Université de Montréal

Essays in Empirical Labor, Housing and Social Network Economics

par
Ngoundjou Nkwinkeum George Aurore Dupin

Département de sciences économiques
Faculté des arts et sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)
en sciences économiques

Octobre, 2020

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée :
Essays in Empirical Labor, Housing and Social Network Economics

présentée par :
Ngoundjou Nkwinkeum George Aurore Dupin

a été évaluée par un jury composé des personnes suivantes :

Raphaël Godefroy,	Directeur de recherche
Mathieu Marcoux,	Membre du jury
Nicholas Lawson,	Examinateur externe
Joshua Lewis,	Président-rapporteur
Michael Huberman,	Représentant du doyen de la FAS

Thèse acceptée le 30/10/ 2020

À mes parents et ma tendre épouse : Ngoundjou Zacharie, Donkeu Naomie Josephin, Betchem Andjike, Yabo Christine et Betchem A Becthem Christine Tabitha

Remerciements

Cette thèse de doctorat est l'aboutissement d'un long voyage jalonné de défis divers tant sur le plan personnel que professionnel. L'expérience a toutefois été très gratifiante, non seulement en raison des connaissances accumulées et des compétences acquises, mais aussi en raison de l'occasion unique que j'ai eu de cotoyer des chercheurs de divers horizons et de collègues très motivés. Tout ce que j'ai vécu au cours de toutes ces années restera à jamais gravé dans ma mémoire.

J'aimerais commencer par remercier tout particulièrement mon directeur de recherche, Raphaël Godefroy. Sans sa grande patience, son soutient constant et ses nombreux encouragements, ce travail n'aurais jamais pu être accompli. J'ai beaucoup appris de lui tant sur le plan humain que professionnel.

Je tiens aussi à remercier Amine Ouazad pour ses commentaires forts utiles qui m'ont permis d'affiner mon travail de recherche. Mes remerciements vont aussi à Nicholas Klein pour le soutien multiforme et les échanges très contsructifs que nous avons eu. J'ai aussi une profonde gratitude envers mes promotionnaires, Hervé Zeida, Marius Pondi, Louphou Coulibaly, et Michel Youmbi, avec qui j'ai passé des moments mémorables. Je suis aussi reconnaissant à tous les chercheurs et enseignants du département pour leurs apports divers tout au long de mon doctorat.

Je suis très reconnaissant à la Société Canadienne d'Hypothèques et de Logement (SCHL) qui m'a accordé l'autorisation d'utiliser un travail de recherche effectué dans le cadre de mon rôle aux fins de compléter les exigences de ma thèse de doctorat. Sous ce rapport, je souhaite exprimer une reconnaissance toute particulière à Josée Paiement et Lois McGrath. Je tiens aussi à remercier mes collègues de la SCHL et co-auteurs du troisième chapitre de cette thèse , à savoir : Judy Zhong et Musah Khalid. Mes remerciements vont aussi à mes collègues qui ont pris de leur temps pour commenter la recherche ainsi effectuée. Il s'agit de : Abra Adamo, Antoine Noubouwo et Mouinica Guivada.

Ma reconnaissance va aussi envers mes proches et amis du Canada, qui ont contribué de diverses manières à la réussite de mon parcours. Parmi eux, je tiens à remercier Laurent Kemoe qui m'a acceuilli et hébergé à mon arrivée au Canada. J'ai aussi une pronfonde et indescriptible gratitude

pour Joel Betchem et Elise Loube Koum qui ont toujours été présents pour moi et mon épouse. Qu'ils trouvent dans ce document l'expression de ma profonde gratitude.

Je tiens aussi exprimer mes remerciements envers le département de sciences économiques de l'Université de Montréal, ainsi que le Centre interuniversitaire de recherche en économie quantitative (CIREQ). Le département et le CIREQ m'ont assuré à la fois le financement de mes études et de ma recherche, mais aussi de bonnes conditions de travail et un enseignement de haut niveau. J'aimerais aussi remercier la Chaire de la fondation J. W. McConnell pour m'avoir accordé une bourse d'étude qui m'a permis d'approfondir la rédaction du premier chapitre de la présente thèse. Enfin, je remercie profondément la faculté des études supérieures de l'Université de Montréal, laquelle m'a accordé une bourse d'excellence qui m'a permis de poursuivre ce long périple.

Résumé

Cette thèse explore les deux sujets suivants: (i) les déterminants de la participation à la population active des personnes âgées aux États-Unis; (ii) le rôle du réseau social dans le niveau de revenus de travail des immigrants récents au Canada.

Plus précisément, le chapitre 1 utilise les données du Health and Retirement Study (HRS) pour fournir des estimations de l'effet causal du prix des maisons au niveau local sur les décisions d'offre de travail des individus âgés aux États-Unis, au cours de la période du boom immobilier 1994-2004. Le prix des maisons est instrumentalisé par des variations spatiales et temporelles de chocs plausiblement exogènes de l'offre de crédit. Les estimations suggèrent que le boom immobilier durant la période considérée pourrait expliquer jusqu'à deux tiers des sorties de la population active des hommes âgés aux États-Unis.

Le chapitre 2 répond à la question: les personnes en bonne santé sont-elles plus susceptibles de rester sur le marché du travail lorsque le chômage augmente? Ce travail fournit des estimations empiriques de l'impact relatif des chocs de chômage au niveau local sur la participation au marché du travail des américains âgés ayant des états de santé hétérogènes. Cette étude montre que la plupart des départs de travailleurs âgés pendant la Grande récession peuvent être attribués à de mauvaises conditions de santé, telles que mesurées, avant la récession.

Enfin, le chapitre 3 analyse l'effet du fait de vivre dans un quartier où la majorité des résidents appartient à des minorités visibles sur les gains des résidents permanents qui sont arrivés à Montréal, Vancouver et Toronto en 2001. Dans l'ensemble, cette recherche documente un effet négatif du regroupement ethnique sur les gains des immigrants récents au Canada et l'explique par "l'effet d'aiguillage".

Mots-clés: Retraite, offre de travail, prix des logements, sécurité sociale, santé, enclave ethnique, réseau social, immigration.

Abstract

This thesis explores the following two topics : (i) determinants of labor force participation at older ages in the US ; (ii) the role of social network in the earnings of recent immigrants in Canada.

Specifically, Chapter 1 provides estimates of the causal effect of local house prices on the labor supply decisions of older workers in the US during the 1994-2004 housing boom period using data from the Health and Retirement Study (HRS). Local house prices are instrumented by plausible exogenous spatial and time-varying credit supply shocks. The estimates suggest that the housing boom accounted for about two thirds of older men labor force exit during the studied period.

Chapter 2 responds to the question : are people in good health more likely to stay in the labor market when unemployment rises? This work provides empirical estimates of the relative impact of local unemployment shocks on the labor force participation of US older workers with heterogeneous health conditions. This study documents that most of the labor force exit of older workers during the Great recession can be attributed to poor health conditions measured before the recession.

Finally, Chapter 3 analyses the effect of living in a neighborhood where the majority of the residents belongs to visible minority groups on the earnings of permanent residents who landed in Montreal, Vancouver and Toronto in 2001. Overall, this research documents a negative effect of ethnic clustering on the earnings of recent immigrants in Canada and explain it by the "referral effect".

Keywords : Retirement, Labor Supply, House Prices, Social Security, Health, Ethnic Enclave, Social network, Immigration.

Contents

Dédicace iii

Remerciements iv

Résumé vi

Abstract vii

Contents viii

Table des matières x

List of Figures xi

Liste des figures xi

List of Tables xii

Liste des tableaux xiii

Introduction 1

1 House Prices and Retirement in US 3

1.1 Introduction 3

1.2 Data and empirical strategy 6

1.2.1 Data 6

1.2.2 Empirical specification 8

1.2.3 Identification strategy: An instrumental variable approach 10

1.3 Results 13

1.3.1 Main result . . . 13
1.4 Robustness checks . . . 18
1.4.1 Robustness to labor demand-based explanations . . . 18
1.4.2 Robustness to the estimation of standard errors . . . 20
1.5 Conclusion . . . 22

2 Labor Market Conditions, Health and Retirement 23
2.1 Introduction . . . 23
2.2 Data and Empirical Specification . . . 26
2.2.1 Data . . . 26
2.2.2 Health measures in the HRS . . . 26
2.2.3 Summary statistics . . . 29
2.2.4 Empirical specification . . . 29
2.2.5 How to interpret β? . . . 30
2.3 Results . . . 32
2.3.1 Main results . . . 32
2.3.2 Heterogeneity . . . 36
2.4 Robustness tests . . . 37
2.4.1 Results using employment-to-population ratio . . . 39
2.4.2 Robustness to the time period considered . . . 40
2.5 Conclusion . . . 42

3 The Effect of Ethnic Clustering on Immigrants Earnings and Housing Conditions 43
3.1 Introduction . . . 43
3.2 Data . . . 46
3.2.1 Census data . . . 46
3.2.2 CEED data . . . 48
3.3 Empirical Strategy . . . 53
3.3.1 Estimating the effects of ethnic clustering . . . 54
3.4 Results . . . 55
3.4.1 Establishing the existence of Social Interactions and the Referral Effect . . . 55
3.4.2 Main results . . . 56
3.4.3 Heterogeneity by sex at birth and age . . . 57
3.4.4 Heterogeneity by years of schooling . . . 58
3.5 Robustness tests . . . 59

3.6 Conclusion 61

Conclusion **63**

Bibliography **64**

Bibliography **65**

Annexes **72**

4.1 Appendix of Chapter 1 72

4.1.1 Additional information regarding some variables in the HRS 72

4.1.2 Summary Statistics 72

4.1.3 Illustration identification strategy : The case of perfectly elastic housing supply curves 74

4.1.4 Additional Results 75

4.2 Appendix of Chapter 2 80

4.2.1 Summary Statistics 80

4.2.2 Other Estimation Results 81

4.3 Appendix of Chapter 3 84

4.3.1 Definition of core housing need 84

4.3.2 Earning trends of permanent residents who landed in Canada in 2001 . . . 84

List of Figures

1.1 Illustration of the identification strategy 12

3.1 Incidence of ethnic clustering in Montreal, Vancouver and Toronto CMAs, 2006 . . 47
3.2 Population living in ethnic enclaves, by visible minority status, 2006 48
3.3 Time series of average earning gaps, by the CMA of landing. 49
3.4 Time series of average earning gaps, by newcomers sex at birth 51
3.5 Time series of the average earnings gap, by newcomers age group 52
4.6 Illustration of the identification strategy 74

List of Tables

1.1 OLS and 2SLS Estimations of the Effect of House Price Growth on the Labor Force Participation of Older Individuals . . . 14
1.2 OLS and 2SLS Estimations of the Effect of House Price Growth on the Labor Force Participation of Older Men . . . 17
1.3 Reduced Form Relationship Between the (County Elasticity of Housing Supply) ×(Banking Deregulation Index) and the Labor Force Participation Dummy . . . 19
1.4 Reduced Form Relationship Between the (County Elasticity of Housing Supply) ×(Banking Deregulation Index) and Local Labour Market Conditionsy . . . 19
1.5 OLS and 2SLS Estimations of the Effect of House Prices on the Labor Force Participation of Older Individuals (*with two-way clustering*) . . . 21

2.1 OLS Estimates of the Effect of the Great Recession on Individuals Labor Supply . . . 32
2.2 OLS Estimates of the Effect of the Great Recession on Individuals Labor Supply . . . 35
2.3 OLS Estimates of the Effect of Local Unemployment Shocks on the Labor Force Participation of Older Workers, by Education, Marital status and Gender. . . . 36
2.4 OLS Estimates of the Relative Impact of Local Unemployment Shocks on the Labor Force Participation of Older Workers with Heterogeneous Mental and Physical Health Conditions 38
2.5 OLS Estimates of the Effect of Local Employment Shocks on the Labor Force Participation of Older Workers. . . . 40
2.6 OLS Estimates of the Effect of Local Unemployment Shocks on the Labor Force Participation of Older Workers, by Time Period . . . 41

3.1 OLS Estimates of the referall effect . . . 55
3.2 OLS Estimates of the Effect of ethnic clustering on the earnings of recent immigrants in Canada . . . 56

3.3 OLS Estimates of the Effect of ethnic clustering on the earnings of recent immigrants in Canada; by sex at birth and age group at landing. . . . 58
3.4 OLS Estimates of the Effect of ethnic clustering on the earnings of recent immigrants in Canada; by years of schooling at landing. . . . 59
3.5 OLS Estimates of the correlation between the number of neighbours from the same country of birth and the knowledge of Canada's official languages at year of landing (i.e. in 2001). . 60
4.6 Detailed information about some variables in the Rand-HRS . . . 72
4.7 Baseline Summary Statistics . . . 73
4.8 OLS and 2SLS Estimations of the Effect of House Price Growth on the Labor Force Participation of Older Individuals . . . 75
4.9 OLS and 2SLS Estimations of the Effect of House Price Growth on the Labor Force Participation of Older Married Individuals . . . 76
4.10 OLS and 2SLS Estimations of the Effect of House Price Growth on the Labor Force Participation of Older Individuals . . . 76
4.11 OLS and 2SLS Estimations of the Effect of House Price Growth on the Housing Consumption of Older Men . . . 77
4.12 OLS and 2SLS Estimations of the Effect of House Price Growth on the Equity-Based Borrowing Behavior of Older Men . . . 77
4.13 OLS and 2SLS Estimations of the Effect of House Price on County Unemployment . . . 78
4.14 OLS and 2SLS Estimations of the Effect of House Price on County Unemployment . . . 78
4.15 OLS and 2SLS Estimations of the Effect of House Price on the Labor Force Participation of Older Individuals (*with wild-cluster bootstrap*) . . . 79
4.16 Baseline Summary Statistics . . . 80
4.17 Effect of the Great Recession on Individuals Labor Supply . . . 81
4.18 Effect of the Great Recession on Individuals Labor Supply . . . 82
4.19 Effect of the Great Recession on Individuals Labor Supply . . . 83
4.20 Average earnings by CMA of residence . . . 84
4.21 Average earnings by sex at birth . . . 85
4.22 Average earnings by age group . . . 86

Introduction

As in many OECD countries, the current US retirement policy fosters labor force participation at older ages. A major shift arose in the early1980s when the financial viability of the Social Security became a question of national interest. One of major reform happened in 1983, also called "Alan Greenspan reform", and was characterized by the increase of the full retirement age from 65 years old to 67, on a phasis basis starting in 2000. Since then, understanding the determinants of the labor force participation of older individuals received a lot of attention from scholars and policy-makers. The first two chapters of this thesis contribute to these efforts by analysing empirically, respectively, the relative importance of housing wealth and health in the labor force participation of older workers in US.

More specifically, Chapter 1 shows that housing wealth is much more important in the labor supply decision of older individuals than previously documented. This chapter provides estimates of the causal effect of local house prices on the labor supply decisions of older workers in the US during the 1994-2004 housing boom period using data from the Health and Retirement Study. Local house prices are instrumented by plausible exogenous spatial and time-varying credit supply shocks. I found that a 10% increase in local house prices rises the probability of older men exiting from the labor force by 16%. Additionally, the estimates suggest that the housing boom accounted for about two thirds of older men labor force exit during the study period. This work also documents the important heterogeneities across gender, age and household level of borrowing constraints. Furthermore, I also found that an increase of local house prices has no effect on the housing consumption and home equity-based-borrowing behavior of older individuals.

The Chapter 2 responds to the following question : are people in good health more likely to stay in the labor market when unemployment rises? This chapter estimates the relative impact of local unemployment shocks on the labor force participation of US older workers with heterogeneous health conditions. I use microdata from the Health and Retirement Study (HRS) in combination with local employment statistics from the Bureau of Labor Statistics (BLS) to document that poor health conditions, as measured just before the Great recession, explain most of the labor force

exit of older individuals in the states of USA which experienced higher unemployment hikes. The observed impact is quantitatively important and much larger than what was documented in previous studies. The results are also consistent across several health measures such as self-reported health status, cognitive ability, depression risk and ability to perform daily activities.

New immigrants in Canada have been identified in the first National Housing Strategy (NHS) as one of the key population groups to target because of their higher housing needs compared to the general population. Given the direct link between housing conditions and earnings, having a better understanding of the determinants of immigrants earnings is now an important topic for affordable housing policies. Chapter 3 analyses the effect of living in a neighborhood where the majority of the residents belongs to visible minority groups on the earnings of permanent residents who landed in Montreal, Vancouver and Toronto in 2001. Firstly, this chapter documents that immigrants who landed in ethnic enclaves had, on average, a persistent lower earnings than their counterparts who landed in non-ethnic enclave neighborhoods. Secondly, the work accounted for self-selection into ethnic enclaves and estimated the effect of the number of close neighbors sharing the same country of birth on the earnings of recent immigrants. Overall, the research in this chapter suggest a negative effect of ethnic clustering and explain it by the "referral effect". However, important differences across the three census metropolitan areas (CMAs) studied, sex at birth, age and education are also documented.

Chapter 1

House Prices and Retirement in US

1.1 Introduction

Since the 1983, Alan Greenspan reform of the US Social Security Retirement policy, fostering labor force participation at older ages is cornerstone to US retirement policies. As a consequence, identifying the determinants of retirement has received a lot of attention in the economic literature. More specifically, the key features of the Social Security Retirement policy, as well as health and health insurance have been studied extensively (see for example, Rust and Phelan (1997), Friedberg (2000), French (2005)). In comparison, a relative small number of studies have examined the link between house prices and retirement in the US. More recently, Zhao and Burge (2017) and Faranham and Sevak (2016) tried to quantify the effect of house price growth on the retirement decision of American older workers. These studies relied on very strong identification hypothesis and found effects close to zero because it is hard to find a good exogenous variation of house prices.

This paper empirically estimates the causal effect of local house prices on the retirement decision of older homeowners in the US between 1994-2004 using public and private microdata from the Health and Retirement Study (HRS). The work also assesses the importance of various types of heterogeneities in an individual's response to home value appreciation. I have found substantial heterogeneities in gender and household borrowing constraints. The identification strategy is based on an instrumental variable approach.

In 1994, the US federal administration enacted the Riegel-Neal Interstate Banking and Branching Efficiency Act (IBBEA) which allowed bank holding companies operating in a specific state to create branches in other states. However, individual states were also given the right to put several restrictions on this interstate branching. Favara and Imbs (2015) documented large spatial and over-time variations in the level of state restrictions on the interstate branching between 1994 and

2005. By comparing treated financial institutions with similar yet untreated ones, they established that the deregulation increased the market value of single-family properties through an the credit expansion which was not driven by local labor demand conditions. They also noted that the deregulation accounted for all the house price boom between 1994 and 2002 and between one-third and half of the market variation during 2003 and 2005. In their work, the IBBEA can be seen as the shock of housing demand in deregulated counties. I, therefore, instrumented the two-year variation of county average house price with the interaction of the time-varying index of state-level deregulation constructed by Rice and Strahan (2010) with the county elasticity of housing supply from Saiz (2010). This identification strategy follows the idea that the IBBEA raised the house prices in heavily land-constrained counties (with low housing supply elasticity) while housing stocks increased in less-constrained counties. I also provided falsification tests which reinforced the exclusion restriction validity of the instrumental variable. To be more specific, there is a direct significant effect of the instrumental variable on the labor supply of homeowners but for renters, I did not find any significant effect. Furthermore, the instrumental variable had no direct effect on local unemployment rates. This ruled-out the possibility that my results might be driven by unobserved local labor demand shocks.

This work proves an important housing wealth effect. I found that a 10% increase in local house prices increased the probability of older individuals exiting from the labor force by 10% overall, and roughly by 16% for older men. However, I found no significant effect of local house prices on the labor supply of older women. In addition, household level of borrowing constraints is an important determinants in this observed effect. All observed labor market exits are concentrated on men from households with high leverage as of baseline. Furthermore, there is no significant effect of local house prices on older individuals housing consumption and home equity-based borrowing.

This paper also contributes to the literature trying to understand the link between two important changes over the past four decades in the US : the increased labor force participation of married women and the actual price increase in housing prices. Theoretically, causality can go in any direction. The increase of relative housing costs can induce households to supply more units of labor by sending wives into the labor force (Warren and Warren (2003)). On the contrary, if married women are more educated and work more, households can bid up for houses where they can find better amenities, and it can raise the relative housing price (Franck and Cook (1995)). This paper shows that the price of housing has no causal effect on the labor force participation of older married women. This result brings a new insight since previous studies have focused most often on the situation of younger married women (see for e.g. Johnson (2014)). Moreover, the results suggest that married women exit from the labor force in response to the exit of their husband, but not directly in response to the housing boom.

This work contributes to the literature studying the effect of house prices on consumption. This research stream is motivated by two core empirical observations. First, the US aggregate consumption represents roughly two thirds of the Gross Domestic Product (GDP). Second, there is a strong positive correlation between house price and consumption Iacoviello (2011), Case et al. (2001). As a consequence, a lot of papers have tried to understand the link between the two variables, especially in regards to important implications for the monetary policy (inflation) and the business cycle. However, an important contradiction seems to emerge with the two key findings from this literature. Empirical estimates of the marginal propensity to consume out of housing wealth (MPC) indicate that the elasticity is significantly larger for older homeowners than for younger ones (see Poterba (2000), Campbell and Coco (2007)[1]). At the same time, Mian and Sufi (2011) documents that while young homeowners significantly increase their borrowing in response to the growth in their home equity to finance actual expenditure and home improvements, their older counterparts' borrowing behavior is less responsive to positive house price shocks. The results of this paper contribute in reconciling those findings. On the one hand, assuming that leisure is a normal good, older workers can increase their consumption of leisure by exiting from the labor force in response to housing wealth capital gains instead of raising their consumption of real outlays through home equity-based borrowing. This is consistent with the predictions of the permanent income hypothesis (see Friedman (1957)). Consistently with that, I have found a large labor supply exit of older men in response to the increase in house prices. On the other hand, older homeowners might be less likely to borrow against the increase in their home equity if they have relatively fewer debts than younger ones. I have found that the observed labor supply of older men primarily comes from individuals in households with lower borrowing constraints.

The paper is organized as follow. Section (3.2) presents data and the empirical strategy. Section (3.4) presents the results, Section(1.4) describes the robustness checks and Section(3.6) gives the conclusion.

[1]Their work is based on microdata from the UK Family Expenditure Survey from 1988-2000.

1.2 Data and empirical strategy

1.2.1 Data

This paper used data from several sources: (i) public and restricted microdata from the Health and Retirement Study (HRS); (ii) US mortgage applications data ; and (iii) counties and states' labor force statistics from the US Bureau of Labor Statistics (BLS).

Microdata

The HRS is a biennial panel data survey on labor supply, income, wealth, health and health insurance for a national representative sample of older US individuals aged above 50 and their household members of any age. Data have been collected every two year since 1992 by the University of Michigan's Institute of Social Research with the funding of the National Institute on Aging (NIA) and the Social Security Administration (SSA). New cohorts of individuals are regularly added to the initial sample. Information have been collected through face-to-face or phone interview.

Labor force participation and retirement in the HRS: In the HRS, an individual's labor force participation and retirement decisions are obtained through several questions that appear at different stages of the interview. As a consequence, a respondent can report working part-time at some point in the interview, but may report being retired at a later stage. A temporary laid off individual may report not working at a earlier stage of the interview, but mention holding a job at a later stage. Therefore, obtaining accurate information regarding respondents' labor force participation status can be time-consuming. This work used an enhanced version of the HRS public file from RAND Corporation, called RAND-HRS file. For each wave of the survey, all available information on a given topic is combined to provide a coherent and unique summary.[2] More precisely, the RAND-HRS file provides a unique variable for the respondent (and one's spouse, if applicable) labor force status during the reference period.[3] The panel A of Table (4.6) gives more details about the RAND-HRS labor force status variable.

