size / stories)	129	462	450
	Fraction of High Rise Building (>=4 stories)	88	16	4

rise buildings (more than four stories), while the fraction of high-rise rental units in Memphis is only 4%.

The New York City and Houston comparison points to the impact of regulation on land use and housing demand. According to Gyourko, Saiz, and Summers (2008), fifteen of the sixteen jurisdictions within New York Primary Metropolitan Statistical Areas (PMSA) have a minimum lot size requirement, and seven have a minimum lot size requirement that is greater than one acre. Meanwhile, among the twelve jurisdictions in the Houston PMSA, nine have a minimum lot size requirement, and only four have a minimum lot size requirement that is greater than one acre.[1] The median lot size for owner-occupied houses in Houston is the same as that in New York City, despite the fact that the population density in New York is eight times higher than that of Houston. Moreover, 60% of owner-occupied houses in Houston have a lot smaller than 1/2 acre compared to 41% in New York City. Only 35% of owner-occupied houses in Houston have a lot larger than 1 acre, which is 20 percentage points lower than that in New York City. These observations suggest that the minimum lot size zoning is important in determining the land input in the construction of houses, and therefore the total land demand.

Change Land Use: 1980-2010 Consistent with a rise in the national-wide land price over the past 50 years (Knoll, Schularick, and Steger, 2017), recently built houses and apartments tend to use less land. Table B.2 summarizes the characteristics of single-family detached houses built in metropolitan areas by the year of construction.

While the median unit size of detached houses has grown steadily, rising from 1900 sqft in the 1980s to 2300 sqft in the 2000s, more recently built houses tend to have smaller lots. The

[1]Memphis has three jurisdictions in the sample, and all of them have a minimum lot size requirement.

Table B.2: characteristics of single-family detached houses by the construction year in Metropolitan areas

Year Built	1980-1990	1990-2000	2000-2009
Median Unit Size (sqf)	1900	2200	2300
Median Lot Size (sqf)	11000	10000	9000
Lot Size Distribution			
Smaller than 1/2 acre	40%	52%	57%
1/2-1 acre	7%	7%	6%
Larger than 1 acre	53%	41%	37%

Note: Author's calculation using data from AHS

fraction of houses with a small lot (less than 1/2 acre) increases from 40% to 57% while the share of houses with large lots (bigger than 1 acre) declines from 53% to 37%.

A similar pattern holds for rental apartments as well. Table B.3 summarizes the characteristics of multifamily buildings by the year of construction. Newly built apartments tend to be larger compared to the old ones. The median unit size for apartments built after 2000 is 10% larger compared to apartments built in the 1980s. However, newly built apartments building are taller, resulting in less land use per apartment compared to old ones. The fraction of high-rise buildings has more than doubled from the 1980s to the 2000s, increasing from 14% to 36%.

Table B.3: Characteristics of Multi-family apartments by the construction year in Metropolitan areas

Year Built	1980-1990	1990-2000	2000-2009
Median Unit Size	845	900	940
Median Stories	2	3	3
Fraction of High-rise (4+ story)	14%	19%	36%
Land usage per unit	350	333	285

Note: Author's calculation using data from AHS
Land usage for apartments is defined by unit size divided by the number of stories.

B.3 List of Sample Cities

This section lists the 181 Metropolian areas used in Chapter 3 alphabetically.

B.4 Definition of Stationary Competitive Equilibrium

As cities are isolated (i.e. households cannot move across cities), each city is a closed economy that is described by a unique competitive equilibrium.

MSA	MSA
Akron, OH	Charleston-N. Charleston, SC
Albany-Schenectady-Troy, NY	Charlotte-Gastonia-Rock Hill, NC/SC
Albuquerque, NM	Chattanooga, TN/GA
Alexandria, LA	Chicago, IL
Allentown-Bethlehem-Easton, PA/NJ	Chico, CA
Altoona, PA	Cincinnati-Hamilton, OH/KY/IN
Amarillo, TX	Cleveland, OH
Ann Arbor, MI	Colorado Springs, CO
Appleton-Oshkosh-Neenah, WI	Columbia, MO
Atlanta, GA	Columbia, SC
Atlantic City, NJ	Columbus, OH
Augusta-Aiken, GA/SC	Corpus Christi, TX
Austin, TX	Dallas-Fort Worth, TX
Bakersfield, CA	Davenport, IA - Rock Island-Moline, IL
Baltimore, MD	Dayton-Springfield, OH
Baton Rouge, LA	Daytona Beach, FL
Beaumont-Port Arthur-Orange, TX	Decatur, IL
Benton Harbor, MI	Denver-Boulder, CO
Billings, MT	Des Moines, IA
Biloxi-Gulfport, MS	Detroit, MI
Binghamton, NY	Duluth-Superior, MN/WI
Birmingham, AL	El Paso, TX
Bloomington-Normal, IL	Erie, PA
Boise City, ID	Eugene-Springfield, OR
Boston, MA/NH	Fayetteville, NC
Bremerton, WA	Fayetteville-Springdale, AR
Brownsville-Harlingen-San Benito, TX	Flint, MI
Buffalo-Niagara Falls, NY	Fort Collins-Loveland, CO
Canton, OH	Fort Myers-Cape Coral, FL
Cedar Rapids, IA	Fort Wayne, IN
Champaign-Urbana-Rantoul, IL	Fresno, CA

