ESSAYS ON URBAN HOUSING AND GENTRIFICATION.

by

CORY BRIAR

A DISSERTATION

Presented to the Department of Economics
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctorate of Philosophy

June 2022

DISSERTATION APPROVAL PAGE

Student: Cory Briar

Title: Essays on Urban Housing and Gentrification.

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctorate of Philosophy degree in the Department of Economics by:

	Chair
Mark Colas	Co-Chair
Bruce Blonigen	Co-Chair
Keaton Miller	Core Member
James Buckley	Institutional Representative

and

Krista Chronister	Vice Provost for Graduate Studies

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded June 2022

DISSERTATION ABSTRACT

Cory Briar

Doctorate of Philosophy

Department of Economics

June 2022

Title: Essays on Urban Housing and Gentrification.

Cities are the locus of the overwhelming majority of economic activity in modern societies; understanding the interactions of urban trends and policies is crucial to achieving desired economic outcomes. This dissertations provides three papers characterizing the mechanisms of change in US cities over the last 30 years. After an introduction in chapter one, chapter two investigates the role of firm-sorting in the process of gentrification in Seattle, Washington between the years of 1990 and 2018, and finds that holding firm distributions fixed at 1990 levels leads to a significant attenuation of the process. Chapter three repeats the analysis of chapter two for the city of Portland, Oregon and compares results with those in Seattle. Chapter three investigates the connection between turnover in the housing market and modern rent control policy in San Francisco, California, and demonstrates the factors at play that lead to the deletion of older structures under said policy from the market.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank the members of my committee for their guidance and mentorship in my time spent writing this dissertation. Special thanks to Jenni Putz, Youssef Ait Benasser, and Chandler Lester for constant advice, moral support, and commiseration. I would also like to thank the Public Works Department of San Francisco for many clear answers to poorly worded questions. Thanks to T.J. Olney, who inspired me to pursue a doctorate in the first place, rest in peace. Sharon Kaplan, thank you as well for constantly shepherding myself and everyone else in the department towards success. And finally, thanks to Daye Thomas, without whom I doubt I would have survived this past year.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

Cities are the locus of roughly 80% of the world's GDP, with an even higher number in more advanced economies (World Bank, 2020); those numbers have been increasing over time, with rising urbanization worldwide. Thus, to understand processes of change in an economy, one has to study the changes that take place in cities. However, the factors that shape cities are irreducibly complex and often overlapping in their causation. Consider a single household deciding where to live in a city. Their decision must factor in housing costs, housing type, commuting distance to work (and where to work, and what level of education to pursue), amenity factors like crime rates, neighbors, and nearby parks, the availability of consumption opportunities, and on and on. One's education and work decisions are tied to one's income, and thereby also related to one's decisions across all of the above factors. Likewise, locating near consumption opportunities may impact the supply of those goods and services as more firms move in to be near clientele. In response to these changes, landowners may decide to raise or lower the price of housing, or withdraw their goods from the housing market entirely, or even renovate, all of which in turn impacts who chooses to live nearby. Causality runs circular, and every action by each economic agent has ripple effects that interact with those of others, leading to unpredictable outcomes. How can we, as economists, policy makers, and residents of these complex dynamic systems, make sense of cities? The answer to that question, like so many in economics, is through the creation of useful models.

1.1 The Modeling of Cities

Economic models of cities date back to the inception of the discipline, but the modern urban literature was undoubtedly kicked off by two seminal papers: Rosen (1979) and Roback (1982). These models considered the locational choices of individual agents. All agents in these models were given utility functions that ranked happiness across locations by incorporating factors such as local rents and wages, and were allowed to choose where to live in response to these stimuli. Ultimately, these models showed that if there are no costs associated with moving, utility in all locations must be equal; were it not, some individuals would relocate to areas of higher utility until rents increase or amenities degrade to restore equality. The next generation of these models began to examine how individuals might respond to the same set of stimuli if each had their own specific preferences across locations. Each agent in these models was given an idiosyncratic taste shock for each location, frequently draw from a Gumbel or Type-I Extreme Value distribution. These shocks meant that there was a simple, closed form probability that a given individual would choose to live in a given area, depending on agglomerative and dissipative forces. These models became the workhorse for much of modern structural urban economics, and those within this dissertation.

The models I develop in my dissertation are designed to characterize interactions between different types of economic actors, such as households, firms, and landowners, that all exist in urban environments. As noted above, the interactions of different players' decisions in common space are ultimately how cities evolve over time; these models are designed to study transition via comparative statics. Two of my models adapt frameworks from the trade literature (Dixit & Stiglitz, 1977) to investigate how households and firms choose to locate near each other and the resulting patterns of development in cities. The final model, simple but novel, seeks to understand how landowners decide if and when to pull their housing from the market in response to sorting households. The two-sided decision making in these models provides a more nuanced and complete depiction of how cities respond to shifts in policy and the fundamental rules of economic interaction therein that simple one-sided sorting models would miss, and allow me to answer more complicated questions about the nature of change in cities.

1.2 Gentrification and Housing Turnover

This dissertation specifically focuses on changes in the residential and commercial composition of neighborhoods and changes in the housing market. The last 50 years have borne dramatic transformations in US cities, including the increasing prevalence of the process known as gentrification. Gentrification occurs in neighborhoods of cities, typically centrally located, which have been historically inhabited by disadvantaged groups in society: those with low income, low educational attainment, of old age, and frequently of color. Older housing which was rundown but cheap featured prominently in these areas. Starting in the 1970's and continuing through the present day, many of these communities have undergone dramatic and rapid transformation as populations of young, high-income, college educate households suddenly inundated downtowns. With these affluent populations came new construction, renovation of old housing, new goods and services from emerging establishments such as bars and gyms, reductions in crime, and so on; gentrification has been alternatively called "urban revival" for some of these benefits. Such rapid change, however, has not been uniformly welcomed by the original residents of gentrifying areas, who cite increasing rents and the erosion of local community as negative outcomes of the process. Others appreciate some of the positive changes, such as reductions in crime and new consumption opportunities. Regardless the normative impact, gentrification remains one of the most prolific forces at work in cities across the world. For instance, Macaig (2015) estimates that a remarkable 29% of Census tracts in the US underwent gentrification between the years of 1990 and 2013.

Gentrification is fundamentally tied to the supply and state of urban housing. One theory for the origin of the process is the replacement of rundown, aging housing, as set forth in Myers and Pitkin (2009). As central city housing, which in older US cities could easily been constructed as far back as the late 1800's, begins to deteriorate, developers replace it with more expensive modern housing. The increased rent of these new structures is less affordable to long-time residents, but attractive to the young, affluent households that come to gentrify these neighborhoods. This is why policies such as rent control, which encourage such housing turnover, can lead to further gentrification in downtown areas despite ostensibly being meant to slow it (Diamond, McQuade, & Qian, 2019a, 2019b). Additionally, as affluent households begin to sort more centrally, they may in fact demand these newer types of housing, further increasing turnover. A full consideration of gentrification therefore necessarily nests an understanding of the evolution of urban housing, and vice-versa.

1.3 Dissertation Outline

The second chapter in this dissertation develops a model of household and firm co-sorting in cities in order to assess the contribution of firms to gentrification. Gentrifying areas of cities are marked by many changes, including changes in amenity values to households resulting from reductions in crime, improvements in school quality, and shifting distributions of firms. These changes are endogenous and self-reinforcing, in that they occur as a particular neighborhood becomes more affluent and also encourage further in-sorting of affluent populations. Past work has lumped all of these amenity changes into simple reduced-form terms which, while effective for capturing these endogenous dynamics, make it impossible to separately tease out the contribution of firm-based and non-firm based amenity shifts to neighborhood change. Separating these effects could be desirable to policy makers hoping to craft policies that encourage economic growth without some of the negative affects associated with gentrification. The model I work with is similar to other intra-city sorting models, treating household neighborhood location as a discrete choice problem based on housing costs, wages, and amenity values. However, I also structurally model differentiated firms' sorting choices; in particular, firms with goods that must be consumed in person (non-tradable) face incentives to locate near households with larger incomes to increase profit. Households, in turn, are incentivized to locate near these firms to minimize time spent traveling to consume their goods. Ultimately, households and firms must sort in relation to each other. I calibrate the model using values from around the literature and the real-world geography of Seattle, Washington. My approach allows me to first decompose the contribution to gentrification of changes in parameter values and neighborhood amenities between 1990 and 2018 that are *orthogonal* to household and firm sorting. Secondly, to illustrate the impact of firm sorting on the gentrification of the city, I fix neighborhood firm distributions to 1990 levels and allow households to sort until a new equilibrium is reached. By common measures, gentrification declines significantly on both the intensive and extensive margin across zip codes under this restriction. These results demonstrate the importance of accounting for firm sorting when seeking to understand changing patterns of development in cities.

Chapter three returns to the model and methodology of the second chapter, adapting the model and calibration to Seattle's twin city, Portland, Oregon. These two cities are alike in climate, culture, and location, and routinely joust for position at the top of "most gentrifying cities in the US" lists (Macaig, 2015). The model performs qualitatively quite similarly to chapter two, producing sorting patterns of households in Portland that mirror equilibria explored in Seattle. This chapter also seeks to compare the results of Seattle and Portland under this simulation, and notes that there are similarities between the two cities under the no-firm-sorting counterfactual. For instance, sans firm sorting, certain neighborhoods that were relatively unremarkable under the observed equilibrium become hotbeds of gentrification. I find that holding firm sorting constant again at 1990 levels has a significantly smaller impact on gentrification in Portland than Seattle, however. Ultimately these results suggest that gentrification is not a uniform process affecting all parts of a city similarly and predictably, but a complex process brought about by the changes in many variables, some firm based and some not.

Finally, chapter four investigates the impacts of San Francisco, California's rent control ordinance on turnover in the city's housing market. Rent control is ostensibly a policy meant to combat rising rents in expensive cities; however, a number of studies have found that such policies can actually *worsen* housing affordability and increase gentrification (Diamond et al., 2019a, 2019b; Kholodilin, Mense, & Michelsen, 2016). For instance, in San Francisco, the rent control law contains an exemption for structures built after 1979 and condominiums, incentivizing landowners to convert older rent controlled structures into new buildings or sell their units as privately owned in order to circumvent regulation. To investigate the impact of the rent control ordinance on such turnover in the housing market, I posit and calibrate a two-sided model of housing supply and demand. On the demand side, households sort across neighborhoods and housing types (including rent-controlled and condominiums) subject to rents and amenities; on the supply side, land owners are endowed with a rent controlled structure and must decide whether or not to convert it to privately owned real estate, subject to costs. Rents on continuing tenancies are constrained to only rise by a certain amount per year, but rents on ending tenancies are unconstrained, a policy known as vacancy decontrol. I access data from the city of San Francisco and US Census Bureau to estimate and calibrate the model and simulate several counterfactual scenarios. I find that when prices are allowed to adjust, the amount of newly available housing from the repeal of vacancy decontrol results in a decrease in housing costs and an increase in conversions. When I hold prices fixed, however, the increase in available housing is tempered by constant rents, leading to fewer conversions. These results partially explain the disparate results from the literature regarding the outcome of deregulation.

CHAPTER II

THE ROLE OF FIRM SORTING IN GENTRIFICATION; A SIMULATION OF SEATTLE, WA

Cities in the past 50 years in the United States have been the stage of massive and fundamental changes in neighborhood wealth, education and race composition. Between the end of WWII and the start of the information technology (IT) revolution, American central cities were depopulating, losing residents of high income to the suburbs. Many explanations have been proposed to explain this phenomenon, from so called "white flight" (Boustan, 2010) and racial tipping point dynamics (Card, Mas, & Rothstein, 2008; Schelling, 1971), to highway construction (Baum-Snow, 2007), to rising incomes among households (Margo, 1992), to the aging of the population and housing stock (Myers & Pitkin, 2009). Rents and home prices were increasing in distance from the city center. But by 1970, this trend began to reverse itself. By 2010, rent and house prices were *decreasing* with distance from the center, and the percentage of central city residents with high income and bachelor's degrees had appreciated greatly (Su, 2020); educated residents were migrating inwards, not outwards.

This process became known as gentrification, under which classes of affluent individuals relocate into formerly poor, run down neighborhoods within cities. With this new affluence comes increases in rent and amenities for these new residents that often push or force the neighborhoods' former residents out. While the precise definition of gentrification varies from study to study, the prevalence of the phenomenon cannot be doubted: using the definitions outlined by Freeman (2005), one of the most cited papers in the gentrification literature, Macaig (2015) estimates that 9% of all Census tracts in the US underwent gentrification in the

1990's; an additional 20% underwent gentrification between 2000 and 2013. In some cities like Portland, Washington D.C., Minneapolis, and Seattle, over 50% of tracts experienced gentrification.[1] Gentrification has been shown to cause various changes to affected neighborhoods, including displacement of former residents through rising rents (Atkinson, Wulff, Reynolds, & Spinney, 2011) shifting employment characteristics (Lester & Hartley, 2014), health impacts (Smith, Breakstone, Dean, & Thorpe, 2020), and changing racial composition (Huante, 2021).

Previous work has noted that gentrification is most likely a self-reinforcing process. As more affluent individuals begin to move in to an area, the amenities of that area begin to change to cater to the new residents. These changes include reductions in crime, improvement in school quality, and greater provision of public goods such as parks. Past work has also shown that affluent residents also provide markets for different types of non-tradable firms, such as gyms, bars, and restaurants (Handbury, 2019; Su, 2020). As areas within cities gentrify, these types of firms tend to pop up to cater to new residents; this in turn provides further incentive for affluent individuals to relocate in gentrifying areas, growing the market for these firms, and so on in a feedback cycle. While other papers have pursued reduced-form characterizations of this process, few structural attempts have been made to disambiguate firm resorting from other changing amenities.

[1] Specifically, Freeman (2005) uses the following criteria to classify a given neighborhood as having been gentrified over the period between two decennial censuses: "1) Be located in the central city at the beginning of the intercensal period; 2) Have a median income less than the median (40th percentile) for that metropolitan area at the beginning of the intercensal period; 3) Have a proportion of housing built within the past 20 years lower than the proportion found at the median (40th percentile) for the respective metropolitan area; 4) Have a percentage increase in educational attainment greater than the median increase in educational attainment for that metropolitan area. 5) Have an increase in real housing prices during the intercensal period" (Freeman, 2005).

In this paper, I build a model of household and firm co-sorting under gentrification. My approach allows me to simulate the locational decisions of both households and firms simultaneously within metropolitan areas; this is desirable, because it allows a researcher to separate out the impact of firms sorting from other endogenous amenity changes such as school quality, racial/income composition of neighborhoods, or reductions in crime. Specifically, I adapt the flexible monopolistic competition model developed by Dixit and Stiglitz (1977) to intra-city household and firm sorting. To access goods from non-tradable firms, households must pay iceberg transportation costs that are increasing in time spent traveling. These shopping costs incentivize households to locate near these firms, and these firms to locate near households to increase demand for their goods. By explicitly modeling the co-location decisions of firms and households, I am able to separately account for firm sorting's impact on the gentrification of the city, in addition to amenity changes that are orthogonal firm sorting. Simply but, my approach allows me to disambiguate between firm- and non-firm related amenity changes in the city.

Accounting separately for the sorting of firms versus other endogenous amenities is important in trying to understand the mechanisms of gentrification for a number of reasons. First, the resorting of firms can have welfare implications for long-time residents. The culture of a neighborhood can be thought of as partly couched in local businesses; these local firms are often crowded out when affluent residents move in and induce other firms to set up shop in the neighborhood. To the extent that the former businesses are locally owned, this firm resorting may also represent a loss of income to original residents. Failing to account for firm sorting in this manner may then lead us to underestimate the welfare impacts of gentrification. Secondly, policies aimed at economic development may have unintended consequences when it comes to firm sorting. For instance, past research has shown that economic opportunity zones, intended to provide assistance to struggling neighborhoods, may actually encourage influxes of businesses catering to gentrifiers from other parts of a city (Hoelzlein, 2019); rather than protecting local businesses, the policies may actually encourage further gentrification. Finally, understanding the role of firm sorting in gentrification could inform public policy aimed at easing displacement pressures for original residents and businesses in gentrifying areas.

After calibrating the model using a variety of data sources and parameter values taken from the literature, I simulate the model using the metropolitan statistical area (MSA) of Seattle-Tacoma-Bellevue, WA (Seattle). Seattle is one of the fastest gentrifying cities in the US, and hence the perfect laboratory to study the process. I first replicate the observed equilibria of 1990 and 2018 in the city in terms of the shares of differently educated households locating in each zip code of the city. This allows me to back out changes in neighborhood amenities that are orthoganal to firm sorting. I document the role that these changing amenities, as well as other parameters, have on gentrification in the city. Then, to investigate the role of firm sorting, I use the calibrated model to hold the distribution of firms within the city constant at 1990 levels. Under this exercise, gentrification decreases by an average of 28% across neighborhoods, as measured in the disproportionate rate of high-skill households moving into the city as compared to low skill households. This decrease in the appreciation of the skill ratio leads to 21% fewer neighborhoods being gentrified in the city. These results suggest that had firms not resorted to serve new clientele, gentrification in the city would have been substantially less intense and widespread in the city.

This paper is related to several literatures within spatial and urban economics, following from the spatial sorting literature sparked by Rosen (1979) and Roback (1982). First and foremost is the emerging mass of research on the process of gentrification. It is worth noting that within this literature, there is no consensus on an exact definition of gentrification; most generally it is defined as the process by which primarily white, young, college educated rich households displace less educated, lower income households in central cities. Explanations for this process usually center around changing downtown amenities for gentrifiers, including reductions in crime (D. Autor, Palmer, & Pathak, 2017; Ellen, Horn, & Reed, 2019), the desire to live near richer neighbors (Guerrieri, Hartley, & Hurst, 2013), and racial disparities in the valuation of downtown amenities (Baum-Snow & Hartley, 2019). Firm sorting is also often cited as a source of appreciating downtown amenities.

While the importance of firm sorting is widely recognized in gentrification, my paper is among the first to explicitly model firm sorting to that end. Many papers in the literature document firm sorting as a set of stylized facts but lump the process in with other endogenous amenities in reduced-form catchall terms during model estimation (Baum-Snow, 2007; Bayer, McMillan, & Rueben, 2004; Card et al., 2008; Couture & Handbury, 2017; Su, 2020). A limited number of papers have delved deeper into the ways in which firms sort to serve new populations of young, educated households. Handbury (2019) finds that product variety offered by firms in cities caters to income-specific preferences of local populations. Spatial resorting is particularly pronounced among non-tradable goods and services in central cities (Couture & Handbury, 2017). Generally, non-tradable goods and services represent a major advantage of density (Couture, Gaubert, Handbury, & Hurst, 2019). Highly educated individuals moving to city centers could also be a result of wanting to be closer to the goods and services offered by downtown firms (Couture, 2013); as hours worked increase for these individuals, not only would they wish to cut down on commuting time but also time spent traveling to consumption opportunities. Inspired by these insights, my model incorporates non-tradable firms which are incentivized by transportation costs to locate near customers. Households of different skills also differ in their preferences for these goods. In this way, my paper brings together two strands of literature: gentrification and firm sorting, as in work by Suárez Serrato and Zidar (2016) and Tsivanidis (2019). My approach also allows me to separately quantify the impacts of firm sorting and other amenities on gentrification; since I account both for firm sorting and other amenity changes, I can alternately hold each constant to determine their contribution. I find that both are independently significant factors effecting the extent of gentrification in the city.