In this work, retirement is principally defined as not working. This is because the unemployment rate is objectively quite low and the labor force participation rate is relatively low among the elderly. That is to say, an individual who is working neither full-time nor part-time is considered as a retiree. In the estimations, the robustness of this definition was tested using an alternative dummy variable equal to one when an individual labor force status is either working or unemployed, and

[2]For more information, please visit https://www.rand.org/labor/aging/dataprod.html

[3]In the HRS, full-time workers are individuals working 35+ hours per week and 36+ weeks per year. Otherwise, a worker is considered as a part-time worker.

zero otherwise. I have found similar results.

Home ownership: As for the labor force status, the HRS provides several pieces of information that needs to be combined to obtain a coherent view regarding the home ownership status of a respondent. I used two RANDFAT-HRS variables to identify homeowners in the sample.[4] Homeowners are defined as individuals who reported living in a house (a single detached home or apartment) owned by themselves.

Geographic information: In each wave, the HRS collects the place of residence of each respondent. This entails state and counties FIPS codes; as well as ZIP5 codes. This information is not available to the public. They have been obtained under a special agreement signed with the University of Michigan's Institute of Social Research. Moreover, whenever needed, geographic level data from other sources have been merged into the HRS microdata using location and year as identifiers.

The sample of interest consist of self-reported homeowners aged 57 to 64 working full-time or part-time as of year 1994. Table (4.7) presents the baseline summary statistics. Those individuals are observed between 1994 and 2004 (6 waves of the survey).

Local Labor Market Conditions

To assess local labor market conditions, this work utilized BLS data on unemployment rates (state and county level), county data on the unemployed and employed, and county' average weekly wages detailed by industry.

Mortgage Applications Data

The Home Mortgage Disclosure Act, enacted by the Congress in 1975 and enforced by the Federal Reserve Board under Regulation C, requires lending institutions to publicly report detailed information on loans. As a first step, all the mortgages transactions reported for the years 1994, 1996 1998, 2000, 2002 and 2004 (135,746,299 transactions) have been collected. Then, the loans originated by savings institutions, mortgage bankers, credit unions and other non-bank lenders were dropped, leaving 62,933,926($\equiv 46\%$) mortgage applications for financial institutions reporting to the Federal Deposit Insurance Corporation (FDIC), the Federal Reserve Board (FRB) and the Office of the Comptroller of the Currency (OCC). In addition, mortgages where borrowers are subsidized by the Federal Housing Authority, the Veterans Administration, and other government programs have also been dropped, reducing the sample to 56,714,944 applications. As only loans for home purchases (20,480,315 applications) have been maintained, mortgages refinancing, home improvement and

[4] Please, find more details in the panels B and C of Table (4.6).

multifamily dwelling purchases have been excluded. Among them, only mortgage applications for owner-occupied as a principal dwelling have been kept, resulting 18,393,089 applications. Finally, only loans originated were considered (9,857,859 $\equiv$ 7% of all the transactions).

1.2.2 Empirical specification

The empirical specification is the following:

$$y_{i,t,c} = \beta \text{HPI}_{t,c} + \delta_t + \delta_t \times b_i + X'_{i,t,c}\gamma + \alpha_i + \varepsilon_{i,t,c} \tag{1.1}$$

Where i is the index of the individual; t is the year and c refers to the county of residence. $y_{i,t,c}$ is a dummy variable equal to one if the individual reports working (full-time or part-time) and zero otherwise. $\text{HPI}_{t,c}$ is the average value of all loans originated as a result of mortgage applications for owner-occupied dwellings in county c during year t. δ_t represents year dummies that control for US national demographic labor force participation trends. b_i are dummies for birth year and the interactions of $\delta_t \times b_i$ which controls cohorts-specific labor force participation trends. X is the vector of other controls. α_i is the individual fixed effect and $\varepsilon_{i,t,c}$ the error term. β is the parameter of interest.

An aggregate measure of house prices is used in this study instead of individuals' self-reported house prices. This have been done because self-reported values are subject to measurement errors. In fact, homeowners tend to systematically overestimate the market value of their primary residence Benitez-Silva et al. (2015).[5] Further, the increases in self-reported house prices can be due to investment in housing such as renovations rather than better market conditions. This issue has motivated the use of aggregate house prices in many studies. Some studies used state-level house price changes (Case et al. (2001), Saxena and Wang (2017)). Others choose more local price indexes such as metropolitan area (MSA) level house price indexes(Faranham and Sevak (2016), Zhao and Burge (2017)) or measures of housing wealth at the zip code level Mian and Sufi (2011). Nonetheless, OLS estimates of β are likely to be biased for another important reason.

Savings and leisure preferences vary from individual to individual. Therefore, an observed correlation between house price and labor force participation decision is likely to be contaminated by these unobserved heterogeneities. Control over individual's fixed effects removes the confounding effects of time-invariant factors. However, there are some unobserved factors that change over time,

[5] Using data from the Health and Retirement Study, he estimated that on average, homeowners overestimate the value of their properties by around 8%

such as local labor demand shocks and anticipations that cannot be fully controlled. To illustrate, a positive local labor demand shock that drives up wages can explain both housing market dynamics and labor force participation decision at older ages. Better wages imply that households can bargain for more expensive houses (or houses located where better amenities exist such as lower crime rates), which might raise the local prices of real estate.[6] If this shock affects occupations held by older individuals , they can find it advantageous to postpone labor market exit since retiring later might imply higher Social Security retirement benefits (French (2005), Rust and Phelan (1997)). On the contrary, if the labor demand shock implies a shift toward higher demand of younger workers, for example due to technological changes, older workers would be more likely to withdraw from the labor market since they are exposed to worsening labor market conditions. This simple arguments means that OLS estimates of the effect of house price growth on the retirement decision of older workers are likely to be biased upward or downward. It is also possible to find no remarkable correlation if the dynamic of the housing market is fully anticipated by households. Previous studies have used different identification strategies.

Some studies estimated the deviation from state or national trends to obtain unexpected house price growth. For example, Faranham and Sevak (2016) in the US used the residuals from the regression of MSA-level price changes on state-specific trends while Disney and Gathergood (2014) in the UK considered deviations from national housing wealth trends. This type of strategy cannot rule out labor demand-based explanations as local labor demand shocks can drive real estate market values far from state or national trends. It is also not helpful in order to compute unanticipated variations in the market value of the real estate because of the persistence in US house price dynamics. For example, Himmelberg et al. (2005) showed that US price-to-rent ratio differences across cities are very persistent over time. Gyourko et al. (2013) also argued that some cities have a very high long-run rates of house prices growth. These findings suggest that workers can anticipate, at least to some extent, the local house price dynamics. Other studies relied on an instrumental variable approach.

Mian and Sufi (2011) instrumented house prices with the Saiz's elasticity of housing supply[7,8,9]. This type of identification strategy has been recently criticized by Davidoff (2016). He

[6] Himmelberg et al. (2005) found that the overall rise of the real price of housing as well as spatial differences of its dynamics cannot be explained by construction costs.

[7] Saiz (2010) constructed this index using topographic measures in order to assess, for each US metropolitan area, the easiness of increasing the housing stock. A low value of the elasticity means that constraints toward building new homes are important.

[8] In their robustness checks the authors also used the Wharton Regulation Index (Gyourko et al. (2008)). This index measures how difficult building new homes in a metropolitan area can be because of the local regulations. The identification issues are similar to those that arise with the elasticity of housing supply.

[9] This identification strategy was also used in Stroebel and Vavra (2015).

documented the correlation between housing supply constraints and determinants of local housing demand, such as historical education levels, immigration and national employment growth in industries which are main drivers of local markets.

Using data from China, Fu et al. (2016) used the average housing capital gains of properties in the neighborhood to instrument individual's own housing wealth gains. Unfortunately, this identification strategy cannot rule out the possibility that some demand factors, such as local productivity shocks, are responsible for both local house wealth changes and labor supply decisions. Therefore, finding a more convincing identification strategy appears to be necessary.

1.2.3 Identification strategy: An instrumental variable approach

To identify a plausible exogenous variation in local house prices, this work is based on the identification strategy proposed by Favara and Imbs (2015). In 1994, the Interstate Banking and Branching Efficiency Act (IBBEA) allowed banks to expand geographically by creating branches across state borders. However, states were also given the right to limit out-of-state branching by setting restrictions on four domains[10]: (i) the minimum age of the institution of acquisition, (ii) allowance of de novo interstate branching, (iii) allowance of interstate branching by acquisition of a single branch or portions of an institution, and (iv) statewide deposit cap on branch acquisition. I used the deregulation index calculated by Rice and Strahan (2010). For each year and each state, the deregulation index is equal to the number of out-of-state branching restrictions.[11] For instance, the index is equal to zero when there are no restrictions at all, one when there is only one restriction put by the state authority and four when all the four restrictions are imposed. In this paper, the index is reverse coded, so high values of the index mean more deregulation.

Because state authorities were allowed to put restrictions on interstate branching, the level of the deregulation as measured by the number of restrictions state authorities imposed a given year substantially change overtime for most individual states as well as between states during the decade that follow (1994 to 2005).

The authors took the advantage of this variation to establish that the deregulation caused the rise of local house prices through exogenous credit expansion. In fact, comparing the financial institutions affected by the deregulation with the unaffected ones comparably similar, the credit expansion was only due to the fact that the treated institutions created new branches in neighboring states and facilitated access to credit to increase their market shares. As a result, the housing

[10] Johnson and Rice (2008) documented that most states exerted their authority regarding interstate branching.

[11] It is impossible to distinguish the effect of each four types of restrictions separately because states used to put or remove several restrictions at once. See Table (1), pp. 868-869 of Rice and Strahan (2010).

demand rose in affected counties, but not homogenously. Since the housing supply elasticity differs across counties because of natural geographical constraints, house prices have increased a lot faster in geographically constrained counties. This paper exploits this result.

Figure (1.1) below illustrates how the identification strategy works in the case of counties where housing supply curves are perfectly inelastic. Let us consider two states indexed 1 and 2, with one county in each state. $D_k, k \in \{1,2\}$ is the housing demand curves for counties 1 and 2 respectively, while $S_k, k \in \{1,2\}$ is the supply curves. The demand curves are parallel and so are the supply curves. This is because the counties are supposed to have the same housing demand elasticity, which implies that the two either share the same labor demand determinants or belong to the same local labor market. Panel A1 represents the situation before the deregulation in state $S1$ and Panel A2 shows the situation after the deregulation. From Favara and Imbs (2015), a deregulation in state $S1$ created an exogenous expansion of credit in the treated state implying a positive housing demand shock in the county 1, but nothing in the county 2. Therefore, there is a house price growth in the treated county holding other factors constant. Given the assumed similarity of housing demand factors in the two counties, the observed difference in house price growth can be considered to be solely driven by the deregulation.

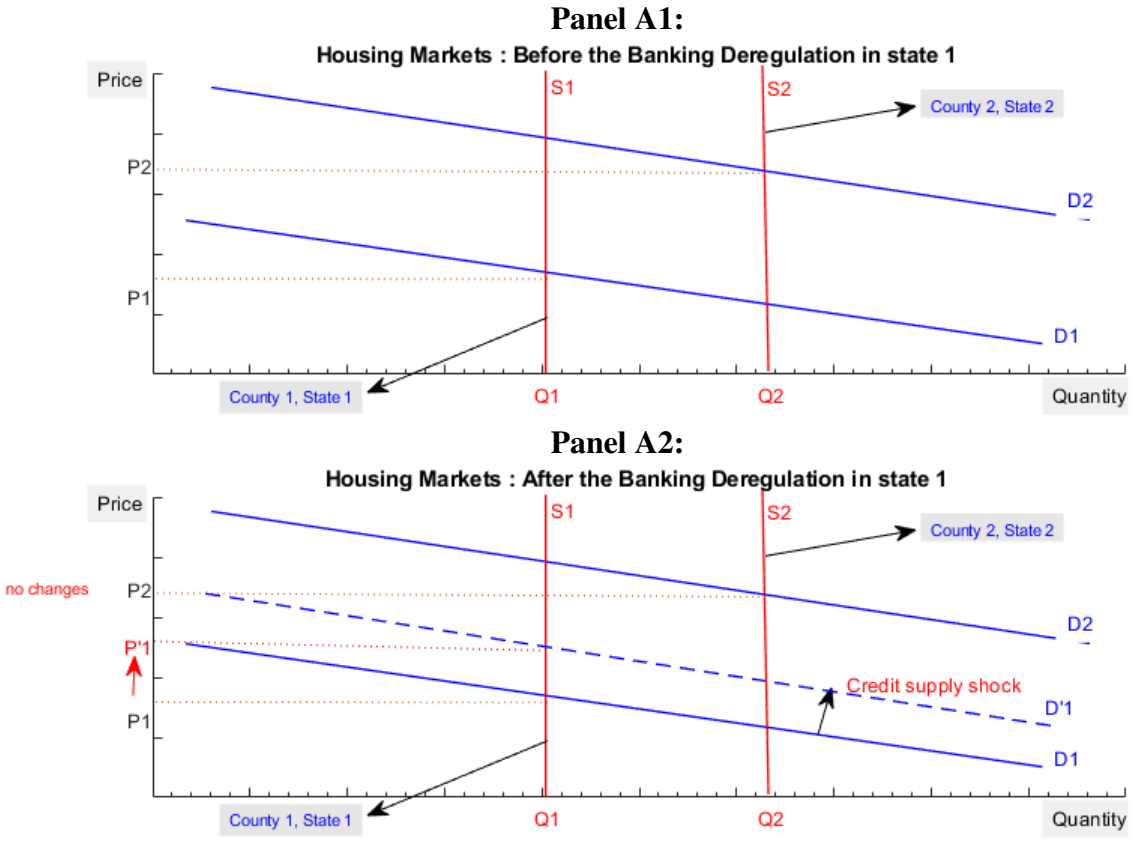
In Figure (4.6), the case of perfectly elastic housing supply curves are considered. In this case, credit expansion rather increases the stock of houses in the treated county. In the untreated county, it is the status quo. The identification hypothesis is, therefore, the following: holding the housing supply constant, the differences in house price growth between more deregulated counties and less deregulated ones can be solely attributed to the credit expansion driven by the banking deregulation.

This, in this paper, I instrumented the local house price growth by the lagged value of the deregulation index interacted with the elasticity of housing supply. Equation (1.1) is estimated using a two-stage least squares (2SLS) approach. In the first stage, the local house price growth variable is regressed on the instrument and the other controls:

$$\text{hpi}_{t,c,s} - \text{hpi}_{t-2,c,s} = (\text{deregulation index})_{t-2,s} \times (\text{elasticity of housing supply})_c + \text{controls} + \alpha_i + \varepsilon_{i,t,c,s} \tag{1.2}$$

In the second stage, the dependent variable of interest is regressed on the instrumented house prices and the controls:

Figure 1.1: Illustration of the identification strategy



$$y_{i,t,c,s} - y_{i,t-2,c,s} = \overline{\text{hpi}_{t,c,s} - \text{hpi}_{t-2,c,s}} + \text{controls} + \alpha_i + \psi_{i,t,c,s} \tag{1.3}$$

To be valid, the instrumental variable used in this work should be relevant and satisfy the exclusion restriction assumption. The next two sections shows results supporting the validity of my identification strategy. More specifically, several robustness test (including a falsification test) are conducted to rule out any alternative labor demand-based explanations, which can invalidate the validity of the instrument.[12]

1.3 Results

1.3.1 Main result

Table (1.1) presents the OLS and 2SLS estimates for the effects of house prices on the labor force participation of older individuals.[13] The sample is constituted by men and women aged between 57 and 64, who were homeowners and working as of year 1994. Local house prices are measured by the average value of all mortgages originated within a county in a given year. The instrumental variable is the interaction between the state index of deregulation (in lag) and the county elasticity of housing supply.

The OLS estimations are negative and significant except for the female subsample, for which the estimate is positive and not significantly different from zero. 2SLS estimates are all greater, in absolute terms, than OLS estimates. This confirms that OLS estimations are biased as argued in the previous section. Besides, 2SLS estimates are all negative (but not significant for women), which means that higher housing wealth implies lower labor force participation at older ages. The first stage estimates confirm the statistical validity of the 2SLS results. They shows a strong and significant correlation between the instrumental variable and the local house price measurements for all the three groups of observations. Furthermore, each F-statistic is greater than the common

[12] The small sample size did not allow me to conduct another interesting robustness test: running the estimations on the sample of individuals living in MSA bordering at least two states such that one can compare treated counties in one state with an untreated one in the same local labor market. In large samples, this identification strategy can provide good results.

[13] I found similar results when measuring the labor force participation with a dummy variable equal to one when the individual was either working or unemployed. See Table (4.8).

accepted threshold of 10 Stock et al. (2002), except for the female subsample of women where the F-stat is 9.380. This may be due to the smaller female sample size.

Table 1.1: OLS and 2SLS Estimations of the Effect of House Price Growth on the Labor Force Participation of Older Individuals

Dependent var. :	**Δ (Working FT or PT)**					
	OLS			**IV**		
	All	**Men**	**Women**	**All**	**Men**	**Women**
ΔLog(County Average Mortgage)	-0.159**	-.328***	0.0368	-1.082**	-1.697***	-0.248
	(0.0605)	(0.0713)	(0.107)	(0.456)	(0.475)	(0.676)
	First Stage Results					
(Elasticity of housing supply)×(Banking Deregulation Index)$_{t-2}$	-	-	-	-0.0029171***	-0.0030548***	-0.0027879***
	-	-	-	(0.000835)	(0.0007925)	(0.0009103)
Kleibergen-Paap F-stat	-	-	-	12.206	14.858	09.380
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)*(year dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Dummies for year of reform change (DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)*(DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
Dependent var (mean)	-0.131	-0.130	-0.132	-0.131	-0.130	-0.132
ΔLog(County Average Mortgage)(mean)	0.051	0.051	0.051	0.051	0.051	0.051
ΔLog(County Average Mortgage) (sd)	0.094	0.094	0.094	0.094	0.094	0.094
# Obs.	7060	3827	3233	7060	3827	3233

***:1%. **:5%. * : 10%. Residuals are clustered at the state level. The sample is constituted by individuals born between 1931 and 1938, who were homeowners and working as of year 1994.

Results for the whole sample (men and women): The OLS and 2SLS estimates are both consistent with a significant negative effect of housing wealth gains on the labor force participation of older individuals. A 10% increase of the average county house price[14] decreases the probability of working by 0.103 i.e. 10.3%. The average change of the housing variable over a two-year period is 5.1%. On average, the labor force participation decreases by $0.051 \times 1.082 = 0.055$. This represents a total decrease of $(\frac{0.055}{0.131}) \times 100 = 42\%$ for the average county. In other words, for the average county, a 42% of older individuals' labor force exit during the period can be attributed to the banking deregulation. This signifies a huge effect.

Results for men: A 10% increase of the average house price decreases the men's probability of working by 0.162 i.e. 16.2%. Considering the average change of the housing variable over a two-year period, the labor force participation decreases by $0.051 \times 1.697 = 0.086$. This represents a total decrease of $(\frac{0.086}{0.130}) \times 100 = 66\%$. This implies that, for the average county, two-thirds of men's labor force exit was due to the deregulation.This huge labor supply exit in response to housing wealth capital gains is consistent with the user cost of housing theory from Morris (2006).

From this theory, as housing wealth grows the user cost of housing also rises[15] reducing housing

[14]Which also corresponds to about the increase of a value corresponding to the standard deviation of the natural logarithm of house prices.

[15]The key driver of the increase of the marginal cost of housing services is the rise of property taxes and costs of maintenance. On the contrary, inflation and home price appreciation decrease the user cost of housing.

capital gains. However, older individuals face the shorter time horizon, ao the wealth capital gains increase with age. This means that a positive housing wealth shock translates into higher wealth capital gains for older individuals than for younger ones. In this context, the permanent-income hypothesis (Friedman (1957)) predicts an increase of the consumption of leisure. The housing wealth effect above is much larger than the findings from previous empirical studies.

Faranham and Sevak (2016) used data from the HRS between 1992 and 2002 and found only 6.5% increase of men's probability of retirement in response to a 10% increase of the local house prices. The authors found that the effect vanished when they controlled for county unemployment rate as a measure of local labor market conditions. This smaller effect they found can be attributed to a specific weakness in their identification strategy. The local house price variable used in their paper was the percentage change of the MSA house price index residualized from state fixed effects. This variable is actually a measure of the deviation of local house prices from the corresponding state average. However, because US house prices are highly persistent (see Himmelberg et al. (2005), and Glaeser and Gyourko (2007)), their estimates are likely to be downward biased. In fact, the strong persistence of local house prices in the US implies that deviations from state average are likely to be anticipated, at least partially, by older workers, so they are already incorporated in their life-cycle decisions such as retirement. Hence, the smaller effect they found does not contradict this paper's results.

Zhao and Burge (2017) found effects that were much smaller. They used data from the HRS between 1991 and 2010, and found that on average, doubling the housing wealth is necessary to obtain a 5% reduction of the labor force participation of the elderly. Assuming linearity, this translatea into a 0.5% rise in the retirement probability for a 10% increase in the housing wealth. They applied two identification strategies. The first one was a difference-in-differences identification strategy where homeowners were compare to renters. However, since the home ownership status is endogenous, the estimates are likely to be biased. In fact, homeowners and renters in these age groups are likely to differ systematically in terms of their time-varying preference for leisure and savings. In the second approach, a linear probability econometric model was estimated in differences. hereafter, self-reported house price was used as a proxy for an individual's housing wealth. However, measurement errors in self-reported house prices are likely to bias the estimated effects downward.

Results for women: Both OLS and 2SLS estimates show no statistically significant effect of house prices on the labor force participation of older women. Nonetheless, the average labor force exit of women is quite similar to that of men; 13.2% versus 13.0%. This suggests that women and men tend to exit from the labor force at the same time. This, therefore, raises a natural question: why is women's labor force participation not responsive to changes in the local house prices?

In the sample, 70% of women are married or partnered. Table (4.9) gives OLS and 2SLS estimates for married individuals. The results display a positive but not significant effect of house prices on the labor force participation of older married women. The positive sign of the estimates is consistent with the stylized fact; there is a positive empirical correlation between local house prices and married women labor force participation in the US Johnson (2014). The observed insignificant effect suggests that while married men retired in response to the boom in the housing market, married women rather stopped working in response to their husband's exit from the labor force. This explanation is consistent with findings from the literature. It has been documented that, the complementary of leisure explains a great part of joint retirement decisions of older married couples in the US (Gustman and Steinmeier (2000),Blau (1998), Michaud et al. (2019)).

However, the reason why married men's labor force participation responded to house prices and not that of married women remains to be explained. In the sample, married men, on average, are more than 5 years older than their spouses. In fact, Table(4.7) displays the average spouse's age of 55.67 years old for married men and of 61.47 years old for married women. Because of the rules of the US Social Security retirement policy, age is a major determinant of men's retirement decisions (French (2005),Rust and Phelan (1997)). This, therefore, implies that married men exited in response to the housing boom as they were older than their wife i.e. closer than their wives to the institutional retirement ages.