MSA	MSA
Gainesville, FL	Little Rock-N. Little Rock, AR
Grand Rapids, MI	Longview-Marshall, TX
Greeley, CO	Los Angeles-Long Beach, CA
Green Bay, WI	Louisville, KY/IN
Greensboro-Winston Salem-High Point, NC	Lubbock, TX
Greenville-Spartenburg-Anderson, SC	Macon-Warner Robins, GA
Hagerstown, MD	Madison, WI
Harrisburg-Lebanon–Carlisle, PA	Mansfield, OH
Hartford-Bristol-Middleton- New Britain, CT	McAllen-Edinburg-Pharr-Mission, TX
Hickory-Morganton, NC	Medford, OR
Houston-Brazoria, TX	Melbourne-Titusville-Cocoa-Palm Bay, FL
Indianapolis, IN	Memphis, TN/AR/MS
Jackson, MS	Miami-Hialeah, FL
Jacksonville, NC	Milwaukee, WI
Janesville-Beloit, WI	Minneapolis-St. Paul, MN
Johnson City-Kingsport–Bristol, TN/VA	Mobile, AL
Johnstown, PA	Modesto, CA
Joplin, MO	Monroe, LA
Kalamazoo-Portage, MI	Montgomery, AL
Kansas City, MO/KS	Nashville, TN
Kileen-Temple, TX	New Haven-Meriden, CT
Knoxville, TN	New Orleans, LA
Lafayette, LA	New York, NY-Northeastern NJ
Lafayette-W. Lafayette, IN	Norfolk-VA Beach–Newport News, VA
Lakeland-Winterhaven, FL	Ocala, FL
Lancaster, PA	Oklahoma City, OK
Lansing-E. Lansing, MI	Olympia, WA
Las Vegas, NV	Omaha, NE/IA
Lexington-Fayette, KY	Orlando, FL
Lima, OH	Pensacola, FL
Lincoln, NE	Peoria, IL

MSA	MSA
Philadelphia, PA/NJ	Santa Rosa-Petaluma, CA
Phoenix, AZ	Sarasota, FL
Pittsburgh, PA	Savannah, GA
Portland, OR/WA	Scranton-Wilkes-Barre, PA
Providence-Fall River-Pawtucket, MA/RI	Seattle-Everett, WA
Provo-Orem, UT	South Bend-Mishawaka, IN
Racine, WI	Spokane, WA
Raleigh-Durham, NC	Springfield, MO
Reading, PA	Springfield-Holyoke-Chicopee, MA
Redding, CA	State College, PA
Reno, NV	Stockton, CA
Richland-Kennewick-Pasco, WA	Syracuse, NY
Richmond-Petersburg, VA	Tampa-St. Petersburg-Clearwater, FL
Riverside-San Bernardino, CA	Toledo, OH/MI
Roanoke, VA	Trenton, NJ
Rochester, NY	Tucson, AZ
Rockford, IL	Tulsa, OK
Saginaw-Bay City-Midland, MI	Tyler, TX
St. Cloud, MN	Utica-Rome, NY
St. Louis, MO/IL	Ventura-Oxnard-Simi Valley, CA
Salem, OR	Vineland-Milville-Bridgetown, NJ
Salinas-Sea Side-Monterey, CA	Visalia-Tulare-Porterville, CA
Salt Lake City-Ogden, UT	Washington, DC/MD/VA
San Antonio, TX	Waterloo-Cedar Falls, IA
San Diego, CA	Wichita, KS
San Francisco-Oakland-Vallejo, CA	Wichita Falls, TX
San Jose, CA	York, PA
Santa Barbara-Santa Maria-Lompoc, CA	Youngstown-Warren, OH/PA
Santa Cruz, CA	