My paper is most closely related to Hoelzlein (2019), which estimates a residential choice model with monopolistically competitive firms. Under counterfactual simulation, that paper finds that placed based "opportunity zones" result in twice as many high-skill households moving into gentrifying neighborhoods. Unlike that paper, however, I use homothetic preferences for households, which are more in line with standard economic assumptions. My results show that the qualitatively dramatic impacts of firm sorting still hold in a constant returns to scale environment.

The rest of this paper is organized as follows: Section 2.1 documents trends in the data which are indicative of the nature and magnitude we observe in modern cities, as well as how firms have changed their sorting behavior over the period of study; Section 2.2 outlines the model in the most general case and describes equilibrium; Section 2.3 details the data I use; in section 2.4, I discuss how I calibrate the model; Section 2.5 decomposes gentrification in Seattle by parameter and amenity changes and firm sorting; Section 2.6 concludes.

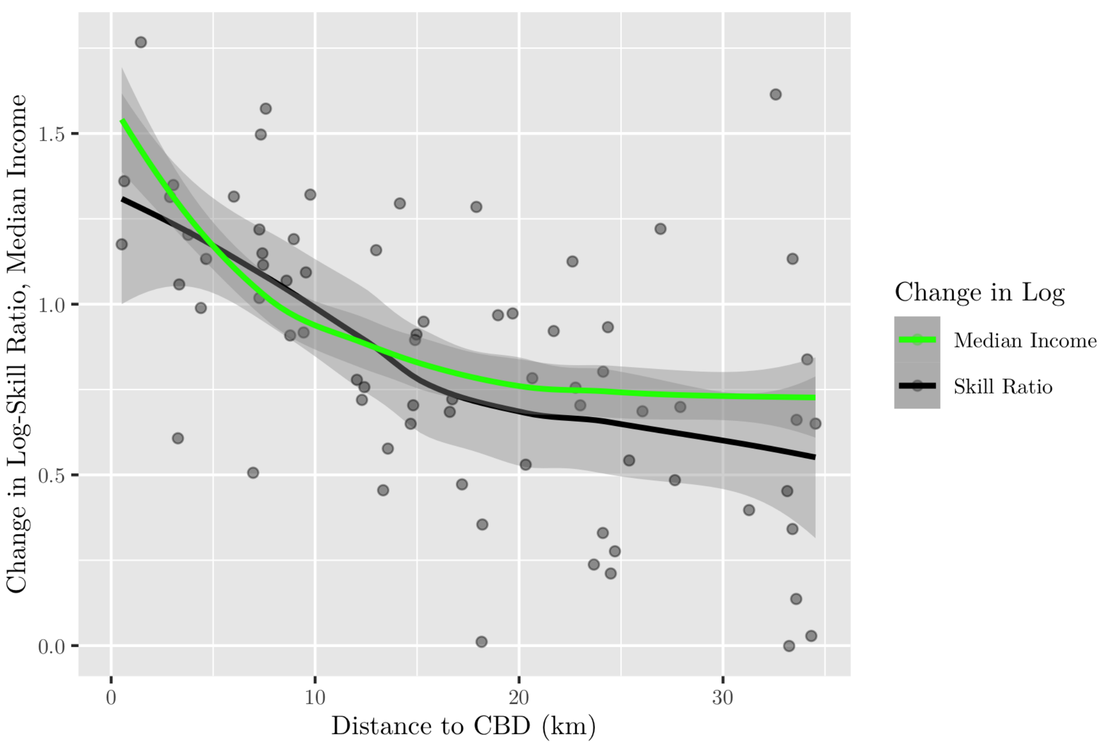
2.1 Descriptive Facts

2.1.1 Gentrification in Seattle. As mentioned previously, gentrification in modern cities seems to be driven by high-skilled households disproportionately sorting towards downtown neighborhoods. The black line in figure 1 plots observed changes in the natural logarithm of the ratio between high and low-skill individuals by distance to the city center between 1990 and 2018;[2] roughly speaking, this is the difference in the percent growth of high- and low-skill individuals in each neighborhood. The curve depicted is fit using a locally estimated scatterplot smoothing (LOESS) process. While the ratio of high- to low-skill households rose overall in Seattle between 1990 and 2018, it is clear that these households are significantly more likely to choose to live downtown; were this line flat it would be evidence of equal sorting of each type of household into every neighborhood. Figure 2, which maps changes in the log-skill ratio, confirms that the main locus of the appreciation is in and around the central downtown area (albeit with some outliers in the Issaquah/Sammamish area). Areas in the south of the city seem to have seen the least appreciation of the skill ratio. These areas are largely residential neighborhoods far from downtown amenities; many

[2] I define high-skill individuals as those with a Bachelor's degree or advanced degree.

also lay under common flight paths for Seattle-Tacoma International Airport, and are probably unattractive to gentrifying populations. Taken together, these two figures provide some sense of the intensive margin of gentrification by capturing disproportionate sorting of high-skill individuals to neighborhoods.

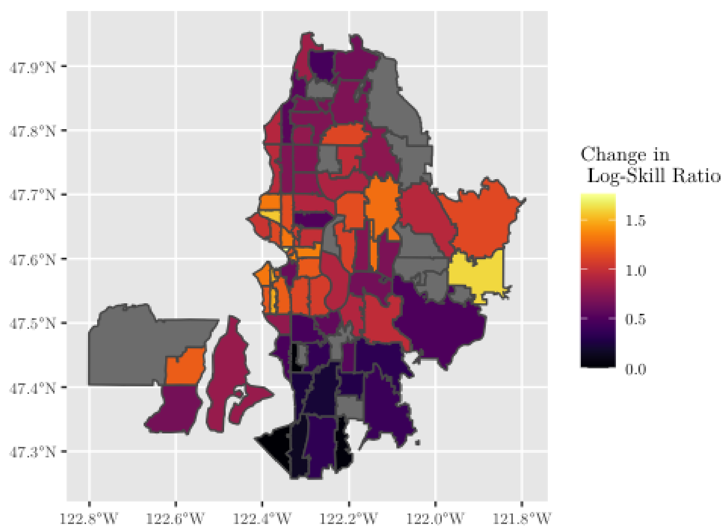
Figure 1.: Change in Log of Skill Ratio and Median Income by Distance to CBD



Dots are observed changes in the log-skill ratio. Curves fit using a LOESS process. Data comes from the US Census Bureau, 1990-2018, accessed via NHGIS.

The green line in figure 1 shows LOESS curve generated by the log of the change in median income for each zip code in the study. This line is quite similar to the black line, peaking around the city center and tapering in the suburbs. Evidently, affluent households have also been centrally sorting over the study window, in a pattern that closely mirrors the changing distribution of skill in the city. Pearson's correlation coefficient between changes in the log-skill ratio and log median income is 0.821, implying substantial correlation. Thus, when we observe changes in the log-skill ratio we can be pretty sure that we are also seeing changes in the median income and other demographic variables associated with gentrification.

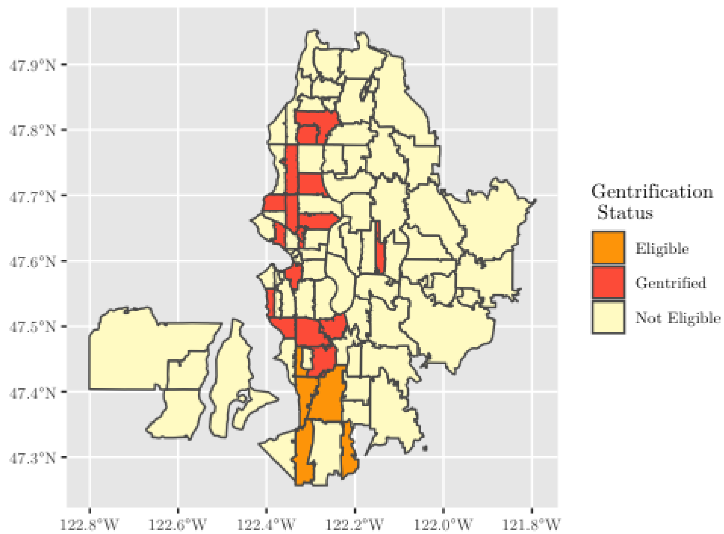
Figure 2.: Change in Log-Skill Ratio



Map filtered to include only zip codes with centroids within 35 kilometers of central business district. Data comes from the US Census Bureau, 1990-2018, accessed via NHGIS.

Figure 3 presents an alternative characterization of gentrification, using Freeman (2005) definitions of gentrification. Here, eligibility and gentrification are binary variables, with a tract only being considered gentrified if it was first eligible under those definitions. Of the 89 zip codes considered in the sample, 21 (24%) were eligible to be gentrified over the study period. Of these eligible zip codes, all but 5 did gentrify over this time period, or about 76%; were only these zip codes considered, Seattle would easily top the list of most gentrifying cities in the country. Once again, we see that gentrification is most heavily concentrated near downtown and to the north thereof. Roughly speaking, 3 depicts the extensive margin of gentrification in Seattle.

Figure 3.: Map of Gentrified zip Codes using Freeman (2005)



Eligibility and gentrification status using the definitions laid out in Freeman (2005). Data comes from the US Census Bureau, 1990-2018, accessed via NHGIS.

2.1.2 Firm Sorting. Using the sectoral definitions outlined in table 1, figure 4 documents changes in the distribution of firms in Seattle between 1994 and 2017 (the widest available range of dates of my data on the distribution of firms; see 2.3 for details). Note that this graph plots the change in the *share* of establishments rather than percent increases; a large increase in the establishment share represents a re-centering of a particular sector towards a new location. While the changes in the tradable sectors are relatively small and un-pronounced, there are clear and large changes in the distribution of non-tradable establishments. Both high- and low-skill non-tradable appear to be sorting more centrally, a trend most dramatically demonstrated by high-skill tradable. Being non-tradable firms, these

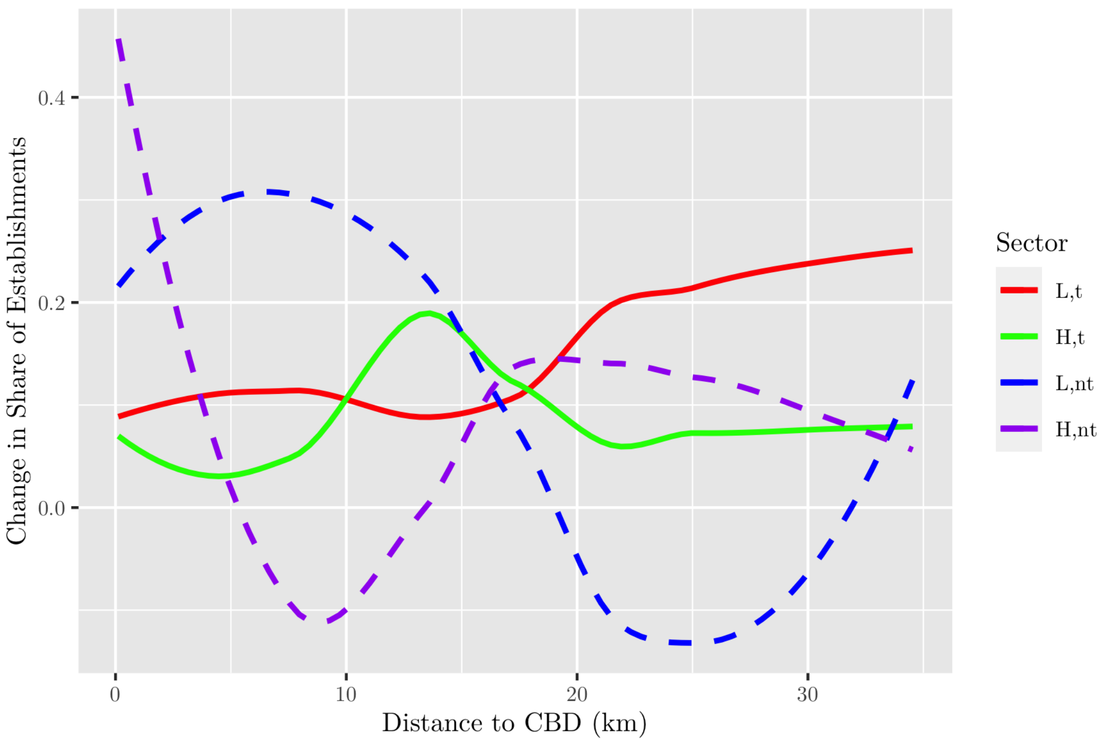
Table 1.: Division of NAICS Sectors by Tradability and Skill Level

	Tradable	Non-Tradable
High-Skill	Information (51)	Education (61)
	Finance, Insurance (52)	Health Care (62)
	Real Estate (53)	
	Professional Services (54)	
	Management of Companies (55)	
Low-Skill	Manufacturing (31-33)	Admin. Support, Waste (56)
	Wholesale Trade (42)	Arts, Entertainment, Recreation (71)
	Transportation, Warehousing (48-49)	Accommodation, Food Services (72)
		Retail (44-45)
		Other Services (81)

establishments need to locate near potential clientele to cut down on commute time to their locations. The relative volatility of these sorting patterns compared to those of tradable firms, who do not face such locational incentives, lends credit to the theory that these non-tradable firms are following increasing affluence in city centers. That said, it is clear that the sorting patterns of the high- and low-skill non-tradable firms is markedly different despite their shared central sorting.

Figure 5 disaggregates the sorting patterns of figure 4, breaking open the non-tradable sectors into their NAICS classifications. While the sorting of the aggregated sectors towards the city center, particularly the low-skill non-tradable, is decisively positive, the changes in individual sectors are much more chaotic and smaller in magnitude. The most pronounced central sorting are in the sectors of Administrative and Support and Waste Management services (56), Health Care and Social Assistance (62), and Accommodation and Food Services (72). Other sectors that we might expect to be associated with influxes of affluence, such as Arts, Entertainment, and Recreation (71) and Other Services (81) seem to be shifting out

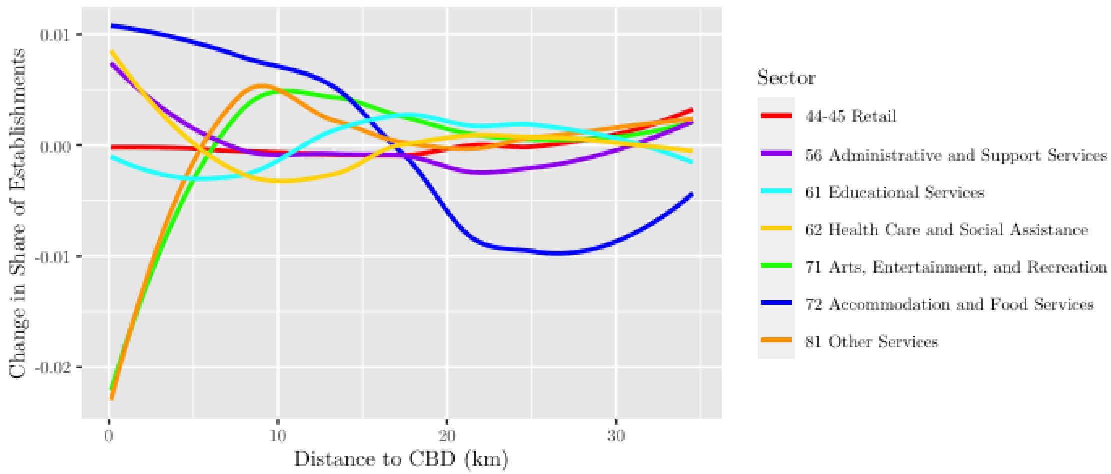
Figure 4.: Change in Share of Total Establishments by Sector



All curves fit using a LOESS process. 'H,nt' refers to the high-skill non-tradable sector, and so on. Dashed lines indicate changes for non-tradable firms, solid lines indicate tradables. Data comes from the US Census Bureau's zipcode Business Patterns Survey, 1994-2017.

of the central city; Retail (44-45) seems virtually unaffected by changing patterns of income. One could imagine further disaggregating these sectors and finding further heterogeneity, such as among gyms, movie theaters, and art galleries within arts, entertainment and recreation, but the picture will only get more complicated. Suffice it to say that while the aggregate non-tradable sectors in 1 are generally sorting towards the city center, many different sectors therein are experiencing different patterns individually.

Figure 5.: Change in Share of Total Establishments by Sector, Non-tradable Detail



All curves fit using a LOESS process. Data comes from the US Census Bureau's zipcode Business Patterns Survey, 1994-2017.

2.2 Model

I now detail the model I posit to investigate the connection between firm sorting and gentrification. I assume that the economy exists in a city divided into locations ℓ. There are two types of economic agents: households $i \in \mathcal{I}$ and firms in $j \in \mathcal{J}$. There are two types of households indexed e, high-skill (H) and low-skill (L); these types are assumed to be exogenous (I do not model the educational choices of households). Firms are divided into sectors indexed by s, each of which has production functions featuring different factor shares for each of the two types of labor. Firms are also labeled with t or nt for whether they are in a sector which is "tradable" or "non-tradable." Each location has a skill-specific amenities and sector specific total factor productivities; households and firms also have idiosyncratic preferences and productivities across land.