Furthermore, it is important to note that the results from Table (1.1) do not rule out the possibility that single older women might have exited from the labor force because of the result of the housing boom. Actually, 2SLS estimates for all women are negative in sign (but not significantly different from zero), but positive in sign for married women (see Table (4.9)). This may suggest that single women and married women responded differently to the housing shock. However, since the sub-sample of single women is very small, it is not possible to rigorously test this with current data.

Heterogeneity in Households' Borrowing Constraints

To examine whether men's labor supply response to house prices shocks vary with the level of their household borrowing constraints, two measures of borrowing constraints have been constructed. These measures are the Loan-To-Value (LTV) and Loan-To-Income (LTI) ratios.[16] A household is considered as having a high leverage (i.e a low ratio) if the ratio used is below its median value. Otherwise, the household is considered of having a low leverage (high ratio). Households have been classified using income and debt values as of year 1994.

[16] See appendix **??** for a detail description of how those variables were computed using HRS data.

Table (1.2) shows that households borrowing constraints played an important role in the labor market decisions of older men during the 1994-2005 housing boom. 2SLS displays no significant effect of local house prices on the labor force participation of older men with low leveraged households (i.e. higher mortgage debts relative to their home value). However, for older men belonging to households with high leverages (i.e. lower mortgage debts relative to their home value), the effect is very significant and higher (in absolute terms) than the average effect for all men. This proves that men who exit from the labor force responding to the housing boom were those from households with relatively lower mortgage debts. This results is consistent with the predictions from the user cost of housing theory Morris (2006).

From this theory, the marginal propensity to consume out of housing gains (MPC) is negatively correlated to the user cost of housing. However, mortgage payments are components of the user cost of housing, so the lower the mortgage payments are, the lower the user cost is. As a consequence, the MPC is higher for households with high leverages than for households with low leverages. In other words, households with smaller mortgage payments benefit more from the housing boom than those with higher mortgage payments. I found similar results when using the LTI ratio instead of the LTV ratio (see Table (4.10)).

Table 1.2: OLS and 2SLS Estimations of the Effect of House Price Growth on the Labor Force Participation of Older Men

Dependent var. :	**Δ (Working FT or PT)**					
	OLS			**IV**		
	All	**High Leverage**	**Low Leverage**	**All**	**High Leverage**	**Low Leverage**
ΔLog(County Average Mortgage)	-0.328***	-0.363**	-0.386***	-1.697***	-2.870***	-0.349
	(0.071)	(0.141)	(0.118)	(0.475)	(0.984)	(0.482)
	First Stage Results					
(Elasticity of housing supply)×(Banking Deregulation Index)$_{t-2}$	-	-	-	-0.0030548***	-0.0031921***	-0.0045716***
	-	-	-	(0.0007925)	(0.0009089)	(0.0008125)
Kleibergen-Paap F-stat	-	-	-	14.858	12.335	31.658
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)*(year dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Dummies for year of reform change (DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)*(DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
Dependent var (mean)	-0.130	-0.126	-0.113	-0.130	-0.126	-0.113
ΔLog(County Average Mortgage)(mean)	0.051	0.051	0.051	0.051	0.051	0.051
ΔLog(County Average Mortgage) (sd)	0.094	0.094	0.094	0.094	0.094	0.094
# Obs.	3827	1077	934	3827	1077	934

***:1%. **:5%. * : 10%. Residuals are clustered at the state level. The sample is constituted by men born between 1931 and 1938, who were homeowners and working as of year 1994

House Price Growth, Housing Consumption and Home Ownership

Table (4.11) reports OLS and 2SLS estimates of the effect of house prices on older men's housing consumption and home ownership status. The variable in the first column is a dummy equal to

1 if the individual reported a second mortgage [17] and zero otherwise. The variable in the second column is a dummy equal to 1 if the individual reported living in an apartment dwelling. Finally, in the third column, the probability of becoming a renter in response to the house price boom is documented. It the 2SLS, it appears that all the estimates are not significantly different from zero. This proves that older men have not changed their housing consumption regarding the housing boom.

House Price Growth and Home Equity-Based Borrowing

Are older men more likely to borrow against home equity increase? Table (4.12) reports 2SLS estimates of the effect of house prices on the probability of having: (i) an equity line of credit or ; (ii) a loan on an equity line of credit. The dependent variables are set to zero for any home equity borrowing behaviors as of 1994. There is no significant effect of housing wealth capital gains on the home equity-based borrowing of older men. Similar results were found for the overall population. These results are consistent with the findings from Mian and Sufi (2011).

1.4 Robustness checks

This section demonstrates that the results stated above are robust to several alternatives explanations or statistical issues.

1.4.1 Robustness to labor demand-based explanations

The main result of this paper is that house prices growth have caused the labor force exit of older individuals in the US. However, this result can be ruled out by labor demand-based explanations. In fact, one can argue that house prices and labor force participation at older ages during the study period were both driven by a third unobserved variable: local labor demand shocks. Below, two arguments are provided to reject this alternative explanation.

First, a falsification test is conducted based on the following argument: if the previous results were caused by unobserved changes of the local labor market conditions instead of the banking deregulation, we should observe a direct effect of the instrumental variable on the labor force participation of older men regardless of their home ownership status. Table (1.3) reports the reduced form estimates of the effect of the instrumental variable on the labor force participation of older

[17] Conditional of not having a second mortgage as of 1994.

individuals in the studied sample. There is no significant effect of the instrumental variable on the labor force participation of older women irrespective to their home ownership status. However, for the whole sample and the subsample of men, there is significant effect for homeowners but not on renters. Furthermore, the estimates for renters are smaller than those for homeowners.

Table 1.3: Reduced Form Relationship Between the (County Elasticity of Housing Supply) ×(Banking Deregulation Index) and the Labor Force Participation Dummy

Dependent var. :	**Δ (Working FT or PT)**					
	Men and Women		**Men**		**Women**	
	Owners	**Renters**	**Owners**	**Renters**	**Owners**	**Renters**
(Elasticity of housing supply)×(Banking Deregulation Index)$_{t-2}$	0.0031575**	0.0028092	0.0051841***	0.000874	0.0006914	0.0033941
	(0.00144)	(0.00198)	(0.00156)	(0.00233)	(0.00197)	(0.00341)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)×(year dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Dummies for year of reform change (DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)×(DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
Dependent var (mean)	-0.131	-0.131	-0.130	-0.115	-0.132	-0.146
# Obs.	7060	1940	3827	907	3233	1033
R2	0.042	0.053	0.047	0.080	0.056	0.077

***:1%. **:5%. * : 10%. Residuals are clustered at the state level. The sample is constituted by individuals born between 1931 and 1938, who were working as of year 1994.

Second, the instrumental variable is regressed on several measures of local labor market conditions. Table (1.4) gives the reduced form estimates of the effect of the instrumental variable on county and state unemployment rates. The results highlihgt no significant effect. In addition, OLS and 2SLS estimates of the effect of house prices on county and state unemployment rates also show no significant effect (see Table (4.13) and Table (4.14)).

Table 1.4: Reduced Form Relationship Between the (County Elasticity of Housing Supply) ×(Banking Deregulation Index) and Local Labour Market Conditionsy

Dependent var. :	**Δ county UR**	**Δ Log(county UR)**	**Δ state UR**	**Δ Log(state UR)**
(Elasticity of housing supply)×(Banking Deregulation Index)$_{t-2}$	-0.0000325	-0.0042716*	-0.0000831	-0.0019564
	(0.000213)	(0.00227)	(0.000156)	(0.00368)
Year dummies	Yes	Yes	Yes	Yes
county dummies	Yes	Yes	Yes	Yes
Dummies for year of reform change (DYRC)	Yes	Yes	Yes	Yes
Dependent var (mean)	-0.001	-0.000	-0.001	-0.019
# Obs.	874	874	874	874
R2	0.678	0.590	0.859	0.837

***:1%. **:5%. * : 10%. Residuals are clustered at the state level. UR ≡ unemployment rate.

1.4.2 Robustness to the estimation of standard errors

In all the previous estimations, standard errors are clustered at the state level. This has been done to take into account the plausible heteroskedascity of unknown form of errors in Equation (1.1). Clustering is a typical process used to ensure that the estimated standard errors are not too smaller than what they should be. Incorrectly smaller standard errors can lead to wrong inference. For example, it can be found the effect of house price growth on the retirement decision of older men to be significantly different from zero while it is not the actual case. As I have done, clustering errors at the state level impose the restriction that errors are correlated within a state but not across states. It means that unobserved factors that might affect both county house price growth and labor force participation decisions of older individuals are likely to be correlated only for individuals living in the same state. However, if individuals living in neighboring states belong to the same local labor market, the estimated standard errors with one-way clustering could be lower than what they should be. Given that more than 26% of my observations belong to metropolitan areas bordering two or more states, this issue appears to be important. To handle the problem I have followed the two-way clustering solution proposed by Cameron et al. (2011), Miglioretti and Heagerty (2006), Thompson (2006) and Thompson (2011).

Two-way clustering: I have clustered errors on both the state and the commuting zone. This means that errors are conditionally uncorrelated unless an individual belong to either the same state or the same commuting zone. A commuting zone is a geographic unit form of a collection of counties that share tight economic relationships (in terms of trade, labor demand and supply, etc.) measured by commuting flows. Commuting zones have been developed by Tolbert and Sizer (1996) to define a local labor market or regional economy.[18] Since then, they have been widely used in the economic literature including influential papers [19](Amior and Manning (2015), Chetty et al. (2014), Author et al. (2013), Restrepo (2015),Yagan (2016)).The interaction of commuting zones and states is motivated by the well-known work of Moulton (1986) and Moulton (1990). He showed that in the presence of grouping variable such as a policy that had the same outcome for all individuals in a given geographic area, clusters should be defined at an aggregate level equal or higher than that of the grouping variable. Since my policy variable, the Rice-Strahan deregulation

[18]The 1990's definition of commuting zone is the one used in this paper. In 1990, 741 commuting zones (CZs) were constructed in the US. However, there were two other sets of commuting zones. In 1980, 768 were constructed and in 2000, the update retained only 709 CZs. Since the deregulation started in 1994, I assume that local labor markets that matter for individuals in my sample are quite well approximated with the 1990's information. Using more recent information might be problematic if the housing boom affects caused large-scale migration that change the relevance of some CZs.

[19]On the contrary of Metropolitan area; where construction is also based on commuting flows, commuting zones have the advantage of covering the entire US.

index, is defined at the state level, I cannot simply cluster my errors at the commuting zone level which is lower than the state level. This assumption is not too restrictive if one assumes that commuting zones are pretty good approximation of local labor markets in the US.

Table (1.5) presents two-stage least squares estimates of the effect of house price growth on the retirement decision of older men. Standard errors obtained from the two-way clustering are very similar to those obtained earlier with one-way clustering.

Table 1.5: OLS and 2SLS Estimations of the Effect of House Prices on the Labor Force Participation of Older Individuals (*with two-way clustering*)

Dependent var. :	**Δ (Working FT or PT)**					
	OLS			**IV**		
	All	**Men**	**Women**	**All**	**Men**	**Women**
ΔLog(County Average Mortgage)	-0.159**	-.328***	0.0368	-1.082**	-1.697***	-0.248
	(0.0634)	(0.0933)	(0.123)	(0.449)	(0.464)	(0.700)
	First Stage Results					
(Elasticity of housing supply)×(Banking Deregulation Index)$_{t-2}$	-	-	-	-0.0029171***	-0.0030548***	-0.0027879***
	-	-	-	(0.0008337)	(0.0008133)	(0.0008778)
Kleibergen-Paap F-stat	-	-	-	12.242	14.107	10.087
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)*(year dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Dummies for year of reform change (DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)*(DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
Dependent var (mean)	-0.131	-0.130	-0.132	-0.131	-0.130	-0.132
ΔLog(County Average Mortgage)(mean)	0.051	0.051	0.051	0.051	0.051	0.051
ΔLog(County Average Mortgage) (sd)	0.094	0.094	0.094	0.094	0.094	0.094
# Obs.	7060	3827	3233	7060	3827	3233

***:1%. **:5%. * : 10%. Residuals are clustered at the state and commuting zone levels. The sample is constituted by individuals born between 1931 and 1938, who were homeowners (of single family-unit) and working as of year 1994.

The number of clusters can be another issue regarding the accuracy of the estimated standard errors. As Cameron and Miller (2014) pointed out, there is a consensus to use 50 clusters for state-year panel data.[20] However, my sample covers only 33 states. To confirm the robustness of the results in this work to the number of clusters, I replicated the main estimates but with the standard errors estimated using the Wild-Cluster bootstrap approach (See Table (4.15)). All the estimates are still significant.

[20]However, Cameron and Miller (2014) in their review of the literature pointed out that in the unbalanced panel case, the right number of clusters could be larger than 50.

1.5 Conclusion

This paper has documented the estimates of the causal effect of local house prices on the labor supply of older individuals in the US during the housing boom period 1994-2004. It has also shown that there are important heterogeneities across gender and households borrowing constraints in the labor supply in the response of older individuals to positive house price shocks. The effect of local house prices on the housing consumption and home equity-based borrowing behavior of older individuals has been examined as well.

An instrumental variable strategy has been used to obtain plausible exogenous variation in local house prices. More precisely, local house prices are instrumented by the interaction of plausible credit supply shocks and housing supply constraints. I have explained that this strategy is robust the alternative explanations and to the weak identification problem.

The results are consistent with an important "housing wealth effect". I found that a 10% increase of the local house prices raise the probability of older workers exiting from the labor force by 10% overall, and by roughly 16% for older men. For the studied sample, the estimates imply that the housing boom accounted for about 40% of the labor market exit of older individuals between 1994 and 2004. For older men, local house prices accounted for 66% of their labor supply exit during the same period. However, I have found no significant effect of local house prices on the labor supply of older women. It has also been found that an increase of local house prices has no effect on older individuals' housing consumption and home equity-based borrowing. Consistent with some theoretical predictions, the majority of the labor supply response comes from men in households with low borrowing constraints.

This paper shows that housing wealth is more important in labor supply decisions of older men than what previous studies have documented. However, two essential questions remain to be answered. First, the number of working months or years lost in response to the housing boom is unknown. Secondly, do older individuals sometimes wait after the beginning of a housing boom before deciding whether to retire or not? If so, how long? Answering these questions is critical to compare the effect of housing wealth to the documented effects of other key determinants of retirement such as the minimum age for claiming social security benefits, the retirement earnings test, health and health insurance.

Chapter 2

Labor Market Conditions, Health and Retirement

2.1 Introduction

Are people in good health more likely to stay in the labor market when unemployment rises? This paper empirically estimates the relative impact of employment shocks on the labor market participation of US older workers with heterogeneous health conditions. This work is important for at least two main reasons.

First, the US current retirement policy fosters labor force participation at older ages, which helps understand that the determinants of retirement are of high interest for policy-makers. Since the early 1980s, the US Social Security retirement policy has been subject to several reforms which aim to ensure the financial viability of the program. Some of these reforms directly incentivized older workers' participation in the labor market. For example, the major reform of 1983 raised the full retirement age (FRA) from the age of 65 to 67, on a phased basis from 2000.[1] Another key change was the removal of the retirement earnings test at or beyond the FRA in 2000. This makes it possible for workers, aged at or beyond the FRA, to collect all their retirement benefits while working.

Second, health has been shown to be a major determinants of labor supply at older ages. However, its relative importance during an economic downturn is much less documented. For example, Dwyer and Mitchell (1999) used the first wave of the Health and Retirement Study (HRS) to document that men in poor health retired between one to two years earlier than their healthy counter-

[1] This component of the reform was recognized to have accounted for 34% of the total net savings of the legislation in the long-run. For further references, see :https://www.ssa.gov/policy/docs/ssb/v70n3/v70n3p1.html

parts. They also observed that health is a more influential determinant of retirement behavior than economic variables. John Bound and Waidmann (1999) using longitudinal data from the HRS, examined how the timing of health shocks affected labor force behavior of older workers. They found that declining health helps explain retirement behavior as current health status do. Their results suggest that studying health in a dynamic framework gives more information about its role in the labor force transitions of older individuals. Other empirical papers support the negative correlation between health and retirement age. It is the case of (Rena M. Conti and Frank (2006)) who documented a negative effect of depression on retirement age. As empirical papers do, structural works have also established the important role of health in the retirement decision of older workers (see for example John Bound and Waidmann (2010), French and Jones (2011), French (2005), Rust and Phelan (1997)). Nonetheless, how older workers with heterogenous health conditions respond to employment shocks is not well known.

For the main estimations, I used the two consecutive waves 2006 and 2008 of the HRS. I considered survey respondents aged between 51 and 69, and who were working as of 2006. The exit of most of older workers from labor market in 2008, in response to the recession related state unemployment hike, seems to be associated with poor health condition as measured more than one year before the recession. This result is consistent across several health measures such as self-reported health status, cognitive ability, depression risk and ability to perform daily activities.

The impact is also quantitatively salient. A 10% increase in the local unemployment rate implies a rise of 22.6% in the labor market exit of workers who had self-reported being in poor health before the recession. Building from simulations done in previous studies, this effect is similar to the increase of about $ 45,200 in the Social Security wealth of the workers studied. To further illustrate, this effect is comparable to around 40% increase of the median Social Security wealth of workers with poor health in the studied sample. This work, therefore, documents a much larger effect of health on the labor force participation of older workers during economic downturn in general, and the Great Recession in particular, compared to what has been found in previous studies.

This paper also confirms the crucial role of the US Social Security policy in the labor market decisions made by older workers as shown in previous studies. In response to local unemployment hike, I found that all observed labor market exits appeared for workers age at least 62, the minimum age for claiming Social Security retirement benefits. In fact, no significant effect was noticed among workers aged below 62.

Nonetheless, I discovered notable differences with respect to educational achievement and marital status. I observed no effect of local unemployment shocks on the labor force participation of college graduates and the individuals who are neither married nor partnered. However, it is interesting to note a strong labor supply response to local employment conditions for married men and

women and non-college graduates in poor health.

This paper contributes to the literature on the retirement behavior of US older workers during the Great Recession. Recent developments regarding this topic were carried out by (Coile and Levine (2009),Coile and Levine (2010), Coile and Levine (2011) and McFall (2011)). All of these studies have two main findings. First, they document that local labor market conditions, as measured by the unemployment rate at the state level, increase the probability of older workers to stop working. Unfortunately, the effect the authors found is not much remarkable. Second, these previous studies have shown that workers retire as soon as they qualify for Social Security retirement benefits (SS). Nonetheless, these works are limited in two areas I have been able to overcome. The studies did not consider health as a source of heterogeneity in the labor supply responses of older workers to local employment shocks. This is probably due to the fact that health was not collected in the CPS[2] data , the data source used in most previous studies. Also, the empirical specifications used in theses researches do not control individuals fixed effects, which can explain their smaller effects.

The paper is organized as follow. The section (2.2) describes the data used and the empirical specification. The section (2.3) presents and discusses the results obtained. In the Section 4, the main robustness tests conducted are presented. Finally, the section (2.5) gives the conclusion.

[2]Current Population Survey

2.2 Data and Empirical Specification

2.2.1 Data

This work uses data from the Health and Retirement Study (HRS). The HRS is a biennial representative survey of older adults (aged over 50) in the US. Since 1992, the HRS provides a rich description of health and health insurance coverage, work and retirement decisions, financial situations and family background of older Americans. It is a longitudinal survey conducted by the University of Michigan Institute for Social Research, funded by the National Institute on Aging and the Social Security Administration.

For the main analyses, I use data from the waves 2006 and 2008 of the survey. The wave 2006 has been chosen as our baseline since its allow us to observe older workers characteristics and outcomes just before the recession which started in the fall of 2007. The wave 2008 leads us to observe individuals labor market outcomes at the beginning of the recession. Nonetheless, for the robustness checks I consider other waves of the survey between 1998 and 2010.

2.2.2 Health measures in the HRS

The HRS is a unique survey because it contains a very rich set of measures of physical health and cognitive and affective functioning.

Measures of physical health

The principal measure of health uses in this paperr is the *self-reported health status*. This variable has been extensively used in the economic literature because it turns out to be a good measure of an individual's working capacity (see for example Lange; and McKee (2012) and Currie (1999)). In the HRS, the question wording is "How would you rate your health at the present time? Would you say it is excellent, very good, good, fair, or poor?". Following what has been done in previous studies, I consider an individual to be in poor/bad health if he/she self-reported his/her health status as good, fair or poor. Otherwise, the person is considered to be in good health. According to this definition, 70.5% of older individuals in our sample self-reported a good health in 2006 (see table ((4.16)).

Furthermore, the panel A of the table (4.16) shows that workers in poor health significantly differ from their counterparts in better health in terms of gender, age, race, marital status and education. To illustrate, the proportion of men in poor health is higher than the corresponding proportion of men in good health. Individuals in poor health are also slightly older than their

counterparts in good health. In terms of race, non-white individuals are more likely to be in poor health than to be in good health (34.1% versus 13.8%). Older workers in poor health are less likely to be married or partnered than their counterparts in good health. Finally, on average individuals in poor health are less educated than those in good health.

The HRS has other measures of physical health. Since 2008, the survey has assessed the physical performance of older individuals according to different dimensions such as lung function, grip strength, balance, walking speed, blood pressure, height, weight, and waist circumference. These assessments have been conducted through several tests. For example, the *peak expiratory air flow* (PEAF) test is used to evaluate a respondents' lung function. It is a measure of obstructive lung disease such as asthma or chronic obstructive pulmonary disease. The *handgrip* is another measure of: (i) the state of general muscle strength; and (ii) the presence of arthritis and other conditions such as degenerative arthritis or rheumatoid arthritis. This measure is considered in the medical literature to be highly predictive of functional limitations and disability at older ages (see Langa and coauthors; (2008)).

Measures of cognition and affective functioning

The HRS contains a rich set of measures of different dimensions of the cognitive and affective functioning[3] : self-reported measures of memory, objectives measures of memory, working memory, mental status, fluid reasoning, vocabulary, dementia, numeracy, quantitative reasoning, verbal reasoning and verbal fluency. For the purpose of this paper, I only considered mental health measures available for individuals aged below 65 as well as for those aged 65 and older. I also focused on mental health measures available in most of the survey waves.

Immediate word recall(IMRC): The interviewer read a list of 10 nouns (e.g. lake, army, car, etc.) to the interviewee, and asks him/her to recall as many words as possible from the list in any order. The interviewer read one of the four non-overlapping lists of 10 nouns to the interviewee. Some precautions are taken to ensure that the test outcome is not biased by the respondent. The initial list is randomly assigned to the respondent. The assignment is made longitudinally to ensure that each respondent is assigned a different set of words in each of four consecutive waves of data collection. The assignment is also made so that two respondents in the same household (e.g. spouses or partners of one another) are not assigned the same set of words in the same or adjacent waves of the survey. The interviewee received a score ranging from zero to ten depending on the number of words he/she recalled.

[3]Find more complete description of mental health measures available in the HRS in Fisher; and Herzog (2005) and McCammon and coauthors (2014)

Delayed word recall(DLRC): After approximately five minutes of asking other survey questions,[4]the respondent is asked to recall the nouns previously presented as part of the immediate recall task. He/She receives a score ranging from 0 to 10 depending on the number of nouns he/she could recall.

The Serial 7's subtraction test (SER7): The interviewer asks the individual to subtract 7 from 100, and continue subtracting 7 from each subsequent number for a total of five trials. The interviewer do not repeat the difference said by the respondent after each trial. It is up to the interviewee to remember the value from the prior subtraction. The score obtained by the respondent range from zero to five, representing the number of successes.