The competitive equilibrium on an island k is defined by the price function and the rent function, $(P_k^*(h^o), R_k^*(h^r))$; land price q_k^*; value functions $V(x)$; the allocation of housing service $h^{o'*}(x), h^{r'*}(x)$ to owners and renters respectively; asset $a'(x)$; consumption $c(x)$; the distribution of people over the state variable x, $g(x)$; the allocation of land $L^{o*}(h^o)$ and material $M^{o*}(h^o)$ to house developers; and the allocation of land $L^{r*}(h^r)$ and material $M^{r*}(h^r)$ to apartment developers; such that on each island k

- Given house price as a function of house size $P_k^*(h^o)$ and rent as a function of apartment size $R_k^*(h^r)$, the functions $V(x), c(x), a'(x), h^o(x), h^r(x)$ solve the maximization problem for households with state variables x specified in Equation 3.5 and Equation 3.7.

- Given the land price q_k^*, material price ϕ_k, and the house price function $P_k^*(h)$, land input $L^{o*}(h^o)$ and material input $M^{o*}(h^o)$ minimize the cost specified in Equation 3.11 for house developers.

- Given the land price q_k^*, material price ϕ_k, and the apartment rent function $R_k^*(h^r)$, land input $L^{r*}(h^r)$ and material input $M^{r*}(h^r)$ minimize the cost specified in Equation 3.12 for apartment developers.

- House market clears. For all house size h^o, the demand for houses of size h^o equals the supply for houses of the same size.

$$H^o = \int \mathbb{1}_{h^o(x)=h^o} g(x)dx \quad \forall h^o \tag{B.1}$$

- Apartment market clears. For all sizes h^r, the demand for apartments of size h^r equals the supply for apartments of the same size. As the technology of apartment construction exhibits constant return to scale, this is equivalent to total demand of apartments equaling total supply of apartments.

$$H^r = \int \mathbb{1}_{h^r(x)=h^r} g(x)dx \quad \forall h^r \tag{B.2}$$

- q_k clears the land market. The total land used for constructing houses and apartments equals the exogenous supply of land.

$$\int L^o(h^o)H^o d(h^o) + \int L^r(h^r)H^r d(h^r) = LS_k \tag{B.3}$$

  where $L^o(h^o)$ is the land used for constructing a house of size h^o. H^o is the number of households who desire owning a house of size h^o given the equilibrium price function

$P^*(h^o)$. $L^o(h^o)$ is the land used for constructing a house of size h^o. H^r is the number of renters who desire renting an apartment of size h^r given the equilibrium rent function $R^*(h^r)$. The first integral represents total land used for house constructions, the second integral shows total land used for apartment constructions, and LS_k is the total residential land supply in city k, which is assumed to be exogenous.

B.5 Computation Detail

The solution is computed numerically for each individual city. The algorithm solves the households' problem backward from the last period of their life by plugging the price function, Equation 3.13, and the rent function, Equation 3.15, into households' problem in Equation 3.5 and Equation 3.7. After solving for the demand for houses and apartments, I aggregate the land use in Equation 3.14 and Equation 3.16 to calculate the total land demand as a function of land price. For each individual city, the minimum lot size $\bar{L_k}$ and the ownership premium θ_k are estimated through the following algorithm.

- Given minimum lot size $\bar{L_k}$

  1. Initialize the model so that prices and rents in 1980 is the same in the model as in the data. Specifically, I solve for land price and material price by inverting the price Equation 3.13 for a standard owner-occupied house of which size is normalized to be 1, i.e. $h^o = 1$, and the rent Equation 3.15 for a standard two-bedroom apartment, of which size is normalized to be 1, i.e. $h^r = 1$.

  2. Solve for prices and rents for all sizes of owner-occupied and rental units, i.e. $P(h^o)$ and $R(h^r)$ for all h^o and h^r, by plugging the land price and material price provided by step 1 back into Equation 3.13 and Equation 3.15.

  3. Search over ownership premium θ_k to minimize distance between model generated homeownership rates g_j and data g_j^0.

$$\Lambda(\bar{L_k}) = \min_{\theta_k} \sum_j \left(\frac{g_j(\theta_k; \bar{L_k}) - g_j^0}{g^0}\right)^2 \tag{B.4}$$

- Loop over minimum lot size $\bar{L_k}$ to minimize $\Lambda(\bar{L_k})$

B.6 Wage of Construction Workers and Building Permits

In this section, I exploit the cross states variation in the relative wages of construction workers compared to workers in other industries and the number of building permits to test whether increasing housing demand affects the labor cost of construction. I use cross states variation instead of cross cities variation due to the concern of sample size. Construction is not a very large sector in the U.S.. The employment share of construction industry is around 4.5% in 2016.[2] As a result, the 1% ACS sample may not have many cities that contain observations for construction workers.

