HAL
open science

Three essays on housing markets and housing policies

Zhejin Zhao

To cite this version:

Zhejin Zhao. Three essays on housing markets and housing policies. Economics and Finance. Université de Lyon, 2018. English. NNT : 2018LYSES033 . tel-02097496

HAL Id: tel-02097496

https://theses.hal.science/tel-02097496v1

Submitted on 12 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NNT : 2018LYSES033

UNIVERSITÉ DE LYON - ECOLE DOCTORALE SCIENCES ECONOMIQUES ET GESTION NN°486

UNIVERSITÉ JEAN MONNET

Groupe d'Analyse et de Théorie Economique

Thèse de Doctorat (NR) de Sciences Economiques

Présentée et soutenue publiquement par

Zhejin ZHAO

20 septembre 2018

en vue de l'obtention du grade de docteur de l'Université de Lyon
délivré par l'Université Jean Monnet

THREE ESSAYS ON HOUSING MARKETS AND HOUSING POLICIES

Jury :

Florence Goffette-Nagot	-	Directrice de recherche CNRS, Université de Lyon, Directrice de thèse
Frédéric Jouneau-Sion	-	Professeur, Université de Lyon 2, Examinateur
Miren Lafourcade	-	Professeur, Université Paris-Sud / Paris-Saclay, Examinatrice
Maëlys de la Rupelle	-	Maître de conférences, Université de Cergy-Pontoise, Examinatrice
Benoît Schmutz	-	Maître de conférences HDR, Ecole Polytechnique et ENSAE, Rapporteur

External reviewer :

Chunbing Xing	-	Professeur, Beijing Normal University (Chine), Rapporteur externe

University of Saint-Étienne is not going to give any approbation or disapprobation about the thoughts expressed in this dissertation. They are only the author's ones and need to be considered such as.

Acknowledgements

Finishing my PhD thesis is a milestone in my life. I am so lucky to have this great opportunity to pursuit a PhD. I would like to thank all those who have supported me in the completion of this thesis.

First and foremost, I would like to express my deepest gratitude to my advisor Florence Goffette-Nagot, for whom I am indebted the most during the PhD period. Before meeting her, I knew nothing about urban economics. It was she that introduced me into the field, which is very cool and powerful to explain many issues in urban areas. She has guided me throughout the past five years with so much care, dedication, patience and impressive intelligence, until the completion of my PhD thesis. During the course of this PhD, I not only got academic training from her, but also learned various traits of having a great personality. I have been impressed by her rigorous thinking, hard-working, academic enthusiasm, patience and responsibility. To be honest, she has changed me a lot. I am very appreciated to have such a kind and conscientious supervisor! I also thank her for her kind co-operation as a coauthor in the writing of the first chapter in this PhD thesis. I look forward to continue working with her after my thesis defense.

I would also like to express my gratitude to Russell Davidson for inviting me to visit the department of Economics, McGill University from November 2014 to May 2015. I learned quite a lot from his econometrics courses. I am also indebted to Saraswata Chaudhuri, Rohan Dutta and Victoria Zinde-Walsh for their permission to let me take part in their courses at McGill University. I wish to extend my special appreciation to Saraswata Chaudhuri who gave me helpful suggestions and encouragement for my first chapter. Many thanks to the PhD students at McGill including Tianyu He, Yang Li and Qi Xu who gave me a lot of support to help me adapt to the environment at McGill.

I would also like to thank the committee members. I am particularly grateful to Benoît Schmutz and Chunbing Xing for accepting to be reviewers on this thesis and to Frédéric Jouneau-Sion, Miren Lafourcade and Maëlys de la Rupelle for being the part of the jury. In addition, I want to thank Sylvie Démurger and Maëlys de la Rupelle as the committee members when I finished the fourth year of my PhD.

This PhD thesis would never haven been completed without the support of several institutions. I acknowledge University of Saint-Étienne and the Ministry of Higher Education, Research and Innovation of France for awarding me a thesis grant for my study in France. I am also grateful to the Region of Rhône-Alpes for awarding me the Explo'ra scholarship to support me to stay at McGill University in Canada. Last but not the least, I highly appreciate the generosity of the GATE Lyon-Saint-Etienne (GATE LSE) for supporting me to attend lots of academic conferences, workshops and summer schools, which is vital to my academic career.

I would like to thank Sonia Paty, the director of GATE, and Marie-Claire Villeval, the previous director of GATE, for providing PhD students with perfect working conditions and research environment. In the past five years, I have greatly benefited from faculties, fellow PhD colleagues and visiting scholars at GATE. First, I would like to say a big thanks to Sylvie Démurger who had introduced Florence Goffette-Nagot to me when I just arrived at Lyon as a master student in 2012. During the process of thesis writing, she gave me a lot of useful suggestions. She is always nice and helpful. Second, I have to mention Pierre-Philippe Combes, who always replied my emails promptly and helped me improve my thesis quality. Although he is quite busy, he spent a lot of time on listening to my presentation and taught me how to think about questions from the perspective of economic theory. I learned a lot from him. Third, I have to mention Samia Badji, Jonathan Goupille-Lebret, Tidiane Ly, Benjamin Monnery and Yohann Trouvé. I enjoyed from discussing with them and got valuable feedbacks on thesis. Fourth, a special thanks to Clément Gorin, who taught me a lot of programming skills, especially in R. In addition, I also wish to express my gratitude to the following persons for their help and advices: Sylvain Chareyron, Damien Cubizol, Frédéric Jouneau-Sion, Yann Kossi, Carl Lin, Philippe Pololmé, and Hui Xu.

I am thankful to administrative staff at GATE. I have to mention Nelly Wirth who always sends me a lot of papers closely related to my thesis, which is quite useful for my thesis writing. Many thanks to Taï Tao, Aude Chapelon, Yamina Mansouri and Béatrice Montbroussous for their supports.

I would also like to thank Jan Rouwenda for helpful suggestions on my second chapter when I participated the summer school in Copenhagen in August 2017. For many helpful comments, I thank the seminar, workshop or conference participants at GATE, the Regional Studies Association Early Career Conference 2015, the 8th Annual SEBA-GATE Workshop 2017, the 11th International Conference on the Chinese Economy 2017, AFSE 2018, and the RUSE workshop 2018.

Thanks to the BNU-GATE master exchange program. Without it, I would not have the chance to come to France. In April 2012, when I studied in the library in Beijing Normal University (BNU), professor Chengyu Yang asked me to come to his office. He told me that there was an exchange

program between BNU and GATE LSE. He strongly encouraged me to applying for it because he visited GATE before and had a good impression on it. Fortunately, I was admitted into the GAEXA master program. Here, I would like to thank especially Chengyu Yang.

As my thesis consists of three empirical chapters, the dataset are very important and indispensable for thesis writing. I thank Loïc Bonneval for sharing the historical data on Lyon and co-working with me in the first chapter. I thank Jianwei Xu for sharing the Chinese Urban Household Survey data, which is the main dataset in the second chapter. I thank Shi Li for sharing the 2000 and 2005 micro census data in China, which is used in the third chapter. I thank Zhilong Li from National Bureau of Statistics of China for helping calculating variable values at the city level in 2010 for the third chapter. I thank Liang Wenquan and Lu Ming for providing land transaction data, although finally I did not use this dataset in the third chapter. Special thanks to Miao Tong for calculating topographic data at the city level using GIS software, which is used to construct one of instrumental variables in the third chapter. Without their contributions, it would not be possible for me to finish the chapters.

I would also like to thank my friends in Lyon who have diversified my life and made me less stressful: Bin Bao, Zhixin Dai, Xinrong Huang, Jinping Li, Tiruo Liu, Lei Mao, Yifan Yang, Jin WU, Meng Wu, Han Zhou, Mengbing Zhu, Min Zhu and Guang Zhu etc. I also have a personal friendship with Tidiana Ly. We had dinner in restaurants regularly.

A big thanks for the support from many friends and researchers in China and Canada: Weiguang Deng, Chuanchuan Hou, Hongquan Lian, Dongran Sun, Dongdong Tian, Tianyu Wang, Shafu Zhang, Yifan Zhang etc. In particular, I am indebted much to Chuanyong Zhang for his trust, encouragement, and his dedication to help me improve my thesis quality.

I would like to thank Dongchen Zheng, the love of my life. Thanks for all that she has done for my family. When I encountered difficulties in research, she always encouraged me and tried to make me continue to move on. In order to make me focus on thesis writing, she did a lot of housework. I am so lucky to have a wife like this. In October 2017, my daughter, Kuoyu Zhao, was born in Lyon. She has been bringing me numberless happiness and huge motivation to pursuit my academic dream. I hope to be an good example for my dear daughter. In addition, I am always indebted for support from my family members, including my mother, my young sister, my parents-in-law and other relatives.

Finally, this thesis is dedicated to the memory of my father, Wusuo Zhao, who passed away in July 2013, four months before I started my PhD.

Contents

1 General Introduction **1**

2 The impact of rent control: investigations on historical data in the city of Lyon **7**

2.1 Abstract 7

2.2 Introduction 7

2.3 Related Literature 8

2.4 Rent control history in Lyon 12

2.5 Data and estimation method 13

2.5.1 Data 13

2.5.2 Descriptive statistics 14

2.5.3 Method 21

2.6 Results 22

2.6.1 The OLS and FE regression results 22

2.6.2 Autoregression problem 24

2.7 Conclusion 25

2A Appendix: Complete regression tables in three adjacent periods 25

3 Age, educational attainment and housing demand in urban China **31**

3.1 Abstract 31

3.2 Introduction 31

3.3 Empirical specification 34

3.4 Data 38

3.5 Regression results 44

3.6 Housing demand and aging 49

3.7 Conclusion 53

3A Appendix A: UHS survey introduction 54

3B Appendix: Descriptive statistics for year separately, 2007-2009 56

3C Appendix C: Unrestricted first-stage regression results, 2007-2009 58

4 Skill intensity ratio and housing prices across Chinese cities 59

4.1 Abstract 59

4.2 Introduction 59

4.3 Literature review 62

4.4 Administrative system in China 64

4.5 Specifications and hypothesis 65

4.6 Data 68

4.7 Descriptive statistics 70

4.8 Empirical results 73

4.8.1 OLS results 73

4.8.2 IV results 75

4.9 Robustness check 77

4.10 Conclusion and further discussions 77

4A Appendices 79

5 General conclusion **83**

Bibliography **87**

List of Tables

2.1 Summary statistics during the period 1890-1968 20

2.2 The regression results between two adjacent periods 23

2.3 Dynamic regression outputs in three adjacent periods 26

2.4 The effects of rent control on rents during the period 1890-1930 27

2.5 The effects of rent control on rents during the period 1914-1948 28

2.6 The effects of rent control on rents during the period 1930-1968 29

3.1 Distribution of the number of observations at the city level in the sample 39

3.2 Summary of annual rents by tenure 40

3.3 Characteristics of housing, household head and households in 2007 and 2009 42

3.4 The first hedonic regression results with homogeneous restrictions, 2007-2009 46

3.5 Selected results of second-stage hedonic regression for computed MWTP of housing characteristics, 2007-2009 47

3.6 Descriptive statistics in the UHS data, 2007-2009 56

3.7 The first hedonic regression results, 2007-2009 58

4.1 Summary statistics for variables in levels 71

4.2 OLS regression results in 2000 and 2010 74

4.3 IV regressions in 2010 76

4.4 First stage regression results in 2010 76

4.5 IV regressions after dropping four first-tier cities 77

A1 Distribution of the number of observations at the prefecture-level city in the micro censuses data 79

A2 Counties or county-level cities which upgraded to urban districts between 2000 and 2010 80

A3 Source of data used in this chapter 81

A4 Industry classification over the period 2000-2010 81

List of Figures

2.1 Ceilings on relative rent increases 13

2.2 The number of flats in each year during the period 1890-1968, Lyon 15

2.3 The number of tenants by categories in each year during the period 1890-1968, Lyon . 15

2.4 Duration and control status of flats during the period 1890-1968 17

2.5 Average rents of flats in Lyon, 1890-1968 18

3.1 The partial and global WTP for selected characteristics of houses along with age . . . 51

3.2 The WTP for a constant-quality house of a representative household by age group . . 53

4.1 Correlations between residuals and log population by year 72

4.2 Correlations among housing prices, population and SIR 72

A1 Example of a typical prefecture-level city 79

Chapter 1

General Introduction

The housing market is always a hot topic for people who live in urban areas and it attracts a lot of interest in both academia and industry. In this dissertation, I investigate three aspects of the housing market and housing policies. In the first essay, I am concerned about the role of government intervention in housing markets. Specifically, I examine the effects of a rent control policy on rents using historical panel data from Lyon, France. In the second essay, I am interested in factors which affect the housing markets. I use micro-level data to investigate how housing demand varies with age in urban China. In addition, I also want to know the impacts of rising housing costs on the local economy. In the third essay, therefore I use census data at the county level to examine the effects of housing costs on the skill intensity ratio across Chinese cities. The research findings about housing markets and housing policies not only explain the phenomenon that happened in France and China, but also apply to other countries and make contributions to the existing literature on housing economics.

The first essay, entitled "The impact of rent control: investigations on historical data in the city of Lyon", reexamines the conventional claims made by economists and policymakers concerning the effects of rent control on rents. As an important form of government intervention, rent control policy has a long history in European countries and the U.S. Generally speaking, there are two types of rent control policies. Rents can be always regulated, regardless of whether the tenant is new. This type of rent control is referred to as the first generation or old-style rent control. The second generation of rent regulation is less tight. Rents are controlled for only when tenants stay in the dwelling.

Theoretically, rent control has ambiguous effects on rents. Basu and Emerson (2003) demonstrate that monopolistic landlords have the motivation to hold rents down to attract a better 'quality' tenant if rents are controlled for. However, Nagy (1997) argues that, under the regulation of the second-generation rent control policy, landlords could set a higher price for new tenants at the beginning in

order to make up for a loss because of a rent freeze in later periods. Early (2000) and Diamond et al. (2018) think that, in the long run, rent control might decrease housing supply, and cause rents to rise even in the controlled sector.

Therefore, it is necessary to examine the effects of rent control on rents. In the literature, a lot of papers have addressed this issue using evidence from cities in the U.S. (See Gyourko and Linneman (1989), Nagy (1997), Sims (2007), Autor et al. (2014) and Diamond et al. (2018)) However, few papers use data from European countries to evaluate the impacts of rent control on rents. Mense et al. (2017) find that the rent cap implemented in Germany decreased rents and housing prices in the regulated sector, but increased them in the unregulated sector. The first essay of this thesis uses unique panel data at the flat level from a property manager's accounting books in Lyon during the period 1890-1968. The estimation results show that rents decrease when flats are controlled for. The first generation rent control policy have larger negative impacts on rents than the second generation.

The housing market is influenced by many factors. The demand-side factors, such as monetary policy, the state of the economy, household income and demographic structure, and housing supply, influence housing prices. Mankiw and Weil (1989) suggest that population aging can cause housing prices to decline in the long run. Ortalo-Magne and Rady (2006) show that the ability of young households to afford the downpayment on their first house and their income are powerful drivers of the housing market. Taylor (2007) claims that excessively expansionary monetary policy during the period from 2002 to 2005 is a cause of a bubble in the housing market in the U.S. Ferreira and Gyourko (2011) show that income of prospective buyers is a fundamental factor of the housing boom across U.S. metropolitan areas from 1993 to 2009. Switching to supply-side factors, Glaeser et al. (2005) and Saiz (2010) show that inelastic housing supply caused by physical and regulatory constraints have positive impacts on housing price appreciation across American cities since 1970s. Liang et al. (2016) argue that the misallocation of construction land supply between coastal and inland regions cause housing prices to rise faster in coastal cities than those in inland cities in China. Dong (2016) uses a sample of 35 major cities in China from 2003 to 2012 to find that both natural and man-made constraints imposed on housing supply are positively correlated to housing price appreciation.

With rising housing demand and limited housing supply, housing prices and rents will unavoidably rise. This is what happened in China in recent years. The high and rising prices in Chinese housing markets have generated global interest. Real prices increased by 160% between 1997 and 2011 according to China Statistics Yearbook. People want to know which factors cause such fast increase and whether the housing price will continue to rise in the future. Considering the difficulty of measuring housing supply at the city level in China, I focus on fundamental factors of housing demand. Specif-

ically, the second essay of this thesis studies how housing demand, quantity and quality of housing services, varies with age at the household level using micro-level data from urban China during the period 2007-2009. Since Mankiw and Weil (1989) explored the effects of demographic structure on housing markets and concluded that the housing demand in the U.S. will decline with population aging in the next 20 years, this question has attracted a lot of attention. However, previous literature does not isolate completely the role of demographic changes and other socio-demographic factors which may affect housing markets.

In the second essay of this thesis, I use the two-stage hedonic price model to investigate the correlation between housing demand and age. The data I use is the Chinese Urban Household Survey (UHS) from 2007 to 2009. This survey is conducted by the National Bureau of Statistics of China (NBSC). The NBSC uses a stratified random sampling to select cities and towns as surveyed regions in the UHS. What I could get access to is the subsample in 16 of 31 provinces from 2007 to 2009, which represents 65% of China's population(NBS, 2013). The UHS survey contains detailed characteristics of housing such as space floor area, housing age, rents and a lot of facilities, and that of households such as age, gender, marital status, income and educational attainment.

Specifically, I calculate the age specific implicit willingness-to-pay for a representative housing in the sample. In the first stage, I regress housing expenditures on all characteristics of a house to get the implicit price for each hedonic characteristic. In the second stage, I regress the calculated implicit prices on all characteristics of housing and socio-demographic characteristics of households as well as year dummies. Finally, I get the partial relationship between housing demand and age, because all socio-demographic characteristics except age have been controlled for. I can get the global relationships between housing demand and age if I do not control for any socio-demographic characteristics of households except age in the second stage, which means that the housing demand of households is allowed to change with characteristics of households. Considering that some socio-demographic characteristics such as marital status change with age naturally, and some characteristics such as educational attainment are constant over the life-cycle of households, I control for socio-demographic characteristics which change with age in the second stage, and then I got the third relationship, i.e. composite relationship between housing demand and age.

The partial and composite curves relating housing demand to age depict flat or slightly negative trends when people are older than 65. However, the global curve between housing demand and age shows that housing demand decreases with age. The overall results suggest that educational attainment largely drives the relationship between housing demand and age, while age does not have negative impacts for the housing demand as long as educational attainment is controlled for.

In the context of China's higher education expansion and rising urbanization rate, this second essay predicts that aggregate housing demand will not drop with population aging.

Because the UHS used in the second essay covers only three years, I cannot disentangle cohort effects from age effects. Instead, I have to assume that cohorts have similar preferences for housing. I have to test the validity of this implicit assumption if more datasets are available in the future. In addition, the UHS mainly focuses on households with local *hukou*, and migrants are under-represented. Hence the computed housing demand of a representative household might be overestimated.

After analyzing the determinants of housing prices, it is natural to think about consequences of a booming housing market. Considering that housing cost is one of the most important components of living cost across cities, what is the impact of housing costs on spatial dispersion of the labor force? The third essay applies the Rosen-Roback model and its extensions to investigate how housing costs affect the skill intensity ratio (SIR) of labor force across cities in China. The SIR is defined as the the ratio of college graduates to non-college graduates among adults aged 25 years old or above. Rosen (1979) creates a model of inter-city wage differences and makes a hypothesis that migration across metropolitan areas should cause equality of indirect utility. In the original model, Rosen (1979) assumes that all workers are homogeneous. The model suggests that wage differentials across cities arise due to relative price or amenity differentials associated with inter-city different characteristics including city size or growth rate. Kim et al. (2009) develop the models above by allowing for heterogeneity in housing preferences and human capital for workers, and they demonstrate that living costs of high- and low-skilled workers have different change rates, which is associated with changes of wage differentials between high- and low- skilled workers within one city.

Following their model, I use China's 2000 and 2010 censuses and corresponding data at the city level to test the model and show that rising average housing prices increase the ratio of skilled to unskilled workers in 2010, when workers' mobility was relaxed. However, effects of average housing prices on the SIR is only significant at the 10% level in 2000, when the *hukou* regulation was tight. As housing prices are endogenous, I use elasticities of housing supply based on land slope and historical housing prices as instruments of average housing prices. I also find that housing prices have significant positive effects on shares of high-skilled workers, but insignificant negative effects on shares of low-skilled workers in 2010 after taking into account the endogeneity issue. This essay confirms the validity of Rosen-Roback model.

This essay is closely related to the research by Broxterman and Yezer (2015), which examine the role of housing cost in determining the SIR across cities using data from the U.S. decennial census over the period 1970-2000. Unlike them, I use a novel instrument for housing prices and focus on the rapidly

rising real estate market using newly available data in China. Because the census survey does not have information on wages, I cannot estimate the effects of housing costs on the wage differences between high-skilled and low-skilled workers using a direct way.

To summarize, separately the three essays represent important additions to their respective literature. Together they help to understand the role of government intervention in housing markets, the deterministic factor of housing markets, and impacts of housing markets on the labor market.

Chapter 2

The impact of rent control: investigations on historical data in the city of Lyon[1]

2.1 Abstract

This chapter reexamines the conventional claims made by economists and policymakers concerning the effect of rent control. We consider the impact of rent control on rents using panel data in Lyon over a 78-years period. Our study is a comprehensive empirical study of different rent control forms using multiple regressions with fixed effects as the main form of analysis. We find that the causal effect between rents and rent control is significantly negative. Furthermore, more restrictive rent control policies cap rents more tightly.

2.2 Introduction

A new housing rent freeze law began to take effect in France in September 2012, which was presented by French Housing Minister Cécile Duflot in order to reduce housing rent increasing trend. This was the first step towards a new system of rent controls proposed by President François Hollande. As rents rise sharply in France, this decree will prevent any increase when a landlord lets a property for the first time or relets it. On the one hand, this regulation is able to cap increases in rents immediately, on the other hand, it discourages investments in the property sector, so that housing supply might be insufficient compared with housing demand in the long run. In this case, the real rents might be higher than if the rent control policy had never been implemented. This policy also has some impacts

[1] This chapter is a joint work with Loïc Bonneval and Florence Goffette-Nagot.

on maintenance expenditures of rental housing units. Landlords spend less on housing maintenance due to rent control, because they are not able to get the same prices as those in the uncontrolled markets. Their rational choice is therefore to decrease their inputs. However, Moon and Stotsky (1993) show that tenants have more incentives to invest in maintenance of their dwelling in order to compensate more or less under rent controls. Could the regulation significantly decrease rents? What are its impacts on maintenance costs? This paper attempts to reexamine the effect of rent control in order to answers these two questions.

Our study tests the effects of rent control on rents using historical data in the city of Lyon over a 78-years period. During the period 1914-1968, the first and the second generation rent control policies were implemented alternatively in Lyon. We find that both forms of regulation can make rents fall in most cases, and the stricter policy has larger effects on rents.

There are some important caveats to our results. First, due to the availability of data, we cannot obtain one third buildings' quality information, so that we have to drop other available buildings' quality information in order to use to the whole sample, which means we do not control all variables affecting rents and maintenance costs. Another limitation is that we only have data about maintenance costs at the building level instead of the flat level. We assume every flat's maintenance cost is decided according to its share of all floor area in the same building.

The reminder of this chapter is organized as follows. The next section presents the literature review. Section 3 describes the rent control history in Lyon and the dataset used. Section 4 introduces the methodology we use in order to process the data. Section 5 reports the results. Section 6 concludes.

2.3 Related Literature

Regulatory intervention in housing markets is broad and deep. Housing markets are governed by planning processes, zoning regulations, land use regulations, financial regulations and numerous other regulations (Turner and Malpezzi, 2003), among which rent control is the most important regulation historically (Gyourko, 2009) .

There are two types of rent control. The rents can be controlled even when the tenant changes. Conventionally, it is referred to as the first generation rent control. This is the type of control that was implemented in Lyon during the 1930-1948 period. This form of rent control was a restricted freeze on nominal rents, that is, the government set absolute ceilings on rents. So it is very strict. Conversely, rent control can work unless the tenant changes, which means rents are regulated only

within a tenancy. We often call this kind of rent control the second generation rent control policy.

Several studies indicate that rent control can reduce rents in the controlled sector. Gyourko and Linneman (1989) regard the constrained rents as a subsidy to the tenant. They find a mean annual subsidy in a rent controlled unit of 27.2% of annual income by analyzing New York's rent control system in 1968. Raess and von Ungern-Sternberg (2002) show that the second generation rent control can limit the owners' abilities to increase rents for a certain contract and leads to lower equilibrium rents, when price discrimination is caused by the existence of product heterogeneity, search costs and switching costs.

Basu and Emerson (2003) also study the effect of the second generation rent control on rents. They think its impact is similar to the first generation rent control policy. Because of inflation and information asymmetries, landlords prefer short-staying tenants to long-staying tenants, but they cannot distinguish which type the tenants are. The long-staying tenants have the incentive to conceal this information to prospective landlords. Considering this point, monopolistic landlords hold price down to attract a better 'quality' tenant (i.e. one who will stay short). Therefore, the second generation rent control can reduce rental levels in a way that mimics old-styled rent control policy. Sims (2007) estimates the effect of rent control in Massachusetts on the rent of renter-occupied apartments. He got the similar conclusion that rent control reduces rents substantially.

However, some studies question whether rent control policy lowers rents in the controlled sector. Nagy (1997) argues that landlords can set a higher price than the rents in an uncontrolled market in order to compensate the impact of the second generation rent control, because the rents will have to remain unchanged until the tenants change. However, as the time goes, the rents paid by tenants in the controlled sector will increase. Therefore, these regulations may change nothing except for altering the timing of rent increases. Nagy uses data from New York City to test for this hypothesis. The author finds that new tenants paid higher rents in controlled sector in New York City compared with those who occupied similar apartments in an uncontrolled sector in 1981. However, tenants in the same controlled sector paid less in 1987.

Some studies even argue that rent control results in higher rents in the controlled sector. Early (2000) shows that the likely long-run effect of rent control is to make the supply decrease and the cost rise, so in the long-run, rents may be higher even in the controlled market due to rent control. He uses the data from New York City in 1996 to test this hypothesis. The results suggest that tenants lost 44 dollars per month for households in rent stabilized apartments and 4 dollars per month for households in more strictly rent controlled housing. The tenants would have been better off in the controlled sector if rent control had never been implemented in New York City. Heffley (1998) also

argues that tenants may benefit from the abolishment of the rent control policy if landlords and tenants can change their economic and location decisions. Diamond et al. (2018) show that the rent control in San Francisco increased the probability of renters staying by 20%, reduced the supply of rental housing by 15%, and led to average rent increase by 5/1%. Finally, they concluded that the rent control policy caused a substantial welfare loss.

Rent control also has an impact on the rents in the uncontrolled sector. If rent control reduces the supply of housing, this will lead to a shortage in the whole housing market. Therefore, the rents will be higher in the uncontrolled markets, too. Tenants have to pay more due to the spillover effect of rent control. This is one of the main arguments against a rent control policy. Fallis and Smith (1984) use the data of Los Angeles, California during 1969-1978 to test the effect of rent control. They find that rent control effectively raised rents in the uncontrolled segments of the markets. Early (2000) draws the conclusion that the fraction of rental units under rent control is positively correlated with the pricing of rental housing in the uncontrolled sector by testing the data of New York City in 1996. Caudill (1993) estimates the effect of New York City's rent controls in 1968 by using both the ordinary least squares regression and frontier method. He finds that rents in the free sector would be lower by 22%-25% if control did not exist. Some economists also draw the same conclusion that rent control leads to higher rents in the uncontrolled sector (Early and Phelps, 1999; Navarro, 1985; Ho, 1992). Autor et al. (2014) find that the unanticipated elimination of rent control in Cambridge, Massachusetts, in 1995 raised housing values of both decontrolled and never-controlled residential properties.

