Essays on Spatial Labor Markets and Public Policies

by

Juliette Fournier

B.S., École Normale Supérieure de Paris (2012)
M.S., Université Paris-Sud/ENS (2014)
M.S., Paris School of Economics/ENS (2015)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2021

Signature of Author . . . .
Department of Economics
May 14, 2021

Certified by . . . .
David H. Autor
Ford Professor of Economics
Thesis Supervisor

Certified by . . . .
Arnaud Costinot
Professor of Economics
Thesis Supervisor

Certified by . . . .
Iván Werning
Robert M. Solow Professor of Economics
Thesis Supervisor

Accepted by . . . .
Amy Finkelstein
John & Jennie S. MacDonald Professor of Economics
Chairman, Department Committee on Graduate Theses

Essays on Spatial Labor Markets and Public Policies

by

Juliette Fournier

Submitted to the Department of Economics
on May 14, 2021, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis consists of three essays on spatial labor markets and public policies. I study successively the interactions of space with job search, demography and housing policy.

In the first essay, I develop a framework to study theoretically and quantitatively the welfare attributes of spatial mismatch, defined as a misalignment between where job seekers reside and suitable employment opportunities. In a quantitative urban model with frictional labor markets, the structure of the city interacts with labor markets because commuting is costly and information about job offers decays with distance. The decentralized equilibrium might feature too much or too little spatial mismatch, depending on whether commuting costs or information decay dominates. When commuting costs prevail, the constrained-efficient allocation may be restored by a mix of moving-to-opportunity and enterprise zone interventions that bring jobs and workers together.

The second essay, joint with David Autor, studies the relationship between population age and population density in the United States. We document the inversion of the rural-urban age gradient between 1950 and 2019. Whereas in 1950, residents in the least dense counties were on average 4.5 years younger than their counterparts in the most dense counties, by 2019 residents of the most rural counties were 2.7 years older than those in the most urban counties, a swing of 7.2 years. We show that sharp temporal changes in age-specific migration rates were the predominant contributor to this reversal.

In the third essay, Hector Blanco and I examine the distributional implications of the shift from public housing to subsidized private housing initiated by the U.S. government over the past few decades. We build a quantitative urban framework where housing assistance complements income taxation to redistribute across workers. We argue that provision of affordable housing involves a trade-off between indirect pecuniary redistribution and direct amenity spin-offs. On the one hand, public housing drives local rents down, while amplifying the spatial concentration of poverty. On the other hand, project- and tenant-based rental assistance enhances the local amenities of subsidized households by promoting mixed-income communities, but pushes private landowners' rents up.

Thesis Supervisor: David H. Autor
Title: Ford Professor of Economics

Thesis Supervisor: Arnaud Costinot
Title: Professor of Economics

Thesis Supervisor: Iván Werning
Title: Robert M. Solow Professor of Economics

Acknowledgments

I could not have hoped for better advisors than David Autor, Arnaud Costinot and Iván Werning. They have always been generous with their time, not only providing invaluable feedback on my work but also offering mentorship and support in innumerable other ways. From my early days in the program, David patiently guided me as I developed my tastes in research. He taught me everything—delineating a good research question, crafting economics papers, and even managing bewildering seminar comments. His intellectual curiosity, attention to real-world issues and creativity spurred my fledgling research and stand as a model for the economist I aspire to be. Arnaud directed me through the meanders of applied general equilibrium theory. His clarity of thought and ability to convert any crude idea into a crisp research project were instrumental to my progress. Iván's deep insights enriched tremendously my work. I admire his thorough approach to theoretical public economics, which I will strive to emulate. Above all, I am grateful to David, Arnaud and Iván for encouraging me to follow my own research path and making it possible to work seamlessly at the intersection of the labor, spatial and public fields.

MIT Department of Economics is famous for its engaging and collaborative environment, from which this thesis benefited greatly. Although officially I had only three advisors, I often felt I had many more given the countless comments and extensive feedback that my work received. First and foremost, Daron Acemoglu challenged me to improve each iteration of my job market paper, diverting me away from unpromising avenues. Jim Poterba was the first person to welcome me to MIT. His contagious enthusiasm for public finance inspired me over years. Josh Angrist taught me how to teach. Conversations with David Atkin and Dave Donaldson, as well as their numerous comments at the Trade Tea, infused this dissertation. Amy Finkelstein's sharp insights amazed me at every lunch and seminar. The eagerness to teach, advise and engage with other people's ideas is a defining trait of MIT culture, and an example for how to treat my future students and colleagues.

Before starting my PhD, I was fortunate to spend a year at Toulouse School of Economics. I thank Jean Tirole for welcoming me there. I learnt from his guidance and his kindness more than he is aware of.

My friends and classmates have been an unfailing support and source of joy during PhD years. First of all, I thank my roommates, Pari and Charles, for the laughters, the gossiping, the cooking and the movie nights. Those five years would have been grim without Clémence's birthday cakes, week-ends in NYC with Charlotte and Pierre, pleasant coffee breaks with Jannis, WhatsApp laughing and commiserating with Lydia, inimitable lunches with May, fruitful discussions with Garima, exchange of TAing experience with Christine. Some classmates became both friends and coauthors; I'm excited to pursue collaborations with Hector, Laura and Emilie. Olivier's advice helped me navigate through the PhD. My officemates, Joe, Chen and Will, made me forget the absence of windows.

I owe my parents everything. I thank my family for their unconditional love and support.

Contents

1 Spatial Mismatch **11**
1.1 Introduction 11
1.2 Toy Model: The Inefficiencies of Spatial Mismatch 17
1.2.1 A Duocentric City Model with Frictional Labor Markets 17
1.2.2 Comparative Statics: Detecting the Channels of Spatial Mismatch 20
1.2.3 Welfare Analysis: Too Much or Too Little Spatial Mismatch? 21
1.3 Background and Data 26
1.3.1 Background 26
1.3.2 Data 29
1.3.3 Some Descriptive Statistics 30
1.4 A Quantitative Model of Spatial Mismatch 32
1.4.1 Environment 32
1.4.2 Equilibrium 36
1.4.3 Welfare Analysis: Extending Toy Model's Results 41
1.5 Model Estimation: Disentangling the Channels of Spatial Mismatch 42
1.5.1 Quantitative Implementation 42
1.5.2 Structural Estimation 44
1.6 Conclusion 51

Appendix to Chapter 1 **59**
1.A Institutional Appendix 59
1.B Omitted Proofs 61
1.B.1 Proofs of Section 1.2 61
1.B.2 Proofs of Section 1.4 65

2 No Country for Young Men: The Inversion of the Rural-Urban Age Gradient in the United States, 1950-2019 **73**
(joint with David Autor)
2.1 Introduction 73
2.2 Trends in Natality, Mortality, and Migration across Time and Space 77

2.3 Decomposing Population Aging . . . 83
2.3.1 The Key Role of Youngsters . . . 83
2.3.2 Assessing the Contributions of Natality, Mortality and Migration . . . 84
2.3.3 Decomposition Results . . . 86
2.4 Conclusion . . . 92

Appendix to Chapter 2 **97**
2.A Additional Figures and Tables . . . 97
2.B Data Appendix . . . 101
2.B.1 Data Sources . . . 101
2.B.2 Variable Construction . . . 101
2.B.3 Sample Selection . . . 103
2.C Omitted Proofs . . . 103
2.C.1 Continuous-Time Results . . . 103
2.C.2 Implementation with Discrete-Time Data . . . 105

3 Redistribution through Housing Assistance **107**
(joint with Hector Blanco)
3.1 Introduction . . . 107
3.2 Background and Data . . . 110
3.2.1 Background . . . 110
3.2.2 Data . . . 112
3.2.3 Descriptive Evidence . . . 113
3.3 A Quantitative Model with Income Taxation and Housing Assistance . . . 118
3.3.1 Environment . . . 118
3.3.2 Equilibrium . . . 121
3.4 Model Estimation: Leveraging Public Housing Demolitions . . . 125
3.4.1 Quantitative Implementation . . . 125
3.4.2 Structural Estimation . . . 127
3.5 Conclusion . . . 129

Appendix to Chapter 3 **133**
3.A Data Appendix . . . 133

List of Figures

1.1.1 Distributions of job seekers and vacancies in Paris and its suburbs 12
1.2.1 Place-based subsidies as a function of commuting costs and information decay . . . 25
1.3.1 Map of the ZFUs 28
1.5.1 Wage premium for commuting 47
1.5.2 Home bias in employment 48
1.5.3 Estimation of information decay κ^{Δ} 50

2.1.1 Mean and median ages of county residents: Rural areas used to be relatively younger and are now relatively older 74
2.2.1 Natality fell relatively and absolutely more in rural counties 78
2.2.2 Raw and age-adjusted mortality fell differentially in urban counties 79
2.2.3 Migration flows differ substantially by decade 81
2.2.4 Migration over the life cycle across decades 82
2.3.1 Decadal contributions of each age bracket to changes in mean county age 84
2.3.2 Decadal contributions of natality, mortality, migration, and initial age distribution to changes in mean county age 87
2.3.3 Contributions of natality, mortality, migration and initial age distribution to changes in the rural-urban age gradient by decade 88
2.3.4 Decadal contribution of natality by mothers' age to changes in mean county age . . 89
2.3.5 Decadal contribution of mortality by age to changes in mean county age 90
2.3.6 Decadal contribution of migration by age to changes in mean county age 91
2.A.1 Map of county population density in 1950 97
2.A.2 Natality over the life cycle across decades 98
2.A.3 Mortality over the life cycle across decades 99
2.A.4 Decadal contributions of initial distribution of the different age brackets to changes in mean county age 100

3.1.1 Number of beneficiaries by housing assistance program 108
3.2.1 Demolitions under HOPE VI were far from offset by new construction 114
3.2.2 Increased exposure to demolitions raises house prices 116

List of Tables

1.3.1 Summary statistics 31
1.5.1 Estimation of κ^d 46
1.5.2 Estimation of commuting costs κ^δ 49
1.A.1 Comparison of exemptions available in ZRUs and ZFUs in 2009 60

3.2.1 Exposition to demolitions coincides with neighborhood composition changes 115
3.2.2 Exposure to demolitions raises house prices more in low-income areas 117
3.4.1 Estimation of housing supply elasticity ζ 128
3.4.2 Estimation of residential mobility σ and local amenity spillovers μ^A 129

Chapter 1

Spatial Mismatch[1]

1.1 Introduction

The most precarious workers—largely unskilled and low-educated—are often confined to impoverished city areas plagued with unemployment. Rather than upgrade their skills, policies intended for urban ghetto residents attempt to bridge the physical distance between them and job opportunities. Prominent examples include enterprise zones, that attract firms to distressed neighborhoods, and moving-to-opportunity designs, that help inhabitants leave high-poverty areas. By enacting such programs, policymakers implicitly acknowledge the responsibility of *spatial mismatch*—defined as a misalignment between where workers reside and suitable employment opportunities—in the adverse labor market outcomes of urban ghetto dwellers. Further, they imply that spatial mismatch entails productive inefficiencies that should be corrected by adequate place-based interventions. Yet the welfare benefits of those policies have not been demonstrated in the past academic literature.

This paper develops a framework to study theoretically and quantitatively the welfare attributes of spatial mismatch. I introduce frictional labor markets into a quantitative urban model. The spatial structure of the city interacts with labor markets because commuting is costly and information about job offers decays with distance. I prove that both workers' choice of residence and vacancy creation are inefficient in the decentralized equilibrium. The constrained-efficient allocation may be restored by a mix of place-based residence and hiring subsidies. I plan to use this framework to quantify the welfare implications of spatial mismatch. I will apply my model to French urban ghettos and leverage a spatial experiment to estimate it. I will evaluate through the lens of my model that the enterprise zone program supposed to undo spatial mismatch in French urban ghettos. Finally, I will explore a range of counterfactual policies designed to tackle spatial mismatch.

[1] I thank the George and Obie Shultz Fund at MIT for financial support. This paper has been prepared by the author under the Alfred P. Sloan Foundation Pre-doctoral Fellowship on the Economics of an Aging Workforce, awarded through the NBER, and the Future of Work Fellowship, sponsored by the OECD. This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the *Investissements d'Avenir* program (reference: ANR-10-EQPX-17 – Centre d'accès sécurisé aux données – CASD). I am grateful to Joy Feng and Jimin Lee for excellent research assistance.

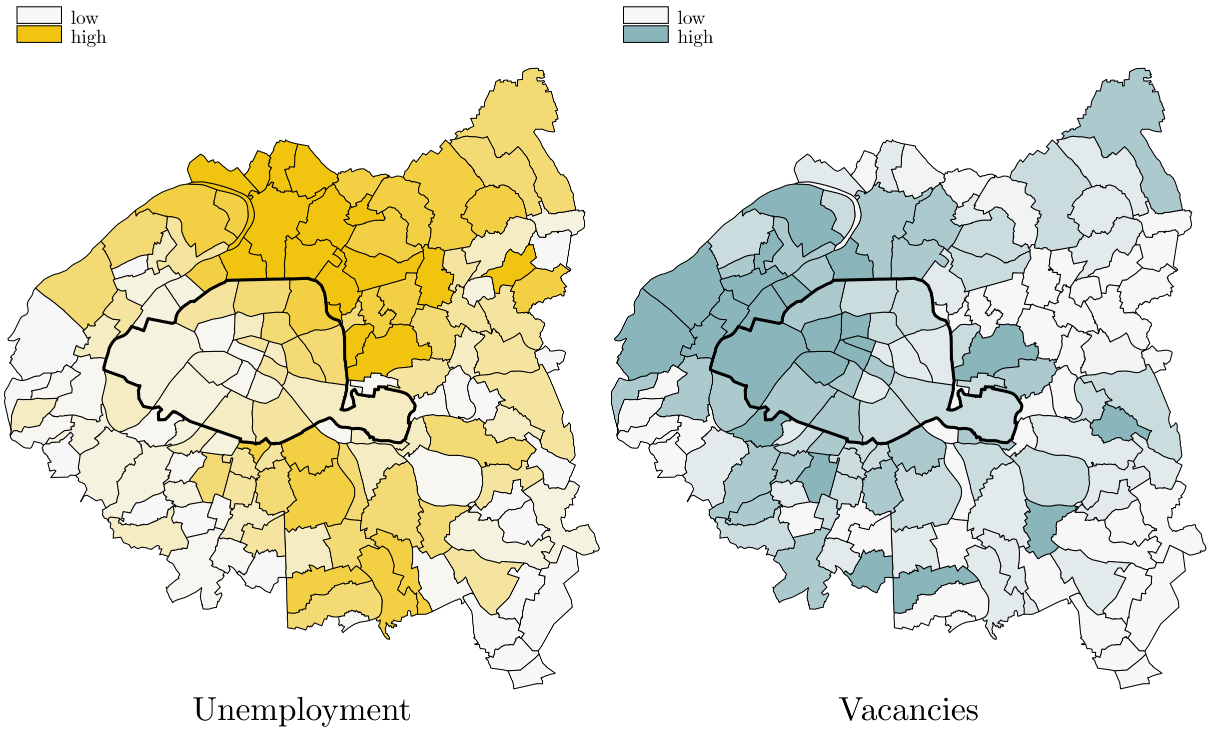


Figure 1.1.1: Distributions of job seekers and vacancies in Paris and its suburbs

Notes. Both variables are scaled by the local labor force. The area outlined in black is Paris.

Source. RP 2006 and ACEMO 2007.

The mechanics of spatial mismatch are illustrated in a stripped-down version of the model highlighting the structural parameters pivotal to the welfare analysis and how to identify them empirically. This toy model is a two-location city with frictional labor markets à la Diamond [1981] and Mortensen and Pissarides [1994]. The city structure interacts with labor markets through two channels: commuting costs, which undermine the surplus of a match, and information decay, that diminish job finding rates. Workers, who are homogenous, choose where to live while recruiters choose where to open vacancies. In equilibrium, there might be *too much* or *too little* spatial mismatch depending on whether commuting costs or information decay dominates. Constrained efficiency may be achieved by a mix of place-based residence and hiring subsidies. In particular, if there is too much spatial mismatch, those subsidies are interpreted as moving-to-opportunity and enterprise zone programs that bring firms and workers together. The structural parameters governing commuting costs and information decay can be identified in the data from the wage premium for commuting and the home bias in employment respectively.

In the decentralized equilibrium, both workers and jobs are misallocated: they are too far apart if commuting costs are large or too close together if information decay is predominant. The benchmark here is the second-best allocation that the planner would achieve subject to commuting costs and information decay. On the worker side, job seekers live too far (resp. too close) from productive locations if commuting costs (resp. information decay) dominate because of a holdup problem in the choice of residence. Indeed, residence choice may be interpreted as an investment in commuting and information access. However, a worker does not capture the full return of his investment because the surplus generated by a match is shared with recruiters through Nash bargaining. Importantly, the *pooling*[2] of workers from different residences searching in a same workplace prevents the local labor market tightness from adjusting and offsetting low wages against high employment. Therefore, job seekers underinvest in commuting access, causing inefficiently low expected surplus of jobs, and overinvest in information access, causing inefficiently high outside options. Symmetrically, on the job side, vacancies are disproportionately created near productive places if commuting costs (resp. information decay) prevail. The intuition is that because of information decay, recruiters in those locations tend to match with workers from which they extract high rents. In a nutshell, there is too much spatial mismatch—workers and jobs are too distant to each other—if commuting costs dominate, and too little spatial mismatch—workers and jobs are too close to each other—if information decay does.

Poor French suburban neighborhoods—the so-called *banlieues*[3]—offer a convenient setting to quantify the welfare attributes of spatial mismatch for two reasons. First, because these areas are particularly exposed to spatial mismatch. They were conceived in the 1950s and 1960s according to an urban planning principle popular at the time—that living centers should be separate from working centers. After the end of postwar economic growth, unemployment rates skyrocketed in these isolated areas. They are now seen as urban ghettos. Second, because the variation induced by an enterprise zone program targeting these neighborhoods—the ZFU policy—facilitates the identification of spatial mismatch.

To assess empirically the extent and the consequences of spatial mismatch, I develop a quantitative framework that extends the toy model by integrating several pertinent realistic features while retaining its key welfare properties. To enable estimation, the quantitative version is comprised of an arbitrary number of neighborhoods and skill groups. It also incorporates two new forces that are relevant for the quantitative analysis: endogenous job acceptance, that directly links job finding rates to commuting costs, and idiosyncratic preferences for locations, that mirror imperfect residential mobility of workers. As in the toy model, the decentralized equilibrium features too much spatial mismatch if commuting costs predominate, and too little if information decay prevails.

[2] This mechanism was first described in Acemoglu [2001]. The term *pooling externalities* was coined by Bilal [2020], who applies them in a spatial context to rationalize enterprise zones.

[3] In France, a *banlieue* is a suburb of a large city. By extension, this term refers low-income housing projects built in the 1950s and 1960s in which mainly immigrants and French of foreign descent reside. Throughout the paper, I will refer to this second meaning when using the word *banlieue*.

The structural estimation of the quantitative model of spatial mismatch involves two main empirical challenges, namely unobserved job matching rates and unobserved skill heterogeneity. Following the comparative statics of the toy model, commuting costs are recovered from the sensitivity of wages to commute times, information decay from the sensitivity of employment shares to commuting times and population mobility from the sensitivity of population to local wages and unemployment rates. A first difficulty is that one cannot infer job *matching* rates from the data, but only job *finding* rates. That is, it is impossible to observe job offers that a worker turns down. Thus, to identify information decay, one needs to adjust for the expected job acceptance given wage level and commuting costs. A second concern is that unobserved skill characteristics of workers might correlate with their ability to commute. I will leverage the exogenous variation in commuting patterns provided by the ZFU policy to estimate credibly model workers' mobility.

In subsequent work, I will apply my model to quantify the welfare implications of spatial mismatch. Armed with my estimated structural model, I will test the spatial mismatch hypothesis in the context of French banlieues. I first plan to assess the extent of spatial mismatch with a model-based measure of job proximity. The arrival rate of job offers of a local unemployed worker provides a natural measure of job proximity which captures the extent of spatial mismatch. Second, I will estimate the employment and welfare consequences of spatial mismatch. Lastly, I will apply the calibrated model to evaluate real-world and counterfactual policies designed to remedy the detrimental effects of spatial mismatch.

Related literature. My paper combines several strands of a vast literature at the intersection of labor, spatial and public economics.

This paper builds on earlier theoretical models of spatial mismatch[4]. The introduction of frictional labor markets into monocentric and duocentric city models rationalized local variations in unemployment and wages by the interaction of housing and labor markets, even absent skill heterogeneity [Brueckner and Martin, 1997; Brueckner and Zenou, 1999, 2003; Zenou, 2009a,b,c,e]. My paper complements this literature by examining the welfare attributes of spatial mismatch and characterizing the policies that would restore efficiency.

My approach combines the insights of theoretical models of spatial mismatch with techniques borrowed from the quantitative urban literature. Early papers in this tradition have underlined the importance of spatial linkages in determining the local response to shocks [Redding and Sturm, 2008; Ahlfeldt et al., 2015; Monte et al., 2018] and in shaping the distribution of economic activity [Allen and Arkolakis, 2014]. More recently, these frameworks have been applied to study the welfare consequences of various urban policies [Allen et al., 2015; Allen and Arkolakis, 2019; Fajgelbaum and Schaal, 2019; Fajgelbaum and Gaubert, Forthcoming; Tsivanidis, 2019; Heblich et al., 2020]. By introducing frictional labor markets into a quantitative urban framework, I bridge the gap between theoretical and the empirical literatures of spatial mismatch. This actionable model allows

[4]See Zenou [2009d] for an exhaustive review.

to test the spatial mismatch hypothesis in the data, to measure its consequences and to evaluate the different policy options available to overcome its effects.

The theoretical basis for place-based policies has been the subject of several recent papers, all focusing on interventions at the local labor market level. Glaeser and Gottlieb [2008], Kline and Moretti [2014], and Kline and Moretti [2014] argue that there is little efficiency ground for place-based policies as agglomeration benefits from subsidizing one location are offset by deagglomeration costs elsewhere. However, Fajgelbaum and Gaubert [Forthcoming] show that there is room for efficient place-based policies with constant spillover elasticities if cross-type spillovers between workers generate inefficient spatial sorting. Gaubert et al. [2020] investigate redistribution as an alternative motive for place-based policies. Most closely related in this literature are Kline and Moretti [2013] and Bilal [2020], that both embed a frictional labor market model à la Diamond [1981] and Mortensen and Pissarides [1994] into a geographic framework. Specifically, Kline and Moretti [2013] show that place-based policies may be rationalized by hiring subsidies implementing the Hosios condition, while Bilal [2020] rationalizes enterprise zones at the level of the local labor market with pooling externalities. Although in practice most enterprise zone programs target derelict neighborhoods within metropolitan areas, this literature has focused on rationales valid at the local labor market level. My paper fills this gap by analyzing a purely urban motive—spatial mismatch—where inefficiencies pertain to the intertwining of housing and labor markets within cities, and focusing on commuting policies.

On the empirical side, most closely related is the body of work assessing empirically the scope of spatial mismatch, which has suffered from several methodological shortcomings[5]. Early papers relied on the comparison of employment, earnings and commuting times of workers living in different parts of the city [Ihlanfeldt and Sjoquist, 1989; Ihlanfeldt and Young, 1994; McLafferty, 1997; Zhang, 1998]. Resulting estimates were hard to interpret as those quantities are endogenously determined. Later papers proposed measures of job proximity to estimate directly the effect of spatial mismatch on local labor market outcomes [Hanson et al., 1997; Rogers, 1997; Immergluck, 1998; Hellerstein et al., 2008]. Regressing wages and unemployment on inaccurate proxies of proxies of spatial mismatch while controlling for well-measured skill composition introduces systematic biases against finding evidence in support of the hypothesis. Indeed, the measures of job proximity that have been used so far are ad hoc. They omit competing labor, rest on jobs rather than vacancies, often oversee the skill segmentation of labor markets and embed rough measurement of distance that do not take into account transport mode. A few more recent papers exploit spatial experiments to estimate the effect of isolation on employment [Popkin et al., 1993; Rosenbaum, 1995; Gore and Herrington, 1997; Rosenbaum and Harris, 2001; Houston, 2001; Miller, 2018; Andersson et al., 2018]. While this strategy provides credible evidence of the effect of spatial isolation on labor market outcomes, it is silent on the extent of spatial mismatch for urban ghettos, as well as on its overall employment and welfare costs.

[5]See Houston [2005] for a critical review of this literature.

By adopting a structural approach, my paper offers a solution to the three main methodological problems encountered by this literature. First, I propose a novel model-based measure of job proximity that accounts properly for competing labor, vacancies, skill segmentation and transport modes. Second, I can purge measures of job proximity from measurement error by using the values determined in equilibrium within the model rather than plugging in the observed variables. Third, the general equilibrium model clarifies the role of endogenous quantities such as transport mode, and allow to distinguish between the extent of spatial mismatch—captured by job proximity—and its consequences on employment and welfare—that depend on endogenous commuting and migration responses.

This paper finally contributes to the vast empirical literature evaluating the impact of urban public policies, in particular enterprise zones and moving-to-opportunity programs. For enterprise zones, findings range between no effect and substantial positive effects on employment. Papers studying large-scale programs at the federal level or in European countries tend to find substantial employment gains in the selected areas [Rathelot and Sillard, 2008; Busso et al., 2013; Kline and Moretti, 2014; Criscuolo et al., 2019]. In contrast, several papers conclude no positive impact of state-level programs in the U.S. context [Bondonio and Engberg, 2000; Bondonio and Greenbaum, 2007; Elvery, 2009; Neumark and Kolko, 2010; Lynch and Zax, 2011]. While those papers take employment growth as the unique criterion of program success, I argue that this gauge is insufficient and even misleading to grasp welfare effects. In parallel, evidence of the effect of moving-to-opportunity programs on adults' labor market outcomes is ambiguous, with most reliable experimental studies concluding that the impact was limited [Harding et al., 2021]. Bergman et al. [2019] [TBD]

The effects of ZFU policy on labor market outcomes have already been studied in several papers, all of which adopt reduced-form approaches [Rathelot and Sillard, 2008; Gobillon et al., 2012; Givord et al., 2013; Briant et al., 2015; Mayer et al., 2015; Givord et al., 2018]. Most closely related is Briant et al. [2015] that shows that the success of the zones crucially depends on spatial integration of targeted areas. Shifting away from labor market outcomes, Poulhès [2015] highlights that ZFUs had a positive impact on commercial property values. Adopting a structural approach allows me evaluate the welfare consequences ZFU policy, with a specific angle—the success in overcoming spatial mismatch.

Layout. The rest of the paper is organized as follows. I illustrate the mechanics of spatial mismatch in section 1.2. I describe the institutional setting and the data that I exploit for my empirical analysis in section 1.3. In section 1.4, I expose a quantitative urban model of spatial mismatch, which I estimate structurally in section 1.5. Finally, section 1.6 concludes.

1.2 Toy Model: The Inefficiencies of Spatial Mismatch

I present a stripped-down version of the model to illustrate the inefficiencies involved by spatial mismatch and the adequate policy responses. I define the channels through which the urban structure interacts with labor markets, namely commuting costs and information decay. I demonstrate how to disentangle them in the data and I clarify their role in the welfare analysis.

The main theoretical result of this section is that the balance between those two forces determines whether there is too much or too little spatial mismatch in the decentralized equilibrium: spatial mismatch is inefficiently high if commuting costs prevail, and inefficiently low otherwise. The misallocation of both workers and jobs stems from a hold-up problem in workers' choice of residence. Underlying this inefficiency is the *pooling*[6] of workers living in different residences while searching in a same workplace. Constrained efficiency may be restored by a mix of residence and hiring subsidies. In particular, if there is too much spatial mismatch, those subsidies can be interpreted as moving-to-opportunity and enterprise zone programs that bring jobs and workers together.

1.2.1 A Duocentric City Model with Frictional Labor Markets

I lay out the main ingredients of a two-location urban model with frictional labor markets. Because neighborhoods are distant, commuting is costly and information about job offers decays across locations. This stylized framework is nested in the full-fledged quantitative model developed in section 1.4.

1.2.1.1 Setup

The city geography is comprised of two neighborhoods, the employment center C and the ghetto G. Productivity is intrinsically higher in C than in G[7].

Homogeneous workers choose their residence between C and the G. They are alternatively employed and unemployed. When unemployed, they search for jobs in both locations. Importantly, I assume that employed workers can't keep their job should they move out to live in another neighborhood. The city is assumed to be closed, meaning that the mass of workers is fixed.

On the production side, recruiters open vacancies in each location. They sell labor inputs to local producers who combine them with intermediate goods. Recruiters enter freely up to zero profit in each location.

Locations C and G are distant from each other. Workers incur a monetary cost when commuting to the other location. Distance is also responsible for information decay hindering job search across locations.

[6] This mechanism was first described in Acemoglu [2001]. The term *pooling externalities* was coined by Bilal [2020], who applies them in a spatial context to rationalize enterprise zones.

[7] Neighborhoods C and G have opposing interpretations in the U.S. and in France. In the U.S., C would be the suburbs and G the inner city. In France, C would be the city center and G the suburbs.

1.2.1.2 Preferences

Instantaneous utility. Each worker consumes tradable goods and housing, and values local amenities. His utility is equal to:

$$a_i u(c, h), \tag{1.2.1}$$

where c and h are the respective amounts of tradable good and housing consumed, and a_i are amenities in location i.

Commuting costs. Commuting costs between residence i and workplace j are denoted d_{ij}. They are paid by employed workers and modeled as a tradable good expense.

I assume that a worker who works in the location where he lives does not incur any cost to commute, and that commuting costs between C and G are symmetric. They are parametrized as follows:

$$d_{ij} = \begin{cases} \bar{d} & \text{if } i \neq j, \\ 0 & \text{if } i = j, \end{cases} \tag{1.2.2}$$

where $\bar{d} \in [0, \infty)$ captures the extent of commuting costs.

1.2.1.3 Technology

Producers. A continuum of identical producers assemble labor inputs into consumption goods in each location. Their constant-returns-to-scale technology Y is common to both locations:

$$Y_j = y_j Y\Big(\sum_{i \in \{C,G\}} L_{ij}, M_j^Y \Big), \tag{1.2.3}$$

with y_j location-specific productivity, L_{ij} the mass of workers living in i and working in j and M_j^Y are intermediate goods. I assume that the employment center is intrinsically more productive than the ghetto:

$$y_C > y_G. \tag{1.2.4}$$

For simplicity, I assume that goods are freely traded across locations[8].

Recruiters. Recruiters post vacancies in each local labor market. Maintaining an open vacancy involves a flow cost, ν, that captures advertisement and other search costs on the side of the recruiter and is common to both locations. They sell labor inputs to producers located in neighborhood j.

[8] While the welfare results still hold with iceberg trade costs, the assumption of free trade eases the proof of comparative statics results.

Information decay. Information about job offers decays with distance, hindering employment search. Information decay between residence i and workplace j is parametrized as:

$$\Delta_{ij} = \begin{cases} \bar{\Delta} & \text{if } i \neq j, \\ 1 & \text{if } i = j, \end{cases} \tag{1.2.5}$$

where $\bar{\Delta} \in [1, \infty)$ captures the extent of information decay induced by distance. There is no information loss for job offers originating from a worker's own residence, so that $\Delta_{ii} = 1$.

Matching. Job seekers and recruiters match on workplace-specific labor markets. The matching technology exhibits constant returns-to-scale. The flow rate of matches is equal to:

$$m_{g,j} = M\Big(\sum_i \Delta_{ij}^{-1} U_i, V_{g,j}\Big), \tag{1.2.6}$$

where U_i is the mass of job seekers living in location i and V_j is the number of vacancies open in location j.

Matches split when a negative productivity shock occurs. Those shocks follow a Poisson process with arrival rate δ.

Developers. Developers provide housing to households in both locations. The presence of a fixed factor, land, induces decreasing returns to scale. Housing supply is given by:

$$H_i = H_i(M_i^H), \tag{1.2.7}$$

where M_i^H represents intermediate goods bought from producers. H_i has constant or decreasing returns-to-scale, which captures the possible presence of a fixed factor, land, used in the production of housing.

1.2.1.4 Closing the Model

Closed city and free entry. The city is assumed to be closed, so that the mass of workers in the city are exogenously fixed. Recruiters, however, enter freely up to the point where the value of entering is zero in each location.

Wage setting and market structure. Matches between recruiters and job seekers generate an intertemporal surplus that is shared through a generalized Nash bargaining process. The total surplus is the sum of the surplus of workers and the surplus of recruiters. The bargaining power of workers is $\beta \in [0, 1]$.

Input and good markets are perfectly competitive and all agents are price-takers.

Unemployment benefits. Unemployed workers receive unemployment benefits b that are formally assumed to be derived from home production. They are constant across locations.

Residential mobility. In this simple model, I assume that employed workers can't keep their jobs if they move out to live in another neighborhood.

Ownership. All workers, irrespective of their employment status or residence, capture the same share of land rents and recruiters' profits.

1.2.2 Comparative Statics: Detecting the Channels of Spatial Mismatch

The spatial structure of the city interplays with labor markets through two distinct channels, commuting costs and information decay. In this simple model, they induce a dichotomy between wages and employment patterns: the former are determined by commuting costs, while the latter are governed by information decay. This section formalizes those positive statements, which allow to disentangle the two forces in the data.

The first proposition describes the effect of commuting costs and information decay on wages and employment.

Proposition 1.1 (Comparative Statics).

1. The higher $\bar{d}$, the higher the wage premium for commuting conditional on residence:

$$\frac{\partial\, w_{GC} - w_{GG}}{\partial\, \bar{d}} > 0, \quad \frac{\partial\, w_{CG} - w_{CC}}{\partial\, \bar{d}} > 0,$$

   where w_{ij} is the wage of workers living in i and employed in j.

2. The higher $\bar{\Delta}$, the higher the home bias in employment:

$$\frac{\partial\, l_{CC}}{\partial\, \bar{\Delta}} > 0, \quad \frac{\partial\, l_{GG}}{\partial\, \bar{\Delta}} > 0,$$

   where l_{ij} is the share of workers employed in j among residents of i.

The intuition for the first point runs as follows. Because of Nash bargaining, residence influences wages through two mechanisms: partial compensation of commuting costs and outside options. The first one, which increases with commuting costs, shows up in cross-workplace wage differentials. The second one affects workers' wages the same wherever they work, so it cancels out when considering the wage premium for commuting.

The second property means that parameter $\bar{\Delta}$ drives the home bias in employment. Indeed, a higher value of $\bar{\Delta}$ implies that job seekers receive relatively less job offers from the other location. One could additionally prove that a stronger home bias also entails that unemployment is relatively higher in G than in C. The reason is that labor market is tighter in C due to difference in intrinsic productivity. Therefore, a stronger home bias also means that unemployment is relatively higher in G than in C.

Note that by assumption commuting costs do not affect directly employment shares in the toy model. Indeed, workers' behavioral response to commuting costs is ruled out, as all job offers are accepted. The quantitative model laid out in section 1.4 relaxes this hypothesis by introducing stochastic job matchings and allowing workers to turn down job offers if the wage net of commuting costs is too low.

Proposition 1.1 suggest a strategy to identify $\bar{d}$ and Δ in the data: $\bar{d}$ from the wage premium for commuting, $\bar{\Delta}$ from the home bias in employment. The estimation of the quantitative model structural parameters in section 1.5 rests on this idea. In the real world, the positive correlation between wages and commute time is partly explained by the fact job seekers turn down low-wage offers from distant employers. This alternative mechanism is incorporated in the full-fledged quantitative model of section 1.4 and requires an adjustment for wage and commuting costs in the estimation of $\bar{\Delta}$.

1.2.3 Welfare Analysis: Too Much or Too Little Spatial Mismatch?

Spatial mismatch entails distortions along two margins: workers' choice of residence and vacancy creation. In the presence of commuting costs and information decay, the decentralized equilibrium is generically inefficient: It features too much spatial mismatch if commuting costs prevail, and too little if information decay dominates. The constrained-efficient allocation may be restored with a mix of place-based residence and hiring subsidies that bring firms and workers closer together in the first case, or keep them away in the second one.

Let me first define the social welfare, $\mathcal{W}$. The objective of the planner is to maximize the expected utility of unemployed and employed workers discounted over time:

$$\mathcal{W} = \int_0^\infty e^{-rt} \sum_{i\in\{C,G\}} L_i \Big[u_i \, a_i u(c_i^U, h_i^U) + \sum_{j\in\{C,G\}} l_{ij} \, a_i u(c_{ij}^E, h_{ij}^E) \Big] dt. \tag{1.2.8}$$

I now lay out succinctly the planning problem, also described in more details in appendix 1.B.

Definition 1.1 (Planner's Problem). The planner's problem is:

$$\max \ \mathcal{W}, \tag{1.2.9}$$

subject to:

(i) spatial mobility constraints;

(ii) tradable good and housing feasibility constraints;

(iii) search and matching constraints;

(iv) labor market clearing;

(v) population constraint.

In this simple environment, two margins may give rise to inefficiencies: workers' choice of residence and recruiters' choice of workplace. Correspondingly, two conditions characterize the constrained-efficient allocation that the planner would achieve subject to the search and matching constraints. They are depicted in the proposition below.

Proposition 1.2 (Constrained-Efficient Allocation). The constrained-efficient allocation solution to the planner's problem 1.1 satisfies the following two conditions.

1. Workers' optimal choice of residence:

$$\omega a_i\Big[u_i u(c_i^U, h_i^U) + \textstyle\sum_j l_{ij} u(c_{ij}^E, h_{ij}^E)\Big] + \sum_j l_{ij} W_j - u_i \sum_j \Delta_{ij}^{-1}\theta_j \nu = u_i x_i^U + \sum_j l_{ij} x_{ij}^E + \Lambda, \tag{1.2.10}$$

   with W_j and θ_j the price of labor inputs and the labor market tightness in location j, x_i^U and x_{ij}^E the expenditures of unemployed and employed workers, ω the Pareto weight and Λ the opportunity cost of a worker.

2. Recruiters' optimal entry:

$$\nu \textstyle\sum_i \Delta_{ij}^{-1} u_i L_i = \frac{1-\mu(\theta_j)}{\theta_j} \sum_i \frac{\lambda_{ij}}{r+\delta+\lambda_i} L_i \Big(u_i\big[W_j - d_{ij} - b + \sum_{j'} \Delta_{ij'}^{-1}\theta_{j'}\nu\big] + \frac{\delta}{r+\delta}\sum_{j'} l_{ij'}\big[W_j - d_{ij} - \big(W_{j'} - d_{ij'}\big)\big]\Big), \tag{1.2.11}$$

   with $\lambda_{ij} = \Delta_{ij}^{-1}\theta_j q(\theta_j)$ the job finding rate in workplace j of a job seeker living in i and $\mu(\theta_j)$ the elasticity of the matching function m with respect to unemployment.

Condition (1.4.30) states that the marginal benefits of an additional worker, minus his marginal costs, must be equalized across residences in the second-best allocation. Benefits comprise the expected utility of a marginal worker, magnified by the Pareto weight ω, and his expected product of labor. Costs include the additional costs of posting vacancies for recruiters and the expected expenditure of the marginal worker. Here, expenditure appears as a cost because of the non-separability between workers' locations and consumption: an additional worker in location i translates into lower consumption of commodities for other workers. Finally, constant Λ is interpreted as the opportunity cost of a worker. It captures the net gains that the marginal worker would yield in other locations. Workers' optimal choice of residence condition is reminiscent of Fajgelbaum and Gaubert [Forthcoming].

The second condition, (1.4.31), characterizes recruiters' optimal entry. It equates the marginal benefits, net of marginal costs, of opening a vacancy workplace by workplace. The left-hand side represents the costs of the direct cost of increasing labor market tightness in θ_j—which is composed of the monetary fee of maintaining a vacancy open, ν, and the number of additional vacancies $\sum_i \Delta_{ij}^{-1} u_i L_i$. The right-hand side reflects the benefits from an additional vacancy in j: decreased costs of unemployment, lower vacancy costs in other locations j' and gains from reallocating workers from j' to j.

Next proposition characterizes the efficiency of the decentralized equilibrium, which may feature too much or too little spatial mismatch.