I run the following regression.

$$log(RelativeWage_{j,2017}) = \alpha + \beta log(Permits_{j,2017}) + \epsilon_j \tag{B.5}$$

where $RelativeWage_{j,2017} = \frac{\text{Mean Wage of Construction Workers}}{\text{Mean Wage of other Workers}}$ is the relative wage of construction workers compared to workers in other industries in state j. The independent variables $Permits_{j,2017}$ is the number of building permits issued in 2017. I use the number of permits for all buildings, 1-unit buildings, 2-unit buildings, buildings with 3 and 4 units, and buildings with 5 units and more as independent variables. The results are presented in the top panel of Table B.4. The data on building permits comes from the Building Permits Survey conducted by the Census. The coefficients on building permits are negative and insignificant, suggesting that relative wages of construction workers do not respond much to increasing housing demand. It is possible that the building permits issued in current year do not represent labor demand for construction as it may take some time to get housing projects started after getting the permits. Therefore, I also use the building permits issued in the previous year as independent variables. The results are presented in the bottom panel of Table B.4. The relative wage of construction workers does not change much towards housing permits issued in the previous year either. These findings are consistent with Gyourko and Saiz (2006), who document large variation in construction costs across housing markets and find that construction costs do not respond significantly to building permits.[3]

[2]Employment by major industry sector provided by BLS: https://www.bls.gov/emp/tables/employment-by-major-industry-sector.htm.

[3]Gyourko and Saiz (2006) use data from the R.S. Means Company. The construction costs include those for materials, labor, and equipment for four different qualities of single unit residences.

Table B.4: Regression of relative wage of construction workers

	Total	1 Unit	2 Units	3 and 4 Units	5 Units or More
	$log(RelativeWage)$	$log(RelativeWage)$	$log(RelativeWage)$	$log(RelativeWage)$	$log(RelativeWage)$
$log(Permits_{2017})$	-0.0225	-0.00255	-0.0101	-0.00107	-0.0235**
	(0.0145)	(0.0265)	(0.0192)	(0.0187)	(0.00919)
Constant	0.133	-0.0590	-0.0277	-0.0767	0.111
	(0.150)	(0.252)	(0.117)	(0.108)	(0.0811)
Observations	51	51	51	51	51
R-squared	0.041	0.001	0.007	0.000	0.070
			Lag of Building Permits		
	Total	1 Unit	2 Units	3 and 4 Units	5 Units or More
	$log(RelativeWage)$	$log(RelativeWage)$	$log(RelativeWage)$	$log(RelativeWage)$	$log(RelativeWage)$
$log(Permits_{2016})$	-0.0224	-0.00297	-0.0140	0.00812	-0.0228**
	(0.0156)	(0.0272)	(0.0225)	(0.0271)	(0.00923)
Constant	0.130	-0.0554	-0.00619	-0.124	0.106
	(0.160)	(0.257)	(0.134)	(0.152)	(0.0831)
Observations	51	51	51	51	51
R-squared	0.039	0.001	0.014	0.005	0.065