2.2.1 Households. Households seek to maximize their own utility, which is derived from the consumption of housing, tradable, and non-tradable goods, amenities, and an idiosyncratic taste shock for where they live. For a household i of type $e \in \{H, L\}$ living in location ℓ and working in ℓ' of the city, utility is given by

$$u_i^e(\ell, \ell') = K_i^{\alpha^e} \left(X_i^{t\beta} X_i^{nt1-\beta}\right)^{1-\alpha^e} \cdot \exp\left(\xi^e(\ell) + \sigma_h \varepsilon_i^h(\ell)\right)$$

where K_i is the amount of housing purchased by the household i, α^e is the budget share of housing for a household of type e, X^{nt} is an aggregation of consumption from all nt firms,

$$X_i^{nt} = \left(\int_{\mathcal{J}^{nt}} x_{ij}^{\frac{1+\zeta}{\zeta}} dj\right)^{\frac{\zeta}{1+\zeta}}$$

and X^t is a similar aggregation of all tradable firms' products. ζ is the price elasticity of demand for each firms' goods. These aggregations are the same as others used throughout the trade literature and due to Dixit and Stiglitz (1977). $\xi^e(\ell)$ is a neighborhood-skill level shared amenity term which is not directly observable to the econometrician.[3] $\varepsilon_i^h(\ell)$ is an idiosyncratic taste shock for location ℓ that is distributed according to a Gumbel distribution with standard deviation σ_h^e. The household faces the budget constraint

$$R(\ell)K_i + \int_{\mathcal{J}^t} p_j x_{ij} dj + \int_{\mathcal{J}^{nt}} p_j x_{ij} \tau(\ell, \ell_j) dj = \left(\sum_s \pi_s^e w_s^e(\ell')\right) \exp(\sigma_w^e \varepsilon_i^w(\ell')) \equiv w_i(\ell') \tag{2.1}$$

$R(\ell)$ is the price of a unit of housing at ℓ, p_j is the price of j's output, and $w_s^e(\ell')$ is the unit wage paid by firms at ℓ' to households of type e in sector s. $\tau(\cdot, \cdot)$ is an iceberg cost for traveling to purchase non-tradable goods from ℓ_j while living in location ℓ. For households the fundamental distinguishing factor between tradable and non-tradable goods is that they must pay $\tau(\ell, \ell_j)$ to obtain non-tradable goods.

[3] In section 2.4.3, I detail how I estimate $\xi^e(\ell)$ for both 1990 and 2018. These unobservable amenities are meant to "soak up" any neighborhood differences and changes (such as changes in crime rate or school quality) that are beyond the scope of this paper to model and estimate.

Ultimately this will mean that households have incentive to locate near tradable firms and vice versa. Household income conditional on place of work $w_i(\ell)$ is partly deterministic and partly stochastic. The deterministic part is equal to the sector-specific sum of wages in place of work ℓ', $w_s^e(\ell')$, weighted by the probability π_s^e that a type e household works in sector s. This deterministic part of income is augmented by $\varepsilon_i^w(\ell')$, an i.i.d. productivity draw from a Gumbel distribution (with standard deviation σ_w^e) that captures the idiosyncratic productivities of households across sector-location combinations. This assumption is made in other papers such as Tsivanidis (2019). Essentially, each location has a base income $\sum_s \pi_s^e w_s^e(\ell')$ that varies slightly with each worker based on their productivity. I assume that π_s^e is exogenous.[4]

Since utility is Cobb-Douglas, we have that

$$X_i^t = (1-\alpha)\beta w_i(\ell')/P^t \quad X_i^{nt} = (1-\alpha)(1-\beta)w_i(\ell')/P^{nt}(\ell) \quad K_i = \alpha w_i(\ell')/R(\ell)$$

Here, $P^{nt}(\ell)$ is the Dixit and Stiglitz (1977) price aggregator that measures cost of living in ℓ,

$$P^{nt}(\ell) = \left(\int_{\mathcal{J}} \big(p_j \tau(\ell, \ell_j)\big)^{1+\zeta} dj \right)^{\frac{1}{1+\zeta}}$$

and P^t is the same aggregator over all firms in the tradable sectors,

$$P^t = \left(\int_{\mathcal{J}_t} p_j^{1+\zeta} dj \right)^{\frac{1}{1+\zeta}}$$

Since there is no transportation cost associated with tradable goods, P^t is constant across locations.

[4] Theoretically, I could model these shocks as being sector-location specific and let households also sort over sectors. This, however, would mean that there would be no existing estimate of the parameter σ_w^e; treating sector choice as exogenous allows me to use the estimate of Tsivanidis (2019).

Taking logs and subtracting constants, I obtain the household's indirect utility function for living at ℓ and working at ℓ':

$$\begin{aligned}\hat{V}_i^e(\ell,\ell') =& \ln\Big(\sum_s \pi_s^e w_s^e(\ell')\Big) - \alpha^e \ln R(\ell) \\ & - (1-\alpha^e)\Big(\beta \ln P^t + (1-\beta)\ln P_m^{nt}(\ell)\Big) + \xi^e(\ell) + \sigma_h \varepsilon_i^h(\ell) + \sigma_w \varepsilon_i^w(\ell')\end{aligned} \tag{2.2}$$

Following Tsivanidis (2019), I assume that households receive their neighborhood preference shocks and decide where to live before they receive their work place productivity shocks, and that they choose where to live in expectation of where they think they will choose to work. This allows me to simplify equation 2.2 by integrating across potential work locations ℓ'. After having received their productivity shocks, the household chooses which location to work; following the logistic choice framework developed in McFadden (1973), the probability that a given household will work in ℓ' is

$$\pi^e(\ell') = \frac{\Big(\sum_s \pi_s^e w_s^e(\ell')\Big)^{\sigma_w}}{\int \Big(\sum_s \pi_s^e w_s^e(\ell'')\Big)^{\sigma_w} d\ell''}$$

Knowing this, households will locate themselves optimally to maximize their expected indirect utility,

$$\begin{aligned}V_i^e(\ell) =& \frac{1}{\sigma_h}\Bigg(\int \pi^e(\ell') \cdot \ln\Big(\sum_s \pi_s^e w_s^e(\ell')\Big) d\ell' \\ & - \alpha^e \ln R(\ell) - (1-\alpha^e)(1-\beta)\ln P^{nt}(\ell) + \xi^e(\ell)\Bigg) + \varepsilon_i^h(\ell) \\ =& \tilde{V}^e(\ell) + \varepsilon_i^w(\ell')\end{aligned} \tag{2.3}$$

Note here that P^t will be the same in all locations and so we may eliminate the $(1-\alpha)\beta \ln P^t$ term from $V_i^e(\ell)$ for sorting purposes. Thus, we have that the probability that the household chooses to live in ℓ is given by

$$\pi^e(\ell) = \frac{\exp(\tilde{V}^e(\ell))}{\int \exp(\tilde{V}^e(\ell''))d\ell''}$$

Finally, I assume there is an outside option which yields utility $V_0 = 0$ for both skill levels. In practice, choosing the outside option corresponds to living in the MSA but not the central city.

2.2.2 Firms. Firms are rational profit maximizers endowed with a sector s; they simultaneously hire labor from households in the city they are located in and provide goods and services to households. Each firm produces a differentiated product y_j subject to production function $y_{js} = B_s(\ell)\mathcal{K}_j^{\kappa_s}\mathcal{L}_{js}^{1-\kappa_s}\exp(\sigma_\epsilon\epsilon_j(\ell))$, a simple Cobb-Douglas function with a multiplicative productivity shock.

$$\mathcal{L}_s = (\theta_s H^{\frac{\rho-1}{\rho}} + (1-\theta_s)L^{\frac{\rho-1}{\rho}})^{\frac{\rho}{\rho-1}}$$

is a constant elasticity of substitution aggregator of labor types with a factor share θ_s that varies by sector; ρ is the elasticity of substitution between labor types, which I assume is shared by all sectors. $\mathcal{K}$ is the amount of land used by the firm (or alternatively, the intensity of land development). Being monopolistically competitive, firms set prices and the quantity of their output.

Recall that the difference between tradable and non-tradable firms is that consumers must commute to non-tradable firms to purchase their goods, incurring shopping costs, but can obtain tradable goods anywhere for the same price. For instance, while a restaurant needs to be located near enough residents of an appropriate level of affluence with taste for their food, a factory that produces cars need not locate immediately next to car buyers. Hence, tradable firms do not need to consider their location in the city relative to the distribution of household incomes, but non-tradable firms do. As a result, I must consider the sorting problems of both separately.

2.2.2.1 Tradable Firms. Tradable firms offer goods that can be purchased costlessly across space, and need not be commuted to by households to access, and so sort to maximize profit across rents, wages, and neighborhood productivities without regard for the distribution of households (except for their impact on these variables). Given the utility functions of households, it can be shown that firms in tradable sectors face demand curves from individual type e households as $x_{ij} = (1-\alpha^e)\beta w_i p_j^\zeta P^{t-(1+\zeta)}$, giving them total market demand curve

$$\mathbf{x}_j^t = \beta p_j^\zeta P^{t-(1+\zeta)} \sum_e (1-\alpha^e) \int N^e(\ell) w^e(\ell) d\ell \tag{2.4}$$

where $N^e(\ell)$ is the number of type e workers living in ℓ and $w^e(\ell)$ is their expected income. This gives tradable firms an inverse demand curve of $p_j = (y_j^t/\mathcal{H}^t)^{1/\zeta}$ where y is the output of a single firm and

$$\mathcal{H}^t = \beta P^{t-(1+\zeta)} \sum_e (1-\alpha^e) \int N^e(\ell) w^e(\ell) d\ell$$

captures the size of potential demand. I assume that firms do not recognize the impact of their own pricing decision on P^t. Dropping sector and firm specific indexing for ease of exposition the profit maximization problem for tradable firms can then be stated as

$$\max_y y^{\frac{1+\zeta}{\zeta}} \mathcal{H}^{t-\frac{1}{\zeta}} - R(\ell)K - w^H(\ell)H - w^L(\ell)L \tag{2.5}$$

$$\text{s.t. } y = B(\ell)\mathcal{K}^\kappa (\theta_s H^{\frac{\rho-1}{\rho}} + (1-\theta_s)L^{\frac{\rho-1}{\rho}})^{\frac{\rho(1-\kappa)}{\rho-1}} \exp(\sigma_\epsilon \epsilon_j(\ell))$$

Solving for conditional factor demands, I obtain:

$$L(y, w^L, w^H, R) = \frac{y}{B}\exp(-\sigma_\epsilon \epsilon)\Big(\frac{1-\theta}{w^L}\Big)^\rho \Big(\frac{1-\kappa}{\kappa}R\Big)^\kappa \Big(c(w^L, w^H)\Big)^{\rho-\kappa}$$
$$H(y, w^L, w^H, R) = \frac{y}{B}\exp(-\sigma_\epsilon \epsilon)\Big(\frac{\theta}{w^H}\Big)^\rho \Big(\frac{1-\kappa}{\kappa}R\Big)^\kappa \Big(c(w^L, w^H)\Big)^{\rho-\kappa}$$
$$\mathcal{K}(y, w^L, w^H, R) = \frac{y}{B}\exp(-\sigma_\epsilon \epsilon)\Big(\frac{\kappa}{R(1-\kappa)}\Big)^{1-\kappa} \Big(c(w^L, w^H)\Big)^{1-\kappa}$$

where $c(w^L, w^H) = (\theta^\rho {w^H}^{1-\rho} + (1-\theta)^\rho {w^L}^{1-\rho})^{\frac{1}{1-\rho}}$ is the unit cost function for the CES labor aggregator. I write the firms' cost function as $C(y, w^L, w^H, R) = y \cdot MC(w^L, w^H, R)$ where

$$MC(w^L, w^H, R, \epsilon) = \frac{1}{B}\exp(-\sigma_\epsilon \epsilon)\Big(\frac{R}{\kappa}\Big)^\kappa \Big(\frac{c(w^L, w^H)}{1-\kappa}\Big)^{1-\kappa}$$

Firms maximize profits by setting marginal revenue equal to marginal cost,

$$y_j(\ell) = \Big(\frac{\zeta}{1+\zeta} MC_s(w_s^L(\ell), w_s^H(\ell), R(\ell), \epsilon_j(\ell))\Big)^\zeta \mathcal{H}^t$$

where the s is to note that marginal costs and wages will vary by sector as well as by location. Firms will choose a price p_j to induce this quantity of demand. Firm j's individual profit at ℓ in sector s will then be

$$\hat{\Pi}_{js}^t(\ell) = -\zeta^\zeta \Big(\frac{1}{1+\zeta} MC_s(w_s^L(\ell), w_s^H(\ell), R(\ell), \epsilon_j(\ell))\Big)^{1+\zeta} \mathcal{H}^t$$

Note that profit will be positive for some values of the equilibrium distributions $w^L(\ell)$, $w^H(\ell)$, $R(\ell)$, and $\mathcal{H}^t$. Firms are competing with each other for their customers income through the tradable price index P^t which is nested inside $\mathcal{H}^t$; otherwise, there is no mechanism by which a market becomes "over-saturated" with firms. Taking logs and subtracting constants from the right hand side, I obtain the firm's indirect profit function for each location ℓ:

$$\begin{aligned}\Pi_{js}^t(\ell) &= \frac{1}{\sigma_\epsilon}\Big(\ln B(\ell) - \kappa_s \ln R(\ell) - (1-\kappa_s)\ln c(w_s^L, w_s^H)\Big) + \epsilon_j(\ell) \\ &= \tilde{\Pi}_s^t(\ell) + \epsilon_j(\ell)\end{aligned}$$

I also subtract the $\mathcal{H}^t$ term since it is constant in all locations ℓ for tradable firms. Following McFadden (1973), this means that the probability that a sector s firm locates in ℓ is

$$\pi_j^t(\ell) \equiv \Pr(\ell = \ell_j) = \frac{\exp(\tilde{\Pi}_s^t(\ell))}{\int \exp(\tilde{\Pi}_s^t(\ell'))d\ell'}$$

2.2.2.2 Non-tradable (nt) Firms. Unlike tradable firms in the section above, non-tradable firms must consider the location of household incomes in the city around them because households must travel to each firm to obtain its goods. The process for finding non-tradable firms' indirect profit is similar to that of tradable but with a few key considerations resulting from this fact. Individual household demand for a non-tradable firm's good is now $x_j = (1-\alpha^e)(1-\beta)w_i(p_j\tau(\ell_i,\ell_j))^{\zeta}P^{nt}(\ell_i)^{-(1+\zeta)}$; therefore the total demand faced by firm j in location ℓ_j

$$\mathbf{x}_j(\ell_j) = (1-\beta)p_j^{\zeta}P^{nt-(1+\zeta)}\sum_{e,s}(1-\alpha^e)\int N^e(\ell)w^e(\ell)\tau(\ell,\ell_j)^{\zeta}d\ell$$

giving each firm an inverse demand curve of $p_j = \mathbf{y}_j{}^{1/\zeta}\mathcal{H}^{nt}(\ell_j)^{-1/\zeta}$ by households where

$$\mathcal{H}^{nt}(\ell) = (1-\beta)\sum_{e}(1-\alpha^e)\int N^e(\ell')w^e(\ell')\tau(\ell,\ell')^{\zeta}P^{nt}(\ell')^{-(1+\zeta)}d\ell',$$

Fundamentally, $\mathcal{H}^{nt}(\ell)$ captures the same essence for non-tradable firms as $\mathcal{H}^t$ does for tradable firms: a factor for scaling the income of households in the model to become a rough proxy of market size. The difference is that since distances to different firms will vary by neighborhood, $\mathcal{H}^{nt}(\ell)$ is location specific. While both types of firms compete with each other across the entire city, non-tradable firms also compete locally to be in high demand areas.[5]

[5] It should be noted that since utility in the outside location (the rest of the MSA) is normalized to 0, non-tradable firms do not directly compete with other non-tradable firms outside

The profit maximization problem and conditional factor demands are identical to those of the tradable firms' down to sector specific production related parameter values. Following from section 2.2.2.1, I arrive at the non-tradable sector s firms' indirect profit function

$$\begin{aligned}\Pi^{nt}_{js}(\ell) &= \frac{1}{\sigma_\epsilon}\left(\ln B(\ell) - \frac{1}{1+\zeta}\ln \mathcal{H}^{nt}(\ell) - \kappa_s \ln R(\ell) - (1-\kappa_s)\ln c(w^L_s, w^H_s)\right) + \epsilon_j(\ell)\\ &= \tilde{\Pi}^{nt}_s(\ell) + \epsilon_j(\ell)\end{aligned}$$

Note here that $\mathcal{H}^{nt}(\ell)$ is not constant across locations and hence does not drop out. This reflects the need by non-tradable firms to locate close to their customer base. As before, there is no "saturation point" for the mass of firms in a given location, but they still compete with each other for households' income via the price aggregator $P^{nt}(\ell)$ inside of $\mathcal{H}^{nt}(\ell)$. Finally, I arrive at the non-tradable firms' probability of locating in any location ℓ:

$$\pi^{nt}_j(\ell) \equiv \Pr(\ell = \ell_j) = \frac{\exp(\tilde{\Pi}^{nt}_s(\ell))}{\int \exp(\tilde{\Pi}^{nt}_s(\ell'))d\ell'}$$

2.2.3 Labor Market.

The wage offered at each location $w^e_s(\ell)$ must be such to entice enough households in expectation of each sector and skill level to ℓ to meet the demand for each labor type. That is, $w^L_s(\ell)$ must solve

$$\int\int_{\mathcal{J}_s(\ell)} L(y_j, w^L_s(\ell), w^H_s(\ell), R(\ell))djds = \int \frac{(w^L(\ell'))^{\sigma^L_w}}{\int (w^L(\ell''))^{\sigma^L_w}d\ell''}d\ell' \tag{2.6}$$

This is an equilibrium condition to ensure clearing in the market for unskilled labor across all locations. An analogous expression holds for high-skilled labor as well. Note that the wage in question, $w^e_s(\ell)$, appears in three places in this equation: in the factor demand, and both $(w^L(\ell'))^{\sigma^L_w}$ and $(w^L(\ell''))^{\sigma^L_w}$.

the central city. This is an obvious simplification of reality meant to place the focus of my analysis on interactions in the central city.

2.2.4 Land Market. One of the key indicators that gentrification may be taking place in a city is swiftly rising land prices near formerly low-rent/income neighborhoods. I assume rents take on an isoelastic formulation,

$$R(\ell) = (\ell)\left(\int_{\mathcal{I}(\ell)} K_i di + \int_{\mathcal{J}(\ell)} \mathcal{K}_j dj \right)^{\psi} \tag{2.7}$$

where ψ is the inverse elasticity of supply for land in city m and $\mathcal{I}(\ell)$, $\mathcal{J}(\ell)$ are the sets of households and firms respectfully locating at ℓ.

2.2.5 Equilibrium. Let $h_s^e(\ell)$ be a density of households of type e and sector s and $f_s(\ell)$ be a density of sector s firms at ℓ. Then given a set of parameters

$$\{\{\alpha^e\}, \eta, \zeta, \chi, \{\delta\}, \{\rho\}, \{\theta_s\}, \{\kappa_s\}, \sigma_h, \sigma_w, \sigma_\epsilon, \{\psi\}, \{\bar{R}(\ell)\}\},$$

and shopping costs $\tau(\ell, \ell')$, an equilibrium in this model is a set of prices $\{w_s^L(\ell), w_s^H(\ell), R(\ell)\}_{s,\ell}$, distributions $\{h_s^L(\ell), h_s^H(\ell), f_s(\ell)\}_{s,\ell}$ and firm specific variables $\{p_j, y_j\}_j$ such that

1. Households choose their location to maximize indirect utility, and consume optimally at that location to maximize utility taking prices and the distributions of prices, firms, and other households as given.