Proposition 1.3 (Efficiency of the Decentralized Equilibrium). Assume the Hosios condition $\mu(\theta_j) = \beta$, with β workers' bargaining power. Then there exists a increasing function $\bar{\Delta}_0 : \bar{d} \mapsto \bar{\Delta}_0(\bar{d})$ satisfying $\bar{\Delta}_0(0) = 0$ and such that:

1. If $\bar{d} \geq 0$ and $\bar{\Delta} = \bar{\Delta}_0(\bar{d})$, the decentralized equilibrium coincides with the constrained-efficient allocation;
2. If $\bar{d} > 0$ and $\bar{\Delta} < \bar{\Delta}_0(\bar{d})$, the decentralized equilibrium features too much spatial mismatch, in the sense that $\frac{L_C^{DE}}{L_G^{DE}} < \frac{L_C^*}{L_G^*}$ and $\frac{\theta_G^{DE}}{\theta_C^{DE}} < \frac{\theta_G^*}{\theta_C^*}$.
3. If $\bar{d} \geq 0$ and $\bar{\Delta} > \bar{\Delta}_0(\bar{d})$, the decentralized equilibrium features too little spatial mismatch, in the sense that $\frac{L_C^{DE}}{L_G^{DE}} > \frac{L_C^*}{L_G^*}$ and $\frac{\theta_G^{DE}}{\theta_C^{DE}} > \frac{\theta_G^*}{\theta_C^*}$.

Intuitively, job seekers live too far (resp. too close) from productive locations if commuting costs (resp. information decay) dominate because of a holdup problem in the choice of residence. Indeed, residence choice may be interpreted as an investment in commuting and information access. However, a worker does not capture the full return of this investment because the surplus of a match that is shared with recruiters through Nash bargaining. Importantly, the *pooling*[9] of workers from different residences searching in a given workplace prevents the local labor market tightness from adjusting. Therefore, job seekers underinvest in commuting access, which raises the expected surplus of a job, and overinvest in information access, which enhance outside options.

Symmetrically, on the job side, too many (resp. too few) vacancies are created near productive places if commuting costs (resp. information decay) dominate. The intuition is that recruiters in those locations tend to match with workers from which they extract high rents because of information decay. In a nutshell, there is too much (resp. too little) spatial mismatch, i.e. workers and jobs are too distant (resp. too close) to each other, if commuting costs (resp. information decay) dominate.

The theoretical possibility of too little spatial mismatch might seem puzzling. Analogously to the social role of unemployment in the standard DMP model, spatial mismatch's role is to facilitate trade at low transaction costs in the presence of information decay. The greater is spatial mismatch is, the less costly this is to fill vacancies because workers' outside options are lower.

The last proposition of this section characterizes the optimal place-based subsidies that restore efficiency in the decentralized equilibrium.

Proposition 1.4 (Optimal Policy). Constrained efficiency may be restored with a mix of two

[9] This mechanism was first described in Acemoglu [2001]. The term *pooling externalities* was coined by Bilal [2020], who applies them in a spatial context.

place-based policies, a residence subsidy, t_i^R, and a hiring subsidy, t_j^H, that can be expressed as:

$$t_i^R = \sum_{j\in\{C,G\}} \frac{\lambda_{ij}}{r+\delta+\bar{\lambda}_i}\Big(\pi_{ij} - \mathbb{E}_{|j}\left[\pi_{i'j}\right] - t_j^H\Big) \tag{1.2.12}$$

$$= -(1-\beta)\sum_{j\in\{C,G\}} \frac{\lambda_{ij}}{r+\delta+\bar{\lambda}_i}\Big(d_{ij} + \Omega_i - \mathbb{E}_{|j}\left[d_{i'j} + \Omega_{i'}\right] - t_j^H\Big) \tag{1.2.13}$$

and:

$$t_j^H = -\mathbb{E}_{|j}\left[t_{i'}^R\right], \tag{1.2.14}$$

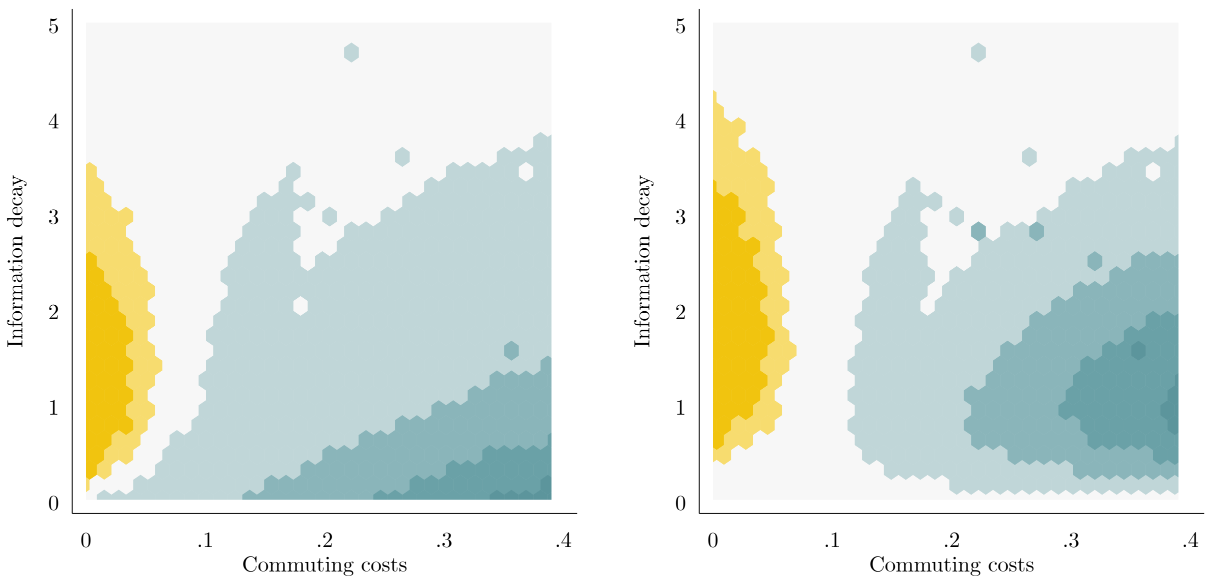
where $\Omega_i = \frac{(r+\delta)b + \sum_j \lambda_{ij}(w_{ij}-d_{ij})}{r+\delta+\bar{\lambda}_i}$ is the outside option of a worker living in i, $\pi_{ij} = (1-\beta)(W_j - d_{ij} - \Omega_i)$ is the profits of a recruiter in j when matched with a worker living in i and $\mathbb{E}_{|j}\left[X_{ij}\right] = \frac{\sum_{i'} \Delta_{i'j}^{-1} U_i X_{i'j}}{\sum_{i'} \Delta_{i'j}^{-1} U_{i'}}$ is the expected value of $X_{i'j}$ over matches for a recruiter in j.

Formula (1.2.12) shows that to achieve the constrained-efficient allocation, the government needs to subsidize locations where workers tend to yield higher profits to recruiters than their coworkers. Equivalently, it must subsidize locations where workers either incur lower commuting costs, or have lower outside options. Factor $1-\beta$ in equation (1.2.13) reflects the holdup problem which disappears as the bargaining power of workers tends to 1, implying full capture of gains from residence investment.

Residence subsidies $(t_i^R)_i$ determine whether there is too much or too little spatial mismatch. $t_C^R = 0$ corresponds to the constrained-efficient case $\bar{\Delta} = \bar{\Delta}_0(\bar{d})$, $t_C^R > 0$ means that there is too much spatial mismatch and $t_C^R < 0$ too little.

Two instruments will generically be needed to restore constrained-efficiency. The reason is that the government needs to move the spatial distribution of workers and recruiters in opposite directions, either to bring them together or to keep them away. One exception is when $\bar{\Delta} = 0$: Then, job creation is optimal conditional on the residence of workers and $t_j^H = 0$ for $j \in \{C, G\}$. In particular, if there is too much spatial mismatch, those subsidies can be interpreted as moving-to-opportunity and enterprise zone programs that bring firms and workers together.

Figure 1.2.1 illustrates the influence of parameters d and Δ on the sign and the magnitude of place-based subsidies. There tend to be too much spatial mismatch when d is large as compared to Δ, implying positive values of MTO subsidies in C. Those areas appear in darker shades of blue on subfigure 1.2.1a. When Δ is positive, those areas also correspond to larger EZ subsidies in G. When Δ is close to zero, residence subsidies in C are enough to restore efficiency. On the contrary, there is too little spatial mismatch when Δ is large as compared to d, resulting in the opposite policies. However, when Δ is very high, there is no more pooling and place-based subsidies are close to zero in all locations. This *geographic* environment without commuting corresponds to the white space at the top of both figures.



(a) Residence subsidies in C

(b) Hiring subsidies in G

Figure 1.2.1: Place-based subsidies as a function of commuting costs and information decay

Notes. Yellow and blue areas correspond respectively to negative and positive subsidies.

Source. Simulations of the toy model for an array of values of $\bar{d}$ and $\bar{\Delta}$.

1.3 Background and Data

In this section, I provide background on the empirical application of this paper: the French *banlieues*. The administrative data available in France allow to examine precisely these poor suburban neighborhoods subject to spatial mismatch.

1.3.1 Background

By their very conception, the French *banlieues* are particularly exposed to spatial mismatch: they were purposefully built away from employment centers. To undo its detrimental effects, an enterprise zone program—the so-called ZFU policy—was implemented in the 1990s.

1.3.1.1 Detrimental Effects of Spatial Mismatch: The Case of French Banlieues

The French *banlieues* are densely-populated suburban areas built in the postwar decades. Their original conception—as living centers separate from working and commercial centers—made them vulnerable to spatial mismatch. Nowadays, they are considered as urban ghettos.

The banlieues are comprised of high-rise apartment blocks hosting a dense population. They were built in the 1950s and 1960s to address the severe housing shortage hitting postwar France. The crisis resulted from the confluence of World War II destruction and of the steady growth of urban population—the product of the baby boom, rural exodus and immigration. To relieve rapidly the pressing need for housing, French government initiated the construction of huge housing projects financed in part by the Marshall Plan and organized through central planning. In the following years, monolithic concrete apartment blocks were erected in the suburbs of large cities across the country.

The urban planning principles that guided the conception of these housing estates made them prone to spatial mismatch. The concept of a *functional city* had been popularized by Le Corbusier [Le Corbusier, 1943]. The architect was advocating the split of cities into three separate centers defined by their function and connected by buses: the living center, the commercial center and the working center. This utopian city inspired the construction of the housing estates—the living centers—in remote suburban areas, away from working centers.

Far from utopian aspirations, the banlieues eventually became urban ghettos. Middle-class households initially occupying the newly-built apartment blocks were soon replaced by immigrants from the former colonies, chiefly from Maghreb. These young workers were encouraged to migrate by the French state and industrials to fill labor shortages. However, after the end of postwar economic growth, unemployment rates skyrocketed in the isolated banlieues. Those neighborhoods are now seen as poverty traps.

1.3.1.2 Undoing Spatial Mismatch: The ZFU Policy

To tackle the detrimental effects of spatial mismatch in French banlieues, policymakers have tried to bring firms into the banlieues by implementing enterprise zone program: the ZFU policy. The induced variation will prove useful to estimate the quantitative model developed in section 1.4.

The ZFU policy is an enterprise zone program enacted in the mid-1990s to break the spatial isolation of French banlieues. In 1996, the Prime Minister Alain Juppé announced the creation of about thirty urban tax-free zones—*Zones Franches Urbaines* or ZFUs in French—as the flagship measure of the so-called *National Plan of Urban Integration*. The program would grant tax credits to firms settling in selected urban areas. It was explicitly targeting the most deprived banlieues.

The exemptions awarded to firms in ZFUs were substantial[10]. The tax package covered the four main taxes paid by corporations in France: the corporate tax, the payroll tax, the local business tax and the real estate tax. The involved exemptions were total for five years, before phasing out over three to nine additional years. Those tax credits were particularly generous, as the corporate tax amounted typically to one third of profits and the payroll tax to about 30% of labor costs at the time.

The tax arrangement was extended over years to nearly one hundred zones that account for more than one million inhabitants. Three generations of ZFUs were successively designated: 38 in 1997, 41 in 2004, and 15 in 2006. As depicted on figure 1.3.1, the ZFUs are spread out over France, with a concentration in Paris region.

[10] See appendix **??** for more details.

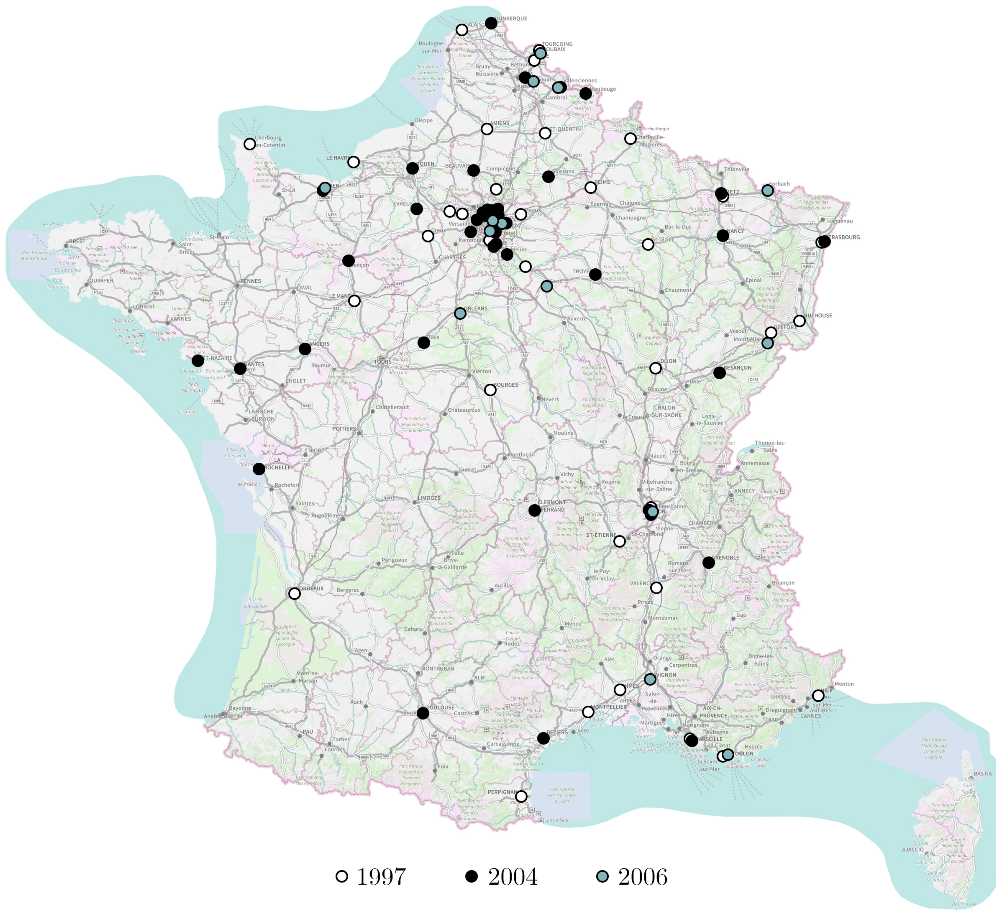


Figure 1.3.1: Map of the ZFUs

Notes. White dots represent the 1st generation of ZFUs (1997), black dots the 2nd generation (2004) and blue dots the 3rd generation (2006).

The ZFU program was designed to bring jobs to the local banlieue population, as can be seen from two distinctive features. First, the payroll tax exemptions only apply to low-wage jobs. Those jobs are the most likely to be occupied by local inhabitants. Second, tax credits were conditional on employing at least 30% of ZFU residents.

The selection process of ZFUs provides a natural group of controls—the ZRUs—to identify empirically their impact. In 1996, the government designated 451 *Zones de Redynamisation Urbaine* (urban renewal zones) as the most distressed French neighborhoods. The three generations of ZFUs were picked among ZRUs. Although ZRUs which did not become ZFUs also benefited from tax exemptions, those were substantially less generous than in ZFUs: their duration was much shorter and their scope limited to the only newly-created firms and jobs.

Two publicly-disclosed criteria were used to select ZFUs among ZRUs. First, the population of ZFUs had to exceed a threshold set to 10,000 for the first generation and reduced to 8,500 for the subsequent ones. Second, the level of priority among the eligible zones was determined by a synthetic index accounting for both the level of deprivation of the zone itself and the capacity of the surrounding city to assist it financially. Specifically, the synthetic index depended positively on zone-level unemployment rate, share of people under 25 and share of people over 15 without any qualification, while being adjusted for the tax potential[11] of the surrounding city and the presence of other distressed urban areas.

1.3.2 Data

This subsection describes the main datasets used in the analysis.

Employment and job search. Panel data from the FH-DADS combines a matched employer-employee database derived from employer's social contribution reports with records from the French unemployment insurance agency, which allows to follow workers over both their employment and unemployment spells. Regarding employment spells, the matched employer-employee data contain information about each worker's income, number of days and hours worked, occupation and industry. Regarding unemployment spells, the unemployment agency records the expected salary and maximal commuting distance reported by the job seeker, as well as all matches with firms that occurred through the agency. They also survey the channels through which job seekers search jobs. The FH-DADS data also provide basic demographic information such as sex, age, citizenship and education. The panel starts in 1988 and follows 1/24th[12] of the labor force, and covers therefore the full period of interest for a large sample of workers.

To track the geographic location of establishments, I use the publicly-available business registrar, *Sirene*, which compiles information on all firms and their establishments. Besides the firm and establishment identifiers, it includes the exact street address of each establishment. I combine those data with the national address database (BAN) that gathers all street addresses in France and their geographic coordinates.

Finally, data on vacancies stem from Ministry of Labor (DARES)'s ACEMO survey. Each semester, all firms with at least 10 employees are required declare the number of vacant positions open. They also have to specify whether the position corresponded to a newly-created job or the replacement of a departing employee.

Geography and transportation. I identify the banlieues using the classification of French most distressed urban areas developed by the administration in charge of the ZFU policy (CGET). This classification establishes three nested levels of priority for policy intervention: ZUSs, ZRUs and

[11] The tax potential of a city is equal to the amount that would yield the four local direct taxes to this municipality if the average national tax rates of each of these taxes were applied.

[12] The sample size was increased to 1/12th after 2002.

ZFUs from the least distressed to the most. The shape files of these urban zones contain the geographic coordinates of the perimeters delineating them.

The BD Topo shape files issued yearly by the French geographic institute IGN comprise traffic routes, transportation facilities as well as and use. Traffic routes include streets and roads. Information on transportation facilities is more limited but include the train and subway stations and major bus hubs. Land use information distinguishes residential and commercial buildings.

The National Transport Survey (ENT) and the Household Commuting Surveys (EMD) offer detailed data on car ownership and transportation modes. These large-scale surveys are conducted periodically in major cities. Besides basic demographics and employment information, respondents are asked their departure and arrival times and points, the distance covered, the travel purpose and the transport mode used. About 1/50th of the population is surveyed overall.

Demographics. Demographic data by place of residence come from the French Population Census (RP), which was used in the selection process of ZFUs. The census was conducted in the years 1990, 1999, and annually since 2006. It provides individual-level demographics, including citizenship and education. It also contains information on employment variables (labor force status, occupation, industry). Finally it includes information on commuting (transportation mode, commuting time) as well as residence (public housing, mobility since the last census). Geographic identifiers are at the block level until 1999, and at the block group level from 2006 onwards.

Real estate. Real estate data come from the National Housing Survey (*ENL*), which provides prices and a meticulous description of dwellings' characteristics. This survey is carried every five years on average, and covers about 1/600th dwelling units. It is comprised of information regarding purchase prices for owner-occupied units and rents for tenant-occupied units. It also includes housing characteristics, such as year of construction, surface area, specific facilities and condition of the unit.

Ownership. I use data from the household budget survey (BDF) to pin down firm and land ownership in the quantitative model.

1.3.3 Some Descriptive Statistics

A first look at the data confirms that the French banlieues are urban ghettos: residents are typically young and uneducated, and suffer from both high unemployed hazards and low wages when employed. Summary statistics also suggest reasons why their residents can't overcome spatial mismatch through commuting or residential mobility.

The demographic profile of banlieue population departs from national averages along three dimensions: age, education and ethnic origin. First, banlieue residents appear distinctively young. Summary statistics displayed in table 1.3.1 show that the share of youth under the age of 25 in

	Banlieues		France
	ZRU	Urban unit	
Demographics			
Population	2,952,000	27,469,000	53,486,000
Mean age	30.5	36.2	37.0
Share under 25	47%	36%	36%
Share without qualification	25%	18%	20%
Share of foreigners	19%	9%	6%
... from Maghreb	53%	38%	36%
... from Subsaharan Africa	8%	7%	6%
Share of French citizens of foreign origin	5%	4%	3%
Labor market outcomes			
Labor force participation	66%	68%	68%
Unemployment rate	20%	11%	11%
Monthly income	1,833	2,321	2,051
Commuting			
Car ownership			
... 1 car or more	77%	85%	90%
... 2 cars or more	24%	40%	48%
Commuting distance (km)	5.7	6.1	6.2
Residence			
Home ownership	18%	47%	58%
Share living in apartments	84%	57%	36%
Share of public housing	68%	23%	15%
Moved since last census (1982-1990)	55%	53%	49%

Table 1.3.1: Summary statistics

Source. 1990 RP and 1993-1995 LFS.

banlieues (47%) is substantially higher than in surrounding urban units (36%) and overall in France (36%). Accordingly, mean age of banlieue inhabitants is 5.7 years below the one of surrounding urban units and 6.5 years below the national mean. Second, the banlieue population is low-educated: among adults over 15 years old who are not currently at school, 25% have not completed any degree—7 percentage points above surrounding urban units. Third, a large share of banlieue population is composed of foreigners (19%) and of French citizens of foreign origin (5%). In comparison, these shares are 6% and 3% respectively in the overall population. Immigrants living in banlieues chiefly hail from Maghreb (53%) and sub-Saharan Africa (8%).

Labor market outcomes of banlieues residents are first and foremost characterized by widespread unemployment and low wages. Unemployment rates reach 20% of the labor force—twice the national rate. When employed, banlieues residents earn monthly incomes that are 500 euros lower than the average of the surrounding urban units. Reported unemployment rates may actually understate the extent of joblessness in French banlieues, which also manifests in weak labor force participation. Despite a relatively young population, labor farce participation in banlieues is 2 percentage points below national participation. Breaking down non-participation by stated cause reveals that the pursuit of studies or retirement can't explain low participation.

Lastly, banlieue residents are unlikely to overcome spatial isolation through commuting and residential mobility. Car ownership is 13-percentage point lower than the national average. Focusing on households with at least two adults in the labor force, only 24% have at least two cars in ZRUs, half of the overall rate. Finally, although residential mobility is slightly higher in banlieues than elsewhere, residential mobility might be hindered by the fact that 68% of the banlieue population lives in public housing.

1.4 A Quantitative Model of Spatial Mismatch

In this section, I develop the quantitative model that I will use to assess the welfare attributes of spatial mismatch. The main departure from the standard quantitative urban literature is the introduction frictional labor markets à la Diamond [1981] and Mortensen and Pissarides [1994]. Commuting costs and information decay due to distance generate unemployment and depress wages in neighborhoods distant from productive locations. The model also incorporates two additional forces that are relevant for the quantitative analysis: endogenous job acceptance, that directly links job finding rates to commuting costs, and idiosyncratic preferences for locations, that mirror imperfect residential mobility of workers. It also features public housing, which was salient in the descriptive statistics of section 1.3, as it acts as a subsidy towards low-productivity locations.

1.4.1 Environment

My starting point is an urban framework where the city is assumed to be a collection of neighborhoods distant from one another. Workers are alternatively employed and unemployed. Distance across neighborhoods affects both employed workers through costs in commuting and unemployed workers through information decay in job search.

After a brief overview of the model setup, I present the primitives of the model: preferences, technology, and taxation. I conclude this subsection with a description of the wage setting mechanism, market structure and ownership over land and firms.

1.4.1.1 Setup

The geography is comprised of a number of locations $i \in \mathcal{I}$ that differ in their fundamental components of amenities and productivity, their land stock and their distances to other locations. Four types of agents step in: workers, recruiters, producers and developers.

A continuum of imperfectly mobile workers choose where to live and have preferences over tradable goods and housing. Each worker belongs to one of a set G of types corresponding to different skill groups. They may be employed or unemployed. When employed, workers pay a monetary cost to commute to work. When unemployed, they search for jobs in the different neighborhoods of the city. Workers cannot move to live in another neighborhood to mitigate commuting costs after they find a job.

Recruiters hire workers and sell labor inputs to good producers and housing developers. They enter freely in each location and post costly type-specific vacancies. The surplus generated by a match between a recruiter and a worker is shared through Nash bargaining. Each match with a recruiter generates a stochastic match output, interpreted as efficiency units of labor. Low realizations may be rejected by workers because of the prospect of a better job match in the future.

Perfectly competitive producers use labor inputs provided by recruiters and intermediate goods to produce commodities that are imperfectly substitutable across locations, following Armington [1969].

Housing provision necessitates the use of labor, intermediate goods and a fixed factor, land, which generates rents. Housing developers are perfectly competitive.

Finally, government levies taxes on recruiters and workers. It also provides housing assistance to poor workers.

1.4.1.2 Preferences

Workers enjoy housing and tradable good consumption. They value local amenities and incur monetary commuting costs when they are employed.

Instantaneous utility. Preferences over housing and tradable goods are captured by a homogeneous utility function u that is common across skill groups. Utility of both unemployed and employed workers is given by:

$$a_{g,i} u(c,h) \varepsilon_{g,i}, \tag{1.4.1}$$

where c and h are the respective amounts of tradable good and housing consumed, $a_{g,i}$ are amenities in location i and $\varepsilon_{g,i}$ are idiosyncratic preferences for residences.

Commuting. To commute between the residence i and the workplace j, employed workers have to pay commuting costs. Those are modeled as a tradable good expenditure, d_{ij}.

Residential mobility. Employed workers cannot move to live in another neighborhood to mitigate commuting costs after they find a job. In other words, a worker cannot keep a job if he moves out. However, unemployed workers can move freely across locations. They have idiosyncratic preferences over residences that are modeled as independent and identically distributed draws $\varepsilon_{g,i}$ and determine residential mobility. Those draws enter multiplicatively in the value function and do not change over time.

1.4.1.3 Technology

Good production and housing development take place in each neighborhood. Beside intermediate goods, producers and developers both use labor inputs provided by recruiters who hire workers on frictional labor markets.

Producers. A continuum of identical producers assemble labor inputs and intermediate goods into consumption goods in each location. They produce:

$$Y_j = Y_j \left((N^Y_{g,j})_{g \in G}, M^Y_j \right), \tag{1.4.2}$$

where Y_j is the production function that is neighborhood-specific and may exhibit constant or decreasing returns to scale, $(N^Y_{g,j})_{g \in G}$ are labor inputs from all skill levels and M^Y_j are intermediate goods.

Production goods are differentiated by neighborhood, following Armington [1969].

Developers. Developers provide housing to workers in all locations. The presence of a fixed factor, land, induces decreasing returns to scale. Housing supply is given by:

$$H_i = H_i((N^H_{g,i})_{g \in G}, M^H_i), \tag{1.4.3}$$

where $(N^H_{g,i})_{g \in G}$ are labor inputs of different skill levels and M^H_i represents intermediate goods.

Recruiters. Recruiters post vacancies in each local labor market. Maintaining an open vacancy involves a flow cost, ν_g, that captures advertisement and other search costs on the side of the firm. Each match between a recruiter and a worker generates a surplus that is sold as type-specific labor input to producers and developers.

Each job-worker match generates a stochastic output $\alpha \geq 0$, specific to the match and interpreted as efficiency units of labor. The α's are distributed according to some exogenous distribution G that does not depend on the neighborhood where the match occurs. The worker observes α after matching with a recruiter, and may turn down job offers if α is too low because of the prospect of a better job match in the future.

Information decay. Distance to local labor markets generates information decay hampering job search. Information decay between residence i and workplace j is equal to $\Delta_{ij} \geq 1$.

Matching. Job seekers and recruiters match on type- and location-specific labor markets. The matching technology M has constant returns to scale. The flow rate of matches is equal to:

$$m_{g,j} = M\left(\sum_i \Delta_{ij}^{-1} U_{g,i}, V_{g,j}\right), \tag{1.4.4}$$

where $U_{g,i}$ is the number of type-g job seekers living in i, Δ_{ij} is information decay between i and j and $V_{g,j}$ is the number of type-g vacancies open in location j.

Matches split when a negative productivity shock occurs. Those shocks follow a Poisson process with arrival rate δ, which does not depend on match productivity α.

1.4.1.4 Taxation and Public Policy

The government grants residence and hiring subsidies, and redistributes across skill groups. Residence subsidies, $(t^R_{g,i})$, are available to eligible groups of workers in a subset of neighborhoods. In the context of the French banlieues, $(t^R_{g,i})$ is interpreted as housing assistance toward low-income families. Hiring subsidies, $\left(t^H_{g,ij}(\alpha)\right)$, may be location- and productivity-specific, which encompasses enterprise zone programs. They are interpreted as taxes when they are negative. Redistribution across skill groups occurs through lump-sum transfers $(T_g)_{g\in G}$ to the different skill groups.

1.4.1.5 Closing the Model

Having described the model's primitives, I specify market structure and ownership in order to close the model. As usual in frictional labor markets, wage is set through Nash bargaining between the worker and the recruiter. All the other prices are determined competitively. Workers of different skill groups own different shares of the land rents and firms' profits. While the population of workers in the city is fixed, recruiters enter freely in each local labor market.

Wage setting, market structure and ownership. Matches between recruiters and job seekers generate an intertemporal surplus that is shared through a generalized Nash-bargaining process. The total surplus is the sum of the surplus of workers and the surplus of recruiters. The bargaining power of workers is $\beta \in [0, 1]$. Production goods, labor inputs and housing markets are perfectly competitive. All agents are price-takers.

Each type-g worker receives a share ω^F_g of firms'—producers' and recruiters'—profits and a share ω^L_g of landlords' rents. Shares add up to 1 across skill groups:

$$\forall k \in \{F, L\}, \quad \sum_g \omega^k_g \bar{L}_g = 1. \tag{1.4.5}$$

Closed city and free entry. The total population of workers of each type g is fixed to an exogenous level, $\bar{L}_g$. Recruiters enter freely up to the point where the value of opening a vacancy is zero in each location and for each skill group.

Timing. Employed workers can't keep their jobs if they move out to live in another neighborhood.

1.4.2 Equilibrium

This subsection lays out the equilibrium behaviors of the different agents in order to define an equilibrium in this model.

1.4.2.1 Workers

Workers' intertemporal utility changes over time as they alternate between employment and unemployment.

Good and housing demands. Having chosen his residence i, a worker chooses optimally his tradable good and housing consumptions to maximize his utility given the local price index of consumption goods, P_i, the prevailing local rent, R_i, and residence subsidies $t^R_{g,i}$.

The worker's budget constraint depends on his employment status. If he is employed, he receives a wage, $w_{g,ij}(\alpha)$, that depends on his match-specific output, α. He also has to pay commuting costs, d_{ij}:

$$P_i c + R_i h \leq w_{g,ij}(\alpha) - P_i d_{ij} + t^R_{g,i}, \tag{1.4.6}$$

where $t^R_{g,i}$ are residence subsidies.

An unemployed worker receives benefits from home production, b, and does not have to pay commuting costs.

$$P_i c + R_i h \leq P_i b + t^R_{g,i}. \tag{1.4.7}$$

Therefore, an employed worker's consumption of tradable good, $c^E_{g,ij}(\alpha)$, and housing, $h^E_{g,ij}(\alpha)$, solve:

$$\max_{c,\,h}\; a_i u(c,h)\varepsilon_{g,i}, \tag{1.4.8}$$

subject to (1.4.6), while an unemployed worker's consumption of tradable good, $c^U_{g,i}$, and housing, $h^U_{g,i}$, solve (1.4.8) subject to (1.4.7).

Value functions. Workers discount the future at rate r and have rational expectations.

In the steady state, the Bellman equation for an employed worker is given by:

$$rJ^E_{g,ij}(\alpha) = a_i u\left(c^E_{g,ij}(\alpha), h^E_{g,ij}(\alpha)\right) + \delta\left(J^U_{g,i} - J^E_{g,ij}(\alpha)\right), \tag{1.4.9}$$

where $J^E_{g,ij}(\alpha)$ (resp. $J^U_{g,i}$) is the value function of an employed (resp. unemployed) worker normalized by the idiosyncratic preference shocks.

This equation states that an employed worker obtains utility $a_i u\left(c^E_{g,ij}(\alpha), h^E_{g,ij}(\alpha)\right)$ today but can lose his job at rate δ and then obtains a negative surplus equal to the difference between his expected lifetime utility if he were to lose his job and his current expected lifetime utility.

The Bellman equation of a job seeker is given by:

$$rJ^U_{g,i} = a_i u\left(c^U_{g,i}, h^U_{g,i}\right) + \sum_{j\in\mathcal{I}} \lambda_{g,ij}\left(\mathbb{E}\left[J^E_{g,ij}(\alpha) \,\middle|\, \alpha \geq \alpha^R_{g,ij}\right] - J^U_{g,i}\right), \tag{1.4.10}$$

where $\alpha^R_{g,ij}$ is the reservation output and $\lambda_{g,ij}$ the job finding rate for a type-g worker living in i who finds a job in j:

$$\lambda_{g,ij} = \left(1 - G(\alpha^R_{g,ij})\right)\Delta^{-1}_{ij}\theta_{g,j} q(\theta_{g,j}). \tag{1.4.11}$$

This second equation states that a job seeker obtains utility $a_i u\left(c^U_{g,i}, h^U_{g,i}\right)$ today, but may find a job in each neighborhood j at rate $\lambda_{g,ij}$ and then obtains a positive surplus equal to the difference between his expected lifetime utility if he is employed in j net of mobility costs, provided that he accepts the job and his current expected lifetime utility.

Job acceptance decisions. After matching with a recruiter, a job seeker decides whether or not to accept the offer after having observed the match-specific output, α. He will accept the offer if and only if the expected lifetime utility that he would get by accepting is higher than the one he currently has being unemployed. Therefore, the reservation output, $\alpha^R_{g,i}$, is implicitly defined by the equality:

$$J^E_{g,ij}\left(\alpha^R_{g,ij}\right) = J^U_{g,i} \tag{1.4.12}$$

That is, an unemployed will accept a job offer if and only if $\alpha \geq \alpha^R_{g,ij}$. Note that $\alpha^R_{g,ij}$ does not depend on a, as there is a dichotomy between residence variables and work variables.

Choice of residence. Unemployed workers pick their residence i to maximize the their value function $rJ^U_{g,i}$, which depends on idiosyncratic residence tastes $\varepsilon_{g,i}$:

$$\max_{i\in\mathcal{I}}\ rJ^U_{g,i}\,\varepsilon_{g,i}. \tag{1.4.13}$$

1.4.2.2 Recruiters

Recruiters enter freely in each labor market. Their intertemporal profits change over time as the job opening is filled or vacant.

Value functions. Risk-neutral recruiters post type-specific vacancies in each neighborhood. The Bellman equation for a filled type-g job with match-specific output α in location j filled by a worker living in i is:

$$rJ^F_{g,ij}(\alpha) = \alpha W_{g,j} - w_{g,ij}(\alpha) + \delta\left(J^V_{g,j} - J^F_{g,ij}(\alpha) + t^H_{g,ij}(\alpha)\right), \tag{1.4.14}$$

with $W_{g,j}$ the price of one unit of type-g labor input in location j.

The interpretation is that a filled vacancy obtains net-of-tax profits today but can split at rate δ and then obtains a negative surplus equal to $J^V_{g,j} - J^F_{g,ij}(\alpha)$.

The Bellman equation for a type-g vacancy in j

$$rJ^V_{g,j} = -P_j\nu_g + q(\theta_{g,j})\mathbb{E}_{|j}\left[J^F_{g,ij}(\alpha) - J^V_{g,j}\right], \tag{1.4.15}$$

where:

$$\mathbb{E}_{|j}\left[J^F_{g,ij}(\alpha) - J^V_{g,j}\right] = \frac{\sum_{i\in\mathcal{I}}\Delta_{ij}^{-1}U_{g,i}\left(1 - G(\alpha^R_{g,ij})\right)\mathbb{E}\left[J^F_{g,ij}(\alpha) - J^V_{g,j}\,|\,\alpha \geq \alpha^R_{g,ij}\right]}{\sum_{i\in\mathcal{I}}\Delta_{ij}^{-1}U_{g,i}}. \tag{1.4.16}$$

The interpretation is the following. A vacant position costs $P_j\nu_g$ per unit of time, but matches with a job seeker at rate $q(\theta_{g,j})$ that depends on labor market tightness for type-g workers in neighborhood j. The expected surplus of a match with a job seeker depends on where the he lives, which affects both the acceptance probability, $1 - G(\alpha^R_{g,ij})$, and the surplus conditional on job acceptance.

Free entry condition. By assumption, recruiters post vacancies up to a point where:

$$J^V_{g,j} = 0. \tag{1.4.17}$$

1.4.2.3 Producers

Producers assemble labor inputs $N^Y_{g,j}$, and intermediate goods M^Y_j into a quantity Y_j of goods. They take prices on the input and output markets as given. To produce Y_j, they solve the following problem:

$$\min_{(N^Y_{g,j})_{g\in G}, M^Y_j} \sum_{g\in G} W_{g,j}N^Y_{g,j} + P_jM^Y_j, \tag{1.4.18}$$

subject to:

$$Y_j\left((N^Y_{g,j})_{g\in G}, M^Y_j\right) \geq Y_j. \tag{1.4.19}$$

1.4.2.4 Developers

Developers use labor inputs $N^H_{g,j}$ and intermediate goods M^H_j to produce housing. They are perfectly competitive and take all prices as given. In each neighborhood j, they solve:

$$\min_{(N^H_{g,j})_{g\in G}, M^H_j} \sum_{g\in G} W_{g,j}N^H_{g,j} + P_jM^H_j, \tag{1.4.20}$$

subject to:

$$H_j((N^H_{g,j})_{g\in G}, M^H_j) \geq H_i. \tag{1.4.21}$$

1.4.2.5 Definition of a Decentralized Equilibrium

Having characterized the equilibrium behavior of each agent, I can now define the decentralized equilibrium of this model. First, I describe the wage-setting mechanism and the equilibrium conditions on the labor, good and housing markets that determine equilibrium prices. I conclude this section with the definition of a decentralized equilibrium.

Wage setting. In each period, the intertemporal surplus generated by a match between a worker and a recruiter is shared through a Nash-bargaining process:

$$w_{g,ij}(\alpha) = \arg\max_{w} \left(J^E_{g,ij}(\alpha) - J^U_{g,i}\right)^{\beta} \left(J^F_{g,ij} - J^V_{g,j}\right)^{1-\beta}, \tag{1.4.22}$$

where value functions $J^E_{g,ij}(\alpha)$ and $J^F_{g,ij}$ depend implicitly on wage w.

Labor market equilibrium. Labor inputs are used by producers and developers. In equilibrium, the number of employees in each neighborhood has to be equal to the total labor inputs used in good production and housing development:

$$\sum_{i\in\mathcal{I}} \alpha^E_{g,ij} l_{g,ij} L_{g,i} = N^Y_{g,j} + N^H_{g,j}, \tag{1.4.23}$$

where $\alpha^E_{g,ij}$ is the average productivity of type-g employees living in i and working in j:

$$\alpha^E_{g,ij} = \mathbb{E}\left[\alpha \,|\, \alpha \geq \alpha^R_{g,ij}\right]. \tag{1.4.24}$$

Good market equilibrium. Goods produced in each neighborhood j are used for consumption by workers and as intermediates by producers and developers. Geography is captured by iceberg trade costs $\chi_{jl} \geq 1$. That is, producers in location j must ship $\chi_{jl} Q_{jl}$ units to location l for Q_{jl} units to arrive. The feasibility constraint for tradable goods implies:

$$Y_j = \sum_{l\in\mathcal{I}} \chi_{jl} Q_{jl}, \tag{1.4.25}$$

where Y_j is the production in location j and Q_{jl} is the sum of goods used by workers, producers and developers in location l.