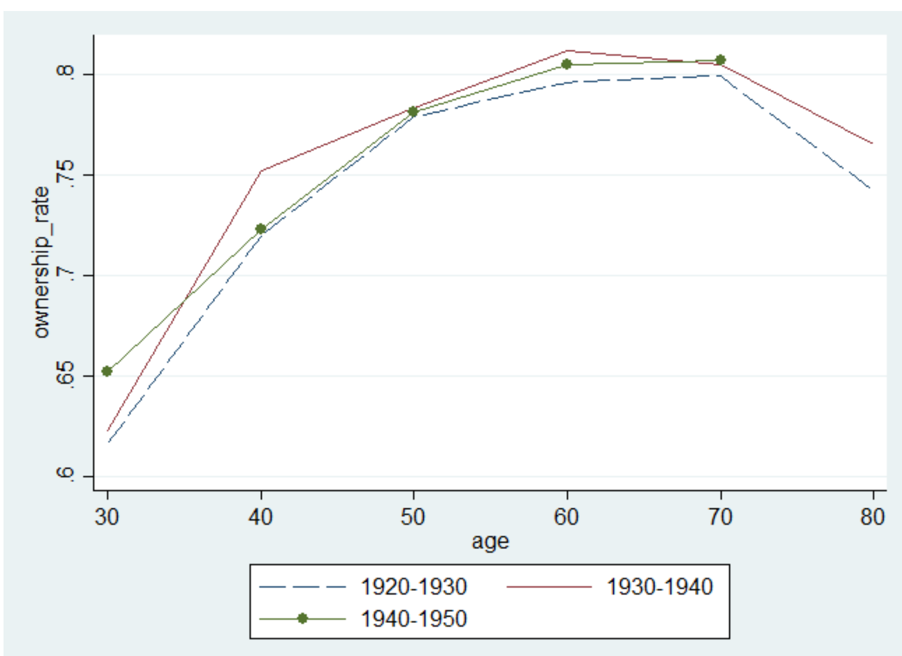
Appendix C

Appendices for Chapter 4

C.1 Cohort Profile

To distinguish the time effect from the cohort effect, we check the ownership profile for different cohorts born between 1920 to 1950 (Figure C.1). Consistent with the cross-sectional comparison, we see a significant increase in the ownership rate among senior households among recently born cohorts.

Figure C.1: Ownership Profile by Cohort



C.2 Construction of Population Structure without Immigration

This section describes the steps to construct the population structure in a scenario where there were no international immigrants since 1970. In each year t, the population of age a and sex x ($x = 1$ for male and $x = 2$ for female), $N^x_{a,t}$, can be calculated by the applying the age, sex, and time specific survival probability $S^x_{a-1,t-1}$ to the population size of age $a-1$ in the previous year for female and male separately.

$$N^x_{a,t} = N^x_{a-1,t-1} S^x_{a-1,t-1} \tag{C.1}$$

The number of new born, $N_{0,t}$, can be calculated by interacting the fertility rate fr_t at time t with the total number of female in their reproductive ages.

$$N_{0,t} = fr_t \sum_{a \in FertileAge} N_{a,t} \tag{C.2}$$

Among the new born, $\frac{1}{2.05}$ are girls and $\frac{1.05}{2.05}$ are boys given the male-female ratio at born is 1.05.

While the fertility rate is available for each individual year, the survival probability is only available every 10 years. For those years in between when the survival rate is not available, we adopt the linear interpolation.

C.3 List of States in the Sample

The balanced panel used in our estimation consists 32 states, including Alabama, Arkansas, California, Colorado, Florida, Georgia, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Nebraska, New Mexico, New York, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, South Carolina, Tennessee, Texas, Virginia, Washington, and Wisconsin.

C.4 Robustness: Model Fit Holding Income Profile Fixed in 1970

Section 4.4 examines the contribution of changing population structure and urbanization on the mean housing price growth from 1970 to 2010. We take the model in 1970 and feed in the population structure and urbanization rate in different years. At the same time, we adjust income profile accordingly. In this section we examine the contribution of changing income profiles on the housing price growth by fixing the income profile at the 1970 level. Figure C.2 shows the results. The dashed line coincides with the connected line, suggesting that change in income profile from 1970 to 2010 has not affected the housing market much.

Figure C.2: Model Fit with Income fixed in 1970

Bibliography

[1] Joseph Gyourko and Albert Saiz. Construction costs and the supply of housing structure. *Journal of regional Science*, 46(4):661–680, 2006.

[2] Joseph Gyourko, Albert Saiz, and Anita Summers. A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index. *Urban Studies*, 45(3):693–729, 2008.

[3] Katharina Knoll, Moritz Schularick, and Thomas Steger. No price like home: Global house prices, 1870-2012. *American Economic Review*, 107(2):331–53, 2017.

Curriculum Vitae

Publications:

Bilateral Liability-Based Contracts in Information Security Outsourcing (joint with Kailung Hui, Pingfan Ke, and Weithoo Yue), Information Systems Research., 30(2), 411-429, 2019.

Name:	Yuxi Yao
Post-Secondary Education and Degrees:	Xiamen University Xiamen, China 2007 - 2011 B.A. Economics 2008 - 2011 B.Sc. Applied Mathematics
	City University of Hong Kong Hong Kong SAR, China 2011 - 2012 M.Sc. Applied Economics
	University of Western Ontario London, Ontario, Canada 2013 - 2014 M.A. Economics 2014 - 2020 Ph.D. Economics
Honours and Awards:	Sir Arthur Currie Memorial Scholarship University of Western Ontario 2016
	Best Graduate Teaching Assistant of the Year University of Western Ontario 2014
	Graduate Fellowship University of Western Ontario 2013 - 2018
Related Work Experience:	Research Assistant City University of Hong Kong 2012 - 2013
	Teaching and Research Assistant The University of Western Ontario 2013 - 2020
	Internship Bank of Canada 2019