2. Firms choose their location to maximize their indirect profit and produce optimally at that location to maximize profit taking the distributions of rents, wages, households, and other firms as given.

3. Wages are set according to 2.6.

4. Rents are formed are formed according to 2.7.

2.3 Data

In this section, I detail the data sources to be used to calibrate the model in section 2.2 and discuss any limitations thereof. The most natural definition of a neighborhood would be Census tracts; however, tracts tend to be redrawn significantly between decennial censuses, making spatial analysis difficult over a wide time frame, such as in this study. Firm level data, which I need to help differentiate the actions of firms impacted by changing income dynamics differently, is also not available at the tract level. Therefore, I choose to use zip code boundaries to define neighborhoods. These have better firm data available and are typically more fixed in their spatial definitions across time. I restrict my spatial analysis to only those zip codes with centroids less than 35 kilometers from the central business district, which I define using the latitude-longitude coordinates of the downtown Seattle pin on Google Earth. I also define an individual as being high-skill if they have a Bachelor's degree or higher and sectors as in table 1.

Most data for this study ultimately come from the US Census Bureau, but are accessed from a variety of sources. From the R package "tidycensus," I obtain counts of high- and low-skill individuals for 1990 and 2018 at the zip code level, as well as boundary files for each. This data come from the Census' American Community Survey for 2018, and the 1990 decennial census. I filter out all individuals less than 25 years of age from these counts. I include all such individuals in the Seattle-Tacoma-Bellevue MSA despite only considering the central city in my analysis, and let the rest of the city be an outside option to account for in- and out-migration under the counterfactual. I turn to the National Historic GIS (NHGIS) database for data on median income, and housing values and construction dates to determine the gentrification status of each neighborhood in figure 3; I also use these to help determine the extensive margin of gentrification under the counterfactual.

Table 2.: Populations by Skill Level

Skill	1990	2018	% Change
High	461,371	1,157,542	151%
Low	1,254,795	1,643,694	31%

Data from US Census Bureau.

I access a large sample of microdata from the Integrated Public Use Microdata Survey (IPUMS), which contains individual Census data on a 1% sample of individuals nationally in the US, which I filter down to just the Seattle-Tacoma-Bellevue MSA. From these data I can observe an individual's age, education, income, and sector of employment. I use these to calculate the probability of being employed in one the aggregate sectors outlined in table 1 conditional on skill level. The 1990 IPUMS file uses the Standard Industrial Classification (SIC) system which was standard at the time; I convert data to NAICS using a crosswalk available from the Census. A few observations are lost in this process.

I obtain data on the distribution of firms from the US Census Bureau's zip Code Business patterns survey (ZBP). This gives me establishment counts by number of employees at the most detailed level possible; I use 2017's ZBP to proxy for 2018 and 1994's for 1990 due to lack of data. Again, 1994 uses the SIC system, so I likewise convert these sectors to NAICS. I then aggregate all establishments up to the 4 sectors from above. As with individuals, I include all firms of each time within the entire MSA even though I restrict spatial analysis to the central city. Table 3 reports establishment counts at the aggregated sector level, the share of high- and low-skill individuals employed by each, and the changes in these numbers over time. The dramatic appreciation of firm counts could be due to the increasing prevalence of information technologies and working from home and self-employment, or missing data from 1994.

Table 3.: Establishment Counts and Share of Workers by Skill-Level By Sector

	1990			2018		
Sector	Count	L-Share	H-Share	Count	L-Share	H-Share
L, t	1,622	0.36	0.26	11,912	0.26	0.15
H, t	3,040	0.31	0.27	39,251	0.15	0.34
L, nt	2,719	0.17	0.15	30,363	0.37	0.21
H, nt	1,078	0.16	0.31	14,088	0.22	0.31

Data from ZBP and IPUMS USA.

To calibrate production side parameters, I rely on two data sets. To obtain $\{\kappa_s\}_s$, the land share for each sector, I turn to the Bureau of Economic Analysis' (BEA) data on gross value added, employment, and worker compensation by sector.[6] I obtain estimates of skilled versus unskilled labor shares $\{\theta_s\}_s$ using wage and hours worked by sector from the CPS dataset. This data is an aggregate of monthly employment data from US labor force survey along with the CPS that records information such as an individual's hours worked per week, salary, and industry of employment.

I assume that each zip code's location can be decently approximated by its geographic centroid. I use this assumption and Google's Travel Matrix to obtain commuting times between neighborhoods, which allows me to enter centroid coordinates and find the time and distance of the optimal travel route between them. I specifically set mode of transit to "driving" and do not model traffic.

[6]The BEA itself obtains this data from the Decennial Census, ACS, and Economic Census carried out by the Census, which is then aggregated by industry and year. See https://www.bea.gov/data/employment/employment-by-industry.

Historical travel times are not available; therefore I use modern travel times as an approximation for both 1990 and 2018.

2.4 Calibration

2.4.1 Household Parameters.

I calibrate the budget shares of housing for high- and low-skilled households to $\alpha^L = 0.337$ and $\alpha^H = 0.321$ based on the Bureau of Labor Statistic's estimate of personal expenditure by education (Foster, 2014). These data come from the US Census Bureau's 2014 Consumer Expenditure Survey and I take to be stable over time. I set the budget share of tradable goods to $\beta = 0.4$ following the findings of Bems (2008); this estimate has also remained relatively stable over time. For the elasticity of product demand I turn to Head and Mayer (2014) and Suárez Serrato and Zidar (2016), setting $\zeta = -2.5$.For the iceberg shopping cost leveled on non-tradable goods, I abuse notation slightly and set $\tau(\ell, \ell') = 1 + \tau \times$travel time from the Google Travel Matrix, where $\tau > 0$. Essentially, τ is a percentage markup over the listed price of goods that increases with the time required to travel to the good's producer. I set $\tau = 0.5$ (a 50% markup for every hour spent traveling). This may seem excessive, but recall that this mark-up is only applied to non-tradable goods and services. For instance, while it would not be unusual for a person to drive an hour-plus to visit a car dealership (tradable, durable), it would be to visit a restaurant (non-tradable, non-durable) when similar near-by options are available. Once the price of transit/gasoline, the value of time, and other similar costs of convenience are accounted for in that hour of travel, the nominal cost could easily approach a 50% mark up in any regard. I set the inverse supply elasticity of housing $\psi = 1.136$ according to Saiz (2010).

I obtain values for the standard deviations of the households' shock terms from two similar papers in the literature. For the neighborhood living preference shocks standard deviation, σ_h^e, I turn to Su (2020), who implicitly calculates the this value when estimating the rent sensitivity parameter of that model; in that model, the rent sensitivity parameter was the ratio of the household budget share and the equivalent standard deviation. Su (2020) estimates these sensitivities to be 0.7950 for high- and 0.4593 for low-skill individuals. Using my calibrated α^e's as the budget share values (the numerator), I can back out the implied standard deviations calculated in that paper. This gives me $\sigma_h^H = 0.4038$ and $\sigma_h^L = 0.6804$. Tsivanidis (2019) is my source for the standard deviations of the neighborhood productivity shocks; that paper estimates Frechét shocks with scale parameters of 2.274 and 3.299 for high- and low-skill workers respectively. Inverting these gives the Gumbel equivalent, or $\sigma_w^H = 0.4397$ and $\sigma_w^L = 0.3031$.

2.4.2 Firm Parameters. To obtain $\{\kappa_s\}_s$, the land share for each sector, I turn to the Bureau of Economic Analysis' (BEA) data on gross value added, employment, and worker compensation by sector. Similarly to Giandrea and Sprague (2017), for each NAICS sector I estimate

$$1 - \kappa_s = \frac{\text{Employee Compensation}}{\text{Gross Value Added}}.$$

The κ_s for each aggregate sector I estimate using an average of individual NAICS sectors weighted by total employment. Luckily data is available for both 1990 and 2018, though the 1990 data uses Standard Industrial Classification (SIC) rather than NAICS codes; as the data is not disaggregated to the 3-digit code level, I have to perform a rough resorting of SIC sectors into NAICS for the 1990 data. Overall, the factor shares for those resorted sectors do not change substantially more than for the others.

I obtain estimates of high- and low-skill labor shares $\{\theta_s\}_s$ using wage and hours worked by sector from the CPS dataset. To convert from salary to hourly wages, I normalize each observation by the usual number of hours worked by that individual in a given week and assume a 50-week work year. θ_s is then computable via

$$\theta_s = \frac{w_s^H H_s^{1/\rho}}{w_s^H H_s^{1/\rho} + w_s^L L_s^{1/\rho}}$$

where here H_s and L_s are the total number of hours worked by individuals of the respective skill types working in sector s in the CPS sample for the years under study. Table 4 reports my calibration of land and skilled labor shares for each NAICS super-sector in both 1990 and 2018.

Table 4.: Calibration of Land share κ_s and High-Skill Labor share θ_s for NAICS Super-Sectors

	1990		2018	
Sector	κ	θ	κ	θ
L, t	0.421	0.342	0.509	0.452
H, t	0.368	0.547	0.365	0.735
L, nt	0.450	0.357	0.557	0.499
S, nt	0.187	0.626	0.195	0.702

From Suárez Serrato and Zidar (2016), I set σ_ϵ, the standard deviation of the firm productivity shocks, to 0.28. Finally, following the exhaustive review of the literature, Card (2009), I set $\rho = 1/0.7$ for all sectors.

2.4.3 Computing Amenities $\xi^e(\ell)$ and Total Factor Productivities $B_s(\ell)$. With all other parameters in hand, I am able to back out the neighborhood unobservable amenities $\xi^e(\ell)$ and firm total factor productivities $B_s(\ell)$. The procedure uses the same insights as in S. Berry, Levinsohn, and Pakes (1995) on the nature of the Gumbel distribution. I start by deriving neighborhood unobservables. Recall that the non-stochastic portion of utility for a given household is given by $\tilde{V}^e(\ell)$; I normalize $V_0^e = 1$ for the outside option. From the data on the distribution of individuals by skill level, I directly observe $\pi^e(\ell)$, the share of total individuals in location ℓ. I can then compute $\tilde{V}^e(\ell)$ as

$$\tilde{V}^e(\ell) = \ln \pi^e(\ell) - \ln \pi_0^e \tag{2.8}$$

where π_0^e is the share of individuals living in the outside option. This provides me a fixed estimate of mean utility in each neighborhood from which I can extract $\xi^e(\ell)$; from equation 2.3, we have that

$$\xi^e(\ell) = \tilde{V}^e(\ell) - \left(\int \pi^e(\ell') \cdot \ln \Big(\sum_s \pi_s^e w_s^e(\ell') \Big) d\ell' - \alpha \ln R(\ell) - (1-\alpha^e)(1-\beta) \ln P^{nt}(\ell) \right)$$

Using the other parameters of the model and my solution algorithm, I can then run the model holding household and firm sorting to observed levels, obtaining the prices $\{P(\ell), w^e(\ell), R(\ell)\}$ that rationalize the observed equilibrium. Applying the above formula then allows me to back out neighborhood amenities.

A very similar operation can be performed for each sector of firm. Here, as above, let $\tilde{\Pi}_s^{nt}(\ell)$ be the non-stochastic portion of a non-tradable firm indirect profit function. Based on the same logic, this shared term can be computed from the data using observed shares of firms, $\tilde{\Pi}_s^{nt}(\ell) = \ln \pi_s^{nt}(\ell) - \ln \pi_{s0}^{nt}$. After running the model with firm and household distributions fixed to obtain model-fitting prices,

$$B_s^{nt}(\ell) = \tilde{\Pi}_s^{nt}(\ell) - \Big(-\frac{1}{1+\zeta} \ln \mathcal{H}^{nt}(\ell) - \kappa_s \ln R(\ell) - (1-\kappa_s) \ln c(w_s^L, w_s^H) \Big).$$

An respective operation holds for tradable firms as well.

2.5 Results

The calculation of neighborhood amenities $\xi^e(\ell)$ and firm productivities $B_s(\ell)$ allows my model to perfectly replicate the observed equilibrium; therefore

I do not present evidence of model fit for a baseline simulation.[7] Instead, I focus on counterfactual simulation in this section. First, I decompose the observed gentrification by parametric change in section 2.5.1. Next, in section 2.5.2, I hold the distribution of firms fixed between the two study years to demonstrate the importance of firm sorting in the gentrification of Seattle.

2.5.1 Decomposition of Gentrification by Parameter Changes.

First, I investigate the extent that the change in key parameter values between 1990 and 2018 impacted gentrification in Seattle. I divide the parameters in question into two sets, production parameters $\{\kappa_s, \theta_s, B_s(\ell)\}_s$ and household amenities $\{\xi^e(\ell)\}_e$. Apart from the sorting of households and firms and the changes in their relative masses, these are the only changes that occur in the model that bring about a different equilibrium between the two years. These exercises clarify the source of change not due to firm sorting. In both simulations, prices, rents, and wages are endogenous.[8]

Figure 6 depicts the simulated changes in the log-skill ratio resulting from holding each set of parameters constant. Unsurprisingly, holding production parameters constant results in only a marginal shift in the distribution of households. These parameters, with the exception of $B_s(\ell)$, change only for the entire city, and thus have little impact on the relative distribution of wages, rents, and prices which are the aspect of firm decisions that impact household sorting.

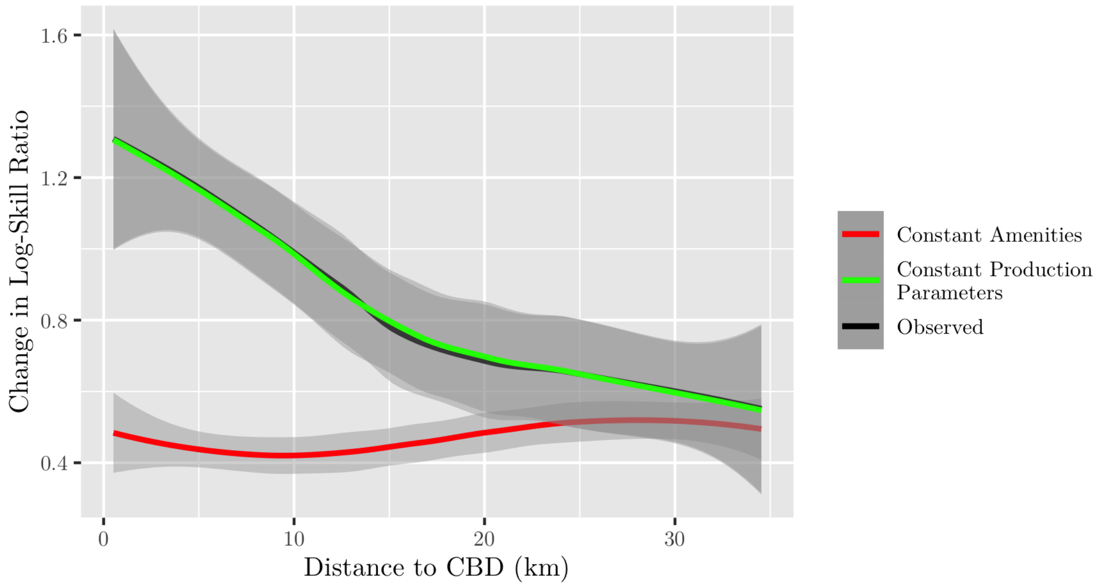
[7]While I do not prove that the equilibrium generated by my model is unique, a general "rule of thumb" is to look to the coefficients on the agglomerative terms in household utility, which in my model are rents and the non-tradable price index. Since the coefficients attached to both of these terms are less than 1, it would be highly unusual to observe multiple equilibria in this model.

[8]It should be noted, however, that these prices do not correspond to observed rents and wages, and that price indices are not observed at all.

Holding the unobservable amenities constant, however, results in significantly less appreciation of the log-skill ratio in the city center compared to the observed equilibrium. Evidently, changing downtown amenities between the study years disproportionately favor high-skilled households in the observed equilibrium; holding these amenities constant at 1990 levels results in disproportionately more sorting of low-skill households towards the city center. As shown in figure 7, while the populations of both skill levels generally increase by more under the fixed amenities scenario, this increase is much larger for low-skill households. This is in line with work by Su (2020), Edlund, Machado, and Sviatschi (2015), who find that appreciating downtown amenities for high skill households are a key driver of gentrification. In the larger body of work on gentrification, non-firm related explanations for the change in downtown amenities for high-skill households include reductions in crime, provision of parks, and improvements in school quality among others. Decomposing the unobservable amenities $\xi^e(\ell)$ is beyond the scope of this paper, but nevertheless, that they follow similar patterns to other estimates in the literature adds credibility to the robustness of my estimates.