Goods are differentiated by origin and aggregated through a homothetic and concave aggregator Q. For now, I assume no further restriction on Q. Feasibility constraint for traded goods implies:

$$Q(Q_{1i}, \ldots, Q_{Ii}) = M^Y_i + M^H_i + \sum_{g\in G} L_{g,i}\Big(u_{g,i} c^U_{g,i} + \sum_{j\in\mathcal{I}} \int_{\alpha \geq \alpha^E_{g,ij}} l_{g,ij}(\alpha) c^E_{g,ij}(\alpha) d\alpha\Big). \tag{1.4.26}$$

This flexible functional form covers in particular perfect substitution as in Rosen [1979] and Roback

[1982]'s seminal models and constant elasticity of substitution (CES) à la Armington [1969], as is standard in economic geography models.

Housing market equilibrium. Housing produced by developers is consumed by workers of different skill groups in each location i:

$$H_i((N^H_{g,i})_{g\in G}, M^H_i) = \sum_{g\in G} L_{g,i}\Big(u_{g,i}h^U_{g,i} + \sum_{j\in\mathcal{I}}\int_{\alpha\geq\alpha^E_{g,ij}} l_{g,ij}(\alpha)h^E_{g,ij}(\alpha)d\alpha\Big). \tag{1.4.27}$$

Definition of a decentralized equilibrium. Before defining a decentralized equilibrium, it is convenient to introduce the definition of an allocation.

Definition 1.2 (Allocation). An *allocation*, $\mathcal{A}$, is the specification at each instant $t \geq 0$ of a partition of workers, $(L^E_{g,ij}(\alpha,t))$ and $(L^U_{g,i}(t))$, associated per capita consumptions of tradable goods and housing, $(c^E_{g,ij}(\alpha,t), h^E_{g,ij}(\alpha,t))$ and $(c^U_{g,i}(t), h^U_{g,i}(t))$, labor inputs used in the production and development sectors, $(N^Y_{g,j}(t))_{g\in G,j\in\mathcal{I}}$ and $(N^H_{g,j}(t))_{g\in G,j\in\mathcal{I}}$, intermediate goods used in the production and development sectors, $(M^Y_{g,j})_{g\in G,j\in\mathcal{I}}$ and $(M^H_{g,j}(t))_{g\in G,j\in\mathcal{I}}$, goods produced and housing developed, $(Y_j(t))_{j\in\mathcal{I}}$ and $(H_j(t))_{j\in\mathcal{I}}$, and labor market tightness in each labor market, $(\theta_{g,j}(t))_{g\in G,j\in\mathcal{I}}$.

Having determined the equilibrium behavior of each agent individually, I now summarize the above conditions to define a decentralized equilibrium.

Definition 1.3 (Decentralized Equilibrium). A *decentralized equilibrium* is an allocation $\mathcal{A}$ such that at each $t \geq 0$:

(i) Workers consume tradable goods and housing to maximize their utility subject to their budget constraint, conditions (1.4.8), (1.4.6) and (1.4.7), and choose their residence optimally when unemployed, condition (1.4.13);

(ii) Recruiters enter freely in each labor market, condition (1.4.17);

(iii) Workers make privately optimal job acceptance decisions, condition (1.4.12);

(iv) Producers choose labor inputs and intermediate goods optimally, conditions (1.4.18) and (1.4.19);

(v) Developers choose labor inputs and intermediate goods optimally, conditions (1.4.20) and (1.4.21);

(vi) Goods are aggregated optimally, condition (1.4.26);

(vii) Wages are determined through Nash bargaining between workers and recruiters, condition (1.4.22);

(viii) Labor, good and housing markets clear, conditions (1.4.24), (1.4.25) and (1.4.27).

1.4.3 Welfare Analysis: Extending Toy Model's Results

The insights of the toy model extend to the quantitative framework, as the balance between commuting costs and information decay determines whether there is too much or too little spatial mismatch.

Welfare of skill group g, $\mathcal{W}_g$, is the expected utility of unemployed and employed workers discounted over time:

$$\mathcal{W}_g = \int_0^\infty e^{-rt} \sum_i \tilde{a}_{g,i} \left[u_{g,i} u(c^U_{g,i}, h^U_{g,i}) + \sum_j \int_\alpha l_{g,ij}(\alpha) u(c^E_{g,ij}(\alpha), h^E_{g,ij}(\alpha)) d\alpha \right] L_{g,i} dt, \tag{1.4.28}$$

where $\tilde{a}_{g,i} = \mathbb{E}\left[\varepsilon_{g,i} \mid i\right] a_{g,i}$ are local amenities adjusted for the average value of idiosyncratic residence preferences.

The planner maximizes the welfare of one group subject to meeting required utility levels for the other groups. A formal definition of the planner's problem is laid out in appendix 1.B.

Definition 1.4 (Planner's Problem). Let $g_0 \in G$. The planner's problem is:

$$\max \ \mathcal{W}_{g_0}, \tag{1.4.29}$$

subject to:

(i) required utility levels for skill groups $g \neq g_0$;

(ii) spatial mobility constraints;

(iii) tradable good and housing feasibility constraints;

(iv) search and matching constraints;

(v) labor market clearing;

(vi) population constraints.

The constrained-efficient allocation is characterized by three conditions. The first two, optimal residence choice of workers and optimal entry of recruiters, generalize those of the toy model. The last one, optimal job acceptance, describes a new margin of inefficiency.

Proposition 1.5 (Constrained-Efficient Allocation). The constrained-efficient allocation solution to the planner's problem 1.4 satisfies the following three conditions.

1. Workers' optimal choice of residence:

$$\begin{aligned} \Lambda_g = \omega_g a_i & \frac{(r+\delta) u(c^U_{g,i}, h^U_{g,i}) + \sum_j \lambda_{g,ij} u(c^E_{g,ij}(\alpha^E_{g,ij}), h^E_{g,ij}(\alpha^E_{g,ij}))}{r + \delta + \bar{\lambda}_{g,i}} \\ & - \frac{(r+\delta) x^U_{g,i} + \sum_j \lambda_{g,ij} x^E_{g,ij}(\alpha^E_{g,ij})}{r+\delta+\bar{\lambda}_{g,i}} + \frac{(r+\delta)\left(P_i b - \sum_j \Delta_{ij}^{-1} \theta_{g,j} P_j \nu\right) + \sum_j \lambda_{g,ij}\left(\alpha^E_{g,ij} W_{g,j} - P_i d_{ij}\right)}{r+\delta+\bar{\lambda}_{g,i}}, \end{aligned} \tag{1.4.30}$$

with $W_{g,j}$ and $\theta_{g,j}$ the price of labor inputs and the labor market tightness for skill group

g in location j, $x^U_{g,i}$ and $x^E_{g,ij}$ the expenditures of unemployed and employed workers, ω_g the Pareto weight and Λ_g the opportunity cost of a worker.

2. Recruiters' optimal entry:

$$\begin{aligned} & P_j \nu_g \sum_i \Delta_{ij}^{-1} u_{g,i} L_{g,i} \\ = & \tfrac{1-\mu(\theta_{g,j})}{\theta_{g,j}} \textstyle\sum_i \tfrac{\lambda_{g,ij}}{r+\delta+\lambda_{g,i}} L_{g,i} \Big(u_{g,i} \left[\alpha^E_{g,ij} W_{g,j} - P_i d_{ij} - P_i b \right] + u_{g,i} \textstyle\sum_{j'} \Delta_{ij'}^{-1} \theta_{g,j'} P_{j'} \nu_g \\ & + \frac{\delta}{r+\delta} \sum_{j'} l_{g,ij'} \left[\alpha^E_{g,ij} W_{g,j} - P_i d_{ij} - \left(\alpha^E_{g,ij'} W_{g,j'} - P_i d_{ij'} \right) \right] \Big) \end{aligned} \tag{1.4.31}$$

with $\lambda_{g,ij} = \big(1 - G(\alpha^R_{g,ij})\big) \Delta_{ij}^{-1} \theta_j q(\theta_j)$ the job finding rate in workplace j of a type-g job seeker living in i and $\mu(\theta_j)$ the elasticity of the matching function m with respect to unemployment.

3. Optimal job acceptance:

$$\alpha^R_{g,ij} W_{g,j} - P_i d_{ij} = \frac{(r+\delta)\Big(P_i b - \sum_{j'} \Delta_{ij'}^{-1} \theta_{g,j'} P_{j'} \nu\Big) + \sum_{j'} \lambda_{g,ij'} \Big(\alpha^E_{g,ij'} W_{g,j'} - P_i d_{ij'}\Big)}{r+\delta+\lambda_{g,i}}. \tag{1.4.32}$$

1.5 Model Estimation: Disentangling the Channels of Spatial Mismatch

I estimate structurally the model described in section 1.4 as per the insights from the toy model's comparative statics. I map the key parameters to the reduced-form counterparts of structural identities, and I leverage the quasi-experimental variation of the ZFU policy to estimate them. Commuting costs are recovered from the sensitivity of wages to commute times, information decay from the sensitivity of employment shares to commuting times and population mobility from the sensitivity of population to local labor market conditions. I validate the estimation process by checking that the estimated model replicates the impact of the ZFU policy on zone-level employment, displacement effects in the neighboring areas and changes in the local population composition.

1.5.1 Quantitative Implementation

This subsection exposes the preliminary steps necessary to take the model developed in section 1.4 to the data.

I specify the functional forms. I then determine the parameter values under which the existence and uniqueness of an equilibrium is granted. I finally describe the full structural estimation procedure.

1.5.1.1 Functional Forms

I start by defining the three key parameters of the model which will be estimated in section 1.5.2.2: κ^d, that captures commuting costs, κ^Δ, that captures information decay, and σ, that captures the residential mobility of workers. I then specify the utility, production and the matching functions which are all Cobb-Douglas and will be calibrated in section 1.5.2.1.

Key parameters. Commuting costs and information decay depend both on the travel time t_{ij} between residence i and workplace j. I assume that commuting costs increase linearly with t_{ij}:

$$d_{ij} = \kappa^d t_{ij}, \tag{1.5.1}$$

where $\kappa^d > 0$. Information decay rises exponentially in t_{ij} at a rate κ^Δ:

$$\Delta_{ij} = \exp\left(\kappa^\Delta t_{ijx}\right). \tag{1.5.2}$$

Idiosyncratic residence draws, $(\varepsilon_{g,i})$ follow a Fréchet distribution with dispersion parameter σ. This parameter captures the strength of idiosyncratic preferences for locations and is inversely proportional to residential mobility.

Utility, production and matching functions. I assume that workers have Cobb-Douglas preferences over tradable goods and housing:

$$u(c,h) = c^\gamma h^{1-\gamma}, \tag{1.5.3}$$

with $\gamma \in (0,1)$.

Producers and developers use both Cobb-Douglas technologies, given respectively by:

$$Y_j = y_j (\bar{N}_j^Y)^{\bar{\alpha}^Y} (M_j^Y)^{\beta^Y}, \tag{1.5.4}$$

and:

$$H_j = h_j (\bar{N}_j^H)^{\bar{\alpha}^H} (M_j^H)^{\beta^H}, \tag{1.5.5}$$

where composite labor $\bar{N}_j^k = \left(\sum_g \alpha_g^k (N_{g,j}^k)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}$ for $k \in \{Y, H\}$. The labor share is $\bar{\alpha}^k = \sum_g \alpha_g^k$ and the elasticity of substitution between labor skill groups is η.

The matching function is also Cobb-Douglas, with $\mu \in (0,1)$ its elasticity with respect to unemployment:

$$M(U,V) = U^\mu V^{1-\mu}. \tag{1.5.6}$$

Finally, the idiosyncratic match outputs follow a Pareto distribution with shape parameter ρ:

$$G(\alpha) = \frac{1}{\alpha^{\rho}}, \quad \alpha \geq 1. \tag{1.5.7}$$

1.5.1.2 Model Inversion

Fundamental location characteristics such as productivities and amenities cannot be directly observed in the data. While the presence of thick market externalities allows for the possibility of multiple equilibria, I am able to recover unique values of intrinsic components of productivities and amenities that rationalize the observed data as a model equilibrium.

This inversion process follows closely the steps outlined in Ahlfeldt et al. [2015] with two twists. First, vacancy costs are additional parameters that need to be calibrated from the data. I assume that vacancy costs, (ν_g), vary by skill but not by location. I calibrate them by targeting the overall unemployment rate of each skill group. Second, there are multiple skill groups, as in Tsivanidis [2019]. I combine those observed data with the model structure to solve for the endogenous variables and back out the unobservable amenities and productivities.

Proposition 1.6 (Model Inversion).

1. Given data on residence by skill group, $(L^R_{g,i})$, total employment by workplace, $(\bar{L}^E_j)$, and overall unemployment rate by skill group, (u_g), in addition to model parameters, there exists unique vectors of labor market tightness $(\theta_{g,j})$, labor input prices, $(W_{g,j})$, and reservation productivities $(\alpha^R_{g,ij})$ that rationalize the observed data as an equilibrium of the model.
2. Given model parameters, data on residence by skill group, $(L^R_{g,i})$ and rent levels (R_i), and vectors of labor market tightness $(\theta_{g,j})$, labor input prices, $(W_{g,j})$, and reservation productivities $(\alpha^R_{g,ij})$, there exist unique vectors of unobservable amenities (a_i) (to scale), productivities (y_j), housing supplies (h_i), and vacancy costs (ν_g) that rationalize the observed data as an equilibrium of the model.

1.5.2 Structural Estimation

I now implement the method exposed in 1.5.1 to estimate structurally the model. I start with calibrated parameters, before switching to the estimation of the key structural parameters that leverages the ZFU policy.

1.5.2.1 Calibrated Parameters

Parameters $\{r, \delta, \beta, \gamma, \mu, \rho, \eta, (\alpha^Y_g), \beta^Y, (\alpha^H_g), \beta^H\}$ are calibrated either directly from the data or to existing values from the literature.

Rates r and δ are set at the quarterly level. Discount rate r is equal to 0.04 to match a 5% annual interest rate. Separation rate is calibrated to match the employment-to-unemployment transition rate which is equal to 1.6% in LFS data. This implies a value of 0.016 for δ.

I set the share of housing expenditure for workers to $1-\gamma = 0.24$ from BDF data, which is close to the commonly used value of 0.3 for the United States. The shape of the Pareto distribution of productivity draws $\alpha^R_{g,ij}$ is estimated from DADS wage data using a Hill's estimator, and is approximately equal to 3. The bargaining power β is set to 0.1 to match the labor share 0.7[13]. This value is close to both macroeconomic estimates [Hagedorn and Manovskii, 2008] and to recent quasi-experimental evidence [Card et al., 2013]. The matching function elasticity comes from Borowczyk-Martins et al. [2013], who find that μ is approximately equal to 0.3.

The elasticity of substitution between labor skill group, η, is set to 1.4 based on the Card [2009]'s review. The shares of labor and equipment correspond to their estimates in Greenwood et al. [1997], renormalized to exclude structures which are absent from the model. Finally, the share of land in the housing development is set to imply a housing supply elasticity of 1.75, as reported by Saiz [2010] in the United States.

1.5.2.2 Estimation Strategy of Key Structural Parameters

To estimate the three main parameters of the model, I proceed in three steps. First, commuting costs κ^d are identified from the sensitivity of wage to changes in travel time. Second, information decay κ^Δ is obtained from projection of job finding rates (adjusted for job acceptance) on travel time. Lastly, residence-specific preferences σ are identified from the response of population to local unemployment and wage moves. The ZFU program yields an exogenous variation in local unemployment and wages that allows to pin down this parameter.

Estimation of commuting costs κ^d. Commuting costs κ^d manifest in cross-location, within-type wage differences through two distinct channels. First, Nash bargaining between the workers and the recruiters implies a monetary compensation of commuting costs as in the toy model exposed in section 1.2. The second mechanism stems from endogenous job acceptance, which has been introduced in the quantitative model of section 1.4. Workers turn down job offers when the wage net of commuting costs is too low. Thus, the reservation wage of a worker is higher for distant locations.

Combining the job acceptance condition (1.4.12) and the Nash bargaining (1.4.22) provides a simple expression of the average wage of type-g workers commuting from i to j. It is given by:

$$w_{g,ij}(\alpha^E_{g,ij}) = \tilde{\rho}\left(\Omega_{g,i} + P_i d_{ij}\right), \tag{1.5.8}$$

where $\Omega_{g,i}$ denotes workers' outside option[14], $P_i d_{ij}$ represent the commuting costs between i and j and $\tilde{\rho} = 1 + \frac{1}{\rho-1}\beta$.

[13] In this model, the labor share is equal to $1 - \frac{1-\beta}{\rho}$.

[14] Formally, $\Omega_{g,i} = \frac{(r+\delta)b + \sum_j \lambda_{g,ij}\left(w_{g,ij}(\alpha^E_{g,ij}) - P_i d_{ij}\right)}{r+\delta+\bar{\lambda}_{g,i}}$.

Intuitively, wages increase in the outside option, augmented by the commuting costs that the worker incurs when he works in j. Higher values of the shape parameter of the productivity draws, ρ, correspond to a flatter productivity distribution, which in turn leads to lower reservation wages and thus lower average wages. The reservation wage responds to the shape of the productivity distribution only to the extent that workers capture a share of the surplus—that is, to a positive value of β.

In the data, commuting costs are identified by the sensitivity of wages to changes in commuting time. The structural relation (1.5.8) is estimated by running the following regression at the individual level:

$$w_{g,ij} = \tilde{\rho}\kappa^d t_{ij} + \omega_{g,i} + \gamma^{d\prime}\,\text{Controls}_g + \varepsilon_{g,ij}, \tag{1.5.9}$$

where t_{ij} is the travel time between locations i and j, $\omega_{g,i}$ is a residence fixed effect and Controls_g are individual characteristics. Here, the error term, $\varepsilon_{g,ij}$, is interpreted as unobserved idiosyncratic productivity draws and unobserved skill heterogeneity.

Table 1.5.1 gives the estimates of κ^d, which is close to 9.5 in the most robust specifications.

	Yearly income					
	(1)	(2)	(3)	(4)	(5)	(6)
Commuting time	53.61***	43.28***	9.486***	8.984***	9.569***	9.349***
	(0.178)	(0.164)	(0.0869)	(0.0866)	(0.181)	(0.181)
Controls						
Sex		×		×		×
Age		×		×		×
Fixed effects						
Year	×	×	×	×	×	×
Individual			×	×	×	×
Resid. by Educ.	×	×			×	×
Observations (in thousands)	9,316	9,316	32,587	32,586	9,275	9,275

Table 1.5.1: Estimation of κ^d

Notes. Yearly income is expressed in euros and adjusted for hours worked. Commuting time is expressed in minutes. Standard errors in parenthesis. * p<0.10, ** p<0.05, *** p<0.01
Source. 1990-2012 FH-DADS.

Estimation of information decay κ^Δ. Information decay κ^Δ governs job finding rates, after adjusting for job acceptance. The job *matching* rate—rate at which workers and recruiters match on

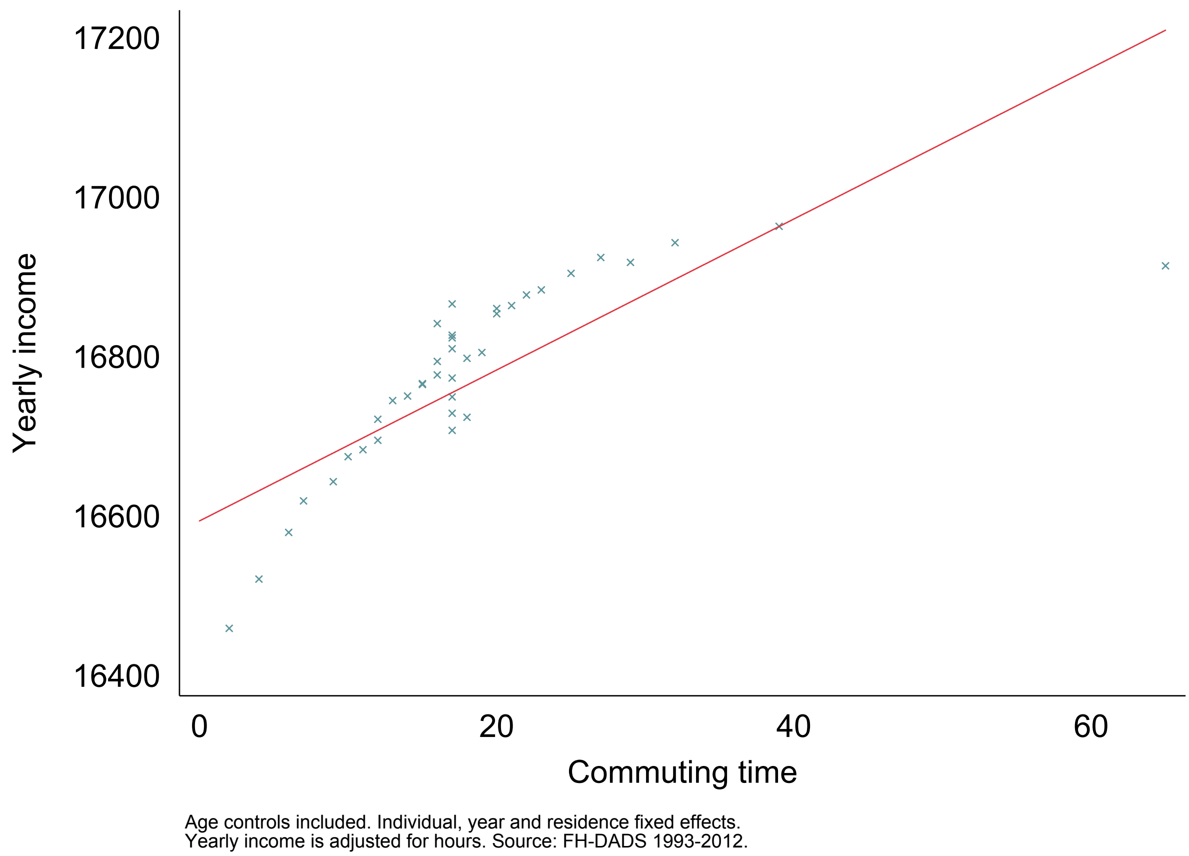


Figure 1.5.1: Wage premium for commuting

Notes. Individual and residence fixed effects and age controls included. Yearly income is adjusted for hours.

Source. FH-DADS 1990-2012.

each local labor market—is fully determined by information decay and local labor market tightness. However, one only observes the job *finding* rate, which is equal to the job matching rate multiplied by the acceptance probability. Therefore, to recover information decay, observed job finding rates have to be adjusted for job acceptance, itself a function of the average wage net of the outside option and commuting costs.

Plugging in $\alpha^R_{g,ij}$ from the job acceptance rule (1.4.12) delivers the following relation for the job finding rates:

$$\log l_{g,ij}(w^E_{g,ij})^\rho = -\kappa^\Delta t_{ij} + \delta_{g,i} + \delta_{g,j} + \varepsilon^\Delta_{g,ij}, \tag{1.5.10}$$

where the average income $w^E_{g,ij}$ is an adjustment factor for job acceptance, $\delta_{g,i}$ are residence fixed effects capturing outside options and $\delta_{g,j}$ are workplace fixed effects which encapsulate labor market tightness and productivity differentials. The error term $\varepsilon^\Delta_{g,ij}$ captures determinants of information decay across neighborhoods that do not vary systematically with travel time.

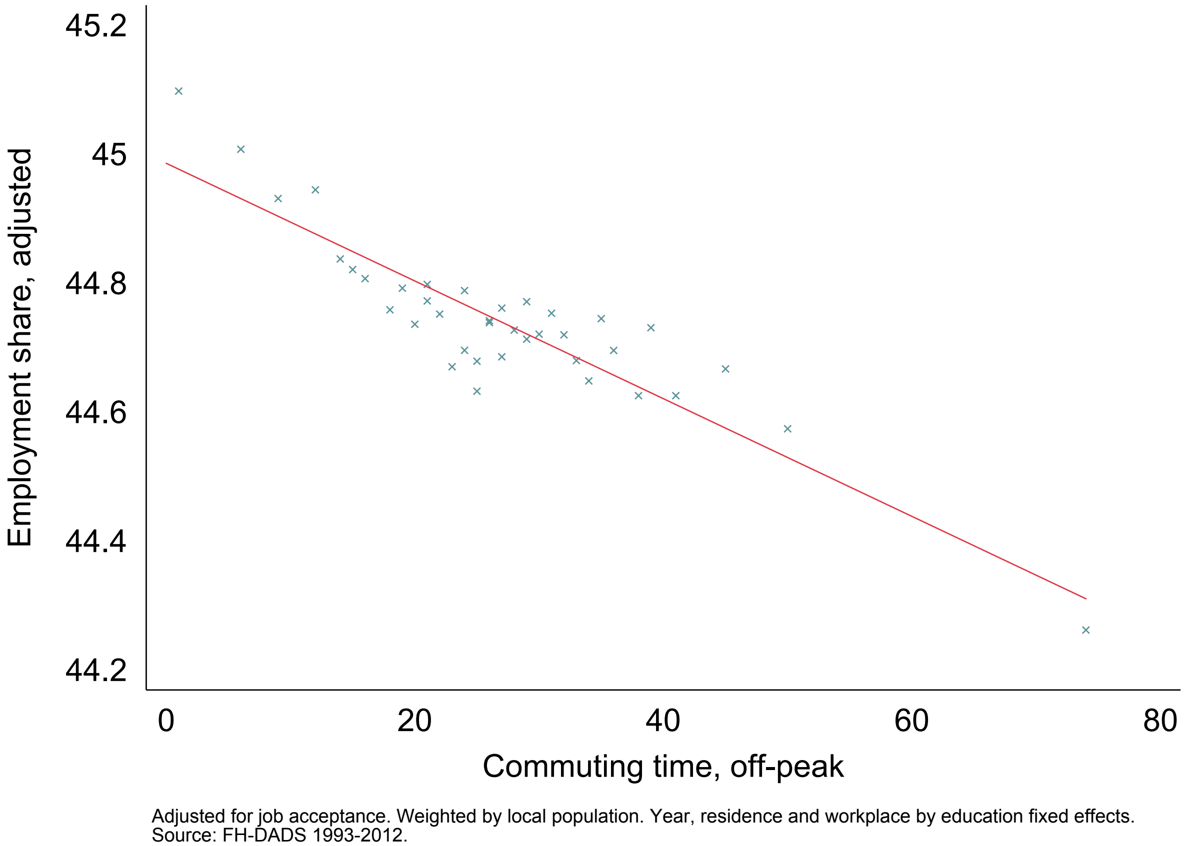


Figure 1.5.2: Home bias in employment

Notes. Individual and residence fixed effects and age controls included. Yearly income is adjusted for hours.

Source. FH-DADS 1990-2012.

Parameter κ^{Δ} is simply retrieved as the projection coefficient in (1.5.10).

	Employment share (log)	
	(1)	(2)
Commuting time	-0.0220***	-0.0174***
	(0.000524)	(0.00103)
Specification		
Adjusted		×
Fixed effects		
Year	×	×
Resid. by Educ.	×	×
Work. by Educ.	×	×
Observations	62,017	58,651

Table 1.5.2: Estimation of commuting costs κ^{δ}

Notes. Commuting time is expressed in minutes. Standard errors in parenthesis. * p<0.10, ** p<0.05, *** p<0.01.
Source. 1990-2012 FH-DADS.

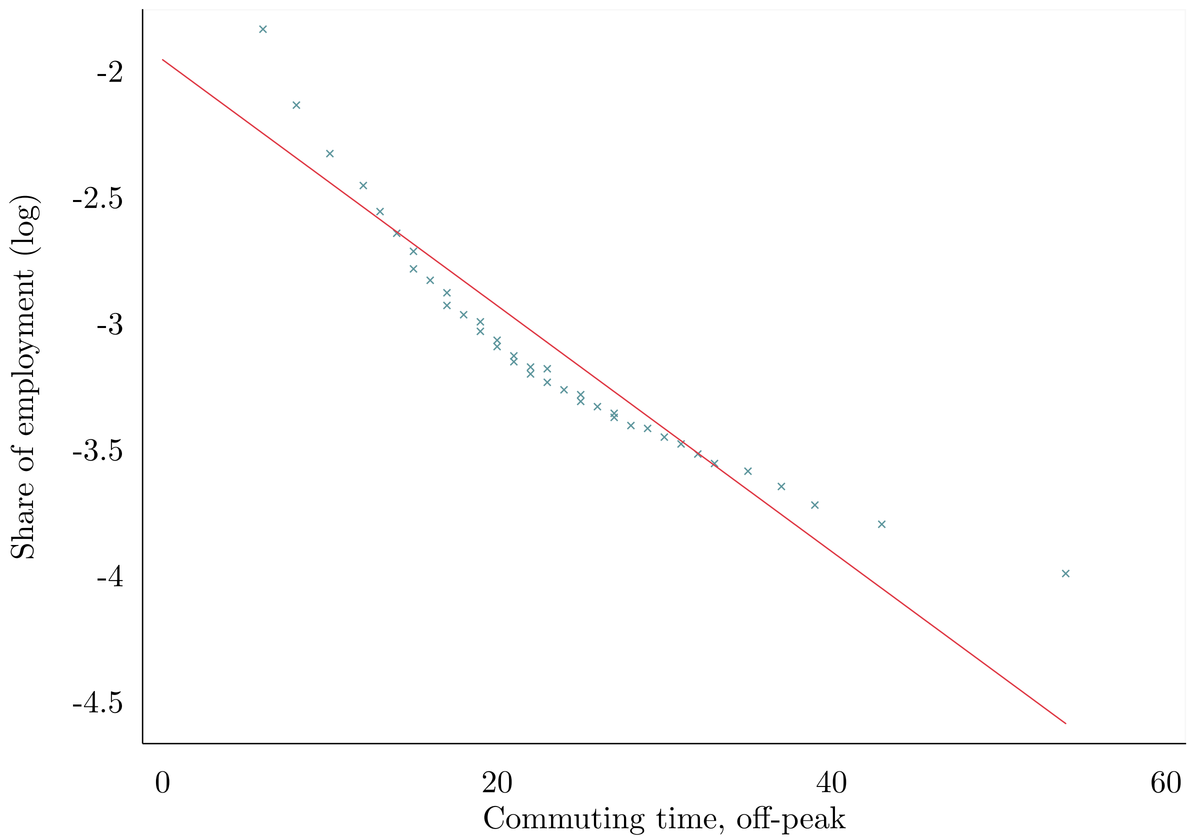


Figure 1.5.3: Estimation of information decay κ^{Δ}

Notes. Residence and workplace fixed effects.

Source. FH-DADS 2006.

1.6 Conclusion

This paper develops a framework to study theoretically and quantitatively the welfare attributes of spatial mismatch. I introduce frictional labor markets into a quantitative urban model. The spatial structure of the city interacts with labor markets because commuting is costly and information about job offers decays with distance. I prove that both workers' choice of residence and vacancy creation are inefficient in the decentralized equilibrium. The constrained-efficient allocation may be restored by a mix of place-based residence and hiring subsidies.

In subsequent work, I will use this framework to quantify the welfare implications of spatial mismatch. I will apply my model to French urban ghettos and leverage a spatial experiment to estimate it. I will evaluate through the lens of my model that the enterprise zone program supposed to undo spatial mismatch in French urban ghettos. Finally, I will explore a range of counterfactual policies designed to tackle spatial mismatch.

Bibliography

Acemoglu, Daron (2001), "Good Jobs versus Bad Jobs," *Journal of Labor Economics*, Vol. 19, No. 1, pp. 1–21.

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf (2015), "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, Vol. 83, No. 6, pp. 2127–2189.

Allen, Treb and Costas Arkolakis (2014), "Trade and the Topography of the Spatial Economy," *Quarterly Journal of Economics*, Vol. 129, No. 3, pp. 1085–1140.

_ and _ (2019), "The Welfare Effects of Transportation Infrastructure Improvements," Working Paper 25487, National Bureau of Economic Research.

_ , _ , and Xiangliang Li (2015), "Optimal City Structure," Preliminary Draft.

Andersson, Fredrik, John C. Haltiwanger, Mark J. Kutzbach, Henry O. Pollakowski, and Daniel H. Weinberg (2018), "Job Displacement and the Duration of Joblessness: The Role of Spatial Mismatch," *Review of Economics and Statistics*, Vol. 100, No. 2, pp. 203–218.

Armington, Paul S. (1969), "A Theory of Demand for Products Distinguished by Place of Production," *Staff Papers (International Monetary Fund)*, Vol. 16, No. 1, pp. 159–178.

Barbanchon, Thomas Le and Franck Malherbet (2013), "An Anatomy of the French Labour Market: Country Case Study on Labour Market Segmentation," Employment Working Paper 142, International Labour Organization.

Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence F. Katz, and Christopher Palmer (2019), "Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice," Working Paper 26164, National Bureau of Economic Research.

Bilal, Adrien (2020), "The Geography of Unemployment," Job Market Paper.

Bondonio, Daniele and John Engberg (2000), "Enterprise Zones and Local Employment: Evidence from the States' Programs," *Regional Science and Urban Economics*, Vol. 30, No. 5, pp. 519–549.

_ and Robert T. Greenbaum (2007), "Do Local Tax Incentives Affect Economic Growth? What Mean Impacts Miss in the Analysis of Enterprise Zone Policies," *Regional Science and Urban Economics*, Vol. 37, No. 1, pp. 121–136.

Borowczyk-Martins, Daniel, Grégory Jolivet, and Fabien Postel-Vinay (2013), "Accounting for Endogeneity in Matching Function Estimation," *Review of Economic Dynamics*, Vol. 16, No. 3, pp. 440–451.

Briant, Anthony, Miren Lafourcade, and Benoît Schmutz (2015), "Can Tax Breaks Beat Geography? Lessons from the French Enterprise Zone Experience," *American Economic Journal: Economic Policy*, Vol. 7, No. 2, pp. 88–124.

Brueckner, Jan K. and Richard W. Martin (1997), "Spatial Mismatch: An Equilibrium Analysis," *Regional Science and Urban Economics*, Vol. 27, No. 6, pp. 693–714.

_ and Yves Zenou (1999), "Harris-Todaro Models with a Land Market," *Regional Science and Urban Economics*, Vol. 29, No. 3, pp. 317–339.

Brueckner, Jan K. and Yves Zenou (2003), "Space and Unemployment: The Labor-Market Effects of Spatial Mismatch," *Journal of Labor Economics*, Vol. 21, No. 1, pp. 242–262.

Busso, Matias, Jesse Gregory, and Patrick Kline (2013), "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *American Economic Review*, Vol. 103, No. 2, pp. 897–947.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro (2019), "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, Vol. 87, No. 3, pp. 741–835.

Card, David (2009), "Immigration and Inequality," *American Economic Review*, Vol. 99, No. 2, pp. 1–21.

_ , Francesco Devicienti, and Agata Maida (2013), "Rent-sharing, Holdup, and Wages: Evidence from Matched Panel Data," *Review of Economic Studies*, Vol. 81, No. 1, pp. 84–111.

Charnoz, Pauline (2018), "Do Enterprise Zones Help Residents? Evidence from France," *Annals of Economics and Statistics*, Vol. 130, pp. 199–225.

Criscuolo, Chiara, Ralf Martin, Henry G. Overman, and John Van Reenen (2019), "Some Causal Effects of an Industrial Policy," *American Economic Review*, Vol. 109, No. 1, pp. 48–85.

Diamond, Peter A. (1981), "Mobility Costs, Frictional Unemployment, and Efficiency," *Journal of Political Economy*, Vol. 89, No. 4, pp. 798–812.

Elvery, Joel A. (2009), "The Impact of Enterprise Zones on Resident Employment: An Evaluation of the Enterprise Zone Programs of California and Florida," *Economic Development Quarterly*, Vol. 23, No. 1, pp. 44–59.

Fajgelbaum, Pablo D. and Cécile Gaubert (Forthcoming), "Optimal Spatial Policies, Geography and Sorting," *Quarterly Journal of Economics*, Vol. 0, No. 0, pp. 000–000.

_ and Edouard Schaal (2019), "Optimal Transport Networks in Spatial Equilibrium," Working Paper 23200, National Bureau of Economic Research.

Gaubert, Cécile, Patrick M. Kline, and Danny Yagan (2020), "Place-Based Redistribution," Working Paper.

Givord, Pauline, Simon Quantin, and Corentin Trevien (2018), "A Long-Term Evaluation of the First Generation of French Urban Enterprise Zones," *Journal of Urban Economics*, Vol. 105, pp. 149–161.

— , Roland Rathelot, and Patrick Sillard (2013), "Place-Based Tax Exemptions and Displacement Effects: An Evaluation of the Zones Franches Urbaines Program," *Regional Science and Urban Economics*, Vol. 43, No. 1, pp. 151–163.

Glaeser, Edward L. and Joshua D. Gottlieb (2008), "The Economics of Place-Making Policies," *Brookings Papers on Economic Activity*, Vol. 39, No. 1, pp. 155–253.

Gobillon, Laurent, Thierry Magnac, and Harris Selod (2012), "Do Unemployed Workers Benefit from Enterprise Zones? The French Experience," *Journal of Public Economics*, Vol. 96, No. 9, pp. 881–892.

— , Harris Selod, and Yves Zenou (2007), "The Mechanisms of Spatial Mismatch," *Urban Studies*, Vol. 44, No. 12, pp. 2401–2427.

Gore, Tony and Alison Herrington (1997), *New Urban Transport Investment and the Sheffield Labour Market : Report of "After" Surveys and Overall Assessment*, Sheffield: School of Urban and Regional Studies, Sheffield Hallam University.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell (1997), "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review*, Vol. 87, No. 3, pp. 342–62.

Hagedorn, Marcus and Iourii Manovskii (2008), "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited," *American Economic Review*, Vol. 98, No. 4, pp. 1692–1706.

Hanson, Susan, Tara Kominiak, and Scott Carlin (1997), "Assessing the Impact of Location on Women's Labor Market Outcomes: A Methodological Exploration," *Geographical Analysis*, Vol. 29, No. 4, pp. 281–297.

Harding, David J., Lisa Sanbonmatsu, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, Matthew Sciandra, and Jens Ludwig (2021), "Evaluating Contradictory Experimental and Non-Experimental Estimates of Neighborhood Effects on Economic Outcomes for Adults," Working Paper 28454, National Bureau of Economic Research.

Heblich, Stephan, Stephen J. Redding, and Daniel M. Sturm (2020), "The Making of the Modern Metropolis: Evidence from London," *Quarterly Journal of Economics*, Vol. 135, No. 4, pp. 2059–2133.

Hellerstein, Judith K., David Neumark, and Melissa McInerney (2008), "Spatial Mismatch or Racial Mismatch?," *Journal of Urban Economics*, Vol. 64, No. 2, pp. 464–479.

Houston, Donald S. (2001), "Testing the spatial mismatch hypothesis in the United Kingdom using evidence from firm relocations," in Pitfield, David E. ed. *Transport Planning, Logistics, and Spatial Mismatch: A Regional Science Perspective*, Vol. 11, London: Pion, pp. 134–151.

_ (2005), "Methods to Test the Spatial Mismatch Hypothesis," *Economic Geography*, Vol. 81, No. 4, pp. 407–434.

Ihlanfeldt, Keith R. and David L. Sjoquist (1989), "The Impact of Job Decentralization on the Economic Welfare of Central City Blacks," *Journal of Urban Economics*, Vol. 26, No. 1, pp. 110–130.

_ and Madelyn V. Young (1994), "Intrametropolitan Variation in Wage Rates: The Case of Atlanta Fast-Food Restaurant Workers," *Review of Economics and Statistics*, Vol. 76, No. 3, pp. 425–433.

Immergluck, Daniel (1998), "Job Proximity and the Urban Employment Problem: Do Suitable Nearby Jobs Improve Neighbourhood Employment Rates?," *Urban Studies*, Vol. 35, No. 1, pp. 7–23.

Institut des Politiques Publiques (2019), "Barèmes IPP."

Kline, Patrick M. and Enrico Moretti (2014), "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *Quarterly Journal of Economics*, Vol. 129, No. 1, pp. 275–331.

Kline, Patrick and Enrico Moretti (2013), "Place Based Policies with Unemployment," *American Economic Review: Papers and Proceedings*, Vol. 103, No. 3, pp. 238–243.

_ and _ (2014), "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs," *Annual Review of Economics*, Vol. 6, No. 1, pp. 629–662.

Le Corbusier (1943), *La Charte d'Athènes*: Plon.