However, the conclusion of effects of rent control on housing markets is still controversial. Considering general equilibrium effects, Hubert (1993) argues that rent control policy might decrease rents in the uncontrolled housing markets because of spillover effects. Rent control acts like a subsidy, which makes housing's demand excess in the controlled sector, so that the housing units' allocation is like a rationing system, which makes tenants reduce their housing consumption. Consequently, the demand also decreases in the uncontrolled sector, so rents might decrease. Heffley (1998) gets a similar conclusion using a spatial equilibrium model of rent control: when rent control is imposed, the rent does not rise in the uncontrolled sector. Although this result depends on the model specification and parameter values, it indicates that the external effect of rent control might be complicated when tenants can change economic and location decisions. Sims (2007) shows that rent control usually reduce maintenance levels. So uncontrolled rental housing located nearby will decrease in value due to the potential negative spillover effects. In this case, the rent might decrease in the uncontrolled sector.

Early and Phelps (1999) argue that the effect of rent control on the uncontrolled sector is ambiguous.

Rent control can reduce the supply of rental housing. They examine the effect of rent control on prices of uncontrolled housing markets by using 1984 to 1986 data from the American Housing Survey, and find that the price in the uncontrolled sector increased since the introduction of rent control. However, these effects decline through time and may disappear after several years.

There are also some debates about the effect of rent control on housing maintenance. Many economists think that landlords allow maintenance expenditures to fall because of the effect of rent control (Navarro, 1985; Albon and Stafford, 1990; Ault and Saba, 1990; Ho, 1992). Since the rent is controlled, the demand for housing excesses the supply. Landlords can make the value of housing fall in order to maximize their profits by decreasing their inputs in maintenance. Gyourko and Linneman (1989) find that a negative relationship between rent control and maintenance do exist. They draw this conclusion by using the data of New York City in 1968. They find that the impacts of rent control are lower in smaller and newer buildings. If the buildings are under ten years old, the impacts disappear. The biggest impacts happen in Manhattan, while the smallest happen in Queens. Under the second generation rent control, rents are controlled unless tenants change. Glaeser (2002) argues that even the second generation rent control gives strong incentives to landlords to keep housing units, the maintenance still is reduced and low quality apartments are created. Arnott and Shevyakhova (2014) argue landlords have incentives to prettify the housing only before the new tenants come in order to attract them, but few incentives to maintain the housing units well during the fixed-duration tenancy. Therefore, the second generation control leads to a decrease in maintenance.

However, Olsen (1988) thinks that the relationship between rent control and maintenance is theoretically ambiguous. Because "*rent control ordinances that increase the ceiling rent on an apartment generously when it is upgraded and decrease it severely when it is allowed to deteriorate will lead to greater landlord maintenance of the unit.*", tenants have incentives to spruce up the housing due to the income effect of rent control. But, the tenants may move before receiving all benefits of maintenance activity. So they are not able to get an unambiguous conclusion about the effect of rent control on maintenance. But Mengle (1985) argues that such rent control ordinances cannot prevent the problems. It is not easy for tenants and administrative agencies to observe the real dynamic maintenance costs change. Some sanctions against landlords who reduce maintenance expenditures do not work sometimes. As Mengle (1985) illustrated, Sims (2007) thinks that rent control does reduce the maintenance in the controlled units although it does not lead to big maintenance failures because the contract about maintenance level can not be complete. For example, the tenant can ask the landlord to repair if water and heat fail, but not for cracked paint.

Moon and Stotsky (1993) argue that landlords input less into maintenance under rent control. But

tenants in fixed long-term duration have more incentives to engage in self-maintenance in order to compensate more or less. They find little evidence that the first generation rent control causes housing quality changes by using data from New York City in the 1970s and 1980s. Kutty (1996) analyzes the effect of rent control on rental maintenance by establishing a dynamic model. She concludes that only the first generation rent control can reduce maintenance unambiguously, while rent control and maintenance may be positively correlated under the second generation rent control, in which rents are decontrolled between the tenancies.

2.4 Rent control history in Lyon

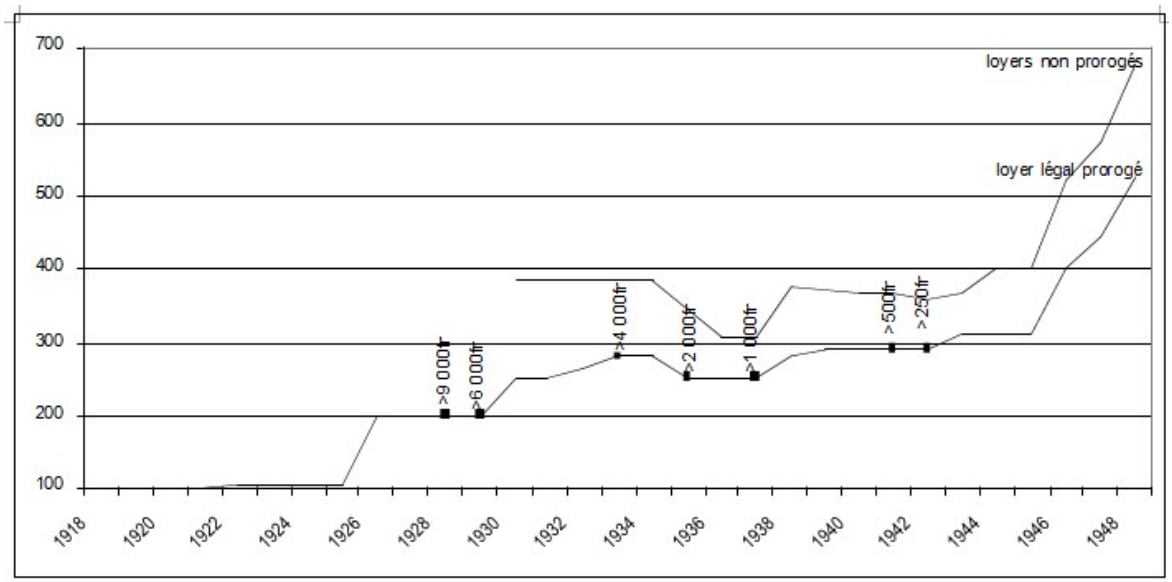
In 1914, due to the burden of World War One, renters were allowed not to pay their rent for periods up to 90 days if the rent was below a ceiling of 600 FF in Lyon. Based on data used in this chapter, this means that an annual average of 45% of housing rents remained unpaid between the second semester of 1914 and the first semester of 1920. This represents about 35% of total rents. After accounting for delayed payments, the total loss of landlords amounted roughly to 20% of their income.

Between the World Wars I and II, a special regime was put in place, in a context where the shortage of housing and the economic situation required to protect renters. A complex system of accumulating successive rules was implemented from 1919 to 1936. It consisted mainly of a ceiling on relative rent increases, as represented on Figure 1.1.

A new law was passed in 1948, which aimed at ending the special rent regime and at increasing the return of housing properties in order to favor housing construction and maintenance. A reference rent was computed based on the housing characteristics such as location, maintenance and quality. Bi-annual increases were applied in order to reach this reference rent by 1955. Continued leases ensured a capped increase of the rent. The incentive for tenants to stay in the apartment or to subcontract was high.

A first generation rent control system was implemented in Lyon during the 1930-1948 period. A second generation rent control policy was in place in Lyon during the periods 1914-1930 and 1948-1968. As we stated earlier, in the second generation rent control, rents can increase with a limited speed in order to ensure a reasonable return on investment to the landlord. Last, it is noted that some flats were out of rent control between 1928 and 1948 when the rent was higher than a certain limit as shown in Figure 2.1. For instance in 1928, rents above 9000fr were not controlled anymore. This limit changed 7 times. The last one is in 1942: rents above 250fr were not controlled anymore.

Figure 2.1: Ceilings on relative rent increases



Note: This figure is taken from Bonneval and Robert (2009). Index 100 corresponds to the rent in 1914. The "loyer légal prorogé" corresponds to leases starting before 1914. Flats went out of this regime progressively when a change of tenant occurred or when the rent reached a ceiling. These ceilings are represented on the figure. For instance, starting on 1928, flats with a rent above 9000 FF are not controlled anymore. Starting from 1930, all rents are capped excepted those which exceed a certain ceiling.

2.5 Data and estimation method

2.5.1 Data source

This study uses data collected from a property manager's accounting books by Bonneval and Robert (2009). The dataset gives information on flat and building characteristics for the period 1890 to 1968. Flat-level variables include whether the flat is used for housing or commercial use, whether the tenant is new or not, rent, number of rooms, floor area, storey, quality category and rent control status. Building-level variables include total rent and maintenance expenditures, location, construction type, construction period, number of floors and floor area. Monetary variables are corrected for inflation using coefficients drawn from Friggit (2002).

Our study focuses on flats used for housing. The original sample size was 32,745. We drop flats of which use is commercial or unknown, that is 8,264 observations. 1438 observations with missing information for floor area or number of rooms are dropped. Flats with missing information for building characteristics are also excluded, resulting in a further reduction in the sample size by 788. The final sample consists of 258 flats and 12,749 observations. We have to correct observations with missing

information on tenant mobility in the 1948-1968 period, that is, 43 observations. In these cases, we consider there was no tenant mobility.

Part of the data is given by semester and we have to aggregate it into yearly data. If one flat is controlled at least one semester, we suppose that the flat is controlled in the whole year. We combine the variable architecture type and construction period into interactive dummy variables, because both variables are highly correlated.

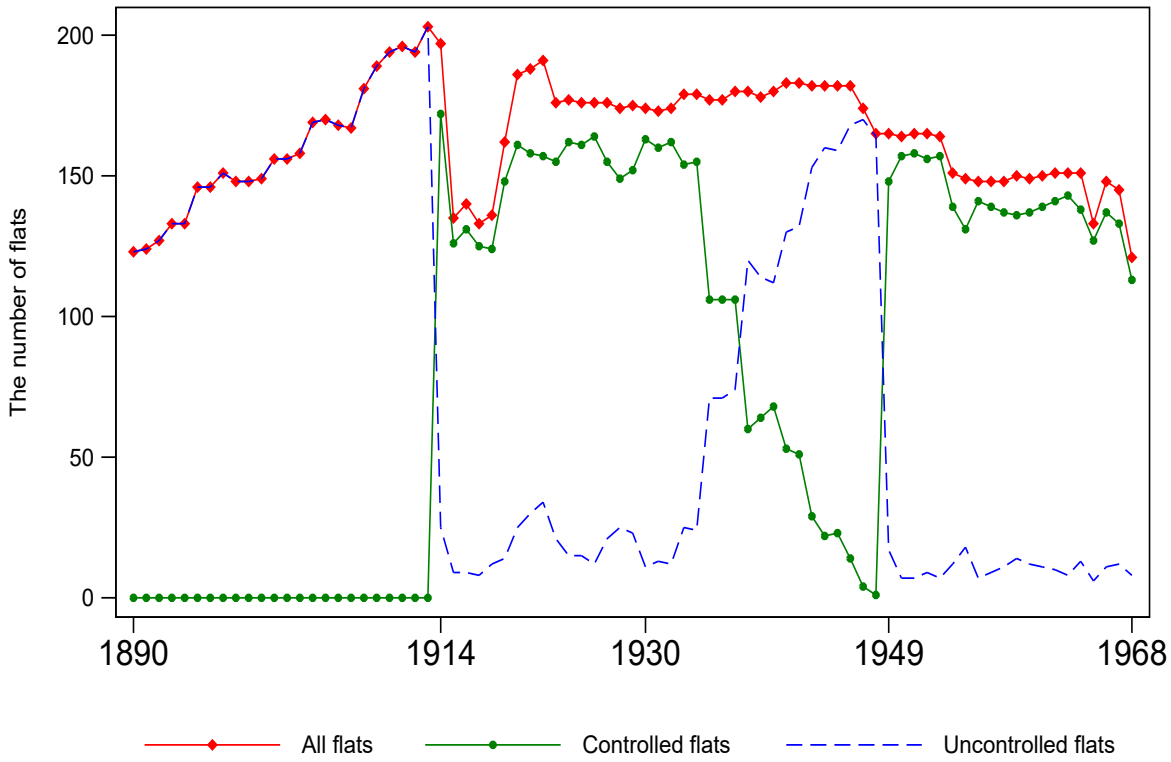
2.5.2 Descriptive statistics

Figure 2.2 plots the evolution of the number of flats in each year during the 1890-1968 period. Between 1890 and 1913, all observable flats were uncontrolled and the number of flats in the data increased from 123 to 200 at the peak of the whole period. This was followed by a reduction in the number of flats between 1914 and 1917, which means some flats disappeared from the survey, and a recovery in 1918 and 1919. The number of flats kept stable from 1920 to 1946. After that, it decreased slightly until 1968.

Starting in 1914, the rent control policy began to take effect and the number of uncontrolled flats plummeted. For example, in 1914, 169 flats were controlled, whereas only 25 flats were still uncontrolled. Reminding that a flat was controlled during 1914-1929 if the tenant changed, we can know that most tenants were the same as those in previous years in this year. The percentage of uncontrolled flats was 11.2% of all flats during 1914-1930. Then, the number of uncontrolled flats rose rapidly until 1948. It means many flats were out of control because the rent was higher than a certain limit.

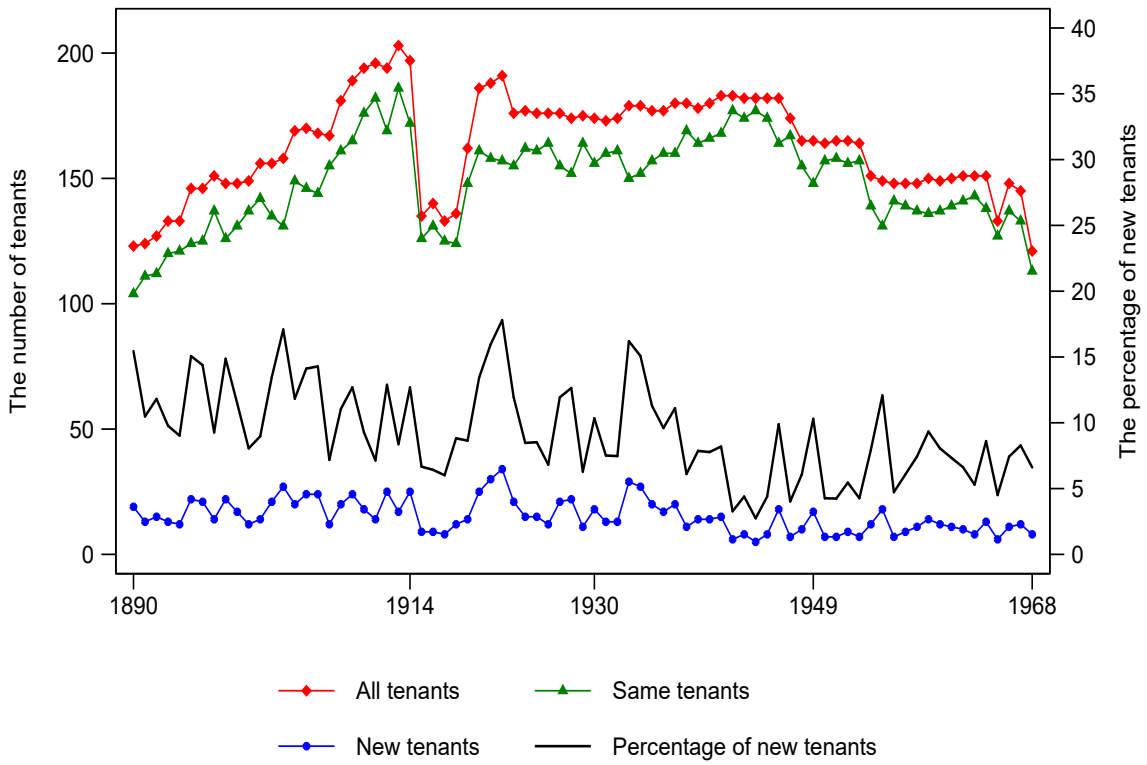
Since 1949, the rent control policy switched to the second generation rent control again and a flat was not out of control unless its tenant was new. The number of uncontrolled flats fell rapidly and then stayed low until the end of this survey. Figure 2.3 shows the number of new and same tenants by year during the period 1890-1968. We find that the share of new tenants is volatile but did not change much in the long run. When the rent control policy started in 1914, the share of new tenants decreased rapidly but then rose after few years. The same situation happened in 1949.

Figure 2.2: The number of flats in each year during the period 1890-1968, Lyon



Source: Loic Bonneval and Francois Robert (2009)

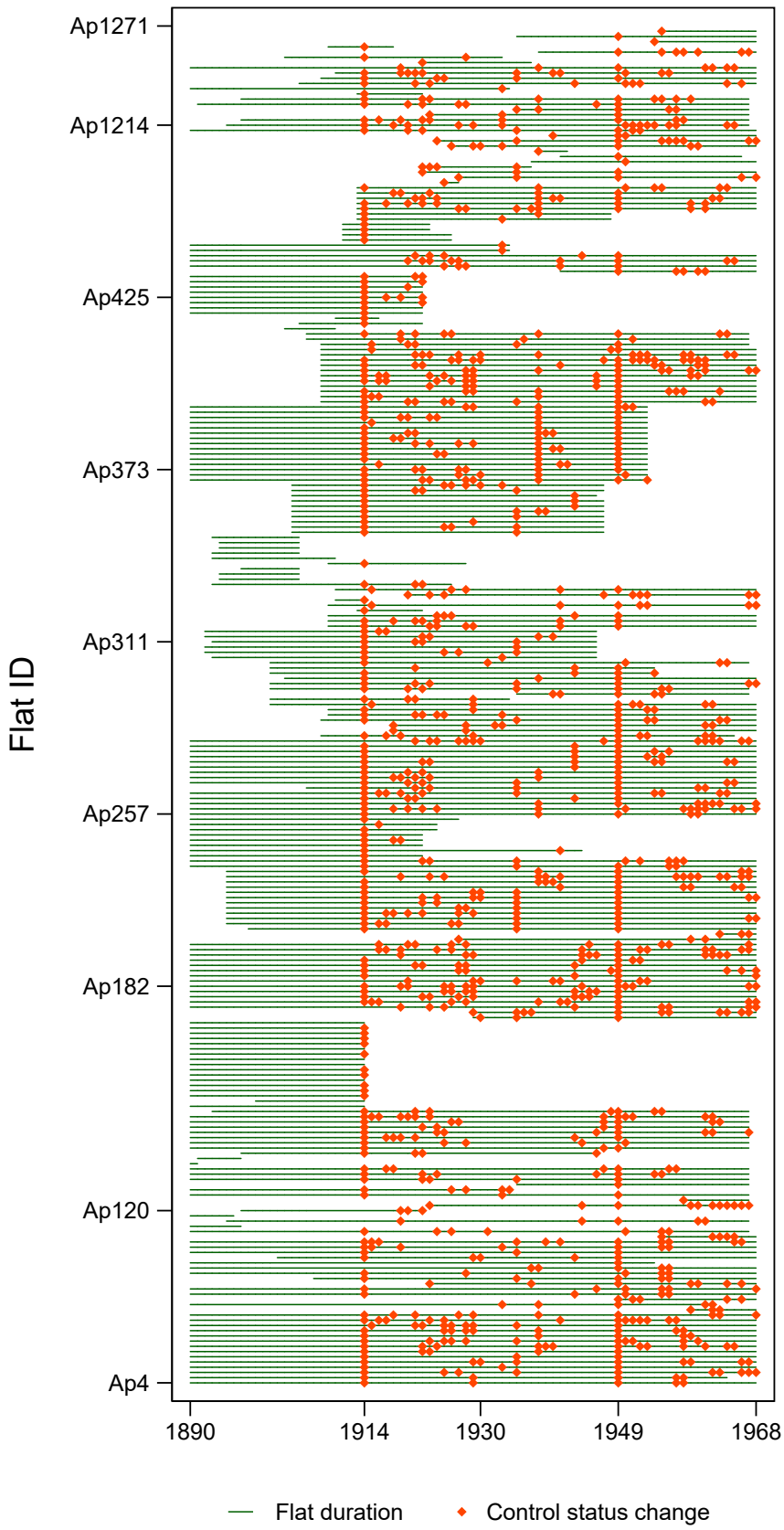
Figure 2.3: The number of tenants by categories in each year during the period 1890-1968, Lyon



Source: Loic Bonneval and Francois Robert (2009)

Figure 2.2 also shows some flats disappeared over time. Figure 2.4 shows this phenomenon more clearly. 48% of flats are present in 1890, the first year of the observation period, and 87% have been observed since at least 1914, while 3% of flats are present only after 1948. The situation is symmetric for the last year of each flat. 47% of flats were observed until 1968 and 66% until 1948, while less than 5% of flats disappeared before 1914. The average number of years a flat is present in the database is 52 years; but it should be noted that the observations are not continuous in 46% of the cases. 54% of flats are observed during the whole period, while the average number of missing years for all flats is 2.7. In addition, 13% flats miss one year and 6% flats miss two years. It should be noted that some flats reappeared after missing many years. So the panel data we use is unbalanced.

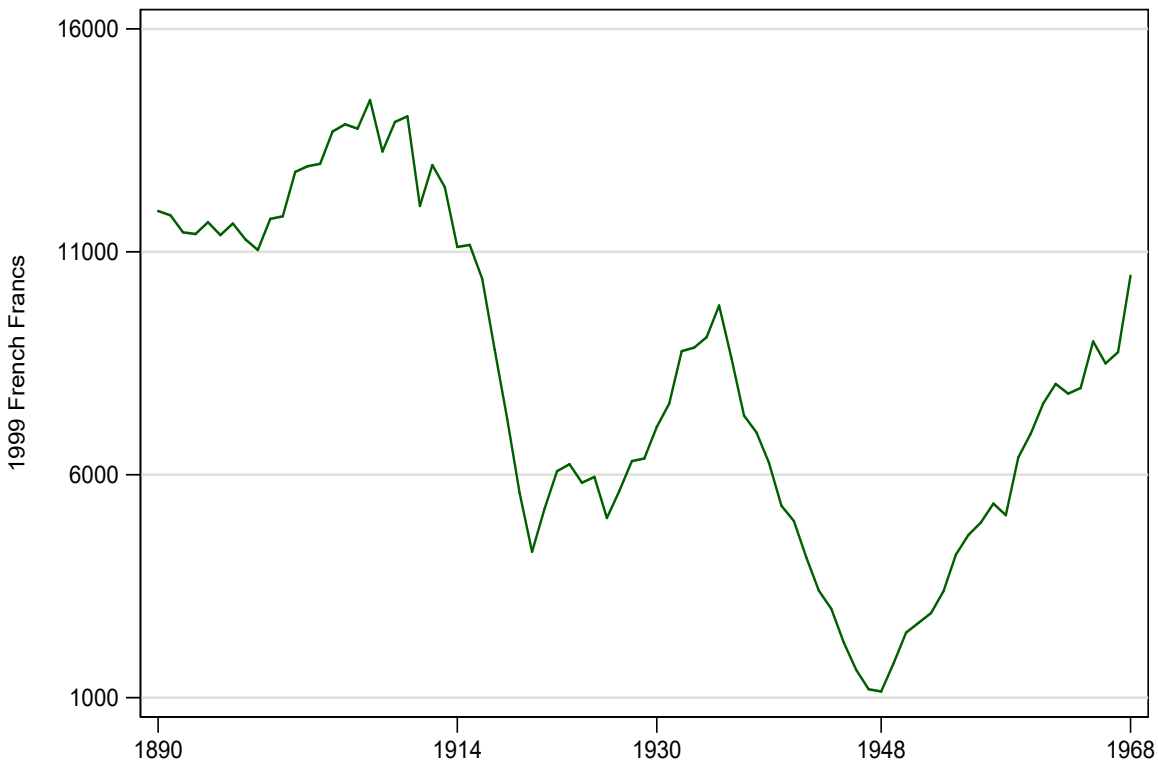
Figure 2.4: Duration and control status of flats during the period 1890-1968



Notes: The red dot means that the control status changed
from uncontrolled to controlled or from controlled to uncontrolled.

Table 2.1 presents descriptive statistics for the variables used in this study by time period and rent control status respectively. Because there was no rent control in Lyon between 1890 and 1914, we just list the variables in the uncontrolled sector. For convenience, all currency amounts are expressed in 1999 French Francs. In the periods 1914-1929 and 1930-1948, the rents in the controlled sector are on average lower than those in the uncontrolled sector. Rents experienced a sharp volatility from 1928 to 1948, as shown in Figure 2.5. In the later period 1949-1968, the rents in the controlled sector are higher than those in the uncontrolled sector. However, the characteristics of dwellings in the controlled and uncontrolled sectors change over time. In the last period 1948-1968, under the second-generation rent control, flats in the controlled sector have more rooms and are at a lower floor. However, in the second period 1914-1929, the controlled sector is further away from the center of Lyon. During the period 1930-1948, under the first generation regulation, the controlled sector is closer to the center of Lyon, but controlled flats have less rooms and are at higher floors. Both dwelling characteristics and rent control impact rents and maintenance expenditures, which means housing characteristics have to be controlled for to test the effect of rent control.

Figure 2.5: Average rents of flats in Lyon, 1890-1968



Source: Loic Bonneval and Francois Robert (2009)

What we are interested in is the relationship between rent control and rents during the whole period. Before 1914, the change of average rents is not large although average rents increased from 11,011 to

15,244 French Francs between 1898 and 1907, and then fell to 12,594 French Francs in 1913 (Figure 2.5). Since 1914, with the implementation of the second generation rent control policy, the average rent dropped drastically until 1919. Then it stayed relatively stable with low volatility until 1926, and then increased again.

In 1930, the stricter rent control regulation began to take effect, while some flats were out of rent control between 1928 and 1948 when the rent was higher than a certain limit. From 1930 to 1935, rents increased rapidly. Beginning in 1936, the rental housing market in Lyon plummeted and the mean rent reached the lowest level in 1948. Rents increased again since then. We hypothesize that the rent control policy partly explains the rent trend in Lyon. We must disentangle the effect of rent control and other factors which can influence rents.