Finally, for both parameter-fixing exercises, I document the change in the distribution of firms from the observed equilibrium. Figure 8 depicts the observed and simulated changes in the distribution of firms under each parameterization. The relative proportions of tradable firms throughout the city are essentially unchanged under the fixed production parameters counterfactual, reflecting the fact that these firms do not need to follow changes in the distribution of residents resulting from these parameter shifts; rather their behavior is governed by the unaltered production parameters. Both types of tradable firms show similar

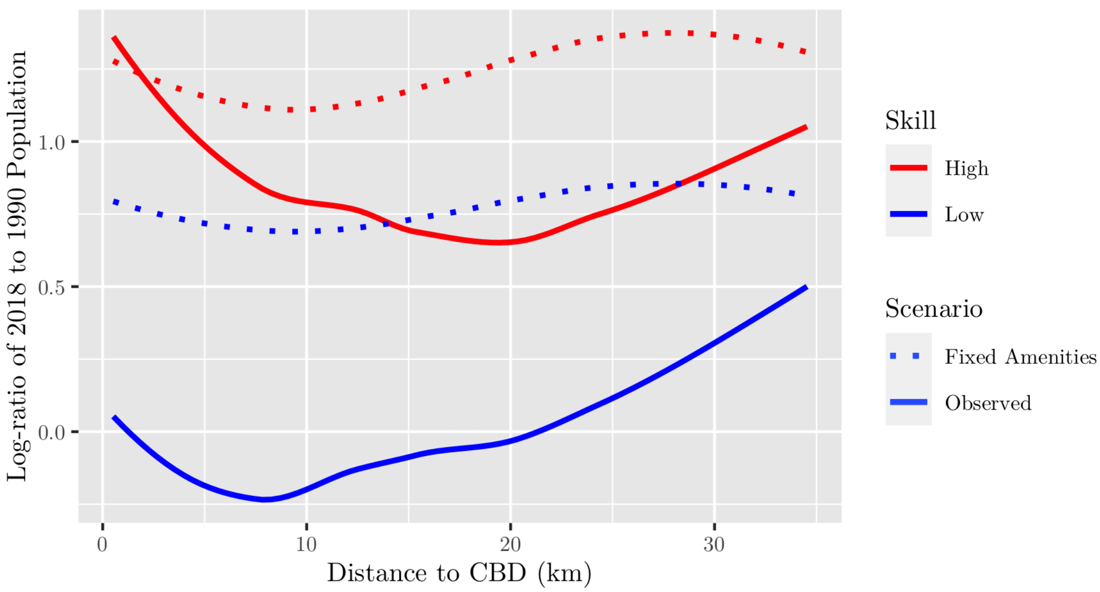
Figure 6.: Change in Log-Skill Ratio With Time-Varying Parameters Held Constant



Simulated changed in the log-skill ratio while holding indicated sets of parameters constant between 1990 and 2018. 95% standard error bands shown by shaded areas. All curves fit using a LOESS process.

patterns when neighborhood amenities are held constant, sorting outside of the downtown area and into the outside option in greater numbers than under the observed equilibrium. This likely is the result of the influx of households of both types under the fixed amenity counterfactual driving up rents; since tradable firms need not locate near their customers, they are pushed out thereby. Conversely, both alternative scenarios lead to wildly different sorting patterns for non-tradable firms. Most notably, the influx of residents under the fixed amenities scenario leads to a dramatic increase in the share of low-skill intensive non-tradable firms locating in the city; this pattern is not replicated by high-skill intense non-tradable firms. This disparity may reflect the fact that disproportionately more low-skill households sort into the city under the fixed amenities scenario, and these firms choose to locate

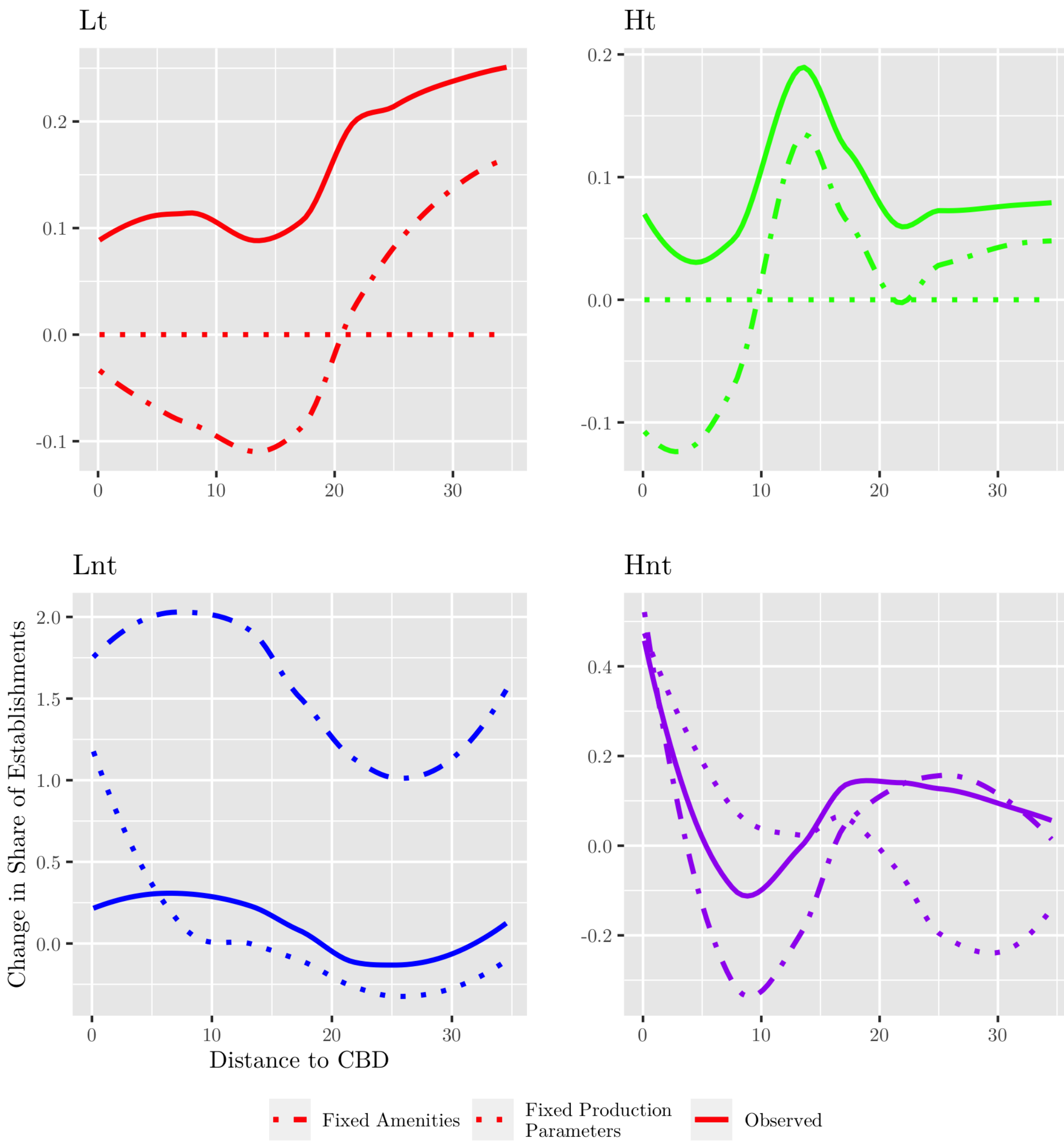
Figure 7.: Log-Ratio of 2018 to 1990, Observed and with Fixed Amenities



Log ratio of 2018 to 1990 populations under observed and fixed-amenities scenarios. All curves fit using a LOESS process.

nearer those workers to cut down on the wages required to entice sufficient labor. Overall, in every scenario, including the observed equilibrium, non-tradable firms exhibit more drastic changes in their distribution between the study years.

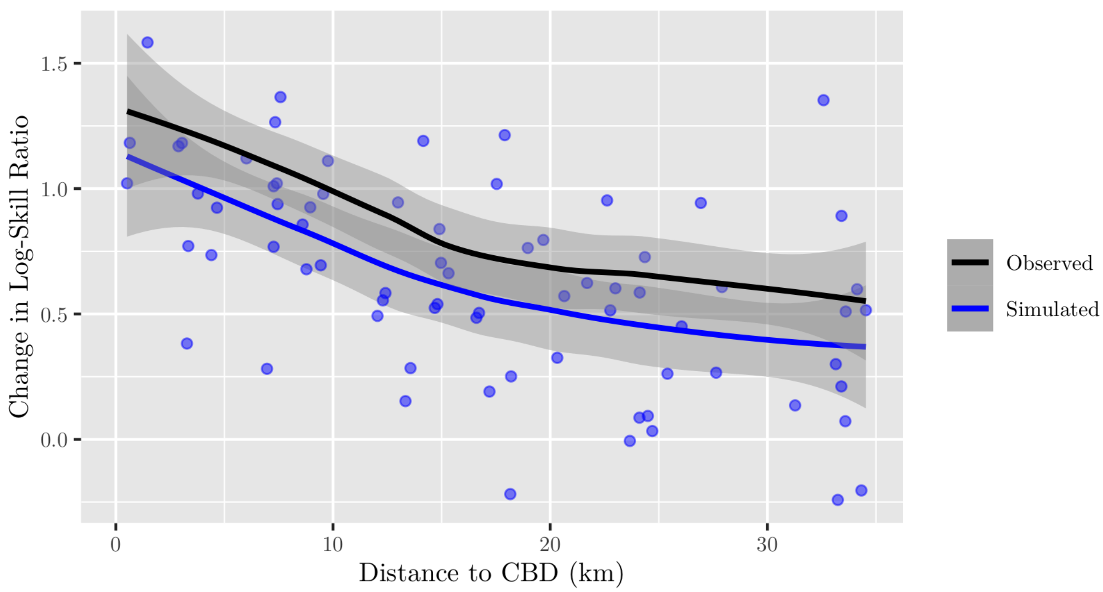
Figure 8.: Change in Distribution of Firms with Time-Varying Parameters Held Constant



Simulated and observed changes in firm distributions while holding indicated sets of parameters constant. All curves fit using a LOESS process.

2.5.2 The Role of Firm Sorting. To determine the impact of firm sorting on the gentrification of Seattle from 1990 to 2018, I now simulate equilibrium with the distribution of firms held constant at 1990 levels. To do this, I allow all parameters, wages, rents, prices, and households to change between 1990 and 2018; I also allow the numbers of each sector of firms to increase to 2018 levels. However, I do not allow firms to sort out of their 1990 neighborhood shares.

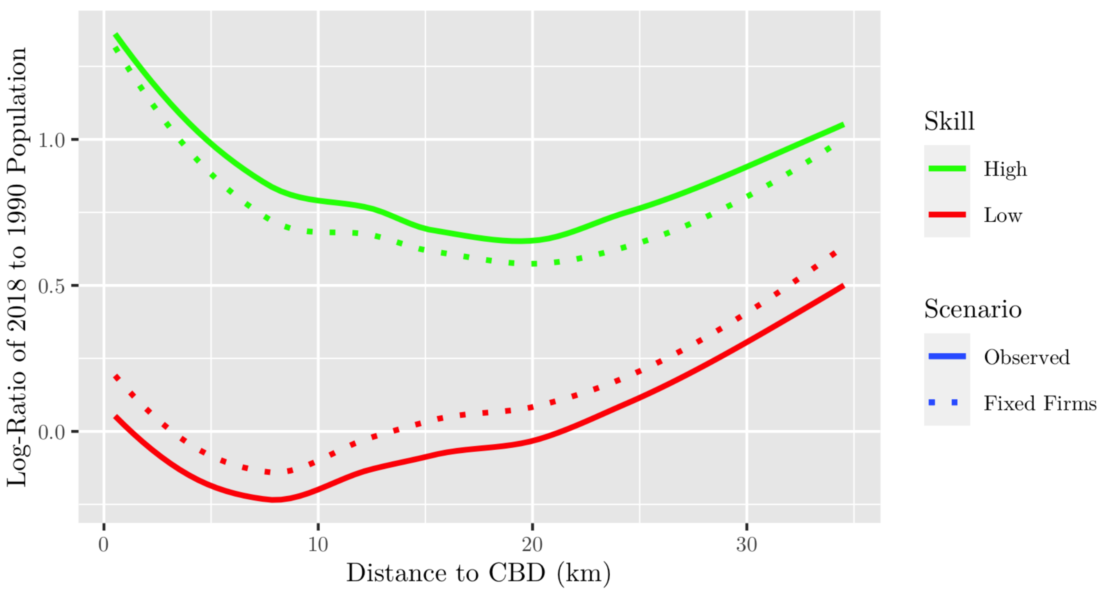
Figure 9.: Change in Log-Skill Ratio Holding Firm Distributions Constant



Observed and simulated change in log-skill ratio when firm sorting is shut down. Curves fit using a LOESS process.

Figure 9 plots the observed and simulated change in the log-skill ratio when the distribution of firms is held constant. The counterfactual simulation closely mirrors the shape of the observed equilibrium, but is consistently lower, demonstrating a decrease in the intensive margin of gentrification. On average, the increase in the log-skill ratio is 28.6% less when firm sorting is shut down. Figure 10 provides clarity on the cause of this decrease in gentrification. When firm sorting is shut down, the model predicts a greater appreciation in the number of low-skill households and a lesser appreciation of high-skill households than under the observed equilibrium. That is, relative to low-skill households, there is less disproportionate sorting of high-skill households into the city.

Figure 10.: Change in Log-Population Ratio Under Observed and Counterfactual Equilibria

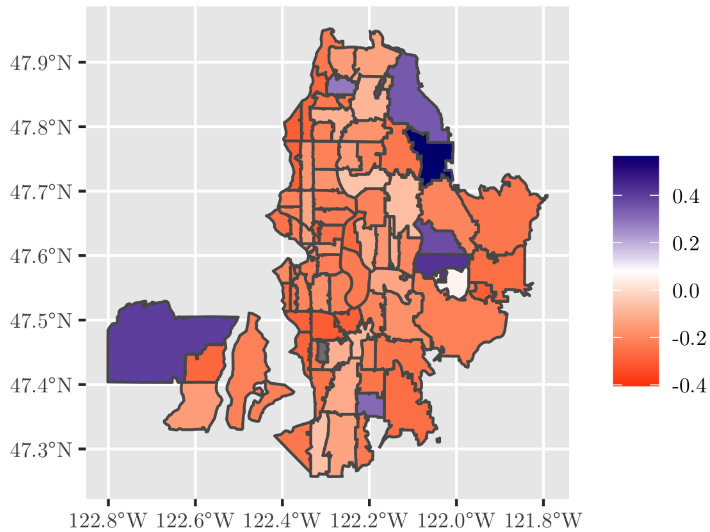


Observed and simulated log-ratio of population between 2018 and 1990 when firm sorting is shut down. Curves fit using a LOESS process.

Figure 11 maps the difference in the appreciation of the log-skill ratio by zip code in the city. As shown in figure 9, the vast majority of neighborhoods in the central city experience less gentrification when firm sorting is shut down. The decrease is particularly acute to the just north and south of the city center, the areas that qualify as gentrifying over the study period according to Freeman (2005) (see figure 3). A few locations near the outskirts of the central city experienced a greater appreciation of the log-skill ratio under the counterfactual than the observed equilibrium (these appear as purple). These represent the increased tendency for high-skill households to sort towards the suburbs and for low skill individuals to sort centrally when firm sorting is shut down.

To get an idea of the change in the extensive margin of gentrification, figure 12 depicts which zip codes underwent gentrification according to the Freeman

Figure 11.: Change in Change of Log-Skill Ratio Holding Firm Distributions Constant



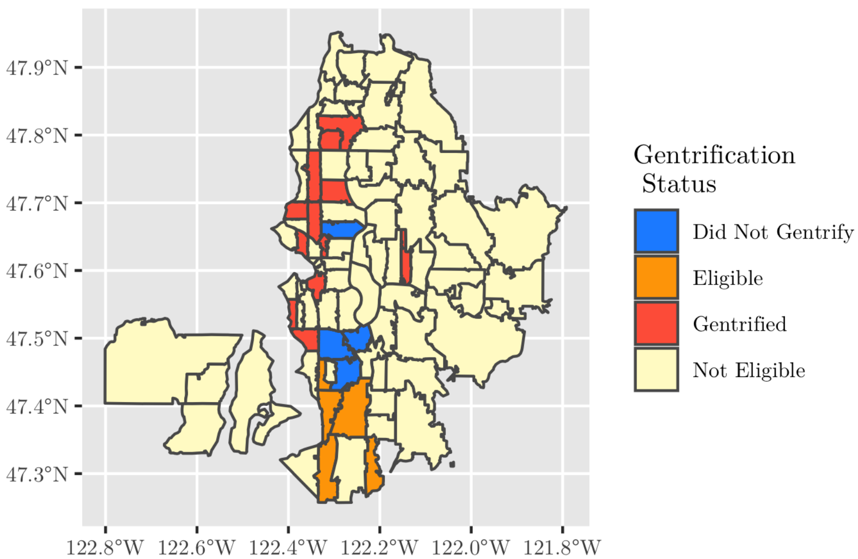
Difference in the change of the Log-skill ratio from the observed equilibrium to the no firm sorting counterfactual.

(2005) definitions. The blue zip codes are those that gentrified under the observed equilibrium but did not gentrify when the distribution of firms was held constant at 1990 levels. Of the 21 zip codes eligible for gentrification, only 15 gentrified under the counterfactual, or 4 less (21%) than observed.

2.6 Conclusion

In this paper, I developed one of the first models of gentrification that embeds endogenous firm sorting. Such models represent a substantive innovation to the literature on gentrification, which heretofore have lumped firm sorting in with the formation of other amenities. My calibrated model demonstrates the importance of firm sorting in the process of gentrification; shutting down changes in the distribution of firms between the years of 1990 and 2018 leads to an average of 28% less gentrification across zip codes in Seattle. Ultimately,

Figure 12.: Gentrification Status of zip Codes under Freeman (2005) with Firm Distributions Held Constant



Eligibility and gentrification status using the definitions laid out in Freeman (2005). "Did not gentrify" denotes neighborhoods that gentrified in the observed equilibrium that do not when firm sorting is held constant.

under commonly used thresholds to determine gentrification, I find that 21% fewer zip codes gentrify in the absence of firm resorting. These numbers represent decompositions of endogenous amenity impacts reported in papers such as Su (2020) and Edlund et al. (2015). I conduct this analysis in a very standard spatial model environment, meaning that these results are likely generalizable to a wide array of economic modeling circumstances. The results of this paper could prove important to policy makers. Hoelzlein (2019) recently showed that policies aimed at reigning in gentrification and promoting local business growth could have counter-productive consequences once firm sorting is accounted for; this paper has shown the importance of firm sorting in gentrification in general, further suggesting that proper accounting for the co-sorting of households and firms is crucial in understanding the dynamics of a changing modern city if we are to craft effective policies.

This work leaves a substantial mantle for future research. Adapting this model to investigate the impact of popular policies, such as rent control, opportunity zones and other place-based policies, or housing vouchers, would be relatively simple to carry-out. Failing to account for the reaction of firms to these policies (and the reactions of households thereto) could lead simulations to mis-characterize counterfactual equilibria. In the very least, this model may be useful as a template for similar inquires regarding gentrification that need accurate accounting of the impacts of firm sorting.

CHAPTER III

FIRM SORTING AND GENTRIFICATION IN PORTLAND, OREGON

Gentrification, the process by which formerly low income central city neighborhoods become occupied by more affluent, highly-educated residents, has been a powerful force in the transformation of urban centers in the past 50 years. These more affluent residents had predominantly lived in the suburbs of large cities prior to 1970, causing downtown housing to be in relatively low demand. However, as these residents began to sort centrally, downtown rents rose, and central city neighborhoods saw numerous pervasive changes: new construction, new goods and services, and reductions in crime. By 2010, many cities in the US had completely flipped the relationship of affluence, housing costs, and distance to downtown. This 50 year transformation has not left former downtown residents unaffected; research on the welfare impacts of gentrification has been extensive but mixed. Some studies have found that the original residents of these neighborhoods are largely ambivalent to the changes brought about by the process (Murdie & Teixeira, 2011); some residents even extol the positive impacts that gentrification has had on their neighborhoods in the form of reduced crime and increased consumption opportunities (Doucet, 2009). Other papers have found that gentrification can have more negative effects, such as displacement (Atkinson, Wulff, Reynolds, & Spinney, 2011) and adverse health impacts (Smith, Breakstone, Dean, & Thorpe, 2020) for former residents. Shifting employment characteristics (Lester & Hartley, 2014) and changing racial composition (Huante, 2021) have also been noted.