Lynch, Devon and Jeffrey S. Zax (2011), "Incidence and Substitution in Enterprise Zone Programs: The Case of Colorado," *Public Finance Review*, Vol. 39, No. 2, pp. 226–255.

Mayer, Thierry, Florian Mayneris, and Loriane Py (2015), "The Impact of Urban Enterprise Zones on Establishment Location Decisions and Labor Market Outcomes: Evidence from France," *Journal of Economic Geography*, Vol. 17, No. 4, pp. 709–752.

McLafferty, Sara (1997), "Gender, Race, and the Determinants of Commuting: New York in 1990," *Urban Geography*, Vol. 18, No. 3, pp. 192–212.

Miller, Conrad (2018), "When Work Moves: Job Suburbanization and Black Employment," Working Paper.

Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg (2018), "Commuting, Migration, and Local Employment Elasticities," *American Economic Review*, Vol. 108, No. 12, pp. 3855–3890.

Mortensen, Dale T. and Christopher A. Pissarides (1994), "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, Vol. 61, No. 3, pp. 397–415.

Neumark, David and Jed Kolko (2010), "Do Enterprise Zones Create Jobs? Evidence from California's Enterprise Zone Program," *Journal of Urban Economics*, Vol. 68, No. 1, pp. 1–19.

Popkin, Susan J., James E. Rosenbaum, and Patricia M. Meaden (1993), "Labor Market Experiences of Low-Income Black Women in Middle-Class Suburbs: Evidence from a Survey of Gautreaux Program Participants," *Journal of Policy Analysis and Management*, Vol. 12, No. 3, pp. 556–573.

Poulhès, Mathilde (2015), "Are Enterprise Zones Benefits Capitalized into Commercial Property Values? The French case," Document de travail G2015/13, Direction des études et Synthèses économiques, INSEE.

Rathelot, Roland and Patrick Sillard (2008), "Zones Franches Urbaines : Quels Effets sur l'Emploi Salarié et les Créations d'Établissements ?," *Economie et Statistique*, Vol. 415, No. 1, pp. 81–96.

Redding, Stephen J. and Daniel M. Sturm (2008), "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review*, Vol. 98, No. 5, pp. 1766–1797.

Roback, Jennifer (1982), "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, Vol. 90, No. 6, pp. 1257–1278.

Rogers, Cynthia L. (1997), "Job Search and Unemployment Duration: Implications for the Spatial Mismatch Hypothesis," *Journal of Urban Economics*, Vol. 42, No. 1, pp. 109–132.

Rosen, Kenneth T. (1979), "A Regional Model of Multifamily Housing Starts," *Real Estate Economics*, Vol. 7, No. 1, pp. 63–76.

Rosenbaum, Emily and Laura E. Harris (2001), "Residential Mobility and Opportunities: Early Impacts of the Moving to Opportunity Demonstration Program in Chicago," *Housing Policy Debate*, Vol. 12, No. 2, pp. 321–346.

Rosenbaum, James E. (1995), "Changing the Geography of Opportunity by Expanding Residential Choice: Lessons from the Gautreaux Program," *Housing Policy Debate*, Vol. 6, No. 1, pp. 231–269.

Saiz, Albert (2010), "The Geographic Determinants of Housing Supply," *Quarterly Journal of Economics*, Vol. 125, No. 3, pp. 1253–1296.

Tsivanidis, Nick (2019), “Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá’s TransMilenio,” Job Market Paper.

Zenou, Yves (2009a), “Endogenous Job Destruction and Job Matching in Cities,” *Journal of Urban Economics*, Vol. 65, No. 3, pp. 323–336.

— (2009b), “Search in Cities,” *European Economic Review*, Vol. 53, No. 6, pp. 607–624.

— (2009c), “Search, Wage Posting and Urban Spatial Structure,” *Journal of Economic Geography*, Vol. 11, No. 3, pp. 387–416.

— (2009d), *Urban Labor Economics*, Cambridge: Cambridge University Press.

— (2009e), “Urban Search Models under High-Relocation Costs. Theory and Application to Spatial Mismatch,” *Labour Economics*, Vol. 16, No. 5, pp. 534–546.

Zhang, Zhongcai (1998), “Indirect Tests of the Spatial Mismatch Hypothesis in the Cleveland PMSA: A Labor Market Perspective,” *Urban Affairs Review*, Vol. 33, No. 5, pp. 712–723.

Appendix

1.A Institutional Appendix

	ZRU	ZFU
Corporate income tax		
Eligibility	only newly-created firms	all firms of at most 50 employees
Exemption	total for 2 years, then declining for 3 years	total for 5 years, then declining for 9 years
Payroll tax		
Eligibility	only newly-created jobs	all firms of at most 50 employees
Exemption	total for 1 year	total for 5 years, then declining for 3 to 9 years
	full exemption up to 1.4 minimum wage, then linearly decreasing; no exemption above 2.4 minimum wage	
Local business tax		
Eligibility	all establishments of at most 150 employees	all establishments of at most 50 employees
Exemption	total for 5 years, then declining for 3 years	total for 5 years, then declining for 3 to 9 years
Property tax		
Eligibility	all firms	
Exemption	5 years of full exemption	

Table 1.A.1: Comparison of exemptions available in ZRUs and ZFUs in 2009

1.B Omitted Proofs

1.B.1 Proofs of Section 1.2

Proposition 1.1 (Comparative Statics).

1. The higher $\bar{d}$, the higher the wage premium for commuting conditional on residence:

$$\frac{\partial\, w_{GC} - w_{GG}}{\partial\, \bar{d}} > 0, \quad \frac{\partial\, w_{CG} - w_{CC}}{\partial\, \bar{d}} > 0,$$

   where w_{ij} is the wage of workers living in i and employed in j.

2. The higher $\bar{\Delta}$, the higher the home bias in employment:

$$\frac{\partial\, l_{CC}}{\partial\, \bar{\Delta}} > 0, \quad \frac{\partial\, l_{GG}}{\partial\, \bar{\Delta}} > 0,$$

   where l_{ij} is the share of workers employed in j among residents of i.

Proof. Because of Nash bargaining, the wage of a worker living in location i and working in j is equal to:

$$w_{ij} = \beta W_j + (1-\beta)\big(Pd_{ij} + \Omega_i\big) \tag{1.B.1}$$

where W_j is the price of labor inputs in workplace j and $\Omega_i = \frac{(r+\delta)Pb+\sum_{j\in\{C,G\}}\lambda_{ij}(w_{ij}-Pd_{ij})}{r+\delta+\lambda_i}$ is the outside option of a job seeker living in i.

Thus the wage premium from commuting from $i \in \{C, G\}$ to $j \neq i$ is given by:

$$w_{ij} - w_{ii} = \beta\big(W_j - W_i\big) + (1-\beta)P\bar{d}. \tag{1.B.2}$$

As P is the numéraire and the only two factors used in good production are labor inputs and intermediate goods[15], the price of labor inputs in each location, (W_j), do not change in equilibrium. Indeed, the system:

$$\begin{cases} y_j Y(N_j, M_j) = Y_j \\ y_j \frac{\partial\, Y}{\partial\, M_j}(N_j, M_j) = 1 \end{cases} \tag{1.B.3}$$

fully determines factor intensities which are independent from the scale Y_j as Y is homogeneous of degree 1. Factor intensities in turn pin down the wage:

$$W_j = y_j \frac{\partial\, Y}{\partial\, N_j}(N_j, M_j), \tag{1.B.4}$$

as $\frac{\partial\, Y}{\partial\, N_j}$ is homogeneous of degree 0.

[15] I assume for simplicity that Y exhibits constant returns to scale.

Therefore, $\frac{\partial W_j}{\partial \bar{d}} = 0$ and equation (1.B.2) implies:

$$\frac{\partial\, w_{ij} - w_{ii}}{\partial \bar{d}} > 0. \tag{1.B.5}$$

To prove the second point, let's first notice that:

$$\begin{aligned}\frac{1}{l_{ii}}\frac{\partial\, l_{ii}}{\partial\, \bar{\Delta}} &= \frac{1}{\lambda_{ii}}\frac{\partial\, \lambda_{ii}}{\partial\, \bar{\Delta}} - \sum_{j\in\{C,G\}} l_{ij}\frac{1}{\lambda_{ij}}\frac{\partial\, \lambda_{ij}}{\partial\, \bar{\Delta}} \\ &= \bar{\Delta}^{-1} l_{i-i} + (1-\mu)\frac{1}{\theta_i}\frac{\partial\, \theta_i}{\partial\, \bar{\Delta}} - (1-\mu)\sum_{j\in\{C,G\}} l_{ij}\frac{1}{\theta_j}\frac{\partial\, \theta_j}{\partial\, \bar{\Delta}},\end{aligned} \tag{1.B.6}$$

for $-i \neq i$.

The general equilibrium effects on labor market tightness, $\left(\frac{\partial\, \theta_j}{\partial\, \bar{\Delta}}\right)_{j\in\{C,G\}}$, together with the derivative of the outside option, $\left(\frac{\partial\, \Omega_i}{\partial\, \bar{\Delta}}\right)_{i\in\{C,G\}}$, are the solutions to the system of four equations obtained by differentiating the free entry condition and the outside option definition[16]:

$$\begin{aligned}\frac{1}{\Omega_i}\frac{\partial\, \Omega_i}{\partial\, \bar{\Delta}} =& \frac{\beta \sum_{j\in\{C,G\}} \lambda_{ij}(W_j - Pd_{ij})\Big((1-\mu)\frac{1}{\theta_j}\frac{\partial\, \theta_j}{\partial\, \bar{\Delta}} - \bar{\Delta}^{-1} 1_{j\neq i}\Big)}{(r+\delta)Pb + \beta \sum_{j\in\{C,G\}} \lambda_{ij}(W_j - Pd_{ij})} \\ &- \frac{\beta\bar{\lambda}_i}{r+\delta+\beta\bar{\lambda}_i}\sum_{j\in\{C,G\}}\frac{\lambda_{ij}}{\bar{\lambda}_i}\Big((1-\mu)\frac{1}{\theta_j}\frac{\partial\, \theta_j}{\partial\, \bar{\Delta}} - \bar{\Delta}^{-1} 1_{j\neq i}\Big), \quad i\in\{C,G\},\end{aligned} \tag{1.B.7}$$

and:

$$\begin{aligned}\mu\frac{1}{\theta_j}\frac{\partial\, \theta_j}{\partial\, \bar{\Delta}} =& -\frac{\sum_i L_i u_i \Delta_{ij}^{-1}\Big(\Big(W_j - Pd_{ij} - \Omega_i\Big)\Big[\bar{\Delta}^{-1}1_{j\neq i} + \sum_{j'}\lambda_{ij'}\Big((1-\mu)\frac{1}{\theta_{j'}}\frac{\partial\, \theta_{j'}}{\partial\, \bar{\Delta}} - \bar{\Delta}^{-1}1_{j'\neq i}\Big)\Big] + \frac{\partial\, \Omega_i}{\partial\, \bar{\Delta}}\Big)}{\sum_i L_i u_i \Delta_{ij}^{-1}\Big(W_j - Pd_{ij} - \Omega_i\Big)} \\ &+ \frac{\sum_i L_i u_i \Delta_{ij}^{-1} P\nu\Big(\bar{\Delta}^{-1}1_{j\neq i} + \sum_{j'}\lambda_{ij'}\Big((1-\mu)\frac{1}{\theta_{j'}}\frac{\partial\, \theta_{j'}}{\partial\, \bar{\Delta}} - \bar{\Delta}^{-1}1_{j'\neq i}\Big)\Big)}{\sum_i L_i u_i \Delta_{ij}^{-1} P\nu}, \quad j\in\{C,G\},\end{aligned} \tag{1.B.8}$$

where I used again the fact that $\frac{\partial W_j}{\partial \bar{\Delta}} = 0$.

The result follows after solving for $\left(\frac{\partial\, \theta_j}{\partial\, \bar{\Delta}}\right)_{j\in\{C,G\}}$ and plugging in into equation (1.B.6).

□

Definition 1.1 (Planner's Problem). The planner's problem is:

$$\max\ \mathcal{W}, \tag{1.2.9}$$

[16]Note that $\Omega_i = \frac{(r+\delta)Pb + \sum_{j\in\{C,G\}}\lambda_{ij}(w_{ij} - Pd_{ij})}{r+\delta+\bar{\lambda}_i} = \frac{(r+\delta)Pb + \beta\sum_{j\in\{C,G\}}\lambda_{ij}(W_j - Pd_{ij})}{r+\delta+\beta\bar{\lambda}_i}$ from Nash bargaining.

subject to:

(i) spatial mobility constraints;

(ii) tradable good and housing feasibility constraints;

(iii) search and matching constraints;

(iv) labor market clearing;

(v) population constraint.

Formal statement.

$$\max \int_0^\infty e^{-rt} \sum_i a_i L_i \left[u_i u(c_i^U, h_i^U) + \sum_j l_{ij} u(c_{ij}^E, h_{ij}^E) \right] dt$$

subject to:

$$L_i u_i J_i^U \geq L_i u_i \mathcal{J} \tag{χ_i}$$

$$\sum_{i \in \mathcal{I}} Y_i((N_i^Y), M_i^Y) \geq M_i^Y + M_i^H + L_i \Big(u_i \big(c_i^U - b\big) + \sum_j l_{ij} \big(c_{ij}^E + d_{ij}\big) \Big) + \nu \sum_l \Delta_{li}^{-1} L_l u_l \theta_i \tag{P_i^*}$$

$$H_i((N_i^H), M_i^H) \geq L_i \Big(u_i h_i^U + \sum_j l_{ij} h_{ij}^E \Big) \tag{R_i^*}$$

$$\left(\Delta_{ij}^{-1} \theta_j q(\theta_j) u_i - \delta l_{ij} \right) L_i = L_i \dot{l}_{ij} + l_{ij} \dot{L}_i \tag{μ_{ij}}$$

$$r L_i u_i J_i^U - L_i u_i a_i u(c_i^U, h_i^U) - L_i u_i \sum_j \Delta_{ij}^{-1} \theta_j q(\theta_j) \left(J_{ij}^E - J_i^U \right) = L_i u_i \dot{J}_i^U \tag{ξ_i^U}$$

$$r L_i l_{ij} J_{ij}^E - L_i l_{ij} a_i u(c_{ij}^E, h_{ij}^E) - L_i l_{ij} \delta \left(J_i^U - J_{ij}^E \right) = L_i l_{ij} \dot{J}_{ij}^E \tag{ξ_{ij}^E}$$

$$\sum_i l_{ij} L_i \geq N_j^Y + N_j^H \tag{W_j^*}$$

$$L_i \geq \left(u_i + \sum_j l_{ij} \right) L_i \tag{ϕ_i}$$

$$\bar{L} \geq \sum_i L_i \tag{Λ}$$

□

Proposition 1.2 (Constrained-Efficient Allocation). The constrained-efficient allocation solution to the planner's problem 1.1 satisfies the following two conditions.

1. Workers' optimal choice of residence:

$$\omega a_i \Big[u_i u(c_i^U, h_i^U) + \textstyle\sum_j l_{ij} u(c_{ij}^E, h_{ij}^E) \Big] + \sum_j l_{ij} W_j - u_i \sum_j \Delta_{ij}^{-1} \theta_j \nu = u_i x_i^U + \sum_j l_{ij} x_{ij}^E + \Lambda, \tag{1.2.10}$$

   with W_j and θ_j the price of labor inputs and the labor market tightness in location j, x_i^U and x_{ij}^E the expenditures of unemployed and employed workers, ω the Pareto weight and Λ the opportunity cost of a worker.

2. Recruiters' optimal entry:

$$\nu \sum_i \Delta_{ij}^{-1} u_i L_i = \tfrac{1-\mu(\theta_j)}{\theta_j} \sum_i \tfrac{\lambda_{ij}}{r+\delta+\bar{\lambda}_i} L_i \Big(u_i \big[W_j - d_{ij} - b + \textstyle\sum_{j'} \Delta_{ij'}^{-1} \theta_{j'} \nu \big] + \tfrac{\delta}{r+\delta} \sum_{j'} l_{ij'} \big[W_j - d_{ij} - \big(W_{j'} - d_{ij'} \big) \big] \Big), \tag{1.2.11}$$

   with $\lambda_{ij} = \Delta_{ij}^{-1} \theta_j q(\theta_j)$ the job finding rate in workplace j of a job seeker living in i and $\mu(\theta_j)$ the elasticity of the matching function m with respect to unemployment.

Proof. See proof of proposition 1.5, which is a generalization of proposition 1.2. □

Proposition 1.4 (Optimal Policy). Constrained efficiency may be restored with a mix of two place-based policies, a residence subsidy, t_i^R, and a hiring subsidy, t_j^H, that can be expressed as:

$$t_i^R = \sum_{j \in \{C,G\}} \frac{\lambda_{ij}}{r+\delta+\bar{\lambda}_i} \Big(\pi_{ij} - \mathbb{E}_{|j} \big[\pi_{i'j} \big] - t_j^H \Big) \tag{1.2.12}$$

$$= -(1-\beta) \sum_{j \in \{C,G\}} \frac{\lambda_{ij}}{r+\delta+\bar{\lambda}_i} \Big(d_{ij} + \Omega_i - \mathbb{E}_{|j} \big[d_{i'j} + \Omega_{i'} \big] - t_j^H \Big) \tag{1.2.13}$$

and:

$$t_j^H = -\mathbb{E}_{|j} \big[t_{i'}^R \big], \tag{1.2.14}$$

where $\Omega_i = \frac{(r+\delta)b + \sum_j \lambda_{ij}(w_{ij} - d_{ij})}{r+\delta+\bar{\lambda}_i}$ is the outside option of a worker living in i, $\pi_{ij} = (1-\beta)(W_j - d_{ij} - \Omega_i)$ is the profits of a recruiter in j when matched with a worker living in i and $\mathbb{E}_{|j}\left[X_{ij}\right] = \frac{\sum_{i'} \Delta_{i'j}^{-1} U_i X_{i'j}}{\sum_{i'} \Delta_{i'j}^{-1} U_{i'}}$ is the expected value of $X_{i'j}$ over matches for a recruiter in j.

Proof. See proof of proposition **??**, which is a generalization of proposition 1.4. □

Proposition 1.3 (Efficiency of the Decentralized Equilibrium). Assume the Hosios condition $\mu(\theta_j) = \beta$, with β workers' bargaining power. Then there exists a increasing function $\bar{\Delta}_0 : \bar{d} \mapsto \bar{\Delta}_0(\bar{d})$ satisfying $\bar{\Delta}_0(0) = 0$ and such that:

1. If $\bar{d} \geq 0$ and $\bar{\Delta} = \bar{\Delta}_0(\bar{d})$, the decentralized equilibrium coincides with the constrained-efficient allocation;

2. If $\bar{d} > 0$ and $\bar{\Delta} < \bar{\Delta}_0(\bar{d})$, the decentralized equilibrium features too much spatial mismatch, in the sense that $\frac{L_C^{DE}}{L_G^{DE}} < \frac{L_C^*}{L_G^*}$ and $\frac{\theta_G^{DE}}{\theta_C^{DE}} < \frac{\theta_G^*}{\theta_C^*}$.
3. If $\bar{d} \geq 0$ and $\bar{\Delta} > \bar{\Delta}_0(\bar{d})$, the decentralized equilibrium features too little spatial mismatch, in the sense that $\frac{L_C^{DE}}{L_G^{DE}} > \frac{L_C^*}{L_G^*}$ and $\frac{\theta_G^{DE}}{\theta_C^{DE}} > \frac{\theta_G^*}{\theta_C^*}$.

Proof. From proposition 1.4, we know that if $\bar{d} > 0$ and $\bar{\Delta} = 0$, then $t_C^R - t_G^R > 0$. If $\bar{d}$ is fixed and $\bar{\Delta}$ is high enough, then $t_C^R - t_G^R < 0$ By continuity of the model, we can define $\bar{\Delta}_0(\bar{d})$ as the value of $\bar{\Delta}$ such that $t_C^R - t_G^R = 0$. □

1.B.2 Proofs of Section 1.4

Definition 1.4 (Planner's Problem). Let $g_0 \in G$. The planner's problem is:

$$\max \ \mathcal{W}_{g_0}, \tag{1.4.29}$$

subject to:

(i) required utility levels for skill groups $g \neq g_0$;
(ii) spatial mobility constraints;
(iii) tradable good and housing feasibility constraints;
(iv) search and matching constraints;
(v) labor market clearing;
(vi) population constraints.

Formal statement.

$$\max \ \mathcal{W}_{g_0}$$

subject to:

$$\int_0^\infty e^{-rt} \sum_i \tilde{a}_{g,i} \Big[u_{g,i} u(c_{g,i}^U, h_{g,i}^U) + \sum_j \int_\alpha l_{g,ij}(\alpha) u(c_{g,ij}^E(\alpha), h_{g,ij}^E(\alpha)) d\alpha \Big] L_{g,i} dt \geq \begin{cases} \mathcal{W}_{g_0} & \text{if } g = g_0 \\ \underline{\mathcal{W}}_g & \text{otherwise} \end{cases} \tag{ω_g}$$

$$\left(u_{g,i} L_{g,i}\right)^{1-\sigma} r J_{g,i}^U \geq u_{g,i} L_{g,i} \mathcal{J}_g \tag{$\chi_{g,i}^J$}$$

$$a_{g,i} \left(u_{g,i} L_{g,i}\right)^{1-\sigma} \geq \tilde{a}_{g,i} u_{g,i} L_{g,i} \mathcal{A}_g \tag{$\chi_{g,i}^A$}$$

$$Y_i\big((N_{g,i}^Y)_g, M_i^Y\big) \geq \sum_j \kappa_{ij} Q_{ij} \tag{p_i^*}$$

$$Q(Q_{1i},\dots,Q_{Ii}) \geq M_i^Y + M_i^H + \sum_g \Big(u_{g,i}\big(c_{g,i}^U - b\big) + \sum_j \int_\alpha l_{g,ij}(\alpha)\big(c_{g,ij}^E(\alpha) + d_{ij}\big)d\alpha\Big)L_{g,i} + \nu \sum_{g,l} \Delta_{li}^{-1}\theta_{g,i}u_{g,l}L_{g,l} \tag{P_i^*}$$

$$H_i\big((N_{g,i}^H)_g, M_i^H\big) \geq \sum_g \Big(u_{g,i}h_{g,i}^U + \sum_j \int_\alpha l_{g,ij}(\alpha)h_{g,ij}^E(\alpha)d\alpha\Big)L_{g,i} \tag{R_i^*}$$

$$\Big(\Delta_{ij}^{-1}\theta_{g,j}q(\theta_{g,j})u_{g,i}g(\alpha)1_{\alpha \geq \alpha_{g,ij}^R} - \delta l_{g,ij}(\alpha)\Big)L_{g,i} = L_{g,i}\dot{l}_{g,ij}(\alpha) + l_{g,ij}(\alpha)\dot{L}_{g,i} \tag{$\mu_{g,ij}(\alpha)$}$$

$$\sum_i \Big(\int_{\alpha \geq \alpha_{g,ij}^R} \alpha g(\alpha)d\alpha\Big)\Delta_{ij}^{-1}\theta_{g,j}q(\theta_{g,j})u_{g,i}L_{g,i} - \delta N_{g,j} = \dot{N}_{g,j} \tag{$\zeta_{g,j}$}$$

$$N_{g,j} \geq N_{g,j}^Y + N_{g,j}^H \tag{$W_{g,j}^*$}$$

$$L_{g,i} \geq \Big(u_{g,i} + \sum_j \int_\alpha l_{g,ij}(\alpha)d\alpha\Big)L_{g,i} \tag{$\phi_{g,i}$}$$

$$\bar{L}_g \geq \sum_i L_{g,i} \tag{Λ_g}$$

$$ru_{g,i}L_{g,i}J_{g,i}^U - u_{g,i}L_{g,i}a_iu(c_{g,i}^U, h_{g,i}^U) - u_{g,i}L_{g,i}\sum_j \Delta_{ij}^{-1}\theta_{g,j}q(\theta_{g,j})\int_{\alpha \geq \alpha_{g,ij}^R} g(\alpha)\left(J_{g,ij}^E(\alpha) - J_{g,i}^U\right)d\alpha = u_{g,i}L_{g,i}\dot{J}_{g,i}^U \tag{$\xi_{g,i}^U$}$$

$$rl_{g,ij}(\alpha)L_{g,i}J_{g,ij}^E(\alpha) - l_{g,ij}(\alpha)L_{g,i}a_iu(c_{g,ij}^E(\alpha), h_{g,ij}^E(\alpha)) - l_{g,ij}(\alpha)L_{g,i}\delta\left(J_{g,i}^U - J_{g,ij}^E(\alpha)\right) = l_{g,ij}(\alpha)L_{g,i}\dot{J}_{g,ij}^E(\alpha) \tag{$\xi_{g,ij}^E(\alpha)$}$$

□

Proposition 1.5 (Constrained-Efficient Allocation). The constrained-efficient allocation solution to the planner's problem 1.4 satisfies the following three conditions.

1. Workers' optimal choice of residence:

$$\Lambda_g = \omega_g a_i \frac{(r+\delta)u(c_{g,i}^U, h_{g,i}^U) + \sum_j \lambda_{g,ij}u(c_{g,ij}^E(\alpha_{g,ij}^E), h_{g,ij}^E(\alpha_{g,ij}^E))}{r+\delta+\bar{\lambda}_{g,i}} - \frac{(r+\delta)x_{g,i}^U + \sum_j \lambda_{g,ij}x_{g,ij}^E(\alpha_{g,ij}^E)}{r+\delta+\bar{\lambda}_{g,i}} + \frac{(r+\delta)\Big(P_ib - \sum_j \Delta_{ij}^{-1}\theta_{g,j}P_j\nu\Big) + \sum_j \lambda_{g,ij}\Big(\alpha_{g,ij}^E W_{g,j} - P_id_{ij}\Big)}{r+\delta+\bar{\lambda}_{g,i}}, \tag{1.4.30}$$

with $W_{g,j}$ and $\theta_{g,j}$ the price of labor inputs and the labor market tightness for skill group g in location j, $x^U_{g,i}$ and $x^E_{g,ij}$ the expenditures of unemployed and employed workers, ω_g the Pareto weight and Λ_g the opportunity cost of a worker.

2. Recruiters' optimal entry:

$$\begin{aligned} & P_j \nu_g \sum_i \Delta_{ij}^{-1} u_{g,i} L_{g,i} \\ &= \tfrac{1-\mu(\theta_{g,j})}{\theta_{g,j}} \textstyle\sum_i \tfrac{\lambda_{g,ij}}{r+\delta+\bar{\lambda}_{g,i}} L_{g,i} \Big(u_{g,i} \left[\alpha^E_{g,ij} W_{g,j} - P_i d_{ij} - P_i b \right] + u_{g,i} \textstyle\sum_{j'} \Delta_{ij'}^{-1} \theta_{g,j'} P_{j'} \nu_g \\ & \quad + \frac{\delta}{r+\delta} \sum_{j'} l_{g,ij'} \left[\alpha^E_{g,ij} W_{g,j} - P_i d_{ij} - \left(\alpha^E_{g,ij'} W_{g,j'} - P_i d_{ij'} \right) \right] \Big) \end{aligned} \tag{1.4.31}$$

with $\lambda_{g,ij} = \big(1 - G(\alpha^R_{g,ij})\big) \Delta_{ij}^{-1} \theta_j q(\theta_j)$ the job finding rate in workplace j of a type-g job seeker living in i and $\mu(\theta_j)$ the elasticity of the matching function m with respect to unemployment.

3. Optimal job acceptance:

$$\alpha^R_{g,ij} W_{g,j} - P_i d_{ij} = \frac{(r+\delta)\Big(P_i b - \sum_{j'} \Delta_{ij'}^{-1} \theta_{g,j'} P_{j'} \nu\Big) + \sum_{j'} \lambda_{g,ij'} \Big(\alpha^E_{g,ij'} W_{g,j'} - P_i d_{ij'}\Big)}{r+\delta+\bar{\lambda}_{g,i}}. \tag{1.4.32}$$

Proof. To obtain the three conditions that characterize the constrained-efficient allocation, I start by laying out the first order conditions derived from the planner's problem 1.4. I then solve for the various Lagrange multipliers. Finally, rearranging the FOCs with respect to local population, $L_{g,i}$, labor market tightness, $\theta_{g,j}$, and reservation productivities, $\alpha^R_{g,ij}$, leads to the desired conditions.

Step 1: First-order conditions

$$\begin{aligned} & r \textstyle\sum_j \int_{\alpha \geq \alpha^R_{g,ij}} l_{g,ij}(\alpha) \mu_{g,ij}(\alpha) d\alpha - \sum_j \int_{\alpha \geq \alpha^R_{g,ij}} l_{g,ij}(\alpha) \dot{\mu}_{g,ij}(\alpha) d\alpha = \omega_g a_i \left[u_{g,i} u\left(c^U_{g,i}, h^U_{g,i}\right) + \sum_j \int_{\alpha \geq \alpha^R_{g,ij}} l_{g,ij}(\alpha) u(c^E_{g,ij}(\alpha), h^E_{g,ij}(\alpha)) d\alpha \right] \\ & - P_i^* \Big(u_{g,i} (c^U_{g,i} - b) + \textstyle\sum_j \int_{\alpha \geq \alpha^R_{g,ij}} l_{g,ij}(\alpha) \big(c^E_{g,ij}(\alpha) + d_{ij} \big) \Big) - R_i^* \Big(u_{g,i} h^U_{g,i} + \sum_j \int_{\alpha \geq \alpha^R_{g,ij}} l_{g,ij}(\alpha) h^E_{g,ij}(\alpha) \Big) + \sum_j \lambda_{g,ij} u_{g,i} \alpha^E_{g,ij} \zeta_{g,j} - \sum_j \Delta_{ij}^{-1} u_{g,i} \theta_{g,j} P_j^* \nu - \Lambda_g - \sigma \chi^U_{g,i} u_{g,i} \mathcal{U}_{g,i} \end{aligned} \tag{$L_{g,i}$}$$

$$0 = a_i l_{g,ij}(\alpha) \left[\omega_g - \xi^E_{g,ij}(\alpha) \right] \frac{\partial u}{\partial c} (c^E_{g,ij}(\alpha), h^E_{g,ij}(\alpha)) - P_i^* l_{g,ij}(\alpha) \tag{$c^E_{g,ij}(\alpha)$}$$

$$0 = a_i l_{g,ij}(\alpha) \left[\omega_g - \xi^E_{g,ij}(\alpha) \right] \frac{\partial u}{\partial h} (c^E_{g,ij}(\alpha), h^E_{g,ij}(\alpha)) - R_i^* l_{g,ij}(\alpha) \tag{$h^E_{g,ij}(\alpha)$}$$

$$0 = a_i u_{g,i} \left[\omega_g - \xi^U_{g,i} \right] \frac{\partial u}{\partial c} (c^U_{g,i}, h^U_{g,i}) - P_i^* u_{g,i} \tag{$c^U_{g,i}$}$$

$$0 = a_i u_{g,i} \left[\omega_g - \xi^U_{g,i} \right] \frac{\partial u}{\partial h} (c^U_{g,i}, h^U_{g,i}) - R_i^* u_{g,i} \tag{$h^U_{g,i}$}$$

$$\left(r\xi_{g,i}^{U}-\dot{\xi}_{g,i}^{U}\right)u_{g,i}=r\xi_{g,i}^{U}u_{g,i}+\bar{\lambda}_{g,i}\xi_{g,i}^{U}u_{g,i}-\delta\sum_{j}\int_{\alpha\geq\alpha_{g,ij}^{R}}l_{g,ij}(\alpha)\xi_{g,ij}^{E}(\alpha)d\alpha+r(u_{g,i}L_{g,i})^{-\sigma}\chi_{g,i}^{U}u_{g,i} \tag{$J_{g,i}^{U}$}$$

$$\left(r\xi_{g,ij}^{E}(\alpha)-\dot{\xi}_{g,ij}^{E}(\alpha)\right)l_{g,ij}(\alpha)=r\xi_{g,ij}^{E}(\alpha)l_{g,ij}(\alpha)+\delta\xi_{g,ij}^{E}(\alpha)l_{g,ij}(\alpha)-\frac{g(\alpha)}{1-G(\alpha_{g,ij}^{R})}\lambda_{g,ij}\xi_{g,i}^{U}u_{g,i} \tag{$J_{g,ij}^{E}(\alpha)$}$$

$$0=\sum_{i}L_{g,i}u_{g,i}\chi_{g,i}^{U} \tag{$\mathcal{U}_{g}$}$$

$$p_{i}^{*}\frac{\partial Y_{i}}{\partial N_{g,i}^{Y}}=W_{g,i}^{*} \tag{$N_{g,i}^{Y}$}$$

$$R_{i}^{*}\frac{\partial H_{i}}{\partial N_{g,i}^{H}}=W_{g,i}^{*} \tag{$N_{g,i}^{H}$}$$

$$P_{j}^{*}\frac{\partial Q}{\partial Q_{ij}}=p_{i}^{*}\kappa_{ij} \tag{Q_{ij}}$$

$$p_{i}^{*}\frac{\partial Y_{i}}{\partial M_{i}^{Y}}=P_{i}^{*} \tag{M_{i}^{Y}}$$

$$R_{i}^{*}\frac{\partial H_{i}}{\partial M_{i}^{H}}=P_{i}^{*} \tag{M_{i}^{H}}$$

$$0=\frac{1-\mu(\theta_{g,j})}{\theta_{g,j}}\sum_{i}\lambda_{g,ij}u_{g,i}L_{g,i}\int_{\alpha\geq\alpha_{g,ij}^{R}}\left(\mu_{g,ij}(\alpha)+\alpha\zeta_{g,j}-\xi_{g,i}^{U}\left(J_{g,ij}^{E}(\alpha)-J_{g,i}^{U}\right)\right)\frac{g(\alpha)}{1-G(\alpha_{g,ij}^{R})}d\alpha-\sum_{i}\Delta_{ij}^{-1}u_{g,i}L \tag{$\theta_{g,j}$}$$

$$0=\omega_{g}a_{i}u(c_{g,i}^{U},h_{g,i}^{U})-P_{i}^{*}\left(c_{g,i}^{U}-b\right)-R_{i}^{*}h_{g,i}^{U}-\sum_{j}P_{j}^{*}\nu\Delta_{ij}^{-1}\theta_{g,j}+\sum_{j}\int_{\alpha\geq\alpha_{g,ij}^{R}}\lambda_{g,ij}\mu_{g,ij}(\alpha)\frac{g(\alpha)}{1-G(\alpha_{g,ij}^{R})}d\alpha+\sum_{j}\lambda_{g,ij}\alpha_{g,ij}^{E}\zeta_{g,j}-\phi_{g,i}-\sigma\chi_{g,i}^{U}\mathcal{U}_{g,i} \tag{$u_{g,i}$}$$

$$r\mu_{g,ij}(\alpha)-\dot{\mu}_{g,ij}(\alpha)=\omega_{g}a_{i}u(c_{g,ij}^{E}(\alpha),h_{g,ij}^{E}(\alpha))-P_{i}^{*}\left(c_{g,ij}^{E}(\alpha)+d_{ij}\right)-R_{i}^{*}h_{g,ij}^{E}(\alpha)-\delta\mu_{g,ij}(\alpha)-\phi_{g,i} \tag{$l_{g,ij}(\alpha)$}$$

$$r\zeta_{g,j}-\dot{\zeta}_{g,j}=W_{g,j}^{*}-\delta\zeta_{g,j} \tag{$N_{g,j}$}$$

$$\xi^U_{g,i}\Big(J^E_{g,ij}(\alpha^R_{g,ij}) - J^U_{g,i}\Big) - \mu_{g,ij}(\alpha^R_{g,ij}) - \alpha^R_{g,ij}\zeta_{g,j} = 0 \qquad (\alpha^R_{g,ij})$$

Step 2: Solving for Lagrange multipliers

⋆ Multipliers $\chi^U_{g,i}$, $\xi^U_{g,i}$ and $\xi^E_{g,ij}$

$$\delta\xi^E_{g,ij}(\alpha)l_{g,ij}(\alpha) = \frac{g(\alpha)}{1-G(\alpha^R_{g,ij})}\lambda_{g,ij}\xi^U_{g,i}u_{g,i}$$

$$\xi^E_{g,ij}(\alpha) = \xi^U_{g,i}$$

$$\chi^U_{g,i} = 0$$

⋆ Expenditure

$$\begin{aligned} x^{E,*}_{g,ij}(\alpha) &\equiv P^*_i c^E_{g,ij}(\alpha) + R^*_i h^E_{g,ij}(\alpha) \\ &= \left[\omega_g - \xi^E_{g,ij}(\alpha)\right] a_i u(c^E_{g,ij}(\alpha), h^E_{g,ij}(\alpha)) \\ &= \left[\omega_g - \xi^U_{g,i}\right] a_i u(c^E_{g,ij}(\alpha), h^E_{g,ij}(\alpha)) \end{aligned}$$

$$\begin{aligned} x^{U,*}_{g,i} &\equiv P^*_i c^U_{g,ij} + R^*_i h^U_{g,ij} \\ &= \left[\omega_g - \xi^U_{g,i}\right] a_i u(c^U_{g,i}, h^U_{g,i}) \end{aligned}$$

⋆ Multipliers ω_g and $\xi^U_{g,i}$

$$\omega_g - \xi^U_{g,i} = \psi_i/a_i$$

$$\xi^U_{g,i} = \omega_g - \psi_i/a_i$$

⋆ Value functions $J^E_{g,ij}(\alpha)$ and $J^U_{g,i}$

$$\sum_j \lambda_{g,ij}\int_{\alpha\geq\alpha^R_{g,ij}}\left(J^E_{g,ij}(\alpha) - J^U_{g,i}\right)\frac{g(\alpha)}{1-G(\alpha^R_{g,ij})}d\alpha = \frac{\sum_j \lambda_{g,ij}a_i\int_{\alpha\geq\alpha^R_{g,ij}}\left(u(c^E_{g,ij}(\alpha),h^E_{g,ij}(\alpha)) - u(c^U_{g,i},h^U_{g,i})\right)\frac{g(\alpha)}{1-G(\alpha^R_{g,ij})}d\alpha}{r+\delta+\bar{\lambda}_{g,i}}$$

$$rJ^U_{g,i} = a_i\frac{(r+\delta)u(c^U_{g,i},h^U_{g,i}) + \sum_j \lambda_{g,ij}\int_{\alpha\geq\alpha^R_{g,ij}} u(c^E_{g,ij}(\alpha),h^E_{g,ij}(\alpha))\frac{g(\alpha)}{1-G(\alpha^R_{g,ij})}d\alpha}{r+\delta+\bar{\lambda}_{g,i}}$$

$$
\begin{aligned}
& (r+\delta)\left(J^E_{g,ij}(\alpha) - J^U_{g,i}\right) \\
= \quad & a_i\left[u(c^E_{g,ij}(\alpha), h^E_{g,ij}(\alpha)) - u(c^U_{g,i}, h^U_{g,i})\right] - \textstyle\sum_{j'} \lambda_{g,ij'} \int_{\alpha' \geq \alpha^R_{g,ij}} \left(J^E_{g,ij'}(\alpha') - J^U_{g,i}\right) \frac{g(\alpha')}{1-G(\alpha^R_{g,ij})} d\alpha' \\
= \quad & a_i u(c^E_{g,ij}(\alpha), h^E_{g,ij}(\alpha)) - a_i \frac{(r+\delta)u(c^U_{g,i}, h^U_{g,i}) + \sum_{j'} \lambda_{g,ij'} \int_{\alpha' \geq \alpha^R_{g,ij'}} u(c^E_{g,ij'}(\alpha'), h^E_{g,ij'}(\alpha')) \frac{g(\alpha')}{1-G(\alpha^R_{g,ij})} d\alpha'}{r+\delta+\bar{\lambda}_{g,i}}
\end{aligned}
$$