Backward counting starting from 20 (BWC20): Respondents are asked to count backward ten consecutive numbers beginning with the number 20. A respondent received a score of 2 if he/she succeeds at the first trial, 1 if the success is obtained at the second trial and 0 if he/she failed after two trials. This is one of the tests used to measure respondent's mental health status.

The four test are conducted through a telephone interview. The sum of the scores from each measure, which ranges from 0 to 27, gives the Telephone Interview for Cognitive Status (TICS) score. The TICS is well-documented in the medical literature, and is used to identify early signs of dementia (The more the score deceases, the more the risk of dementia increases). In this paper, I use the TICS to measure workers cognitive health in the robustness tests.

Self-rated Memory (present)(SRMP): the wording is "How would you rate your memory at the present time?Would you say it is excellent, very good, good, fair, or poor?". In this study, an individual is considered to have a poor/bad memory if he/she self-reported his/her memory status as good, fair or poor. Otherwise, the person is considered to have a good memory.

Measuring Affective Functioning: The HRS depression symptoms measure is a subset of the Center for Epidemiologic Studies Depression scale (CESD). The original measure contains 20 items designed to assess the level of depressive symptoms in epidemiologic studies of various populations. Due to interview time constraints, the HRS includes a shortened version of the CESD. This version contains only 8 symptoms since the wave 2 (1994) of the survey.

[4]For example, in 1998, questions about the depressive episodes, backward counting, and the serial 7's subtraction test were administered between the two recall tasks. The questions between immediate and delayed word recall tasks varied to some extent across survey waves.

Question wording for CESD	
Question wording	Symptoms
Now think about the past	felt depressed.
week and the feelings you	felt that everything was an effort
have experienced. Please tell	sleep was restless
me if each of the following	could not get going
was true for you much of the	felt lonely
time this past week.Much of	enjoyed life
the time during the past	felt sad
week, you....Would you say	was happy
yes or no?	

The items are summed to yield a total score ranging from 0 to 8. The items wording in the positive direction are scored in reverse. The cut-off point of 3 or more is sometimes used in the psychological literature to identify individuals with high risk of depression. This study uses the same cut-off.

2.2.3 Summary statistics

The table (4.16) presents the summary statistics (demographics, income and wealth), and labor market outcomes of older workers in the sample studied. For our analysis, I focus on individuals aged 51-69 and in the labor force as of 2006. This choice therefore excluded from the analysis older workers who have a little attachment to the labor force.

The panel B of the table shows that workers in poor health have lower earnings and wealth than their counterparts in good health.The earnings of workers in poor health is 30% lower than those of workers in good health. Older workers in poor health also have little financial assets. Their median capital income is only $48, and their median non-housing financial assets is just $2000. This implies that a job loss would be more likely to push them into early claiming of SS since they don't have enough savings or capital income to maintain their lifestyle during economic downturn.

The panel C of the table features that older workers in poor health have lower labor market outcomes than those in good health.They are twice more likely to be unemployed than healthier workers. It appears in the panels D and E that they are also overrepresented in non labor market activities during the Great Recession (retired and not working, disabled or homemaking).

2.2.4 Empirical specification

To respond to the research question, I use the following empirical specification :

$$Y_{i,s,t} = \alpha_i + \gamma UR_{i,s,t} + \beta(UR_{i,s,t} \times H_{i,0}) + \text{controls} + \varepsilon_{i,s,t} \tag{2.1}$$

Where s is the index of the state where individual i lives at baseline, $Y_{i,s,t}$ is the labor market outcome at time t, the α_i are individuals' fixed effects, $UR_{i,s,t}$ is the unemployment rate in the state s and at time t. $H_{i,0}$ is the individual's baseline health status. Controls are age and year dummies. β is the parameter of interest.

2.2.5 How to interpret β?

The implicit idea of this work is estimating the effect of being displaced on the labor force participation of older workers with heterogeneous health conditions. Let us consider the following equation:

$$Y_{i,s,t} = \beta_i \times d_{i,s,t} + u_{i,s,t} \tag{2.2}$$

Where $d_{i,s,t}$ is a dummy taking the value 1 if individual i is a displaced worker at time t, and $u_{i,t}$ is the individual heterogeneity term. It is assumed here that the effect of a displacement is heterogeneous among workers. Let us also assume that the labor supply of older workers at the extensive margin (participation) depends on their working capacity. This implies:

$$\begin{aligned} \beta_i &= \text{function}(\text{baseline working capacity}) \\ &= \gamma_0 + \gamma_1 \times (\text{baseline working capacity}) \\ &= \gamma_0 + \gamma_1 \times (\text{health status}_{i,0}) \end{aligned} \tag{2.3}$$

In equation (2.2) an individual's working capacity is measured by his/her health status at baseline. One can view it as a dummy that take 1 if the individual is in poor health condition (or low working capacity).

Finally, let us assume that the likelihood of being displaced depends on local labor market conditions. Then I obtain the following equation:

$$
\begin{aligned}
d_{i,s,t} &= \text{function}(\text{Local Economic Conditions}, \zeta_{i,s,t}) \\
&= \delta \text{UR}_{i,s,t} + \text{controls} + \zeta_{i,s,t} \\
&\text{Where} \quad \zeta_{i,s,t} \quad \text{is an error term}
\end{aligned} \tag{2.4}
$$

Pooling all equations (2.2), (2.3), and (2.4) together, the main empirical specification in equation (2.1) is obtained. I can, therefore, interpret $\beta = \gamma_1 \times \delta$ as the relative impact of being displaced during the recession, on the labor force participation of workers with limited working capacity (poor health).

2.3 Results

This section conveys the results obtained. I first present and discuss the main estimation results. Second, I analyze the interactions between the US Social Security Policy, health and local labor market conditions in shaping the labor supply decision of American older workers. Finally, I examine other sources of heterogeneity such as the sex at birth, education and marital status.

2.3.1 Main results

Table (2.1) reports Ordinary Least Squares (OLS) estimates of the effect of the state unemployment rate on the labor force participation of older workers with various self-reported health statuses in 2006. The columns (1) and (4) report estimated results when no controls are included (except the constant term). In the columns (2) and (5), several controls are included in the regressions. They are: dummies for the state of residence as of 2006, age dummies, year dummies, sex at birth and a dummy for being white/Caucasian. Finally, the columns (3) and (6) report results with individuals fixed effects, age and year dummies included as controls.

Table 2.1: OLS Estimates of the Effect of the Great Recession on Individuals Labor Supply

Dependent var. :	**Retired and not Working**					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
State Unemployment Rate (UR)	2.854***	0.541	0.498	2.324***	0.073	-0.705
	(0.531)	(0.608)	(0.643)	(0.475)	(0.650)	(0.738)
Self-reported bad health in 2006[1]				-0.034	-0.030	
				(0.024)	(0.025)	
(Self-reported bad health in 2006)*UR				1.209**	1.121**	2.870**
				(0.542)	(0.551)	(1.357)
Constant	Yes	Yes	No	Yes	Yes	No
State dummies	No	Yes	No	No	Yes	No
Year dummies	No	Yes	Yes	No	Yes	Yes
Age dummies	No	Yes	Yes	No	Yes	Yes
Oth. Demographics	No	Yes	No	No	Yes	No
Ind. fixed effects	No	No	Yes	No	No	Yes
Dependent var (mean)	0.063	0.063	0.063	0.063	0.063	0.063
# Obs.	7806	7806	7806	7806	7806	7806
R2	0.020	0.103	0.559	0.025	0.107	0.560
Sef-reported poor health in 2006 (mean)	0.452	0.452	0.452	0.452	0.452	0.452
State Unemployment Rate in 2006 (mean)	0.046	0.046	0.046	0.046	0.046	0.046
State Unemployment Rate in 2008 (mean)	0.058	0.058	0.058	0.058	0.058	0.058

*** : 1%. ** : 5%. * : 10%. Residuals are clustered at the state level. Other demographics: sex at birth and race.
[1] People are classified as being in bad health if they self-reported their health as good, fair or poor. Otherwise, they are considered to be in good health i.e. if they self-reported their health as very good or excellent. The category "good" corresponds to the median of the variable in the sample studied.

The column (6) of the table gives the main results. I observe that local unemployment rates have no direct effect on older workers' labor force participation. However, the effect is positive and strongly significant for older workers who self-reported a bad health status in 2006. This result has a remarkable implication: as the local unemployment rises, only workers with poor health conditions exit from the labor force.

The estimates also imply a quantitatively meaningful effect. $\hat{\beta}_{OLS} \times \Delta U_{s,t} = 2.870 \times (5.8 - 4.6) = 3.44\%$. Reporting that on average labor force exit in 2008 retirement rate means : $\frac{3.44}{6.30} \times 100 = 54.60\%$. However, workers who self-reported a bad health in 2006 account for 45.20% of the sample. Therefore the magnitude of the effect in the sample is: $0.546 \times 0.452 \equiv 25\%$. This is a very noteworthy effect. It is five times higher than the 5% documented in Coile and Levine (2007). Our effects are large because of the sudden and drastic unemployment shock following the Great recession. For example, the average annual state unemployment rate increased from 4.6% in 2006 to 5.8% in 2008; an increase of 26%. However, I estimated the same empirical specification, but in other time periods and found effects of similar magnitude than what documented by the aforementioned authors.

Coile and Levine (2007) also simulated that an increase of $10,000 of the Social Security wealth implies a rise of 5% of the retirement/claiming decision. Assuming some linearity, the effects documented above would be similar to an increase of $50,000 of the Social Security wealth of older workers in poor health i.e., more than 40% increase of the median of their Social Security wealth as estimated in 2004 (see the panel B of table (4.16)).

A natural question that arises is why do older workers in poor health retire? They are many potential explanations : having been unemployed for a long period i.e. the discouraged worker effect, having received a retirement package i.e. strong incentive to leave, poor anticipation about future labor market outcomes or health.

Our results cannot be explained by the discouraged worker effect. This is supported by two key arguments. First of all, the unemployment rate is very low among older workers in poor health. In 2008, only 3.8% of unhealthy older workers in our sample reported being unemployed and 5.7% in 2010. Second, I estimated the same econometric model with two different dependent variables, out of the labor force (see table (4.18)) and not working i.e. out of labor force or unemployed (see table (4.17)). If the discouraged worker effect were considerable enough in our context, I would observe a significant difference in the magnitude of the effect of interest. However, the relative impact estimated is quite similar across the two dependent variables. This suggests that most of the transitions were from job to retirement i.e. being out of the labor force.

Have older workers in poor health retired because they have anticipated that their health conditions would deteriorate sooner due to the result of the deterioration of local labor market con-

ditions? I have not found any evidence to support it. For example, I estimated the effect of state unemployment rates on the cognitive ability (measured by the log of the TICS) of older workers in our sample.[5] Overall, OLS estimates showed a negative and significant effect, but after controlling for individual fixed effects, the effect turned out to be non significant and with the opposite sign.

Altogether, the results above prove that older workers in poor health mostly retire since they were anticipating poor labor market outcomes as a consequence of the Great recession. However, I cannot rule out the possibility that they may have received incentives to quit their job such as retirement packages.[6]

The role of the Social Security retirement policy

Why have some workers in poor health retired and others have not? One key reason comes from certain important features of the US Social Security retirement policy (SSRP), which are likely to interplay with local employment shocks in a way that affects the labor supply decision of older workers.

First, it is not possible to collect Social Security old-age retirement benefits before reaching the age 62[7] . This implies a huge incentive to stay in the labor market even after a job loss for workers aged below 62. Second, even for workers older than 62 years old, claiming their benefits before the normal retirement age (NRA) is highly penalized. Collecting benefits right at age 62 implies a reduction of up to 30% of the monthly benefits a worker would be eligible to get at his/her full-retirement age (FRA). Rust and Phelan (1997) and French (2005) documented that the SSRP has a huge affect on individuals' labor force participation because it encourages them to work up to the FRA and not after. In addition to the incentives from the SSRP, older workers may remain in the labor force up to age 65 in order to qualify for Medicare. The importance of Medicare in retirement behaviour of older men is well-documented, particularly in studies that rely on structural models (French and Jones (2011), John Bound and Waidmann (2010), and Rust and Phelan (1997)).

To empirically assess how the SSRP interacts with local labor market conditions to affect the labor force participation decision of older workers in poor health, I apply an approach which is similar to what (Coile and Levine (2007)) did. The columns (4), (5) and (6) of the table (2.2) present the estimating results. There are two key observations.

[5]The results are available upon request

[6]There is a question in the HRS regarding why people left their previous job. However, the non-response rate is very high which did not allow us to examine whether a significant proportion of older workers received retirement packages.

[7]There is an exception for workers who become disabled. In this case, they will receive their full-retirement benefits regardless of their age.

First, there is no significant effect of local unemployment shocks on the labor supply of workers in poor health aged below 62. I only observe significant effects among workers aged 62 and above. I interpret this finding as an empirical evidence that the minimum age for claiming Social Security old-age benefits has a powerful discouraging effect.

Second, the estimated effects for the subgroup of workers aged between 62 and 64 is quite similar to the effects for their counterparts aged 65 and above. This suggests that people claim their Social Security old-age benefits as soon as they can. In fact, if workers aged between 62 and 64 were more likely to wait to attain their age of full benefits (which is at least 65 years old depending on the cohort), I would observe an significantly lower effect among individuals aged 62-64 compared to those aged 65 and above. This result is consistent with the findings from (Rust and Phelan (1997)).

Table 2.2: OLS Estimates of the Effect of the Great Recession on Individuals Labor Supply

Dependent var. :	**Retired and not Working**					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Unemployment Rate (UR)	2.324***	0.073	-0.705	2.324***	0.080	-0.391
	(0.475)	(0.650)	(0.738)	(0.475)	(0.650)	(0.719)
Self-reported bad health in 2006[1]	-0.034	-0.030		-0.026	-0.028	
	(0.024)	(0.025)		(0.024)	(0.024)	
(Self-reported bad health in 2006)*UR	1.209**	1.121**	2.870**			
	(0.542)	(0.551)	(1.357)			
(Self-reported bad health in 2006)*UR*(age 55-61)				0.212	0.701	1.274
				(0.514)	(0.522)	(1.184)
(Self-reported bad health in 2006)*UR*(age 62-64)				1.990***	1.908***	3.350***
				(0.480)	(0.450)	(1.163)
(Self-reported bad health in 2006)*UR*(age 65-69)				2.614***	1.574**	3.812**
				(0.644)	(0.649)	(1.588)
Constant	Yes	Yes	No	Yes	Yes	No
State dummies	No	Yes	No	No	Yes	No
Year dummies	No	Yes	Yes	No	Yes	Yes
Age dummies	No	Yes	Yes	No	Yes	Yes
Oth. Demographics	No	Yes	No	No	Yes	No
Ind. fixed effects	No	No	Yes	No	No	Yes
Dependent var (mean)	0.063	0.063	0.063	0.063	0.063	0.063
# Obs.	7806	7806	7806	7806	7806	7806
R2	0.025	0.107	0.560	0.050	0.110	0.562
Self-reported bad health in 2006 (mean)	0.452	0.452	0.452	0.452	0.452	0.452
Self-reported bad health in 2006, age 55-61 (mean)	0.437	0.437	0.437	0.437	0.437	0.437
Self-reported bad health in 2006, age 62-64 (mean)	0.474	0.474	0.474	0.474	0.474	0.474
Self-reported bad health in 2006, age 65-69 (mean)	0.470	0.470	0.470	0.470	0.470	0.470
State Unemployment Rate in 2006 (mean)	0.046	0.046	0.046	0.046	0.046	0.046
State Unemployment Rate in 2008 (mean)	0.058	0.058	0.058	0.058	0.058	0.058

*** : 1%. ** : 5%. * : 10%. Residuals are clustered at the state level. Other demographics: sex at birth and race.

[1] People are classified as being in bad health if they self-reported their health as good, fair or poor. Otherwise, they are considered to be in good health i.e. if they self-reported their health as very good or excellent.

2.3.2 *Heterogeneity*

The table (2.3) covers the OLS estimates of the relative impact, by education, marital status and gender, of local unemployment rates on the labor force participation of older workers with heterogeneous health conditions.

There are outstanding differences depending on the education background of older workers. There is no significant effect for college graduates whereas for non-college graduates, the effect is significant and quantitatively important. These differences in the labor supply response between college and non-college graduated workers are consistent with previous findings about the Great Recession. For example, Farber (2011) documented that less educated workers were more likely to lose their job than more educated ones. He also pointed out that the prospects for reemployment were very low, especially among people unemployed people for a very long period (35 weeks). Furthermore, reemployed workers suffered from serious decrease in earnings. Our results highlight that women in poor health are more likely to withdraw from the labor market after a job loss than men. This is also consistent with what documented in Farber's paper. He observed that women job losers were more likely to withdraw from the labor force than men (17% versus 10%). Finally, I observe no significant effect for not married/partnered individuals, but significant ones for both married men and women who represented 74% of the individuals in our sample.

Table 2.3: OLS Estimates of the Effect of Local Unemployment Shocks on the Labor Force Participation of Older Workers, by Education, Marital status and Gender.

Dependent var. :	**Retired and not Working**							
	By Education Attainment		**By Marital Status**		**By Gender**		**By Gender & Marital Stat.**	
	College Grad.(CG)	**Non CG**	**Not Married**	**Married**	**Men**	**Women**	**Married men**	**Married women**
State Unemployment Rate (UR)	0.189	-1.083	-1.680	-0.326	-0.894	-0.525	-0.529	0.005
	(1.897)	(0.926)	(1.761)	(0.684)	(1.063)	(0.906)	(1.160)	(1.022)
(Self-reported bad health in 2006)[1]*UR	1.521	3.085**	2.502	3.124***	2.209	3.315**	2.832**	3.320**
	(1.799)	(1.323)	(2.894)	(1.026)	(1.406)	(1.616)	(1.293)	(1.254)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ind. fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dependent var (mean)	0.050	0.068	0.066	0.061	0.060	0.065	0.061	0.062
# Obs.	2415	5391	2059	5745	3548	4258	2974	2771
R2	0.560	0.565	0.565	0.563	0.555	0.571	0.562	0.571
Self-reported bad health in 2006 (mean)	0.303	0.518	0.489	0.439	0.441	0.441	0.443	0.434

*** : 1%. ** : 5%. * : 10%. Residuals are clustered at the state level. Other demographics: sex at birth and race. [1] People are classified as being in bad health if they self-reported their health as good, fair or poor. Otherwise, they are considered to be in good health i.e. if they self-reported their health as very good or excellent.

2.4 Robustness tests

This section deals with the results from the main robustness tests conducted. More specifically, I clarify how sensible the main results documented above are to : the use of the self-reported health status, the measure of local labor market conditions and the time period considered.

Robustness to the use of more objectives measures of health

The estimates presented so far have used self-reported health status. However, even if self-reported health measures have been widely used in the economic literature as a proxy of health, they can have different kinds of measurement errors. One key issue with self-reported health measures is the *justification bias* : individuals out of the labor force are more likely to overreport poor health conditions. I controlled for this type of bias by considering: (i) the self-reported status in 2006 and (ii) employed workers. The second key issue is that self-reported health measures are likely to measure not only an individual's health condition, but also their skills. In our sample, only 13.5% of older workers who self-reported bad health were college graduates compared to 31.0% of their counterparts who self-reported a good health status (See table (4.16)). We, therefore, estimate the same econometric model, but with more objective measures of mental and physical capacity in older workers health conditions in 2006. The table (2.4) indicates the results.

First, estimation with either mental health or physical health measures is consistent with what was found using self-reported health status. The only exception in our table is the one with high blood pressure with which I found no significant labor supply responses to employment shocks.[8]

Second, the magnitude of the effect tends to increase with the severity of the disease. The magnitude of the effect is around 50% higher for individuals with severe cognitive problems than for those with moderate cognitive problems. The effect is also larger for workers with high risk of depression at baseline. Limited physical working capacity is associated with higher labor supply response. It is particularly what is observed from individuals with back problems (Lift/Carry) or with difficulties in performing daily life activities.

Finally, I do not observe any effect for individuals with problems such as high blood pressure or respiratory problems.[9] It can be explained by the fact that it is the underlying working capacity that matter, but not simply the health status. In other words, individuals with the aforementioned health conditions might still be able to have a similar productivity than comparable workers without these conditions.

[8] I found similar results with expiratory air flow and grip strength. The results are available upon request.

[9] I also found no effect by using obesity

Table 2.4: OLS Estimates of the Relative Impact of Local Unemployment Shocks on the Labor Force Participation of Older Workers with Heterogeneous Mental and Physical Health Conditions

	Health measures						
	Cognitive prob.[1]	**Sev. CP [2]**	**Depression[3]**	**Memory[4]**	**Lift/Carry[5]**	**ADL [6]**	**HBP [7]**
Dependent var. :	**Retired and not Working**						
State Unemployment Rate (UR)	-0.175	0.150	0.008	0.184	0.173	0.142	0.220
	(0.746)	(0.590)	(0.620)	(0.649)	(0.627)	(0.627)	(0.830)
(Poor health in 2006)*UR	1.923**	2.941***	3.202*	1.754**	4.411*	6.622**	2.118
	(1.739)	(0.986)	(1.589)	(0.857)	(2.238)	(2.685)	(1.710)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ind. fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dependent var (mean)	0.063	0.062	0.063	0.063	0.062	0.063	0.066
# Obs.	7806	7685	7806	7806	7774	7806	3108
R2	0.559	0.563	0.560	0.559	0.559	0.561	0560
Poor health in 2006 (mean)	0.386	0.150	0.162	0.183	0.093	0.055	0.259
State Unemployment Rate in 2006 (mean)	0.046	0.046	0.046	0.046	0.046	0.046	0.046
State Unemployment Rate in 2008 (mean)	0.058	0.058	0.058	0.058	0.058	0.058	0.058

*** : 1%. ** : 5%. * : 10%. Residuals are clustered at the state level. Other demographics: sex at birth and race.

[1] People are classified as having cognitive problems if their score for TICS is lower than the median.

[2] People are classified as having important cognitive problems if their score for TICS is lower than the quantile 25%.

[3] *Depression* is a dummy that takes 1 for individuals in high risk of having depression, a score for CESD >= 3.

[4] *Memory* is the self-reported poor memory status variable.

[5] *Lift/Carry* is a dummy that takes 1 if a worker has problem with lifting or carrying heavy objects.

[6] *ADL* is a dummy that takes 1 if the worker has difficulties performing daily life activities.

[7] *HBP* is a dummy that takes 1 if a worker has high blood pressure (systolic >=120).

Is self-reported health status simply a proxy for education or occupation?

As self-reported health status has a strong negative correlation with education and labor market outcomes, some people may argue that this health measure is simply a proxy for education or occupation. This would mean that older workers did not withdraw from the labor force because of health problems, but because they were employed in the occupations most exposed to the recession. The table (4.19) presents several estimates of the analyzed interactions between local employment shock, education and workers occupations.

The second column of the table shows that older workers in poor health who withdrew from the labor market as the result of the recession are mainly high school graduates. This is consistent with the findings from (Farber (2011)). Farber documented that, during the Great Recession in the US, the higher rate of job loss was noticed among less educated workers than among college graduates and others.[10] Moreover, Zago (2015) found that individuals who moved out from the labor force during the Great Recession were mainly middle-skilled and low-skilled workers.