This paper investigates the role that firm sorting has on gentrification. Past work has noted that gentrification is a self-reinforcing process; with surging affluence in city centers come changes in amenities, such as reductions in crime and improvements in school quality, which further incentivize the central sorting of affluent households. The re-sorting of firms to serve these populations has also been prominently noted (Handbury, 2019; Su, 2020). However, most papers aggregate all firm and non-firm related amenities into reduced-form amenity terms. There are several reasons policy makers might be interested in the separate contribution of firm-sorting to gentrification. First, firms play a crucial role in establishing the culture of a given neighborhood, especially when locally owned by residents. The rising rents of gentrification and influxes of new firms may crowd out long-time community fixtures and locally owned businesses and cause neighborhoods to feel less like home to long time residents. Additionally, place-based policies aimed at economic development may actually have adverse impacts when considered in the context of gentrification; in particular, economic opportunity zones that provide incentive for businesses to locate in impoverished areas made lead to further gentrification if they encourage influxes of wealthy residents to follow (Hoelzlein, 2019). Crafting policies which support existing businesses will require an understanding of how households and firms choose to co-locate throughout a city.

In this paper, I repeat the analysis of chapter two of this dissertation in a novel environment. I develop a model of household and firm sorting in central cities that characterizes the interactions of each across space. Households choose where to locate in the city subject to rents, wages, amenities, and shopping costs accrued from purchasing goods from non-tradable firms across space. Firms locate in response to household decisions and compete monopolistically via a Dixit and Stiglitz (1977) framework. The first chapter of this dissertation calibrated the model for the metropolitan area of Seattle, WA and found that firm sorting accounted for a significant portion of gentrification in the city. This paper recalibrates the model for the Portland-Vancouver-Hillsboro metropolitan statistical area (Portland MSA, or simply Portland) between the years of 1990 and 2018 and conducts a number of counterfactual simulations. First, I document the role of changing firm-side parameters of production, including labor intensities by skill level and neighborhood total factor productivities. I then depict the impact of changing household amenities by running the model for 2018 with amenities fixed at 1990 levels; this leads to a serious attenuation of gentrification in the city. Finally, I freeze firms to their 1990 distribution throughout the city and allow households to sort. I find that gentrification, measured by appreciation of the logarithm of the ratio between college- and non-college-educated individuals, decreases by an average of 17% across zip codes in the city. Using the definitions provided by Freeman (2005), a commonly accepted threshold for gentrification, this leads to a 9.5% reduction in the number of zip codes gentrifying over the study period. I compare these results with those of Seattle; while firm-sorting appears to have a larger contribution to gentrification in Seattle than Portland, both cities exhibit similar impacts of the process.

This paper is related to the literature on gentrification in urban environments. Past work has proposed myriad causes for the process. Su (2020) and Edlund, Machado, and Sviatschi (2015) study the increasing prevalence of long working hours and rising values of time among college-educated young professionals; they hypothesize that these individuals may choose to relocate downtown, closer to their jobs, to cut down on commute time. Other papers focus on amenity growth, such as from reductions in crime (D. Autor et al., 2017; Ellen et al., 2019) and preferences over neighbors (Baum-Snow, 2007; Guerrieri, Hartley,

& Hurst, 2013). The replenishment of central cities' aging housing stock has also been explored (Myers & Pitkin, 2009). Some papers have also investigated the connection between public policies such as rent control (Diamond, McQuade, & Qian, 2019a, 2019b) and placed based opportunity zones Hoelzlein (2019).

I focus on the role of firm sorting in particular as a source of relocation incentive to affluent college-educated households. While the importance of firm sorting has been noted throughout the literature, very little work has been done to structurally account for its impacts. To this point, most papers in the literature have relied on reduced-form endogenous amenity terms to model the process (Bayer et al., 2004; Card et al., 2008; Diamond, 2016; Handbury, 2019; Su, 2020). Often, the rise in the ratio of college- to non-college-educated households is associated with the growth of non-tradable firms such as restaurants, bars, and gyms as well reductions in crime, improvements in school quality, and other such amenity changes. Endogenous amenities appear to play a significant role in gentrification; Su (2020) in particular finds that holding city amenity levels constant at pre-gentrification levels results in a 45% decrease in gentrification in US cities. While useful for the sake of simplicity, these reduced-form terms do not allow researchers to separate out the contributions of individual amenity factors. By explicitly modeling the decisions of firms in a city, I am able to back out the contribution of firm resorting in gentrification via counterfactual simulation.

This paper also contributes to the literature on firm decision making in spatial environments. The model I develop draws on recent work using Dixit-Stiglitz monopolistic competition in such environments to study how firms compete for households money and labor (Suárez Serrato & Zidar, 2016; Tsivanidis, 2019). I go further than these papers, however, by differentiating between tradable and non-tradable goods firms. Tradable goods can generally be purchased near costlessly across space; firms producing these goods need not locate close clientele to win their income. Non-tradable firms, such as bars, restaurants, gyms, and galleries, must be commuted to by individuals wishing to purchase their services, meaning that these firms need to be accessible to their customer base; this provides them incentive to sort towards populations of affluence. Handbury (2019) notes that non-tradable firms, as a result, tend to cater to the tastes of nearby residents. This, in turn, provides incentive for more households that favor those goods to locate in those neighborhoods, causing even more firm resorting, and so on. By considering the decision making of differently-tradable firms, I demonstrate not only the impact of firm sorting on households but also how different sectors respond to changes in cities.

My paper is most related to Hoelzlein (2019), which posits a structural model of Los Angeles with household and firm co-sorting. Households are specifically assumed to possess non-homothetic preferences across monopolistically competitive firm goods, in line with recent work by Couture and Handbury (2017) and Handbury (2019). Under counterfactual simulation, the model shows that place based policies that encourage firm growth in disadvantaged areas lead to twice as many college-educated households moving into these neighborhoods. There is also evidence of significant interdependence in the locational decisions of households and firms. I confirm this interdependence in an environment with homothetic preferences, demonstrating that similar incentives exist under constant returns to scale. That is, I show that homothetic preferences are not necessary to explain firms' role in gentrification.

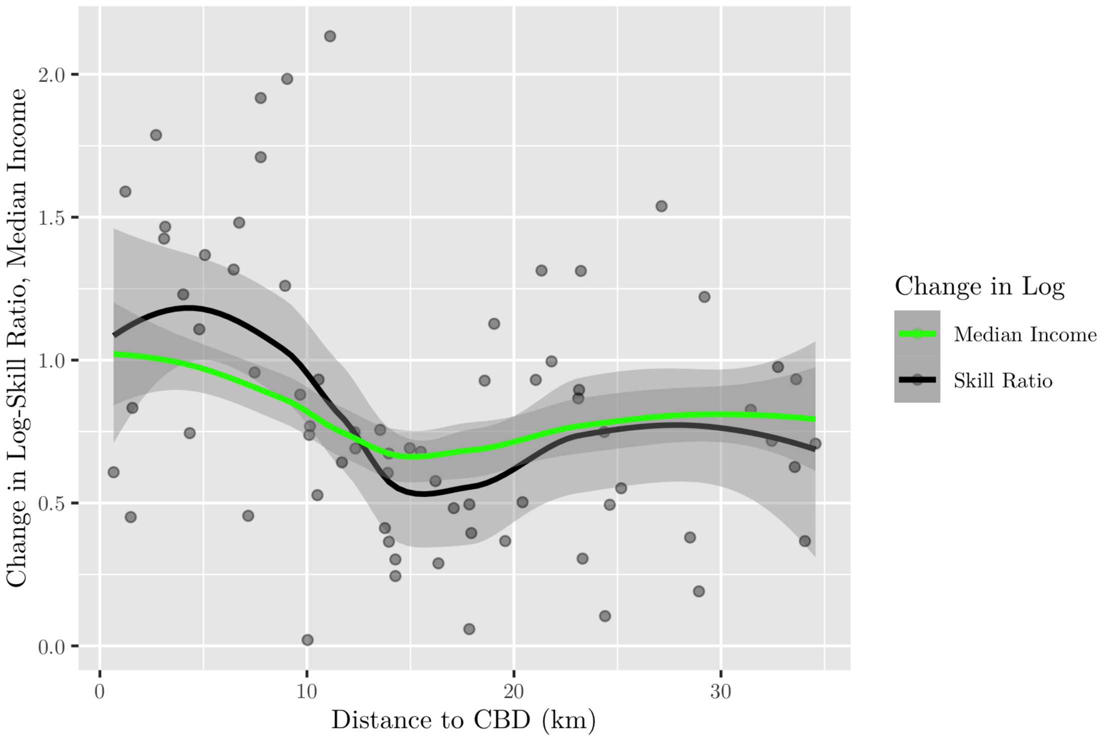
The rest of this paper is organized as follows: Section 3.1 documents trends in Portland associated with gentrification; Section 3.2 briefly outlines the key details of the model;[1] Section 3.3 details the data I use; in section 3.4, I discuss how I calibrate the model; Section 3.5 decomposes gentrification in Seattle by parameter and amenity changes and firm sorting; Section 3.6 concludes.

3.1 Descriptive Facts

3.1.1 Gentrification in Portland. Figure 13 plots two common measures of gentrification for the city of Portland. The black line in this figure represents the change in the logarithm of the ratio of college to non-college educated individuals between the years of 1990 and 2018. Greater appreciations of this ratio are considered to be evidence of more drastic gentrification. The curvature of this line suggests that college-educated households are not sorting uniformly, even as their overall relative mass increases compared to non-college educated households; rather, they appear to favor sorting towards downtown areas near the city center. This is a trend noted in many other papers on gentrification (Edlund et al., 2015; Su, 2020). The green line in figure 13 depicts changes in median income between the study years. Clearly, changes in affluence in the city are following changes in the log-skill ratio; the two have a Pearon's R of 0.838, suggesting a high degree of correlation. The appeal of the log-skill ratio as a measure of gentrification is thus demonstrated, as it appears to summarize multiple changes associated with gentrification (such as rising affluence). Figure 13 suggests that affluence in the city is increasing overall, but disproportionately towards the city center.

[1] A full treatment of the model is given in the first chapter in this dissertation.

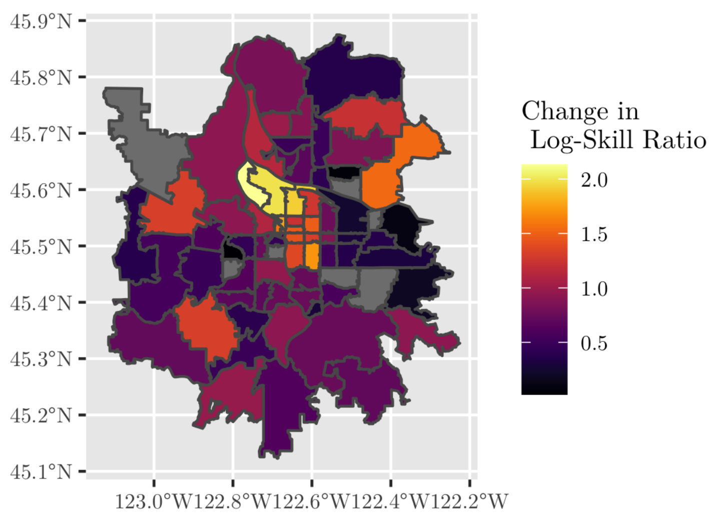
Figure 13.: Change in Log of Skill Ratio and Median Income by Distance to CBD



Dots are observed changes in the log-skill ratio. Curves fit using a LOESS process. Data comes from the US Census Bureau, 1990-2018, accessed via NHGIS.

Figures 14 and 15 provide geographical documentation of two alternative measures of gentrification in Portland. Specifically, figure 14 shows the change in the log-skill ratio by zip code; the appreciation of this ratio is most pronounced in the central downtown area, as indicated in figure 13. These areas correspond to the Portland neighborhoods known as the Pearl District, Elliot, Downtown, and Lloyd. Other scattered hot spots exist throughout the city's more suburban neighborhoods towards the edge, but it is clear that gentrification is centered in these neighborhoods in terms of greatest intensity. Figure 15 depicts gentrification using the methodology of Freeman (2005); this methodology treats gentrification as a binary outcome, and so it gives some sense of the "extensive" margin of gentrification.[2] Once again, gentrification appears to be clustered in the central downtown area, with a few outlying areas also experiencing the transition. Of all 24 zip codes eligible, 21, or 88% actually gentrify. This places Portland among the fastest gentrifying cities in the US during this time period (Macaig, 2015).

Figure 14.: Change in Log-Skill Ratio



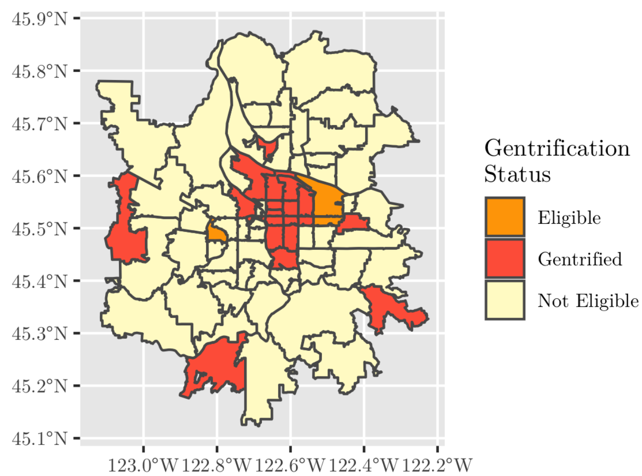
Map filtered to include only zip codes with centroids within 35 kilometers of central business district. Data comes from the US Census Bureau, 1990-2018, accessed via NHGIS.

3.1.2 Firm Sorting.

For the purposes of this paper, I aggregate 19 of the 24 North American Industrial Classification System (NAICS) sectors into four

[2] Specifically, Freeman (2005) uses the following criteria to classify a given neighborhood as having been gentrified over the period between two decennial censuses: "1) Be located in the central city at the beginning of the intercensal period; 2) Have a median income less than the median (40th percentile) for that metropolitan area at the beginning of the intercensal period; 3) Have a proportion of housing built within the past 20 years lower than the proportion found at the median (40th percentile) for the respective metropolitan area; 4) Have a percentage increase in educational attainment greater than the median increase in educational attainment for that metropolitan area. 5) Have an increase in real housing prices during the intercensal period" (Freeman, 2005).

Figure 15.: Map of Gentrified zip Codes using Freeman (2005)



Eligibility and gentrification status using the definitions laid out in Freeman (2005). Data comes from the US Census Bureau, 1990-2018, accessed via NHGIS.

broad super-sectors similarly to Eckert, Ganapati, and Walsh (2019). Each of these super-sectors is labeled with a simple shorthand that records whether the super-sector is high- or low-skill labor intensive (H versus L) and a tradability (tradable t versus non-tradable nt). Table 5 documents which NAICS sectors I assign to each super-sector. I use these super-sectors throughout this paper for convenience. I omit the NAICS sectors for Agriculture, Forestry, Fishing and Hunting (11), Mining, Quarrying, and Oil and Gas (21), Extraction, Utilities (22), Construction (23), and Public Administration (92) from consideration.

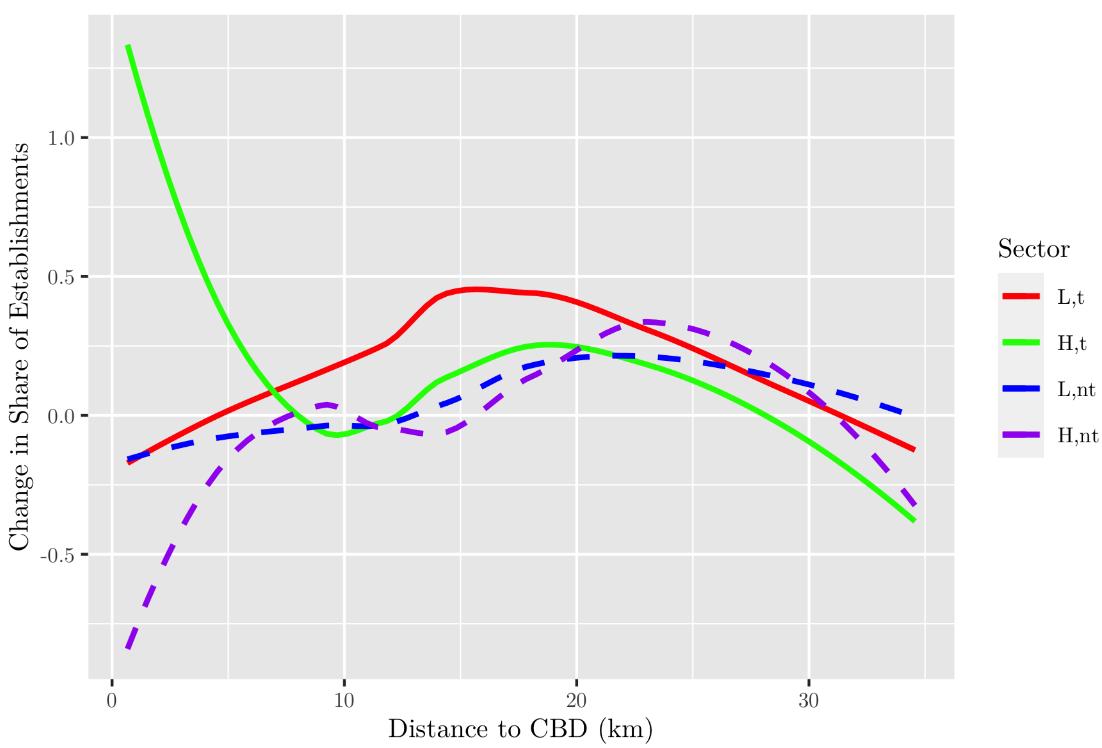
Figure 16 depicts changes in the distribution of these super-sectors by distance to the city center; specifically, it shows how the share of firms in each super-sector has shifted between the study years. A number of previously mentioned papers (Couture, 2013; Couture et al., 2019; Couture & Handbury, 2017) have found that the sorting of non-tradable firms towards downtown areas

Table 5.: Division of NAICS Sectors by Tradability and Skill Level

	Tradable	Non-Tradable
High-Skill	Information (51)	Education (61)
	Finance, Insurance (52)	Health Care (62)
	Real Estate (53)	
	Professional Services (54)	
	Management of Companies (55)	
Low-Skill	Manufacturing (31-33)	Admin. Support, Waste (56)
	Wholesale Trade (42)	Arts, Entertainment, Recreation (71)
	Transportation, Warehousing (48-49)	Accommodation, Food Services (72)
		Retail (44-45)
		Other Services (81)

reflects changing patterns of affluence associated with gentrification, a trend which I documented in the second chapter of this dissertation for Seattle, WA. The logic is that, since non-tradable firms need to locate near their clientele in order to serve them consistently, these firms would wish to capitalize on rising incomes in downtown gentrifying areas. However, in Portland, it appears that the downtown hot spot of gentrification is an area these firms are moving out of. In particular, the high-skill intensive non-tradable sector (in purple) makes a pronounced dip, indicating that these firms have been actually sorting *away* from the gentrifying city center. Instead, all 4 super-sectors of firms seem to be shifting towards an area about 20 kilometers outside of the city center. Figure 13 shows that there is a marked increase in the change in the log-skill ratio and log-median income at about this distance from the city center, suggesting that these firms may be still reacting to shifting patterns of affluence and skill in the city. The remarkable increase in the share of high-skill intensive tradable firms downtown is also worth noting, as firm sorting theories would predict that these firms would be unaffected by shifting incomes (since their goods can be accessed equally by all households regardless of location). This may reflect growth in the tech sectors and other professional services in downtown, which was particularly pronounced during the decade of 2000-2010, in part due to the influence of downtown universities such as Portland State University (Mayer, 2005, 2006). Some recent work has suggested that high-skill households may sort centrally to cut down on time spent commuting to work (Costa, 2000; Edlund et al., 2015; Su, 2020); the centralization of these high-skill tradable firms may thus provide incentive for gentrifiers to move closer to the central business district beyond considerations of access to consumption opportunities

Figure 16.: Change in Share of Total Establishments by Sector



All curves fit using a LOESS process. 'H,nt' refers to the high-skill non-tradable sector, and so on. Data comes from the US Census Bureau's zipcode Business Patterns Survey, 1994-2017.