⋆ Value functions $J^F_{g,ij}$ and $J^V_{g,j}$

$$
q(\theta_{g,j})\mathbb{E}_j\left[J^F_{g,ij}(\alpha) - J^V_{g,j}\right] = \mathbb{E}_j\left[\pi_{g,ij}(\alpha) - P^*_j \nu\right] - (r+\delta)\mathbb{E}_j\left[J^F_{g,ij}(\alpha) - J^V_{g,j}\right]
$$

$$
rJ^V_{g,j} = \frac{(r+\delta)P^*_j \nu + q(\theta_{g,j})\mathbb{E}_j\left[\pi_{g,ij}(\alpha)\right]}{r+\delta+q(\theta_{g,j})}
$$

$$
\begin{aligned}
(r+\delta)\mathbb{E}_j\left[J^F_{g,ij}(\alpha) - J^V_{g,j}\right] \quad &= \quad \mathbb{E}_j\left[\pi_{g,ij}(\alpha) - P^*_j \nu\right] - q(\theta_{g,j})\mathbb{E}_j\left[J^F_{g,ij}(\alpha) - J^V_{g,j}\right] \\
&= \quad \pi_{g,ij}(\alpha) - \frac{(r+\delta)P^*_j \nu + q(\theta_{g,j})\mathbb{E}_j\left[\pi_{g,ij}(\alpha)\right]}{r+\delta+q(\theta_{g,j})}
\end{aligned}
$$

⋆ Multipliers $\phi_{g,i}$ and $\mu_{g,ij}$

$$
\begin{aligned}
\phi_{g,i} \quad = \quad & \omega_g a_i \frac{(r+\delta)u(c^U_{g,i}, h^U_{g,i}) + \sum_j \lambda_{g,ij} \int_{\alpha \geq \alpha^R_{g,ij}} u(c^E_{g,ij}(\alpha), h^E_{g,ij}(\alpha)) \frac{g(\alpha)}{1-G(\alpha^R_{g,ij})} d\alpha}{r+\delta+\bar{\lambda}_{g,i}} \\
& - \frac{(r+\delta)x^U_{g,i} + \sum_j \lambda_{g,ij} \int_{\alpha \geq \alpha^R_{g,ij}} x^E_{g,ij}(\alpha) \frac{g(\alpha)}{1-G(\alpha^R_{g,ij})} d\alpha}{r+\delta+\bar{\lambda}_{g,i}} \\
& + \frac{(r+\delta)\left(P^*_i b - \sum_j \Delta^{-1}_{ij} \theta_{g,j} P^*_j \nu\right) + \sum_j \lambda_{g,ij} \int_{\alpha \geq \alpha^R_{g,ij}} \left(\alpha W^*_{g,j} - P^*_i d_{ij}\right) \frac{g(\alpha)}{1-G(\alpha^R_{g,ij})} d\alpha}{r+\delta+\bar{\lambda}_{g,i}}
\end{aligned}
$$

$$
\begin{aligned}
& (r+\delta)\left(\mu_{g,ij}(\alpha)+\alpha\zeta_{g,j}\right) \\
= \quad & \omega_g a_i \left[u(c^E_{g,ij}(\alpha), h^E_{g,ij}(\alpha)) - \frac{(r+\delta)u(c^U_{g,i},h^U_{g,i})+\sum_{j'}\lambda_{g,ij'}\int_{\alpha'\geq\alpha^R_{g,ij'}} u(c^E_{g,ij'}(\alpha'),h^E_{g,ij'}(\alpha'))\frac{g(\alpha')}{1-G(\alpha^R_{g,ij'})}d\alpha'}{r+\delta+\bar{\lambda}_{g,i}} \right] \\
& - \left[x^E_{g,ij}(\alpha) - \frac{(r+\delta)x^U_{g,i}+\sum_{j'}\lambda_{g,ij'}\int_{\alpha'\geq\alpha^R_{g,ij'}} x^E_{g,ij'}(\alpha')\frac{g(\alpha')}{1-G(\alpha^R_{g,ij'})}d\alpha'}{r+\delta+\bar{\lambda}_{g,i}} \right] \\
& + \left[\alpha W^*_{g,j} - P^*_i d_{ij} - \frac{(r+\delta)\left(P^*_i b - \sum_{j'}\Delta^{-1}_{ij'}\theta_{g,j'}P^*_{j'}\nu\right)+\sum_{j'}\lambda_{g,ij'}\int_{\alpha'\geq\alpha^R_{g,ij'}}\left(\alpha' W^*_{g,j'} - P^*_i d_{ij'}\right)\frac{g(\alpha')}{1-G(\alpha^R_{g,ij'})}d\alpha'}{r+\delta+\bar{\lambda}_{g,i}} \right] \\
= \quad & (r+\delta)\left(\omega_g - \psi_i/a_i\right)\left(J^E_{g,ij}(\alpha) - J^U_{g,i}\right) + \left[\alpha W^*_{g,j} - P^*_i d_{ij} - \frac{(r+\delta)\left(P^*_i b - \sum_{j'}\Delta^{-1}_{ij'}\theta_{g,j'}P^*_{j'}\nu\right)+\sum_{j'}\lambda_{g,ij'}\int_{\alpha'\geq\alpha^R_{g,ij'}}\left(\alpha' W^*_{g,j'} - P^*_i d_{ij'}\right)\frac{g(\alpha')}{1-G(\alpha^R_{g,ij})}d\alpha'}{r+\delta+\bar{\lambda}_{g,i}} \right]
\end{aligned}
$$

Step 3: Deriving optimality conditions

⋆ Workers' optimal residence choice

↪ Substitute for multipliers in the FOC with respect to population, $L_{g,i}$.

$$
\begin{aligned}
\Lambda_g = \omega_g a_i \frac{(r+\delta)u(c^U_{g,i},h^U_{g,i})+\sum_j \lambda_{g,ij} u(c^E_{g,ij}(\alpha^E_{g,ij}),h^E_{g,ij}(\alpha^E_{g,ij}))}{r+\delta+\bar{\lambda}_{g,i}} - \frac{(r+\delta)x^U_{g,i}+\sum_j \lambda_{g,ij} x^E_{g,ij}(\alpha^E_{g,ij})}{r+\delta+\bar{\lambda}_{g,i}} \\
+ \frac{(r+\delta)\left(P^*_i b - \sum_j \Delta^{-1}_{ij}\theta_{g,j}P^*_j\nu\right) + \sum_j \lambda_{g,ij}\left(\alpha^E_{g,ij}W^*_{g,j} - P^*_i d_{ij}\right)}{r+\delta+\bar{\lambda}_{g,i}}
\end{aligned}
$$

⋆ Recruiters' optimal entry

↪ Substitute for multipliers in the FOC with respect to labor market tightness, $\theta_{g,j}$.

$$
\begin{aligned}
& P_j \nu_g \sum_i \Delta^{-1}_{ij} u_{g,i} L_{g,i} \\
= \; & \frac{1-\mu(\theta_{g,j})}{\theta_{g,j}} \sum_i \frac{\lambda_{g,ij}}{r+\delta+\bar{\lambda}_{g,i}} L_{g,i} \Big(u_{g,i}\left[\alpha^E_{g,ij}W_{g,j} - P_i d_{ij} - P_i b\right] + u_{g,i}\sum_{j'}\Delta^{-1}_{ij'}\theta_{g,j'}P_{j'}\nu_g \\
& + \frac{\delta}{r+\delta}\sum_{j'} l_{g,ij'}\left[\alpha^E_{g,ij}W_{g,j} - P_i d_{ij} - \left(\alpha^E_{g,ij'}W_{g,j'} - P_i d_{ij'}\right)\right] \Big)
\end{aligned}
$$

⋆ Optimal job acceptance

↪ Substitute for multipliers in the FOC with respect to reservation productivity, $\alpha^R_{g,ij}$.

$$
\alpha^R_{g,ij}W_{g,j} - P_i d_{ij} = \frac{(r+\delta)\left(P_i b - \sum_{j'}\Delta^{-1}_{ij'}\theta_{g,j'}P_{j'}\nu\right) + \sum_{j'}\lambda_{g,ij'}\left(\alpha^E_{g,ij'}W_{g,j'} - P_i d_{ij'}\right)}{r+\delta+\bar{\lambda}_{g,i}}.
$$

□

Chapter 2

No Country for Young Men: The Inversion of the Rural-Urban Age Gradient in the United States, 1950-2019

joint with David Autor

2.1 Introduction

It is well known that residents of rural areas in the U.S. are on average older than those living in suburbs or cities Glasgow and Brown [2012], and this fact is frequently invoked to explain rural-urban differences in partisan voting patterns Scala et al. [2015]; Rodden [2019], healthcare utilization Meara et al. [2004]; Keehan et al. [2017], and economic dynamism Maestas et al. [2016]; Karahan et al. [2019]; Jones [2020]; Boehm and Siegel [2021]. Given its prominence in the data and in popular understanding, one might assume that the negative rural-urban age gradient is a long-standing and perhaps immutable feature of the economic landscape. We document that the opposite is true: the striking inverse relationship between population age and population density in contemporary America reflects a stark reversal of demographic patterns prevailing seven decades earlier. In fact, this age-density gradient dramatically inverted between 1950 and 2010, and has since stabilized. In this article, we explicate the demographic trends that gave rise to this striking inversion and discuss their implications for economics and policy going forward.

Figure 2.1.1 depicts the inversion of the rural-urban age gradient between 1950 and 2019 by plotting the mean and median age of residents of the contiguous United States by county in each decade. In this and all subsequent figures, counties are ordered by their 1950 population densities,

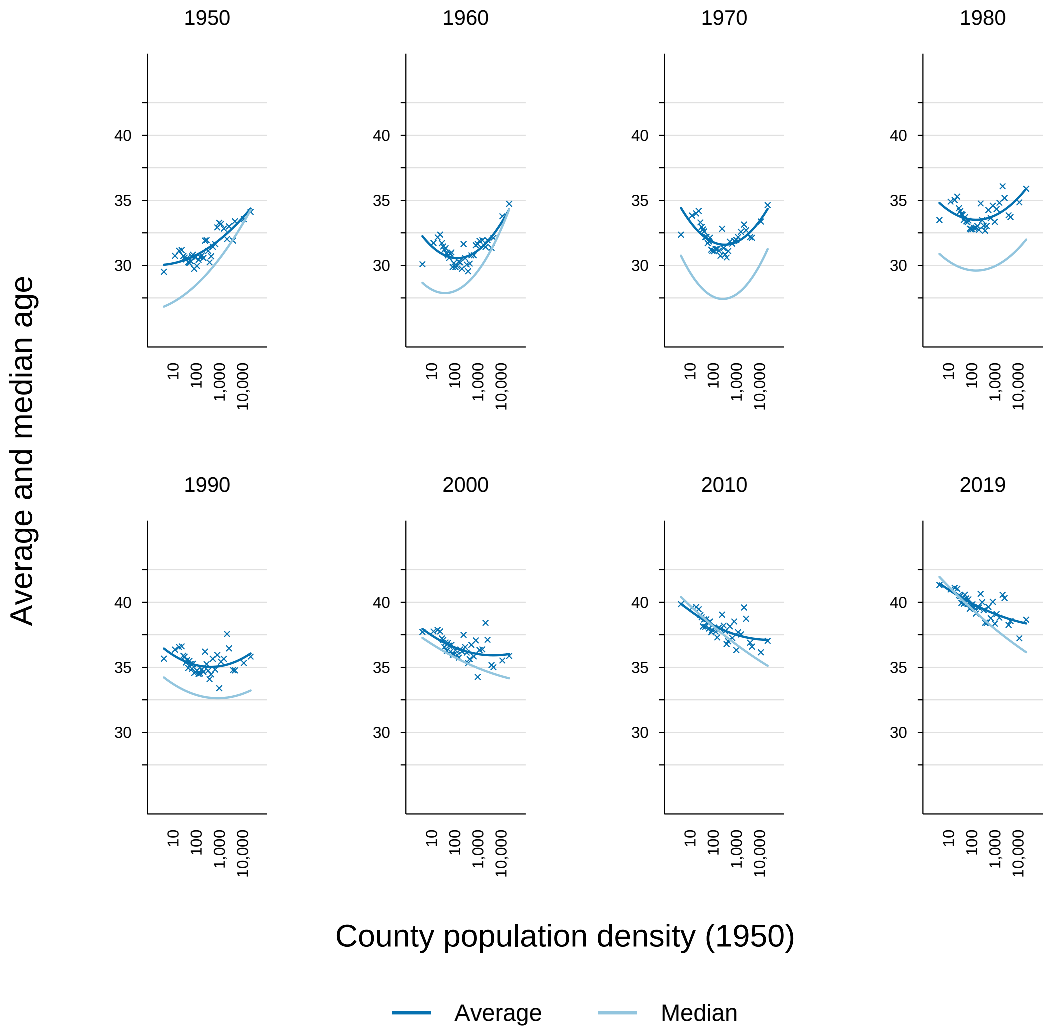


Figure 2.1.1: Mean and median ages of county residents:
Rural areas used to be relatively younger and are now relatively older

Notes. Each graph represents the mean and the median age in a given year. The mean age is represented by the markers and darker blue line, and the median age by the light blue line. Counties are ordered by their 1950 population densities and weighted by their 1950 population. Each plotted point contains approximately 2.5 percent of 1950 population.

an ordering that is highly stable across decades.[1] In 1950, residents in the densest (most urban)

[1] Each plotted point in Figure 2.1.1 corresponds to the population-weighted mean age in 40 bins, each corresponding to the set of counties ordered by 1950 population density that contain approximately 2.5 percent of 1950 population. Due to uneven county sizes, some points contain slightly more or less than 2.5 percent. The cross-county ranking of population density is highly stable across decades, exceeding 0.95 across each successive decade and 0.91 between the two nonadjacent decades of 1950 and 2019. Plotted lines correspond to a population-weighted regression of mean or median county age on an intercept and the natural logarithm of population density and its square.

counties were on average four-and-a-half years older than those in the least dense (most rural) of counties. Within two decades, this upward sloping relationship was supplanted by a distinct U-shape in which the least and most dense counties possessed among the oldest populations. Two decades further on, in 1990, the age-density gradient was almost entirely flat. And by 2010, it had inverted, so that urban areas were distinctly younger than rural areas. This gradient then remained stable over the ensuing decade through 2019 (when our data end). Whereas in 1950, residents in the least dense counties were on average 4.5 years younger than their counterparts in the most dense counties (with both sets comprising 2.5 percent of 1950 population), by 2019 residents of the most rural counties were 2.7 years older than those in the most urban counties, a swing of 7.2 years. This urban-rural age inversion is even starker when measured using median rather than mean population age, as also seen in Figure 2.1.1 .[2]

As an accounting matter, the reversal of the age-density gradient could stem from three proximate causes: a differential fall in natality in rural versus urban areas, which would reduce the stock of young rural relative to urban residents[3]; a differential fall in mortality in rural versus urban areas, which would increase the stock of older rural relative to urban residents[4]; and differential out-migration of younger residents or in-migration of older residents that shift the balance of young and old across rural and urban areas. By harmonizing historical data on the exact age by race by sex distribution of all U.S. counties for the period 1950 and 2019 and applying multiple sources of reported data at various levels of aggregation to calculate age-specific natality and mortality by place, race, and sex, we explore the role of each of these channels.

We have three principle findings: Declining (relative) age-adjusted rural natality contributes modestly to the reversal of the age-density gradient; declining (relative) age-adjusted urban mortality modestly works (modestly) in the opposite direction; while sharp temporal changes in age-specific migration rates are the predominant contributor. This latter force is in turn driven by the behavior of prime-age adults ages 25–54, who are consequential to changing age structure for two distinct reasons: first, they have exhibited distinctly different choices about where to spend their prime working years in different eras—in cities versus suburbs versus non-metropolitan areas; second, their mobility patterns largely determine (and almost perfectly predict) the mobility of children and youth ages 0–17, who in many cases are their dependents. Logically, the geographic distribution of children has substantial leverage on mean population age across locations since children are almost necessarily far younger than average. By contrast, the mobility of most other age groups has followed a relatively stable life-cycle pattern during these seven decades: young adults ages 18–24 have flowed strongly towards cities; while older working-age adults 55–64, and retirement-age adults ages 65+, have flowed in the opposite direction. Thus, fluctuations in the migration patterns of

[2] In 1950, the median age in the least dense and most counties that each comprised 2.5 percent of population was 27.0 years and 33.8 years, respectively. In 2019, these values were 41.9 and 36.5 years, corresponding to a swing in the median age gap of 12.2 years.

[3] Natality refers to live births per population and is distinct from fertility, which refers to conceptions.

[4] Assuming that mortality is concentrated among those of above-average age, it reduces mean population age.

prime-age adults and their children drive the reversal of the age-density gradient between 1950 and 2010 seen in Figure 2.1.1.

Although the inversion of the rural-urban age gradient took six decades to unfold, it does not reflect a smooth, continuous reallocation of prime-age adults and dependent-age children from rural to urban areas. Rather, the data reveal three distinct eras of migration. The decades of 1950 through 1970 witnessed rapid suburbanization, in which prime-age adults, along with dependent-age children 0–17, flowed from both urban and non-metropolitan areas towards mid-density locations (i.e., suburbs). Subsequently, prime age adults and dependent-age children flowed out of urban areas and uniformly downward along the urban-rural gradient during a period of urban decline between 1970 and 1990. Most recently, from 1990 to the present—a period that some scholars have dubbed an 'American urban revival' Couture and Handbury [2020]—prime-age adults and their dependent-age children urbanized. Counterintuitively, the falling (relative) age of urban areas since 1990 was not driven by movement of prime-age adults and dependent children *into* these locations; rather, it is driven by a fall in their *exit rate*, a point to which we return below.

To our knowledge, the stark reversal of the population age-density gradient in the United States is previously unremarked in scholarly literature and its causes are unstudied. Yet, the relationship between population density and population age—and the factors shaping this relationship—is critical for multiple areas of social science, public health, and public policy. In economics, it is widely hypothesized that younger populations generate dynamism in the local labor markets in which they reside, as measured by business formation, output of original ideas, or even (more mundanely) employment rates Shimer [2001]; Maestas et al. [2016]; Karahan et al. [2019]; Jones [2020]; Boehm and Siegel [2021]. In political science, the over-representation of older Americans in rural areas is understood to interact with the structure of the U.S. electoral system to reduce the power of cities in electoral politics Scala et al. [2015]; Rodden [2019]. In public health, age is the largest driver of healthcare expenditure Meara et al. [2004]; Keehan et al. [2017], and the distribution of the elderly across locations determines the citing of medical facilities and the provision of care services. In the realm of policy, the demographics of cities and towns affect their needs for public services (e.g., schools, recreational facilities, public transportation) as well as the capacity of governments to raise revenue to fund these services Murdock et al. [2015]; Butler and Yi [2019] since retirees primarily draw down savings rather than generating new income.

As fundamentally, the changing shape of the age-density gradient illuminates the economic forces shaping the character of places and the set of residents attracted to them. Indeed, one widely noted puzzle that our findings may help to resolve is that even as U.S. urban wage differentials have increased in the past three decades, the geographic mobility of U.S. workers has fallen Molloy et al. [2011, 2016]. This evidence is often taken to indicate that workers are failing to take advantage of earnings gains available to movers Ganong and Shoag [2017]; Kaplan and Schulhofer-Wohl [2017]; Hoxie et al. [2019]; Hsieh and Moretti [2019]. Our findings suggests an alternative interpretation: Falling mobility may itself be a response to prime-age adults and their dependent children *not* leaving cities in mid-adulthood as they have done in previous decades. Specifically, in the decades prior to the 2000s, young adults age 18–24 flowed rapidly towards cities while prime-age adults ages 25–54 flowed outward. Since 1990, however, the outward flow of prime-age adults has largely halted while urban inflows of young adults have continued unabated. The net effect is that younger adults and dependent-age children are increasingly represented in urban areas. Thus, paradoxically, the *falling* net mobility of prime-age adults (and dependent-age children) may represent an improvement in the quality of urban life that has slowed out-migration of prime-age adults and their families.[5]

2.2 Trends in Natality, Mortality, and Migration across Time and Space

Before decomposing changes in mean age across space, we visually summarize the main dynamic forces determining changes in the age structure, beginning with natality. Natality has fallen steeply since 1950, as seen in Figure 2.2.1, which reports a bin-scatter of annual births per thousand women by county among woman ages 15–44 at decadal frequencies between 1950 and 2018.[6] The secular decline in natality has not been uniform across geographies. In 1950, age-adjusted natality was more than 50 percent higher in rural than urban counties, ranging from 133 births per 1,000 women in the lowest density counties to 87 births per women in the most urban counties.[7] Natality declined near monotonically over the next seven decades, but the absolute and proportionate falls were far steeper in rural areas. By 2018, age-adjusted natality had fallen by 55 percent in rural counties (from 133 to 73 births per thousand women) and by only 38 percent in urban counties (from 87 to 54 births per thousand women). Changes in population age structure make only a modest contribution to this fall (Figure 2.2.1), but observed (raw) natality has fallen by more than age-adjusted natality in non-urban counties due to rising relative ages in these locations (Figure 2.1.1).

While the changing geography of natality contributes to differential aging in rural areas, trends in mortality work in precisely the opposite direction. As shown in Figure 2.2.2, age-adjusted mortality was approximately 20 percent higher in urban than rural counties in 1950 (10.7 versus 9.0 deaths per thousand residents), but this differential declined and then reversed sign over the next seven decades.[8] As of 2018, age-adjusted mortality was 20 percent *lower* in urban than rural counties (3.6 versus 4.5 deaths per thousand residents). Since mortality is concentrated among the elderly, the reversal of the density-mortality gradient contributes to a relative *rise* in age in urban counties,

[5]Papers by Edlund et al. [2015]; Diamond [2016]; Ellen et al. [2019]; Su [2019]; Baum-Snow and Hartley [2020] emphasize the rising attractiveness of cities for college-educated adults between 1980 and 2010 stemming from improvements in urban amenities, rising incomes, and a concomitant increase in the opportunity cost of commuting. Moretti [2013] offers a contrasting view.

[6]Each panel corresponds to natality in the indicated year. We do not report multi-period averages since these obscure the temporal fluctuations.

[7]Age-adjusted natality is calculated by assigning each county the 1950 U.S. age distribution of women ages 15-44 in all periods

[8]Mortality statistics are age-adjusted by assigning each county the 1950 U.S. aggregate age, sex, and race distribution in all periods

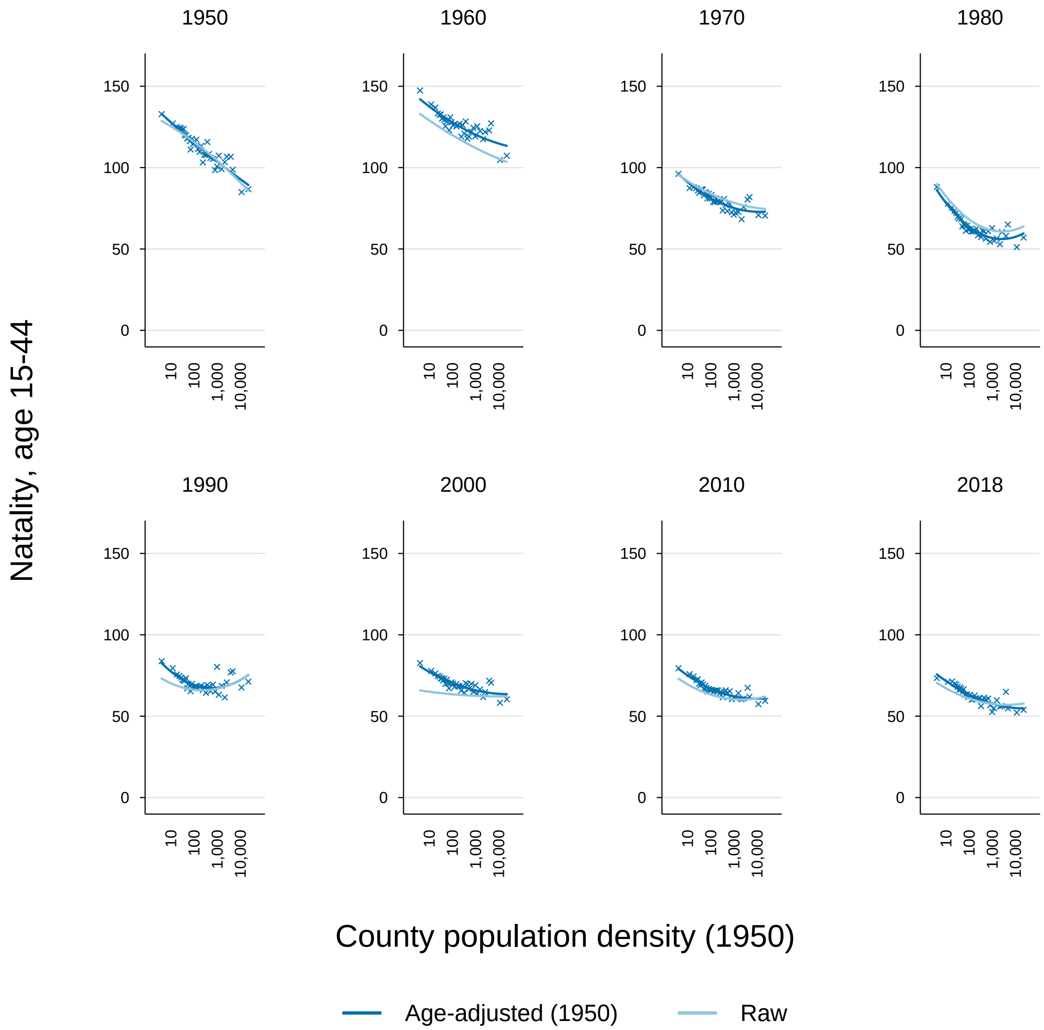


Figure 2.2.1: Natality fell relatively and absolutely more in rural counties

Notes. Each graph represents the number of births per 1,000 women ages 15-44 in a given year. Data represented by the markers and darker blue line are age-adjusted using 1950 women age distribution. Data represented by the light blue line are not age-adjusted and corresponds to the current age distribution. Counties are ordered by their 1950 population densities and weighted by their 1950 population. Each plotted point contains approximately 2.5 percent of 1950 population.

opposite to the case for fertility. As Figure 2.2.2 also underscores, age-adjusted mortality trends diverge substantially from observed (raw) mortality trends over these seven decades, both overall and across geographies: the aggregate aging of the population, which drives mortality rates upward, masks a very substantial fall in age-constant mortality. Simultaneously, the observed (raw) change in mortality in urban versus rural areas *exceeds* the age-adjusted change because it does not account for the fall in the relative age of urban residents.

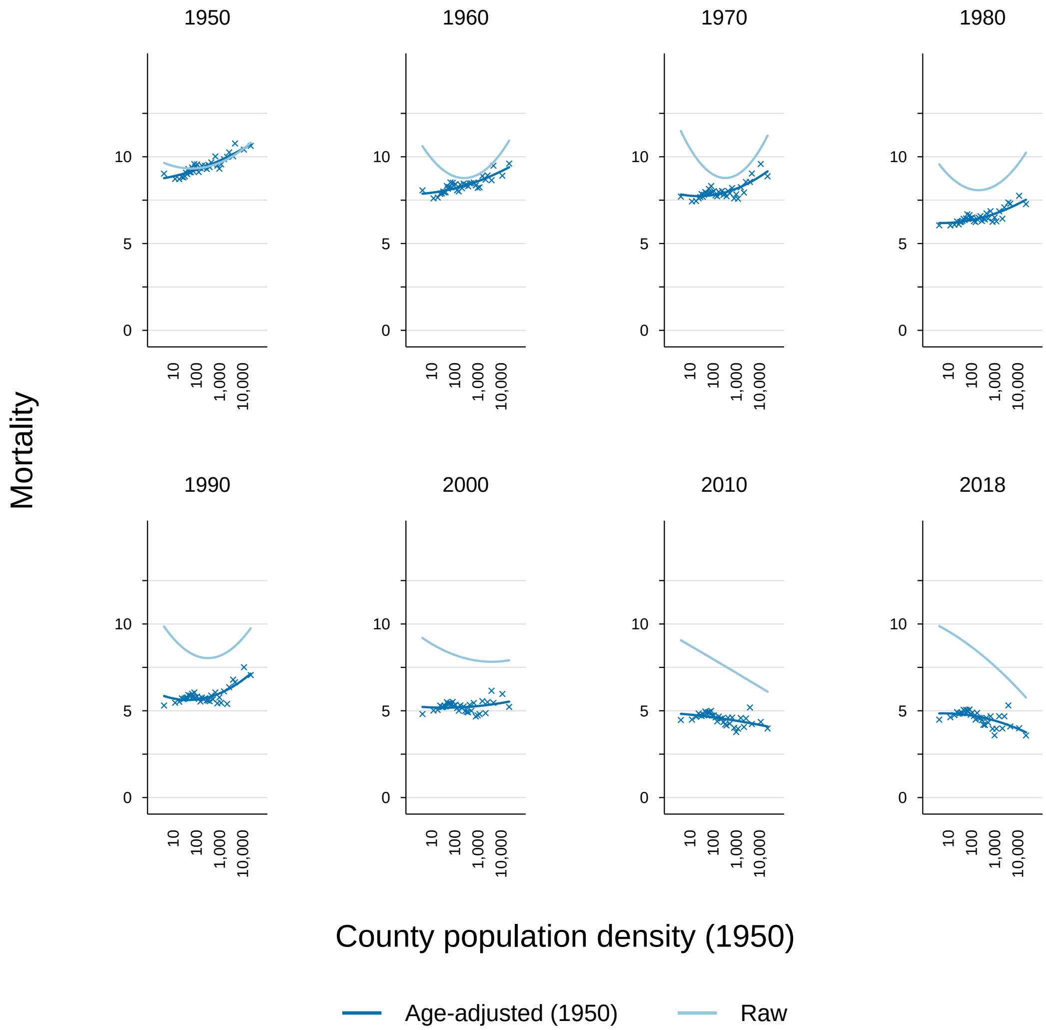


Figure 2.2.2: Raw and age-adjusted mortality fell differentially in urban counties

Notes. Each graph represents the number of deaths per 1,000 residents in a given year. Data represented by the markers and darker blue line are age-adjusted using 1950 age distribution. Data represented by the light blue line are not age-adjusted and corresponds to the current age distribution. Counties are ordered by their 1950 population densities and weighted by their 1950 population. Each plotted point contains approximately 2.5 percent of 1950 population.

Figure 2.2.3 plots the third component of our decomposition, net country migration rates, defined as net migration inflows or outflows of population per thousand county residents after accounting for natality and mortality.[9] We again plot raw and age-adjusted net migration rates, where the

[9]Gross population flows across U.S. counties by age, sex, and race are not available on a consistent basis for most of this time interval. We calculate *net* migration as the residual change in the count of county residents in each sex, age, and race category (accounting for aging) net of births and deaths by age, sex, and race, which are known.

latter assigns each county the 1950 U.S. aggregate age, sex, and race distribution. The geography of migration flows differs sharply across decades. During the 1950s, net population flowed out of rural and urban areas towards mid-density locations (i.e., suburbs). During the 1970s, migration flowed from high to low-density locations, a trend that slackened but did not change sign during the 1990s. In the 2010s, net migration flows trended in the reverse direction—from low to high-density counties (though this trend was quite shallow). In 2018, the most recent year available, population again flowed out of urban areas in net, though the outflow rate was modest compared to earlier decades. Thus, consistent with Molloy et al. [2011, 2016], we see a step fall in net cross-county migration in recent decades.

Comparing the decadal trends in migration in Figure 2.2.3 with the changing rural-urban age gradient in Figure 2.1.1 hints at a critical regularity: net migration flows predict relative changes in population age. When net migration flows from rural and urban areas towards suburban areas during the 1950s and 1960s, suburban areas become relatively younger in the ensuing decades; when net migration from urban areas slows in the 1990s, 2000s, and 2010s, urban areas become relatively younger. This correlation between migration inflows and relative age reductions will tend to arise if net migration is concentrated among the young (assuming these impacts are not offset by countervailing trends in fertility and mortality). Figure 2.2.4 provides additional detail on these trends by plotting net migration flows by decade (1950, 1970, 1990, and 2010) for each of six age brackets: 0–17, 18–24, 25–39, 40–54, 55–65, and 65+. In all decades: (1) young adults ages 18–24 tend to flow from low to high-density counties; and (2) older adults ages 55–64 and 65+ tend to flow from higher to lower-density counties. Both patterns likely reflect life-cycle migration motives, the former due to moves for college-attendance and job-seeking, and the latter due (in part) to moves for retirement.

In contrast to these regularities, Figure 2.2.4 reveals two sharp, parallel shifts in migration patterns, one concentrated among prime-age adult, ages 25–39 and 40–54, and the other among minor children ages 0–17. Migration flows among these groups are inverse U-shaped in the 1950s—indicating suburbanization of both urban and rural residents; downward sloping and in the 1970s, indicating de-urbanization; and increasingly flat in the 1990s and 2010s, meaning little net migration. Logically, migration patterns of children follow those of prime-age adults since, presumably, the prime-age movers are their parents. This pattern suggests that changes in migration patterns of prime-age adults may be key to understanding the inversion of the urban-rural age-density gradient. We formally confirm this next.

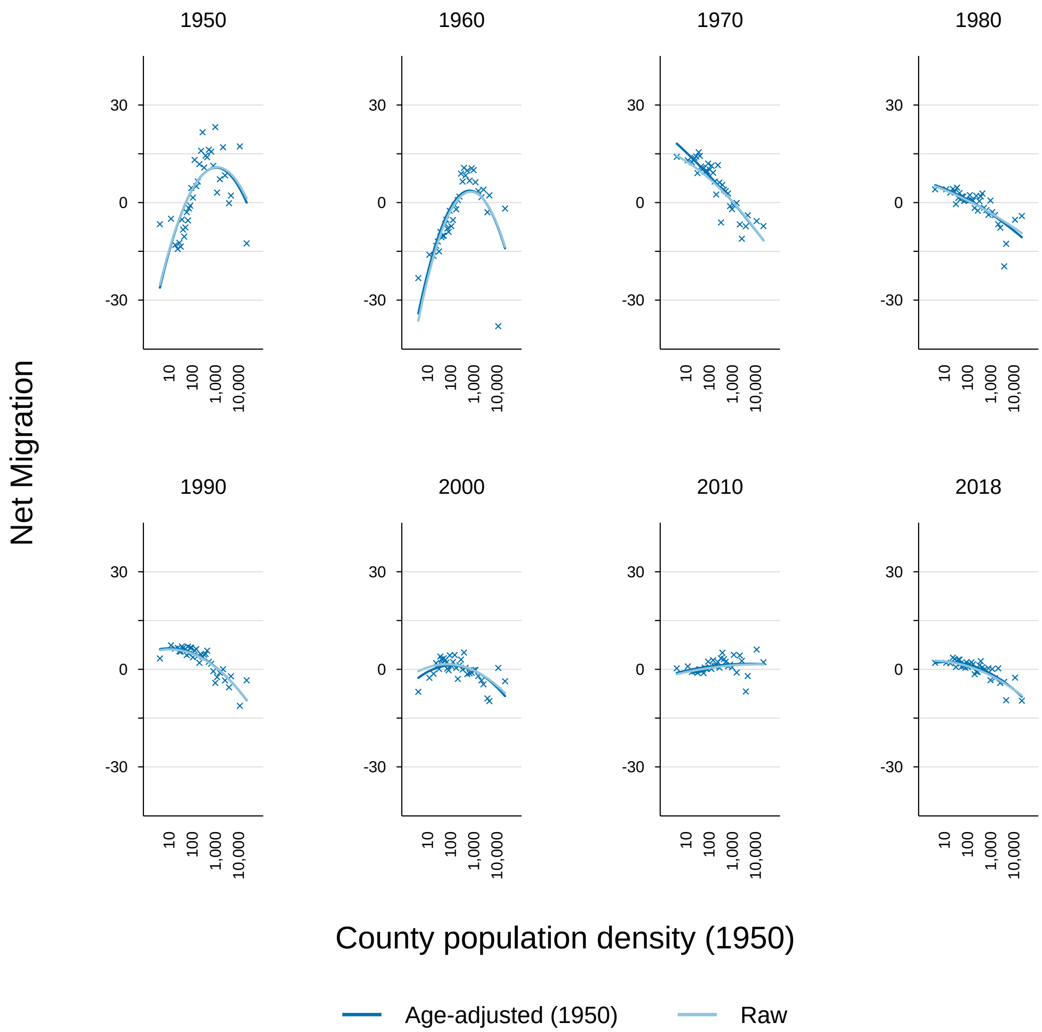


Figure 2.2.3: Migration flows differ substantially by decade

Notes. Each graph represents the population gains/losses per 1,000 residents net of natality and mortality per 1,000 residents in a given year. Data represented by the markers and darker blue line are age-adjusted using 1950 age distribution. Data represented by the light blue line are not age-adjusted and corresponds to the current age distribution. Counties are ordered by their 1950 population densities and weighted by their 1950 population. Each plotted point contains approximately 2.5 percent of 1950 population.

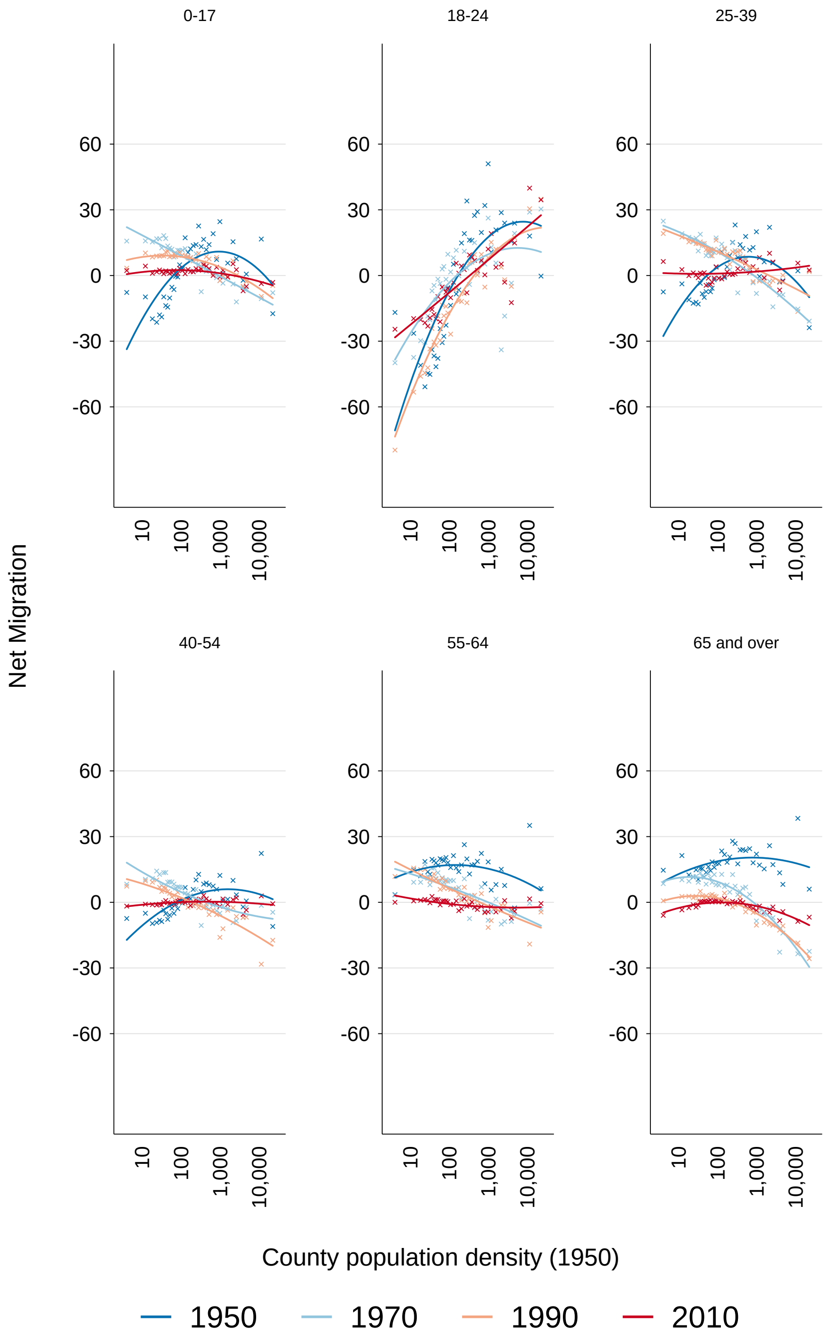


Figure 2.2.4: Migration over the life cycle across decades

Notes. Each graph represents the population gains/losses per 1,000 residents net of natality and mortality per 1,000 residents in a given age bracket for various decades. Counties are ordered by their 1950 population densities and weighted by their 1950 population. Each plotted point contains approximately 2.5 percent of 1950 population.