In the third column of the table, I introduced in the initial econometric model interactions between local unemployment rates and workers educational achievement I observe that the effect of interest, an interaction between health status and local unemployment rates, remains significant and its magnitude does not change remarkably. This result implies that self-reported health status is not a proxy for education achievement.

In the column (4), the interactions between workers' baseline occupations and local unemployment rates are introduced in the regression. The idea is to examine whether health is simply a proxy for occupation in the sense that individuals in poor health retired because they were overrepresented in jobs more vulnerable to employment shocks. I did the same in the column (5) but used industry instead of occupation. In the column (6), I controlled both. In all the cases, the effect of interest remains significant and the magnitude of the effect does not change remarkably.

Overall, these results confirm that health as a specific and important role in the labor force participation decision of older workers exposed to a high risk of unemployment regardless of their education, occupation or industry.

2.4.1 Results using employment-to-population ratio

Unemployment rates are not always good measures of local labor market conditions for at least two reasons. First, the unemployment rate can remain high while employment rate rises if the transitions from out of the labor force to employment are important. This may happen to be the

[10] Another important finding of Farber's paper is the strong cyclicality of displacement among less educated workers.

case if discouraged workers find a job. Second, the low unemployment rate can be misleading. In fact, unemployment rate can decrease simply as long-term unemployed individuals stop active job search and become discouraged. Given the aforementioned limitations of the unemployment rate as a measure of the local labor market conditions, I used the employment-to-population ratio instead. The table (2.5) shows that the results are very similar to what I found using the unemployment rate.

Table 2.5: OLS Estimates of the Effect of Local Employment Shocks on the Labor Force Participation of Older Workers.

Dependent var. :	**Retired and not Working**					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
1-Employment-to-Pop. ratio (EPR)	0.334**	0.762	0.822	0.201	0.652	-0.837
	(0.146)	(0.510)	(0.703)	(0.149)	(0.513)	(1.060)
Self-reported bad health in 2006				-0.067	-0.067	
				(0.051)	(0.048)	
(Self-reported bad health in 2006)*EPR				0.255*	0.257*	3.755***
				(0.141)	(0.131)	(1.246)
Constant	Yes	Yes	No	Yes	Yes	No
State dummies	No	Yes	No	No	Yes	No
Year dummies	No	Yes	Yes	No	Yes	Yes
Age dummies	No	Yes	Yes	No	Yes	Yes
Oth. Demographics	No	Yes	No	No	Yes	No
Ind. fixed effects	No	No	Yes	No	No	Yes
Dependent var (mean)	0.063	0.063	0.063	0.063	0.063	0.063
# Obs.	7806	7806	7806	7806	7806	7806
R2	0.002	0.103	0.559	0.006	0.107	0.562
Self-reported bad health in 2006 (mean)	0.495	0.495	0.495	0.495	0.495	0.495

*** : 1%. ** : 5%. * : 10%. Residuals are clustered at the state level. Other demographics: sex at birth and race.
[1] People are classified as being in bad health if they self-reported their health as good, fair or poor. Otherwise, they are considered to be in good health i.e. if they self-reported their health as very good or excellent.

2.4.2 Robustness to the time period considered

I have documented so far that local unemployment hikes resulting from the Great Recession declined the labor force participation of older workers in poor health conditions. Is this result confined to the Great Recession, or rather a more general trend? To answer this question, I estimated the same empirical specification (equation (2.1)) in different periods before the Great Recession. The table (2.6) summarizes the results obtained.

The first column presents the "benchmark" results i.e. for the period 2006-2008. In the second column, the study period is extended to cover the whole Great Recession. The effects of interest are qualitatively similar, but quantitatively smaller. This smaller effect reflects the fact that most of the labor force exit of older workers in poor health arose at the beginning of the recession. The

same pattern is also observable during the recession of 2001 (variation between 2000 and 2002 in our data), but with effects of lower magnitude. The estimated effects over the periods 2004-2006, 2002-2004 and 1998-2000 are negative. However, the unemployment rate declined during these time periods. For example, the state average unemployment rate felt from 4.4% in 1998 to 3.9% in 2000. Taking this into account, I still observe a positive correlation between local unemployment rates and the labor force exit of older workers in poor health. I can, therefore, conclude that the patterns observed during the Great Recession is a trend which seems to be more general over various time periods.

Table 2.6: OLS Estimates of the Effect of Local Unemployment Shocks on the Labor Force Participation of Older Workers, by Time Period

Dependent var. :	**Retired and not Working**					
	2006-2008	**2006-2010**	**2004-2006**	**2002-2004**	**2000-2002**	**1998-2000**
State Unemployment Rate (UR)	-0.705	-0.598	2.212**	1.129	-1.785	0.553
	(0.738)	(0.560)	(0.901)	(2.273)	(1.652)	(2.151)
(Self-reported bad health at baseline)*UR	2.870**	1.820***	-2.919**	-2.026	2.263***	-2.605
	(1.357)	(0.362)	(0.907)	(2.436)	(0.750)	(1.700)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes
Ind. fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Dependent var (mean)	0.063	0.122	0.089	0.080	0.081	0.074
# Obs.	7806	11389	6009	7241	8601	9445
R2	0.560	0.573	0.569	0.570	0.575	0.565
Self-reported bad health at baseline (mean)	0.452	0.449	0.451	0.445	0.426	0.483
State Unemployment Rate at baseline (mean)	0.046	0.046	0.053	0.057	0.039	0.044
State Unemployment Rate at year 2 (mean)	0.058	0.096	0.046	0.038	0.057	0.039

*** : 1%. ** : 5%. * : 10%. Residuals are clustered at the state level. Other demographics: sex at birth and race.

[1] People are classified as being in bad health if they self-reported their health as good, fair or poor. Otherwise, they are considered to be in good health i.e. if they self-reported their health as very good or excellent.

2.5 Conclusion

Are people in good health more likely to stay in the labor market when unemployment rises? This paper has used microdata from the Health and Retirement Study (HRS) and states unemployment statistics from the US Bureau of Labor Statistics (BLS) to empirically estimate the relative impact of unemployment shocks on the labor force participation of older workers with various health conditions during the Great Recession.

This paper documents that most of the labor force exit of older individuals, in the states of USA which experienced higher unemployment hikes, was due to poor health conditions before the recession. The observed impact is much larger than what was documented in previous studies. The results are also consistent across multiple health measures such as self-reported health status, cognitive ability, depression risk and ability to perform daily activities.

Nonetheless, the results of this study cannot be interpreted in a causal sense for at least two reasons. First, the documented results are likely to reflect differences in work abilities among workers with various health outcomes because health is well-documented to be correlated with different measures of skills. Second, the health measures used in this paper are subject to measurement errors that I did not have proper control over. Yet I have proved that the observed impact does not disappear after controlling education achievement. It implies the results of this study are not driven solely by differences in abilities among workers with different health outcomes. Moreover, as I have found similar results with self-reported health measures and more objective measures, the impact of measurement errors in this study might be limited.

This paper contributes to the extensive literature on the determinants of labor market outcomes at older ages. As the US as well as other OECD countries are implementing several incentives to encourage labor force participation among older individuals, the results from this work raise the key role of older workers' health in their responses to these policies, especially during economic downturns.

Chapter 3

The Effect of Ethnic Clustering on Immigrants Earnings and Housing Conditions

Disclaimer

Preliminary incomplete. Please Do Not Cite or Circulate without the clear consent of Canada Mortgage and Housing Corporation (CMHC). The research in this chapter has been done in collaboration with two colleagues at CMHC : Musah Khalid and Judy Zhong.

3.1 Introduction

The goal of this study is to answer the following questions: what is the effect of living in an ethnic enclave (a neighborhood where the majority of the residents belongs to a visible minority group) on the earnings of recent immigrants in Canada? What are the implications on their ability to afford housing that meets their needs? Answering those questions will reveal whether and how, in the context of the National Housing Strategy (NHS), the ethnic composition of the neighborhood should be taken into account, either when it comes to decisions on where to build new affordable housing units, or the allocation, to eligible households, of affordable housing renewed with NHS' funds[1] .

[1]The latter is not actually part of existing NHS policies. But this study will examine whether it is an appropriate policy recommendation regarding the results that will be obtained.

The NHS is a $55 billion 10-year program enacted by the federal government of Canada in November 2017 with the main goal of reducing chronic homelessness by half and reducing housing need for up to 530,000 households.[2] To achieve this ultimate goal, the NHS is designed to promote and fund programs and projects that aim to provide affordable housing solutions to households in need. Newcomers (including refugees)[3] are one of the NHS priority populations who experience housing need (and related vulnerabilities).[4] The NHS was also designed with the assumption that better housing conditions are important means of improving social and economic inclusion of the targeted populations. To this end, several efforts aimed at promoting mix-income housing and improving proximity to jobs and local amenities (e.g.: proximity to health care facilities) have been initiated .[5] Understanding the role of the ethnic composition of the neighborhood on newcomers' labor market and housing outcomes also contributes to these efforts.

Why should policy-makers care about the ethnic composition of neighborhoods in the design of these policies? They should consider it for two main reasons. Firstly, there are strong theoretical and empirical support of the role of social interactions on individuals' labor market outcomes through the referral effect [6] which may have important implications for housing affordability given the direct link between household earnings and consumption. Calvó-Armengol and Jackson (2004) developed a model where individuals received information on job opportunities through their social network. One key result is the positive correlation between the size of the network and the probability of finding a job if unemployed or a better job if employed. They also show that the quality of networks can explain labor force participation differences between individuals from different races. The theoretical predictions of this model have motivated many empirical investigations since it was proposed. Patacchini and Zenou (2012) using micro-data from UK find a strong positive correla-

[2]There are different definition of housing needs. One concept often uses in Canada is core housing need. Find more details in the appendix (4.3.1).

[3]"Immigrants or refugees who have been in Canada for a short time, usually less than 5 years". See of Canada (2017).

[4]In 2011, recent immigrants households were more likely to be in core housing need (29.6%) whereas the incidence of core housing need was 12.5% for all households in Canada.of Canada (2011)

[5]To this end, CMHC and Statistics Canada have joined their forces on several initiatives to fill key data gaps. These concerted efforts have led to the release of: (i) the first nation-wide Proximity Measures Database; a database informing on how close Canadians leave to selected services (including transit stops) and jobs; (ii) the Income Divergence Index (D-index) on a mapping visualization tool. This indicator informs on the extent to which Canadians from urban areas live in mixed-income neighborhoods.

[6]There are many mechanisms that can be used to explain correlation between individuals socioeconomic outcomes and some characteristics of their neighborhood. Ham et al. (2012) group the mechanisms identified in the literature in four main categories : social-interactive mechanisms (e.g.: social networks, peer effects, role models, etc.), environmental mechanism (e.g.: exposure to violence, physical surroundings, toxic exposure), geographical mechanism (spatial mismatch, access to public services), and institutional mechanism (e.g.: stigmatization, local institutional resources, local market actors). This paper focus only on the role of social networks.

tion between the percentage of a given ethnic group living nearby and the probability of finding a job through social contacts. They also find that the effect decays rapidly with distance. Bayer et al. (2008) use US census data to show that individuals residing in the same city block are more likely to work together than those in nearby blocks. They also find that the estimated referral effect is stronger when individuals are similar in socio-demographic characteristics. Secondly, ethnic clustering is a well established characteristics of immigration in Canada (see Zucchi (2007)). The phenomenon is also quantitatively important. For example, Hiebert (2009) using data from the 2006 census, shows that 26.2% of individuals from a racialized community in Toronto, live in a minority enclave (a census tract where members from racialized communities represent at least 70% of the population).

This paper analyses administrative data on permanent residents who landed in Montreal, Vancouver and Toronto in 2001; for which earnings are observed up to 2016. We observed that immigrants who landed in neighborhoods with more than 50% of the population self-identifying as visible minority groups (ethnic enclaves) earned less, on average, than their counterparts who landed in non-ethnic enclave neighborhoods. The earnings gap between the two groups also increased over time.

We also test and use the existence of the referral effect to establish the link between immigrants' earnings and their exposure to neighbors from the same country of birth. Overall, we found that, for immigrants who landed in ethnic enclaves, an increase of the number of their neighbors having the same country of birth, is associated with a decrease their earnings on average. However, we also observed important heterogeneities across CMAs, sex at birth, age and education. For example, we observed no effect of ethnic clustering on the earnings of immigrants who landed in Montreal, a positive effect for those in Vancouver and a negative effect in Toronto. Furthermore, ethnic clustering has a negative effect on women's earnings but no effect for men.

This paper also contributes to the literature on immigrant-Canadian-born wage gap in Canada by pinpointing an additional mechanism based on the ethnic-related social interactions at the neighborhood level. A number of studies advocate that immigrant-Canadian-born wage gap in Canada is problematic. Oreopoulos (2011) highlighted the role of employers' discrimination against immigrant skills in English to explain why skilled workers with foreign-sounding names perform less well in the Greater Toronto Area (GTA) labor market than their Canadian-born counterparts. Frenette and Morissette (2003) documented the difficulty of immigrants to have their foreign labor market experience and education recognized in the Canadian labor market to explain wage-gap between recent cohorts of immigrants and Canadian-born. More recently, analysis of the Canadian 2016 census suggested that the wage gap is tremendously large in particular when comparing im-

migrants not speaking English at home with Canadian-born.[7] However, none of these results can completely explain why immigrants are paid less than comparable Canadian-born.

The paper is organized as follow. Section (3.2) presents data and earning trends observed. The section (3.3) describes the empirical strategy used in this study. The Section (3.4) presents the results, section(3.5) shows the robustness checks and section(3.6) gives the conclusion.

3.2 Data

This study uses two sources of data : (i) the Census; and (ii) the Canadian Employer-Employee Dynamics Database (CEEDD).

3.2.1 Census data

Data from the Census 2006 have been used in order to classify each census tract[8] from Montreal, Vancouver and Toronto Census Metropolitan Areas (CMAs) as an ethnic enclave or not.

For each census tract, we computed the total population and the population of individuals belonging to a visible minority group[9] as defined in the Census. An ethnic enclave is then defined as a census tract with more than 50% of its population composed by individuals identifying as visible minorities. Which this cut-off of 50%, ethnic enclaves in this study encompass the enclaves of type III -VI as defined by Hiebert (2009) (Pages 9-10). This cut-off also allow us to have a higher number of ethnic enclaves than what we would have obtained with a more restrictive definition.

There are notable differences between the three CMAs in terms of both the incidence of ethnic enclaves and their relative importance.

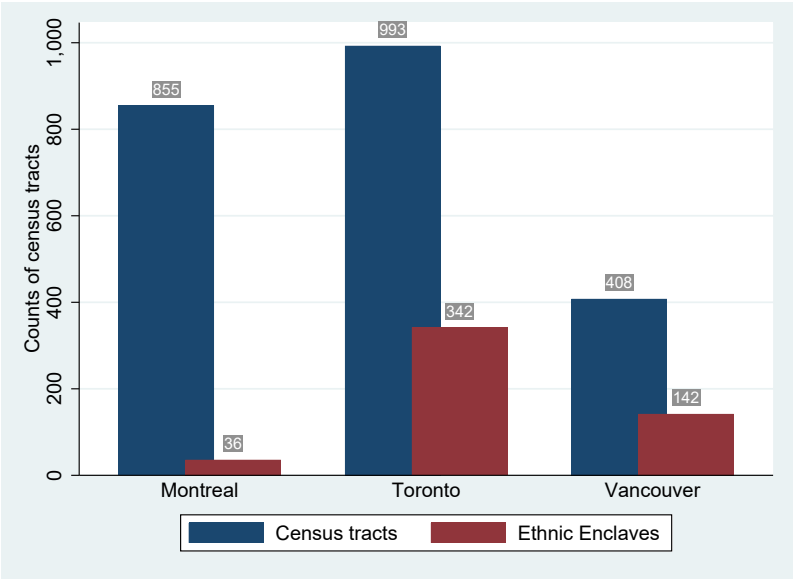
[7] https://www.macleans.ca/news/canada/new-figures-show-just-how-big-canadas-immigrant-wage-gap-is/

[8] "Census tracts (CTs) are small, relatively stable geographic areas that usually have a population between 2,500 and 8,000 persons. They are located in census metropolitan areas and in census agglomerations that had a core population of 50,000 or more in the previous census.

A committee of local specialists (for example, planners, health and social workers, and educators) initially delineates census tracts in conjunction with Statistics Canada. Once a census metropolitan area (CMA) or census agglomeration (CA) has been subdivided into census tracts, the census tracts are maintained even if the core population subsequently declines below 50,000." https://www150.statcan.gc.ca/n1/pub/92-195-x/2011001/geo/ct-sr/def-eng.htm

[9] "Visible minority population refers to the visible minority group to which the respondent belongs. The Employment Equity Act defines visible minorities as persons, other than Aboriginal peoples, who are non-Caucasian in race or non-white in colour. Categories in the visible minority population variable include Chinese, South Asian, Black, Filipino, Latin American, Southeast Asian, Arab, West Asian, Korean, Japanese, Visible minority, n.i.e. (n.i.e. means not included elsewhere), Multiple visible minority, and Not a visible minority". https://www12.statcan.gc.ca/census-recensement/2006/ref/rp-guides/visible_minority-minorites_visibles-eng.cfm

Figure 3.1: Incidence of ethnic clustering in Montreal, Vancouver and Toronto CMAs, 2006

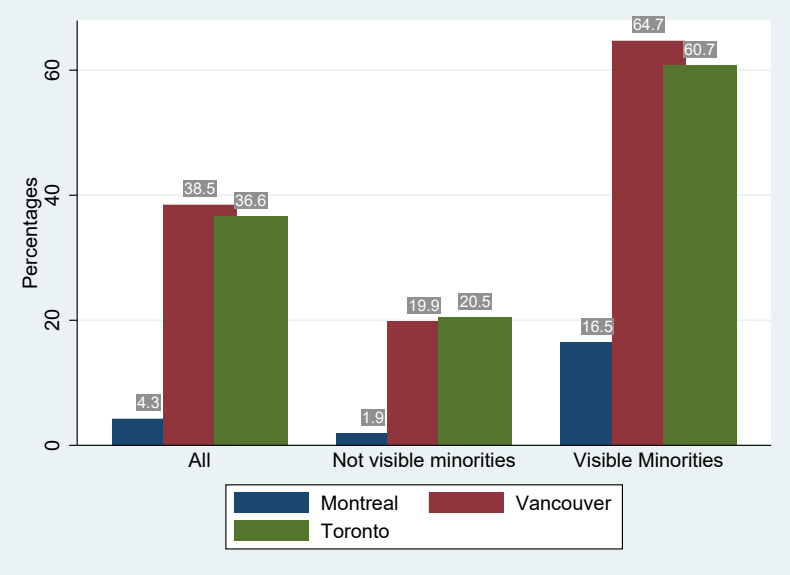


Source: CMHC, computed from Statistics Canada Census 2006.

The figure (3.1) shows a quite small incidence of ethnic clustering in Montreal. Among the 855 census tracts identified in the data, only 36 of them had more than 50% of their population belonging to visible minority groups. Therefore in Montreal, only 4.2% of neighborhoods can be considered as ethnic enclaves. However, in Vancouver and Toronto the picture is quite different. Toronto had much more ethnic enclaves than Vancouver; 342 compared to 142. Nonetheless, the proportion of this type of neighborhoods is slightly higher in Vancouver. Ethnic enclaves represented 34.8% of neighborhoods in Vancouver whereas the corresponding proportion is 34.4% for Toronto. These differences are also reflected in the distribution, within each CMA, of the population across ethnic enclave neighborhoods and neighborhoods that are not ethnic clusters.

It appears from figure (3.2) that Montreal differs significantly from Vancouver and Toronto with respect to the distribution of the population across census tracts. In Montreal only 4.3% of the population lived in ethnic enclaves in 2006. Furthermore, the proportion of individuals not belonging to visible minority groups who lived in an ethnic enclave was close to eight times lower (2%) than the proportion of people from visible minorities who lived in this type of neighborhood. On the contrary, in Vancouver and Toronto it is a different picture : the proportion of individuals not belonging to visible minorities who lived in an ethnic enclave is only three times lower than the proportion of people from visible minorities who lived in this type of neighborhoods. Besides, in Vancouver and Toronto, a very high proportion of individuals from visible minorities lived in

Figure 3.2: Population living in ethnic enclaves, by visible minority status, 2006



Source: CMHC, computed from Statistics Canada Census 2006.

ethnic clusters; respectively 64.7% and 60.7%.

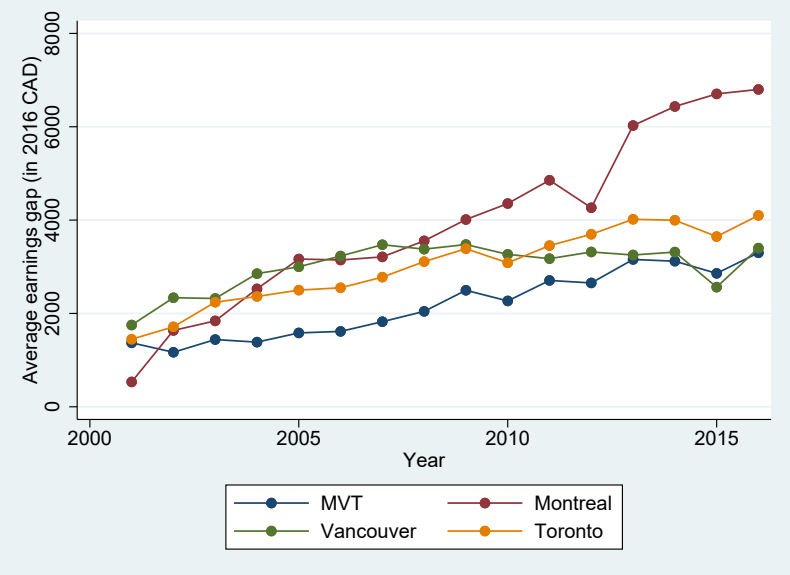
Altogether, the figures (3.1) and (3.2) confirm that ethnic clustering is an important phenomenon in the three largest CMAs of Canada.

3.2.2 CEED data

The CEED is an employer-employee linkage environment updated and maintained by Statistics Canada. The CEED links a total of nine databases : T1 individual file, T1 family file, T4 file (per employment), T4 and record of employment file (T4ROE), unincorporated business and owner file, incorporated business owner file, National Accounts Longitudinal Microdata File (NALMF), an Export file (TEC) and the Longitudinal immigrants and temporary residents file (IMDB). For this study, we linked the IMDB to tax files using individuals' unique identifier. This allow us to have altogether, landing variables (e.g: sex at birth, country at birth, knowledge of official languages, etc.), geographical identifiers for the CMA and dissemination area of residence each year as well as income variables retrieved from tax files.

The sample studied is constituted by permanent residents who landed in Canada in 2001, who were aged between 20 to 49, and who landed in Montreal, Vancouver and Toronto CMAs. They are observed up to 2016.

Figure 3.3: Time series of average earning gaps, by the CMA of landing.



All the results are significant at 1%. Find more details in table(4.20). MVT ≡ Montreal, Vancouver and Toronto.

Source: CMHC, computed from Statistics Canada Canadian Employer-Employee Dynamics Database, vintage 2017.