3.2 Model

Here, I outline the key features of my model; a more step by step development of the model can be found in the second chapter of this dissertation. I assume that the city is divided into a set of discrete neighborhoods and an outside option indexed ℓ and is populated in each time period by a mass of households and firms who optimize statically. Households i are exogenously assigned a type $e \in \{H, L\}$, high- or low-skill (corresponding to having completed a bachelor's degree versus not), and must decide where in the city to live and work to maximize utility. Firms are assigned a sector s and a tradability t (tradable) or nt (non-tradable) and seek to locate themselves to maximize profit. In the model, non-tradable goods must be traveled to in order to be purchased, incurring an iceberg "shopping cost," whereas tradable goods can be purchased for the same cost regardless of location. Both firms and households choose where they live in reaction to the other; households want to locate near non-tradable firms to cut down on consumption costs and (non-tradable) firms wish to locate near households to increase the market size for their products. Rents and wages are set to clear the market competitively.

3.2.1 Households.

For a household i of type $e \in \{H, L\}$ living in location ℓ and working in ℓ', utility is given by

$$u_i^e(\ell, \ell') = K_i^{\alpha^e} \left(X_i^{t\,\beta} X_i^{nt\,1-\beta}\right)^{1-\alpha^e} \cdot \exp\left(\xi^e(\ell) + \sigma_h^e \varepsilon_i^h(\ell)\right)$$

where K_i is the amount of housing purchased by the household i, X^t is a Dixit-Stiglitz aggregator of goods from tradable firms $j \in \mathcal{J}^t$,

$$X_i^t = \left(\int_{\mathcal{J}^t} x_{ij}^{\frac{1+\zeta}{\zeta}} dj\right)^{\frac{\zeta}{1+\zeta}},$$

and X^{nt} is an analogous aggregator for non-tradable goods. $\xi^e(\ell)$ are neighborhood amenities and $\varepsilon_i^h(\ell)$ is an idiosyncratic neighborhood taste shock for households drawn from a Gumbel distribution with scale σ_h^e. α^e is the budget share of housing for type-e households, β is the consumption share of tradable, and ζ is the price elasticity of demand for each good. The household faces the budget constraint

$$R(\ell)K_i + \int_{\mathcal{J}^t} p_j x_{ij} dj + \int_{\mathcal{J}^{nt}} p_j x_{ij} \tau(\ell, \ell_j) dj = \Big(\sum_s \pi_s^e w_s^e(\ell') \Big) \exp\left(\sigma_w^e \varepsilon_i^w(\ell') \right) \quad (3.1)$$

$R(\ell)$ is rent in ℓ, p_j is the price of firm j's output, and $\tau(\ell, \ell_j)$ is the iceberg shopping cost associated with traveling from location ℓ to ℓ_j (the location of j). Household income is composed of a base income that is a weighted sum across sectors in the household's place of work ℓ' and a multiplicative productivity shock $\varepsilon_i^w(\ell')$ which is drawn from a Gumbel distribution with scale σ_w^e.[3]

I assume, as in Tsivanidis (2019), that households receive their work productivity shocks after having decided where to live, such that they choose where to live in expectation of their productivity shocks. Conditional on income, the probability of working in ℓ' is given by

$$\pi_w^e(\ell') = \frac{\Big(\sum_s \pi_s^e w_s^e(\ell') \Big)^{1/\sigma_w^e}}{\int \Big(\sum_s \pi_s^e w_s^e(\ell'') \Big)^{1/\sigma_w^e} d\ell''}$$

Knowing this, households will locate them selves optimally to maximize their expected indirect utility,

$$\begin{aligned} V_i^e(\ell) =& \frac{1}{\sigma_h} \Bigg(\int \pi_w^e(\ell') \cdot \ln \Big(\sum_s \pi_s^e w_s^e(\ell') \Big) d\ell' \qquad (3.2) \\ & - \alpha^e \ln R(\ell) - (1-\alpha^e)(1-\beta) \ln P^{nt}(\ell) + \xi^e(\ell) \Bigg) + \varepsilon_i^h(\ell) \\ =& \tilde{V}^e(\ell) + \varepsilon_i^w(\ell') \end{aligned}$$

[3]In detail, $w_s^e(\ell')$ is the wage for sector s for e households in ℓ' and π_s^e is the probability of working in sector s conditional on being skill level e. I assume that π_s^e is exogenous.

Where

$$P^{nt} = \left(\int [p_j \tau(\ell, \ell_j)]^{1+\zeta} dj \right)^{\frac{1}{1+\zeta}}$$

is the Dixit-Stiglitz price aggregator. This price aggregator is increasing both in the price of individual goods and also the distance, such that households have an incentive to live near tradable firms to increase utility. Following McFadden (1973), we have that the probability that the household chooses to live in ℓ is given by

$$\pi^e(\ell) = \frac{\exp(\tilde{V}^e(\ell))}{\int \exp(\tilde{V}^e(\ell''))d\ell''}.$$

Households can also choose to live in the outside option, which yields a constant utility of 0; this corresponds to living in the surrounding area rather than downtown.

3.2.2 Firms. As mentioned previously, I assume that firms are divided into four sectors: high-skill tradable, low-skill tradable, high-skill non-tradable, and low-skill non-tradable. Sectors s are endowed. Firms hire labor, produce their good, and set price, competing monopolistically. Each firm j produces a differentiated product y_j subject to production function $y_{js} = B_s(\ell)\mathcal{K}_j^{\kappa_s}\mathcal{L}_{js}^{1-\kappa_s}\exp(\sigma_\epsilon \epsilon_j(\ell))$, a simple Cobb-Douglas function with a multiplicative productivity shock where $\mathcal{K}$ is land and

$$\mathcal{L}_s = \left(\theta_s H^{\frac{\rho-1}{\rho}} + (1-\theta_s)L^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}}$$

is a constant elasticity of substitution aggregator of labor types with a factor share θ_s that varies by sector; ρ is the elasticity of substitution between labor types. $\epsilon_j(\ell)$ is a neighborhood specific idiosyncratic productivity shock drawn from a Gumbel distribution with scale σ_ϵ.[4] Firms in the non-tradable sectors face a different problem than those in tradable sectors because households must pay

[4]You guessed it.

shopping costs to purchase their good; this means that households are more likely to purchase from nearby non-tradable firms and hence these firms have incentive to locate nearby their customer base. A household located at ℓ_i will demand $x_j = (1-\alpha^e)(1-\beta)w_i(p_j\tau(\ell_i,\ell_j))^{\zeta}P(\ell_i)^{-(1+\zeta)}$ goods from a non-tradable firm in ℓ. The non-tradable firm therefore solves

$$\max_y y^{\frac{1+\zeta}{\zeta}}\mathcal{H}^{nt}(\ell)^{-\frac{1}{\zeta}} - R(\ell)\mathcal{K} - w^H(\ell)H - w^L(\ell)L \tag{3.3}$$
$$\text{s.t. } y = B(\ell)\mathcal{K}^{\kappa}\big(\theta_s H^{\frac{\rho-1}{\rho}} + (1-\theta_s)L^{\frac{\rho-1}{\rho}}\big)^{\frac{\rho(1-\kappa)}{\rho-1}}\exp(\sigma_\epsilon\epsilon_j(\ell))$$

where H and L are the amounts of labor bought from high- and low-skill households respectively, and

$$\mathcal{H}^{nt}(\ell) = (1-\beta)\sum_e(1-\alpha^e)\int N^e(\ell')w^e(\ell')\tau(\ell,\ell')^{\zeta}P(\ell')^{-(1+\zeta)}d\ell'$$

aggregates household demands across locations to capture market size. Solving 3.3, plugging in optimal values of output, land, and labor, and taking logs gives the non-tradable firms' indirect profit function:

$$\Pi^{nt}_{js}(\ell) = \frac{1}{\sigma_\epsilon}\left(\ln B(\ell) - \frac{1}{1+\zeta}\ln\mathcal{H}^{nt}(\ell) - \kappa_s\ln R(\ell) - (1-\kappa_s)\ln c(w^L_s,w^H_s)\right) + \epsilon_j(\ell) \tag{3.4}$$
$$= \tilde{\Pi}^{nt}_s(\ell) + \epsilon_j(\ell)$$

where $c(w^L,w^H) = (\theta^{\rho}w^{H^{1-\rho}} + (1-\theta)^{\rho}w^{L^{1-\rho}})^{\frac{1}{1-\rho}}$ is the unit cost function for the CES labor aggregator. Note here that generally $\zeta < -1$, meaning that Π^{nt}_j is increasing in market size $\mathcal{H}^{nt}(\ell)$, reflecting that non-tradable firms can increase profit by moving closer to residents, particularly affluent ones, *ceterus paribus.*

tradable firms have a slightly easier problem; they solve

$$\max_y y^{\frac{1+\zeta}{\zeta}}\mathcal{H}^{t^{-\frac{1}{\zeta}}} - R(\ell)\mathcal{K} - w^H(\ell)H - w^L(\ell)L \tag{3.5}$$
$$\text{s.t. } y = B(\ell)\mathcal{K}^{\kappa}(\theta_s H^{\frac{\rho-1}{\rho}} + (1-\theta_s)L^{\frac{\rho-1}{\rho}})^{\frac{\rho(1-\kappa)}{\rho-1}}\exp(\sigma_\epsilon\epsilon_j(\ell)),$$

the only difference being that $\mathcal{H}^t$, analogous to $\mathcal{H}^{nt}(\ell)$ for non-tradable firms, is not location dependent, reflecting that households do not pay shopping costs on tradable goods. The resulting indirect profit function for tradable firms is

$$\Pi^t_{js}(\ell) = \frac{1}{\sigma_\epsilon}\Big(\ln B(\ell) - \kappa_s \ln R(\ell) - (1-\kappa_s)\ln c(w^L_s, w^H_s)\Big) + \epsilon_j(\ell) \tag{3.6}$$

$$= \tilde{\Pi}^t_s(\ell) + \epsilon_j(\ell). \tag{3.7}$$

The only difference between the indirect profit functions 3.4 and 3.6 is the omission of market size $\mathcal{H}^t$ from 3.6.

Since ϵ is distributed Gumbel, the probability that a firm will locate in a given neighborhood is

$$\pi^t_j(\ell) \equiv \Pr(\ell = \ell_j) = \frac{\exp(\tilde{\Pi}_s(\ell))}{\int \exp(\tilde{\Pi}_s(\ell'))d\ell'}.$$

3.2.3 Equilibrium. I model the market for labor as being competitive. This means that the income promised to workers of each skill type, $w^e(\ell)$, must be enough to attract workers of each type to meet the total labor demand in location ℓ. For example, $w^L_s(\ell)$ must solve

$$\int\int L(y_j, w^L_s(\ell), w^H_s(\ell), R(\ell))djds = \int N^L(\ell')\frac{(w^L(\ell'))^{1/\sigma^L_w}}{\int (w^L(\ell''))^{1/\sigma^L_w}d\ell''}d\ell', \tag{3.8}$$

which is simply the equating of the demand for (left-hand side) and supply of (right-hand side) labor. A similar condition holds for high-skill labor. Note here that, with the denominator on the right hand side of the equation, each location in the city is competing for households' labor with every other location in the city.

I assume that rents are formed via an isoelastic formula that aggregates household and firm demand for land in ℓ

$$R(\ell) = \left(\int_{\mathcal{I}(\ell)} K_i di + \int_{\mathcal{J}(\ell)} \mathcal{K}_j dj\right)^{\psi} \tag{3.9}$$

where ψ is the inverse elasticity of supply for land in city m and $\mathcal{I}(\ell)$, $\mathcal{J}(\ell)$ are the sets of households and firms respectfully locating at ℓ. I omit the standard multiplicative constant term in the formula to keep rents as simple as possible; my calibration method for amenities and total factor productivities makes its inclusion unnecessary.

An equilibrium in this model is a set of prices $\{w_s^L(\ell), w_s^H(\ell), R(\ell)\}_{s,\ell}$, distributions $\{h_s^L(\ell), h_s^H(\ell), f_s(\ell)\}_{s,\ell}$ and firm specific variables $\{p_j, y_j\}_j$ such that

1. Households choose their location to maximize indirect utility, and consume optimally at that location to maximize utility taking prices and the distributions of firms and other households as given.

2. Firms choose their location to maximize their indirect profit and produce optimally at that location to maximize profit taking the distributions of rents, wages, households, and other firms as given.

3. Wages are set according to equation 3.8.

4. Rents are formed are formed according to equation 3.9.

3.3 Data

I begin by defining several of the qualitative variables in 3.2. Ideally, I would define neighborhoods according to Portland's official neighborhood designations. However, spatial data are not available at this geographic level. The next best option would be the Census tract, but tracts are frequently redrawn between decennial censuses, making consistent spatial comparisons difficult; additionally, data on the distribution of firms is not available at the tract level. Instead, I choose to treat zip codes as neighborhoods in the city. zip codes tend to be more stable over time, and also have much better firm data. I define the boundaries of the city to be all zip codes with geographic centroids within 35 kilometers of the coordinates of the downtown Portland pin on Google Earth. The remainder of the Portland MSA I lump into the outside option. I define an individual as being high-skill if they have a Bachelor's degree or higher and sectors as in table 5.

I obtain boundary files for the zip codes in the study as well as counts of individuals of each skill level from the tidycensus package in R, which provides data from the Census' American Community Survey is an easy to access format. I include all individuals in the MSA despite only explicitly modeling the inner 35 kilometers of the city, allowing me to account for net in- or out-migration under each tested counterfactual. Individuals living in the MSA but not the main city are considered to be occupying the outside option. I remove all individuals of less than 25 years of age. Table 6 provides the resulting population counts by skill level and year. I turn to the National Historic GIS (NHGIS) database for data on median income, housing values, and housing construction dates to determine the gentrification status of each neighborhood in figure 15; I also use these to help determine the extensive margin of gentrification under the counterfactual.

Table 6.: Populations by Skill Level

Skill	2018	1990	% Change
High	689,672	244,871	182%
Low	1,139,002	843,196	35%

Data from US Census Bureau.

For data on employment by sector, I use microdata from the Integrated Public Use Microdata Survey (IPUMS), a national annual sample of 1% of the US population. After filtering down to the Portland MSA, I can compute the probability of being employed in each NAICS sector conditional on skill level;

I then aggregate these probabilities up to the super-sector level as defined in table 5. The IPUMS file for 1990 uses an older sectoral classification system, the Standard Industrial Classification (SIC). A complete reconciling of SIC and NAICS is impossible; nonetheless, I convert as many SIC codes from individuals in 1990 as possible to NAICS. A few observations are lost in this process.

I turn to the Census Bureau's zip Code Business Patterns Survey (ZBP) for data on firm counts. This data reports the number of firms by sector and number of employees at the zip code level. As with individuals, I include all firms active in each study year within the entire MSA even though I restrict spatial analysis to the central city. Table 7 reports establishment counts at the aggregated sector level, the share of high- and low-skill individuals employed by each, and the changes in these numbers over time. Note that I use the 2017 ZBP for 2018 and the 1994 ZBP for 1990, as these are the closest years in which the data is available. Once again, the data from the 1994 uses SIC codes, and so I convert as many as possible to NAICS before aggregation; some codes are unable to be reconciled, and I omit them from consideration. This leads to what is probably a significant under-reporting of firm establishment counts in 1990, as reflected in the table.

Table 7.: Establishment Counts and Share of Workers by Skill-Level By Sector

	1990			2018		
Sector	Count	L-Share	H-Share	Count	L-Share	H-Share
L, t	654	0.36	0.26	9,401	0.26	0.15
H, t	463	0.31	0.27	19,551	0.15	0.34
L, nt	4,132	0.17	0.15	26,636	0.37	0.21
H, nt	382	0.16	0.31	9,956	0.22	0.31

Data from ZBP and IPUMS USA.

My calibration of the firm-side parameters relies on two data sets. To estimate the land share for each sector, $\{\kappa_s\}_s$, I turn to the Bureau of Economic Analysis' (BEA) data on gross value added, employment, and worker compensation by sector.[5] For $\{\theta_s\}$, I use data from the Census' Current Population Survey (CPS), accessed via IPUMS, which provides records of individuals' wages, hours worked per week, and weeks worked per year by industry.

Finally, I obtain travel times between zip codes from the Google Travel Time Matrix API. This API allows researchers to obtain estimates for distance and travel time between locations via requests to Google Maps' directions feature. I assume that locations within a given zip code can be decently approximated by the zip code's centroid and let commute times between any two neighborhoods be given by the travel time between their centroids. I specifically select "driving" as the mode of travel and do not model travel times accounting for traffic. Unfortunately only modern day travel times are available; I use these travel times for both 1990 and 2018.

3.4 Calibration

The vast majority of the calibrations I use in this paper are the same as in chapter two of this dissertation. I set household housing budget shares $\alpha^L = 0.337$ and $\alpha^H = 0.321$ based on the Bureau of Labor Statistic's estimate of personal expenditure by education (Foster, 2014), and assume both skill levels consume the same share of tradable goods, $\beta = 0.4$ Bems (2008). I set the price elasticity of demand for all goods equal to $\zeta = -2.5$ as in Head and Mayer (2014) and Suárez Serrato and Zidar (2016). I assume that the iceberg shopping cost function

[5] The BEA itself obtains this data from the Decennial Census, ACS, and Economic Census carried out by the Census, which is then aggregated by industry and year. See https://www.bea.gov/data/employment/employment-by-industry.