Table 2.1: Summary statistics during the period 1890-1968

	1890-1913	1914-1929		1930-1948		1949-1968	
	Uncontrolled	Controlled	Uncontrolled	Controlled	Uncontrolled	Controlled	Uncontrolled
Rent	12,611	6,834	7,014	4,494	6,618	5,852	5,395
	(13,222)	(8,672)	(9,265)	(3,293)	(8,964)	(5,942)	(5,191)
The number of rooms	3.29	3.55	3.37	2.82	4.06	3.56	3.19
	(1.63)	(1.82)	(1.83)	(1.24)	(1.88)	(1.91)	(1.69)
Surface per room	22.97	22.04	21.49	20.52	22.72	22.01	22.49
	(10.24)	(9.67)	(8.19)	(8.33)	(7.75)	(8.09)	(9.19)
The number of rooms							
Ground floor or 1st	0.17	0.21	0.18	0.14	0.25	0.18	0.24
	(0.38)	(0.41)	(0.38)	(0.35)	(0.44)	(0.38)	(0.43)
2nd-4th Floor	0.58	0.59	0.58	0.52	0.63	0.59	0.44
	(0.49)	(0.49)	(0.49)	(0.50)	(0.48)	(0.49)	(0.50)
5th-6th Floor	0.25	0.20	0.24	0.34	0.12	0.23	0.32
	(0.43)	(0.40)	(0.43)	(0.47)	(0.32)	(0.42)	(0.47)
Construction type and period							
Haussmannian*1871-1914	0.39	0.41	0.36	0.37	0.45	0.42	0.38
	(0.49)	(0.49)	(0.48)	(0.48)	(0.50)	(0.49)	(0.49)
Ancient*1871-1914	0.20	0.16	0.11	0.11	0.16	0.10	0.06
	(0.40)	(0.36)	(0.31)	(0.31)	(0.36)	(0.30)	(0.24)
Haussmannian*before 1871	0.10	0.11	0.06	0.09	0.12	0.10	0.14
	(0.30)	(0.31)	(0.24)	(0.29)	(0.33)	(0.29)	(0.35)
Ancient*before 1871	0.31	0.33	0.47	0.43	0.27	0.39	0.42
	(0.46)	(0.47)	(0.50)	(0.50)	(0.45)	(0.49)	(0.49)
District							
District 1st	0.19	0.24	0.37	0.34	0.27	0.36	0.41
	(0.39)	(0.43)	(0.48)	(0.47)	(0.44)	(0.48)	(0.49)
District 2nd	0.42	0.37	0.30	0.34	0.32	0.28	0.30
	(0.49)	(0.48)	(0.46)	(0.48)	(0.47)	(0.45)	(0.46)
District 3rd	0.18	0.14	0.17	0.11	0.16	0.16	0.09
	(0.38)	(0.35)	(0.37)	(0.31)	(0.37)	(0.36)	(0.29)
District 5th	0.08	0.07	0.08	0.09	0.06	0.09	0.07
	(0.27)	(0.26)	(0.26)	(0.28)	(0.24)	(0.28)	(0.25)
District 6th	0.13	0.18	0.09	0.12	0.19	0.12	0.13
	(0.34)	(0.38)	(0.28)	(0.32)	(0.39)	(0.32)	(0.33)
Distance to the city center	759	787	749	769	790	742	680
	(370)	(360)	(335)	(349)	(350)	(311)	(305)
Observations	3807	2365	291	1482	1845	2753	206

2.5.3 Method

Methodologically, we exploit the rent control status variation of dwellings in our database and use instrumental variables to identify the causal effect of rent control on rents at the flat level. The basic regression model in levels is the following:

$$Iny_{it} = \alpha_i + \lambda_t + \beta_1 rc1_{it} + \beta_2 rc2_{it} + \gamma x_i + \varepsilon_{it} \tag{2.1}$$

where $\ln y_{it}$ represents the logarithm of annual rent of flat i in the year t, α_i a flat-specific intercept, λ_t is a vector of year dummies, $rc1_{it}$ takes value 1 for a flat controlled by the first generation (strong) rent control policy, $rc2_{it}$ takes value 1 for a flat which is controlled by the second generation (moderate) rent control policy, x_i are flat's characteristics.

The flat fixed effects allow for time-constant, flat-specific factors that affect the level of rents, while the year dummies control for time variations in the level of rents.

We then write the model in first-differences, so that unobserved time-invariant factors that are specific to each flat and affect the level of rents are differenced out:

$$\Delta Iny_{it} = \Delta\lambda_t + \beta_1 \Delta rc1_{it} + \beta_2 \Delta rc2_{it} + \varepsilon_{it} \tag{2.2}$$

In some additional specifications, we control extra variables, such as building-level dummies, building-specific linear time trends, different time trends for different districts within Lyon, or an interaction between the construction type (or period) and a time trend.

Throughout all models, we cluster standard errors by buildings. To address the possible existence of autocorrelation and moving averages, I also show regression results using maximum likelihood estimation of a model with ARMA. The main coefficient of interest is β. It can be interpreted as the causal effect of rent control on the flat rent.

Despite controlling for building trends and several variables at the flat level, estimation of β in regression model (2) by ordinary least squares (OLS) may suffer an endogeneity bias. Because rent control status is not randomly assigned to housing units, it may be correlated with some omitted variables. For example, in the period 1914-1927, the flat was controlled unless the tenant was new. Suppose that, for some reasons such as the improvement of local amenities, a flat becomes more attractive during this period. Consequently, the original tenant would continue to stay here or a new tenant would have to pay a higher rent. This would induce an upward bias in the OLS estimates of

β . Another problem is the reverse causation. For example, in the years 1928 and 1929, either if the rent exceeded a certain limit or the tenant was new, the flat was out of control. If the flat rent was relatively high, the tenant would be likely to move out. This means that the magnitude of rents can change the rent control status during this period.

2.6 Main regression results

There are three types of rent control policies in Lyon in the whole period. The rent ceiling policy, which was implemented from 1928 to 1948, overlapped with other rent control measures. The government carried out the second generation rent control policy in 1928 and 1929 as well as the first generation thereafter until 1948. In 1928 and 1929, there were only 16 observations of 12 flats which were uncontrolled by the second generation rent control policy, while there were 343 observations for 173 flats in 1928 and 1929, as well as 2,261 observations of 221 flats in 1913-1929. Therefore, we can roughly attribute the effect of rent control to the second generation rent control policy during the 1913-1929 period. From 1930 to 1948, both the first generation rent control policy and rent ceiling policy were implemented, so that we can regard the effect of rent control as their aggregated effect.

Therefore, the whole period 1890-1968 is divided into four periods: 1890-1913, 1914-1929, 1930-1948 and 1949-1968. In order to avoid very heterogeneous economic contexts, we estimate the effect of rent control between two adjacent periods one by one.

2.6.1 The OLS and FE regression results

Table 2.2 shows the main regression results between two adjacent periods. The control variables contain building type, building construction period, average surface per room, number of rooms, floor, distance to the city center and year dummies in the OLS regressions and year dummies in the FE regressions.[2]

[2]See the complete regression tables in Appendix A1-A3.

Table 2.2: The regression results between two adjacent periods

Period	Explanatory Variable	OLS	FE	RE
	1st generation rent control			
1890-1914-1930	2nd generation rent control	-0.119***	-0.111**	-0.113**
		(0.037)	(0.028)	(0.028)
	Observations	6463	6463	6463
	Adjusted R^2	0.712	0.650	0.708
	1st generation rent control	-0.390***	-0.062***	-0.076***
		(0.053)	(0.021)	(0.023)
1914-1930-1948	2nd generation rent control	-0.110***	-0.090***	-0.093***
		(0.038)	(0.023)	(0.023)
	Observations	5983	5983	5983
	Adjusted R^2	0.803	0.794	0.779
	1st generation rent control	-0.437***	-0.187***	-0.199***
		(0.059)	(0.044)	(0.045)
1930-1948-1968	2nd generation rent control	-0.100***	-0.094***	-0.093***
		(0.039)	(0.023)	(0.023)
	Observations	6286	6286	6286
	Adjusted R^2	0.865	0.871	0.858

Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Control variables contain building type, building construction period, average surface per room, number of rooms, floor, distance to the city center and year dummies in the OLS and FE regressions.

Firstly, we get the results for the effect of rent control in 1890-1929. In the first period 1890-1914, there was no rent control. Starting from 1914, rents were controlled unless tenants changed. The estimated coefficient for our main explanatory variable is -0.151 in the basic specification and it is highly significant. The estimated coefficient in the fixed effect specification is -0.129. Based on these estimates, the coefficient of interest is around -0.140. [3]

Then we estimate the effect of regulation in 1914-1948. In the first period 1914-1929, rents were controlled unless tenants changed. After 1929, the rent control policy became stricter. Even if the tenant

[3]The average surface per room is positively correlated with rents. The higher the number of rooms, the higher the rent. Similarly, the higher the floor is, the lower the rent. Distance to the city center does not have any significant impact on rents.

were new, rents were still controlled unless rents exceed one certain ceiling. The second column in Table 2 represents regression results in this period. We can find that the effect of the second generation rent control is still negative and significant, but it is smaller in 1914-1948 compared to the previous period. Perhaps the reason is that many flats with high rents were out of control in 1930-1948. The OLS regression shows that the first generation rent control is able to decrease rents by 46.8%. However, the magnitude of the effect of the rent control policy is lower in the FE estimation, which means unobserved variables correlated to the rent control status influence rents.

Lastly, we examine the relationship between rent control and rents in 1930-1968. In the first period 1930-1948, rents were controlled even after a change of tenant. After 1948, a new rent control law was passed. If when tenant changes, the rent is not controlled anymore. In this period, the FE regression results show that the effect of the first generation rent control policy is twice larger in magnitude compared to the second generation rent control policy. The characteristics of housing units have similar influence as in the previous periods.

In sum, both rent control policies can decrease rents, while the first generation (strong) rent control has bigger impacts. It is noted that flats' unobserved characteristics have substantial impacts on the estimation.

2.6.2 Autoregression problem

In the long period, error terms of the estimated model may be correlated serially. In fact, the correlation coefficient between $\ln rent$ and lagged $\ln rent$ is above 0.9. Considering the possible autoregression problem, we use FGLS with AR (1) to reestimate the effect of rent control on rents, with the following specification:

$$\ln y_{it} = \alpha_i + \rho \ln y_{i,t-1} + \lambda_t + \beta_1 rc1_{it} + \beta_2 rc2_{it} + \gamma x_i + \varepsilon_{it} \tag{2.3}$$

where lny_{it} represents logarithm of rents, ρ the effects of lagged logarithm of rents, α_i a flat-specific intercept, λ_t year dummies, $rc1_{it}$ which takes value 1 if the flat is controlled by the first generation (strong) rent control policy, $rc2_{it}$ which takes value 1 if the flat is controlled by the second generation (moderate) rent control policy, x_i flat characteristics.

The equation is the same as the random effect linear model with an an AR(1) disturbance. Here we suppose there is correlation between errors and unobservable variables.

Table 2.3 shows that both types of rent control policy have negative impacts on rents in Lyon, and

the first generation rent control policy has larger impacts compared to the second. The coefficients of control variables are as expected. The old building with old style have negative influences on rents. Rents increase with the number of rooms. Flats at higher floors have lower rents.

2.7 Conclusion and further studies

Few studies have been carried out in Europe to test for the impact of a rent control policy. Our results complement a recent literature that aims to understand the consequences of rent control implemented in Europe. In our study, we estimate the effect of a restrictive regulation (the first generation rent control) and a moderate regulation (the second generation rent control) in the housing market. In the rent control history of Lyon, regulation forms changed from being moderate to being restrictive and then move back to being moderate. This article finds that both forms of regulation can make rents decrease in most cases.

However, due to the limitation of the data, all flats were controlled at least one year during the observation period. Therefore, we could not observe the externalities of rent control on the uncontrolled rental market. As our results show, the rent control policy reduces rents in the controlled sector. This will lead to a decrease of supply in the whole housing market. It might be that rents will be higher in the uncontrolled part of the market due to spillover effects.

This is an ongoing work and we have not finished estimating the effect of rent control on rents, not to mention the effect of rent control policy on maintenance costs.

In further research, we aim at extending this analysis. First, endogeneity of rent control for a given flat could be dealt with with instrumental variables. Second, we would like to use complementary databases to control macroeconomic variables such as the annual interest and average income in each district of Lyon.

2A Appendix: Complete regression tables in three adjacent periods

Table 2.3: Dynamic regression outputs in three adjacent periods

	(1) 1890-1914-1930 β / SE	(2) 1914-1930-1948 β / SE	(3) 1930-1948-1968 β / SE
Actually controlled under the 1st generation	0.000	–0.151***	–0.145***
	(.)	(0.022)	(0.023)
Actually controlled under the 2nd generation	–0.152***	–0.127***	–0.013
	(0.023)	(0.027)	(0.022)
The number of rooms	0.286***	0.254***	0.273***
	(0.026)	(0.022)	(0.023)
Surface per room	–0.000	0.010*	0.023***
	(0.004)	(0.006)	(0.005)
2nd-4th Floor	–0.166**	–0.233***	–0.084
	(0.068)	(0.068)	(0.057)
5th-6th Floor	–0.466***	–0.665***	–0.483***
	(0.104)	(0.101)	(0.072)
Construction type and period			
Ancient*1871-1914	–0.190**	–0.116	–0.080
	(0.094)	(0.075)	(0.084)
Haussmannian*before 1871	0.448*	0.231	0.335***
	(0.233)	(0.199)	(0.128)
Ancient*before 1871	–0.590***	–0.684***	–0.308***
	(0.160)	(0.163)	(0.102)
District			
District 2nd	–0.447**	–0.493***	–0.412***
	(0.194)	(0.157)	(0.114)
District 3rd	–0.359*	–0.133	0.022
	(0.205)	(0.132)	(0.092)
District 5th	0.205	0.259**	0.183**
	(0.149)	(0.109)	(0.084)
District 6th	–0.316	–0.462***	–0.388***
	(0.210)	(0.165)	(0.131)
Distance to the city center	–0.000	–0.000	0.000
	(0.000)	(0.000)	(0.000)
Year	Yes	Yes	Yes
Observations	3721	3761	4887

Note: Standard errors in parentheses *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

Table 2.4: The effects of rent control on rents during the period 1890-1930

	(1) OLS	(2) FE	(3) RE
	β / SE	β / SE	β / SE
Actually controlled under the 2nd generation	–0.119***	–0.111***	–0.113***
	(0.037)	(0.028)	(0.028)
The number of rooms	0.305***		0.295***
	(0.042)		(0.031)
Surface per room	0.009		0.008
	(0.006)		(0.006)
2nd-4th Floor	–0.156***		–0.210***
	(0.053)		(0.065)
5th-6th Floor	–0.527***		–0.589***
	(0.072)		(0.083)
Construction type and period			
Ancient*1871-1914	–0.106		–0.077
	(0.141)		(0.164)
Haussmannian*before 1871	–0.041		–0.006
	(0.328)		(0.334)
Ancient*before 1871	–0.339*		–0.343**
	(0.179)		(0.160)
District			
District 2nd	–0.143		–0.209
	(0.198)		(0.224)
District 3rd	–0.157		–0.280
	(0.216)		(0.219)
District 5th	0.291**		0.279**
	(0.120)		(0.116)
District 6th	0.029		–0.055
	(0.191)		(0.199)
Distance to the city center	–0.000		–0.000
	(0.000)		(0.000)
Year	Yes	Yes	Yes
Observations	6463	6463	6463
R^2	0.712	0.650	0.708

Note: Standard errors in parentheses *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. Reference groups are 0-1st floor, Ancient*before 1871, District 1st.

Table 2.5: The effects of rent control on rents during the period 1914-1948

	(1) OLS	(2) FE	(3) RE
	β / SE	β / SE	β / SE
Actually controlled under the 1st generation	–0.390***	–0.062***	–0.076***
	(0.053)	(0.021)	(0.023)
Actually controlled under the 2nd generation	–0.110***	–0.090***	–0.093***
	(0.038)	(0.023)	(0.023)
The number of rooms	0.263***		0.293***
	(0.038)		(0.037)
Surface per room	0.015***		0.008
	(0.005)		(0.006)
2nd-4th Floor	–0.155***		–0.234***
	(0.052)		(0.066)
5th-6th Floor	–0.570***		–0.672***
	(0.050)		(0.091)
Construction type and period			
Ancient*1871-1914	–0.042		–0.096
	(0.071)		(0.155)
Haussmannian*before 1871	0.013		–0.018
	(0.232)		(0.328)
Ancient*before 1871	–0.447***		–0.385*
	(0.147)		(0.214)
District			
District 2nd	–0.273		–0.238
	(0.167)		(0.233)
District 3rd	–0.053		–0.305
	(0.131)		(0.222)
District 5th	0.284***		0.209*
	(0.082)		(0.113)
District 6th	–0.180		–0.225
	(0.150)		(0.212)
Distance to the city center	–0.000		–0.000
	(0.000)		(0.000)
Year	Yes	Yes	Yes
Observations	5983	5983	5983
R^2	0.803	0.794	0.779

Note: Standard errors in parentheses *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. Reference groups are 0-1st floor, Ancient*before 1871, District 1st.

Table 2.6: The effects of rent control on rents during the period 1930-1968

	(1) OLS	(2) FE	(3) RE
	β / SE	β / SE	β / SE
Actually controlled under the 1st generation	–0.437***	–0.187***	–0.199***
	(0.059)	(0.044)	(0.045)
Actually controlled under the 2nd generation	–0.100**	–0.094***	–0.093***
	(0.039)	(0.023)	(0.023)
The number of rooms	0.253***		0.261***
	(0.038)		(0.038)
Surface per room	0.021***		0.021***
	(0.004)		(0.004)
2nd-4th Floor	–0.096*		–0.114**
	(0.053)		(0.052)
5th-6th Floor	–0.434***		–0.553***
	(0.037)		(0.062)
Construction type and period			
Ancient*1871-1914	–0.095		–0.066
	(0.065)		(0.056)
Haussmannian*before 1871	0.293*		0.317**
	(0.163)		(0.155)
Ancient*before 1871	–0.323***		–0.433***
	(0.097)		(0.121)
District			
District 2nd	–0.388***		–0.499***
	(0.103)		(0.109)
District 3rd	0.025		–0.069
	(0.098)		(0.103)
District 5th	0.229***		0.247***
	(0.063)		(0.073)
District 6th	–0.271**		–0.442***
	(0.120)		(0.135)
Distance to the city center	0.000		0.000*
	(0.000)		(0.000)
Year	Yes	Yes	Yes
Observations	6286	6286	6286
R^2	0.865	0.871	0.858

Note: Standard errors in parentheses *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. Reference groups are 0-1st floor, Ancient*before 1871, District 1st.

Chapter 3

Age, educational attainment and housing demand in urban China

3.1 Abstract

China is rapidly aging and experiencing a booming real estate market, so people are concerned about if total housing demand will decrease because of population aging in the future. To address this issue, this chapter explores how housing demand varies with age using micro-level data from urban China in the period 2007-2009. The results show that the willingness-to-pay for a constant-quality house will decrease slightly or keep constant after household heads become old, when educational attainment is controlled for. They imply that educational attainment is one of deterministic factors on housing demand. Therefore the total housing demand will not decline although population aging, because the current middle-aged generation has higher educational attainment than the current old generation. In the context of China's higher education expansion and fast urbanization, this chapter predicts that aggregate housing demand will not drop with population aging.

3.2 Introduction

In the recent years, China has witnessed rapidly rising housing prices. The residential housing prices have been rising by 12% annually from 2003 to 2011 in the 35 major Chinese cities, which surpasses the fast-growing per capita GDP in the same period. During this period, China's population has been aging rapidly. According to China 2010 census data, the proportion of people above 60 has

increased from 10.45% to 13.31% from 2000 to 2010. Further, China will become the most aged society in the world by 2030, according to a report released by the Chinese Academy of Social Sciences (CASS) in 2010.[1] This chapter tries to examine the relationship between aging and housing demand in urban China. In particular, this paper explores how housing demand varies with age at the household level using micro-level data from China. Housing demand is defined as housing service flows consumed, i.e. housing expenditure, which equals quantities of housing services consumed times unit price. Throughout this chapter, I control for the quality of houses to analyze the relationship between housing demand and aging. Therefore, housing demand reflects the willingness to pay for a constant-quality house. This chapter concludes that housing demand in China will not decrease much with aging in the long run.

Mankiw and Weil (1989) is the first paper to investigate the effects of demographic changes on housing markets. Using 1970 census data for the US, they estimate the effect of age on the demand for housing per capita and find that housing demand increases with age for residents aged 20-30, and then starts to decline for those aged 40 and more. Furthermore, they do not find any relationships between the demand for housing and the quantity of housing measured as the national stock of housing at the macro level. Yet, changes in housing demand have substantial impacts on housing prices, which implies housing supply is inelastic, and suggests that housing prices will fall with population aging in the U.S.

However, the conclusion of their paper triggers a number of criticisms. First, their model is not general enough to predict trends of housing prices. Hendershott (1991) argues that Mankiw and Weil's model explains housing prices in the U.S. in the 1950s and 1960s well, but cannot predict real prices in the 1970s and 1980s, not to mention the forecast of the 1990s. He also points out that the forecast of housing prices decline results from negative time trends in their estimation equation. Furthermore, applying Mankiw and Weil estimation equations to postwar data for Canada, Engelhardt and Poterba (1991) conclude that Mankiw and Weil's model cannot forecast housing prices in Canada although the shape of age-housing demand for Canada is similar to the one for the U.S. Second, Mankiw and Weil (1989) ignore other determinants of housing prices except age. Swan (1995) argues that Mankiw and Weil (1989) do not consider real income and real interest rates which also determine the demand for housing in addition to adult population. After correction for that, he still finds great impacts of demographic changes on house prices. Green and Hendershott (1996) keep housing quality constant by relating real expenditures to hedonic characteristics. Then they compute how the willingness to pay for a constant-quality housing units varies only with age or as well as with other demographic and economic characteristics. Finally, they find a positive effect of age on hous-

[1] Retrieved from http://en.people.cn/90001/90776/90882/7137446.html

ing demand after holding other characteristics constant. The reason why the conclusion of Mankiw and Weil (1989) is wrong is that they ignore the effect of other factors such as education and income. Eichholtz and Lindenthal (2014) confirm this result using data collected from the English Housing Condition Survey (EHCS). Third and perhaps more important, housing supply appears to be elastic in the long run, so that changes in housing demand induced by demographic changes will affect quantity of housing instead of housing prices. Holland (1991) shows that increasing housing demand driven by demographic changes causes the growth of the real residential investment, but is not the major reason for increased housing prices in the U.S, which implies that the supply of housing is elastic and the housing market is efficient. Engelhardt and Poterba (1991) find a negative or statistically insignificant association between housing demand induced by demography and house prices in Canada, in contrast to the conclusion of Mankiw and Weil (1989). For Japan, Ohtake and Shintani (1996) find that demographic change affects housing prices significantly in the short run, but has an insignificant effect on housing prices in the long run. Therefore, this paper focuses on investigating the effect of demographic changes on housing demand rather than housing prices.

I highlight two micro channels through which age may affect housing demand. Households' characteristics, such as size, wealth, employment status, physical fitness and income, are expected to change over the life cycle, so I call them life-cycle variables. First, if life-cycle variables are controlled for, there could be an effect of age on housing demand because of changing preference over the life-cycle. For instance, middle aged people prefer to own more living space than old retirees with similar incomes, because the latter does not have enough energy to take care of big houses. Second, the life-cycle variables also change over the life cycle, which could influence housing demand.

The life-cycle model concludes that elderly have lower housing demand. But in China, the elderly still have a large demand for housing due to the following factors. First, since the Chinese housing reform started in 1988, many elderly still live in the old-styled housing units which are usually allocated by state-owned enterprises. Floor space per capita is small and housing quality is low. These elderly try to improve their living conditions. Therefore, it is necessary to examine how housing demand varies with age in China. Second, with the booming Chinese economy and rapid education improvement, the average income of current generations is much higher than previous those, so the housing demand will also increase a lot in the long run.

Many researchers study the impacts of demographic changes on housing demand and housing prices in North America and Europe, but they have not reached a consensus. The conclusion of Mankiw and Weil (1989) is supported by the following research. For example, Ermisch (1996) uses micro-level data to show that the growth rate of aggregate housing demand decelerates substantially as a con-

sequence of population aging in six British agglomerations using micro data. Lindh and Malmberg (2008) find residential construction rate decreases with aging in Sweden and imply that population aging has a negative association with housing demand. Levin et al. (2009) report a strong negative effect of aging on house prices by comparing Scotland and England. Takáts (2012) uses a macroeconomic panel data including 22 advanced countries between 1970 and 2009 to forecast that population aging will decrease housing prices significantly in the next 40 years. However, some researchers disagree. Neuteboom and Brounen (2007) predict both housing consumption and homeownership to rise with Dutch population aging because housing demand of seniors is increasing and ownership rates among young households also rise. Green and Hendershott (1996) and Eichholtz and Lindenthal (2014) also find housing demand will increase with demographic changes. Using several waves of surveys in the U.S., Green and Lee (2016) examine the relationship between demographics and housing demand after controlling for variables related to age and cohort variables, and find that housing demand will not lower along with aging because the educational attainment and income levels of the current and future old persons are relatively higher than that of previous cohorts.

However, there are few empirical papers which emphasize the role of demographic changes on housing demand in developing countries. To the best of my knowledge, only Li and Shen (2013) study the relationship between population aging and housing consumption in China using aggregate data. As housing prices in China have risen fast since 2003, one could wonder whether China is in a housing bubble. (See Wu et al. (2012), Glaeser et al. (2017)) This chapter contributes to the understanding of the Chinese booming housing markets from the perspective of demographics and offer new evidence on this debated issue as few examines this question in the emerging countries.

The remainder of this chapter is organized as follows: The next section develops the empirical specification. Section 3.4 then turns to the interpretation and descriptive statistics of the micro data used in this chapter. Section 3.5 presents the regression results. Section 3.6 presents figures on the relationship between housing demand and aging. Section 3.7 concludes this chapter.

3.3 Empirical specification

This paper uses a two-stage hedonic price model to investigate how housing demand varies with age. This model was originally developed by Rosen (1974), then applied in Green and Hendershott (1996), and extended by Eichholtz and Lindenthal (2014) and Green and Lee (2016). The first stage consists of a hedonic function of housing price with respect to each characteristic of a housing unit in order to get the implicit price for each hedonic characteristic. In the second stage, the implicit prices

above are regressed on socio-demographic characteristics of the household and all hedonic characteristics of the housing as well as year dummies. This step aims at estimating the relationship between the implicit marginal price of each housing characteristic and household's age, after controlling for housing quality and other household's characteristics. Then, the age-specific willingness-to-pay for a constant-quality housing unit can be calculated given implicit marginal prices of housing characteristics and a vector of associated household characteristics. This method enables to estimate the pure age effect on housing demand because other related socio-demographical variables are controlled for. Alternatively, if one does not control for socio-demographics except age in the second stage, all socio-demographical variables are allowed to change with age, which allows to compute the global age effect on housing demand.

In the first stage, the hedonic price model links a household's housing expenditure with n characteristics of the housing unit where the household lives.

$$lnq_j = \alpha + \sum_{i=1}^{n} \beta_i lnz_{ij} + \varepsilon_j \tag{3.1}$$

where q_j denotes the housing expenditure of household j, z_{ij} the ith hedonic characteristic of the housing unit, and ε_j the error term.

Taking the partial derivative of q_j with respect to z_{ij} in Equation (3.1):

$$\frac{1}{q_j} q_{ij} = \frac{\beta_i}{z_{ij}} \tag{3.2}$$

where $q_{ij} \equiv dq_j/dz_{ij}$.

Then, multiplying both sides of Equation (3.2) by q_j and using the estimated regression coefficient $\hat{\beta}_i$ of Equation (3.1), I get the value of implicit marginal price ($\widetilde{q}_{ij}$) of the hedonic characteristic i for household j as follows:

$$\widetilde{q}_{ij} = \hat{\beta}_i \frac{q_j}{z_{ij}} \tag{3.3}$$

I estimate Equation (3.1) separately for each year in the dataset to allow the estimated parameters $\hat{\beta}_i$ to be year-specific, which makes the marginal implicit prices vary over years. Here, I follow Green and Lee (2016) and use the log-log regression model in the first-stage regression, which not only assumes that the utility function is separable and has a constant elasticity β_i for the ith hedonic characteristic,

but also allows to impose linear homogeneity on the hedonic function q_j.[2] It should be noted that the linear homogeneity restriction is necessary, albeit ad hoc, to compute the total demand for an entire house.[3] Under this restriction condition, the total quantity of housing services for household j can be computed as:

$$q_j = \sum_{i=1}^{n} \widetilde{q}_{ij} z_{ij} \tag{3.4}$$

by Euler's theorem, where $\widetilde{q}_{ij}$ is the computed marginal price for the ith housing attribute of household j obtained from Equation (3.3).