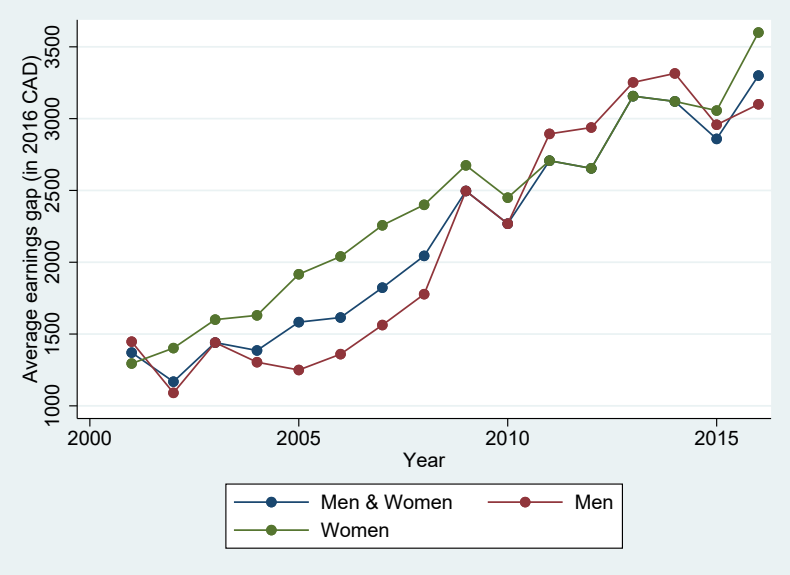
The figure (3.3) documents the annual average earning differences between newcomers who landed in non-ethnic enclave neighborhoods and those who landed in ethnic enclaves (average earnings gap thereafter) as of 2001. Three main findings can be highlighted : (i) permanent residents who landed in ethnic enclaves earned less, on average, than their counterparts who landed in non-ethnic enclave neighborhoods; (ii) the observed average earnings gap tend to increase over time; (iii) the three CMAs studied display similar trends even though the average earnings gap increased faster in Montreal than in the other CMAs.

The table (4.20) gives more details on data from the figure (3.3). In 2001, the average earnings of newcomers living in ethnic enclaves was respectively $6400 in the three CMAs (also called MVT), $7007 in Montreal, $6398 in Vancouver and $6474 in Toronto. Their counterparts who landed in non-ethnic enclaves earned on average, an additional $1371 in MVT, $533 in Montreal, $1752 in Vancouver and $1447 in Toronto. Fifteen years later, i.e. in 2016, the average earnings of newcomers living in ethnic enclaves had increased significantly to become respectively $48,900 in MVT, $43,200 in Montreal, $45,900 in Vancouver and $50,300 in Toronto. However, the earning penalty had also increased. In 2016, newcomers who landed in non-ethnic enclaves in 2001 earned on average, an additional $3300 in MVT, $6800 in Montreal, $3400 in Vancouver and $4100 in Toronto. Clearly, Montreal display the largest average earnings gaps between individuals who landed in non-ethnic enclaves and those who landed in ethnic enclaves.

The figure (3.3) presents the evolution of the average earning gaps between permanent residents who landed in non-ethnic enclave neighborhoods and those who land in ethnic enclaves as of 2001 by the sex at birth. Either men or women who landed in ethnic enclaves earned less on average than their counterparts who landed in non ethnic enclaves. Furthermore, the earning penalty increased steadily overtime. Even though, the earning penalty was higher for women within the first ten years in Canada, the gap tended to become closer (in level) to that of men after.

The table (4.21) shows the estimates used to plot figure (3.4). In 2001, the average earnings of men and women living in ethnic enclaves was respectively, $7007 and $5788. This was lower by $1447 and $1295. In 2016, men who landed in ethnic enclaves earned $54,500 on average and women $43,200.

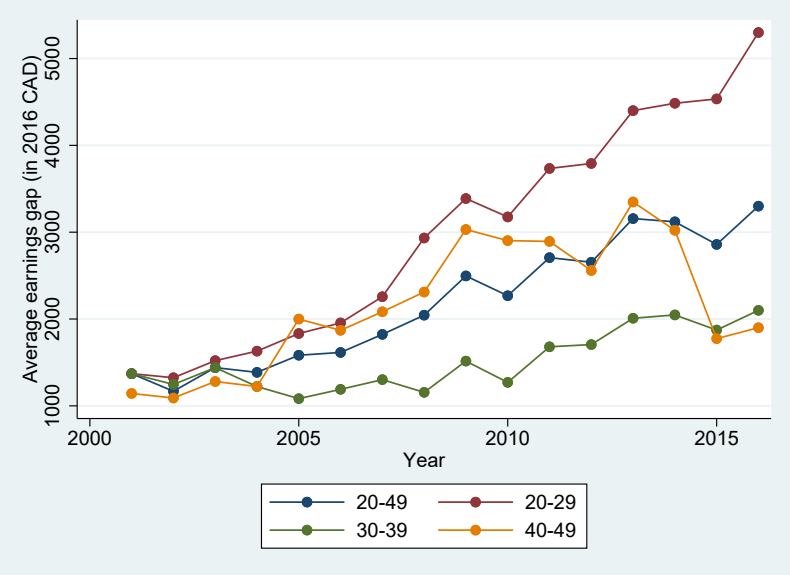
Figure 3.4: Time series of average earning gaps, by newcomers sex at birth .



All the results are significant at 1%. Find more details in table(4.21).

Source: CMHC, computed from Statistics Canada Canadian Employer-Employee Dynamics Database, vintage 2017.

Figure 3.5: Time series of the average earnings gap, by newcomers age group



All the results are significant at 1%. Find more details in table(4.22).

Source: CMHC, computed from Statistics Canada Canadian Employer-Employee Dynamics Database, vintage 2017.

The figure (3.5) shows the dynamics of the average earnings gaps between permanent residents who landed in non-ethnic enclave neighborhoods and those who landed in ethnic enclaves by the age group of the newcomer as of 2001. Individuals who landed as young adults (age 20-29) had the largest differences in average earnings and the gap displayed the fastest growth.

The results presented above document strong and significant differences between the earnings of newcomers who landed in ethnic enclaves and the earnings of their counterparts who landed in non-ethnic enclave neighborhoods. The observed differences (in level) tend to increase over time. Furthermore, similar trends are observed regardless of individuals sex at birth, CMA of residence and age group at landing. However, because newcomers might have self-selected themselves into neighborhoods with different ethnic compositions, the observed earnings gaps can be driven by differences in skills between individuals and not only by the ethnic composition of the neighborhoods. It is therefore, not possible to interpret the results above in a causal sense. The next sections, we will address this issue by trying to isolate the effect of country of birth based social networks on newcomers' earnings.

3.3 Empirical Strategy

The main challenge to answering the research question is how to correct for self-selection of immigrants into ethnic enclaves. This self selection implies that OLS estimates are likely to be biased because the decision on where to live is likely to be correlated with time-varying unobserved abilities to succeed in the Canadian labor market.[10]

Furthermore, whether OLS estimates are upward or downward biased depends on the specific context. To illustrate, immigrants with high unobserved skills (not exclusively college graduates' immigrants) could choose to live in ethnic neighborhood because being considered as role models or taking leadership roles in their community provides them an additional satisfaction. In this case, OLS estimates will produce an upward biased positive correlation between the number of peers in the neighborhood and the earnings. On the contrary, immigrants with unobserved low skills could choose to look for housing within enclaves to enhance their economic opportunities through their peers. In this second case, OLS estimates will produce a downward biased negative correlation. Clearly, any correlation between neighborhood's ethnic composition and individuals' earnings cannot be interpreted directly as causal.

To overcome this issue, the empirical strategy follows the idea developed in Bayer et al. (2008)

[10] example of some unobserved skills: motivation, ability to learn English or French, ability to adapt to local customs.

and used thereafter by Hellerstein et al. (2011), Hellerstein et al. (2014), and Schmutte (2015). The identification relies on the hypothesis that even if individuals self-select themselves into ethnic enclaves, they have no control on where they will find the housing they are looking for within the enclave because the housing market is very tight in more granular levels of geography. For example a new Chinese immigrant can definitely choose to live in Scarborough in Toronto. But whether he will finally live either on Shepard Avenue or on Midland Avenue depends on the availability of the type of housing he is looking for; something on which he has less or even no control over. Thus, the number of immigrants from China in his building or even in his immediate neighborhood is likely to be beyond any control on his part.

This paper uses the following empirical strategy in order to estimate the effects of ethnic clustering on immigrants' earnings.

3.3.1 Estimating the effects of ethnic clustering

To identify the contemporaneous effects of social networks on immigrant's earnings we will use the following empirical specification :

$$y_{i,d,c,t} = \alpha_i + \delta_t + \beta_0 Z_{i,c,t} + \beta_1 Z_{i,c,t} \times S_{d,t} + \beta_2 S_{d,t} + X'_{i,c,t}\gamma + \varepsilon_{i,d,c,t} \tag{3.1}$$

Where i is the index of the individual, d is the index of the dissemination area (DA), c is the index of the census tract (CT) and t is the index of the year. $y_{i,c,t}$ is the individual's i log-wage and $Z_{i,c,t}$ is a dummy equal to 1 if the individual live in an ethnic enclave and zero otherwise. α_i is the individual fixed-effect. δ_t are year dummies; they control for national trends in the outcomes of workers. X is a vector of controls that include dummies for education attainment and experience up to order three. $S_{d,t}$ is the number of individuals (in log) having the same country of birth than i and who live in the same DA.

β_0 is supposed to measure the "average" earnings differences between individuals living in enclave and those living outside after controlling for potential confounders. β_1 is the parameter of interest. It measures the average earning differences between individuals living in the same ethnic enclave that is due to their differential exposure to their peers. $\beta_{0,OLS}$ is likely to be biased because immigrants workers from visible minority might self-select themselves into ethnic neighborhoods based on time-varying characteristics that have a direct effect on their performances in the labor market but that are unobserved by the researcher. However, $\beta_{1,OLS}$ is an unbiased estimate of the true β_1 under the identification hypothesis which imply that $Cov(S_{d,t}, \varepsilon_{i,d,c,t}) = 0$. We test that

hypothesis in this paper.

3.4 Results

This section first establishes the existence of the "referall effect" in the context of this study, and then present estimates of the effect of ethnic clustering on recent immigrants' earnings. We also analyse heterogeneities by CMA of residence, sex at birth, age and years of schooling.

3.4.1 Establishing the existence of Social Interactions and the Referral Effect

To validate the existence of social interactions and the referral effect we estimate the following empirical specification using OLS:

$$naics_{j,d,t} = \alpha_j + \alpha_d + \delta_t + \alpha_0 S_{j,d,t} + X'_{j,d,t}\gamma + \eta_{j,d,t} \tag{3.2}$$

Where j is the index of country of birth, α_j is the country of birth fixed effect and α_d is the dissemination area fixed effect. $naics_{j,d,t}$ is the number of NAICS having at least two people from the same country, living in the same dissemination area working together. We look at the NAICS instead of at a more disaggregated level, such as the firm level, in order to insure a sufficient number of observations. α_0 is the parameter of interest. Under the hypothesis of social interactions and referral effect, we should observe that $\hat{\alpha}_{0,OLS} > 0$. Which means that, higher is the number of individuals from a given country in a neighborhood, higher is the probability that those individuals will find jobs through referral.

Table 3.1: OLS Estimates of the referall effect

Dependent var. :	**Log(Number of 4-digit NAICS with more than two people having the same country of birth)**
Log(Number of neighbours having the same country of birth)	+ n/a ***
	(n/a)
	Controls
Year dummies	Yes
Dissemination area dummies	Yes
Cluster var. dummies	Yes
Country of birth dummies	Yes
Constant term	Yes
Dependent var (mean)	n/a
Log(Number of neighbours having the same country of birth) (mean)	n/a
# Obs.	n/a

***:1%. **:5%. * : 10%. n/a ≡ figures cannot be disclosed for the moment; the vetting process by Statistics Canada is ongoing. Residuals are clustered by CMA and year.

Source: CMHC, computed from Statistics Canada Canadian Employer-Employee Dynamics Database, vintage 2017.

The OLS estimate in the table (3.1) above is positive and strongly significant; which suggest the existence of the "referall effect" : immigrants are very likely to find jobs through their social network (here composed by neighbours who have the same country of birth).

3.4.2 Main results

The table (3.2) presents OLS estimates of the effect of ethnic clustering on the earnings of permanent residents who landed in Canada in 2001. The first column gives the results for the whole sample constituted by immigrants who lived in Montreal, Vancouver and Toronto at the year of landing. The following columns present the estimates for each of the three CMAs. The key parameter of interest is the effect of the second variable i.e the interaction between ethnic enclave dummy and the number of country-mates (in log).

Table 3.2: OLS Estimates of the Effect of ethnic clustering on the earnings of recent immigrants in Canada

Dependent var. :	**Log(Annual earnings)**			
	MVT	**Montreal**	**Vancouver**	**Toronto**
Log(Number of neighbours having the same country of birth) $(S_{d,t})$	- n/a	+ n/a	- n/a ***	- n/a
	(n/a)	(n/a)	(n/a)	(n/a)
I{Dummy = 1 if lived in an ethnic enclave in 2001} $\times$ $(S_{d,t})$	- n/a **	- n/a	+ n/a ***	- n/a ***
	(n/a)	(n/a)	(n/a)	(n/a)
	Controls			
Year dummies	Yes	Yes	Yes	Yes
Dissemination area dummies	Yes	Yes	Yes	Yes
Cluster var. dummies	Yes	Yes	Yes	Yes
Experience (= age - years of schooling - 6)	Yes	Yes	Yes	Yes
Experience^2	Yes	Yes	Yes	Yes
Experience^3	Yes	Yes	Yes	Yes
Constant term	Yes	Yes	Yes	Yes
Dependent var (mean)	n/a	n/a	n/a	n/a
Log(Number of neighbours having the same country of birth) (mean)	n/a	n/a	n/a	n/a
# Obs.	n/a	n/a	n/a	n/a

***:1%. **:5%. * : 10%. n/a $\equiv$ figures cannot be disclosed for the moment; the vetting process by Statistics Canada is ongoing. Residuals are clustered by CMA and year.

Source: CMHC, computed from Statistics Canada Canadian Employer-Employee Dynamics Database, vintage 2017.

For immigrants who lived in ethnic enclaves, higher is the number of their country-mates in their close neighborhoods, lower will be their earnings. Given that we have established the existence of the "referall effect" in this context, the observed negative effect can be attributed to a lower quality social network i.e. immigrants who lived in ethnic enclaves hold low paid jobs and therefore share with newcomers job opportunities available where they work. However, the results are not similar across CMAs.

We found no significant effect of ethnic clustering on the earnings of immigrants who landed in Montreal. In Vancouver, we observe a positive and strongly significant effect. This mean that

among recent immigrants who lived in ethnic enclaves, those who had more country-mates in their close neighborhood had better outcomes than their counterparts who had much less country-mates living closer. We also observe that there is significant and negative direct effect of the number of country-mates (in log) on immigrants' earnings. We interpret this as the negative self-selection i.e. immigrants who lived closer to country-mates have lower earning potential than those who live farther. For Toronto, the results show a clear negative effect of ethnic clustering on the earnings of recent immigrants.

Altogether, the results above suggest that the effect of ethnic clustering might be different across communities [11] as the three CMAs attract immigrants from different parts of the world and with different backgrounds. The results would depend on the labor market outcomes of previous cohorts of immigrants. In a community where the previous cohorts of immigrants had achieved better outcomes, newcomers will also achieve better outcomes compared to newcomers belonging to a communities with historical poor labor market outcomes.

3.4.3 Heterogeneity by sex at birth and age

Table (3.3) presents the OLS estimates of the effect of ethnic clustering on the earnings of permanent residents who landed in Canada in 2001 by sex at birth and age group.

The estimates show different patterns between men and women. There is no significant effect of ethnic clustering on men's earnings. However, for women the effect is significant and negative. This suggest that ethnic clustering has a much more negative impact on women labor market outcomes than on men. The results are also heterogeneous by age group (as of 2001).

We observe no significant effect of ethnic clustering on the earnings of immigrants who were aged below 40 at landing. However, for those who were aged between 40 and 49 years old at the landing year, there is a significant negative correlation between the number of country-mates and their earnings. This suggests that immigrants who rely on the country of birth based social network to find job opportunities are relatively older immigrants.

[11] A community constituted here by people having the same country of birth.

Table 3.3: OLS Estimates of the Effect of ethnic clustering on the earnings of recent immigrants in Canada; by sex at birth and age group at landing.

Dependent var. :	**Log(Annual earnings)**				
	Sex at birth		**age group in 2001**		
	Male	**Female**	**20-29**	**30-39**	**40-49**
Log(Number of neighbours having the same country of birth) ($S_{d,t}$)	- n/a *	- n/a	- n/a	- n/a ***	- n/a
	(n/a)	(n/a)	(n/a)	(n/a)	(n/a)
I{Dummy = 1 if lived in an ethnic enclave in 2001} $\times$ ($S_{d,t}$)	- n/a	- n/a ***	- n/a	+ n/a	- n/a ***
	(n/a)	(n/a)	(n/a)	(n/a)	(n/a)
	Controls				
Year dummies	Yes	Yes	Yes	Yes	Yes
Dissemination area dummies	Yes	Yes	Yes	Yes	Yes
Cluster var. dummies	Yes	Yes	Yes	Yes	Yes
Experience (= age - years of schooling - 6)	Yes	Yes	Yes	Yes	Yes
Experience^2	Yes	Yes	Yes	Yes	Yes
Experience^3	Yes	Yes	Yes	Yes	Yes
Constant term	Yes	Yes	Yes	Yes	Yes
Dependent var (mean)	n/a	n/a	n/a	n/a	n/a
Log(Number of neighbours having the same country of birth) (mean)	n/a	n/a	n/a	n/a	n/a
# Obs.	n/a	n/a	n/a	n/a	n/a

***:1%. **:5%. * : 10%. n/a $\equiv$ figures cannot be disclosed for the moment; the vetting process by Statistics Canada is ongoing. Residuals are clustered by CMA and year.

Source: CMHC, computed from Statistics Canada Canadian Employer-Employee Dynamics Database, vintage 2017.

3.4.4 Heterogeneity by years of schooling

Table (3.4) presents the estimates of the effect of ethnic clustering on recent immigrants' earning by years of schooling as of 2001. Clearly, there are important heterogeneities based on schooling outcomes of immigrants before their landing in Canada. We observe a negative and significant effect of ethnic clustering on the earnings of very low-skilled immigrants (i.e. with 9 years of schooling or less) and high skilled immigrants (13 years of schooling or more). However, for immigrants with 10-12 years of schooling, the effect of ethnic clustering is positive and significantly different from zero.

Table 3.4: OLS Estimates of the Effect of ethnic clustering on the earnings of recent immigrants in Canada; by years of schooling at landing.

Dependent var. :	**Log(Annual earnings)**		
	Years of schooling in 2001		
	0-9	**10-12**	**13 and +**
Log(Number of neighbours having the same country of birth) ($S_{d,t}$)	+ n/a **	- n/a **	- n/a
	(n/a)	(n/a)	(n/a)
I{Dummy = 1 if lived in an ethnic enclave in 2001} $\times$ ($S_{d,t}$)	- n/a ***	+ n/a **	- n/a ***
	(n/a)	(n/a)	(n/a)
	Controls		
Year dummies	Yes	Yes	Yes
Dissemination area dummies	Yes	Yes	Yes
Cluster var. dummies	Yes	Yes	Yes
Experience (= age - years of schooling - 6)	Yes	Yes	Yes
Experience^2	Yes	Yes	Yes
Experience^3	Yes	Yes	Yes
Constant term	Yes	Yes	Yes
Dependent var (mean)	n/a	n/a	n/a
Log(Number of neighbours having the same country of birth) (mean)	n/a	n/a	n/a
# Obs.	n/a	n/a	n/a

***:1%. **:5%. * : 10%. n/a $\equiv$ figures cannot be disclosed for the moment; the vetting process by Statistics Canada is ongoing. Residuals are clustered by CMA and year.

Source: CMHC, computed from Statistics Canada Canadian Employer-Employee Dynamics Database, vintage 2017.

3.5 Robustness tests

This section presents the results from the falsification tests conducted to validate the identification strategy used in this paper.

The identification strategy relies on the hypothesis that, for each immigrant, the number of close neighbors (i.e. neighbors living in the same dissemination area) who have the same country of birth is not correlated with any determinant of earnings. This means that an immigrant can choose to live in an ethnic enclave census tract, but once this choice is made, the immigrant has no control on the number of his/her close neighbours who have the same country of birth.

To validate the identification hypothesis, we estimated a modified version of equation (3.1), but with a measure of the knowledge of the official languages of Canada as the dependent variable. Host country language acquisition has been shown to be an important determinant of immigrants' earnings in many OECD countries (see for example : Bonikowska et al. (2008),Chiswick (1991),Chiswick and Miller (1991), Berman et al. (2003),Dustmann and Fabbri (2003)). Therefore, if the identification hypothesis is false, we should observe, among immigrants who lived in ethnic enclaves, a significant correlation between the number of close neighbors and their language skills.

Table (3.5) presents the OLS estimates of the correlation between the number of neighbors from the same country of birth and four dummy variables representing a measure of the knowledge

of English and/or French at year of landing. $I\{\text{Dummy = 1 if lived in an ethnic enclave in 2001}\} \times (S_{d,t})$ is the independent variable of interest. The estimates give the correlation between the number of close neighbours and a measure of the knowledge of English and/or French for immigrants who lived in ethnic enclave census tracts as of 2001. In each of the four cases, we observe no significant correlation. These results are consistent with the identification hypothesis used in this paper.

Table 3.5: OLS Estimates of the correlation between the number of neighbours from the same country of birth and the knowledge of Canada's official languages at year of landing (i.e. in 2001).

Dependent var. :	**Dummies for the knowledge of Canada's official languages**			
	English at least	**English and French**	**English only**	**French only**
Log(Number of neighbours having the same country of birth) $(S_{d,t})$	- n/a ***	- n/a	- n/a ***	- n/a
	(n/a)	(n/a)	(n/a)	(n/a)
I{Dummy = 1 if lived in an ethnic enclave in 2001} $\times$ $(S_{d,t})$	- n/a	+ n/a	- n/a	+ n/a
	(n/a)	(n/a)	(n/a)	(n/a)
	Controls			
Year dummies	Yes	Yes	Yes	Yes
Dissemination area dummies	Yes	Yes	Yes	Yes
Country of birth dummies	Yes	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes	Yes
Experience (= age - years of schooling - 6)	Yes	Yes	Yes	Yes
Experience^2	Yes	Yes	Yes	Yes
Experience^3	Yes	Yes	Yes	Yes
Constant term	Yes	Yes	Yes	Yes
Dependent var (mean)	n/a	n/a	n/a	n/a
Log(Number of neighbours having the same country of birth) (mean)	n/a	n/a	n/a	n/a
# Obs.	n/a	n/a	n/a	n/a

***:1%. **:5%. * : 10%. n/a $\equiv$ figures cannot be disclosed for the moment; the vetting process by Statistics Canada is ongoing. Residuals are clustered by the country of birth. The sample is restricted to observations as of 2001.

Source: CMHC, computed from Statistics Canada Canadian Employer-Employee Dynamics Database, vintage 2017.

3.6 Conclusion

This paper studies the effect of ethnic clustering on the earnings of permanent residents who landed in Montreal, Vancouver and Toronto in 2001.

Firstly, we documented a strong and persistent difference between the average earnings of newcomers who landed in neighborhoods with more than 50% of the population belonging to a visible minority group (also called ethnic enclaves) and their counterparts who landed in non-ethnic enclaves. Immigrants who landed in ethnic enclaves earned less on average than their counterparts who landed in non-ethnic enclave neighborhoods. Furthermore, the earnings gap between the two groups also increased over time. This observation remains across the three CMAs, sex at birth, and age groups.