2.3 Decomposing Population Aging

Having described the trends in natality, mortality and migration across U.S. counties, we now formalize a method to assess the contribution of each of these forces to differential aging.

2.3.1 The Key Role of Youngsters

The contributions of each age brackets to population aging can be calculated using the decomposition derived by Preston et al. [1989]. We present this decomposition below and then apply it to seven decades of U.S. county-level population data.

Let $n(a,t)$ denote the total population of age a at time t and $p(a,t) = \frac{n(a,t)}{\int_0^\infty n(\tilde{a},t)d\tilde{a}}$ be the share of the population of age a. The mean age in the population defined as:

$$A(t) \equiv \frac{\int_0^\infty n(a,t)ada}{\int_0^\infty n(a,t)da} = \int_0^\infty p(a,t)ada. \tag{2.3.1}$$

To measure the contribution of each age group to the change in mean age over time, we build on this simple expression of the change in population mean age:

$$\frac{dA}{dt}(t) = \int_0^\infty p(a,t)r(a,t)[a - A(t)]da, \tag{2.3.2}$$

with $r(a,t)$ the growth rate of population of age a.

Equation (2.3.2) states that each age group's contribution to overall population aging is proportional to its leverage, $a - A(t)$, its growth rate, $r(a,t)$, and its initial share in the population, $p(a,t)$. Intuitively, population age rises if groups above average age $A(t)$ grow relatively faster than those below average age and declines if groups below average age grow relatively faster than those above average age.

To implement the population age decomposition in Figure 2.3.1, we group counties into twenty vigintiles (depicted as bars) ordered by 1950 county population density, with each vigintile containing approximately five percent of the 1950 population.[10] We remove the overall change in mean population age in each figure, so that each bar represents the contribution by age group to the change in average age within the vigintile *relative* to the aggregate change.

Figure 2.3.1 reveals the central role of age groups with the most leverage, and in particular children ages 0-17, in relative aging across counties. Changes in the share of youngsters in the population appear to explain most of the relative rejuvenation of suburbs in the 1950S and 1960s, as well as the aging of cities in the 1970s. In more recent decades, older adults ages 55–64 and 65+ counteracted the impact of kids, leading to lower variation in aging across county densities.

[10] Some vigintiles contain slightly more or slightly less than five percent of population since counties are of uneven sizes. Because approximately 10 percent of U.S. population in 1950 resided in New York county, which contains Manhattan, this county occupies the upper two vigintiles.

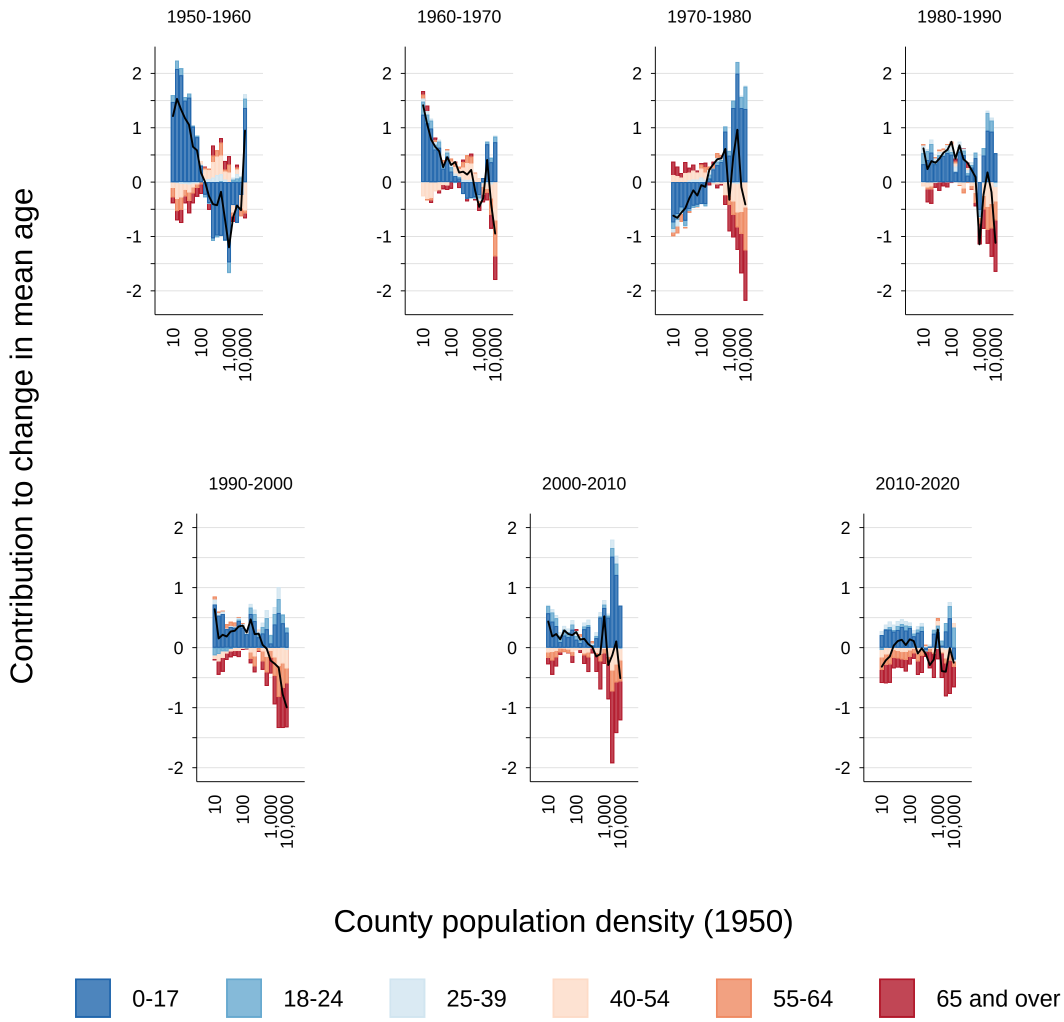


Figure 2.3.1: Decadal contributions of each age bracket to changes in mean county age

Notes. Each graph represents the decadal contributions of each age bracket to changes in mean county age to aging applying formula (2.3.2). Counties are ordered by their 1950 population densities and weighted by their 1950 population. Each plotted bar contains approximately 2.5 percent of 1950 population.

2.3.2 Assessing the Contributions of Natality, Mortality and Migration

As an accounting matter, changes in the age structure of the population are fully determined by initial age structure and three dynamic forces: natality, mortality, and net migration. We build on expression (2.3.2) to recover the contributions of each of those components.

The age schedule of population growth rate, $r(a,t)$, is fully determined by natality, mortality, and net migration, along with the age distribution at time t. Specifically, $r(a,t)$ is equal to natural aging, captured by minus the semi-elasticity of population with respect to age, $-\frac{1}{n(a,t)}\frac{\partial n(a,t)}{\partial a}$ if

$a > 0$, minus the mortality rate at age a, plus the net migration rate. If $a = 0$, $r(a,t)$ is simply equal to the ratio of births to newborns, plus natural aging set to -1 by convention.

Integrating over age groups gives a decomposition of aging into three components, corresponding to the population averages of natality, mortality and migration contributions. Indeed, the effect of natural aging on mean age $A(t)$ is mechanically equal to 1: with no births, no mortality and no migration, the population (and each of its members) ages by 1 over one year. It follows that aging is fully determined by natality, mortality and migration.

Formally, the change in mean age can be expressed as:

$$\frac{dA}{dt}(t) = 1 + \partial_b A - \partial_d A + \partial_m A \tag{2.3.3}$$

where we define:

$$\partial_b A = -\int_0^\infty p(a,t)\beta(a,t)A(t)da, \tag{2.3.4}$$

$$\partial_d A = \int_0^\infty p(a,t)\delta(a,t)(a - A(t))da, \tag{2.3.5}$$

and:

$$\partial_m A = \int_0^\infty p(a,t)\mu(a,t)(a - A(t))da, \tag{2.3.6}$$

with $\beta(a,t) = \frac{1}{n(a,t)}\frac{\partial b}{\partial t}(a,t)$, $\delta(a,t) = \frac{1}{n(a,t)}\frac{\partial d}{\partial t}(a,t)$ and $\mu(a,t) = \frac{1}{n(a,t)}\frac{\partial m}{\partial t}(a,t)$ the respective natality rate, mortality rate and migration rate for age a.

We go one step further to disentangle the effects of differentials in natality, mortality and migration rates across counties from the mechanical effect of the initial age distribution. In fact, decomposition (2.3.3) confounds the impact of each component through differences in the rates, $\beta(a,t)$, $\delta(a,t)$ and $\mu(a,t)$, with composition effects due to variations in $p(a,t)$. To distinguish between the two, we define a fourth component, named initial age distribution, which corresponds to the predicted change in mean age given initial county population distribution when applying the U.S. natality, mortality and migration rates, $\beta^{US}(a,t)$, $\delta^{US}(a,t)$ and $\mu^{US}(a,t)$. We define the initial distribution component as:

$$\partial_p A = \int_0^\infty p(a,t)\Big[-\beta^{US}(a,t)A(t) + \big(-\delta^{US}(a,t) + \mu^{US}(a,t)\big)(a - A(t))\Big]da, \tag{2.3.7}$$

and the residual natality, mortality and migration components as:

$$\partial_{b-b^{US}} A = -\int_0^\infty p(a,t)\big(\beta(a,t) - \beta^{US}(a,t)\big)A(t)da, \tag{2.3.8}$$

$$\partial_{d-d^{US}} A = \int_0^\infty p(a,t)\big(\delta(a,t) - \delta^{US}(a,t)\big)(a - A(t))da, \tag{2.3.9}$$

and:

$$\partial_{m-m^{US}} A = \int_0^\infty p(a,t)\big(\mu(a,t) - \mu^{US}(a,t)\big)(a - A(t))da. \tag{2.3.10}$$

Those four components sum up to the mean age change:

$$\frac{dA}{dt}(t) = 1 + \partial_p A + \partial_{b-b^{US}} A - \partial_{d-d^{US}} A + \partial_{m-m^{US}} A. \tag{2.3.11}$$

2.3.3 Decomposition Results

To implement the population age decomposition in Figure 2.3.2, we group counties into twenty vigintiles (depicted as bars) ordered by 1950 county population density, with each vigintile containing approximately five percent of the 1950 population.[11] We remove the overall change in mean population age in each figure, so that each bar represents the contribution of natality, mortality, and net migration to the change in average age within the vigintile *relative* to the aggregate change. Each vigintile in turn contains four distinct stacked bars representing the separate contributions of natality, mortality, net migration and initial age distribution. Bars with negative heights represent factors that reduce county relative age. Those with positive heights correspond to factors that increase country relative age. The net effect of these three factors in each vigintile, equal to the sum of the three bars, is marked with a black line.

Migration plays a critical role in the (relative) aging of rural areas throughout these seven decades. In each ten year interval between 1950 and 2019, migration adds between 0.25 and 1.25 years to the relative age of the most rural counties—meaning either that younger people moved out or older people moved in. (We explore this next.) Surprisingly, a second factor that contributes modestly but persistently to rural aging in all decades except for 1970–1980 is natality. Although natality among women ages 15-44 is consistently higher in rural counties (Figure 2.2.1), there are relatively few women in these age brackets in rural areas, and moreover, their relative scarcity rises with across decades. Finally, mortality reduces relative age in rural areas in all periods. In early decades, age-adjusted mortality is somewhat *lower* in rural areas, but older adults are sufficiently over-represented that aggregate rural mortality is still relatively high. Over subsequent decades, age-adjusted mortality falls in all locations, but it falls by less in rural areas. In combination with the aging of rural locations, mortality plays an increasingly substantial role in retarding the aging of rural locations, reducing relative age by 0.19 years in the 1950s, by 0.45 years in the 1970s, by 0.62 years in the 1990s, and by 0.53 years between 2010 and 2019. Thus, the aging of rural areas should be understood as the net effect of three persistent forces: (1) net out-migration of relatively young residents; (2) falling natality among women ages 15-44, combined with a shrinking share of women in this age bracket; and (3) comparatively elevated rural mortality, which slows the rate of population aging through attrition of the old.

To synthesize the information in Figure 2.3.2, it is useful to consider the following decomposition

[11] Some vigintiles contain slightly more or slightly less than five percent of population since counties are of uneven sizes. Because approximately 10 percent of U.S. population in 1950 resided in New York county, which contains Manhattan, this county occupies the upper two vigintiles.

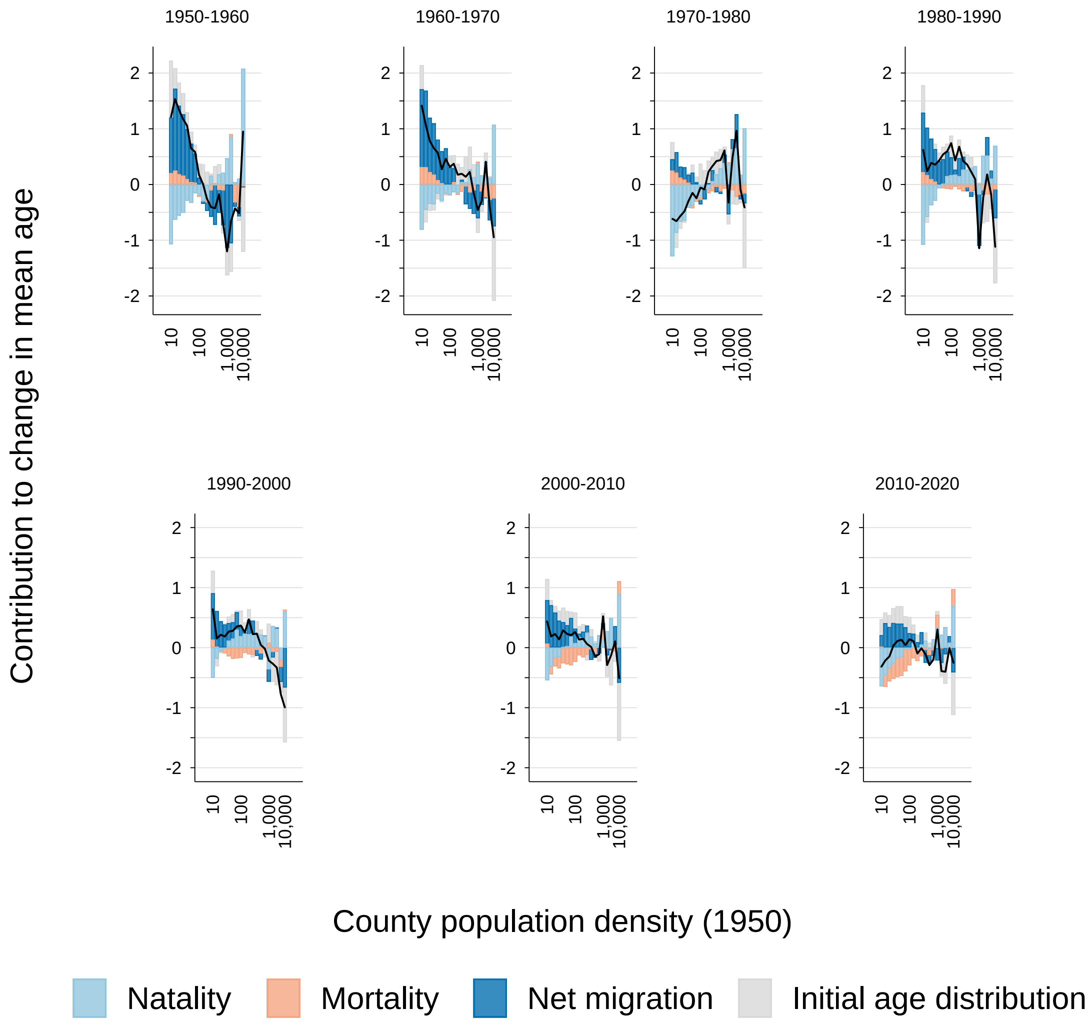


Figure 2.3.2: Decadal contributions of natality, mortality, migration, and initial age distribution to changes in mean county age

Notes. Each graph represents the decadal contributions of natality, mortality, migration, and initial age distribution to changes in mean county age to aging applying formula (2.3.11). Counties are ordered by their 1950 population densities and weighted by their 1950 population. Each plotted bar contains approximately 2.5 percent of 1950 population.

of variance in mean age change:

$$\mathbb{V}\text{ar}\left[\frac{dA}{dt}\right] = \mathbb{C}\text{ov}\left[\frac{dA}{dt}, \partial_{b-b^{US}}A\right] + \mathbb{C}\text{ov}\left[\frac{dA}{dt}, -\partial_{d-d^{US}}A\right] + \mathbb{C}\text{ov}\left[\frac{dA}{dt}, \partial_{m-m^{US}}A\right] + \mathbb{C}\text{ov}\left[\frac{dA}{dt}, \partial_{p}A\right]. \tag{2.3.12}$$

Changes in mean age and natality, mortality, migration and initial age distribution components are first averaged within each of the forty vigintiles of 1950 population density before applying formula (2.3.12). Figure 2.3.3 illustrates the recent slackening of relative changes in mean age across counties. It also confirms the key role of migration, especially in the 1950s and 1960s, while underlying the importance of natality in the 1970s.

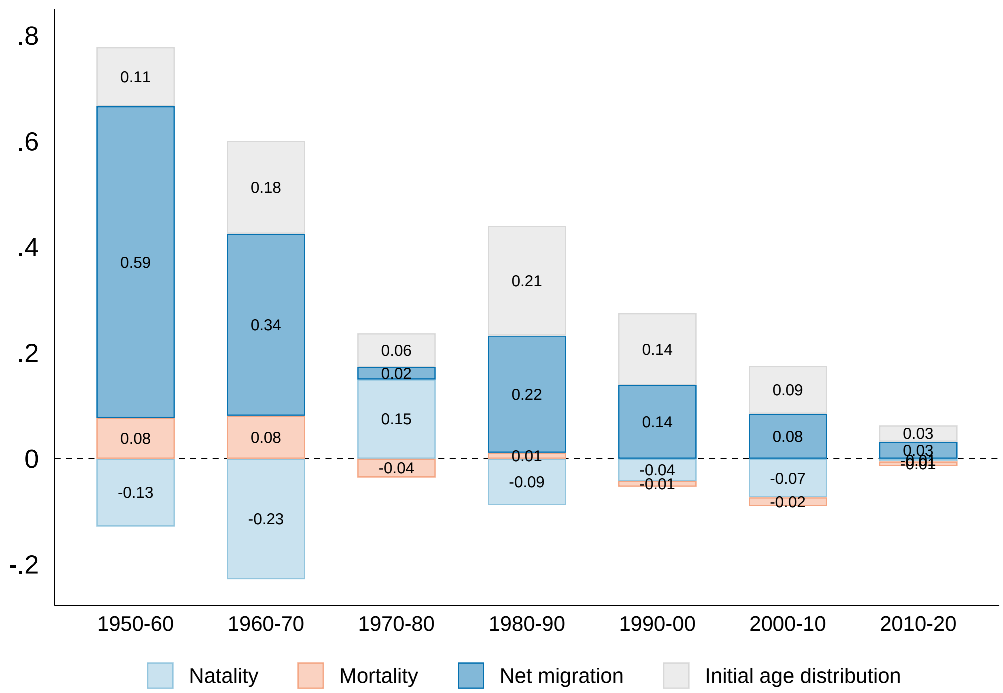


Figure 2.3.3: Contributions of natality, mortality, migration and initial age distribution to changes in the rural-urban age gradient by decade

Notes. Each bar represents the contributions of natality, mortality, migration and initial age distribution to changes in the rural-urban age gradient by decade applying formula (2.3.12). Counties are ordered by their 1950 population densities and weighted by their 1950 population. Changes in mean age and natality, mortality, migration and initial age distribution components are first averaged within forty vigintiles, which each contains approximately 2.5 percent of 1950 population, before applying formula (2.3.12).

Having explored the magnitude of each primary force to relative aging across space, we now zoom in to unbundle the role of natality, mortality and migration by age. Figures 2.3.4, 2.3.5, and 2.3.6 represent their decadal contributions broken down by age. Figure 2.3.4 shows that high fertility of young women ages 15-24 is a lasting factor limiting the relative aging of rural areas. Figure 2.A.2 corroborates this finding: While fertility of young women ages 15-24 decreased uniformly over decades, it has invariably been higher in rural areas, with a differential of about 50 births per 1,000 women. Figure 2.3.5 supports the limited role of mortality. Finally, figure 2.3.6 demonstrates the key role of groups with highest leverage: children ages 0-17 and older people above 55.

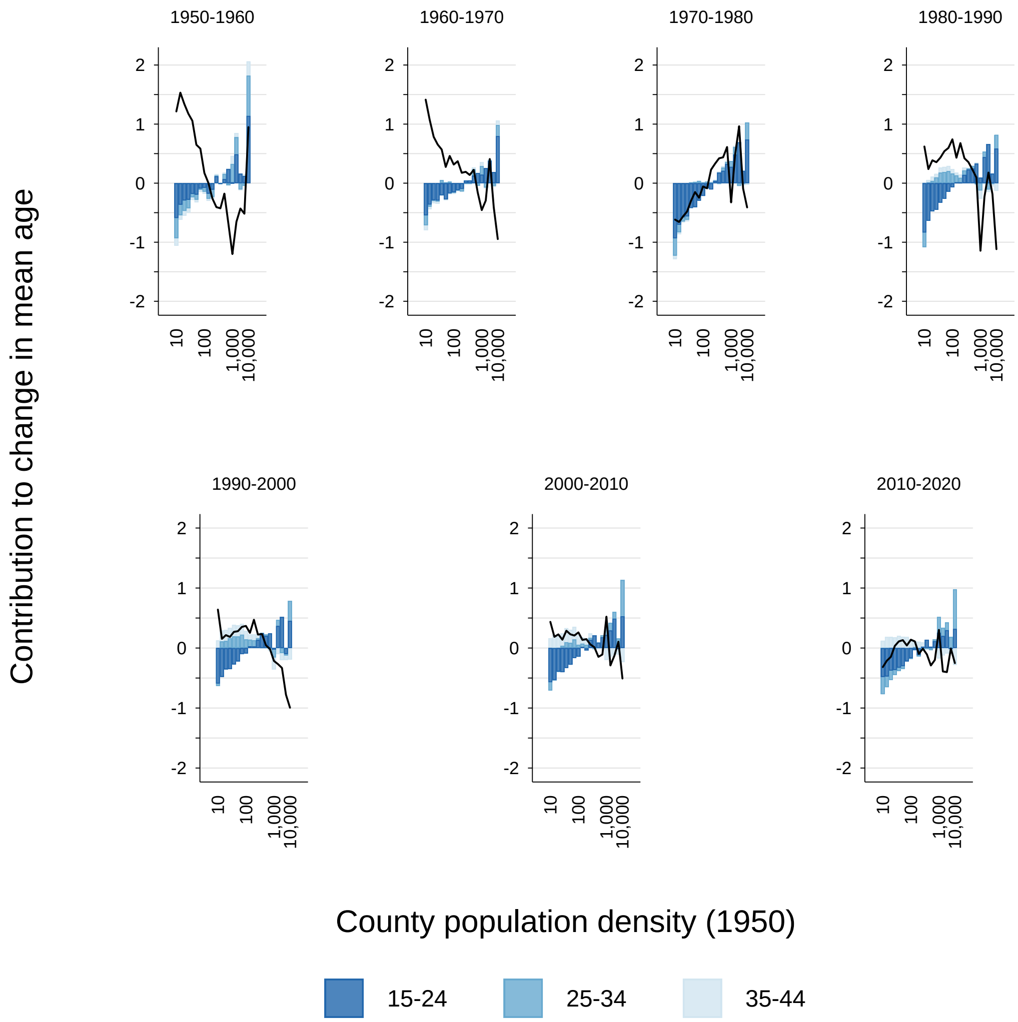


Figure 2.3.4: Decadal contribution of natality by mothers' age to changes in mean county age

Notes. Each graph represents the decadal contributions of natality by mothers' age bracket to changes in mean county age to aging applying formula (2.3.8). Counties are ordered by their 1950 population densities and weighted by their 1950 population. Each plotted bar contains approximately 2.5 percent of 1950 population.

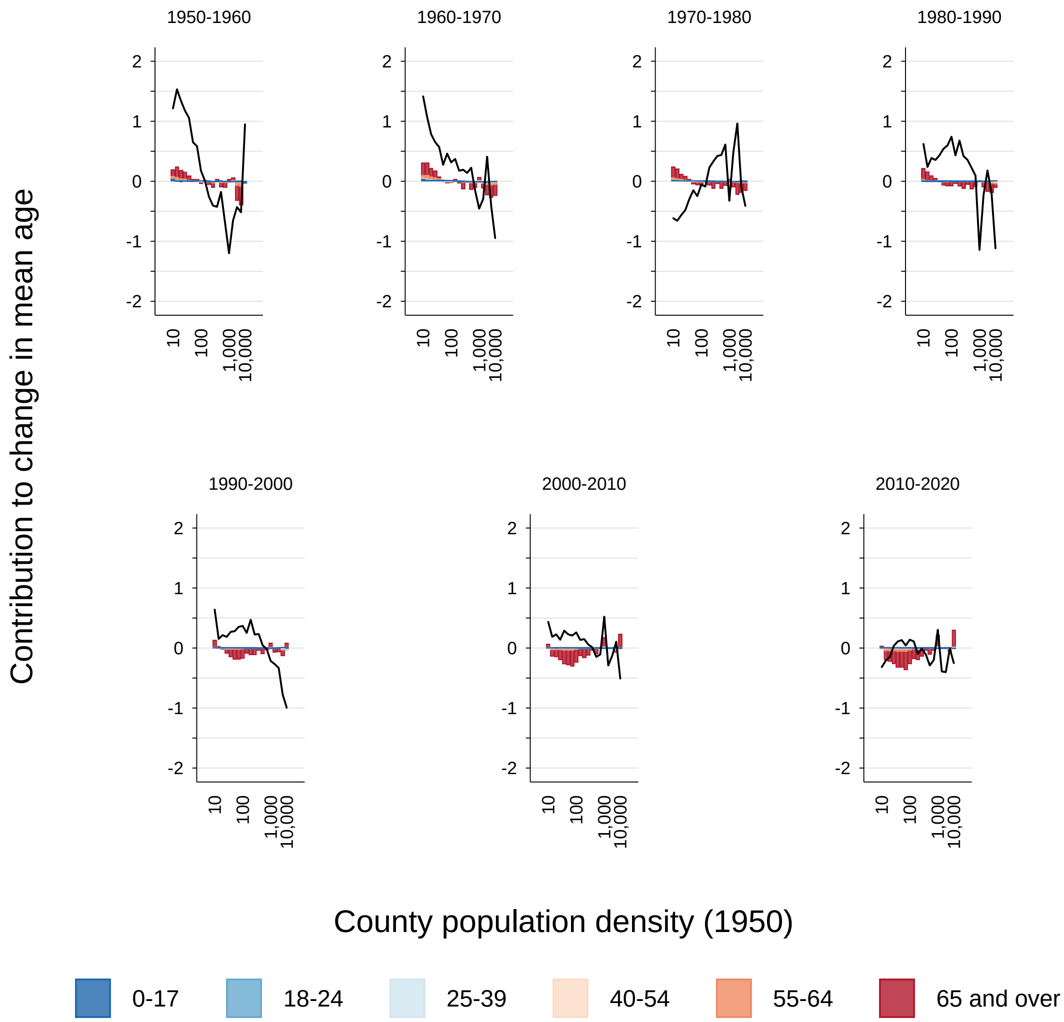


Figure 2.3.5: Decadal contribution of mortality by age to changes in mean county age

Notes. Each graph represents the decadal contributions of mortality by age bracket to changes in mean county age to aging applying formula (2.3.9). Counties are ordered by their 1950 population densities and weighted by their 1950 population. Each plotted bar contains approximately 2.5 percent of 1950 population.

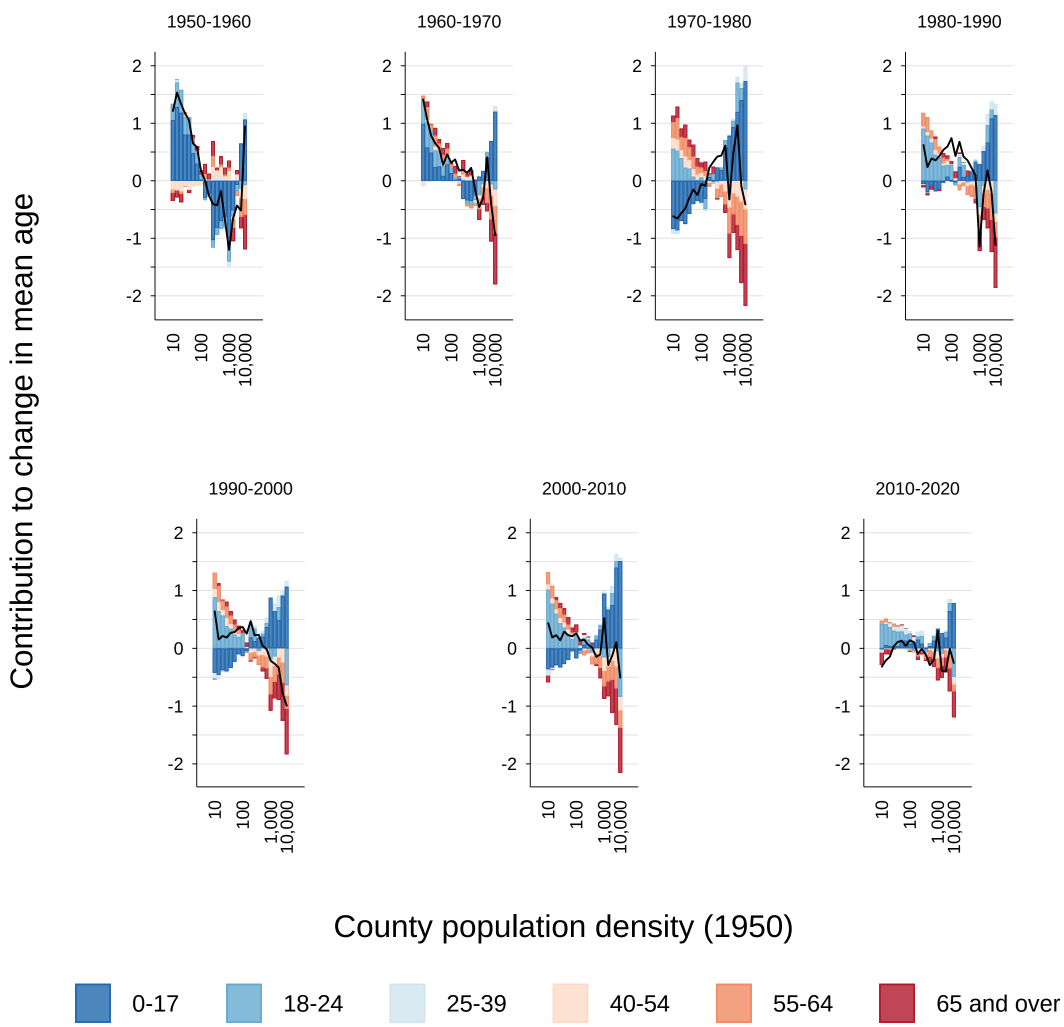


Figure 2.3.6: Decadal contribution of migration by age to changes in mean county age

Notes. Each graph represents the decadal contributions of net migration by age bracket to changes in mean county age to aging applying formula (2.3.10). Counties are ordered by their 1950 population densities and weighted by their 1950 population. Each plotted bar contains approximately 2.5 percent of 1950 population.

2.4 Conclusion

In this paper, we have documented the inversion of the rural-urban age gradient between 1950 and 2019. Whereas in 1950, residents in the least dense counties were on average 4.5 years younger than their counterparts in the most dense counties, by 2019 residents of the most rural counties were 2.7 years older than those in the most urban counties, a swing of 7.2 years.

We showed that sharp temporal changes in age-specific migration rates were the predominant contributor to this reversal. In particular, migration patterns of children, ages 0-17, appear to have determined most of relative aging across counties, together with migration of older people aged 55 and over. Natality and mortality differentials across counties played a limited role, with the exception of the 1970s when higher birth rates in rural counties contributed to their relative rejuvenation.

Bibliography

Bailey, Martha, Karen Clay, Price Fishback, Michael R. Haines, Shawn Kantor, Edson Severnini, and Anna Wentz (2018), "U.S. County-Level Natality and Mortality Data, 1915-2007."

Baum-Snow, Nathaniel and Daniel Hartley (2020), "Accounting for central neighborhood change, 1980–2010," *Journal of urban economics*, Vol. 117, pp. 1032–28.

Boehm, Michael and Christian Siegel (2021), "Make yourselves scarce: The effect of demographic change on the relative wages and employment rates of experienced workers."

Butler, Alexander W and Hanyi Yi (2019), "Aging and public financing costs: Evidence from US municipal bond markets," *Available at SSRN 3301648*.

Couture, Victor and Jessie Handbury (2020), "Urban revival in America," *Journal of Urban Economics*, Vol. 119, pp. 1032–67.

Diamond, Rebecca (2016), "The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000," *American Economic Review*, Vol. 106, No. 3, pp. 479–524.

Edlund, Lena, Cecilia Machado, and Maria Micaela Sviatschi (2015), "Gentrification and the rising returns to skill,"Technical report, National Bureau of Economic Research.

Ellen, Ingrid Gould, Keren Mertens Horn, and Davin Reed (2019), "Has falling crime invited gentrification?" *Journal of Housing Economics*, Vol. 46, pp. 1016–36.

Ganong, Peter and Daniel Shoag (2017), "Why has regional income convergence in the US declined?" *Journal of Urban Economics*, Vol. 102, pp. 76–90.

Glasgow, Nina and David L Brown (2012), "Rural ageing in the United States: Trends and contexts," *Journal of Rural Studies*, Vol. 28, No. 4, pp. 422–431.

Hoxie, Philip, Daniel Shoag, and Stan Veuger (2019), "Moving to Density: Half a Century of Housing Costs and Wage Premia from Queens to King Salmon," *AEI Paper & Studies*.

Hsieh, Chang-Tai and Enrico Moretti (2019), "Housing constraints and spatial misallocation," *American Economic Journal: Macroeconomics*, Vol. 11, No. 2, pp. 1–39.

Jones, Charles I (2020), "The end of economic growth? Unintended consequences of a declining population,"Technical report, National Bureau of Economic Research.

Kaplan, Greg and Sam Schulhofer-Wohl (2017), "Understanding the long-run decline in interstate migration," *International Economic Review*, Vol. 58, No. 1, pp. 57–94.

Karahan, Fatih, Benjamin Pugsley, and Ayşegül Şahin (2019), "Demographic origins of the startup deficit,"Technical report, National Bureau of Economic Research.

Keehan, Sean P, Devin A Stone, John A Poisal, Gigi A Cuckler, Andrea M Sisko, Sheila D Smith, Andrew J Madison, Christian J Wolfe, and Joseph M Lizonitz (2017), "National health expenditure projections, 2016–25: price increases, aging push sector to 20 percent of economy," *Health Affairs*, Vol. 36, No. 3, pp. 553–563.

Maestas, Nicole, Kathleen J Mullen, and David Powell (2016), "The effect of population aging on economic growth, the labor force and productivity,"Technical report, National Bureau of Economic Research.

Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles (2020), "IPUMS National Historical Geographic Information System," Version 15.0.

Meara, Ellen, Chapin White, and David M Cutler (2004), "Trends in medical spending by age, 1963–2000," *Health Affairs*, Vol. 23, No. 4, pp. 176–183.

Molloy, Raven, Christopher L Smith, and Abigail Wozniak (2011), "Internal migration in the United States," *Journal of Economic perspectives*, Vol. 25, No. 3, pp. 173–96.

_ , Riccardo Trezzi, Christopher L Smith, and Abigail Wozniak (2016), "Understanding declining fluidity in the US labor market," *Brookings Papers on Economic Activity*, Vol. 2016, No. 1, pp. 183–259.

Moretti, Enrico (2013), "Real wage inequality," *American Economic Journal: Applied Economics*, Vol. 5, No. 1, pp. 65–103.

Murdock, Steve H, Michael E Cline, Mary Zey, Deborah Perez, and P Wilner Jeanty (2015), "Effects of Demographic Change on Selected Economic Factors Impacting the Public and Private Sectors in the United States," in *Population Change in the United States*: Springer, pp. 63–91.

DCCPS, Surveillance Research Program National Cancer Institute (2021), "Surveillance, Epidemiology, and End Results (SEER) Program Populations (1969-2019)."

Preston, Samuel H., Christine Himes, and Mitchell Eggers (1989), "Demographic Conditions Responsible for Population Aging," *Demography*, Vol. 26, No. 4, pp. 691–704.

Rodden, Jonathan A. (2019), *Why Cities Lose: The Deep Roots of the Urban-Rural Political Divide*, New York: Basic Books.

Scala, Dante J, Kenneth M Johnson, and Luke T Rogers (2015), "Red rural, blue rural? Presidential voting patterns in a changing rural America," *Political Geography*, Vol. 48, pp. 108–118.

Shimer, Robert (2001), "The impact of young workers on the aggregate labor market," *The Quarterly Journal of Economics*, Vol. 116, No. 3, pp. 969–1007.

Su, Yichen (2019), "The rising value of time and the origin of urban gentrification."

Appendix

2.A Additional Figures and Tables

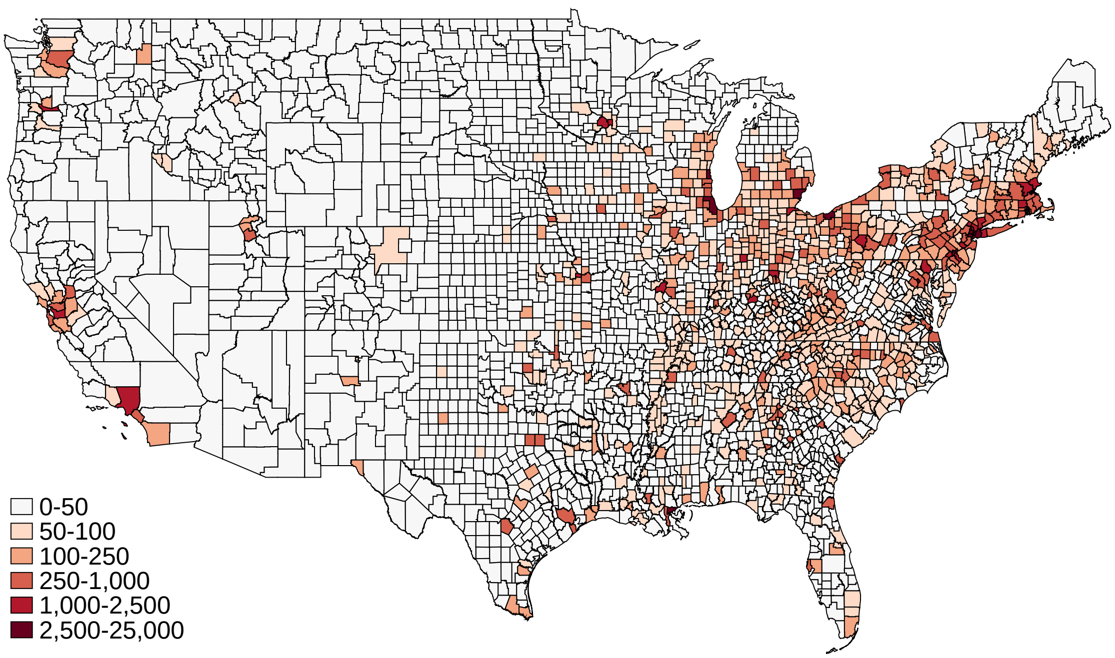


Figure 2.A.1: Map of county population density in 1950

Notes. The map represents six distinct levels of county population density in 1950.