In the second stage, I relate q_{ij} to the vector of age, socio-demographic characteristics of households, housing characteristics and time-dummies:

$$\widetilde{q}_{ij} = \gamma_{i0} + \sum_{a=1}^{A} \lambda_{ai} age_{aj} + \boldsymbol{\theta_i X_j} + \boldsymbol{\gamma_i Z_j} + \boldsymbol{\delta_i Y} + \mu_{ij} \tag{3.5}$$

where $\widetilde{q}_{ij}$ is calculated in Equation (3.3) with coefficients of Equation (3.1), $\boldsymbol{Z_j}$ the vector of housing characteristics of household j, $\boldsymbol{X_j}$ the vector of socio-demographic characteristics except age of household j, age_{aj} the a_{th} age group dummy of head of household j, $\boldsymbol{Y}$ the vector of year dummies, μ disturbance term that is independently and normally distributed.

It should be noted that the marginal implicit willingness-to-pay for the i_{th} characteristic of housing not only relies on socio-demographical characteristics of the household, but also depends on the level of the housing hedonic characteristic q_{ij} and other characteristics of housing. For example, the willingness to pay for an extra unit of housing floor space depends not only on the current amount of housing but also on the number of rooms. This is why $\boldsymbol{Z}_j$ in included into the right-hand side of Equation (3.5)

Considering that the object of this paper is to link the demand for a constant-quality housing unit to household's age, I decide to use the OLS approach to analyze the relationship between the implicit prices of a constant-quality housing unit and age. The constant-quality housing unit is defined as a housing unit with mean values of housing characteristics of all households in the sample used in this paper, so it could be regarded as given and exogenous to household's characteristics. As long as

[2]The three most frequently used models in the related literature are the log-log, the semi-log and the linear. It is impossible to impose homogeneity restriction on the semi-log model, although housing characteristics containing zero values can be treated in the semi-log model rather than the log-log model. The constant returns of scale assumption could be imposed on the linear model, but the elasticity of substitution of attributes varies in it. (Diewert, 2003)

[3]In the latter section, I compare the first-stage regression result with homogeneity restriction and that without the restriction, and do not find substantial differences between coefficients of variables.

effects of unobservable variables on the demand for a constant-quality housing do not change with age, the effect of age on the constant-quality housing demand does not suffer much from the omitted variable bias.

Then, borrowing the coefficients estimated in Equation (3.5), I can compute for each age v a representative household's marginal willingness-to-pay (MWTP) for the ith characteristic of the constant-quality house following the equation below:

$$\hat{q}_{iv} = \hat{\gamma}_{i0} + \sum_{i=1}^{n} \hat{\gamma}_i \bar{z}_i + \sum_{i=1}^{m} \hat{\theta}_i \bar{x}_i + \hat{\lambda}_{iv} + \hat{\delta}_i \overline{Y} \tag{3.6}$$

where $\bar{z}_i$ represents the ith characteristic of the constant-quality house, $\bar{x}_i$ the ith element of the vector of m socio-demographic characteristics of a representative household, $\hat{\lambda}_{iv}$ indicates the estimated age coefficient of the household head of the representative household v, $\overline{Y}$ the average year in the sample, and $\hat{\delta}_i$ the year coefficient estimated for the ith characteristic. The representative household is defined as a household with average socio-demographic characteristics except age.

I also compute the total willingness-to-pay (WTP) for the constant-quality housing unit of the representative household aged v years old. It is convenient and necessary to impose the constraint that the hedonic function is homogeneous of degree one, that is, $\sum_{i=1}^{n} \beta_i = 1$, because it implies the marginal price of the ith characteristic of housing equals its average price, which is the case in the general equilibrium. Then Euler's Theorem shows that the WTP of a representative household headed by a person aged v years old for the constant-quality house can be expressed as dot products of a vector of n housing characteristics with its corresponding vector of implicit prices for each housing attribute as follows:

$$q_v = \sum_{i=1}^{n} \hat{q}_{iv} \bar{z}_i \tag{3.7}$$

where $\hat{q}_{iv}$ is the computed value obtained from Equation (3.6).

If all demographic variables except age are included into $\boldsymbol{X_j}$ in Equation (3.5), then λ_{ai} could be interpreted as the pure effect of age on the MWTP for the housing characteristic q_i. Therefore, q_{iv} calculated in Equation (3.6) varies only with λ_{iv}, which changes across age groups. Finally, the total WTP of a representative household for the constant-quality house q_v, calculated from Equation (3.7), also only varies with the household's age.

However, it is also possible to estimate the second-stage regression without any socio-demographic characteristics in Mankiw and Weil (1989):

$$q_{ij} = \gamma_{i0} + \boldsymbol{\gamma_i Z_j} + \sum_{a=1}^{A} \lambda_{ai} age_{aj} + \boldsymbol{\delta_i Y} + \mu_{ij} \tag{3.8}$$

which means that λ_{ai} can be interpreted as the global age effect on the MWTP for q_i, in other words, all effects of variables associated with age are incorporated into λ_{ai}. I will compare the pure age effect on the WTP for a house and the global age effect to demonstrate how it is important to distinguish the role of age from that of other socio-demographical variables when analyzing the effect of age on housing demand.

Some socio-demographical variables such as marital status, income and household size change with age naturally, while other variables such as gender and educational attainment keep constant over the household life-cycle. Therefore, following Eichholtz and Lindenthal (2014), I add age-constant variables such as gender and educational attainment into $\boldsymbol{X_j}$ in Equation (3.5), but do not control for variables which change with age. Because housing demand might change across generations, it would be useful to control for cohort effects when analyzing the effect of age on the WTP for a house. However, as the UHS used in this paper covers only three successive years, households of same age are also the same cohort in the sample, I cannot disentangle age effect and cohort effect.

3.4 Data

The primary data used in this paper are from the annual Urban Household Survey (UHS), which is conducted by the National Bureau of Statistics of China (NBSC) aiming to know about living situation of urban households. The NBS samples urban households in all 31 province-level regions of Mainland China and investigates detailed characteristics of households, cash income, non-cash revenues, and expenditures.[4] The data used in the estimation covers 16 provinces in the period 2007-2009. Housing expenditures are measured by annual rents. The UHS contains two types of rents: actual paid rent and imputed rent. In general, renters in private housing are asked to declare amount of money paid for rent per month, and owner-occupiers are asked to estimate their imputed rent. Households renting public housing are asked to report both actual paid rents and imputed market rents, because their actual paid rents are subsidized by the government or state-owned enterprises therefore much lower than imputed rents. I decide to measure housing expenditures using actual rents paid by households if they rent from private housing and imputed rents otherwise. To take into account the possible measurement error of imputed rents, I control for housing tenure in the

[4]See the detailed introduction to the UHS in Appendix A.

estimations. The UHS data provides six-digit address (district-level) code from 2007 to 2009.[5] I got the longitude and latitude coordinates from city and district names by using the Baidu (China's leading search engine) Maps , then I calculated the distance from each district center to the center of the corresponding city in order to control for relative locations of housing within each city.

There are 93,518 observations at the household level in cities in the UHS dataset from 2007 to 2009. I exclude 160 households headed by a person aged over 85 years-old, because they might not be the actual decision makers within their households. In addition, 11 observations are discarded because they have more than 10 household members. After excluding 1,119 observations with missing values, there are 149 cities and 92,417 observations in the sample. These observations are representative of 63% of the population living in urban districts in China, according to the China Yearbook for Regional Economy. In the sample, all cities have at least 40 observations each year and there are more than 100 observations for 85% of cities shown in Table 3.1. The NBS uses the stratified random sampling method to select households as samples of the UHS.

Table 3.1: Distribution of the number of observations at the city level in the sample

	No.	Col %	Cum %
The number of observations			
40-69	38	8.5	8.5
70-99	30	6.7	15.2
100-199	220	49.2	64.4
200-399	114	25.5	89.9
400-599	26	5.8	95.7
600-	19	4.3	100.0
Total	447	100.0	

The UHS survey classifies housing tenures into six types as shown in Table 3.2. Annual rents are deflated by the urban province-specific CPI index. Renters account for 9.6% of the sample, while 89.0% households are owner-occupiers.[6] Owning a private housing by housing reform means that households bought their housing units from public housing during the housing reform process starting in 1998. The government usually provides a substantial discount to people working in state-owned institutions or enterprises to encourage people to buy public housing. Real estate housing is defined that the housing is constructed by a real estate agency. The average imputed rent for households own-

[5]During the period 2002-2006, the UHS data only provide four-digit address (city-level) code.

[6]The homeownership rate in the UHS sample is higher than that calculated using 2010 census, because few migrants are surveyed in the UHS and migrants have lower homeownership rate than local people.

ing real estate housing is substantially higher than that of other tenure types, which might be caused by the relative higher housing quality of real estate housing compared to other housing tenure.

Table 3.2: Summary of annual rents by tenure

	Mean	SD	Median
Renting public housing (7.15%)	7212	7206	4988
Renting private housing (2.46%)	6948	7159	4800
Owning private housing (6.87%)	7848	8462	4834
Owning private housing by housing reform (50.57%)	6920	7023	4742
Owning real estate housing (31.59%)	9989	9395	6276
Other types (1.37%)	6668	7524	3940
Total (100%)	7971	8088	5030
Observations	92417		

Note: All amounts are adjusted for inflation in 2009 ¥. The fourth type (owning private housing by housing reform) means that households got their home ownership by buying public housing during the housing reform process starting in 1998. The government usually provides a substantial discount to people working in state-owned institutions or enterprises to encourage people buy public housing. Yang and Chen (2014) illustrate the detail of housing reform in China.

The characteristics of housing in the sample include floor space area, housing age, housing tenure, housing types, presence of drinking water, toilet and bathroom, heating facilities and fuel sources. Location characteristics include distance to the city center, city size, employment shares across sectors, population density, the number of students in colleges, greening ratio, pupil-teacher ratio, and the number of books in public library at the city level. The housing type is divided into six categories: low-quality apartment, apartment with single bedroom, apartment with two bedrooms, apartment with three bedrooms, apartment with four or more bedrooms, and an independent building. A low-quality apartment is defined as a housing unit with shared kitchen, dinning room and toilet and other types of houses with poor facilities. An independent building is defined as a building with high-standard facilities, which has more than 100 square meters of floor area and its own kitchen, dinning room, and toilet. Characteristics of cities are from the China City Statistics Yearbook.

With respect to characteristics of households, I use age of household head to represent the age of household. Age is divided into 12 five-year age groups. The other socio-demographic characteristics

of households include household size, the number of other housing units owned, householder's gender, marital status and educational attainment. They also include non-housing income, which is defined as gross income for owner-occupiers, because owner-occupiers do not need to pay money for their imputed rents. While for household renting private housing, the non-housing income is equal to the gross income net of rents. The non-housing income of households renting from public housing should equal their gross income less than actual paid rents.

Following the strategy of Green and Lee (2016), I add one to housing age, and assign e (mathematical constant) to dummy variables if the dummy variable equals one, and assign one to dummy variables otherwise to avoid the "log zero value" issue. Thus, when I apply the log-log model in the hedonic price model, dummy variables and variables whose values equal zero or positive numbers can be included into the first-stage regression equation.

During the period 2007-2009, half of households are replaced in the sample each year, so that all households are changed every two years. Table 3.3 summarizes descriptive statistics of characteristics of housing, household head and household in 2007 and 2009.[7] To test whether mean values of variables changed significantly from 2007 to 2009, I use t-test to compare means of continuous or dichotomous variables, and use Pearson's χ^2 test to compare means of categorical variables with more than 2 categories. As shown in Table 3.3, average housing expenditure increased from ¥6,324 to ¥9,610 from 2007 to 2009, which indicates that China is experiencing a booming housing market, albeit the great recession in 2008 had negative impacts on the housing market. The floor space area also kept increasing. The aging of house increased by 0.8 years from 2007 to 2009. Consistent with the housing reform and booming housing markets in China, the proportion of households owning real estate housing increased from 24% in 2007 to 33% in 2009, while the share of households renting public housing and owning private housing by housing reform decreased during the period 2007-2009. Regarding types of housing, about 80% of households lived in an apartment with two or three bedrooms. During this period, more than 83% of households drunk tap water. Facilities of the toilet, bathroom, heating system also improved on average. More households used clean energy such as pipe gas. In addition, the average distance of housing to the city center decreased from 8.235 km to 7.394 km during the period 2007-2009. In sum, the housing quality of urban China has greatly improved in this period. To control for characteristics of neighborhood within a city, I include four variables at the city level: city size, green space ratio, student-teacher ratio in primary schools, and the number of books in public libraries. The average city size increased from 3.458 million to 3.696 million in three years. The green space ratio and the number of books in public libraries also increased, whereas the student-teacher ratio in primary schools decreased.

[7] Table 3.6 in Appendix B reports descriptive statistics for the sample each year.

Table 3.3: Characteristics of housing, household head and households in 2007 and 2009

	2007		2009		Difference		Pearson's χ^2	P-value
	Mean	SD	Mean	SD	Mean	SD		
Sample size	27949		32027					
Characteristics of housing								
Housing expenditure (2009 Yuan)	6,324	(7,344)	9,610	(9,844)	3,286***	(8.472)		
Floor area (square meter)	77.825	(35.264)	81.365	(42.161)	3.540***	(0.566)		
House age	16.084	(16.181)	16.900	(11.958)	0.816***	(0.340)		
Tenure (%)							955.998***	(0.000)
Renting public housing	0.082	(0.274)	0.072	(0.258)				
Renting private housing	0.015	(0.121)	0.032	(0.176)				
Owning private housing	0.057	(0.232)	0.070	(0.254)				
Owning private housing by housing reform	0.589	(0.492)	0.485	(0.500)				
Owning real estate	0.242	(0.428)	0.330	(0.470)				
Other types	0.015	(0.122)	0.012	(0.110)				
Types of housing (%)							103.836***	(0.000)
Low quality apartment	0.095	(0.293)	0.092	(0.290)				
Apartment with single bedroom	0.058	(0.233)	0.065	(0.247)				
Apartment with two bedrooms	0.519	(0.500)	0.491	(0.500)				
Apartment with three bedrooms	0.286	(0.452)	0.298	(0.457)				
Apartment with four or more bedrooms	0.032	(0.176)	0.034	(0.182)				
An independent building	0.011	(0.106)	0.019	(0.136)				
Drinking water (%)							176.658***	(0.000)
Tap water	0.838	(0.369)	0.869	(0.338)				
Mineral/Purified/Filtered water	0.082	(0.274)	0.078	(0.268)				
Well/Spring/River/Lake/Other water	0.080	(0.271)	0.054	(0.226)				
Toilet and bathroom (%)							42.005***	(0.000)
None	0.031	(0.172)	0.027	(0.163)				
Both	0.790	(0.407)	0.811	(0.392)				
Have toilet but no bathroom	0.155	(0.362)	0.141	(0.348)				
Public facilities	0.024	(0.154)	0.021	(0.143)				
Heating facilities (%)							0.951	(0.813)
None	0.210	(0.408)	0.212	(0.409)	0.002	(0.058)		
Air conditioning	0.299	(0.458)	0.298	(0.457)	-0.001	(0.061)		
Heating	0.419	(0.493)	0.420	(0.494)	0.001	(0.064)		
Other	0.072	(0.258)	0.070	(0.255)	-0.002	(0.046)		
Fuel source (%)							63.711***	(0.000)
Pipe gas	0.595	(0.491)	0.626	(0.484)				
Liquid gas	0.337	(0.473)	0.308	(0.462)				
Coal or other	0.068	(0.251)	0.067	(0.249)				
Characteristics of cities								
Distance to the city center (km)	8.235	(17.274)	7.394	(10.188)	-0.841***	(0.338)		
City size (in millions)	3.458	(4.086)	3.696	(4.229)	0.238***	(0.185)		
Green space ratio (%)	37.960	(7.922)	42.987	(27.681)	5.027***	(0.414)		
Student-teacher ratio	17.720	(3.548)	16.800	(3.339)	-0.920***	(0.168)		
# Books in public libraries per 1000 persons	12.447	(11.883)	14.623	(14.428)	2.176***	(0.330)		

continued on next page

Table 3.3 (*continued*)

	2007		2009		Difference		Pearson's χ^2	P-value
	Mean	SD	Mean	SD	Mean	SD		
Characteristics of household head								
Age	49.347	(11.621)	49.497	(12.339)	0.149	(0.314)		
Male	0.694	(0.461)	0.687	(0.464)	-0.007*	(0.062)		
Marriage rate (%)	0.935	(0.247)	0.914	(0.280)	-0.021***	(0.047)		
Educational attainment (%)							115.748***	(0.000)
Up to primary schooling	0.060	(0.237)	0.067	(0.251)				
Middle schooling	0.259	(0.438)	0.272	(0.445)				
High schooling	0.351	(0.477)	0.326	(0.469)				
Vocational training	0.206	(0.405)	0.190	(0.392)				
Bachelor degree or higher	0.125	(0.330)	0.145	(0.352)				
Characteristics of household								
Household size	2.858	(0.798)	2.839	(0.867)	-0.019***	(0.083)		
Number of other houses owned by households	0.109	(0.357)	0.134	(0.387)	0.026***	(0.055)		
Non-housing income (2009 Yuan)	47,869	(37,782)	58,039	(45,275)	10,169***	(18.531)		

Note: T-test for difference beteween the variables in 2007 and those in 2009. Person's χ^2 for the hypothesis that the rows and columns in a two-way table are independent. * $p < 0.1$, ** $p < 0.5$, *** $p < 0.01$. Yuan is the base unit of Chinese currencies.

Characteristics of household head in 2009 differ in their characteristics in 2007 in term of gender, marriage rate and educational attainment. The average age of household heads in 2009 is 49.497, which is older than their age in 2007 whereas the difference is not significant. More than 68% of household head is male from 2007 to 2009, which implies that males are mainly responsible for the economic well-being of households in urban China. The marriage rate is very high and exceeds 90% in the sample, but the marriage rate has declined by 2.1% during the period 2007-2009. The fraction of household head with bachelor's degree is slightly higher in 2009 compared to 2007.

Examining characteristics of households, unsurprisingly, the household size declined from 2.858 to 2.839 because of the decreasing fertility rate. The average number of other housing units owned by households besides the main housing unit increased from 0.109 to 0.134 and the non-housing income kept increasing from ¥47,869 to ¥58,039.

In sum, most characteristics of housing, household head and households listed in Table 3.3 changed significantly from 2007 to 2009. Therefore, it is necessary to estimate the first-hedonic regression model year by year.

3.5 Regression results

I estimate Equation (3.1) by OLS with robust standard errors for the years 2007 to 2009 separately. I impose the homotheticity restriction to the estimated coefficients in the first stage hedonic housing price model in order to be able to calculate the total housing demand.[8] In the hedonic housing model, I control for cities' amenities and province fixed effects to examine impacts of characteristics of local housing markets on housing expenditures. I control for distance to the city center to capture locations of housings.

Table 3.4 shows that most estimated coefficients have expected signs in each year. The mean implicit marginal prices (MIMP) for each housing characteristic, the mean q_{ij} in Equation (3.3), are also reported in Table 3.4. A percentage increase in floor area raises the housing expenditure by 0.625% to 0.743% depending on year. The mean implicit marginal price of floor space area is between ¥66 and ¥80 from 2007 to 2009. The housing expenditure will decrease by around between 0.013% and 0.033% if age of house increases by 1%. The average implicit marginal price of house age decreased from ¥7 to ¥27. One should be cautious is interpreting coefficients and mean implicit marginal prices of tenure categories as measurement errors of housing expenditures might matter. Compared with renting public housing, households prefer to pay more if they can rent private housing or own their dwelling. For example, households would like to pay a 42.8% higher rent if they rent private housing rather than public housing in 2007, which is equivalent to ¥2683. The types of housing have positive and significant effects on housing expenditure as expected. Residents are willing to pay on average an additional ¥1202 a year for living in an independent building in 2007, when compared to a low-quality building. Then, the premium households are willing to pay for an independent building, when a low-quality building is taken as reference, increased to ¥2530 in 2008, and further to ¥3265 in 2009. It is surprising and difficult to explain that spring, river, lake or other water have positive effects on housing expenditure, while drinking mineral, purified or filtered water has inconsistent effects over years, taking tap water as reference to drinking water, drinking well. Pipe gas as the fuel source can make households pay more for housing, compared with liquid gas, coal or other fuel. Naturally, toilet, bathroom and heating facilities have positive effects on housing expenditure.

When housing is closer to the city center, it is more attractive conditioning on other housing characteristics. Households are willing to pay (on average) an additional ¥246 a year when the distance of housing to the city center decreases by one kilometer. Focusing on characteristics of cities, large city size, lower student-teacher ratio, more books in public libraries per capita have positive effects

[8] I also estimate the first hedonic housing price model without the homogenous restriction in Table 3.7, and most coefficients of variables of interest do not show substantial differences from those obtained in Table 3.4. The coefficients of dummy variables within each group change, but their relative magnitude do not change a lot.

on housing expenditures as expected. Green space ratio also has a positive and significant effect on housing expenditure except in 2007.

In the second stage hedonic housing price model, I relate the implicit marginal price for each housing characteristic, computed based on Equation (3.3), to all housing characteristics and socio-demographic variables of households, as shown in Equation (3.5). To calculate how the WTP on a constant-quality house of a representative household changes with age throughout the whole period, I use the pooled cross-sectional data set after controlling for year dummies.

Table 3.5 in each column presents estimated coefficients from selected estimation regression results of the second stage hedonic model, where I regress implicit marginal prices of housing characteristics on housing characteristics, household head, and households as well as year dummies. The results show that male as a household head relative to female is willing to pay more for heating facilities, cities with more population, cities with larger green space ratio, and pay less for additional floor area and older housing. Compared with male as a household head, female has a relatively greater willingness-to-pay for dwellings closer to the city center. As expected, everything else being equal, the non-housing income is positively correlated with the MWTP for favorable housing characteristics (such as floor area, an independent building, toilet and bathroom facilities, heating facilities, green space ratio, number of books in public library), and negatively associated with the computed implicit marginal price of inferior housing characteristics (e.g. housing age, distance to the city center and student-teacher ratio in primary schools). Educational attainment affects the MWTP for housing characteristics via two effects. The first is that as a proxy for permanent income, the degree of educational attainment can predict future incomes of households. Conditioning on current income, household head with a higher degree would prefer to consume more housing services. For example, Column 1 in Table 3.5 shows that household head with a bachelor degree or higher is willing to pay ¥5.922 for one additional unit of housing area, compared to household head attending only primary schools or below, holding other things constant. This effect has been documented in related research (see (Green and Hendershott, 1996), (Eichholtz and Lindenthal, 2014) and (Green and Lee, 2016)). Here what I want to emphasize is the preference heterogeneity effect. As stated in Diamond (2016), high-educated workers are relatively more sensitive to amenity levels than low-educated workers. Table 3.5 shows that more educated people are willing to pay more for closer distance to the city center, bigger city size, larger green space ratio, smaller student-teacher ratio in primary schools, larger number of books in public library per capita (compared to less educated people).

Table 3.4: The first hedonic regression results with homogeneous restrictions, 2007-2009

Depedent variable: *Log(rents)*	2007			2008			2009		
	Coefficent	SE	MIMP	Coefficent	SE	MIMP	Coefficent	SE	MIMP
Log (Floor area (square meter))	0.743***	(0.018)	66	0.697***	(0.015)	83	0.625***	(0.014)	80
Log (Housing age)	–0.013	(0.010)	–7	–0.019**	(0.008)	–15	–0.033***	(0.007)	–27
Tenure categories (ref. renting public housing)									
Renting private housing	0.428***	(0.039)	2683	0.409***	(0.025)	3489	0.237***	(0.022)	2240
Owning private housing	0.094***	(0.028)	572	0.157***	(0.023)	1303	0.086***	(0.020)	788
Owning private housing by housing reform	0.027	(0.018)	116	0.120***	(0.016)	747	0.055***	(0.013)	381
Owning real estate	0.108***	(0.020)	546	0.222***	(0.018)	1489	0.147***	(0.014)	1069
Other types	0.117***	(0.038)	732	0.193***	(0.030)	1666	0.193***	(0.029)	1844
Housing types (ref. a low quality building)									
Apartment with single bedroom	–0.115***	(0.030)	–706	–0.000	(0.024)	–1	–0.000	(0.020)	–1
Apartment with two bedrooms	–0.062**	(0.025)	–269	0.038*	(0.020)	234	0.038**	(0.017)	259
Apartment with three bedrooms	–0.007	(0.025)	–36	0.099***	(0.021)	679	0.105***	(0.018)	792
Apartment with four or more bedrooms	0.046	(0.036)	282	0.173***	(0.029)	1459	0.214***	(0.025)	1985
An independent building	0.192***	(0.047)	1202	0.297***	(0.042)	2530	0.347***	(0.038)	3265
Drinking water (ref. tap water)									
Mineral/Purified/Filtered water	–0.067***	(0.017)	–401	–0.029*	(0.015)	–241	0.070***	(0.013)	632
Well/Spring/River/Lake/Other water	0.079***	(0.015)	465	0.070***	(0.015)	578	0.102***	(0.013)	933
Toilet and bathroom facilities (ref. none)									
Both	0.184***	(0.035)	526	0.319***	(0.032)	1215	0.227***	(0.028)	937
Have toilet but no bathroom	0.174***	(0.035)	1033	0.311***	(0.032)	2560	0.173***	(0.028)	1587
Public facilities	0.125***	(0.042)	783	0.350***	(0.036)	3022	0.157***	(0.030)	1495
Heating facilities (ref. none)									
Air conditioning	0.032**	(0.014)	158	0.095***	(0.015)	676	0.021*	(0.013)	166
Heating	0.237***	(0.021)	1072	0.407***	(0.020)	2457	0.294***	(0.017)	1990
Other	0.068***	(0.023)	417	0.214***	(0.022)	1810	0.155***	(0.019)	1449
Fuel source (ref. pipe gas)									
Liquid gas	–0.220***	(0.010)	–1174	–0.078***	(0.010)	–581	–0.101***	(0.008)	–826
Coal or other	–0.176***	(0.017)	–1091	–0.014	(0.017)	–120	–0.097***	(0.015)	–914
Log (Distance to the city center (km))	–0.097***	(0.004)	–253	–0.106***	(0.004)	–322	–0.075***	(0.004)	–246
Characteristics of cities									
Log (City size)	0.201***	(0.006)	711	0.284***	(0.006)	1214	0.327***	(0.006)	1508
Log (Green space ratio (%))	–0.096***	(0.020)	–16	0.216***	(0.024)	48	0.245***	(0.022)	55
Log (Student-teacher ratio)	–0.720***	(0.028)	–282	–0.616***	(0.027)	–336	–0.666***	(0.025)	–411
Log (# Books in public libraries per 1000 persons)	0.135***	(0.006)	119	0.245***	(0.006)	239	0.225***	(0.005)	252
Constant	7.013***	(0.136)	7	5.421***	(0.137)	5	6.083***	(0.130)	6
Province FEs (16 provinces)	Yes			Yes			Yes		
Observations	27844			32153			31721		

Note: * $p < 0.1$, ** $p < 0.5$, *** $p < 0.01$. Standard errors are heteroskedasticity consistent. The dependent variable is the annual rents. MIMP represents the mean implicit marginal price of all household surveyed in each year for each hedonic characteristic. Owning a private housing by housing reform means that households bought their housing units from public housing during the housing reform process starting in 1998. Real estate housing is defined that the housing is constructed by a real estate agency.