Secondly, this study also documented the existence of the "referral effect" i.e. immigrants tend to use their social network (here composed by neighbors having the same country of birth) to find jobs. In fact, we observed a positive correlation between the number of immigrants from a given country in a neighborhood and the likelihood of two immigrants from the same country of birth and living in the same neighborhood to work in the same industry group.

Finally, we estimated the effect of ethnic clustering on the earnings of recent immigrants. Overall, for immigrants who lived in ethnic enclaves, we observed a strong negative correlation between the number of neighbors having the same country of birth and their earnings. However, there are important heterogeneities across CMAs, sex at birth, age group and education. Across CMAs, we observed no effect of ethnic clustering on immigrants' earnings in Montreal, a positive effect in Vancouver and a negative effect in Toronto. Furthermore, the effect of ethnic clustering is negative for women and immigrants aged 40-49 years old at landing. But we found no effect on men's earnings as well as on the earnings of younger immigrants. In terms of education, ethnic clustering has a negative effect on immigrants with either less than 10 years of schooling or 13 years of schooling or more. But for those with 10-12 years of schooling the effect is positive.

As next steps, we will improve this work in two important areas. Firstly, the measure of the social network used in this study is subject to measurement errors because immigrants who landed in Canada before 1980 are not observed in the data. We will address this issue by using an instrumental variable approach. Secondly, some important source of heterogeneities such as immigration status (e.g.: refugees, economic immigrant, etc.) and ethnicity (e.g.: Black, south asian, west asian, arabic, etc.) haven't been explored. We will also analyse heterogeneities along those lines.

Nevertheless, this work has two notable limitations. First of all, the social network considered in this paper is restricted to the network composed by immigrants from the same country of birth. However, it is not clear that this social network is the most important for immigrants' labour market

outcomes in Canada. Furthermore, we only consider here employed immigrants. However, the social network might also have an important effect on the income of immigrants who are business owners. The later two areas represent important research topics for affordable housing policy.

Conclusion

This thesis proposed three essays on two topics with important implications for public policies : (i) determinants of labor force participation at older ages ; (ii) the role of social network in the earnings of recent immigrants. Chapter 1 quantified the causal effect of housing wealth on the labor force participation decisions of older workers in US. Chapter 2 answered the question : are people in good health more likely to stay in the labor market when unemployment rises? Finally, Chapter 3 estimated the effect of ethnic clustering on the earnings of recent immigrants in Canada.

This thesis researches contribute to the literature by bringing new insights. Chapter 1 used plausible exogenous credit supply shocks to document a strong and positive causal effect of housing wealth on the labor force exit of older workers during the 1994-2004 housing boom period in the US. With US as well as other OECD countries experiencing simultaneously housing booms in many of their major city centres, a declining labor force participation of baby boomers, and an increase of the share of seniors in their populations, the results from Chapter 1 suggest that the housing boom may accelerate the financial imbalance of public retirement funds in many developed countries because of early labor market exits. Chapter 2 documented that when local unemployment rises, older workers who exit from the labor force are mainly those having poor health conditions. This result raises the importance of a good health policy for a long lasting participation of older workers in the labor market. Finally, Chapter 3 controlled for self-selection into ethnic enclaves and documented, overall, a negative effect of ethnic clustering on the earnings of recent immigrant in Canada. This work emphasizes the importance of a good access to job opportunities for a better economic integration of immigrants.

This thesis opens several areas for future researches. Chapter 1 raised at least two important outstanding questions. The number of months or years lost as the result of a given increase of individuals' housing wealth is not known. The timing between the start of a housing boom in a region and the decision to stop working by older homeowners is not also clear. Answering these questions is critical to compare the relative importance of housing wealth, in the labor supply decisions of older workers, to other determinants of retirement such as health and health insurance,

to name a few. Chapter 2 documented empirically the important role of health in the labor supply decisions of older workers, especially during economic downturns. However, the estimates do not have a causal interpretation. This represents a key topic for future research. Finally, the research in Chapter 3 considered only the social network composed by immigrants from the same country of birth and living nearby. However, whether this social network is the most important for immigrants labor market outcomes in Canada is not clear. Furthermore, this work haven not analyzed the role of social networks in the outcomes of immigrants who are business owners. These outstanding questions represent two important areas for future researches.

Bibliography

AMIOR, M. AND A. MANNING (2015): "The Persistence of Local Joblessness," *Centre for Economic Performance, LSE CEP Discussion Papers*.

AUTHOR, D. H., D. DORN, AND G. H. HANSON (2013): "The China Syndrome : Local Labor Markets Effects of Import Competition in the United States," *American Economic Review*, 103.

BAYER, P., S. L. ROSS, AND G. TOPA (2008): "Place of Work and Place of Residence : Informal Hiring Networks and Labor Market Outcomes," *Journal of Political Economy*, 116.

BENITEZ-SILVA, H., S. EREN, F. HEILAND, AND S. JIMENEZ-MARTIN (2015): "How well do individuals predict the selling prices of their homes?" *Journal of Housing Economics*, 12–25.

BERMAN, E., K. LANG, AND E. SINIVER (2003): "Language-skill complementarity: returns to immigrant language acquisition," *Labour Economics*, 265–290.

BLAU, D. M. (1998): "Labor Force Dynamics of Older Married Couples," *Journal of Labor Economics*, 16, 595–629.

BONIKOWSKA, A., D. A. GREEN, AND W. C. RIDDELL (2008): "Literacy and the Labour Market: Cognitive Skillsand Immigrant Earnings," *Statistics Canada*, Catalogue no. 89-552-M.

CALVÓ-ARMENGOL, A. AND M. O. JACKSON (2004): "The Effects of Social Networks on Employment and Inequality," *American Economic Review*, 94, 426–454.

CAMERON, A. C., J. G. GELBACH, AND D. L. MILLER (2011): "Robust Inference with Multi-Way Clustering," *Journal Business and Economic Statistics*, 29, 238–49.

CAMERON, A. C. AND D. L. MILLER (2014): "A Practitioner's Guide to Cluster-Robust Inference," *The Journal of Human Ressources*, 50.

CAMPBELL, J. Y. AND J. F. COCO (2007): "How do house prices affect consumption? Evidence from micro data," *Journal of Monetary Economics*, 54, 591–621.

CASE, K. E., R. J. SHILLER, AND J. M. QUIGLEY (2001): "Comparing Wealth Effects : The Stock Market Versus the Housing Market," *NBER Working Paper*.

CHETTY, R., N. HENDREN, P. KLINE, AND E. SAEZ (2014): "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States," *Quarterly Journal of Economics*, 194, 1553–1623.

CHISWICK, B. R. (1991): "Speaking, Reading, and Earnings among Low-skilled Immigrants," *Journal of Labor Economics*, 2, 897–921.

CHISWICK, B. R. AND P. W. MILLER (1991): "The Endogeneity between Language and Earnings:International Analyses," *Journal of Labor Economics*, 2, 246–288.

COILE, C. C. AND P. B. LEVINE (2007): "Labor market shocks and retirement : Do government programs matter?" *Journal of Public Economics*, 91, 1902–1919.

——— (2009): "The Market Crash and Mass Layoffs : How the Current Economic Crisis May Affect Retirement." *National Bureau of Economic Research*, 15395.

——— (2010): "Reconsidering Retirement : How Losses and Layoffs affect Older Workers," *Washington DC : Brooking Institution Press.*

——— (2011): "Recessions, Retirement, and Social Security," *American Economic Review*, 101, 23–28.

CURRIE, J. (1999): *Health, Health Insurance and the Labor Market*, Elsevier Science, 1st edition ed.

DAVIDOFF, T. (2016): "Supply Constraints Are Not Valid Instrumental Variables for Home Prices Because They Are Correlated With Demand Factors," *Critical Finance Review*, 5, pp. 177–206.

DISNEY, R. AND J. GATHERGOOD (2014): "House Prices, Wealth Effects and Labor Supply," *Institute For Fiscal Studies Working Paper*.

DUSTMANN, C. AND F. FABBRI (2003): "Language Proficiency and Labour Market Performance of Immigrants in the UK," *Economic Journal*, 695–717.

DWYER, D. S. AND O. S. MITCHELL (1999): "Health problems as determinants of retirement : Are self-rated measures endogenous?" *Journal of Health Economics*, 18, 173–193.

FARANHAM, M. AND P. SEVAK (2016): "Housing Wealth and Retirement," *CESifo Economic Studies*, 26–46.

FARBER, H. S. (2011): "Job Loss in the Great Recession : Historical Perspective From the Displaced Worker Survey, 1984-2010," Working Paper 17040.

FAVARA, G. AND J. IMBS (2015): "Credit Supply and the Price of Housing," *American Economic Review*, 105, 958–92.

FISHER;, M. B. O. G. G. AND A. R. HERZOG (2005): *Documentation of Cognitive Functioning Measures in the Health and Retirement Study*, HRS/AHEAD Documentation Report, dr-006 ed.

FRANCK, R. H. AND P. J. COOK (1995): *The Winner-take-all Society.*, New York: Free Press.

FRENCH, E. (2005): "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour," *Review of Economic Studies*, 72, 395–427.

FRENCH, E. AND J. B. JONES (2011): "The Effects of Health Insurance and Self-Insurance on Retirement Behavior," *Econometrica*, 79, 693–732.

FRENETTE, M. AND R. MORISSETTE (2003): "Will they ever converge? Earnings of immigrants and Canadian-born workers over the last two decades," *Analyical Studies Branch Research paper series-Statistics Canada*, 11F0019MIE.

FRIEDBERG, L. (2000): "The Labor Supply Effects of the Social Security Earnings test," *The Review of Economics and Statistics*, 82, 48–63.

FRIEDMAN, M. (1957): "The Permanent Income Hypothesis," *Princeton University Press*.

FU, S., Y. LIAO, AND J. ZHANG (2016): "The effect of housing wealth on labor force participation : Evidence from China," *Journal of Housing Economics*, 59–69.

GLAESER, E. AND J. GYOURKO (2007): "Arbitrage in Housing Markets," *NBER Working Paper*.

GUSTMAN, A. L. AND T. L. STEINMEIER (2000): "Retirement in Dual-Career Families : A Structural Model," *Journal of Labor Economics*, 18, 503–545.

GYOURKO, J., C. MAYER, AND T. SINAI (2013): "Superstar Cities." *American Economic Journal : Economic Policy*, 5, 167–199.

GYOURKO, J., A. SAIZ, AND A. A. SUMMERS (2008): "A New Measure of the Local Regulatory Environment for Housing Markets : The Wharton Residential Land Use Regulatory Index," *Urban Studies*, 693–729.

HAM, M. V., D. MANLEY, N. BAILEY, L. SIMPSON, AND D. MACLENNAN (2012): *Neighbourhood Effects Research : New Perspectives*, Springer Dordrecht Heidelberg London New York.

HELLERSTEIN, J. K., M. J. KUTZBACH, AND D. NEUMARK (2014): "Do labor market networks have an important spatial dimension?" *Journal of Urban Economics*, 79, 39–58.

HELLERSTEIN, J. K., M. P. MCINERNEY, AND D. NEUMARK (2011): "Neighbors and coworkers: The importance of residential labor market networks," *Journal of Labor Economics*, 29, 659–695.

HIEBERT, D. (2009): "Exploring Minority Enclave Areas in Montreal, Toronto, and Vancouver," *Research and Evaluation, Citizenship and Immigration Canada*.

HIMMELBERG, C., C. MAYER, AND S. TODD (2005): "Assessing High House Prices : Bubbles, Fundamentals, and Misperceptions," *Journal of Economic Perspectives*, 67–92.

IACOVIELLO, M. (2011): "Housing Wealth and Consumption," *International Finance Discussion Papers*.

JOHN BOUND, MICHAEL SCHOENBAUM, T. R. S. AND T. WAIDMANN (1999): "The dynamic effects of health on the labor force transitions of older workers," *Labour Economics*, 6, 179–202.

JOHN BOUND, T. S. AND T. WAIDMANN (2010): "Health, economic resource and the work decisions of older men," *Journal of Econometrics*, 156, 106–129.

JOHNSON, C. AND T. RICE (2008): "Assessing a decade of Interstate Bank Branching," *Washington and Lee Law Review*, 73–127.

JOHNSON, W. R. (2014): "House prices and female labor force participation," *Journal of Urban Economics*, 82, 1–11.

LANGA, E. C. H. G. K. AND COAUTHORS; (2008): *Documentation of Physical Measures, Anthropometrics, and Blood Pressure in the Health and Retirement Study*, HRS Documentation Report, dr-011 ed.

LANGE;, F. AND D. MCKEE (2012): "The Evolution of Latent Health over the Life Course," .

MCCAMMON, G. G. F. J. J. M. R. J. AND COAUTHORS (2014): *New Measures of Fluid Intelligence in the HRS*, HRS Documentation Report, dr-027 ed.

MCFALL, B. H. (2011): "Crash and Wait? The Impact of the Great Recession on the Retirement Plans of Older Americans," *American Economic Review*, 101, pp. 40–44.

MIAN, A. AND A. SUFI (2011): "House Prices, Home Equity-Based Borrowing, and the US Household Leverage Crisis," *American Economic Review*, 2132–2156.

MICHAUD, P.-C., A. V. SOEST, AND L. BISSONETTE (2019): "Understanding joint retirement," *Journal of Economic Behavior and Organization*.

MIGLIORETTI, D. L. AND P. J. HEAGERTY (2006): "Marginal Modeling of Nonnested Multilevel Data using Standard Software," *American Journal of Epidemiology*, 165, 453–63.

MORRIS, E. D. (2006): "Examining the Wealth Effect from Home Price Appreciation," *Working Paper*.

MOULTON, B. R. (1986): "Random Group Effects and the Precision of Regression Estimates," *Journal of Econometrics*, 385–97.

——— (1990): "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units," *Review of Economic and Statistics*, 72, 334–38.

OF CANADA, G. (2011): "The Housing Conditions of Recent Immigrants Households," *Canada Mortgage and Housing Corporation*, Census-National Household Survey Housing Series.

——— (2017): "Canada's National Housing Strategy," *Department of Families, Children and Social Development*.

OREOPOULOS, P. (2011): "Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes," *American Economic Journal : Economic Policy 3*, 148.

PATACCHINI, E. AND Y. ZENOU (2012): "Ethnic Networks and employment Outcomes," *Regional Science and Urban Economics*, 42, 938–949.

POTERBA, J. M. (2000): "Stock Market Wealth and Consumption," *Journal of Economic Perspectives*, 14, 99–118.

RENA M. CONTI, E. R. B. AND R. G. FRANK (2006): "Early Retirement and Public Disability Insurance Applications : Exploring the Impact of Depression," *National Bureau of Economic Research*, Working paper 12237.

RESTREPO, P. (2015): "Skill Mismatch and Structural Unemployment," *Massachusetts Institute of Technology Job Market Paper*.

RICE, T. AND P. E. STRAHAN (2010): "Does Credit COmpetition Affect small-Firm Finance?" *Journal of Finance*, 861–889.

RUST, J. AND C. PHELAN (1997): "How social security and medicare affect retirement behavior in a world of incomplete markets," *Econometrica*, 65, 781–831.

SAIZ, A. (2010): "The geographic determinants of housing supply," *Quaterly Journal of Economics*, 125, 1253–1296.

SAXENA, K. AND P. WANG (2017): "How do house prices affect household consumption growth over the life cycle," *Working Paper*.

SCHMUTTE, I. M. (2015): "Job Referral Networks and the Determination of Earnings in Local Labor Markets," *Journal of Labor Economics*, 33, 1–32.

STOCK, J., J. WRIGHT, AND M. YOGO (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Methods of Moments," *Journal of Business and Economic Statistics*, 518–29.

STROEBEL, J. AND J. VAVRA (2015): "House Prices, Local Demand, and Retail Prices," *NBER*.

THOMPSON, S. (2006): "Simple Formulas for Standard Errors that Cluster by Both Firm and Time," *SSRN paper. http://ssrn.com/abstract=914002*.

——— (2011): "Simple Formulas for Standard Errors that Cluster by Both Firm and Time," *Journal of Financial Economics*, 99, 1–10.

TOLBERT, C. M. AND M. SIZER (1996): "US Commuting Zones and Labor Market Areas : A 1990 update," *Economic Research Service Staff Paper*.

WARREN, E. AND T. A. WARREN (2003): *The Two-income Trap : Why Middle-class Mothers and Fathers are going broke*, New York: Basic Books.

YAGAN, D. (2016): "The Enduring Employment Impact of Your Great Recession Location," *University of California, Berkeley Working Paper*.

ZAGO, R. (2015): "Skill Mismatch, Job Polarization and the Great Recession," *IZA*.

ZHAO, L. AND G. BURGE (2017): "Housing Wealth, Property Taxes and Labor Supply among the Elderly," *Journal of Labor Economics*, 35.

ZUCCHI, J. (2007): *A history of ethnic enclaves in Canada*, Canada Historical Association.

Annexes

4.1 Appendix of Chapter 1

4.1.1 Additional information regarding some variables in the HRS

Table 4.6: Detailed information about some variables in the Rand-HRS

Panel A : Labor Force Status Variable	
Value	Value Label
1	Working full-time
2	Working part-time
3	Unemployed
4	Retired and working part-time
5	Retired and not working
6	Disabled
7	Not in the labor force
Panel B : Type of Home[1]	
Value	Value Label
1	Mobile Home
2	One-Family House
3	Two-Family House//Duplex
4	Appartment//Townhouse
7	Other (Specify)
8	Don't Know
9	Refused
Panel B : Home ownership status[2]	
Value	Value Label
1	Own
2	rent
7	Other (Specify)
8	Don't Know
9	Refused

Wave 1994 : 1.Variable w607. 2.Variable w607.

4.1.2 Summary Statistics

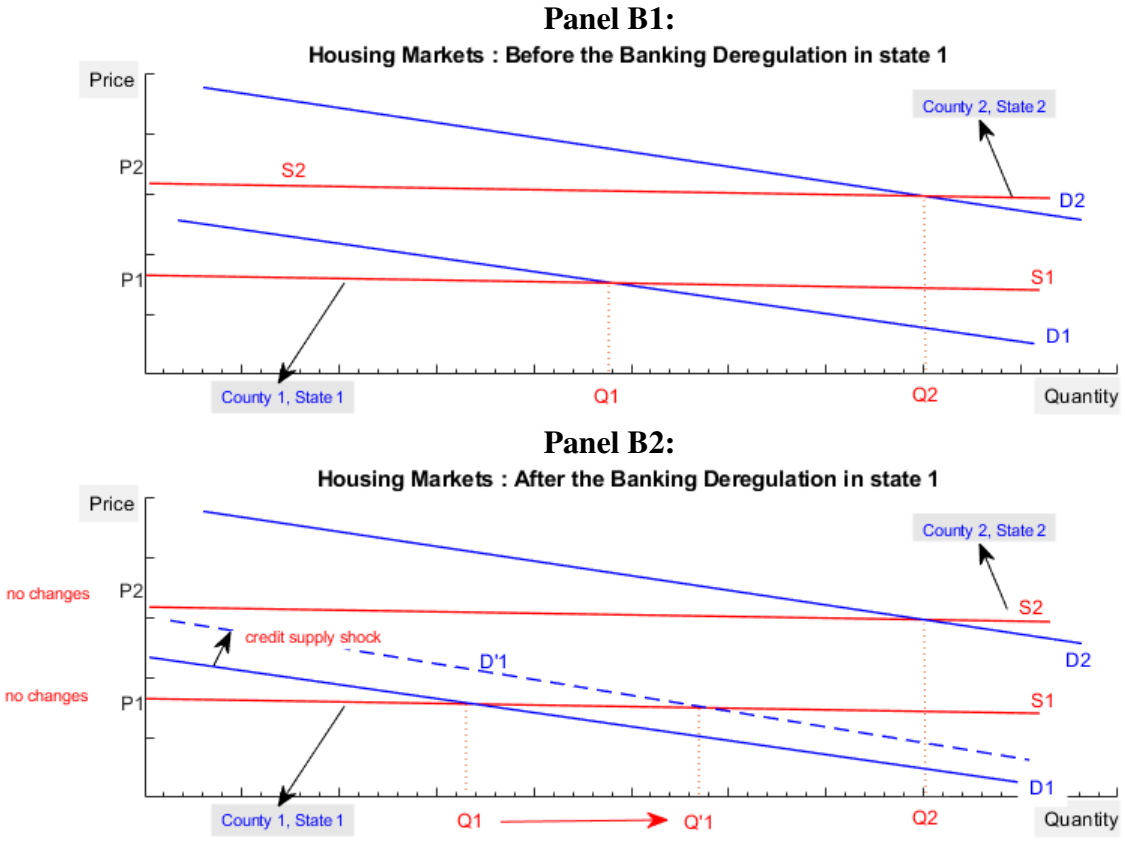
Table 4.7: Baseline Summary Statistics

	All	Men)	Women
Panel A : Demographics			
%Women	45.79	-	-
% white	83.41	85.76	80.64
% Married/partenered	82.00	92.11	70.03
Age (in years)	59.75	59.82	59.66
Spouse's age (in years)	57.95	55.67	61.47
% < High school	16.46	18.19	14.41
% High school grad.	36.02	32.19	40.55
% Some college	22.63	21.22	24.31
% College and +	24.89	28.40	20.72
Panel B : Income and financial situations (Baseline medians)			
	All	Men	Women
Annual earnings 1993 ($)	25,000	32,000	17,200
Household total inc. ($)	50,000	55,424	52,535
Household total Assets ($)	174,500	178,000	170,000
Household total non-housing assets ($)	83,200	90,000	78,000
Self-report. house value($)	100,000	100,000	96,000
Self-report. mortgage value($)	39,000	40,000	36,000
Self-report. value of savings($)	7000	7000	6700
Resp. Social Security wealth at NRA in 1992 ($)	105,500	118,800	85,900
Loan-To-Value Ratio (%)	35.75	35.72	36.07
Loan-To-Income Ratio (%)	67.10	60.72	72.58
Panel C : Labor Force Participation			
	(All)	(Men)	(Women)
% Working FT	73.46	82.21	63.10
% Spouse working	65.08	64.33	66.18
#Individuals	1577	869	708
% Attrition as of 2004	18.90	21.40	15.82

PT ≡ Part-time. LF ≡ Labor Force. NRA ≡ Normal Retirement Age.

4.1.3 *Illustration identification strategy : The case of perfectly elastic housing supply curves*

Figure 4.6: Illustration of the identification strategy



4.1.4 Additional Results

Table 4.8: OLS and 2SLS Estimations of the Effect of House Price Growth on the Labor Force Participation of Older Individuals

Dependent var. :	**Δ (Working or unemployed)**					
	OLS			**IV**		
	All	**Men**	**Women**	**All**	**Men**	**Women**
ΔLog(County Average Mortgage)	-0.133382**	-0.329603***	0.0972479	-0.868546*	-1.5542***	0.0601126
	(0.0552)	(0.0660)	(0.0957)	(0.4490)	(0.4270)	(0.7320)
	First Stage Results					
(Elasticity of housing supply)×(Banking Deregulation Index)$_{t-2}$	-	-	-	-0.0029171***	-0.0030548***	-0.0027879***
	-	-	-	(0.000835)	(0.0007925)	(0.0009103)
Kleibergen-Paap F-stat	-	-	-	12.206	14.858	09.380
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)*(year dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Dummies for year of reform change (DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)*(DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
Dependent var (mean)	-0.129	-0.128	-0.131	-0.129	-0.128	-0.131
ΔLog(County Average Mortgage)(mean)	0.051	0.051	0.051	0.051	0.051	0.051
ΔLog(County Average Mortgage) (sd)	0.094	0.094	0.094	0.094	0.094	0.094
# Obs.	7060	3827	3233	7060	3827	3233

***:1%. **:5%. * : 10%. Residuals are clustered at the state level. The sample is constituted by individuals born between 1931 and 1938, who were homeowners and working as of year 1994.