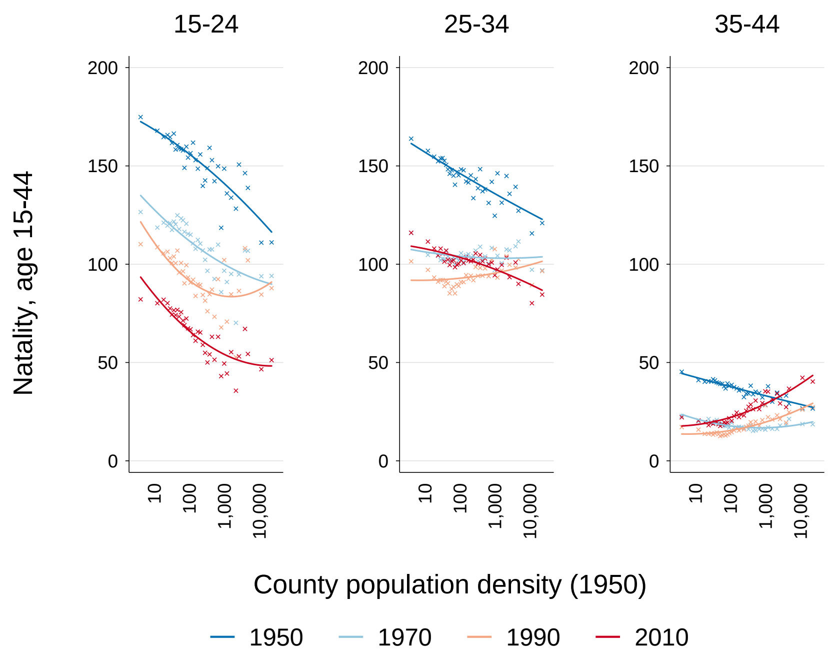


Figure 2.A.2: Natality over the life cycle across decades

Notes. Each graph represents the number of births per 1,000 women in a given age bracket for various decades. Counties are ordered by their 1950 population densities and weighted by their 1950 population. Each plotted point contains approximately 2.5 percent of 1950 population.

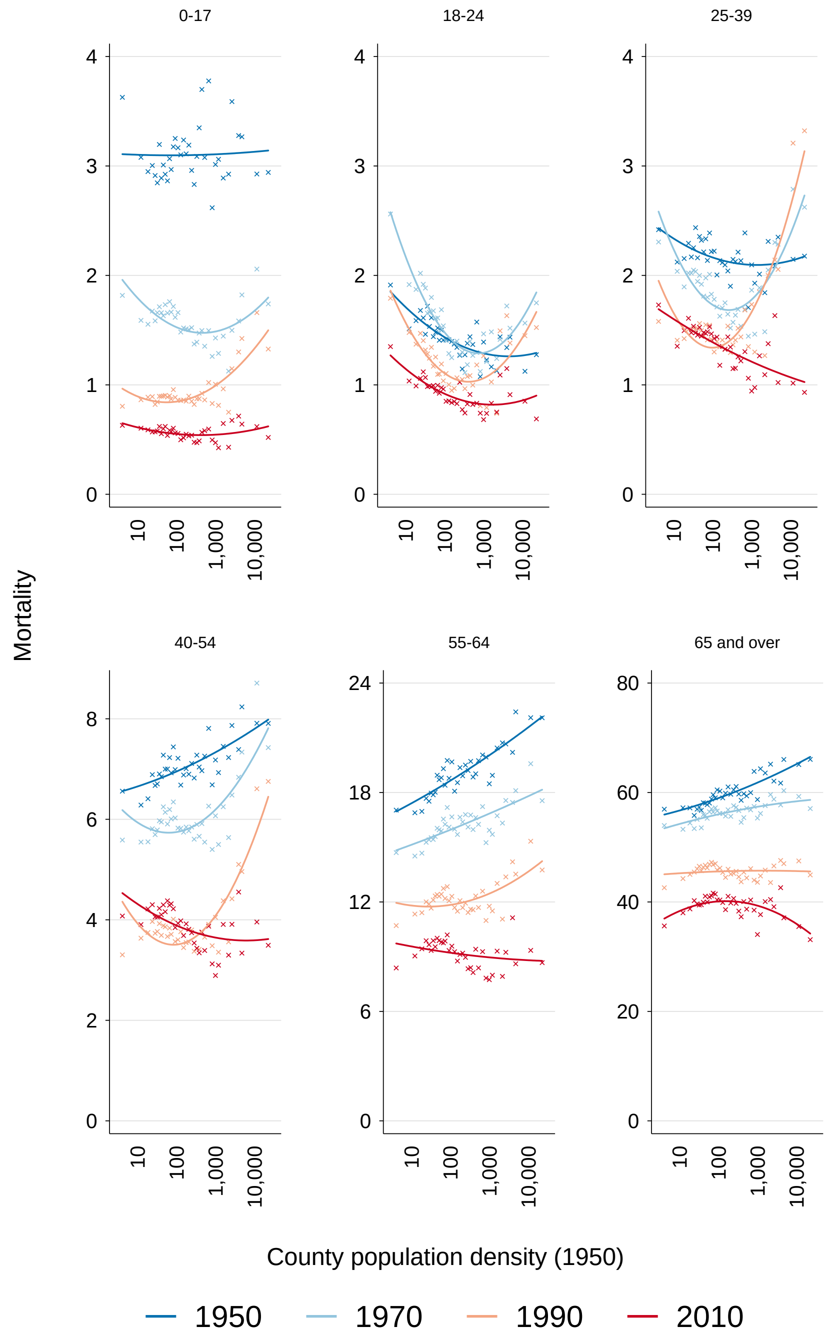


Figure 2.A.3: Mortality over the life cycle across decades

Notes. Each graph represents the number of deaths per 1,000 residents in a given age bracket for various decades. Counties are ordered by their 1950 population densities and weighted by their 1950 population. Each plotted point contains approximately 2.5 percent of 1950 population.

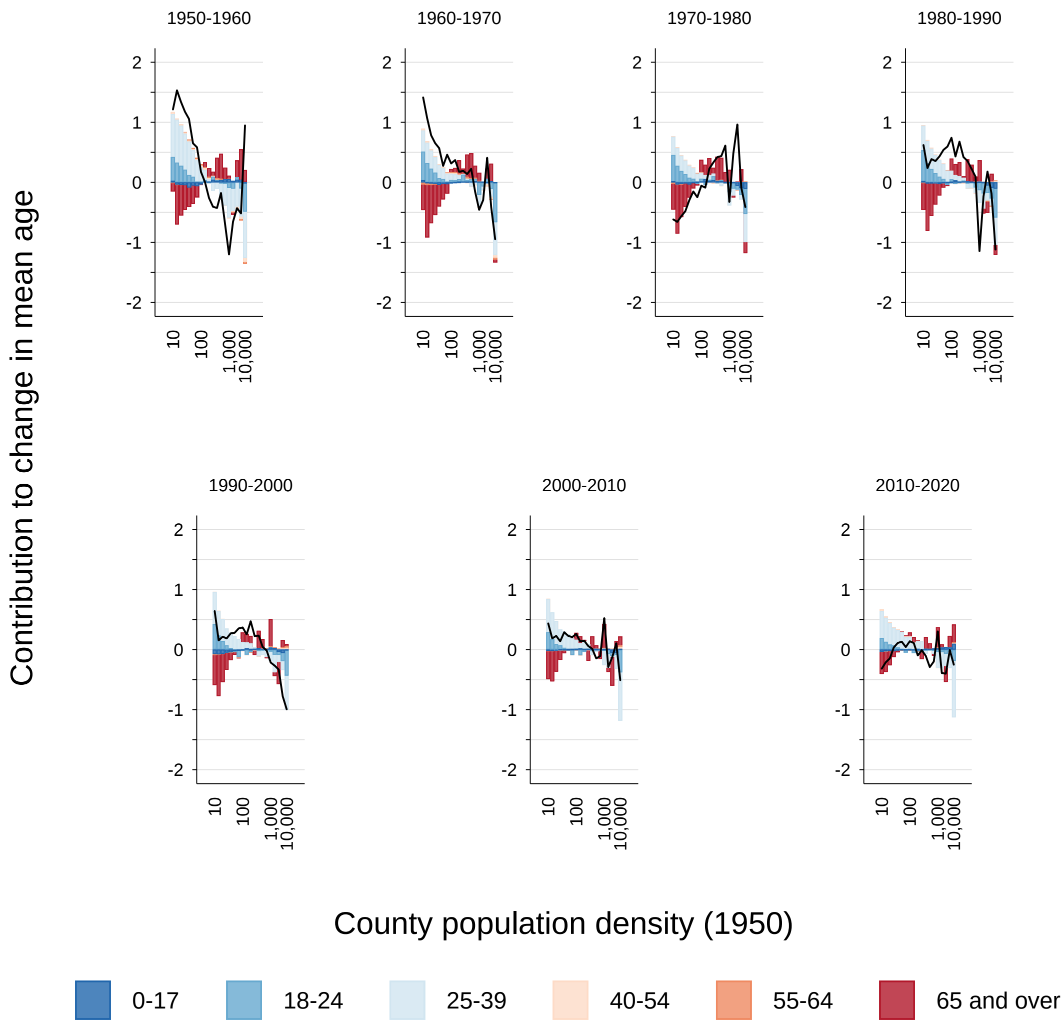


Figure 2.A.4: Decadal contributions of initial distribution of the different age brackets to changes in mean county age

Notes. Each graph represents the decadal contributions of initial distribution of the different age brackets to changes in mean county age to aging applying formula (2.3.7). Counties are ordered by their 1950 population densities and weighted by their 1950 population. Each plotted bar contains approximately 2.5 percent of 1950 population.

2.B Data Appendix

2.B.1 Data Sources

Population. We use two distinct sources for population data: SEER data from the National Cancer Institute (1969–2019), and U.S. Census data downloaded from NHGIS website (1950 and 1960). The SEER data are yearly bridged-race (White, Black and Other) estimates by single year of age at the county level based on Census Bureau intercensal data. NHGIS data for 1950 and 1960 provide county-level population counts by sex, race (White and Non-White), and age (single-year brackets from 0 to 20 years old, 5-year brackets for other age groups).

Natality and mortality. Natality and mortality data come from the National Center for Health Statistics, and are complemented with data assembled by Bailey et al. [2018] for earlier years. Micro-data files are available for natality over the period 1968-2018. They provide information about county of residence, mother's race and age and child's sex and race. A sample of 50 percent of births are recorded through 1988; the files are exhaustive afterwards. From 1950 to 1967, only data aggregated at the county level exist. We use Bailey et al. [2018], which does not contain any breakdown other than county of residence until 1959, when child's race is added. Similarly, micro-data files for mortality only cover the period 1959–2018. The files are exhaustive for all those years and contain county of residence, age, sex and race variables. Over 1950–1958, we use Bailey et al. [2018] data which are aggregated at the county level with no breakdown.

To impute age, sex and race variables in the earlier years of natality and mortality datasets, we use state-level data from the Vital Statistics of the United States (VSUS) publications. Those tabulations give number of births and deaths broken down by (mother's) age, sex and race.

County areas. We obtain the land area of each county from the County and City Data Book, 1947-1977 (downloaded from ICPSR website).

2.B.2 Variable Construction

Population. Both SEER and NHGIS data group some age brackets, which we unbundle using piecewise cubic Hermite interpolation. Specifically, people aged 85 and over are grouped in the SEER data, while NHGIS data only contain 5-year age bins for people over 20. To impute population by single age group from 0 to 99 years old, we write the cumulative distribution function of the age distribution by sex and race and then interpolate missing values using piecewise cubic Hermite interpolation (`pchipolate` in Stata).

Data imputations in the 1950s and 1960s. Both population data and vital statistics are incomplete in the 1950s and 1960s: No intercensal population estimate is available for those two decades; natality and mortality data are available yearly, but do not report a breakdown by age, sex and race.

To overcome these limitations, we first impute missing years and demographics in the 1960s, followed by the 1950s. For each decade, we proceed in five steps:

1. We impute population by age, sex and race of cohorts that are at least 10 years of age at the end of the decade (meaning they were born by the start of the decade). We apply log interpolation while allowing the growth rate to vary over time along a linear trend:

$$\log(pop_{t+H}/pop_t) = \sum_{h=0}^{H-1} r_{t+h}, \tag{2.B.1}$$

   with:

$$r_{t+h} = r_t + h\rho, \tag{2.B.2}$$

   where h is the elapsed time index from equation (2.B.1) above. We calibrate ρ and r_0 using Census population for each county, age, race and sex bin at the start and end of the decade as well as 5 years after the end of the decade.

2. Using state-level birth rates by mothers' age and race and children's sex, combined with intercensal distribution of women imputed in the previous step, we allocate the total births per county to women of different age and race groups.
3. Combining imputed birth counts by sex and race with the end-of-the-decade population counts, we impute population for cohorts who are less than 10 years old at the end of the decade applying log interpolation.
4. We allocate death counts in the 1950s given estimated population for each bin in each year.
5. Finally, we obtain county-level net migration as the residual of population changes by age group net of deaths as:

$$nmig_{t+1} = pop_{t+1} - pop_t + dcount_{t+1}, \tag{2.B.3}$$

   if age is positive and:

$$nmig_{t+1} = pop_{t+1} - bcount_{t+1} + dcount_{t+1}, \tag{2.B.4}$$

   for newborns.

Time-consistent counties. We inventory all county boundary changes that occurred in the United States since the 1950s by bringing together several sources:

1. Notes provided by David Dorn (University of Zurich), "FIPS County Code Changes": https://www.ddorn.net/data/FIPS_County_Code_Changes.pdf
   which report changes across commuting zones (which are clusters of contiguous counties) but

not across counties within commuting zones;

2. U.S. Census, “Substantial Changes to Counties and County Equivalent Entities: 1970-Present":
   https://www.census.gov/geo/reference/county-changes.html
   which enumerates only *substantial* changes, defined as: county boundary changes affecting an estimated population of 200 or more people; changes in boundaries of at least one square mile where no estimated population was provided but research indicated that the affected population may have been 200 people or more; or “large" annexations of unpopulated territory (10 square miles or more);
3. Atlas of Historical County Boundaries,“Consolidated Chronology" files:
   https://publications.newberry.org/ahcbp/
   which is exhaustive. We drop changes that are described as “small". In Tennessee, we drop the many reported changes “to accommodate local resident(s)", which are minute changes made when county boundaries cross private property boundaries.

We build a time-consistent set of counties, coded as FIPS5018, by identifying the connected components in the set of all substantial county changes using a depth-first search algorithm. The 3,010 components are our unit of analysis throughout the paper. We label each component with the FIPS code of the county that was the most populated in the 1990 Census.

2.B.3 Sample Selection

We focus on contiguous United States. In particular, we exclude the states of Alaska and Hawaii, and all other off-shore insular areas, including American Samoa, the U.S. Virgin Islands, Northern Mariana Islands, Guam and Puerto Rico.

2.C Omitted Proofs

2.C.1 Continuous-Time Results

Proposition 2.1. The change in mean age $A(t)$ is given by:

$$\frac{dA}{dt}(t) = \int_0^\infty p(a,t)r(a,t)[a - A(t)]da, \tag{2.C.1}$$

with $r(a,t) = \frac{1}{n(a,t)}\frac{\partial n}{\partial t}(a,t)$ the growth rate of population aged a over time.

Proof. By definition, mean age $A(t)$ is equal to:

$$A(t) \equiv \frac{\int_0^\infty n(a,t)ada}{N(t)} = \int_0^\infty p(a,t)ada, \tag{2.C.2}$$

with $N(t) = \int_0^\infty n(a,t)da$ and $p(a,t) = \frac{n(a,t)}{N(t)}$.

Differentiating with respect to t, we obtain:

$$\begin{aligned}\frac{dA}{dt}(t) &= \int_0^\infty \frac{\frac{\partial n}{\partial t}(a,t)N(t) - n(a,t)\frac{dN}{dt}}{N(t)^2} a da \\ &= \int_0^\infty p(a,t)[r(a,t) - \bar{r}(t)] a da \\ &= \int_0^\infty p(a,t)r(a,t) a da - \bar{r}(t)A(t) \\ &= \int_0^\infty p(a,t)r(a,t)[a - A(t)] da.\end{aligned} \tag{2.C.3}$$

□

Proposition 2.2. The change in mean age can be decomposed into:

$$\frac{dA}{dt}(t) = 1 + \partial_b A - \partial_d A + \partial_m A \tag{2.C.4}$$

where we define:

$$\partial_b A \equiv -\int_0^\infty p(a,t)\beta(a,t)A(t)da,$$

$$\partial_d A \equiv \int_0^\infty p(a,t)\delta(a,t)[a - A(t)]da,$$

and:

$$\partial_m A \equiv \int_0^\infty p(a,t)\mu(a,t)[a - A(t)]da,$$

with $\beta(a,t) = \frac{1}{n(a,t)}\frac{\partial b}{\partial t}(a,t)$, $\delta(a,t) = \frac{1}{n(a,t)}\frac{\partial d}{\partial t}(a,t)$ and $\mu(a,t) = \frac{1}{n(a,t)}\frac{\partial m}{\partial t}(a,t)$ the respective natality rate, mortality rate and migration rate for age a.

Proof. Let's first notice that:

$$r(a,t) = -\frac{1}{n(a,t)}\frac{\partial n}{\partial a}(a,t) - \delta(a,t) + \mu(a,t) + \frac{1}{n(0,t)}\left(\int_0^\infty \beta(\tilde{a},t)n(\tilde{a},t)d\tilde{a} - n(0,t)\right)\mathbb{1}_0. \tag{2.C.5}$$

with $\mathbb{1}_0$ the Dirac distribution in 0.

We define natural aging as:

$$\nu(a,t) = -\frac{1}{n(a,t)}\frac{\partial n}{\partial a}(a,t) - \mathbb{1}_0. \tag{2.C.6}$$

Substituting for $r(a,t)$ in equation (2.C.1), we get that:

$$\frac{dA}{dt}(t) = \partial_n A + \partial_b A - \partial_d A + \partial_m A \tag{2.C.7}$$

where:

$$\partial_n A \equiv \int_0^\infty p(a,t)\nu(a,t)[a - A(t)]da.$$

Now, let's show that $\partial_n A = 1$. The intuition runs as follows. Component $\partial_n A$ corresponds to the effect on mean age of the aging of the existing population. If everyone ages by one year, than the average age also increases by one year.

Formally:

$$\begin{aligned}
\partial_n A &= \int_0^\infty p(a,t)\nu(a,t)[a - A(t)]da \\
&= -\int_0^\infty p(a,t)\frac{1}{n(a,t)}\frac{\partial n}{\partial a}(a,t)[a - A(t)]da + p(0,t)A(t) \\
&= -\frac{1}{N(t)}\int_0^\infty \frac{\partial n}{\partial a}(a,t)[a - A(t)]da + p(0,t)A(t) \\
&= -\frac{1}{N(t)}\left[0 - \big(-A(t)\big)n(0,t) - \int_0^\infty n(a,t)da\right] + p(0,t)A(t) \\
&= 1.
\end{aligned} \tag{2.C.8}$$

□

2.C.2 Implementation with Discrete-Time Data

In practice, we observe population, births, deaths and migration on a yearly basis and with discrete values of age. With non-continuous data, the results of the previous section do not hold exactly.

In particular, if we define growth rates $r(a,t) = \frac{n(a,t+1) - n(a,t)}{n(a,t)}$ and $\bar{r}(t) = \frac{N(t+1) - N(t)}{N(t)}$, then we have:

$$\begin{aligned}
\Delta A(t) &\equiv A(t+1) - A(t) \\
&= \sum_{a=0}^{\infty} ap(a,t+1) - \sum_{a=0}^{\infty} ap(a,t) \\
&= \sum_{a=0}^{\infty} a\frac{n(a,t+1)N(t) - n(a,t)N(t+1)}{N(t)N(t+1)} \\
&= \frac{1}{1+\bar{r}(t)}\sum_{a=0}^{\infty} ap(a,t)[r(a,t) - \bar{r}(t)] \\
&= \frac{1}{1+\bar{r}(t)}\sum_{a=0}^{\infty} p(a,t)r(a,t)[a - A(t)].
\end{aligned} \tag{2.C.9}$$

If we defined growth rates with population at $t+1$ at the denominator, the unwanted factor $\frac{1}{1+\bar{r}(t)}$ would drop, but we would lose another useful property. Specifically, keeping time t at the denominator allows us to define discrete-time natural aging as:

$$\nu(a,t) \equiv -\frac{1}{n(a,t)}\big(n(a,t) - n(a-1,t)\big) \tag{2.C.10}$$

for $a > 0$ and:

$$\nu(0,t) \equiv -1, \tag{2.C.11}$$

which satisfies:

$$\begin{aligned}
\Delta_n A &= \sum_{a=0}^{\infty} p(a,t)\nu(a,t)[a - A(t)] \\
&= -\frac{1}{N(t)} \left(\sum_{a=1}^{\infty} \big(n(a,t) - n(a-1,t)\big)[a - A(t)] - n(0,t)A(t) \right) \\
&= -\frac{1}{N(t)} \left(\sum_{a=1}^{\infty} \big(n(a,t) - n(a-1,t)\big)a \right) \\
&= -\frac{1}{N(t)} \left(\sum_{a=1}^{\infty} n(a,t)a - \sum_{a=1}^{\infty} n(a-1,t)(a-1) - \sum_{a=1}^{\infty} n(a-1,t)) \right) \\
&= 1.
\end{aligned} \tag{2.C.12}$$

This property underpins the decomposition:

$$\Delta A = 1 + \Delta_b A - \Delta_d A + \Delta_m A, \tag{2.C.13}$$

which would not hold exactly with the alternative definition of growth rates.

In practice, we resolve this trade-off by using the average population between t and $t+1$ at the denominator of growth rates:

$$r(a,t) = \frac{n(a,t+1) - n(a,t)}{\frac{1}{2}\big(n(a,t) + n(a,t+1)\big)}, \quad \bar{r}(t) = \frac{N(t+1) - N(t)}{\frac{1}{2}\big(N(t) + N(t+1)\big)}, \tag{2.C.14}$$

and similarly for ν, β, δ and μ.

Chapter 3

Redistribution through Housing Assistance

joint with Hector Blanco

3.1 Introduction

"In policy circles, [housing] vouchers were known as a 'public private partnership'. In real estate circles, they were known as a 'win'."

Evicted, Matthew Desmond (2016)

The past few decades saw a dramatic change in the provision of housing assistance to low-income families in the United States. The federal government gradually shifted away from constructing public housing through local public housing authorities (PHAs) toward subsidizing private housing. An emblematic measure of this transition is the HOPE VI program, initiated in 1992, which led to the demolition of hundreds of public housing developments. Concurrently, tenant-based rental assistance—such as Section 8 Housing Choice Voucher Program—and project-based rental assistance—including the Low Income Housing Tax Credits (LIHTC)—have expanded considerably (see Figure 3.1.1). Those programs leave a larger role to private developers and property owners, who allegedly capture a substantial share of the benefits intended to disadvantaged households [Desmond, 2016]. However, the academic literature has so far largely ignored the distributive implications of the decline of public housing.

This paper builds a framework to characterize theoretically and quantitatively the redistributive implications of housing assistance. First, we argue that public housing may in theory improve redistribution efficiency when income taxation does not dissociate wage and rental incomes. We then estimate structurally our model leveraging demolition of public housing as an exogenous shocks to pin down key elasticities. In future work, we intend to apply this model to assess the incidence of the shift from public housing to subsidized private housing and to benchmark our results against an optimally-designed housing assistance program.

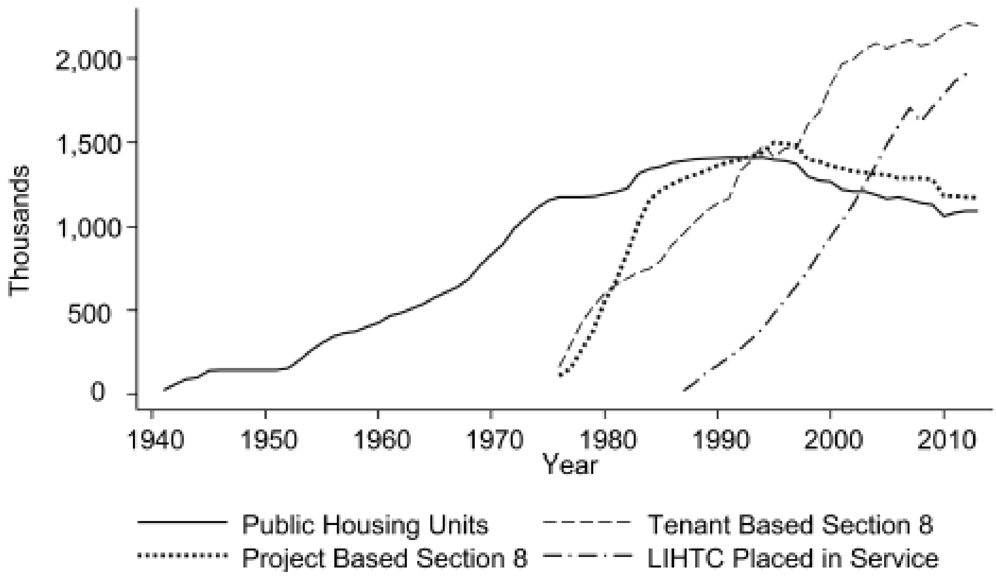


Figure 3.1.1: Number of beneficiaries by housing assistance program
Source. Collinson et al. [2019]

Housing assistance programs feature a trade-off between indirect pecuniary redistribution and direct amenity effects. On the one hand, public housing increases the stock of housing units, which drives local rents down, but amplifies the spatial concentration of poverty, lowering local amenities for recipients. On the other hand, voucher and LIHTC programs improve the local amenities of subsidized households, while pushing private landowners' rents up.

Focusing on seven major metropolitan areas[1], we establish two facts that underpin this trade-off. First, for each of these cities, the demolition of public housing has not been compensated by the construction of new housing units, whether subsidized or not. Therefore, we expect that the transition benefited local landowners through higher rents. Second, the demographic composition of census tracts affected by demolitions changed drastically. In particular, the education level and per capita income increased more than twice as fast as in unaffected areas. Poor households may have benefited from lower spatial concentration of poverty through enhanced amenities.

To quantify the distributional consequences of housing assistance programs, we develop a quantitative urban model with redistributive policies and endogenous amenities. The main departure from the existing literature is the introduction of two types of tax instruments, non-linear income taxation and housing assistance, that are used by the government to redistribute across house-

[1] Namely, Atlanta, Baltimore, Chicago, Memphis, Newark, Pittsburgh and Washington, DC.

holds. We assume that the set of taxes that can be imposed on the workers are limited to be a function of their total income, but not their type, following Mirrlees [1971]. This restriction creates a trade-off between redistribution and productive efficiency. As in Naito [1999], housing assistance programs complement non-linear income taxation by affecting the pre-taxation income distribution—specifically, the distribution of land rents.

When estimating structurally our model, the main challenge is to disentangle amenity effects from housing demand and supply elasticities. To estimate the housing supply elasticity, we leverage the HOPE VI program as a shock to housing demand on the private housing market. On the demand side, we use two instruments to estimate both the housing demand elasticity and the amenity spillovers. The first instrument is the the LIHTC eligibility rules, which induce an exogenous variation affecting neighborhood composition, and the second is the distance to new rapid transit lines.

Related literature. Our paper combines tools from public and urban economics to study housing policy.

We propose a novel interpretation of the redistributive virtues of in-kind transfers which builds on insights from the optimal taxation literature. Early theoretical work argued that in-kind transfers could improve redistribution absent lump-sum transfers. Nichols and Zeckhauser [1982] point at targeting efficiency, while Coate et al. [1994] are the first to explore pecuniary externalities as a rationale for in-kind transfers. We provide a general equilibrium framework to think about those so-called pecuniary externalities. Our main point is that public ownership of some fixed factor–land in the case of housing assistance–is the underlying reason why in-kind transfers can enhance redistribution. The formal argument resembles Naito [1999]'s seminal paper and recent work by Costinot and Werning [2018]: Assuming that the set of taxes that can be imposed on the workers are limited to be a function of their total income, but not their type, housing assistance programs complement non-linear income taxation by affecting the pre-taxation income distribution—specifically, the distribution of land rents.

This paper adds to a growing empirical literature demonstrating both indirect the pecuniary effects and the direct amenity spin-offs of housing assistance programs. Work in progress by Blanco [2021] suggests that public housing demolitions in Chicago induced significant price increases in local housing markets. Susin [2002] and Collinson et al. [2019] show that rents for low-income households increased in areas with more housing vouchers, and Susin [2002] argues that the overall rent increase is considerably larger than the subsidy to housing voucher beneficiaries. Diamond and McQuade [2019] find that new LIHTC buildings increase prices in low-income neighborhoods and decrease them in high-income areas. Diamond and McQuade [2019], as well as Davis et al. [2019a] and Davis et al. [2019b], estimate dynamic discrete choice models of location choice that include households that value endogenous characteristics such as demographic composition and median income in the neighborhood to rationalize households' location decisions in response to changes in housing policies.

We borrow techniques from the quantitative urban literature to quantify the aforementioned trade-off between pecuniary effects and amenity spin-offs of housing assistance programs. This body of work highlights the role of endogenous amenities in explaining households' location choice within a city [Ahlfeldt et al., 2015a; Couture et al., 2019; Tsivanidis, 2019b]. Recently, Gaubert et al. [2021] introduced optimal taxation into a quantitative spatial model to rationalize place-based redistribution. Our work introduces a second tax instrument, housing assistance, that is used to redistribute across households through pecuniary effects, but also through endogenously-determined local amenities.

Layout. The rest of the paper is organized as follows. Section 3.2 provides an overview of housing assistance in the United States and descriptive evidence of the major changes in housing assistance provision since the early 1990s. In section 3.3, we develop a quantitative urban model with non-linear income taxation and housing assistance. We estimate its key structural parameters in section 3.4. Finally, section 3.5 concludes.

3.2 Background and Data

We give a brief overview[2] of the history of the three main housing assistant programs—public housing, housing vouchers and LIHTC—and their relevant characteristics. A combination of administrative datasets and real estate transactions from the Zillow database allows to investigate the impact of the major changes in the mix of housing assistance programs at the census tract level since the early 1990s. We conclude this section with descriptive facts mirroring the trade-off between pecuniary effects and amenity spinoffs.

3.2.1 Background

3.2.1.1 Historical Overview: From Public Housing to Subsidized Private Housing

Although public housing used to be the main housing assistance program until the early 1990s, tenant-based vouchers and privately-owned subsidized housing, mainly in the form of LIHTC, experienced an enormous increase in the past few decades.

Public housing was introduced in the late 1930s as a solution to housing affordability, but it proved unsustainable by the late 1980s due to the poor maintenance of the buildings. The Housing Act 1937 aimed at providing affordable housing for low-income people by constructing high-rise buildings that were to be managed by public housing authorities (PHAs). The initial intention was for the federal government to pay for the construction of public housing development and for PHAs

[2] For a more comprehensive description of housing assistance programs, see Collinson et al. [2019] and Vale and Freemark [2012].

to be in charge of their maintenance through rent revenues. However, PHAs failed to upkeep the public housing developments and, by the end of the 1980s, most of the high-rise buildings were in a very poor condition and concentrated high levels of poverty and crime.

As a result, Congress approved the HOPE VI program in 1993, which demolished around 8% of the nation's public housing stock. Under this program, PHAs could apply for funds to demolish, revitalize or rehabilitate public housing developments that were considered to be "severely distressed". This resulted in 100,000 demolished units, 11,000 rehabilitated units and approximately 90,000 new projected units. Out of the latter, 13% are market rate units and the remaining 87% are affordable housing units, which might include public housing, LIHTC or other subsidized housing. An immediate result of the HOPE VI program was the displacement of many families living in public housing. Around half of them were relocated to other public housing, whereas the other half were relocated to housing vouchers.

By the time HOPE VI was approved, housing vouchers had become the largest housing assistance program by number of beneficiaries, and the Low-Income Housing Tax Credit (LIHTC) program had been introduced in 1986 to boost the production of mixed-income multifamily housing. Figure 3.1.1 illustrates the dramatic shift in the mix of housing assistance programs over the past few decades. As programs like HOPE VI downsized public housing, there are now twice as many beneficiaries from either tenant-based vouchers or the LIHTC program as from public housing.

3.2.1.2 Program Characteristics: Concentrated Poverty versus Mixed-Income Communities

We now summarize some relevant characteristics of each of the three main housing assistance program.

Eligibility and the level of subsidized rents is determined similarly in each of the housing assistance programs. Eligibility is based on the total annual gross income and the family size. PHAs tier income limits defined as a percentage of the Area Median Income (AMI), and reserve some units for the poorest households. The rent level is fixed at 30% of the monthly adjusted income of the family, the rest being covered by PHAs up to a rent ceiling (known as *Fair Market Rent*).

The main argument in favor of vouchers and LIHTC over public housing is that the former programs give the opportunity to beneficiaries to live in better neighborhoods without concentrating low-income individuals in high-rise buildings. Voucher recipients are free to choose any housing that meets the requirements of the program and are not limited to units located in subsidized housing projects. At least 20 percent of the tenants in LIHTC projects must earn less than 50 percent of the area median gross income (AMGI) or, alternatively, at least 40 percent of them must earn less than 60 percent of AMGI. The other units tend to be occupied by middle-income households.

As a compensation for welcoming low-income households and curbing rents, private owners participating in subsidized-housing programs receive tax credits. To boost the production of mixed-income multifamily housing, the LIHTC program allocates federal tax credits based on population to the states, which in turn award these credits to developers of qualified projects. Developers can sell these credits to investors to raise equity capital for their projects and reduce the amount of capital they would otherwise have to borrow. Hence, investors receive a dollar-for-dollar credit against their federal tax liability for a period of 10 years as long as the qualified project complies with the program guidelines[3]. In exchange for the credit benefits, developers must not only welcome a substantial share of low-income households, but also restrict rents, including utility allowance, in low-income units to 30 percent of the relevant income limit for a minimum affordability period of 30 years.

3.2.2 Data

We bring together various datasets to obtain a comprehensive picture of the evolution of housing assistance programs, house prices and socio-demographic characteristics at the 2010 census tract level for the period 1990-2010.

Housing assistance programs. The first dataset compiles information from several sources to cover all public housing buildings active at any point between 1995 and 2018. We use data coming from 1996 HUD-951 forms[4], which contains a snapshot of public housing building addresses, units and geographical coordinates for developments in that year, complemented by similar dataset for the year 2018.

Administrative data on the HOPE VI program are issued by the Department of Housing and Urban development (HUD). The data report the magnitude and timing of the demolitions, as well as new construction for developments linked to a HOPE VI *revitalization grant*—those that involved some reconstruction. For *demolition grants*, we use publicly-available data containing the award year and the number of demolished units at the project level. For the city of Chicago, we also include the list of non-HOPE VI demolished public housing units provided by the Chicago Housing Authority.

Lastly, data on the LIHTC and tenant-based vouchers come from two additional sources. First, the public LIHTC database which contains address-level information on LIHTC-financed projects for the period 1987-2019. Second, the Picture of Subsidized Households includes the number of households per program and census tract during the period 1993-2019, with some discontinuously over the time period.

House prices. Zillow's Ztrax data on real estate transactions includes information regarding all real estate transactions in the main U.S. metropolitan areas starting in the early 1990s, as well as property characteristics recorded from local tax assessor's data. For each house sale, the transaction

[3] A more detailed description can be found in Diamond and McQuade [2019]

[4] These are forms that public housing authorities (PHAs) were required to report to the Department of Housing and Urban development (HUD) containing information on all of their public housing buildings. This dataset is publicly available in the HUD website.

dataset contains a transaction id, address, sale date, sale price, mortgage information, foreclosure status and other information collected by the local tax assessor. We merge this with other property characteristics that Zillow acquired from local assessors' offices.

We clean the data to include only residential arms-length transactions and eliminate outliers. For the former, we restrict the sample to property transactions with a residential use and drop those signaled as intra-family transactions. For the latter, we eliminate outliers by excluding transactions in the top percentile of the yearly price distribution.

Our main outcome of interest is constructed as a quality-adjusted house price index at the census tract level, following Baum-Snow and Han [2020] and Blanco [2021]. The house price index ρ_{ict} is obtained from the regression:

$$\ln \mathrm{P}_{hict} = \rho_{ict} + \alpha_m + \gamma' \mathbf{X}_{hict} + u_{hict}. \tag{3.2.1}$$

The left-hand side is the logarithm of the sale price of property h in census tract i in county c in year t. α_m are month-of-sale fixed effects that capture seasonality in sale prices, whereas $\mathbf{X}_{hict}$ is a vector of property characteristics, including building type, building age dummies, lot size, lot size squared, number of stories, number of bedrooms and roof cover type[5]. We define the house price index as the census tract-county-year fixed effects in the regression above, ρ_{ict}.

Other data sources. Two additional sources complete our data. First, local data on demographic, socioeconomic and housing characteristics at the census tract level come from the decennial census for years 1990, 2000, 2010, downloaded from National Historical Geographic Information System (NHGIS). Second, we exploit a shapefile of the rail transit network in the United States in the years 1980, 1990, 2000 and 2004 from the National Transportation Atlas Database[6] to construct an instrument for neighborhood composition changes based on subway station openings.

3.2.3 Descriptive Evidence

We document the shift from public housing to subsidized private housing and how it affected the most exposed areas in a group of seven cities[7]. We uncover three main facts: First, the decline in public housing was far from offset by the rise tenant-based and project-based subsidized private housing; second, the local demographic composition changed dramatically following public housing demolitions; third, demolition exposure is associated with price hikes.

The HOPE VI program led to a sharp reduction in public housing units, driving down housing supply. While the program financed the demolition of approximately 48,000 units in these cities,

[5]Since some property characteristics are missing from some transactions, we generate dummy variables for missing values for each property characteristic except building type (which is never missing) and re-code missing values as zeros. In the regression, we include a term interacting each characteristic's missing dummy variable with building type to flexibly account for heterogeneity in that characteristic across property types when data is missing.

[6]We gratefully acknowledge Nathaniel Baum-Snow for sharing these data, used in Baum-Snow et al. [2005].

[7]Namely, Atlanta, Baltimore, Chicago, Memphis, Newark, Pittsburgh and Washington, DC.

only around 6,600 units (14%) were rebuilt as public housing. The bulk of the new construction relied on other types of affordable (private) housing, of which almost 13,000 units were built. Figure 3.2.1 plots these numbers by city. Although this pattern repeats across cities, Chicago is a notable outlier, accounting for almost half of demolished units in the sample and with only 7% of them being regenerated as public housing.

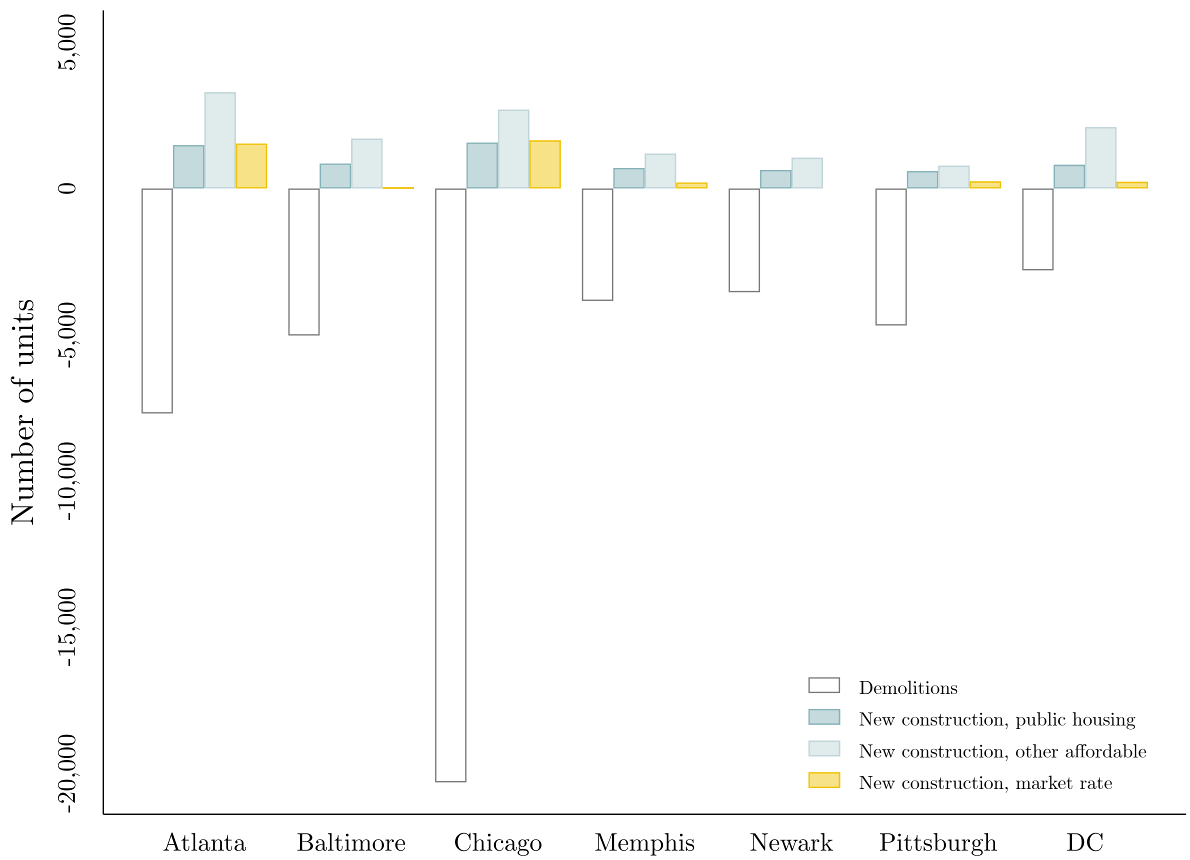


Figure 3.2.1: Demolitions under HOPE VI were far from offset by new construction

Notes. The barplot show the total number of units demolished and constructed between 1995 and 2011 as reported in HOPE VI administrative data. The category "New construction, public housing" is a conservative estimate: it also includes new construction labeled as a mix between public housing and other affordable housing. The category "New construction, other affordable" includes both the Low Income Housing Tax Credit units and units generally labeled as affordable.