Table 3.5: Selected results of second-stage hedonic regression for computed MWTP of housing characteristics, 2007-2009

	Floor area	Housing age	An independent building (vs. a low quality apartment)	Well/Spring/River/Lake/Other (vs. tap water)	Both toilet and bathroom (vs. none)	Heating (vs. none)
Male	–0.904**	–0.533***	14.564	2.673	–4.243	42.514***
	(0.450)	(0.201)	(8.940)	(4.076)	(6.038)	(13.538)
Household size	–0.364	0.831***	–19.132***	–7.215***	–6.432	–17.588
	(0.225)	(0.147)	(5.841)	(2.757)	(3.992)	(10.971)
Non-housing income (in thousands)	0.085***	–0.074***	3.536***	1.792***	2.049***	6.007***
	(0.013)	(0.007)	(0.344)	(0.179)	(0.239)	(0.697)
Educational attainment (ref. primary school or below)						
Middle schooling	3.812***	1.068***	14.694	10.315*	36.843***	46.679**
	(0.669)	(0.337)	(13.645)	(6.234)	(10.107)	(22.335)
High schooling	3.525***	2.194***	–0.634	–0.113	32.813***	29.500
	(0.652)	(0.331)	(13.765)	(6.371)	(9.506)	(22.406)
Vocational training	5.034***	2.133***	33.846**	8.954	58.639***	62.217**
	(0.731)	(0.407)	(16.890)	(7.896)	(11.598)	(27.638)
Bachelor degree or higher	5.922***	–0.849*	109.367***	47.613***	100.924***	156.282***
	(0.864)	(0.507)	(21.204)	(10.156)	(14.106)	(35.458)
Age group (ref. (18-25))						
(26-30)	7.106**	–4.808***	159.636***	82.933***	93.476**	161.623
	(3.187)	(1.375)	(57.819)	(24.306)	(40.901)	(100.758)
(31-35)	3.956	–3.205**	146.801***	80.011***	72.321*	134.709
	(3.026)	(1.252)	(54.004)	(23.017)	(38.075)	(92.746)
(36-40)	3.592	–2.642**	138.092***	81.142***	75.695**	136.780
	(3.009)	(1.239)	(53.420)	(22.755)	(37.956)	(91.592)
(41-45)	3.685	–2.296*	147.104***	85.071***	89.779**	152.379*
	(3.052)	(1.243)	(53.654)	(22.834)	(38.858)	(92.148)
(46-50)	2.455	–1.269	91.452*	63.097***	58.449	84.034
	(3.010)	(1.243)	(53.478)	(22.797)	(38.027)	(91.740)
(51-55)	3.078	–0.599	92.279*	55.730**	61.364	60.021
	(3.032)	(1.244)	(53.467)	(22.787)	(38.124)	(91.864)
(56-60)	3.377	–0.496	108.558**	67.513***	89.969**	86.511
	(3.044)	(1.259)	(54.074)	(23.103)	(38.528)	(92.590)
(61-65)	2.800	–0.724	156.396***	85.540***	96.236**	124.005
	(3.045)	(1.268)	(54.447)	(23.246)	(38.580)	(93.015)
(66-70)	2.089	0.015	133.845**	77.587***	98.528**	114.503
	(3.041)	(1.273)	(54.457)	(23.310)	(38.785)	(92.954)
(71-75)	3.237	0.087	168.396***	93.593***	106.173***	158.126*
	(3.081)	(1.294)	(55.797)	(24.025)	(39.109)	(94.029)
(76-)	3.709	0.257	174.271***	92.345***	116.824***	166.119*
	(3.153)	(1.335)	(57.395)	(24.866)	(40.346)	(96.821)
Characteristics of housing	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Observations	92417	92417	92417	92417	92417	92417
Adjusted R^2	0.529	0.354	0.560	0.550	0.434	0.518

Note: * $p < 0.1$, ** $p < 0.5$, *** $p < 0.01$. Standard errors are heteroskedasticity consistent. The dependent variable in each column is the implicit marginal price of one housing characteristic. To save space, I do not report coefficients of other housing characteristics, the number of old family members, the number of children, whether the household owns other dwellings, martial status, year dummies and province dummies.

Table 3.5 (*continued*)

	Liquid gas (vs. pipe gas)	Distance to the city center	Population (in millions)	Green space ratio	Student-teacher ratio in primary schools	# Books in public library per capita
Male	2.411	10.605**	22.855**	0.574*	0.116	–1.627
	(6.448)	(4.876)	(10.387)	(0.303)	(2.151)	(1.910)
Household size	10.091***	2.443	40.126***	–0.474**	4.254***	1.605
	(3.773)	(2.880)	(13.629)	(0.192)	(1.421)	(1.232)
Non-housing income (in thousands)	–2.273***	–0.729***	1.912***	0.096***	–0.916***	0.054*
	(0.256)	(0.123)	(0.287)	(0.010)	(0.094)	(0.032)
Educational attainment (ref. primary school or below)						
Middle schooling	5.488	–29.267***	76.056***	0.345	–4.273	14.029***
	(8.158)	(6.851)	(22.601)	(0.437)	(3.041)	(3.463)
High schooling	19.527**	–34.036***	136.837***	0.344	2.533	15.209***
	(8.569)	(7.175)	(26.808)	(0.448)	(3.056)	(3.425)
Vocational training	10.268	–47.535***	166.093***	1.187**	–3.708	20.449***
	(10.958)	(8.677)	(26.294)	(0.547)	(3.918)	(3.863)
Bachelor degree or higher	–46.584***	–68.083***	163.980***	4.014***	–24.909***	25.744***
	(14.066)	(10.369)	(27.524)	(0.673)	(5.054)	(4.504)
Age group (ref. (18-25))						
(26-30)	–54.845	–35.070*	80.766	2.759	–42.241***	18.396
	(39.680)	(20.563)	(53.881)	(1.905)	(13.446)	(11.492)
(31-35)	–43.114	–69.854***	35.683	2.944	–38.353***	14.817
	(37.404)	(20.361)	(49.499)	(1.792)	(12.421)	(10.655)
(36-40)	–51.477	–60.886***	45.636	2.669	–37.987***	8.702
	(37.068)	(19.746)	(49.615)	(1.772)	(12.295)	(10.501)
(41-45)	–53.497	–59.945***	87.224*	3.271*	–40.385***	9.142
	(37.248)	(19.374)	(51.130)	(1.781)	(12.382)	(10.501)
(46-50)	–46.713	–62.789***	59.853	1.289	–31.198**	2.059
	(37.327)	(19.535)	(49.027)	(1.776)	(12.356)	(10.465)
(51-55)	–35.376	–59.483***	63.529	1.427	–29.769**	0.763
	(37.431)	(19.688)	(48.915)	(1.777)	(12.377)	(10.540)
(56-60)	–66.964*	–76.915***	95.611*	1.266	–38.316***	–4.757
	(37.931)	(20.255)	(49.603)	(1.798)	(12.549)	(10.542)
(61-65)	–82.877**	–95.365***	120.788**	3.071*	–47.999***	–8.573
	(38.110)	(21.142)	(54.643)	(1.814)	(12.581)	(10.621)
(66-70)	–67.995*	–67.715***	161.021***	3.078*	–44.153***	–10.457
	(37.996)	(20.689)	(61.044)	(1.817)	(12.581)	(10.636)
(71-75)	–88.463**	–76.256***	117.544**	3.566*	–50.363***	–11.670
	(38.672)	(22.547)	(54.582)	(1.855)	(12.873)	(10.729)
(76-)	–86.748**	–88.760***	125.346**	4.503**	–53.919***	–11.377
	(40.266)	(23.662)	(53.397)	(1.928)	(13.457)	(11.067)
Characteristics of housing	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Observations	92417	92417	92417	92417	92417	92417
Adjusted R^2	0.525	0.106	0.256	0.529	0.577	0.277

Note: * $p < 0.1$, ** $p < 0.5$, *** $p < 0.01$. Standard errors are heteroskedasticity consistent. The dependent variable in each column is the implicit marginal price of one housing characteristic. To save space, I do not report coefficients of other housing characteristics, the number of old family members, the number of children, whether the household owns other dwellings, martial status, year dummies and province dummies.

The coefficients for age groups show that the MWTP of households for housing characteristics change over their life-cycle. Compared to households head aged 18-25 years old, households head aged between 26 and 30 years old have relatively greater willingness to pay for additional floor area, newer dwelling, higher quality building, better sanitary facilities, heating facilities and better amenities, ceteris paribus. The household head aged over 60 years old value an independent building more than household head aged between 18 and 25 years old. Understandably, old households have a higher willingness-to-pay for toilet and bathroom facilities and heating facilities than young households as they have more time to stay home. Households aged older than 30 years old will pay less for the marginal price of the student-teacher ratio, which implies that old households care more about education of their children. With aging, households prefer to pay less for green space ratio. When people are old, especially after people get retired after 60 years old, households favor living closer to city center and big cities. This might imply that city center and big cities can supply more resources needed for old people. For example, reputed hospitals usually locate in inner cities and in big cities. The MWTP for the number of books in public library per capita do not change over age of household head significantly, although households aged below 55 years old value it more.

3.6 Housing demand and aging

Using coefficients obtained in Table 3.5, I compute the partial WTP for each characteristic of houses along with age following Equation (3.6). Specifically, I calculate housing demand for each characteristic of the constant-quality housing by age of household head after controlling for time trends and socio-demographic characteristics of a representative household except age. If I do not control for socio-demographic characteristics of a representative household, then the computed housing demand varies with age of household head and other socio-demographic characteristics. So it is called as the global WTP for each characteristic of houses along with age, because it depends on socio-demographic characteristics of households.

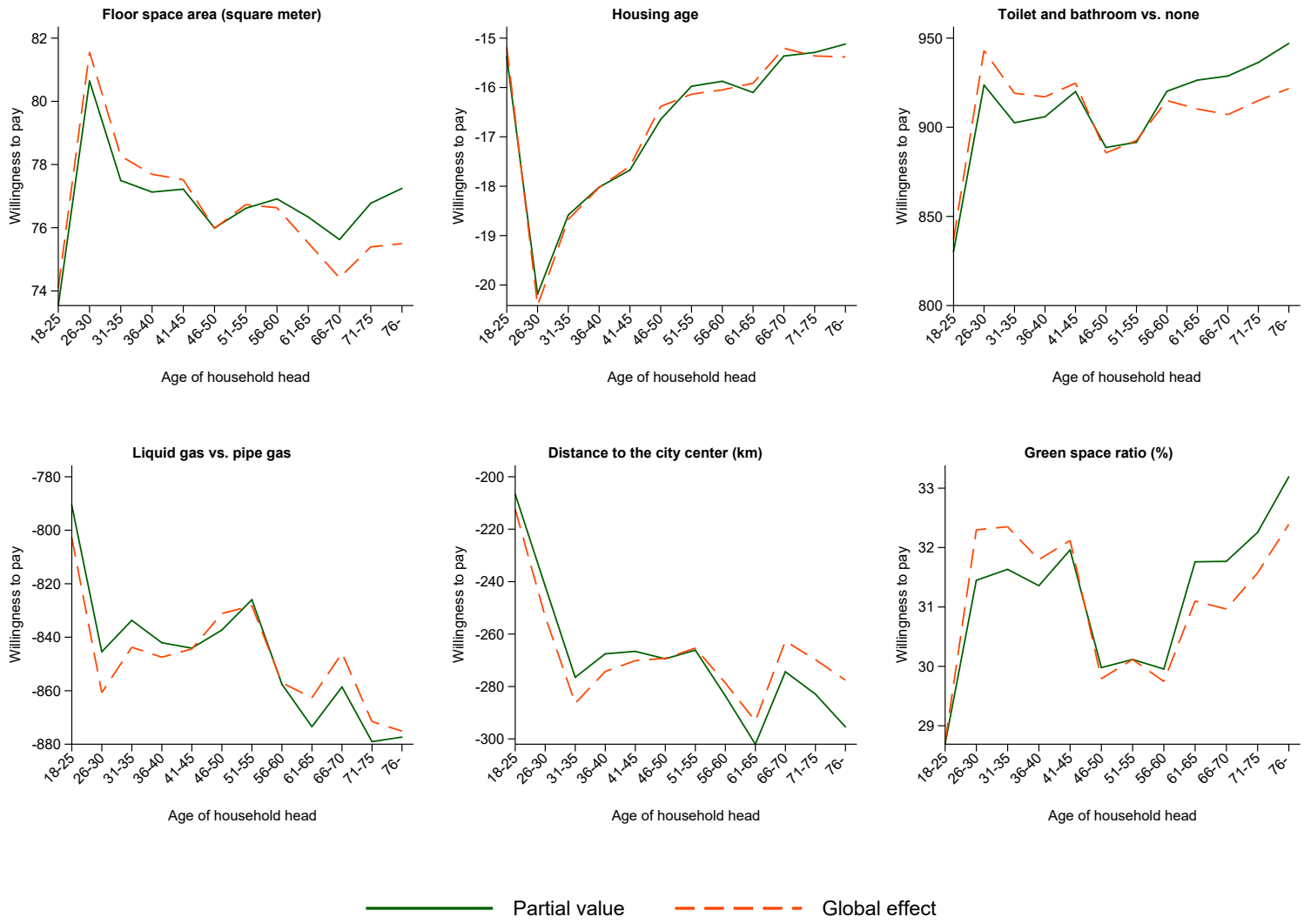
Figure 3.1 shows figures for the partial and global WTP by age for selected housing characteristics of a representative household in the sample along with age. The partial WTP for floor space area, for example, rises from ¥74 to ¥81 from age 18-25 to 26-30. Then it falls to ¥77 at age 31-35 before remaining relatively constant after age 35. The global WTP for floor space area does not show much differences from the partial one before age 60. However, after heads of households are older than 60, i.e. people get retired, the partial WTP for floor space area is still stable, while the global WTP for floor space area decreases. When people are retired, their income decreases dramatically, which

lessens the WTP for floor space area. In addition, the elderly in the sample are less educated: only 9.75% of household head aged over 60 years old received Bachelor's degree or above, while 13.58% of household head have at least a college degree in the whole sample. Therefore, after controlling for socio-demographic characteristics of a representative household except age, the partial WTP for floor space area does not decrease when household head are older than 60. The partial and global WTP for housing age are almost the same over the life-cycle of a household. Households are willing to pay more to avoid additional housing age, an inferior housing attribute, when household head at age 26-30. The WTP for a newer housing decreases with aging after household head are older than 30 years old. On average, the WTP for toilet and bathroom facilities does not change much after heads of households are older than 30. If all socio-demographic characteristics except age are controlled for, the partial WTP for toilet and bathroom facilities increases when households are older than 60, which is easy to understand, because the elderly are less mobile and have more demand for indoor toilet and bathroom facilities. Compared to pipe gas, households are willing to pay less for the liquid gas as the energy, especially when people are old. Households are willing to pay to live closer to the city center. When people get old, the WTP for closer distance from the house to the city center increases, which might be caused by the fact that most health facilities locate close to the city center in China. Regarding the green space ratio, the partial WTP of a representative household starts out at ¥29 and rises to 31.5 when heads of households are aged 26-30. It then remains around this level, followed by an increase when households' heads are aged from 36-40 through 41-45. The partial WTP for green space ratio then has a dramatic decline for middle-aged households. For retirees, the partial WTP for green space ratio exhibits a sharp rebound. Without controlling for socio-demographic characteristics of a representative household, the global WTP shows a similar pattern with the partial one, while the volatility is lower.

One of common feature of the six figures is that a representative household's partial and global WTP has sizable fluctuations between age 18-25 and 31-35. Specifically, the WTP for favorable housing characteristics, such as floor space area, toilet and bathroom facilities and green space ratio, increases from age 18-25 to 26-30 and then declines, while it for inferior housing characteristics, such as housing age, liquid gas (compared to pipe gas) and distance to the city center, decreases from age 18-25 to 26-30 and then increases.

Another common feature of the six figures is that the partial WTP has disparities with the global one after household head are older than 60, although this phenomenon is not so obvious for the figure of WTP for housing age. In particular, the partial WTP for favorable housing characteristics is larger than the global WTP after household heads get retired at age 60, and it for inferior housing characteristics is smaller than the global WTP when people are older than 60.

Figure 3.1: The partial and global WTP for selected characteristics of houses along with age



Note: The partial WTP describes the pure age effect on WTP when all demographic variables (income, educational attainment, marital status, etc.) are kept constant over the life-cycle. The global WTP shows the global age effect on WTP for housing characteristics of house when all demographic variables change with age.

It is necessary to look at the relationship between total housing demand and age. Using Equation (3.7) and (3.5) as well as (3.8), I compute the partial and global WTP for each age group. The solid line in Figure 3.2 represents the partial WTP of a household with average characteristics based on the constant-quality housing units in the sample, and the dashed line represents the global WTP. In addition, I also compute the composite WTP for each group after adding age-constant variables such as gender and educational attainment into the vector of socio-demographic characteristics in Equation (3.5). Figure 3.2 shows the three types of the relationship between the total housing demand and age. All WTP reaches the peak when the household head is aged 26-30. It might be caused by competition from marriage markets.(Wei et al., 2017) Due in part to the one-child policy, the sex ratio is as high

as 120 to 100 by 2005. In order to find a suitable partner, men have to win against the status competition by consuming more housing. The average age at first marriage is 26 in urban China during the period 2007-2009, according to the 2010 census.[9] Therefore, the housing demand is highest when household head is aged 26-30. It should be noted that young people usually can get transfers from their parents to consume housing in order to get married, as indicated in Lan (2018). The housing demand of young people depends on not only their own incomes, but also wealth of their parents.

The partial, composite and global WTP curves share a similar trend before household head aged 65. After household head's age is 65 or above, the difference between partial WTP and global one becomes larger and larger. The partial WTP of a constant-quality house decreases slightly for elderly older than 65, and it remains stable after household head is older than 71. However, the global WTP of a constant-quality house declines gradually for elderly older than 65. The composite WTP of a constant-quality house keeps a similar pattern with the partial one. The results imply that socio-demographic variables of households have significant impacts on housing demand. When socio-demographic variables are allowed to change with age, the global WTP curve shows that housing demand decreases when household heads are older than 65. However, when all socio-demographic variables except age are controlled for, the partial WTP curve shows that housing demand does not change much after household heads become elderly. When only age-constant variables including gender and educational attainment are controlled for, the composite WTP curve shows a similar pattern with the partial one. Considering that gender has much less impacts on the housing demand than educational attainment as shown in Table 3.5, I can conclude that educational attainment, which is a proxy for permanent income, is one of key factors deciding housing demand. In the sample, 30.3% of households aged 26-30 have at least a Bachelor's degree, and 20.6% of households aged 36-40 have at least a Bachelor's degree, while only 10.9% of households aged above 65 are college-educated. Therefore, the big gap in educational attainment across generations could explain the difference between the partial and global WTP curves. The results also imply that the housing demand will not decrease when the current middle-aged generation becomes old, become their educational attainment is higher than the current old generation.

Since 1998, China has experienced expansion of higher education. The college acceptance rate has increased from 34% in 1998 to 75% in 2017.[10] It is reasonable to predict that the aggregate housing demand will not decrease with population aging in China in the future.

[9]Data source: `http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexch.htm`
[10]The college here contains three-year colleges and four-year universities.

Figure 3.2: The WTP for a constant-quality house of a representative household by age group

Note: The partial WTP describes the pure age effect on WTP when all demographic variables (income, educational attainment, marital status, etc.) are kept constant over the life-cycle. The global WTP shows the global age effect on WTP for housing characteristics of house when all demographic variables change with age. The composite effect shows the age effect on WTP when all demographic variables except educational attainment and gender change with age.

3.7 Conclusion

The population of China is rapidly aging. At the same time, China has been experiencing a booming real estate market. Naturally, many people are concerned about whether the aggregate housing demand will decrease because of aging in the long run. This chapter uses micro-level data to investigate how housing demand changes with age in urban China and highlights the role of educational attainment in housing demand. The results are in line with the findings in Eichholtz and Lindenthal (2014) and Green and Lee (2016). They show that housing demand will not decrease much when the current middle-aged generation becomes old, because their educational attainment level is higher than the current old generation. The difference in levels of educational attainment could largely explain the difference between the partial and global WTP of a constant-quality house.

This chapter contributes to the existing related literature by adding evidence from China, a rapidly aging country with skyrocketing housing costs. Previously, some papers use macro-level data to examine how housing demand changes with age in China, however, they cannot disentangle the age effect and other socio-demographic variables' effects. (Li and Shen, 2013) In this chapter, I use three ways to compute the relationship between housing demand and age. By comparing the partial, composite and global WTP with respect to age, I find that household's educational attainment can largely affect housing demand, especially for elderly.

The findings have important policy implications for China and other countries with aging population. Despite an aging population, China's total housing demand is not likely to decline in the long run because educational attainment is higher for young generation. The improving human capital level can ensure total housing demand. Because of massive rural-urban migration, the number of households in urban areas increases over time, therefore the aggregate housing demand might increase in urban China in the future.

The cross-sectional data, UHS, does not allow me to disentangle age effects from cohort effects. Hence an implicit assumption is that cohorts have similar preferences on housing demand. In the future research, it is necessary to test the validity of this assumption. In addition, the UHS data include a disproportionately small percentage of migrants, which might influence accurateness of computed housing demand.

3A Appendix: UHS survey introduction

The survey design of the UHS is comparable to that of the Current Population Surveys (CPS) that is widely used to in the study of population aging and housing markets in the U.S (Ge and Yang, 2014). The NBS uses the stratified random sampling to select cities and towns as samples of the UHS. First, sorting by population size, it divides all cities and counties into 3 strata: large/medium sized cities (province-level cities and prefecture-level cities); county-level cities and counties. Notes that cities in the UHS refer to only urban districts (*Shixiaqu* in Chinese) of them and counties refer to the town where the county government locates (*Chengguanzhen* in Chinese), which could be regarded as small cities. To simplify expressions, I do not distinguish cities and counties afterwards. Second, it allocates sample size by the share of each strata's population. Third, sorting in descending order of average annual income at the city level, it calculates the cumulative number of population of cities one by one, and then uses distance sampling method to sample the required number of cities according to the sample size. The NBS examined the distribution of sampled cities and verified that the sample is

representative to the main survey indicators at the province-level.

Households are sampled in two steps. First, the NBS draws a "big sample" to provide sampling frames and basic information for the regular survey every three years. In the "big sample" survey, it use stratified sampling, two or multi-stage sampling and PPS methods to select samples in the selected cities. Second, a "small sample" is sampled from the "big sample" following a proportional allocation strategy which is decided by household size, the number of employed household members and income. Selected households are required to record their detailed consumption expenditure.

In order to strengthen representative of the sample and reduce the long-term burden of bookkeeping, during the period 2007-2009, half of selected households are replaced each year and all households should be rotated every two years.

The UHS has not been publicly available yet. What I used in this paper to is a subsample covering 16 provinces during the period 2007-2009, in which Heilongjiang and Liaoning locates in the northeastern region; Beijing, Shanghai, Shandong, Jiangsu and Guangdong locate in the eastern region; Anhui, Jiangxi, Shanxi, Hubei and Henan in the central region; Gansu, Chongqing, Sichuan and Yunnan in the western region. The 16 provinces are scattered around the country and represent 65% of China's population. (NBS, 2013)

The UHS (2007-2009) survey provides information on locations of all households at the county level (6-digit code) if they live in county-level cities or counties. For people living in a prefecture-level city, the UHS survey did not distinguish districts within cities during the period 2007-2009.

The housing markets are highly dependent among subdivisions of one urban district, because subdivisions stick together spatially and are directly administrated by the city-level government, so I aggregate households of all subdivisions within one prefecture-level city to the urban district level for the UHS.