Table 4.9: OLS and 2SLS Estimations of the Effect of House Price Growth on the Labor Force Participation of Older Married Individuals

Dependent var. :	**Δ (Working FT or PT)**					
	OLS			**IV**		
	All	**Men**	**Women**	**All**	**Men**	**Women**
ΔLog(County Average Mortgage)	-0.15824**	-0.339098***	0.112287	-0.968132**	-1.69845***	0.196235
	(0.0685)	(0.0700)	(0.1440)	(0.4390)	(0.4970)	(0.9770)
	First Stage Results					
(Elasticity of housing supply)×(Banking Deregulation Index)$_{t-2}$	-	-	-	-0.0028872***	-0.0030856***	-0.002505***
	-	-	-	(0.0008399)	(0.0008100)	(0.0009125)
Kleibergen-Paap F-stat	-	-	-	11.818	14.513	07.535
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)*(year dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Dummies for year of reform change (DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)*(DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
Dependent var (mean)	-0.132	-0.130	-0.135	-0.132	-0.130	-0.135
ΔLog(County Average Mortgage)(mean)	0.051	0.051	0.051	0.051	0.051	0.051
ΔLog(County Average Mortgage) (sd)	0.094	0.094	0.094	0.094	0.094	0.094
# Obs.	5789	3525	2264	5789	3525	2264

***:1%. **:5%. * : 10%. Residuals are clustered at the state level.The sample is constituted by married or partnered individuals born between 1931 and 1938, who were homeowners and working as of year 1994.

House Price and Borrowing Constraints

Table 4.10: OLS and 2SLS Estimations of the Effect of House Price Growth on the Labor Force Participation of Older Individuals

Dependent var. :	**Δ (Working FT or PT)**					
	OLS			**IV**		
	All M	**High Leverage**	**Low Leverage**	**All**	**High Leverage**	**Low leverage**
ΔLog(County Average Mortgage)	-0.328***	-0.231*	-0.539***	-1.697***	-2.035**	-0.926
	(0.071)	(0.114)	(0.097)	(0.475)	(0.929)	(0.602)
	First Stage Results					
(Elasticity of housing supply)×(Banking Deregulation Index)$_{t-2}$	-	-	-	-0.0030548***	-0.0031261***	-0.0047199***
	-	-	-	(0.0007925)	(0.0008883)	(0.0008351)
Kleibergen-Paap F-stat	-	-	-	14.858	12.384	31.944
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)*(year dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Dummies for year of reform change (DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)*(DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
Dependent var (mean)	-0.130	-0.121	-0.123	-0.130	-0.121	-0.123
ΔLog(County Average Mortgage)(mean)	0.051	0.051	0.051	0.051	0.051	0.051
ΔLog(County Average Mortgage) (sd)	0.094	0.094	0.094	0.094	0.094	0.094
# Obs.	3827	1074	973	3827	1074	973

***:1%. **:5%. * : 10%. Residuals are clustered at the state level. The sample is constituted by individuals born between 1931 and 1938, who were homeowners and working as of year 1994.

House Price and Housing Consumption

Table 4.11: OLS and 2SLS Estimations of the Effect of House Price Growth on the Housing Consumption of Older Men

Dependent var. :	**Δ (Working FT or PT)**					
	OLS			**IV**		
	Δ(2nd Mortg.)	**Δ(Appartment)**	**Δ(Renter)**	**Δ(2nd Mortg.)**	**Δ(Appartment)**	**Δ(Renter)**
ΔLog(County Average Mortgage)	-0.044	-0.078**	-0.048	-0.107	-0.363	-0.103
	(0.062)	(0.036)	(0.030)	(0.143)	(0.238)	(0.153)
	First Stage Results					
(Elasticity of housing supply)×(Banking Deregulation Index)$_{t-2}$	-	-	-	-0.0049758***	-0.0030338***	-0.0030513***
	-	-	-	(0.0010204)	(0.000798)	(0.0007948)
Kleibergen-Paap F-stat	-	-	-	23.776	14.452	14.740
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)*(year dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Dummies for year of reform change (DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
(Birth year dummies)*(DYRC)	Yes	Yes	Yes	Yes	Yes	Yes
Dependent var (mean)	-0.130	-0.121	-0.123	-0.130	-0.121	-0.123
ΔLog(County Average Mortgage)(mean)	0.051	0.051	0.051	0.051	0.051	0.051
ΔLog(County Average Mortgage) (sd)	0.094	0.094	0.094	0.094	0.094	0.094
# Obs.	3827	1074	973	3827	1074	973

***:1%. **:5%. * : 10%. Residuals are clustered at the state level. The sample is constituted by men born between 1931 and 1938, who were homeowners and working as of year 1994.

House Price Growth and Equity-Based Borrowing

Table 4.12: OLS and 2SLS Estimations of the Effect of House Price Growth on the Equity-Based Borrowing Behavior of Older Men

Dependent var. :	**Δ (Working FT or PT)**			
	OLS		**IV**	
	Δ(Have an ELC)	**Δ(Have a loan on the ELC)**	**Δ(Have an ECL)**	**Δ(Have a loan on the ELC)**
ΔLog(County Average Mortgage)	-0.001521	0.0452723	0.0905358	0.38465
	(0.0594)	(0.0642)	(0.4430)	(0.4310)
	First Stage Results			
(Elasticity of housing supply)×(Banking Deregulation Index)$_{t-2}$	-	-	-0.0029698***	-0.0029698***
	-	-	(0.0008146)	(0.0008146)
Kleibergen-Paap F-stat	-	-	13.293	13.293
Year dummies	Yes	Yes	Yes	Yes
(Birth year dummies)*(year dummies)	Yes	Yes	Yes	Yes
Dummies for year of reform change (DYRC)	Yes	Yes	Yes	Yes
(Birth year dummies)*(DYRC)	Yes	Yes	Yes	Yes
Dependent var (mean)	0.002	0.001	0.002	0.001
ΔLog(County Average Mortgage)(mean)	0.051	0.051	0.051	0.051
ΔLog(County Average Mortgage) (sd)	0.094	0.094	0.094	0.094
# Obs.	3639	3639	3639	3639

***:1%. **:5%. * : 10%. Residuals are clustered at the state level. The sample is constituted by men born between 1931 and 1938, who were homeowners and working as of year 1994. ECL ≡ Equity line of credit.

House Prices and Local Unemployment

Table 4.13: OLS and 2SLS Estimations of the Effect of House Price on County Unemployment

Dependent var. :	**Δ county UR**	**Δ Log(county UR)**	**Δ county UR**	**Δ Log(county UR)**
	OLS		**IV**	
ΔLog(County Average Mortgage)	0.0013816	0.0026285	0.0059421	0.781174
	(0.00684)	(0.0833)	(0.0343)	(0.531)
Year dummies	Yes	Yes	Yes	Yes
County dummies	Yes	Yes	Yes	Yes
Dummies for year of reform change (DYRC)	Yes	Yes	Yes	Yes
Dependent var (mean)	-0.001	-0.000	-0.001	-0.000
# Obs.	874	874	874	874

***:1%. **:5%. * : 10%. Residuals are clustered at the state level. UR ≡ unemployment rate.

Table 4.14: OLS and 2SLS Estimations of the Effect of House Price on County Unemployment

Dependent var. :	**Δ state UR**	**Δ Log(state UR)**	**Δ state UR**	**Δ Log(state UR)**
	OLS		**IV**	
ΔLog(County Average Mortgage)	-0.0028166	-0.0830189	0.0151887	0.357777
	(0.00388)	(0.0852)	(0.0249)	(0.604)
Year dummies	Yes	Yes	Yes	Yes
County dummies	Yes	Yes	Yes	Yes
Dummies for year of reform change (DYRC)	Yes	Yes	Yes	Yes
Dependent var (mean)	-0.001	-0.019	-0.001	-0.019
# Obs.	874	874	874	874

***:1%. **:5%. * : 10%. Residuals are clustered at the state level. UR ≡ unemployment rate.

Estimates using the wild-cluster bootstrap approach

Table 4.15: OLS and 2SLS Estimations of the Effect of House Price on the Labor Force Participation of Older Individuals (*with wild-cluster bootstrap*)

Dependent var. :	**Δ (Working FT or PT)**			
	OLS		**IV**	
	All	**Men**	**All**	**Men**
ΔLog(County Average Mortgage)	-0.159**	-.328***	-1.082**	-1.697**
	Wild cluster bootstrap statistics			
t[1]	-2.6215	-4.5855	-2.3755	-3.5715
Prob $>\|t\|$[2]	0.0270	0.0040	0.0400	0.0210
Year dummies	Yes	Yes	Yes	Yes
(Birth year dummies)*(year dummies)	Yes	Yes	Yes	Yes
Dummies for year of reform change (DYRC)	Yes	Yes	Yes	Yes
(Birth year dummies)*(DYRC)	Yes	Yes	Yes	Yes
Dependent var (mean)	-0.131	-0.130	-0.131	-0.130
ΔLog(County Average Mortgage)(mean)	0.051	0.051	0.051	0.051
ΔLog(County Average Mortgage) (sd)	0.094	0.094	0.094	0.094
# Obs.	7060	3827	7060	3827

***:1%. **:5%. * : 10%. The sample is constituted by individuals born between 1931 and 1938, who were homeowners (of single family-unit) and working as of year 1994. [1] : z statistics for IV. [2] : Prob $>|z|$ for IV.

4.2 Appendix of Chapter 2

4.2.1 Summary Statistics

Table 4.16: Baseline Summary Statistics

	(All)	(Good health)	(Poor health)	(Diff.(Good-Poor)
Panel A : Demographics				
Men	0.460	0.446	0.493	-0.047***
Age (in years)	59.212	59.044	59.615	-0.571***
Non white	0.197	0.138	0.341	-0.203***
Married/partenered	0.737	0.754	0.695	+0.058***
< High school	0.101	0.051	0.220	-0.169***
High school grad.	0.316	0.281	0.399	-0.117***
Some college	0.273	0.284	0.246	0.038***
College+	0.310	0.383	0.135	0.247***
Panel B : Income and financial situations (Baseline medians)				
Annual earnings 2005 ($)	26000	30000	20833	-
Household total inc. ($)	64824	74718	49000	-
HH total Assets ($)	220000	277700	105500	-
HH total non-housing assets ($)	70000	104000	25000	-
HH non-housing Financial assets($)	11000	19000	2000	-
HH capital income($)	760	1472.5	48	-
Resp. SS wealth at age 62 ($)	130200	137150	113750	-
Resp. SS wealth at nra ($)	142250	150100	124250	-
Resp. SS wealth at age 70 ($)	141400	148700	125900	-
Pension wealth ($)	95056.170	102461	79067	-
Panel C : labor market outcomes (Baseline)				
Working	0.805	0.807	0.802	0.004
Retired and working	0.161	0.165	0.152	+0.013
Unemployed	0.033	0.028	0.045	-0.017***
Retired and not working	-	-	-	-
Disabled	-	-	-	-
Not in the LF	-	-	-	-
#Individuals	3953	2788	1165	-
Panel D : labor market outcomes (2008)				
Working	0.681	0.692	0.654	0.037***
Retired and working	0.144	0.153	0.123	+ 0.030***
Unemployed	0.028	0.024	0.038	-0.015***
Retired and not working	0.125	0.114	0.152	-0.038***
Disabled	0.007	0.003	0.015	-0.012***
Not in the LF	0.015	0.014	0.016	-0.003
#Individuals	3863	2747	1116	-
Attrition rate (%)	02.277	01.470	04.206	-02.736***
Panel E : labor market outcomes (2010)				
Working	0.520	0.538	0.473	0.065***
Retired and working	0.160	0.167	0.142	0.025
Unemployed	0.051	0.048	0.057	-0.008
Retired and not working	0.251	0.229	0.306	-0.077***
Disabled	0.006	0.005	0.006	-0.000
Not in the LF	0.013	0.012	0.016	-0.004
Earnings in 2009 (2005 $)	46224.890	50537.440	33552.340	17015.100***
#Individuals	3587	2583	1004	-
Attrition rate (%)	07.145	05.970	10.036	-04.066***

4.2.2 Other Estimation Results

Dependent var : not working (Unemployed + Out of the LF)

Table 4.17: Effect of the Great Recession on Individuals Labor Supply

Dependent var. :	**Not working (Unemployed + Out of the LF)**					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Unemployment Rate (UR)	3.381***	1.107*	1.109	2.624***	0.411	-0.576
	(0.555)	(0.633)	(0.715)	(0.495)	(0.747)	(0.814)
Self-reported bad health in 2006				-0.039	-0.036	
				(0.034)	(0.034)	
(Self-reported bad health in 2006)*UR				1.720**	1.665**	4.017**
				(0.714)	(0.733)	(1.680)
Constant	Yes	Yes	No	Yes	Yes	No
State dummies	No	Yes	No	No	Yes	No
Year dummies	No	Yes	Yes	No	Yes	Yes
Age dummies	No	Yes	Yes	No	Yes	Yes
Oth. Demographics	No	Yes	No	No	Yes	No
Ind. fixed effects	No	No	Yes	No	No	Yes
Dependent var (mean)	0.104	0.104	0.104	0.104	0.104	0.104
# Obs.	7806	7806	7806	7806	7806	7806
R2	0.018	0.074	0.602	0.026	0.081	0.604
Self-reported bad health in 2006 (mean)	0.452	0.452	0.452	0.452	0.452	0.452
State Unemployment Rate in 2006 (mean)	0.046	0.046	0.046	0.046	0.046	0.046
State Unemployment Rate in 2008 (mean)	0.058	0.058	0.058	0.058	0.058	0.058

***:1%. **:5%. * : 10%. Residuals are clustered at the State level. Oth. demographics: sex and race.

Dependent var : Out of the Labour Force

Table 4.18: Effect of the Great Recession on Individuals Labor Supply

Dependent var. :	**Out of the Labour Force**					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Unemployment Rate (UR)	3.465***	0.800	0.750	2.765***	0.145	-0.890
	(0.599)	(0.537)	(0.601)	(0.472)	(0.648)	(0.766)
Self-reported bad health in 2006				-0.047	-0.047	
				(0.032)	(0.033)	
(Self-reported bad health in 2006e)*UR				1.599**	1.565**	3.911***
				(0.675)	(0.698)	(1.252)
Constant	Yes	Yes	No	Yes	Yes	No
State dummies	No	Yes	No	No	Yes	No
Year dummies	No	Yes	Yes	No	Yes	Yes
Age dummies	No	Yes	Yes	No	Yes	Yes
Oth. Demographics	No	Yes	No	No	Yes	No
Ind. fixed effects	No	No	Yes	No	No	Yes
Dependent var (mean)	0.073	0.073	0.073	0.073	0.073	0.073
# Obs.	7806	7806	7806	7806	7806	7806
R2	0.026	0.108	0.558	0.032	0.113	0.561
Self-reported bad health in 2006 (mean)	0.452	0.452	0.452	0.452	0.452	0.452
State Unemployment Rate in 2006 (mean)	0.046	0.046	0.046	0.046	0.046	0.046
State Unemployment Rate in 2008 (mean)	0.058	0.058	0.058	0.058	0.058	0.058

***:1%. **:5%. * : 10%. Residuals are clustered at the State level. Oth. demographics: sex at birth and race.

Results when controlling education and occupation

Table 4.19: Effect of the Great Recession on Individuals Labor Supply

Dependent var. :	**Retired and not Working**					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
State Unemployment Rate (UR)	-0.565	-0.560	-1.095	-2.035*	-4.814*	-4.741
	(0.766)	(0.774)	(1.279)	(1.075)	(2.701)	(3.084)
(Self-reported bad health in 2006)*UR	2.965**		2.843**	2.478**	2.564**	2.295**
	(1.333)		(1.178)	(1.178)	(1.179)	(1.124)
(< High School)*(Self-reported bad health in 2006)*UR		2.062				
		(2.846)				
(High School Grad.)*(Self-reported bad health in 2006h)*UR		3.654***				
		(1.271)				
(College Grad.)*(Self-reported bad health in 2006)*UR		1.561				
		(2.139)				
(< High School)*UR			0.687			
			(3.062)			
(High School Grad.)*UR			0.864			
			(1.278)			
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes
Ind. fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
(Occupations dummies)*UR	No	No	No	Yes	No	Yes
(Industries dummies)*UR	No	No	No	No	Yes	Yes
Dependent var (mean)	0.062	0.062	0.062	0.059	0.045	0.045
# Obs.	7809	7809	7809	7464	5125	5123
R2	0.561	0.561	0.561	0.563	0.556	0.562

*** : 1%. ** : 5%. * : 10%. Residuals are clustered at the state level. Other demographics: sex at birth and race.

4.3 Appendix of Chapter 3

4.3.1 Definition of core housing need

A household is in **core housing need** (CHN) if its housing is below one or more of the adequacy, suitability and affordability standards, and it would have to spend 30% or more of its before-tax household income to access local housing that meets all three standards.
Adequate housing does not require any major repairs, according to residents.
Suitable housing has enough bedrooms for the size and makeup of resident households, according to National Occupancy Standard (NOS) requirements.
Affordable housing costs less than 30% of before-tax household income. For renters, shelter costs include, as applicable, rent and payments for electricity, fuel, water and other municipal services. Households assessed for core housing need comprise private, non-farm, non-band, non-reserve households with incomes greater than zero and shelter cost-to-income ratios (STIRs) less than 100%.

4.3.2 Earning trends of permanent residents who landed in Canada in 2001

Table 4.20: Average earnings by CMA of residence

Year	MVT		Montreal		Vancouver		Toronto	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
2001	6474	1371***	7007	533***	6398	1752***	6474	1447***
2002	12697	1168***	10516	1636***	11607	2337***	13087	1714***
2003	15606	1441***	12965	1841***	14406	2321***	16167	2241***
2004	18420	1386***	14590	2527***	16627	2853***	19235	2364***
2005	21248	1583***	16748	3166***	19248	3000***	22247	2500***
2006	24396	1615***	19891	3145***	22356	3230***	25416	2550***
2007	27781	1823***	23440	3212***	25698	3473***	28823	2778***
2008	31110	2044***	26666	3555***	29155	3378***	32088	3111***
2009	32006	2496***	27994	4012***	30045	3477***	32987	3388***
2010	34750	2268***	30122	4355***	32028	3266***	36020	3085***
2011	37715	2707***	33420	4854***	34821	3174***	38928	3454***
2012	40377	2654***	36206	4265***	37533	3317***	41609	3696***
2013	42376	3157***	37306	6026***	39411	3252***	43715	4018***
2014	44847	3120***	39290	6435***	42020	3315***	46115	3997***
2015	47025	2859***	41011	6704***	44363	2563***	48307	3648***
2016	48900	3300***	43200	6800***	45900	3400***	50300	4100***

***:1%. **:5%. * : 10%. MVT ≡ Montreal, Vancouver and Toronto. Results are produced from tax data of permanent residents who landed in Montreal, Vancouver and Toronto in 2001. However, Individuals with earnings below the quantile 5% of the earning distribution where excluded from the calculation of the average earnings; this was done to avoid bias from extremely low values that may arise from individuals with only a weak attachment to the labour market. Column (a) : average earnings of individuals living in an ethnic enclave. Column (b) : average earning gap ≡ (average earnings of individuals not living in enclave) - (average earnings of individuals living in an ethnic enclave).

Source: CMHC, computed from Statistics Canada Canadian Employer-Employee Dynamics Database, vintage 2017.

Table 4.21: Average earnings by sex at birth

Year	All		Men		Women	
	(a)	(b)	(a)	(b)	(a)	(b)
2001	6474	1371***	7007	1447***	5788	1295***
2002	12697	1168***	14411	1091***	10516	1402***
2003	15606	1441***	17767	1441***	12965	1601***
2004	18420	1386***	21110	1304***	15242	1630***
2005	21248	1583***	24497	1250***	17581	1916***
2006	24396	1615***	28222	1360***	20146	2040***
2007	27781	1823***	32035	1563***	23180	2257***
2008	31110	2044***	35821	1778***	26133	2400***
2009	32006	2496***	36285	2496***	27459	2675***
2010	34750	2268***	39286	2268***	30032	2450***
2011	37715	2707***	42382	2894***	32767	2707***
2012	40377	2654***	45305	2938***	35164	2654***
2013	42376	3157***	47542	3252***	37019	3157***
2014	44847	3120***	50210	3315***	39290	3120***
2015	47025	2859***	52447	2958***	41307	3056***
2016	48900	3300***	54500	3100***	43200	3600***

***:1%. **:5%. * : 10%. Results are produced from tax data of permanent residents who landed in Montreal, Vancouver and Toronto in 2001. However, Individuals with earnings below the quantile 5% of the earning distribution where excluded from the calculation of the average earnings; this was done to avoid bias from extremely low values that may arise from individuals with only a weak attachment to the labour market. Column (a) : average earnings of individuals living in an ethnic enclave. Column (b) : average earning gap $\equiv$ (average earnings of individuals not living in enclave) - (average earnings of individuals living in an ethnic enclave).

Source: CMHC, computed from Statistics Canada Canadian Employer-Employee Dynamics Database, vintage 2017.

Table 4.22: Average earnings by age group

Year	All		20-29		30-39		40-49	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
2001	6474	1371***	6626	1371***	6398	1371***	6398	1142***
2002	12697	1168***	12463	1324***	12853	1246***	12619	1091***
2003	15606	1441***	15206	1521***	16006	1441***	15526	1281***
2004	18420	1386***	17768	1630***	18991	1223***	18257	1223***
2005	21248	1583***	20414	1833***	22247	1083***	20414	2000***
2006	24396	1615***	23461	1955***	25671	1190***	22951	1870***
2007	27781	1823***	26653	2257***	29344	1302***	25958	2084***
2008	31110	2044***	29688	2933***	33155	1156***	28622	2311***
2009	32006	2496***	30936	3388***	34056	1516***	28796	3031***
2010	34750	2268***	33842	3176***	36836	1270***	31120	2903***
2011	37715	2707***	36594	3734***	40048	1680***	33607	2894***
2012	40377	2654***	39429	3791***	42651	1706***	36206	2559***
2013	42376	3157***	41420	4400***	44768	2009***	37689	3348***
2014	44847	3120***	43872	4485***	47285	2047***	39973	3022***
2015	47025	2859***	46138	4535***	49391	1873***	41899	1775***
2016	48900	3300***	48000	5300***	51400	2100***	43800	1900***

***:1%. **:5%. * : 10%. Results are produced from tax data of permanent residents who landed in Montreal, Vancouver and Toronto in 2001. However, Individuals with earnings below the quantile 5% of the earning distribution where excluded from the calculation of the average earnings; this was done to avoid bias from extremely low values that may arise from individuals with only a weak attachment to the labour market. Column (a) : average earnings of individuals living in an ethnic enclave. Column (b) : average earning gap $\equiv$ (average earnings of individuals not living in enclave) - (average earnings of individuals living in an ethnic enclave).

Source: CMHC, computed from Statistics Canada Canadian Employer-Employee Dynamics Database, vintage 2017.