Table 3.2.1 shows that neighborhoods exposed to public housing demolitions experienced substantial demographic changes, pointing at plausible local amenity effects. Between 1990 and 2010, tracts with demolitions increased their education levels and per capita income by more than twice the average remaining tract[8].

[8] Tach and Emory [2017] provide a detailed analysis of how the demographic composition of these neighborhoods changed after the implementation of the HOPE VI program

	Pre-demo: 1990		Post-demo: 2010		Change (%)	
	Other	Demo	Other	Demo	Other	Demo
Population	3,577	3,047	3,744	2,137	5	-30
Black share	0.30	0.83	0.35	0.73	15.28	-12.10
Educ: $\geq$ bachelor	0.23	0.08	0.33	0.23	44.33	193.26
Per capita income	16,176	6,503	30,859	20,244	91	211
Housing units	1,496	1,323	1,667	1,116	11	-16
Demolished units	0	423	0	423		
Observations	2826	123	2826	123	1	1

Table 3.2.1: Exposition to demolitions coincides with neighborhood composition changes

Notes. This table reports the mean of several census tract characteristics for census tracts affected by demolitions ("Demo") and remaining tracts in the included counties ("Other"). The included counties belong to the cities of Atlanta, Baltimore, Chicago, Memphis, Newark, Pittsburgh and Washington, DC.

Finally, we show that exposure to public housing demolitions is associated with local housing price hikes, which suggests a joint effect of deprived housing supply and improved amenities. To proceed, we regress the change in the house price index between the early 1990s and 2010 in census tract i on a demolition exposure index, defined as as follows:

$$\text{Demolition exposure}_i = \frac{1}{H_i^{1990}} \sum_j \frac{1}{\exp d_{ij}} \times \text{ Demolished units}_i$$

where H_i^{1990} is the baseline number of housing units in the tract and d_{ij} is the distance between census tracts i and j. That is, the index captures the number of demolished units weighted by their distance to census tract i. After controlling for several baseline characteristics, the relationship between house price changes and demolition exposure is positive and very significant (Figure 3.2.2).

To provide suggestive evidence of a supply channel, Table 3.2.2 reproduces the regression above but also interacts the demolition exposure index with the median household income tercile of the census tract (within each city). The table shows that, for a given level of demolition exposure, house prices increased by more in low-income census tracts. This fact, which is robust to including several control variables, suggests that housing in census tracts that competed directly with public housing suffered higher increases due to a reduction in public supply. Conversely, richer census tracts also experiencing higher house prices may be explained by richer households valuing more amenity changes.

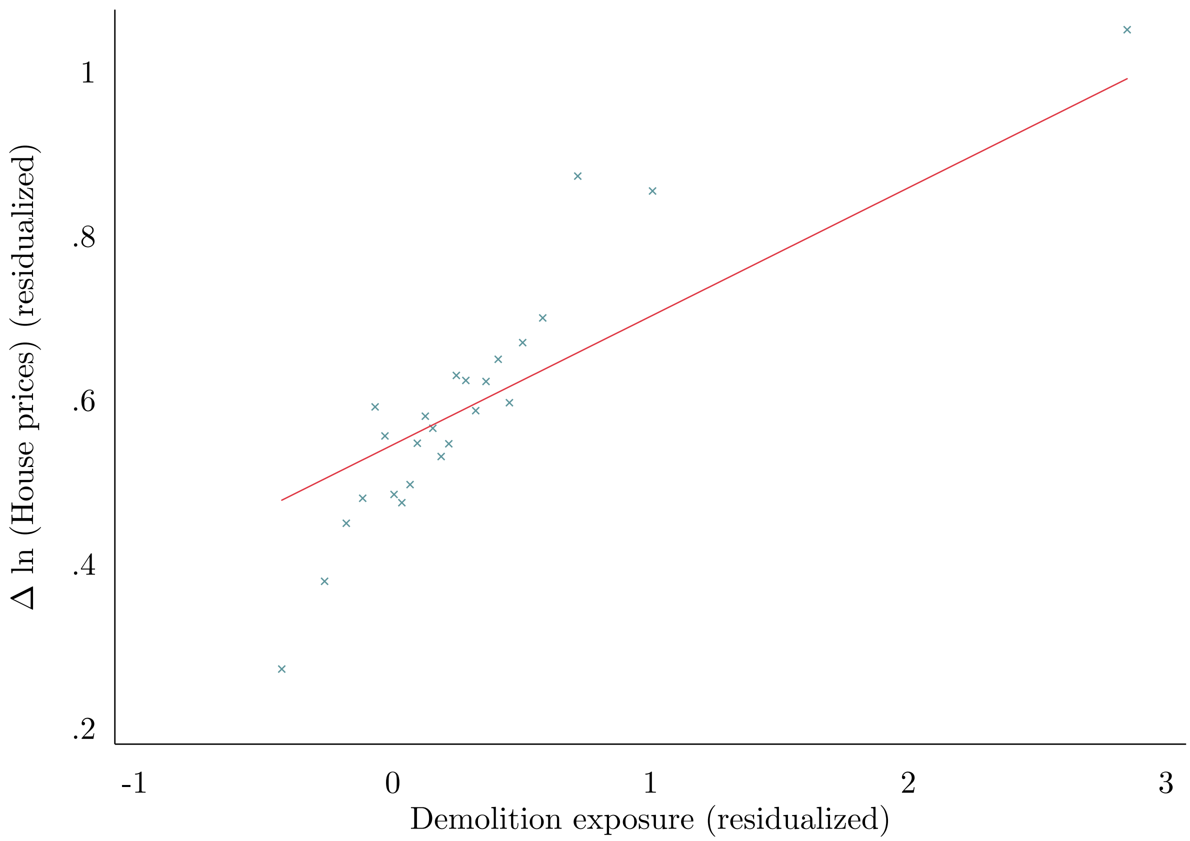


Figure 3.2.2: Increased exposure to demolitions raises house prices

Notes. The figure is a binned scatter plot of the increase in (the logarithm of) house prices between 1990 and 2010 on the demolition exposure index, after residualizing them for several baseline characteristics in 1990 (education levels, income per capita, black share and the number of housing units in 1990), the change in the share of housing units owned by the public sector, and county fixed effects.

	(1)	(2)	(3)
Demolition exposure	0.074***	0.073***	0.086***
	(0.019)	(0.020)	(0.018)
Demolition exposure × Low Income	0.117**	0.120***	0.131***
	(0.036)	(0.030)	(0.030)
Demolition exposure × High Income	0.041	0.040	0.050**
	(0.025)	(0.026)	(0.018)
Baseline prices	No	Yes	Yes
Baseline census chars.	No	No	Yes
County FE	Yes	Yes	Yes
N	2837	2837	2837

Table 3.2.2: Exposure to demolitions raises house prices more in low-income areas

Notes. All columns include fixed effects for low and high income census tracts. Baseline census characteristics contain education levels, black shares and the number of housing units in 1990. The included counties belong to the cities of Atlanta, Baltimore, Chicago, Memphis, Newark, Pittsburgh and Washington, DC. Standard errors in parenthesis, clustered at the county level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.3 A Quantitative Model with Income Taxation and Housing Assistance

We develop a quantitative urban model to analyze the welfare consequences of housing assistance programs. The main departure from the existing literature is the introduction of two types of tax instruments, non-linear income taxation and housing assistance, that are used by the government to redistribute across households. We assume that the set of taxes that can be imposed on the workers are limited to be a function of their total income, but not their type, following Mirrlees [1971]. This restriction creates a trade-off between redistribution and productive efficiency. In the spirit of Naito [1999], housing assistance programs complement non-linear income taxation by affecting the pre-taxation income distribution—specifically, the distribution of land rents.

3.3.1 Environment

The starting point of this model is an urban framework where the city is assumed to be a collection of neighborhoods distant from each other. This city is populated by heterogeneous workers who supply labor elastically and own land. Redistribution is achieved through two distinct policies: non-linear income taxation and means-tested housing assistance.

3.3.1.1 Setup

The city is comprised of $\mathcal{I}$ locations or neighborhoods, that differ in their fundamental levels of amenity and productivity, their land supply and their distances to other locations. Three types of agents step in: workers, producers and developers.

Households, indexed by Θ, are heterogeneous in skill, preferences over locations, land ownership and family characteristics. A type-Θ individual chooses a location i in which to live and a location j in which to work. He derives utility from the consumption of tradable goods and residential floorspace, but incurs a disutility from supplying labor. Poor households spend a relatively higher share of their revenues on residential floorspace.

Good production and floorspace development occur in the various neighborhoods of the city. Both sectors are perfectly competitive. Producers assemble labor supplied by the different skill groups, commercial floorspace and intermediate goods into tradable goods. Developers use labor, intermediate goods and an additional fixed factor, land, to develop residential and commercial floorspace which is rented to workers and producers respectively.

Government redistributes across households with two policy instruments: non-linear income taxation and means-tested housing assistance. We assume that these policies are restricted to be a function of their total income, but not their type, following Mirrlees [1971]. In particular, the government can't disentangle between labor supply and skill, and does not observe the composition of income between labor wages and land rents. As a result, a redistribution-efficiency trade-off arises.

As in Naito [1999], housing assistance programs complement non-linear income taxation by affecting the pre-taxation income distribution. The government implements three distinct programs: public housing, subsidized private housing and tenant-based vouchers. In equilibrium, housing assistance programs redistribute from landlords towards low-skilled households through two channels. First, directly by providing subsidized housing to the poorest workers. Second, indirectly by distorting equilibrium prices on the private housing market.

3.3.1.2 Preferences

Workers have weakly separable preferences between consumption and labor supply. They derive utility from consuming tradable good, c, and residential floorspace, h, and local amenities, a_i, but experience disutility from supplying labor, n. The utility of worker Θ is given by:

$$U(\Theta) = u(V(\Theta), n(\Theta), a_i; \Theta), \tag{3.3.1}$$

$$V(\Theta) = v(c(\Theta), h(\Theta)), \tag{3.3.2}$$

where $V(\Theta)$ is the sub-utility that worker Θ derives from consumption of tradable goods $c(\Theta)$ and housing $h(\Theta)$, and $n(\Theta)$ is his labor supply. We assume that the both utility functions $u(\cdot; \Theta)$ and $v(\cdot)$ are quasi-concave and strictly increasing.

Workers are distinguished by their multi-dimensional type $\Theta = (\theta, \varepsilon, \omega, \xi)$ with distribution F. The parameter θ indexes the household's skill, ε is a vector of idiosyncratic preference shocks for living in each location $i \in \mathcal{I}$, ω captures land ownership and ξ family characteristics, e.g. family size, that affect government redistributive preferences.

3.3.1.3 Technology

Tradable good production and floorspace development take place in each neighborhood.

Good production. Producers assemble tradable goods in every neighborhood j. They use labor inputs, $N_j^Y = \big(n_j^Y(\theta)\big)$, intermediate goods, M_j^Y, and commercial floorspace, H_j^Y. Their production technology Y_j is neighborhood-specific:

$$Y_j = Y_j\big(N_j^Y, M_j^Y, H_j^Y\big). \tag{3.3.3}$$

It exhibits constant returns to scale.

Floorspace development. Developers provide residential and commercial floorspace to households and producers in all neighborhoods. The presence of a fixed factor, land, induces decreasing returns to scale. Land ownership is split between the private households and the government.

Floorspace supply in sector $s \in \{P, G\}$ is given by:

$$H_{j,s} = H_{j,s}\big(N^H_{j,s}, M^H_{j,s}\big), \tag{3.3.4}$$

with $N^H_{j,s} = \big(N^H_{j,s}(\theta)\big)$ labor inputs and $M^H_{j,s}$ are intermediate goods, for $j \in \mathcal{I}$.

Amenity and productivity spillovers. Amenities in neighborhood i depend on the local distribution of types:

$$a_i = \bar{a}_i a(\mathbb{E}_{|i}[\theta]), \tag{3.3.5}$$

where $\bar{a}_i$ is the fundamental level of amenities in location i, a is some function of average skill of workers living in i, $\mathbb{E}_{|i}[\theta]$.

Similarly, we assume that productivity in neighborhood j is a function of the local distribution of workers:

$$y_j = \bar{y}_j y(\mathbb{E}_{|j}[\theta]), \tag{3.3.6}$$

where $\bar{y}_j$ is the fundamental level of productivity in location j, y is some function of the average skill of workers employed in j, $\mathbb{E}_{|j}[\theta]$.

3.3.1.4 Taxation and Public Policy

To redistribute across workers, the government implements two types of policy instruments: non-linear income taxation and means-tested housing assistance.

Income taxation. The government levies a non-linear income tax over total income. A worker of type Θ will retain:

$$x(\Theta) - T\big(x(\Theta)\big) \tag{3.3.7}$$

where $x(\Theta)$ is total income, which is comprised of labor income, as well as rental incomes from capital and land:

$$x(\Theta) = w(\theta)n(\Theta) + \sum_j \omega_j \Pi_j. \tag{3.3.8}$$

Here, $w(\theta)$ denotes worker's wage, $n(\Theta)$ his labor supply, ω_j the shares of land used in location j floorspace development that he owns, and Π_j the total land rents in location j.

By assumption, both lump-sum transfers and factor-specific linear taxes are ruled out, so that neither the Second Welfare Theorem nor linear taxation results à la Diamond and Mirrlees [1971a,b] apply.

Housing assistance. Government provides a housing assistance through three different means-tested programs: public housing (P), subsidized private housing (S), and vouchers (V). The first one, public housing, is a rent subsidy that is only available to workers renting residential floorspace developed on government-owned land. The other two, subsidized private housing and vouchers, are rent subsidies which are only available to households renting housing on the private market. The three different programs are modeled as ad valorem subsidies. Finally, we assume that subsidies from different programs can't be combined, so that a household only receives the most advantageous subsidy he is entitled to.

The total rent subsidy received for program $\pi \in \{P, S, V\}$, $\tau^{\pi}_{i,s}(x_{ij}(\Theta), \xi)$, may depend workers' income, $x_{ij}(\Theta)$, some observable characteristics such as family status, household size or age, ξ, neighborhood of residence, i, and sector $s \in \{P, G\}$. Specifically, we assume that $\tau^{P}_{i,P} \equiv 0$ and $\tau^{S}_{i,G} = \tau^{V}_{i,G} \equiv 0$ so that subsidized private housing and vouchers are only available to tenants on the private market. Voucher subsidies do not depend on residence, so $\tau^{V}_{i,G} \equiv \tau^{V}_{G}$. When a household is not eligible for a program, we simply write that he does not receive any subsidy, that is, $\tau^{\pi}_{i,s}(x_{ij}(\Theta), \xi) = 0$.

3.3.1.5 Closing the Model

Having described the model's primitives, we specify ownership and market structure in order to close the model. Ownership of fixed factors is split between the government and private households. Wages and prices are determined competitively. The city is assumed to be closed.

Ownership. Each type-Θ worker owns a share ω_j of privately-owned land used to produce floorspace in location j. Shares add up to 1 so:

$$\int \omega_j dF(\Theta) = 1, \quad j \in \mathcal{I}. \tag{3.3.9}$$

Land used to produce public housing is entirely owned by the government.

Market structure. Production, labor and housing markets are perfectly competitive. All agents are price-takers.

Closed city. The mass of workers of each type Θ is equal to $f(\Theta)$.

3.3.2 Equilibrium

This subsection lays out the equilibrium behaviors of the different agents in order to define an equilibrium in this model.

3.3.2.1 Workers

Workers choose their residence, workplace, labor supply and consumption of goods and housing to maximize their utility.

Good and housing demands. Having chosen a residence i, a workplace j and a labor supply level $n_{ij}(\Theta)$, a worker chooses optimally his consumption of tradable goods, c, and of residential floorspace, h, given the local price index of consumption goods, P_i, and the local rent he has to pay, $R_i(\Theta)$, defined as:

$$R_{ij}(\Theta) \equiv \min \left\{ \Big(1 - \tau_{i,s}^{\pi}(x_{ij}(\Theta), \xi)\Big) R_{i,s}^{R} \;\Big|\; s \in \{P, G\}, \pi \in \{P, S, V\} \right\}, \tag{3.3.10}$$

where $R_{i,s}$ is the prevailing rent in the housing market of sector $s \in \{P, G\}$.

The worker's budget constraint is thus given by:

$$P_i c + R_{ij}(\Theta) h \leq x_{ij}(\Theta) - T\big(x_{ij}(\Theta)\big). \tag{3.3.11}$$

Conditional on residence i and income $x_{ij}(\Theta)$, workers pick their good and housing consumptions, $c_{ij}(\Theta)$ and $h_{ij}(\Theta)$, by solving:

$$V_{ij}(\theta) \equiv \max_{c,h} \; v(c, h), \tag{3.3.12}$$

subject to (3.3.11).

Labor supply. Conditional on their residence, i, and workplace, j, workers choose their labor supply, $n_{ij}(\Theta)$, to maximize their utility:

$$U_{ij}(\Theta) \equiv \max_{n_{ij}(\Theta)} \; U(V_{ij}(\Theta), n_{ij}(\Theta); \Theta). \tag{3.3.13}$$

Choice of residence and workplace. Workers pick their residence $i(\Theta)$ and workplace $j(\Theta)$ to maximize their utility given :

$$\max_{i,j \in \mathcal{I}} \; U_{ij}(\Theta). \tag{3.3.14}$$

3.3.2.2 Producers

Producers assemble labor inputs, $N_j^Y = \big(N_j^Y(\theta)\big)$, intermediate goods, M_j^Y, and commercial floorspace, H_j^Y into a quantity Y_j of goods, taking prices on input and output markets as given.

To produce a quantity Y_j of goods, they solve the following cost minimization problem:

$$\min_{N_j^Y, M_j^Y, H_j^Y} \; \int w_j(\theta) N_j^Y(\theta) d\theta + P_j M_j^Y + R_i^C H_j^Y, \tag{3.3.15}$$

subject to:

$$Y_j\big(N_j^Y, M_j^Y, H_j^Y\big) \geq Y_j. \tag{3.3.16}$$

3.3.2.3 Developers

Developers of both sectors $s \in \{P, G\}$ use labor inputs $N^H_{j,s} = \big(N^H_{j,s}(\theta)\big)$ and intermediate goods $M^H_{j,s}$ to produce floorspace, and sell it to workers and producers. They take prices of inputs and outputs as given.

Input demands. Developers minimize the cost of inputs:

$$\min_{N^H_{j,s},\, M^H_{j,s}} \int w_{j,s}(\theta) N^H_{j,s}(\theta) d\theta + P_j M^H_{j,s}, \tag{3.3.17}$$

subject to:

$$H_{j,s}\big(N^H_{j,s}, M^H_{j,s}\big) \geq H_{j,s}. \tag{3.3.18}$$

Floorspace use allocation. Floorspace built on privately-owned land may be used for either residential or commercial floorspace. Developers choose the fraction λ_i if floorspace allocated to residential use to maximize profits. They allocate floorspace to its most profitable use, so that:

$$\begin{cases} \lambda_i \in (0,1) & \Rightarrow \quad R^R_{i,P} = R^C_{i,P}, \\ R^R_{i,P} > R^Y_{i,P} & \Rightarrow \quad \lambda_i = 1, \\ R^W_{i,P} < R^Y_{i,P} & \Rightarrow \quad \lambda_i = 0. \end{cases} \tag{3.3.19}$$

3.3.2.4 Definition of a Decentralized Equilibrium

Having characterized the equilibrium behavior of each agent, we define the decentralized equilibrium of this model. We describe the equilibrium conditions on the labor, good and housing markets, before giving the formal definition of a decentralized equilibrium.

Labor market equilibrium. Labor is used by both producers and developers. In equilibrium, the total labor inputs each skill θ used by producers and developers in each workplace j has to be equal to sum over i of labor supplied in j by i residents:

$$\sum_i n_{ij}(\theta) L_{ij}(\theta) = N^Y_j(\theta) + N^H_{j,P}(\theta) + N^H_{j,G}(\theta). \tag{3.3.20}$$

Good market equilibrium. Goods produced in each location j are used for consumption by workers and as intermediates by producers and developers. Geography is captured by iceberg trade frictions $\chi_{jl} \geq 1$. That is, producers in location j must ship $\chi_{jl} Q_{jl}$ units to location i for Q_{jl} units to arrive. The feasibility constraint for tradable goods implies:

$$Y_j \geq \sum_l \chi_{jl} Q_{jl}, \tag{3.3.21}$$

where Y_j is the production in location j and Q_{jl} is the sum of goods used by workers, producers and developers in location l.

Tradable goods are differentiated by origin and aggregated through a homothetic and concave aggregator Q. Feasibility constraint for traded good good imposes:

$$Q(Q_{1i}, \ldots, Q_{Ii}) = M_i^Y + M_i^H + \sum_j \int c_{ij}(\theta) L_{ij}(\theta) d\theta, \tag{3.3.22}$$

for each location i.

This flexible functional form covers in particular perfect substitution as in Rosen [1979] and Roback [1982]'s seminal model and constant elasticity of substitution (CES) à la Armington [1969], as in standard economic geography models.

Floorspace market equilibrium. Both private and public floorspace markets are in equilibrium. We also assume that developers cannot price discriminate across workers, so that the residential rent $R_{i,s}^R$ is the same for all units of a same sector.

Floorspace produced by developers is divided between residential and commercial floorspace in the private sector, so that:

$$H_{i,P}\big(N_{i,P}^H, M_{i,P}^H\big) = H_i^Y + \int h_{ij,P}(\Theta) L_{ij,P}^R(\Theta) d\Theta. \tag{3.3.23}$$

Floorspace built on public land is reserved for residential use, so:

$$H_{i,G}\big(N_{i,G}^H, M_{i,G}^H\big) = \int h_{ij,G}(\theta) L_{ij,G}^R(\Theta) d\Theta. \tag{3.3.24}$$

Definition of a decentralized equilibrium. Before defining a decentralized equilibrium, we introduce the convenient definition of an allocation.

Definition 3.1 (Allocation). An *allocation*, $\mathcal{A}$, is the specification of a partition of workers, $(i(\Theta), j(\Theta))_\Theta$, associated per capita consumptions of tradable goods and housing, $(c_{ij}(\Theta), h_{ij}(\Theta))_{\Theta,i,j\in\mathcal{I}}$, labor inputs used in the production and development sectors, $(N_j^Y(\Theta))_{\Theta,j\in\mathcal{I}}$ and $(N_{j,s}^H(\Theta))_{\Theta,j\in\mathcal{I},s\in\{P,G\}}$, intermediate goods used in the production and development sectors, $(M_j^Y)_{j\in\mathcal{I}}$ and $(M_{j,s}^H)_{j\in\mathcal{I},s\in\{P,G\}}$, floorspace used in the production of tradable goods, $(H_j^Y)_{j\in\mathcal{I}}$, goods produced and floorspace developed, $(Y_j)_{j\in\mathcal{I}}$ and $(H_{j,s})_{j\in\mathcal{I},s\in\{P,G\}}$.

Having determined the equilibrium behavior of each agent individually, we now summarize the above conditions to define a decentralized equilibrium.

Definition 3.2 (Decentralized Equilibrium). A *decentralized equilibrium* is an allocation $\mathcal{A}$ such that:

(i) Workers consume tradable goods and housing to maximize their utility subject to their budget constraint, conditions (3.3.12), (3.3.11), and choose their residence optimally, condition (3.3.14);

(ii) Producers choose labor inputs and intermediate goods optimally, conditions (3.3.15) and (3.3.16);

(iii) Developers choose labor inputs and intermediate goods optimally, conditions (3.3.17) and (3.3.18), and allocate optimally floorspace between residential and commercial use, (3.3.19);

(iv) Goods are aggregated optimally, condition (3.3.22);

(v) Labor, good and housing markets clear, conditions (3.3.20), (3.3.21), (3.3.23) and (3.3.24).

3.4 Model Estimation: Leveraging Public Housing Demolitions

We lay out the main steps to estimate structurally the model developed in section 3.3. We map the key parameters of this model to reduced-form counterparts of structural identities. We leverage public housing demolitions as a quasi-experimental shock to estimate them.

3.4.1 Quantitative Implementation

This subsection exposes the preliminary steps necessary to take the model developed in section 3.3 to the data.

We specify the functional forms and describe the full structural estimation procedure.

3.4.1.1 Functional Forms

We start by defining the three key parameters of the model which will be estimated in section 3.4.2.2: ζ, the elasticity of floorspace supply, σ, that captures the residential mobility of workers, and μ^A, which captures the strength of residential spillovers. We then specify the utility and production which are all Cobb-Douglas and will be calibrated in section 3.4.2.1.

Key parameters. We assume that the floorspace production function is:

$$H_{j,s} = \bar{h}_{j,s}(M^H_{j,s})^{\frac{\zeta}{1+\zeta}}, \tag{3.4.1}$$

with $\bar{h}_{j,s}$ the fundamental floorspace supply and ζ the floorspace supply elasticity.

Idiosyncratic residence draws, $(\varepsilon_{g,i})$ follow a Fréchet distribution with dispersion parameter σ and are independent of the other components of Θ. This parameter captures the strength of idiosyncratic preferences for locations and is inversely proportional to residential mobility.

The amenity spillovers are:

$$a(\Theta^R_i) = (\bar{\theta}_i)^{\mu^A}, \tag{3.4.2}$$

with $\bar{\theta}_i$ the mean skill level in neighborhood i and μ^A the strength of residential spillovers.

Utility, production and matching functions. We assume that workers have Stone-Geary preferences over tradable goods and housing:

$$u(c, h) = c^{\gamma}(h - \bar{h})^{1-\gamma}, \tag{3.4.3}$$

with $\gamma \in (0, 1)$.

Producers use a Cobb-Douglas technology given by:

$$Y_j = y_j(\bar{N}_j^Y)^{\alpha^Y}(M_j^Y)^{\beta^Y}, \tag{3.4.4}$$

where composite labor $\bar{N}_j^Y = \int_\theta n_{g,j}(\theta)L_j^E(\theta)d\theta$. The labor share is α^Y.

Finally, the skill draws θ follow a Pareto distribution with shape parameter ρ:

$$G(\theta) = \frac{1}{\theta^\rho}, \quad \theta \geq 1, \tag{3.4.5}$$

while the ω_j's are uniform over $[0, 1]$ and χ is non-random.

3.4.1.2 Model Inversion

Fundamental location characteristics such as productivities, amenities and housing supplies cannot be directly observed in the data. While the presence of local amenity and productivity spillovers allows for the possibility of multiple equilibria, we are able to recover unique values of intrinsic components of productivities, amenities and housing supplies that rationalize the observed data as a model equilibrium.

This inversion process follows closely the steps outlined in Ahlfeldt et al. [2015b] and Tsivanidis [2019a]. We combine those observed data with the model structure to solve for the endogenous variables and back out the unobservable amenities, productivities and housing supplies.

Proposition 3.1 (Model Inversion).

1. Given data on residence by type, $(L_i^R(\Theta))$, total employment by workplace, $(\bar{L}_j^E)$, in addition to model parameters, there exists a unique vector of labor input prices, $(w_j(\theta))$ that rationalizes the observed data as an equilibrium of the model.
2. Given model parameters, data on residence by type, $(L_i^R(\Theta))$ and rent levels $(R_{i,s})$, and a vectors of labor input prices, $(w_j(\theta))$, there exist unique vectors of unobservable amenities $(a_i(\Theta))$ (to scale), productivities (y_j) and housing supplies $(h_{i,s})$ that rationalize the observed data as an equilibrium of the model.

3.4.2 Structural Estimation

We now implement the method exposed in 3.4.1 to estimate structurally the model. We start with calibrated parameters, before switching to the estimation of the key structural parameters.

3.4.2.1 Calibrated Parameters

Parameters $\{\gamma, \rho, \alpha^Y, \beta^Y\}$ are calibrated either directly from the data or to existing values from the literature.

We set the share of housing expenditure for workers to the commonly used value of $1-\gamma = 0.3$ for the United States. The shape of the Pareto distribution of skill draws θ is estimated from wage data using a Hill's estimator, and is approximately equal to 2.

The shares of labor and equipment correspond to their estimates in Greenwood et al. [1997], renormalized to exclude structures which are absent from the model.

3.4.2.2 Estimation of Key Structural Parameters

Housing supply elasticity ζ. Demolition of public housing through the Hope VI program can be interpreted as a demand shock for private housing developers: It increased the demand for private housing, which led to house price increases in affected areas. Our estimation of housing supply elasticity, ζ, is therefore similar to suggested evidence presented in section 3.2.3. We regress the change in the house price index between the early 1990s and 2010 in census tract i on a demolition exposure index, defined as as before:

$$\text{Demolition exposure}_i = \frac{1}{H_i^{1990}} \sum_j \frac{1}{\exp d_{ij}} \times \text{Demolished units}_i$$

where H_i^{1990} is the baseline number of housing units in the tract and d_{ij} is the distance between census tracts i and j. That is, the index captures the number of demolished units weighted by their distance to census tract i. After controlling for several baseline characteristics, the relationship between house price changes and demolition exposure is positive and very significant, as shown in figure 3.2.2.

We use demolition exposure as an instrument to estimate the housing supply elasticity, ζ. Results are reported in table 3.4.1. We find that $\zeta \simeq 0.44$.

Residential mobility σ and local amenity spillovers μ^A. To estimate the demand side, we use two different instruments that affect both the supply of housing and the local skill composition. Our first instrument relies on LIHTC eligibility rules. We exploit the fact that HUD defines Qualifying Census Tracts (QCTs) for the LIHTC programs with a discontinuous rule. Specifically, the census tracts must fulfill on of these two requirements: (1) tract median income is below 60% of the area median income, or (2) poverty rate is above 25%. We follow Davis, Gregory and Hartley (2021), and

	(1) OLS	(2) 1st stage	(3) 2nd stage	(4) 1st stage	(5) 2nd stage
Δ ln R (ζ)	0.08***		0.51***		0.44***
	(0.01)		(0.09)		(0.06)
Demo exposure		0.11***		0.16***	
		(0.01)		(0.01)	
F-stat		93.49		189.7	
R-squared	0.17	0.27		0.37	0.03
Observations	2,833	2,833	2,833	2,833	2,833
County FE	Y	Y	Y	Y	Y
Controls	Y	N	N	Y	Y

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

Table 3.4.1: Estimation of housing supply elasticity ζ

combine these two rules into a single index of QCT eligibility E_i, where $E_i = \max\{\text{poverty rate}_i - 0.25, 0.6 - \text{median income index}_i\}$, where the median income index is the ratio of the median income in the tract over the area median income. Our instrument is defined as $1_{E_i>0} + E_i + 1_{E_i>0}E_i$ as the instrument. In our context, we think that this instrument changes both prices [Diamond and McQuade, 2019] and neighborhood composition by shifting housing supply.

Our second instrument is distance to rapid transit lines. Baum-Snow and Han [2020] estimate the impact of the distance to new rail transit lines (mostly built in the 1980s) on their usage and housing values. They document that decreasing distance to transit from 3 to 1km away increases rents by $19 per month and housing values by around $5,000. Kahn [2007] shows that increased access to rapid transit lines increases gentrification, which we can use as a shifter of the neighborhood composition. We use the change in the distance (in km) to a rapid transit line from 1980 to 2004–we include the period before 1990 because it is where most of the variation is happening).

Results are reported in table 3.4.2. We find that $\sigma \simeq 0.25$ and $\mu^A \simeq 0.35$.

	(1) OLS	(2) IV
Δ ln R (ε^s)	0.08*	-0.25***
	(0.04)	(0.05)
Δ ln(Low educ)	-0.06***	-0.35***
	(0.02)	(0.13)
R-squared	0.12	
Observations	2,815	2,207
County FE	Y	Y
Controls	N	Y

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

Table 3.4.2: Estimation of residential mobility σ and local amenity spillovers μ^A

3.5 Conclusion

We developed a model to assess the redistributive implications of housing assistance. Housing assistance programs feature a trade-off between indirect pecuniary redistribution and direct amenity effects. Public housing may improve redistribution efficiency, but at the expense of lower local amenities for low-income households. We estimated the structural parameters of our model leveraging demolition of public housing. In future work, we intend to measure the incidence of the change in housing assistance programs that occurred in the U.S. over the last decades.

Bibliography

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf (2015a), "The Economics of Density: Evidence From the Berlin Wall," *Econometrica*, Vol. 83, No. 6, pp. 2127–2189.

—, —, —, and — (2015b), "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, Vol. 83, No. 6, pp. 2127–2189.

Armington, Paul S. (1969), "A Theory of Demand for Products Distinguished by Place of Production," *Staff Papers (International Monetary Fund)*, Vol. 16, No. 1, pp. 159–178.

Baum-Snow, Nathaniel and Lu Han (2020), "The Microgeography of Housing Supply*," *Working Paper*.

—, Matthew E. Kahn, and Richard Voith (2005), "Effects of Urban Rail Transit Expansions: Evidence from Sixteen Cities, 1970-2000 [with Comment]," *Brookings-Wharton Papers on Urban Affairs*, pp. 147–206.

Blanco, Hector (2021), "Pecuniary Effects of Public Housing Demolitions: Evidence from Chicago," *Working Paper*.

Coate, Stephen, Stephen Johnson, and Richard Zeckhauser (1994), "Pecuniary Redistribution through In-Kind Programs," *Journal of Public Economics*, Vol. 55, No. 1, pp. 19 – 40.

Collinson, Robert, Ingrid Gould Ellen, and Jens Ludwig (2019), "Reforming Housing Assistance," *The ANNALS of the American Academy of Political and Social Science*, Vol. 686, No. 1, pp. 250–285.

Costinot, Arnaud and Iván Werning (2018), "Robots, Trade, and Luddism: A Sufficient Statistic Approach to Optimal Technology Regulation," NBER Working Papers 25103, National Bureau of Economic Research.

Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst (2019), "Income Growth and the Distributional Effects of Urban Spatial Sorting," Working Paper 26142, National Bureau of Economic Research.

Davis, Morris, Jesse Gregory, and Daniel Hartley (2019a), "The Long-Run Effects of Low-Income Housing on Neighborhood Composition," *Working Paper*.

—, —, —, and Kegon Tan (2019b), "Neighborhood Effects and Housing Vouchers," *Working Paper*.

Desmond, Matthew (2016), *Evicted: Poverty and Profit in the American City*, New York: Crown.

Diamond, Peter A. and James A. Mirrlees (1971a), "Optimal Taxation and Public Production I: Production Efficiency," *The American Economic Review*, Vol. 61, No. 1, pp. 8–27.

— and — (1971b), "Optimal Taxation and Public Production II: Tax Rules," *The American Economic Review*, Vol. 61, No. 3, pp. 261–278.

Diamond, Rebecca and Tim McQuade (2019), "Who Wants Affordable Housing in Their Backyard? An Equilibrium Analysis of Low-Income Property Development," *Journal of Political Economy*, Vol. 127, No. 3, pp. 1063–1117.

Gaubert, Cecile, Patrick M. Kline, and Danny Yagan (2021), "Place-Based Redistribution," Working Paper 28337, National Bureau of Economic Research.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell (1997), "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review*, Vol. 87, No. 3, pp. 342–62.

Kahn, Matthew E. (2007), "Gentrification Trends in New Transit-Oriented Communities: Evidence from 14 Cities That Expanded and Built Rail Transit Systems," *Real Estate Economics*, Vol. 35, No. 2, pp. 155–182.

Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles (2020), "IPUMS National Historical Geographic Information System," Version 15.0.

Mirrlees, J. A. (1971), "An Exploration in the Theory of Optimum Income Taxation," *The Review of Economic Studies*, Vol. 38, No. 2, pp. 175–208.

Naito, Hisahiro (1999), "Re-Examination of Uniform Commodity Taxes Under a Non-Linear Income Tax System and Its Implication for Production Efficiency," *Journal of Public Economics*, Vol. 71, No. 2, pp. 165–188.

Nichols, Albert L. and Richard J. Zeckhauser (1982), "Targeting Transfers through Restrictions on Recipients," *American Economic Review, Papers and Proceedings*, Vol. 72, No. 2, pp. 372–377.

Roback, Jennifer (1982), "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, Vol. 90, No. 6, pp. 1257–1278.

Rosen, Kenneth T. (1979), "A Regional Model of Multifamily Housing Starts," *Real Estate Economics*, Vol. 7, No. 1, pp. 63–76.

Susin, Scott (2002), "Rent vouchers and the price of low-income housing," *Journal of Public Economics*, Vol. 83, No. 1, pp. 109 – 152.

Tach, Laura and Allison Dwyer Emory (2017), "Public Housing Redevelopment, Neighborhood Change, and the Restructuring of Urban Inequality," *American Journal of Sociology*, Vol. 123, No. 3, pp. 686–739.

Tsivanidis, Nick (2019a), "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá's TransMilenio," Job Market Paper.

— (2019b), "Evaluating the Impact of Urban Trasit Infrastructure: Evidence from Bogota's TransMilenio," *Working Paper*.

Vale, Lawrence and Yonah Freemark (2012), "From Public Housing to Public-Private Housing," *Journal of the American Planning Association*, Vol. 78, pp. 379–402.

Zillow (2017), "ZTRAX: Zillow Transaction and Assessor Dataset, 2017-Q4."

Appendix

3.A Data Appendix

To obtain a comprehensive picture of the number of public housing units demolished in every census tract, we match developments demolished under the HOPE VI program to their geolocated addresses in the 1996 HUD-951 form public file. For each city included in the sample, we follow these steps:

1. **Match developments in HOPE VI with 1996 HUD-951 form**. HOPE VI administrative data only provides the name and, in some cases, the HUD project number of the public housing development. For this reason, we use the 1996 HUD-951 form public file to associate them to geolocated addresses, which allows us to assign each demolished unit to a particular census tract.
    - For "revitalization" grants, HOPE VI administrative data provides the HUD project number of the development, which is also indicated in HUD-951 forms. Thus, we match on this number.
    - For "demolition only" grants, we only obtained a list of development names. We proceed as follows. First, we manually discard those developments already counted in a "revitalization" grant. Second, we use an algorithm that matches development names in the "demolitions only" grant list with similar development names in the 1996 HUD-951 form public file, within the same city. To do this, we use the package "matchit" in Stata. Finally, we manually revise all of the matches.
    - In the case of Chicago, we include the list of non-HOPE VI demolished public housing addresses, which was provided by the Chicago Housing Authority through a FOIA request.
    - We merge the three datasets above to obtain the full list of demolished public housing addresses.
2. **Compute the number of demolished units per address**. Not all of the developments were fully demolished, thus, we use the following method to compute the number of demolished units per address:

- First, count as demolished all units that appear in the 1996 HUD-951 form public file[9].
- Second, we manually check the developments that were partially demolished and change the number of demolished units to reflect the actual number under HOPE VI.

[9] In this file, sometimes the total number of units in geolocated addresses for a development is less than the actual number of units. We solve this by assigning every geolocated address a proportional number of the missing geolocated units until obtaining the total number of units in the development.