3B Appendix: Descriptive statistics for year separately, 2007-2009

Table 3.6: Descriptive statistics in the UHS data, 2007-2009

	2007		2008		2009	
	Mean	SD	Mean	SD	Mean	SD
Characteristics of housing						
Housing expenditure (2009 Yuan)	6324	(7344)	8704	(10014)	9610	(9844)
Floor area (square meter)	77.8	(35.3)	79.8	(39)	81.4	(42.2)
Housing age	16.1	(16.2)	16.5	(25.3)	16.9	(12)
Tenure (%)						
Renting public housing	.0816	(.274)	.0768	(.266)	.0717	(.258)
Renting private housing	.0148	(.121)	.0345	(.182)	.0321	(.176)
Owning private housing	.0569	(.232)	.073	(.26)	.0695	(.254)
Owning private housing by housing reform	.589	(.492)	.492	(.5)	.485	(.5)
Owning real estate	.242	(.428)	.307	(.461)	.33	(.47)
Other types	.0151	(.122)	.0159	(.125)	.0122	(.11)
Types of housing (%)						
Low quality apartment	.0946	(.293)	.0999	(.3)	.0924	(.29)
Apartment with single bedroom	.0577	(.233)	.0686	(.253)	.0654	(.247)
Apartment with two bedrooms	.519	(.5)	.492	(.5)	.491	(.5)
Apartment with three bedrooms	.286	(.452)	.287	(.452)	.298	(.457)
Apartment with four or more bedrooms	.0319	(.176)	.0327	(.178)	.0344	(.182)
An independent building	.0114	(.106)	.0197	(.139)	.019	(.136)
Driking water (%)						
Tap water	.838	(.369)	.867	(.34)	.869	(.338)
Mineral/Purified/Filtered water	.082	(.274)	.0786	(.269)	.0776	(.268)
Well/Spring/River/Lake/Other water	.0801	(.271)	.0546	(.227)	.0538	(.226)
Toilet and bathroom (%)						
None	.0307	(.172)	.0285	(.166)	.0273	(.163)
Both	.79	(.407)	.793	(.405)	.811	(.392)
Have toilet but no bathroom	.155	(.362)	.155	(.362)	.141	(.348)
Public facilities	.0243	(.154)	.0228	(.149)	.0208	(.143)
Heating facilities (%)						
None	.21	(.408)	.217	(.412)	.212	(.409)
Air conditioning	.299	(.458)	.286	(.452)	.298	(.457)
Heating	.419	(.493)	.421	(.494)	.42	(.494)
Other	.0715	(.258)	.0755	(.264)	.0698	(.255)
Fuel source (%)						
Pipe gas	.595	(.491)	.61	(.488)	.626	(.484)
Liquid gas	.337	(.473)	.317	(.465)	.308	(.462)
Coal or other	.0678	(.251)	.0731	(.26)	.0666	(.249)

Table 3.6 (*continued*)

	2007		2008		2009	
	Mean	SD	Mean	SD	Mean	SD
Characteristics of cities						
Distance to the city center (km)	8.24	(17.3)	7.39	(10.2)	7.39	(10.2)
City size (in millions)	3.46	(4.09)	3.7	(4.23)	3.7	(4.23)
Green space ratio (%)	38	(7.92)	41.5	(29.5)	43	(27.7)
Student-teacher ratio	17.7	(3.55)	17.4	(3.55)	16.8	(3.34)
# Books in public libraries per 1000 persons	12.4	(11.9)	14.3	(13.8)	14.6	(14.4)
Characteristics of household head						
Age	49.3	(11.6)	49.1	(12.6)	49.5	(12.3)
Male	.694	(.461)	.69	(.462)	.687	(.464)
Marriage rate (%)	.935	(.247)	.908	(.288)	.914	(.28)
Educational attainment(%)						
Up to primary schooling	.0598	(.237)	.076	(.265)	.0673	(.251)
Middle schooling	.259	(.438)	.277	(.448)	.272	(.445)
High schooling	.351	(.477)	.324	(.468)	.326	(.469)
Vocational training	.206	(.405)	.186	(.389)	.19	(.392)
Bachelor degree or higher	.125	(.33)	.136	(.343)	.145	(.352)
Household size	2.86	(.798)	2.85	(.883)	2.84	(.867)
Number of other houses owned by households	.109	(.357)	.108	(.356)	.134	(.387)
Non-housing income (2009 Yuan)	47869	(37782)	51102	(38600)	58039	(45275)
Observations	27949		32441		32027	

3C Appendix C: Unrestricted first-stage regression results, 2007-2009

Table 3.7: The first hedonic regression results, 2007-2009

Depedent variable: *Log(rents)*	2007			2008			2009		
	Coefficent	SE	MIMP	Coefficent	SE	MIMP	Coefficent	SE	MIMP
Log (Floor area (square meter))	0.749***	(0.018)	66	0.716***	(0.015)	85	0.627***	(0.014)	81
Log (Housing age)	–0.029***	(0.010)	–16	–0.042***	(0.008)	–33	–0.050***	(0.007)	–41
Tenure categories (ref. renting public housing)									
Renting private housing	0.361***	(0.040)	2265	0.315***	(0.025)	2690	0.172***	(0.022)	1628
Owning private housing	0.021	(0.028)	126	0.048**	(0.023)	398	0.013	(0.020)	118
Owning private housing by housing reform	0.001	(0.018)	5	0.059***	(0.016)	367	0.015	(0.012)	106
Owning real estate	0.069***	(0.020)	347	0.141***	(0.018)	945	0.096***	(0.014)	696
Other types	0.041	(0.038)	254	0.057**	(0.029)	496	0.095***	(0.029)	905
Housing types (ref. a low quality building)									
Apartment with single bedroom	–0.168***	(0.030)	–1034	–0.069***	(0.024)	–577	–0.067***	(0.020)	–619
Apartment with two bedrooms	–0.113***	(0.025)	–491	–0.024	(0.020)	–149	–0.020	(0.017)	–133
Apartment with three bedrooms	–0.063**	(0.026)	–318	0.026	(0.021)	180	0.045**	(0.018)	339
Apartment with four or more bedrooms	–0.029	(0.036)	–175	0.078***	(0.029)	654	0.140***	(0.025)	1299
An independent building	0.095**	(0.047)	594	0.178***	(0.043)	1517	0.248***	(0.039)	2335
Drinking water (ref. tap water)									
Mineral/Purified/Filtered water	–0.082***	(0.017)	–491	–0.054***	(0.015)	–441	0.058***	(0.013)	521
Well/Spring/River/Lake/Other water	0.071***	(0.015)	417	0.049***	(0.015)	408	0.093***	(0.013)	856
Toilet and bathroom facilities (ref. none)									
Both	0.183***	(0.035)	523	0.288***	(0.031)	1097	0.236***	(0.028)	975
Have toilet but no bathroom	0.177***	(0.035)	1051	0.285***	(0.031)	2350	0.178***	(0.028)	1635
Public facilities	0.047	(0.042)	296	0.164***	(0.036)	1412	0.054*	(0.030)	516
Heating facilities (ref. none)									
Air conditioning	0.007	(0.014)	35	0.057***	(0.015)	404	–0.009	(0.013)	–72
Heating	0.074***	(0.022)	334	0.088***	(0.024)	529	0.081***	(0.020)	549
Other	–0.053**	(0.023)	–325	–0.021	(0.025)	–177	–0.009	(0.021)	–84
Fuel source (ref. pipe gas)									
Liquid gas	–0.235***	(0.010)	–1252	–0.108***	(0.009)	–805	–0.124***	(0.008)	–1021
Coal or other	–0.214***	(0.017)	–1324	–0.077***	(0.016)	–652	–0.147***	(0.015)	–1380
Log (Distance to the city center (km))	–0.088***	(0.004)	–230	–0.091***	(0.004)	–279	–0.065***	(0.004)	–213
Characteristics of cities									
Log (City size)	0.162***	(0.007)	573	0.211***	(0.007)	901	0.280***	(0.006)	1290
Log (Green space ratio (%))	–0.106***	(0.020)	–18	0.281***	(0.024)	63	0.101***	(0.022)	23
Log (Student-teacher ratio)	–0.599***	(0.028)	–235	–0.372***	(0.028)	–203	–0.571***	(0.025)	–353
Log (# Books in public libraries per 1000 persons)	0.121***	(0.006)	106	0.217***	(0.006)	212	0.211***	(0.005)	237
Constant	7.233***	(0.138)	7	5.403***	(0.134)	5	6.976***	(0.133)	7
Province FEs	Yes			Yes			Yes		
Observations	27844			32153			31721		
Adjusted R^2	0.493			0.604			0.677		

Note: * $p < 0.1$, ** $p < 0.5$, *** $p < 0.01$. Standard errors are heteroskedasticity consistent. The dependent variable is the annual rents. MIMP represents the mean implicit marginal price of each hedonic characteristic. Owning a private housing by housing reform means that households bought their housing units from public housing during the housing reform process starting in 1998. Real estate housing is defined that the housing is constructed by a real estate agency. I do not impose the homothetic restriction on the estimated coefficients in Equation 3.1.

Chapter 4

Skill intensity ratio and housing prices across Chinese cities

4.1 Abstract

Using China's census data in 2000 and 2010, this chapter shows that housing costs are positively related to the ratio of high-skilled to low-skilled workers (skill intensity ratio). More specifically, the average housing price has significant positive effects on the share of high-skilled workers, but negative and insignificant impacts on the share of low-skilled workers in 2010. This result is robust to the use of both land slope variables and historical housing prices as the instruments of housing prices. However, the skill intensity ratio is less affected by housing price in 2000 when the workers' mobility was tightly regulated. This analysis offers new evidence from China for the test of the Rosen-Roback model, under the conditions that workers can move freely and the income elasticity of housing demand is less than one.

4.2 Introduction

The substantial increase in the ratio of college to high school graduates from 2000 to 2010 has been accompanied by a dramatic increase in housing cost in China. This is in line with the findings in the U.S. (Moretti 2004, Shapiro 2006). Moretti (2013) has argued that college educated workers significantly driven by skill-intensive labor demand are drawn to cities with high cost of housing. A causal relationship between the two phenomena can be in either directions. In Gyourko et al. (2013), educated workers disproportionately sort into cities in unique locations, creating "superstar cities"

with higher housing costs. Diamond (2016) shows that cities with higher shares of college graduates experienced higher wages and endogenously increase amenities, which causes costs of housing to increase. But an inverse direction of causation could also be plausible.

In this chapter, I attempt to identify impacts of housing costs on the skill intensity ratio (SIR), which is measured by the ratio of college-educated people (high-skilled workers) to non-college educated (low-skilled workers) for all adults aged 20 years old or above, across Chinese cities using census data from 2000 to 2010. The OLS regression results show that housing prices have significant positive impacts on the SIR across Chinese cities in 2010, while it has only marginally significant impacts on the SIR in 2000 when the worker's mobility was tightly regulated.[1] I use shares of naturally unavailable land, measurements of housing supply constraints, and historical housing prices in 2000, as instruments for housing price in 2010. The 2SLS regression results show that housing costs measured by average housing prices at the city level have positive and significant impacts on the SIR in 2010. Particularly, housing prices have positive and significant effects on the share of high-skilled workers in total adults aged over 20 in 2010, while they have negative significant effects on the share of non-college educated workers.

Of course, disentangling the causal impact of housing costs on SIR is complicated by many endogeneity concerns. First, as mentioned at the beginning of this chapter, a reverse causality exists. Second, common factors such as education expansion may drive both housing costs and the SIR.[2] Broxterman and Yezer (2015) use the Bartik's employment share index (Bartik 1991) and the export price index developed by Pennington-Cross (1997) and Larson (2011) as exogenous labor demand shock to be instrumental variables of housing costs. However, these instruments could be linked to the SIR via other channels. For example, innovative industries have higher geographical concentration than manufacturing industries nowadays. More and more skilled workers are attracted to cities in which there are already large employment shares in innovative industries (Moretti 2012). Thus, the exclusion restriction condition might be violated. Liang et al. (2016) use land transactions per capita as instrumental variable for housing prices. Although the total amounts of land transaction is allocated by the central government, the local government can decide the share of land used for real estate development via intervention, which explicitly relies on housing price (Zhang et al. 2017).

Differing from the previous literature, I use both land unavailability, i.e. shares of naturally unavailable land, and historical housing prices as instruments for housing prices. When the land supply

[1]The coefficient of log housing prices on the log SIR is significant at the 10% level in the cross-section regression using data in 2000.

[2]Chen and Zhang (2016) show that China's higher education expansion has increased the shares of college educated workers dramatically and the housing prices increase by between 4.6% and 7.9% when the share of college educated people increases by 1%.

is constrained by exogenous natural conditions, the inelastic housing supply will cause faster housing price growth following demand shocks. The identification assumption behind the use of this instrument is that the housing supply constraint affects the SIR only through its effects on housing costs after controlling for related factors such as population, employment shares across sectors and amenities. As the local housing supply varies much across cities in China, this chapter is able to identify the causal effect of housing costs on the SIR by controlling for common factors and avoid the reverse causality. In addition, I use historical housing prices in 2000 as another instrumental variable for housing prices in 2010. The housing price in 2000 is the earliest data I can find for the housing price in available public datasets. I believe that with 10 years of time gap, the housing price in 2000 is unlikely to affect the share of high-educated workers in total demand in 2010 except through the channel of housing prices in 2010.

Housing costs might have different effects on wages of high-skilled and low-skilled workers, which could in turn influence the SIR across cities. Assuming that the income elasticity of housing demand is well below one, Kim et al. (2009) developed a spatial equilibrium model based on the Rosen-Roback model (Rosen 1974, Roback 1982) and demonstrated that changes in wage differentials between high-skilled and low-skilled workers are associated with different growth rates of living costs between different types of workers in one city. The key mechanism is as follows: if the income elasticity of housing demand is well below unity, rising housing prices will have relative larger negative impacts on low-skilled workers than on high-skilled workers, because housing cost accounts for a larger fraction of incomes for low-skilled workers. Under the spatial equilibrium condition, low-skilled workers must be paid relatively higher wages in order to compensate for higher living costs, otherwise they would move to other cities with lower housing costs. It follows that, firms have motivation to employ a larger fraction of high-skilled workers because of the decline of their relative wages. Therefore the SIR rises due to higher housing cost. Using tabulations of decennial census data in the U.S. over 1980-2000, they find that rising housing prices have a negative and significant effect on wage gap between high-skilled and low-skilled workers, which confirms the validity of their theoretical explanations.

This chapter complements the literature on the Rosen-Roback spatial equilibrium model. Rosen (1979) created a model of inter-city wage differences and made a hypothesis that migration across metropolitan areas should cause equality of indirect utility. In the original model, Rosen (1979) assumed that all workers are homogeneous. The model suggests that wage differentials across cities arise due to relative price or amenity differentials associated with inter-city different characteristics including city size or growth rate. This model was extended to workers varying with education levels by Beeson (1991) and the essential conclusion still holds. This chapter assumes that workers are

heterogeneous in terms of skills, hence housing demand, and provides new evidence for the validity of Rosen's model and for the understanding the role of housing price in labor demand composition.

In addition, as far as I know, this is the first essay to collect all Chinese cities' topographical data to construct a reasonable instrument for housing prices. I obtain the digital elevation data (30m resolution) from the ASTER Global Digital Elevation Model Version 2 and calculate shares of land slope above 30% after excluding bodies of water and areas outside national boundaries in a 35 km radius. Following the insight of Saiz (2010), Dong (2016) collects data of geographical land constraints of 35 major cities in China. Shi et al. (2016) also calculate the housing supply elasticity of 35 major cities using data from the China Data Center of the University of Michigan. But no one has calculated shares of unavailable land for all Chinese cities before. The topographical data can be used to explore the social and economic impacts of housing prices in China in the future.

The remaining of this chapter is organized as follows: the next section briefly discusses the previous literature. The section 4.4 describes the details of the administrative system in China and geographical scales of variables used in the paper. Section 4.5 presents the identification strategy and justifies the validity of the instruments. Section 4.6 introduces the data used in this chapter. Section 4.7 shows summary statistics for variables in levels. Then, section 4.8 offers empirical results which confirm the role of variation in housing cost in influencing changes in the SIR of Chinese cities. Section 4.9 shows robustness checks after dropping four mega cities from the sample. Finally, section 4.10 concludes.

4.3 Literature review

Housing costs exert indirect effects on the SIR through influencing wage differentials between high-skilled and low-skilled workers at the city level. Moretti (2004) studies externalities of human capital and finds that the rising proportion of college-educated workers is able to raise wages of high-school dropouts more than that of college-educated workers, which implies a strong complementarity between high-skilled and low-skilled workers. In other words, the rising SIR is able to reduce differences in wages between high- and low- skilled workers. Furthermore, Moretti (2013) finds that the real wage differences between high-skilled and low-skilled workers have increased less than the nominal differences, because high-skilled workers need to pay more for housing costs as more and more college graduates concentrate in big cities associated with high cost of housing in recent years. Black et al. (2009) use the 1990 U.S. census to find that returns to education are relatively low in big cities such as San Francisco and Seattle, because high housing prices countervail relatively high nominal wages. Kim et al. (2009) also use U.S. censuses to investigate the effect of housing prices on wage

differentials among workers with or without bachelor's degree and conclude that high housing costs decrease the Skill Wage Ratio (SWR).[3]

More recently, Broxterman and Yezer (2015) study how housing costs indirectly influence the SIR using U.S. censuses data from 1970 to 2000. Their estimation results confirm the prediction of the Rosen's extended model and draw the conclusion that high housing costs have positive effects on the SIR. However, as far as I know, no evidence from developing countries including China has been used to test this empirical prediction.

Several other possible causes including city size, industry mix, local amenities and taxes have been used to explain spatial variation of the SIR across cities in the literature. Elvery (2010) tests whether skill intensity increases with city size based on data extracted from the Occupational Employment Statistics survey in the U.S.. He finds that firms in big cities employ a more skill intensive mix of labor than similar firms in small cities. In addition, he also finds that workers disproportionately concentrate in high-skilled industries in big cities. But both population and industry mix account for a small part of skill intensity differences across cities. Similarly, Hendricks (2011) estimates that differences in the sector composition explain only 20% of the skill intensity variation across cities. Using U.S. data from 1980 to 2010, Brinkman (2014) finds that industry mix can account for only a tiny proportion of differences in the relative demand for skills across cities. Therefore, population and industry mix are not all the factors to explain why the skill intensity ratio varies largely across cities. Gyourko et al. (2013) think that cities with local delicate amenities, which are based on topography and climate, are attractive disproportionately to high-income households. As the income level is highly correlated with educational level, high-quality amenities are able to result in a higher SIR. Conversely, amenities such as crime and congestion have negative impacts on the SIR. The demand for local public services such as quality of public schools varies with household educational levels. The local taxes might have different impacts on residents with different incomes generally based on education levels.

Rosen's extended model relies on the assumption that the income elasticity of housing demand is well below unity. Several empirical papers have found that this assumption holds true in China. Zheng and Liu (2005) use simple OLS and WLS specifications to estimate the income elasticity of housing demand using micro survey data from Beijing, and the results show that the income elasticity for owner-occupiers is 0.86, which may be biased because of unobservable characteristics. Based on a household survey conducted in Shanghai, Chen and Jin (2014) use a two-stage model to alleviate the problem of sample selection induced by the endogeneity of tenure choice, and find the permanent

[3]It should be noted that Wheeler (2004) finds a negative and significant relationship between population and the SWR at the city level, but without controlling for housing prices, which may cause a biased estimation result.

income elasticity of housing demand for owner-occupiers in Shanghai to be between 0.375 and 0.447, which is far below unity. While Chow and Niu (2015) utilize a simultaneous equations framework to find the income elasticity of housing demand in urban China to be approximately 1 in the long run, Murray and Sun (2017) develop a derived demand framework and find the elasticity of housing services demand with respect to permanent income to be 0.72. Given that Chow and Niu (2015) use aggregate annual data at the national level from 1987 to 2012 and Murray and Sun (2017) employ household data from Chinese Health and Nutrition Surveys during a long period (1991-2006), one can consider the latter result as more credible.

4.4 Administrative system in China

Within China, there are four tiers of administrative divisions: the province-, prefecture-, county- and township- levels. Until the end of 2015, Mainland China had 31 province-level regions, 334 prefecture-level regions, 2580 county-level regions and 39789 village-level regions (Yearbook et al., 2016). Specifically, the province-level divisions consist of 4 municipalities, 23 provinces and 5 autonomous regions; the prefecture-level divisions consist of 291 prefecture-level cities, 10 prefectures, 30 autonomous prefectures and 3 leagues. It should be noted that the municipalities and prefecture-level cities have similar administrative structure, except that the former is directly under the control of the central government, so I treat municipalities as a special type of prefecture-level cities and also call them prefecture-level cities for brevity. In my analysis, I focus exclusively on the prefecture-level cities for two reasons. Firstly, unlike prefecture-level cities, other types of prefecture-level regions do not include any cities. Secondly, no available dataset for amenities could cover these regions.

A typical prefecture-level city consists of three types of county-level regions: (1) urban districts (*Shixiaqu* in Chinese), (2) county-level cities (*Xianjishi* in Chinese) and (3) counties (*Xian* in Chinese) . Figure A1 represents an example of a typical prefecture-level city in China. Baoding, located in Hebei Province, consists of 5 urban districts, 4 county-level cities and 15 counties in 2015. It should be noted that prefecture-level cities in China are expansive regions, covering both urban areas and surrounding rural areas in the sense that they are actually equivalent to departments in France or counties in the U.S. (Chan, 2007). Prefecture-level city's urban districts, which are contiguous areas representing a "city proper" (Roberts et al., 2012), constitute the urban core of this prefecture-level city, although, in many cases, rural and agricultural residents make a non-negligible share of total population within it. Urban districts usually occupy a fraction of the prefecture-level city's area as shown in Figure A1. Two separate statistics, one aggregated for the urban districts level and another covering the whole

administrative areas of the prefecture-level city, are reported in the China City Statistics Yearbook. In my analysis, I use the former statistics of prefecture-level cities since the geographical areas counted is close to the conventional conception of the city in western countries.

Ideally, all variables should be generated at the level of urban districts. However, the real estate price is only available from the Chinese Regional Economic Statistical Yearbook series at the level of whole prefecture-level city, so I have to use it instead. As urban districts are the economic center of the prefecture-level city, real estate transactions within urban districts represent large shares of the total real estate markets in the whole prefecture-level city, the real estate price measure at the level of prefecture-level city should not have much discrepancies with it at the level of urban districts. Zheng and Kahn (2013) calculated the correlation of the real estate price measure with the quality-controlled hedonic price index compiled by the Institute of Real Estate Studies at Tsinghua University,[4] and the results show the correlation coefficient is high (0.90) for 35 big cities for which both price data are available.

4.5 Specifications and hypothesis

My empirical model aims to explore effects of housing prices on the share of high-skilled/low-skilled workers and the SIR at the city level. Following Broxterman and Yezer (2015), I first describe how rising housing prices affect the share of high-skilled workers at the city level using the following regression equation:

$$lnSHHW_{it} = \alpha_h lnHC_{it} + \beta_h lnPop_{it} + \boldsymbol{\gamma_h} Empl_{it} + \boldsymbol{\delta_h} InAmen1_{it} + \boldsymbol{\lambda_h} ShAmen2_{it} + \epsilon_{iht} \tag{4.1}$$

where $SHHW_{it}$ represents the share of high-skilled workers in adults in city i at time t, HC_{it} average housing price, P_{it} population, $Empl_{it}$ a vector of employment shares across industries, $Amen1_{it}$ a vector of amenities which are amounts, $Amen2_{it}$ represents a vector of amenities which are shares, and ϵ_{iht} a specific error which varies with city and time. The coefficient of interest, α_h, is interpreted as the percentage change in the share of high-skilled workers corresponding to a percent increase in housing price. Bacolod et al. (2009) show that large cities have skill-biased productivity advantages, so a larger city attracts disproportionately more skilled workers. Elvery (2010) finds that the skill intensity of human capital increases with city size, even conditioning on the industry mix. Hendricks (2011) finds that differences in the industry mix explain 20% of the skill intensity variation across

[4] Zheng et al. (2010) introduce the details of the compiling methodology of this quality-controlled hedonic price index.

cities.[5] Local amenities are also factors attracting high-skilled workers. That is why I control for city size, industry mix and amenities in Equation (4.1). Local taxes effects, local unobserved amenities and public goods provided by local governments, which are correlated to both housing cost and the SIR, are incorporated into ϵ_{iht}. Noted that information about average housing quality is not available, so it will be captured in ϵ_{iht}. A similar equation can be written for the share of low-skilled workers LW_{it}.

$$lnSHLW_{it} = \alpha_l lnHC_{it} + \beta_l lnPop_{it} + \gamma_l Empl_{it} + \delta_l InAmen1_{it} + \lambda_l ShAmen2_{it} + \epsilon_{ilt} \tag{4.2}$$

It then follows:

$$lnSIR_{it} = lnSHHW_{it} - lnSHLW_{it} = \alpha lnHC_{it} + \beta lnPop_{it} + \gamma Empl_{it} + \delta InAmen1_{it} + \lambda ShAmen2_{it} + \epsilon_{ilt} \tag{4.3}$$

where $\alpha = \alpha_h - \alpha_l$, $\beta = \beta_h - \beta_l$, $\gamma = \gamma_h - \gamma_l$, $\delta = \delta_h - \delta_l$, $\lambda = \lambda_h - \lambda_l$, and $\epsilon_{it} = \epsilon_{iht} - \epsilon_{ilt}$, in which h and l indicate high-skilled and low-skilled workers respectively. The parameter of interest is α, which measures impacts of housing prices on the SIR.

I conduct a cross-sectional regression for each year separately, not in time variations, because obtaining stable coefficients are much more difficult in time-variant data than in cross-section data. For instance, if the estimated coefficient is not stable over time, i.e. different for 2000 and 2010, it would be impossible to obtain something very consistent in time difference. Since 2000, China has been experiencing rapid economical and social development, and a lot of factors used in this chapter have changed from 2000 to 2010. Actually, cross-sectional regression results shown in the following reflect that the estimated coefficients of interest are indeed not consistent over time.

Given that the income elasticity of housing demand is well below unity, the heterogeneous Rosen-Roback model predicts that the parameter α_h in Equation (4.1) should be positive, α_l in Equation (4.2) should be negative, and α in Equation (4.3) should be positive. Firms will indeed employ a larger share of high-skilled workers and a smaller share of low-skilled workers because housing costs account for a smaller proportion of incomes for high-skilled workers than for low-skilled workers.

As discussed in the introduction section, estimating Equation (4.3) by OLS is likely to result in a

[5] It is equivocal that how large the effect of industry mix on the SIR. For instance, using U.S. data from 1980 to 2010, Brinkman (2014) finds that industry mix can account for only a tiny proportion of differences in the relative demand for skills across cities.

biased estimation result of a causal effect of housing prices on the SIR because reverse causality issues and possible omitted variables.[6]

To address the endogeneity concerns, I use both land unavailability and historical housing prices to instrument for current housing prices. Saiz (2010) points out that the land unavailability measures the land supply elasticity because it is positively linked with housing construction costs. If a city's land slope is steep, its land supply will be less than that of cities with flat land slope when facing an exogenous national housing demand shock, so the housing prices will increase faster. In the literature, using data in 35 Chinese major cities from 1998 to 2009, Wang et al. (2012) present that the supply elasticity decreases approximately 0.15 percentage points when undevelopable Land share increases 1 percentage. Thus when an exogenous housing demand happens, areas with geographic barriers experience faster growth rate of house price. I also use housing prices in 2000 to instrument for housing prices in 2010. As the public available data on housing prices at the city level start from 2000, I can only conduct a 2SLS regression in 2010 using historical housing prices as one of instruments for housing prices. Naturally, housing prices in 2000 can predict the housing prices in 2010 well. The formal housing market began in 1998 in China due to a housing reform conducted by the Chinese government, therefore the housing market was still at the early stage by 2000. The housing prices in 2000 are not affected by the socio and economics development from 2000 to 2010 which could affect shares of high-skilled or low-skilled workers at the city level. In sum, the housing prices in 2000 and land unavailability together can predict the housing prices in 2010 well. Given historical housing prices, the land supply elasticity measured by the land unavailability is lower, housing prices increase faster and the current housing prices will be higher.

Following Saiz (2010) and Dong (2016), I use the share of land slope above 30 degree within a 35 km radius from each city centroid to measure the land unavailability. I got the digital elevation data (30m resolution) from the ASTER Global Digital Elevation Model Version 2 and compute shares of land slope above 30% within the 35km radius that are lost to bodies of water and areas outside national boundaries. The radius I choose to calculate the degree of land unavailability is larger than the actual radius of all cities except for Beijing, Shanghai, Guangzhou and Shenzhenm, which is smaller than 50 km used for all U.S. metropolitan areas by Saiz (2010). The use of the 35km raidus makes the measure of land unavailability exogenous on urban sprawl.

Although the real interest rate is often used to instrument for a national housing cycle in the literature (Mian and Sufi (2011) and Chaney et al. (2012)), I found that the aggregate movements of house prices in China does not follow the fluctuations in real interest rates. From 2002 to 2013, the average increase

[6]The endogeneity issue is the same when the dependent variable is log of share of high-skilled or low-skilled workers.

in the real interest rate was -0.33 percentage points, while real house prices increased by 7.9-13.1 percent annually in most cities.(Fang et al., 2015) In addition, china's interest rate is modulated by People's Bank of China. Since the bank lending mortgage rate is identical across different regions and any changes in the rate are released simultaneously, I believe there is a similar impact for mortgage rate on housing supply decision across regional markets.

4.6 Data

Information on share of high-skilled workers, share of low-skilled workers, SIR, population and industry mix are obtained from two waves of census data. Specifically, I use the 2000 and 2010 county-level tabulations of censuses released by the National Bureau of Statistics of China (NBS). I prefer to use the aggregate census data rather than the 1/1000 micro sample of 2000 and 2010 censuses, which I could get access to, because of two reasons. First, the four-digit address code provided in the micro sample can only identify prefecture-level units but not county-level units, so I cannot restrict the sample to individuals living in only urban districts of prefecture-level cities. If I use the urban-rural identifier to identify urban areas in the micro sample, the urban area of county-level cities and urban districts would be included into urban areas. Second, perhaps more important, when I restrict the micro sample to individuals living in urban areas, the sample size for small prefecture-level cities is limited. Table A1 shows that 75% of cities have less than 31 college educated individuals in the one thousand 2000 micro census, which implies insufficient precision in estimating the SIR in a given city.

Conventionally, the SIR should be defined as the ratio of college graduates to non-college graduates among adults aged 25 years old or above. However, the county-level tabulations of censuses report only population by education background for people aged 6 years old or above, thus I have to revise the definition of the SIR. Considering that only a small proportion of the population aged under 20 is college-educated, the total number of non college-educated people minus population aged 6-19 could be regarded as the number of non college-educated people aged 20 years old or above. But another problem is that the county-level tabulations of censuses show population by gender and five-year age groups such as 0-4, 5-9, 10-14, etc. I calculate the share of people aged 5 years old in the cities' population based on micro-level censuses, and then compute the population aged 6-19 in the censuses at the county-level following the formula: $pop_{(6,19)} = (male_{(5,9)} + male_{(10,14)} + male_{(15,19)} + female_{(5,9)} + female_{(10,14)} + female_{(15,19)}) * (1 - share_age_5)$.[7] Therefore, the SIR in this chapter is defined as the ratio of individuals who attending 4-year formal universities to someone

[7]Where $male_{(5,9)}$ denotes the number of male aged 5-9, $male_{(10,14)}$ denotes the number of male aged 10-14, etc. and $share_age_5$ denotes the share of population aged 5 years old.

else among people aged 20 years old or above. Using the micro censuses data in 2000 and 2010, it can be shown that the SIR among people aged 20 years old or above is highly correlated with that among people aged 25 years old above at the province level and the coefficient of correlation is 0.94.[8] The share of high-skilled workers is defined as the ratio of individuals who attending 4-year formal universities to people aged 20 years old or above, and the share of low-skilled workers is defined as the ratio of individuals who are not college-educated to people aged 20 years or above.

The county-level tabulations of censuses also provide the number of employees across industrial sectors, in which employment shares are calculated as the number of employees divided by the total working population at the city level.[9] Consistent with Broxterman and Yezer (2015), I control for the employment share of manufacturing, trade & catering services, and professional services. In particular, professional services contains finance, research & environmental protection, and transportation & communication.

The total revenue and sale quantity of real estate for each prefecture-level city are drawn from the China Statistical Yearbook for Regional Economy (2001 and 2011), which is the only available source of data on housing price in China at the prefecture level during the period 2000-2010. Average real estate prices are calculated by dividing total revenue of real estate by sale quantity of real estate. In effect, the more appropriate variable is the average residential housing price, because real estate buildings include buildings for business use, which reflect labor demand rather than the consumer side. But information on average residential housing price was not recorded in 2000, so I have to use average real estate prices as proxies of housing demand.[10] Nevertheless, the correlation coefficients between prices of real estate and housing prices are larger than 0.99 in 2010,[11] thus it is reasonable to infer that they are also highly correlated in 2000.

Due to fast urbanization in recent years, many counties or county-level cities were upgraded to urban districts of prefecture-level cities. Table A2 listed all county-level unities which were transformed to urban districts during the period 2000-2010. If geographical boundaries of urban districts are not held constant over time, the measurement of variables used in 2010 will have errors. I take geographical definitions in 2000 census and apply them to urban districts areas in 2010 census.

[8]Source: Author's calculation based on the one thousand micro censuses data of China in 2000 and 2010

[9]It should be noted that the China City Statistical Yearbook series also provide data on population and the number of employment at the urban districts level but it omits people lacking local *hukou* registration when counting population and does not take most self-employed or private sector's workers into employment across industry. Therefore, unlike some prior studies (such as Au and Henderson (2006) and Liang et al. (2016)), I obtain information on population and industry mix from micro censuses instead of China City Statistical Yearbook.

[10]It is noted that the sales of commercial real estate is largened 10 times by error for all cities of Hunan Province in 2010. I correct this error.

[11]Source: Author's calculation based on the China Statistical Yearbook for Regional Economy (2001)

Table A3 summarizes the definition, geographical scale, source, whether the geographical boundary is adjusted of all variables that I use in the paper. It is noted that variables drawn from the Statistical Yearbook for Regional Economy and City Statistical Yearbook are reported directly either at the urban district level or prefecture level, so I cannot process these variables by controlling for boundary changes over time. In addition, the employment classification changed in 2002, which makes employment data across industrial sectors inconsistent before and after 2002. Table A4 reclassifies employment sectors into 13 categories in 2000 and 2010.

Finally, to construct balanced panel data at the city level, I restrict my dataset to 257 prefecture-level cities. Six prefecture-level cities including Baoshan, Chaohu, Jiayuguan, Jinchang, Lhasa and Sanya are dropped, because they are attributable to missing values of real estate prices or amenities.

4.7 Descriptive statistics

Table 4.1 shows descriptive statistics for all variables in levels in 2000 and 2010. The average SIR increased from 0.038 to 0.088 during the period 2000-2010, while the coefficient of variation (CV) did not change much. The rapid increasing rate of SIR is partly caused by China's higher education expansion since 1998. The share of high-skilled workers in adults aged 20 years old or above is 0.035 in 2000, and it became 0.078 in 2010. Because the shares of high-skilled and low-skilled workers are computed depending on the estimated share of people aged 5 years old from the aggregate census at the county level as I mentioned in the data section, their values have measurement errors. For example, the sum of shares of high-skilled and low-skilled workers does not exactly equal one. In some cases, the computed share of low-skilled workers exceeds unity. I regress the residuals, which equals one minus the sum of shares of high-skilled workers and low-skilled, on log population for year separately, and do not find significant correlations between them as shown in Figure 4.1.

At the same time, the average housing price increased by more than 3 times in the decade. The housing price appreciation rate can be comparable with the housing price indices constructed based on sequential new housing sales within the same development project at the city level in Fang et al. (2015). They find that the housing price of the second- and third-tier cities rose about three-fold between 2003 and 2013.

Due to accelerated urbanization rate, the average population at the city level increased from 0.94 million to 1.32 million from 2000 to 2010. The employment shares by industries show that the manufacturing share declined, but professional service share, trade and catering services share increased.

Table 4.1: Summary statistics for variables in levels

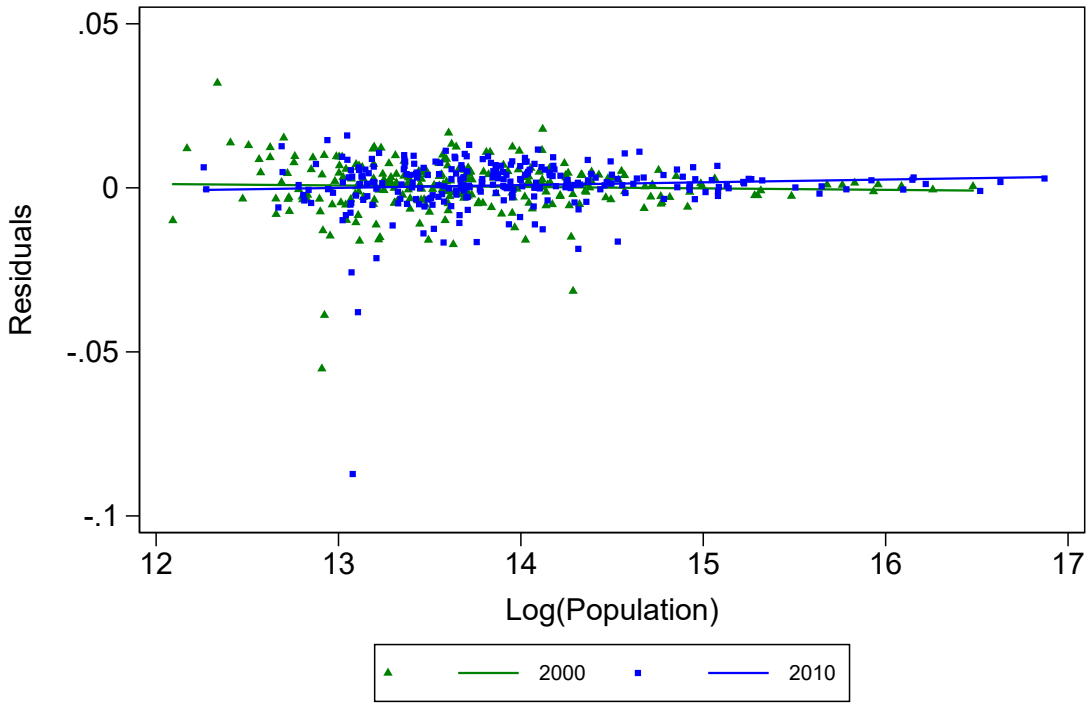
	2000						2010					
	Mean	Min	Median	Max	SD	CV	Mean	Min	Median	Max	SD	CV
Skill intenity ratio	0.038	0.003	0.030	0.157	0.030	0.792	0.088	0.013	0.075	0.331	0.055	0.631
Share of high-skilled	0.035	0.003	0.029	0.136	0.027	0.749	0.078	0.013	0.070	0.249	0.044	0.563
Share of low-skilled	0.964	0.865	0.970	1.030	0.027	0.028	0.921	0.750	0.928	1.044	0.045	0.049
House price	1274	469	1119	5718	651	0.511	3872	1725	3094	19170	2407	0.622
Population (in thousands)	944	127	592	11473	1292	1.368	1324	161	755	18402	2009	1.518
Share of employment												
Professional service	0.072	0.011	0.071	0.162	0.029	0.404	0.098	0.025	0.096	0.191	0.035	0.356
Manufacturing	0.213	0.022	0.201	0.756	0.122	0.572	0.191	0.024	0.164	0.743	0.118	0.619
Trade and catering services	0.125	0.035	0.126	0.247	0.046	0.369	0.186	0.077	0.182	0.333	0.053	0.286
Amenities												
Green space ratio (%)	28.591	2.000	30.200	61.600	10.074	0.352	39.284	2.410	39.940	68.510	6.823	0.174
Student-teacher ratio	20.558	8.936	19.638	63.601	5.283	0.257	17.544	0.959	17.473	26.973	3.866	0.220
# Books in public libraries per 1000 persons	0.556	0.010	0.400	3.771	0.492	0.886	0.667	0.025	0.507	3.542	0.562	0.844
# Doctors per 1000 persons	2.602	0.485	2.506	10.070	1.217	0.468	2.734	0.519	2.535	10.737	1.179	0.431
Number of college students in schools	20211	0	6666	251935	40098	1.984	81996	0	33941	881433	141869	1.730
Observations	257						257					

Note: The Skill intensity ratio (SIR) is defined as the ratio of individuals who attending 4-year formal universities to someone else among people aged 20 years old or above. The share of high-skilled workers is defined as the ratio of individuals who attending 4-year formal universities to population aged 20 years old or above. The share of low-skilled workers is defined as the ratio of individuals who are not college-educated to population aged 20 years old or above. College students in schools contain students who are attending three-year or four-year universitiies, and postgraduates on campus. Unit: Housing price: Yuan/m^2; Population: 1 million Yuan; Share of employment: %; Green space ratio: %. Yuan is the name of a unit of the Chhinse currency.

In the first decade of the 21st century, the average green space ratio of Chinese cities increased from 28.6% to 39.3%, which reflects that the environment improved. The student-teacher ratio in primary schools decreased from 20.6 to 17.5, which means that there are more teachers per pupil in primary schools on average. The number of books in public libraries per 1000 persons, the number of doctors per 1000 persons and the number of college students in schools also increased from 2000 to 2010, which indicates that the cultural facilities, health facilities, and higher education resources expanded.

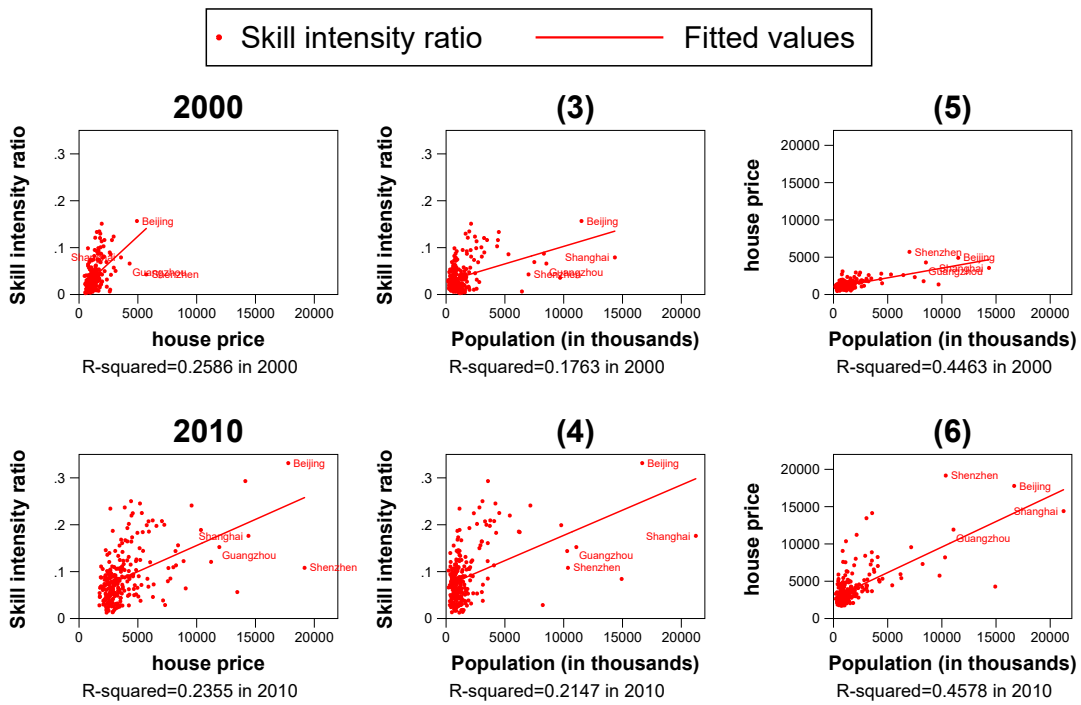
To show correlations among SIR, house price and population in 2000 and 2010 respectively, I draw scatter plots and regression lines in Figure 4.2. Consistent with the related theory, the SIR has a positive association with housing prices. I also find a positive relations between the SIR and population at the city level in 2000 and 2010. In addition, housing prices are also positively related with population as expected. Overall, the figures indicate that the three variables are positively correlated each other. It suggests that I have to control for population when I study the effects of house prices on the SIR.

Figure 4.1: Correlations between residuals and log population by year



Note: Residuals equal one minus the sum of shares of high-skilled workers and low-skilled.

Figure 4.2: Correlations among housing prices, population and SIR



Note: Residuals equal one minus the sum of shares of high-skilled workers and low-skilled.

4.8 Empirical results

4.8.1 OLS results

Baseline results from OLS estimates in Table 4.2 report how changes in housing prices are correlated with changes in the share of high-skilled workers, and that of low-skilled workers, and the SIR in 2000 and 2010 separately. The coefficients for the three dependent variables are shown in separate groups in the table.

Column (1) in Panel A shows that housing prices have significant and positive correlations with the share of high-skilled workers in 2000 if nothing else is controlled for. After controlling for population, the coefficient of log housing prices does not change much as shown in Column (2) of panel A. Again, adding amenities into the regression, the results presented in Column (3) of panel A exhibit that the coefficient of interest is reduced to 0.424, which is in line with expectation. However, when I control for employment shares across sectors further, the effect of housing prices on the log share of high-skilled workers becomes 0.204 and significant only at the 10% level in 2000 reported in Column (4) of panel A. Specifically, the rising employment shares in trade and catering services and professional service increases the share of high-skilled workers, while the rising employment share in manufacturing decreases the share of high-skilled workers. These results except trade and catering services employment shares are well understood, because manufacturing sector requires for non-college-educated workers, and professional services sector needs more college-educated workers.

Surprisingly, after controlling for population, employment shares across sectors and amenities, housing prices have no impacts on the share of low-skilled workers in 2000 as shown in Column (8) of Panel A. Finally, the coefficient of log housing prices on log SIR is 0.204% and significant only at the 10% level in 2000 shown in Column (12) of Panel A. Population has positive relationship with the SIR, consistent with the findings in Elvery (2010). Overall, effects of housing prices on the shares of high-skilled workers, low-skilled workers, and the SIR are not so significant in 2000. One potential reason is that the *hukou* regulation was tight by 2000 and mobility costs were high.(Chan and Zhang, 1999) The *hukou* system is like an internal passport system within China. People's *hukou* status (local/non-local, urban/rural) is tied to their rights to housing, schooling, health care and employment opportunities. Before 2000, it was very difficult to change *hukou* status from rural to urban or from one region to another. People could move across cities without changing their *hukou* but they can benefit from little local public services at the destination place.(Au and Henderson, 2006) The large mobility costs across cities violate the assumption of Rosen-Roback model, therefore hous-

Table 4.2: OLS regression results in 2000 and 2010

	Log share of high-skilled				Log share of low-skilled				Log SIR			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: 2000												
Log(Housing prices)	1.003***	1.000***	0.424***	0.204*	–0.037***	–0.028***	–0.008*	0.000	1.040***	1.028***	0.432***	0.204*
	(0.120)	(0.136)	(0.103)	(0.105)	(0.005)	(0.005)	(0.004)	(0.004)	(0.124)	(0.140)	(0.107)	(0.108)
Log (Population)		0.003	0.025	0.105**		–0.008***	–0.012***	–0.015***		0.011	0.037	0.120***
		(0.060)	(0.042)	(0.041)		(0.002)	(0.002)	(0.002)		(0.062)	(0.043)	(0.043)
Manufacturing				–0.334				0.040***				–0.374
				(0.305)				(0.012)				(0.312)
Trade and catering services				3.451***				–0.199***				3.650***
				(0.968)				(0.045)				(0.998)
Professional service				3.780**				–0.135**				3.915**
				(1.514)				(0.064)				(1.565)
Amenities	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Observations	257	257	257	257	257	257	257	257	257	257	257	257
Adjusted R^2	0.279	0.277	0.723	0.764	0.268	0.297	0.580	0.653	0.282	0.279	0.722	0.764
Panel B: 2010												
Log(Housing prices)	0.594***	0.417***	0.263***	0.220***	–0.058***	–0.031***	–0.020**	–0.022***	0.652***	0.448***	0.283***	0.242***
	(0.072)	(0.089)	(0.076)	(0.080)	(0.008)	(0.009)	(0.008)	(0.008)	(0.080)	(0.097)	(0.083)	(0.086)
Log (Population)		0.150***	0.041	0.061		–0.022***	–0.015***	–0.016***		0.172***	0.056	0.077*
		(0.051)	(0.046)	(0.041)		(0.005)	(0.004)	(0.004)		(0.055)	(0.049)	(0.044)
Manufacturing				–0.439*				0.087***				–0.526*
				(0.265)				(0.022)				(0.281)
Trade and catering services				1.663**				–0.215***				1.878**
				(0.761)				(0.055)				(0.807)
Professional service				4.292***				–0.262***				4.553***
				(1.289)				(0.095)				(1.369)
Amenities	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Observations	257	257	257	257	257	257	257	257	257	257	257	257
Adjusted R^2	0.193	0.214	0.572	0.662	0.247	0.316	0.553	0.669	0.200	0.223	0.575	0.668

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

ing costs do not have strong relationships with the share of high-skilled workers, that of low-skilled workers and the SIR in 2000. After 2000, Chinese government conducted some *hukou* reforms and relaxed the *hukou* restriction on workers' mobility, which allows for more elastic labor supply responses to housing costs.

Column (4) of panel B in Table 4.2 reports that one percentage increase in housing prices leads to 0.220% increase in the share of high-skilled workers in 2010, after controlling for log population, employment shares across sectors and amenities. Meanwhile, one percentage increase in housing prices is related to 0.022% decrease in the share of low-skilled workers in 2010. As indicated in Equation (4.3), when housing prices increase by 1%, the SIR will increase by 0.242% in 2010, as predicted in the Rosen-Roback model.

The coefficients of housing prices on the log SIR change from 0.204 to 0.242 during the period 2000-2010, and the corresponding significant levels changed from 10% to 1%. The coefficients of log population decreased from 0.120 to 0.077 in time time period and the significant degree also changed. In addition, the coefficients of employment shares across sectors also changed over time. Overall, the results in Table 4.2 show that coefficients are not stable over time, suggesting that it is not proper

to use the first-difference model to examine effects of housing prices on the share of high-skilled workers, low-skilled workers, and the SIR respectively. Therefore, I insist to use the cross-sectional regression in this chapter.

4.8.2 IV results

I use both historical housing prices in 2000 and topographical factor (share of land slope above 30 degree within a 35km radius) as instruments of housing prices in 2010 in Table 4.3. The land slope is one of determinate factors of housing supply elasticity. When the local housing supply elasticity is inelastic, national housing demand shock could translate into faster housing prices increases. If historical housing prices are given, the housing prices in levels will be higher. The exclusion restriction is that the historical housing prices and housing supply elasticity affect the share of high-skilled workers, that of low-skilled workers, and the skill intensity ratio only through the channel of current housing prices.

The two instruments are relevant to housing prices in 2010 as shown in Table 4.4. The IV results show that housing prices have positive and significant impacts on the share of high-skilled workers and the SIR, but negative and insignificant impacts on the share of low-skilled workers in 2010. When housing prices increase by 1%, the share of high-skilled workers will increase by 0.920%, the share of low-skilled workers will decrease by 0.022%, and the SIR will increase by 0.327%. The first-stage F statistics is quite large, and the values of Hansen J statistics indicate that the null hypothesis on validity of instruments is not rejected at the 5% level.

The first-stage regression in Table 4.4 presents a significant and positive association between land unavailability and growth rate of house prices. When the share of undevelopable land increases by one percentage, the growth rate of house price will increase by 1.188 percentage points. Housing prices in 2000 is positively related to the housing prices in 2010. In addition, the R square in Column (4) is 0.729, implying that housing supply constraints and other controlled variables can explain the volatility of house price in a large extent. The value of F statistics for the excluded instrument is 38.147, which exceeds Stock and Yogo (2005) threshold on the maximum IV estimator biased at the 15 percent level. The results for Hansen J Statistics show that the overidentification restrictions are valid.

Table 4.3: IV regressions in 2010

	Log Share of high-skilled			Log Share of low-skilled			Log SIR		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log(Housing prices)	0.920***	0.439***	0.305**	–0.065***	–0.029***	–0.022*	0.986***	0.469***	0.327**
	(0.170)	(0.129)	(0.142)	(0.013)	(0.011)	(0.013)	(0.182)	(0.139)	(0.152)
Log (Population)	–0.031	–0.013	0.038	–0.010*	–0.012***	–0.016***	–0.021	–0.001	0.054
	(0.068)	(0.054)	(0.052)	(0.005)	(0.005)	(0.005)	(0.072)	(0.057)	(0.056)
Manufacturing			–0.558*			0.087***			–0.645**
			(0.298)			(0.025)			(0.318)
Trade and catering services			1.551**			–0.215***			1.766**
			(0.736)			(0.056)			(0.783)
Professional service			4.192***			–0.262***			4.453***
			(1.269)			(0.094)			(1.350)
Constant	–9.997***	–7.281***	–7.003***	0.516***	0.305***	0.302***	–10.513***	–7.586***	–7.305***
	(1.080)	(0.847)	(0.880)	(0.089)	(0.078)	(0.080)	(1.160)	(0.914)	(0.941)
Amenities	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	257	257	257	257	257	257	257	257	257
Adjusted R^2	0.132	0.562	0.660	0.267	0.549	0.669	0.144	0.566	0.667
First-stage F Statistics	76.876	67.552	38.145	76.876	67.552	38.147	76.876	67.552	38.147
Hansen J Statistics	2.224	2.07	2.956	1.887	1.319	3.395	2.218	2.049	3.057

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

Note: The instrumental variables are housing prices in 2000 and the share of land slope above 30 degrees within a 35km radius from a city centroid. Hansen J statistics is used for testing over-identifying restrictions.

Table 4.4: First stage regression results in 2010

	(1)	(2)	(3)
Log(Housing prices in 2000)	0.667***	0.683***	0.540***
	(0.054)	(0.059)	(0.064)
The share of land slope above 30 degree within a 35 km radius	1.226***	1.219***	1.188***
	(0.406)	(0.400)	(0.380)
Log (Population)	0.163***	0.142***	0.153***
	(0.026)	(0.027)	(0.026)
Manufacturing			0.901***
			(0.182)
Trade and catering services			0.511
			(0.464)
Professional service			0.342
			(0.668)
Constant	2.292***	2.158***	2.991***
	(0.303)	(0.362)	(0.410)
Amenities	No	Yes	Yes
Observations	257	257	257
Adjusted R^2	0.673	0.688	0.729
First-stage F Statistics	76.876	67.552	38.145

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

Table 4.5: IV regressions after dropping four first-tier cities

	Log Share of high-skilled			Log Share of low-skilled			Log SIR		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log(Housing prices)	1.016***	0.489***	0.358**	–0.072***	–0.034***	–0.025*	1.087***	0.522***	0.383**
	(0.185)	(0.139)	(0.155)	(0.014)	(0.012)	(0.014)	(0.197)	(0.149)	(0.167)
Log (Population)	–0.022	–0.007	0.047	–0.011**	–0.012**	–0.016***	–0.012	0.005	0.064
	(0.069)	(0.054)	(0.052)	(0.005)	(0.005)	(0.005)	(0.074)	(0.058)	(0.055)
Manufacturing			–0.632**			0.085***			–0.717**
			(0.313)			(0.026)			(0.335)
Trade and catering services			1.485**			–0.212***			1.697**
			(0.743)			(0.057)			(0.791)
Professional service			4.222***			–0.248***			4.470***
			(1.287)			(0.094)			(1.368)
Constant	–10.821***	–7.703***	–7.467***	0.570***	0.334***	0.327***	–11.391***	–8.037***	–7.794***
	(1.229)	(0.943)	(0.997)	(0.095)	(0.080)	(0.086)	(1.315)	(1.014)	(1.068)
Amenities	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	253	253	253	253	253	253	253	253	253
Adjusted R^2	0.091	0.547	0.650	0.219	0.531	0.652	0.102	0.551	0.655
First-stage F Statistics	63.359	56.172	31.381	63.359	56.172	31.381	63.359	56.172	31.381
Hansen J Statistics	2.299	2.190	3.058	2.055	1.613	3.726	2.299	2.180	3.174

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

Note: The instrumental variables are housing prices in 2000 and the share of land slope above 30 degrees within a 35km radius from a city centroid. Hansen J statistics is used for testing over-identifying restrictions. The four first-tier cities contain Beijing, Shanghai, Guangzhou and Shenzhen.

4.9 Robustness check

The existence of some megacities including Beijing, Shanghai, Guangzhou and Shenzhen might bias the regression results, because these cities usually faced more tight *hukou* regulations. Following Fang et al. (2015), I run the regressions again after dropping them from the sample. Table 4.5 reports similar patterns with the previous tables, suggesting that the regression results in Section 4.8.2 are robust and stable. The magnitudes of coefficients of house price on the SIR and the number of high-skilled workers in Table 4.5 become larger than those in Table 4.3, which implies that the effects of housing prices on the SIR and the number of high-skilled workers are smaller in four mega cities.

4.10 Conclusion and further discussions

The main goal of this chapter is to apply for the extensions of the Rosen-Roback model to examine the relationship between housing costs and the SIR. The empirical results show that the SIR will increase 0.327% when average housing price increases one percentage in 2010, when *hukou* regulation was relaxed. Specifically, the share of high-skilled workers in population is positively related to average housing price, while the share of low-skilled workers in population is negatively related to average housing price. These results confirms validity of the Rosen-Roback model.

In this chapter, I empirically identify the impact of housing price to share of high-skilled workers, that of low-skilled workers, and the SIR using land slope variables and historical housing prices as instruments drawing on insights about identifying local land supply shocks.

The results can be explained with the spatial equilibrium model. Assuming that the income elasticity of housing demand is less than one, and that housing preferences are non-homothetic, rising housing costs cause the share of housing costs on income increases faster for low-skilled workers than high-skilled. Therefore, firms have to pay relatively more for low-skilled workers in order to compensate housing costs of low-skilled workers. Therefore, the SIR rises with rising housing costs within cities. The findings imply that the assumption that housing preferences are homothetic is problematic. When one analyzes consequences of rising housing costs, it is necessary to consider the heterogeneity of different skilled types of workers.

One important caveat to the chapter is that I cannot examine the relationship between housing costs and wage differentials between high-skilled workers and low-skilled due to limitation of data. In the future, I hope to test the validity of Rosen-Roback model directly if I have suitable dataset.

4A Appendices

Figure A1: Example of a typical prefecture-level city

Note: Baoding consists of 5 urban districts (Jingxiu, Lianchi, Mancheng, Qingyuan, and Xushui), 4 county-level cities (Dingzhou, Zhuozhou, Anguo and Gaobeidian) and 15 counties (Yi, Laiyuan, Dingxing, Shunping, Tang, Wangdu, Laishui, Gaoyang, Anxin, Xiong, Rongcheng, Quyang, Fuping, Boye and Li). This map is sourced from WikiPedia.

Table A1: Distribution of the number of observations at the prefecture-level city in the micro censuses data

	Mean	SD	Min	Q1	Median	Q3	P90	Max
2000								
College-educated	42.74	98.81	0	7	15	31	86	1062
Non college-educated	717.73	952.00	86	264	414	723	1441	8955
2010								
College-educated	135.73	311.20	3	24	48	101	296	3291
Non college-educated	992.42	1410.84	111	327	513	956	2169	12139

Note: There are 257 cities in the sample each year. I restrict the sample to individuals living in urban districts of these cities. All variables refer to the number of them in each city. The number of college educated and non college educated are calculated based on adults aged 20 years old or above.

Table A2: Counties or county-level cities which upgraded to urban districts between 2000 and 2010

Province	Prefecture-level City	Previous County Code	Previous County Name	Current County Code	Current County Name	Date of Boundary Change
Beijing		110224	Daxing County	110115	Daxing District	Jan 09, 2001
Beijing		110226	Pinggu County	110117	Pinggu District	Jan 09, 2001
Beijing		110227	Huairou County	110116	Huairou District	Jan 09, 2001
Tianjin		120224	Baodi County	120115	Baodi District	Mar 22, 2001
Hebei	Tangshan	130282	Fengnan City	130207	Fengnan District	Feb 01, 2002
Jilin	Baishan	220625	Jiangyuan County	220605	Jiangyuan District	Jun 05, 2006
Heilongjiang	Harbin	230121	Hulan County	230111	Hulan District	Feb 04, 2004
Heilongjiang	Harbin	230181	Acheng City	230112	Acheng District	Aug 15, 2006
Shanghai		310226	Fengxian County	310120	Fengxian District	Jan 09, 2001
Jiangsu	Xuzhou	320323	Tongshan County	320312	Tongshan District	Sep 05, 2010
Jiangsu	Changzhou	320483	Wujin City	320412	Wujin District	Apr 03, 2002
Jiangsu	Nantong	320683	Tongzhou City	320612	Tongzhou District	Mar 23, 2009
Jiangsu	Yancheng	320928	Yandong County	320903	Yandong District	Dec 18, 2003
Jiangsu	Zhenjiang	321121	Dantu County	321112	Dantu District	Apr 03, 2002
Jiangsu	Suqian	321321	Suyu County	321311	Suyu District	Jan 15, 2004
Zhejiang	Hangzhou	330181	Xiaoshan City	330109	Xiaoshan District	Feb 02, 2001
Zhejiang	Hangzhou	330184	Yuhang City	330110	Yuhang District	Feb 02, 2001
Zhejiang	Ningbo	330227	Yin County	330212	Yinzhou District	Feb 01, 2002
Zhejiang	Quzhou	330821	Yin County	330803	Qujiang District	Dec 10, 2001
Fujian	Putian	350321	Putian County	350304	Licheng District	Feb 01, 2002
Fujian	Putian	350321	Putian County	350305	Xiuyu District	Feb 01, 2002
Shandong	Jinan	370123	Changqing County	370113	Changqing District	Jun 06, 2001
Henan	Luohe	411123	Yancheng County	411104	Zhaoling District	Sep 07, 2004
Henan	Luohe	411123	Yancheng County	411103	Yancheng District	Sep 07, 2004
Hubei	Yichang	420521	Yichang County	420506	Yiling District	Mar 22, 2001
Hubei	Xiangfan	420621	Xiangyang County	420607	Xiangyang District	Aug 31, 2001
Hubei	Suizhou	421303	Suidu District	421321	Sui County	Mar 23, 2009
Guangdong	Shaoguan	440221	Qujiang County	440205	Qujiang District	May 29, 2004
Guangdong	Zhuhai	440421	Dongmen County	440403	Dongmen District	Apr 04, 2001
Guangdong	Shantou	440582	Chaoyang City	440514	Chaonan District	Jan 29, 2003
Guangdong	Shantou	440582	Chaoyang City	440513	Chaoyang District	Jan 29, 2003
Guangdong	Shantou	440583	Chenghai City	440515	Chenghai District	Jan 29, 2003
Guangdong	Foshan	440681	Shunde City	440606	Shunde District	Dec 08, 2002
Guangdong	Foshan	440682	Nanhai City	440605	Nanhai District	Dec 08, 2002
Guangdong	Foshan	440683	Sanshui City	440607	Sanshui District	Dec 08, 2002
Guangdong	Foshan	440684	Gaoming City	440608	Gaoming District	Dec 08, 2002
Guangdong	Jiangmen	440782	Xinhui City	440705	Xinhui District	Jun 22, 2002
Guangdong	Maoming	440923	Dianbai County	440903	Maogang District	Jan 22, 2001
Guangdong	Huizhou	441381	Huiyang City	441303	Huiyang District	Mar 06, 2003
Guangxi	Nanning	450121	Yongning County	450109	Yongning District	Sep 15, 2004
Guangxi	Nanning	450121	Yongning County	450108	Liangqing District	Sep 15, 2004
Chongqing		500221	Changshou County	500115	Changshou District	Dec 25, 2001
Sichuan	Chengdong	510123	Wenjiang County	510115	Wenjiang District	Apr 14, 2002
Sichuan	Chengdong	510125	Xindong County	510114	Xindong District	Nov 15, 2001
Shanxi	Xian	610121	Changan County	610116	Changan District	Jun 02, 2002
Shanxi	Tongchuan	610221	Yao County	610204	Yaozhou District	Jun 18, 2002
Shanxi	Baoji	610321	Baoji County	610304	Chencang District	Mar 01, 2003
Ningxia	Wuzhong	640321	Zhongwei County	640502	Shapotou District	Dec 31, 2003

Table A3: Source of data used in this chapter

Variables	Definition	Geographical scale	Source	Boundary adjustment
SIR	The ratio of college-educated individuals (four-year college) to non college-educated individuals among people aged 15 years or older	Urban districts	County-level tabulations of census	Yes
Population		Urban districts	County-level tabulations of census	Yes
Industry mix	The ratio of employment in each industry to total employment in all industries	Urban districts	County-level tabulations of census	Yes
Average real estate price	Total revenue of real estate divided by sale quantity of real estate	Prefecture	Statistical Yearbook for Regional Economy	No
Amenities		Urban districts	China City Statistical Yearbook	No

Note: Industry mix only includes employment shares in manufacturing, trade & catering services, and professional services. In particular, professional services contains finance, research & environmental protection, and transportation & communication. Amenities include green space ratio, teacher-student ratio for primary schools, the number of books in public libraries, the number of doctors, and the number of college students in schools.

Table A4: Industry classification over the period 2000-2010

No.	Industry classification	2000	2010
1	Agriculture	Farming, Forestry, Animal Husbandry and Fisheries	Agriculture, Forestry, Animal Husbandry and Fisheries
2	Mining	Mining and quarrying	Mining
3	Manufacturing	Manufacturing	Manufacturing
4	Utilities	Production and supply of electric power, gas and tap water	Production and Supply of Electricity, Gas and Water
5	Construction	Construction	Construction
6	Transportation and Communication	Transport, storage, post and telecommunication services	Storage and postal services
7	Trade and catering services	Wholesale and retail trades & catering services	Information Transmission, Computer Service and Software Information transfer and wholesale and retail trades Accommodations and catering services
8	Finance	Finance and insurance	Finance
9	Real estate	Real estate	Real estate
10	Social services	Social services	Leasing and business services Resident and other services
11	Research and environmental protection	Scientific Research and Polytechnic Services	Scientific Research, Technical Services, and Geological Prospecting
12	Social welfare and cultural activities	Geological Prospecting and Water Conservancy	Management of water conservancy, environment and public facilities
		Health Care, Sports & Social Welfare	Health, Social Securities and Social Welfare
		Education, Culture and Arts, Radio, Film and Television	Education
			Culture, Sports and Entertainment
13	Public agency	Government agencies, Party Agencies and Social Organizations	Public Management and Social Organization
		Others	International Organization

Chapter 5

General conclusion

This thesis contains three essays on housing markets and housing policies. In chapter 2, I evaluate a housing policy. Specifically, I examine the effects of rent control on rents using historical data from Lyon during the period 1890-1968. Then I turn to explore the effects of one of deterministic factors on housing markets in the long run. In chapter 3, I estimate how housing demand, one side of the housing market, changes with age using micro-level data from China. And I am also concerned about the interaction between housing markets and labor markets. In chapter 4, I study how housing costs affect the ratio of high-skilled to low-skilled workers, the skill intensity ratio (SIR), across Chinese cities using the Rosen-Roback model.

Most of the existing literature evaluates the effects of rent control on housing markets in the U.S., in particular in New York City, Boston and San Francisco. However, few studies have estimated the effects of rent control on housing markets in Europe. Borrowing a unique dataset collected from a property manager's accounting books by Bonneval and Robert (2009), in Chapter 2, I estimate effects of two forms of rent control (one tight regulation and another moderate regulation) on rents in Lyon from 1890 to 1968. The results show that both forms of rent control can make rents decrease in most cases, where the tight regulation causes rents to decrease more. This chapter complements the literature on rent control policy evaluations and add new evidence from France.

In Chapter 3, I estimate the effect of the age of household head on housing demand, quantity and quality of consumed housing services, using micro-level data from urban China during the period 2007-2009. The econometric analysis reveals that the housing demand decreases when household head is above 60 years old if socio-demographic characteristics of households except age are not controlled for. The result is in line with the findings in Mankiw and Weil (1989). However, if I control for socio-demographic characteristics of households including income, educational attainment

and marital status in the analysis, the housing demand does not decrease after people are retired. Furthermore, if I control for socio-demographic variables which are constant over the life-cycle of households such as gender and educational attainment, and do not control for variables which vary with age such as income and marital status, the housing demand still does not decrease much after people are older than 60 years. The results reflect that educational attainment strongly affects housing demand. The cross-sectional data shows that if housing demand decreases with age, it is because the educational attainment of current elderly is lower than that of current young people, rather than because their preference for housing demand is lower. When the current young people become old, the aggregate housing demand will not drop because they have higher educational attainment and incomes than the previous generations. In this chapter, I emphasize the importance of the effects of educational attainment on housing demand. In the context of Chinas higher education expansion, the chapter predicts that aggregate housing demand will not decline with population aging in urban China.

Chapter 4 investigates the impacts of housing costs on the SIR across cities in China. This chapter uses China's census data in 2000 and 2010 to show that housing costs are positively related to the SIR. Specifically, the average housing prices have significant positive effects on the SIR in 2010. When housing prices increase by one percentage point, the SIR increases by 0.327%. The result is robust to the use of share of unavailable land and historical housing prices as instruments of housing prices. However, the SIR is less affected by average housing prices in 2000, when workers' mobility was tightly regulated. This chapter suggests that rising housing costs change the labor demand composition across Chinese cities when workers' mobility is relaxed.

Of course, these chapters have some limitations, which indicate directions for future research. In Chapter 2, the data have some limitations. First, all flats were controlled at least once during the whole period. Therefore, I cannot use the difference-in-difference strategy to estimate the effect of rent control on rents. In addition, this limitation does not allow me to observe spillover effects of rent control on the uncontrolled rental markets. Second, maintenance costs are collected at the building level rather than at the flat level. If I want to investigate how rent control affects maintenance costs of flats, I need more assumptions to disaggregate the maintenance costs from the building level to the flat level. Third, characteristics of households are not available in the sample. Thus I cannot analyze the effects of rent control on tenants and landlords' behaviors, not to mention welfare effects. In further research, I would like to incorporate other datasets to control for macroeconomic variables such as annual interest rate and average income at the district level in Lyon. I also hope to find an instrumental variable related to rent control status to deal with the endogeneity issue.

Chapter 3 also has an identification problem. A household's unobservable preference simultaneously affect both prices and quantities of housing services so that it would be better to find instrumental variables to avoid endogeneity issues, as suggested by Bartik (1987). In the literature, geographical classifications are commonly used as instruments (Bartik (1987); Green and Hendershott (1996); Eichholtz and Lindenthal (2014)). However, the suggested instruments are found to be weak in many papers. I hope to find a better instrument to address this issue in the future. In addition, I only get the relationship between housing demand and age at the household level at the moment. In the future, I would like to compute aggregate age-specific housing demand in the next several years if predicted demographic structure is known. This simulation could give a reference for the government and households.

In Chapter 4, I examine the indirect effect of housing costs on the labor market using data at the city level in China. The Rosen-Roback model predicts that the wage gap between high-skilled workers and low-skilled workers is negatively related to housing costs. Then firms have the motivation to employ a relatively large share of high-skilled workers in response to the changes in the wage ratio of high-skilled to low-skilled workers. The census data do not record skill-specific wages, so I cannot test the Rosen-Roback model in a direct way. In the future, I hope to obtain average wage for high- and low-skilled workers at the city level from other surveys to test the Rosen-Roback model directly. In addition, I would like to examine how housing costs affect the SIR for migrants as migrants are more mobile than local people.

References

Albon, R. P. and D. C. Stafford (1990). Rent control and housing maintenance. *Urban Studies 27*(2), 233–240.

Arnott, R. and E. Shevyakhova (2014). Tenancy rent control and credible commitment in maintenance. *Regional Science and Urban Economics 47*, 72–85.

Au, C.-C. and J. V. Henderson (2006). Are chinese cities too small? *The Review of Economic Studies 73*(3), 549–576.

Ault, R. and R. Saba (1990). The economic effects of long-term rent control: The case of new york city. *The Journal of Real Estate Finance and Economics 3*(1), 25–41.

Autor, D. H., C. J. Palmer, and P. A. Pathak (2014). Housing market spillovers: Evidence from the end of rent control in cambridge, massachusetts. *Journal of Political Economy 122*(3), 661–717.

Bacolod, M., B. S. Blum, and W. C. Strange (2009). Skills in the city. *Journal of Urban Economics 65*(2), 136–153.

Bartik, T. J. (1987). The estimation of demand parameters in hedonic price models. *Journal of political Economy 95*(1), 81–88.

Bartik, T. J. (1991). Who benefits from state and local economic development policies?

Basu, K. and P. M. Emerson (2003). Efficiency pricing, tenancy rent control and monopolistic landlords. *Economica 70*(278), 223–232.

Beeson, P. E. (1991). Amenities and regional differences in returns to worker characteristics. *Journal of Urban Economics 30*(2), 224–241.

Black, D., N. Kolesnikova, and L. Taylor (2009). Earnings functions when wages and prices vary by location. *Journal of Labor Economics 27*(1), 21–47.

Bonneval, L. and F. Robert (2009, December). L'immobilier de rapport : forme urbaine, rapports de locations et stratégies patrimoniales. Le cas de Lyon (1860-1990). Technical report.

Brinkman, J. (2014). The supply and demand of skilled workers in cities and the role of industry composition.

Broxterman, D. A. and A. M. Yezer (2015). Why does skill intensity vary across cities? the role of housing cost. *Regional Science and Urban Economics 55*, 14–27.

Caudill, S. B. (1993). Estimating the costs of partial-coverage rent controls: A stochastic frontier approach. *The Review of Economics and Statistics*, 727–731.

Chan, K. W. (2007). Misconceptions and complexities in the study of china's cities: Definitions, statistics, and implications. *Eurasian Geography and Economics 48*(4), 383–412.

Chan, K. W. and L. Zhang (1999). The hukou system and rural-urban migration in china: Processes and changes. *The China Quarterly 160*, 818855.

Chaney, T., D. Sraer, and D. Thesmar (2012). The collateral channel: How real estate shocks affect corporate investment. *The American Economic Review 102*(6), 2381–2409.

Chen, B. and C. Zhang (2016). Human capital and housing prices in chinese cities (in chinese). *Social Sciences in China (Zhongguo Shehui Kexue) 5*, 44–65.

Chen, J. and M. Jin (2014). Income elasticity of housing demand in china: micro-data evidence from shanghai. *Journal of Contemporary China 23*(85), 68–84.

Chow, G. C. and L. Niu (2015). Housing prices in urban china as determined by demand and supply. *Pacific Economic Review 20*(1), 1–16.

Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980–2000. *The American Economic Review 106*(3), 479–524.

Diamond, R., T. McQuade, and F. Qian (2018). The effects of rent control expansion on tenants, landlords, and inequality: Evidence from san francisco. Technical report, National Bureau of Economic Research.

Diewert, W. E. (2003). Hedonic regressions. a consumer theory approach. In *Scanner data and price indexes*, pp. 317–348. University of Chicago Press.

Dong, Y. (2016). A note on geographical constraints and housing markets in china. *Journal of Housing Economics 33*, 15–21.

Early, D. and J. Phelps (1999). Rent regulations' pricing effect in the uncontrolled sector: An empirical investigation. *Journal of Housing Research 10*(2), 267–285.

Early, D. W. (2000). Rent control, rental housing supply, and the distribution of tenant benefits. *Journal of Urban Economics 48*(2), 185–204.

Eichholtz, P. and T. Lindenthal (2014). Demographics, human capital, and the demand for housing. *Journal of Housing Economics 26*, 19–32.

Elvery, J. A. (2010). City size and skill intensity. *Regional Science and Urban Economics 40*(6), 367–379.

Engelhardt, G. V. and J. M. Poterba (1991). House prices and demographic change: Canadian evidence. *Regional Science and Urban Economics 21*(4), 539–546.

Ermisch, J. (1996). The demand for housing in britain and population ageing: microeconometric evidence. *Economica*, 383–404.

Fallis, G. and L. B. Smith (1984). Uncontrolled prices in a controlled market: the case of rent controls. *The American Economic Review 74*(1), 193–200.

Fang, H., Q. Gu, W. Xiong, and L.-A. Zhou (2015). Demystifying the chinese housing boom. In *NBER Macroeconomics Annual 2015, Volume 30*. University of Chicago Press.

Ferreira, F. and J. Gyourko (2011). Anatomy of the beginning of the housing boom: Us neighborhoods and metropolitan areas, 1993-2009. Technical report, National Bureau of Economic Research.

Friggit, J. (2002). L'évolution sur longue période du prix des logements. *note de synthèse du SES (Service des études et de la statistique), Ministére de l'Equipement*.

Ge, S. and D. T. Yang (2014). Changes in china's wage structure. *Journal of the European Economic Association 12*(2), 300–336.

Glaeser, E., W. Huang, Y. Ma, and A. Shleifer (2017). A real estate boom with chinese characteristics. *Journal of Economic Perspectives 31*(1), 93–116.

Glaeser, E. L. (2002). Does rent control reduce segregation?

Glaeser, E. L., J. Gyourko, and R. E. Saks (2005). Why have housing prices gone up? *American Economic Review 95*(2), 329–333.

Green, R. and P. H. Hendershott (1996). Age, housing demand, and real house prices. *Regional Science and Urban Economics 26*(5), 465–480.

Green, R. K. and H. Lee (2016). Age, demographics, and the demand for housing, revisited. *Regional Science and Urban Economics 61*, 86–98.

Gyourko, J. (2009). Rethinking federal housing policy: How to make housing plentiful and affordable.

Gyourko, J. and P. Linneman (1989). Equity and efficiency aspects of rent control: An empirical study of new york city. *Journal of urban Economics 26*(1), 54–74.

Gyourko, J., C. Mayer, and T. Sinai (2013). Superstar cities. *American Economic Journal: Economic Policy 5*(4), 167–99.

Heffley, D. (1998). Landlords, tenants and the public sector in a spatial equilibrium model of rent control. *Regional Science and Urban Economics 28*(6), 745–772.

Hendershott, P. H. (1991). Are real house prices likely to decline by 47 percent? *Regional Science and Urban Economics 21*(4), 553–563.

Hendricks, L. (2011). The skill composition of us cities. *International Economic Review 52*(1), 1–32.

Ho, L. S. (1992). Rent control: its rationale and effects. *Urban Studies 29*(7), 1183–1189.

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

Hubert, F. (1993). The impact of rent control on rents in the free sector. *Urban Studies 30*(1), 51–61.

Kim, D., F. Liu, and A. Yezer (2009). Do inter-city differences in intra-city wage differentials have any interesting implications? *Journal of Urban Economics 66*(3), 203–209.

Kutty, N. K. (1996). The impact of rent control on housing maintenance: A dynamic analysis incorporating european and north american rent regulations. *Housing Studies 11*(1), 69–88.

Lan, L. (2018). Understanding the saving behavior of chinese households: Intergenerational transfer and housing. *Job Market Paper*.

Larson, W. D. (2011). Housing and labor market dynamics in growing versus declining cities. *Available at SSRN 1702488*.

Levin, E., A. Montagnoli, and R. E. Wright (2009). Demographic change and the housing market: evidence from a comparison of scotland and england. *Urban Studies 46*(1), 27–43.

Li, M. and K. Shen (2013). Population aging and housing consumption: A nonlinear relationship in china. *China & World Economy 21*(5), 60–77.

Liang, W., M. Lu, and H. Zhang (2016). Housing prices raise wages: Estimating the unexpected effects of land supply regulation in china. *Journal of Housing Economics 33*, 70–81.

Lindh, T. and B. Malmberg (2008). Demography and housing demand-what can we learn from residential construction data? *Journal of Population Economics 21*(3), 521–539.

Mankiw, N. G. and D. N. Weil (1989). The baby boom, the baby bust, and the housing market. *Regional science and urban economics 19*(2), 235–258.

Mengle, D. (1985). The effect of second generation rent controls on the quality of rental housing. *Federal Reserve Bank of Richmond Working Paper* (85-5).

Mense, A., C. Michelsen, and K. A. Cholodilin (2017). Empirics on the causal effects of rent control in germany. Technical report, FAU Discussion Papers in Economics.

Mian, A. and A. Sufi (2011). House prices, home equity–based borrowing, and the us household leverage crisis. *The American Economic Review 101*(5), 2132–2156.

Moon, C.-G. and J. G. Stotsky (1993). The effect of rent control on housing quality change: a longitudinal analysis. *Journal of Political Economy*, 1114–1148.

Moretti, E. (2004). Human capital externalities in cities. *Handbook of regional and urban economics 4*, 2243–2291.

Moretti, E. (2012). *The new geography of jobs*. Houghton Mifflin Harcourt.

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

Murray, M. P. and G. Sun (2017). The demand for space in china. *Journal of Urban Economics 98*, 214–222.

Nagy, J. (1997). Do vacancy decontrol provisions undo rent control? *Journal of Urban Economics 42*(1), 64–78.

Navarro, P. (1985). Rent control in cambridge, mass. *The Public Interest* (78), 83.

NBS (2013). China statistical yearbook 2010.

Neuteboom, P. and D. Brounen (2007). Demography and housing demand–dutch cohort evidence. *Erasmus University Working Paper*.

Ohtake, F. and M. Shintani (1996). The effect of demographics on the japanese housing market. *Regional Science and Urban Economics 26*(2), 189–201.

Olsen, E. O. (1988). What do economists know about the effect of rent control on housing maintenance? *The journal of real estate finance and economics 1*(3), 295–307.

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

Pennington-Cross, A. (1997). Measuring external shocks to the city economy: An index of export prices and terms of trade. *Real Estate Economics 25*(1), 105–128.

Raess, P. and T. von Ungern-Sternberg (2002). A model of regulation in the rental housing market. *Regional Science and Urban Economics 32*(4), 475–500.

Roback, J. (1982). Wages, rents, and the quality of life. *The Journal of Political Economy*, 1257–1278.

Roberts, M., U. Deichmann, B. Fingleton, and T. Shi (2012). Evaluating china's road to prosperity: A new economic geography approach. *Regional Science and Urban Economics 42*(4), 580–594.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy 82*(1), 34–55.

Rosen, S. (1979). Wage-based indexes of urban quality of life. In P. Mieszkowski and M. Straszheim (Eds.), *Current Issues in Urban Economics*, pp. 74–104. Baltimore: John Hopkins University Press.

Saiz, A. (2010). The geographic determinants of housing supply. *quarterly Journal of Economics 125*(3).

Shapiro, J. M. (2006). Smart cities: quality of life, productivity, and the growth effects of human capital. *The review of economics and statistics 88*(2), 324–335.

Shi, J., Y. Wang, W. Wu, et al. (2016). The crowding-out effect of real estate markets on corporate innovation: Evidence from china.

Sims, D. P. (2007). Out of control: What can we learn from the end of massachusetts rent control? *Journal of Urban Economics 61*(1), 129–151.

Stock, J. and M. Yogo (2005). *Testing for Weak Instruments in Linear IV Regression*, pp. 80–108. New York: Cambridge University Press.

Swan, C. (1995). Demography and the demand for housing a reinterpretation of the mankiw-weil demand variable. *Regional Science and Urban Economics 25*(1), 41–58.

Takáts, E. (2012). Aging and house prices. *Journal of Housing Economics 21*(2), 131–141.

Taylor, J. B. (2007). Housing and monetary policy. Technical report, National Bureau of Economic Research.

Turner, B. and S. Malpezzi (2003). A review of empirical evidence on the costs and benefits of rent control. *Swedish Economic Policy Review* (10).

Wang, S., S. H. Chan, and B. Xu (2012). The estimation and determinants of the price elasticity of housing supply: Evidence from china. *Journal of Real Estate Research 34*(3), 311–344.

Wei, S.-J., X. Zhang, and Y. Liu (2017). Home ownership as status competition: Some theory and evidence. *Journal of Development Economics 127*, 169–186.

Wheeler, C. H. (2004). On the distributional aspects of urban growth. *Journal of Urban Economics 55*(2), 371–397.

Wu, J., J. Gyourko, and Y. Deng (2012). Evaluating conditions in major chinese housing markets. *Regional Science and Urban Economics 42*(3), 531–543.

Yang, Z. and J. Chen (2014). *Housing affordability and housing policy in urban China*. Springer.

Yearbook, C. S. et al. (2016). National bureau of statistics of china. *China Statistical Yearbook*.

Zhang, J., J. Fan, and J. Mo (2017). Government intervention, land market, and urban development: Evidence from chinese cities. *Economic Inquiry 55*(1), 115–136.

Zheng, S. and M. E. Kahn (2013). China's bullet trains facilitate market integration and mitigate the cost of megacity growth. *Proceedings of the National Academy of Sciences 110*(14), E1248–E1253.

Zheng, S., M. E. Kahn, and H. Liu (2010). Towards a system of open cities in china: Home prices, fdi flows and air quality in 35 major cities. *Regional Science and Urban Economics 40*(1), 1–10.

Zheng, S. and H. Liu (2005). Income elasticity of housing demand in china: model, estimation and forecast. *China Civil Engineering Journal 7*, 122–126 (in Chinese).