$\tau(\ell, \ell') = 1 + \tau \times$ travel time and set $\tau = 0.5$, reflecting a 50% markup in cost per hour of travel to non-tradable firms. Intuitively, this means that an individual would be indifferent between visiting an establishment immediately outside their residence and an identical establishment an hour away if the latter were half as expensive. I set household preference shock standard deviations to $\sigma_h^H = 0.7950$ and $\sigma_h^L = 0.4593$ and productivity shocks to $\sigma_w^H = 0.4397$ and $\sigma_w^L = 0.3031$, again drawing on calibrations from Su (2020) and Tsivanidis (2019).

On the firm side, I repeat the calculations in chapter two to obtain firm production parameters $\{\kappa_s, \theta_s\}_s$ for both 1990 and 2018 using simple calculations from the BEA and CPS data. Since I assess these parameters at the national level (so as to maximize the power of my estimates) there are no changes from the second chapter. For convenience, the calibration of these parameters is summarized in table 8. I also set the standard deviation of firm productivity shocks, σ^F, to 0.28 as in Suárez Serrato and Zidar (2016), and use Card (2009) to set $\rho = 1/0.7$ for all sectors.

Following Saiz (2010), I set the elasticity of rent with respect to land demand to $\psi = 0.935$.

Table 8.: Calibration of Land share κ_s and Skilled Labor share θ_s for NAICS Super-Sectors

	1990		2018	
Sector	κ	θ	κ	θ
L, t	0.421	0.342	0.509	0.452
H, t	0.368	0.547	0.365	0.735
L, nt	0.450	0.357	0.557	0.499
S, nt	0.187	0.626	0.195	0.702

With these parameters in hand, I am able to set about obtaining the remaining parameters of the model, household neighborhood amenities $\xi^e(\ell)$, and total factor productivities $B_s(\ell)$. Again, I use the same methodology as in chapter two to obtain these, which I briefly outline here. To obtain $\xi^e(\ell)$, I first compute mean utilities experienced by households in each neighborhood using

$$\tilde{V}^e(\ell) = \ln \pi^e(\ell) - \ln \pi_0^e$$

where $\pi^e(\ell)$ is the share of type-e households living in neighborhood ℓ and π_0^e is the share in the outside option (non-downtown MSA). This is similar to the simple shared utility case outlined in S. Berry et al. (1995). Likewise, I compute shared firm profit as

$$\tilde{\Pi}^{t/nt} = \ln \pi^{t/nt}(\ell) - \pi_0^{t/nt}$$

where notation is organized similarly. Under the factual simulation, the distribution of households and firms will follow sorting patterns generated from these indirect utilities and profits; therefore, to obtain $\xi^e(\ell)$ and $B_s(\ell)$, I can force the observed masses households and firms into each neighborhood and search for equilibrium to find the rents, wages, and firm prices that rationalize the observed equilibrium. I can then extract $\xi^e(\ell)$ and $B_s(\ell)$ by taking these obtained prices and manipulating the formulas for $\tilde{V}^e(\ell)$ and $\tilde{\Pi}^{t/nt}$ given in section 3.2.

3.5 Results

As with Seattle, the inclusion of amenities $\xi^e(\ell)$ and total factor productivities $B_s(\ell)$ allow my model to perfectly match the observed equilibrium. Therefore, shape-fitting comparisons are arbitrary, and I instead focus on counterfactual simulations. In section 3.5.1 I experiment with holding production parameters and amenities constant to determine their contribution to gentrification.

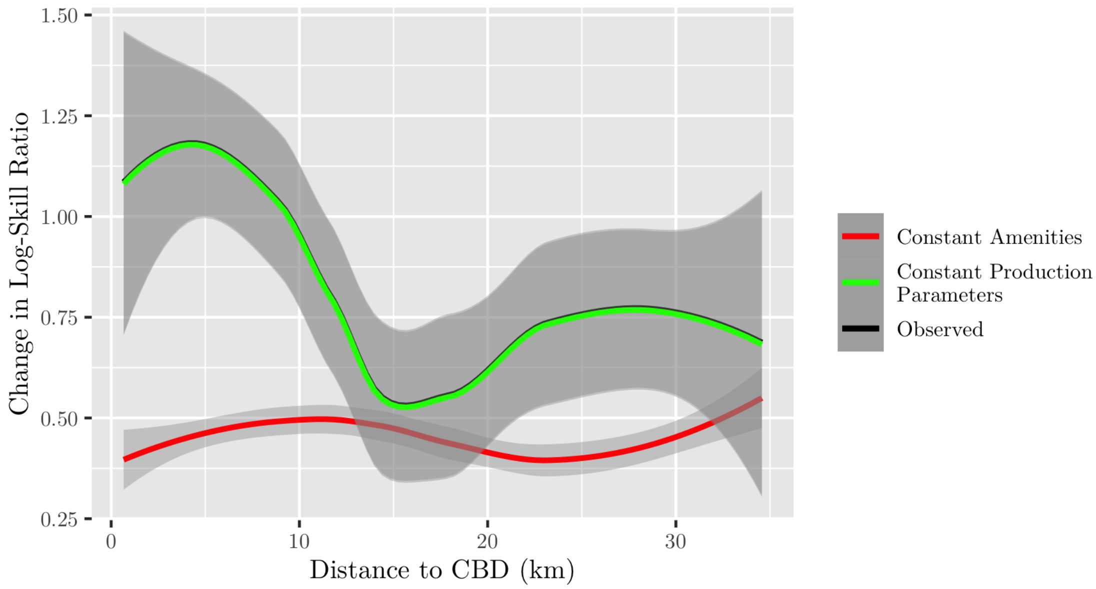
Then, in section 3.5.2, I demonstrate the contribution of firm sorting by holding fixed the distribution of firms in Portland and comparing the resulting equilibrium to the observed one. Finally, in section 3.5.3, I compare key differences between Seattle and Portland in these simulations.

3.5.1 Decomposition of Gentrification by Parameter Changes.

The first series of counterfactual simulations I run document the impacts of parameter changes between the study years of 1990 and 2018. On the firm side, these include the parameters κ_s, θ_s, and $B_s(\ell)$; on the household side, they include $\xi^e(\ell)$. Figure 17 documents the change in the log-skill ratio in the case that each set of parameters (on the firm and household side) are held constant. Holding production parameters constant leads to a minimal change from the observed equilibrium on the household side. The black line, which represents the observed equilibrium, is barely even visible behind the fixed parameter equilibrium in green. The production parameters do impact the ways that firms set prices and choose to locate but evidently not enough to significantly change the locational incentives for households. Overall, these firm side parameter changes over the study window seem to have had little impact on the extent of gentrification.

Holding neighborhood amenities $\xi^e(\ell)$ fixed, however, leads to a drastically different equilibrium with far less gentrification than the observed equilibrium. This scenario, represented by the red line in figure 17, results in an average decrease in the change in the log-skill ratio of about 8.6% across neighborhoods, and an average total decrease of 0.31. The importance of neighborhoods amenities is documented in Su (2020), who estimates that endogenous amenity changes account for about 45% of gentrification across US cities; his estimates lump firm-sorting into amenities, however, and so my estimate constitutes a decomposition of that

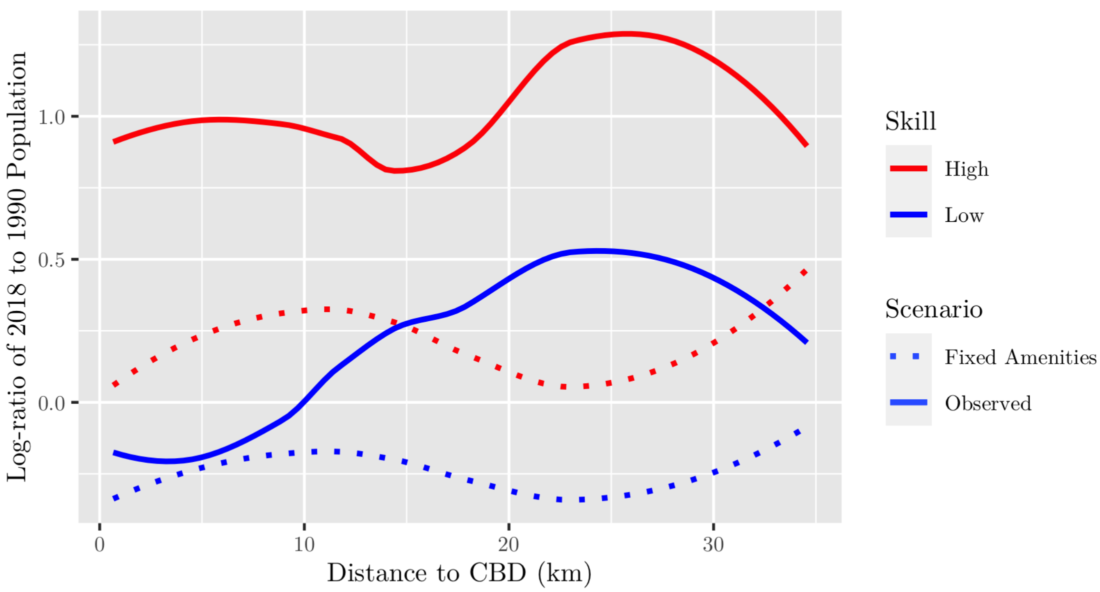
Figure 17.: Change in Log-Skill Ratio With Time-Varying Parameters Held Constant



Simulated changed in the log-skill ratio while holding indicated sets of parameters constant between 1990 and 2018. 95% standard error bands shown by shaded areas. All curves fit using a LOESS process.

estimate into non-firm based amenities. Amenities remain a significant driver of gentrification. A second portrayal of how amenities impact the distribution of households throughout the city can be found in figure 18, which presents the log-ratio of population per neighborhood by skill level and scenario from 2018 to 1990. An increase in this ratio means that relatively more households of a certain skill level are moving into an area. Under fixed amenities, the number of households of both skill levels decreases significantly, as both lines corresponding to the fixed amenities scenario lie entirely below their solid line counterpart (which represent the observed equilibrium). Apparently, amenity growth in the city has fueled net in-migration for both skill levels. Note, however, that the decrease in the central sorting of high-skill households is much more pronounced than it is for low-skill households under the fixed amenities scenario. Thus, while both skill levels experienced amenity growth between 1990 and 2018, high-skill households enjoyed far greater amenity growth, leading them to sort centrally. This is in line with work by Edlund et al. (2015) and Su (2020).

Figure 18.: Log-Ratio of 2018 to 1990, Observed and with Fixed Amenities



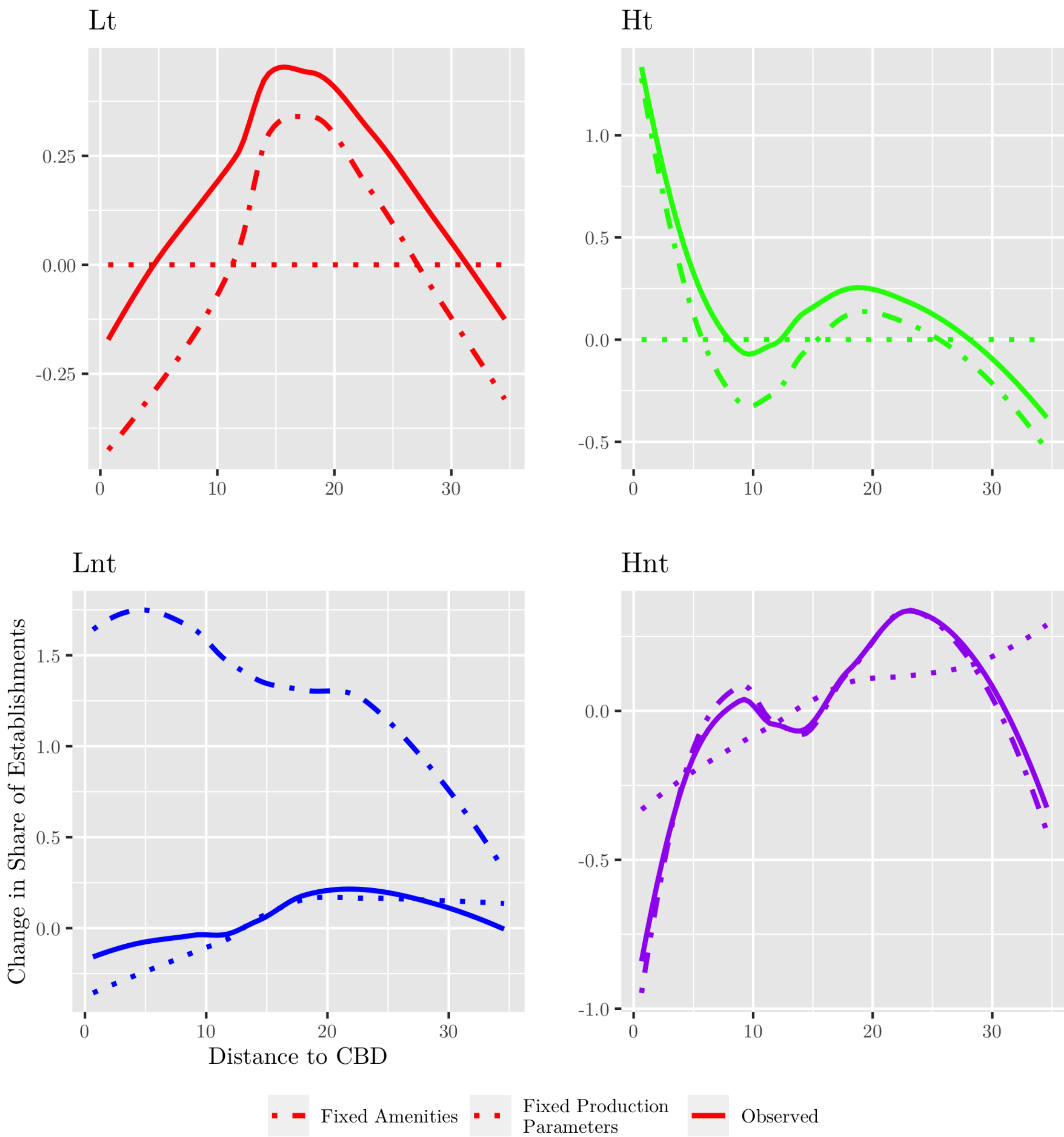
Log ratio of 2018 to 1990 populations under observed and fixed-amenities scenarios. All curves fit using a LOESS process.

Figure 19 documents changes in firm sorting under the fixed parameter scenarios. As would be expected, the fixed production parameter simulation lead to no difference in the sorting patterns of tradable firms. Recall that these firms are insensitive to the locations of their customers, and sort themselves based on rents, wages, and total factor productivities (which are held constant here). From figure 17, we know that households have not really adjusted their sorting patterns at all, and hence wages and rents are similar to the baseline scenario. In short, nothing has really changed for these firms between 1990 and 2018. Non-tradable firms do experience some resorting, albeit relatively minor in the case of low-skill intensive non-tradable. When I allow parameters of production to change however, and instead hold amenities constant, however, the shape of each each log-ratio curve more closely matches the observed scenario curve for non-Lnt firms. Curves for both types of non-tradable firms take on the shape of their observed scenario curves, shifted down slightly; this is a somewhat unexpected result, as figure 18 shows that net population is lower when amenities are fixed, meaning lower rents. I found a similar result for Seattle under the same scenario. For Hnt-firms, sorting is practically identical to the observed equilibrium under fixed amenities. However, for low-skill intensive non-tradable firms, the number of firms under the fixed amenities swells drastically, another pattern that is repeated from Seattle. This may be the result of more favorable labor market conditions, where relatively more low-skill households remain in the city than high-skill households relative to the observed equilibrium, leading to a lower wage differential between skill levels.

3.5.2 The Role of Firm Sorting. I now discuss the results of this paper's main simulation, which holds the relative distribution of firms in the city fixed at 1990 neighborhood shares. To do this, I use the profits computed in section 3.4 to generate sorting probabilities for firms of each sector. However, I allow the number of firms to increase to 2018 levels; I consider this a more "fair" simulation than holding the measure of firms to 1990 levels, because it allows firms to grow in proportion to population. This way, I can be sure that any changes in gentrification are from changing distributions of firms rather than new firms entering the market. I also allow wages, rents, and prices to adjust to instill equilibrium.

Figure 20 presents the results of fixing the distribution of firms to 1990 shares. The change in the log-skill ratio is on average lower than in the observed equilibrium, meaning that the relative relocation of firms had an average positive

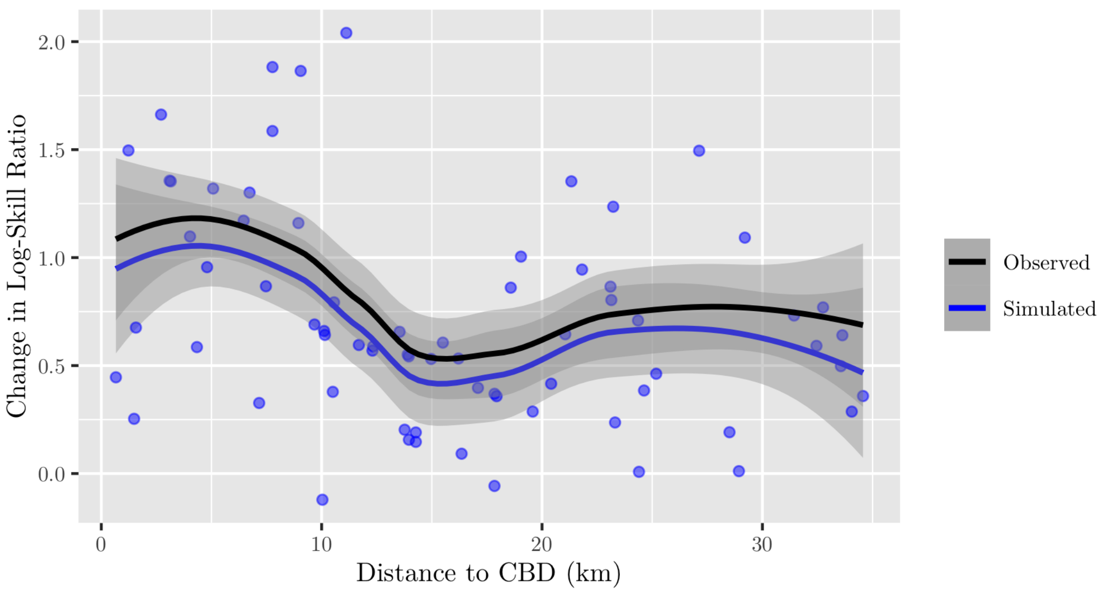
Figure 19.: Change in Distribution of Firms with Time-Varying Parameters Held Constant



Simulated and observed changes in firm distributions while holding indicated sets of parameters constant. All curves fit using a LOESS process.

impact on gentrification in the city. As with Seattle, the shape of the log-skill curve for the fixed firm scenario closely mirrors the shape of the observed equilibrium, but with a slight negative shift. On average, this exercise yields an increase in the

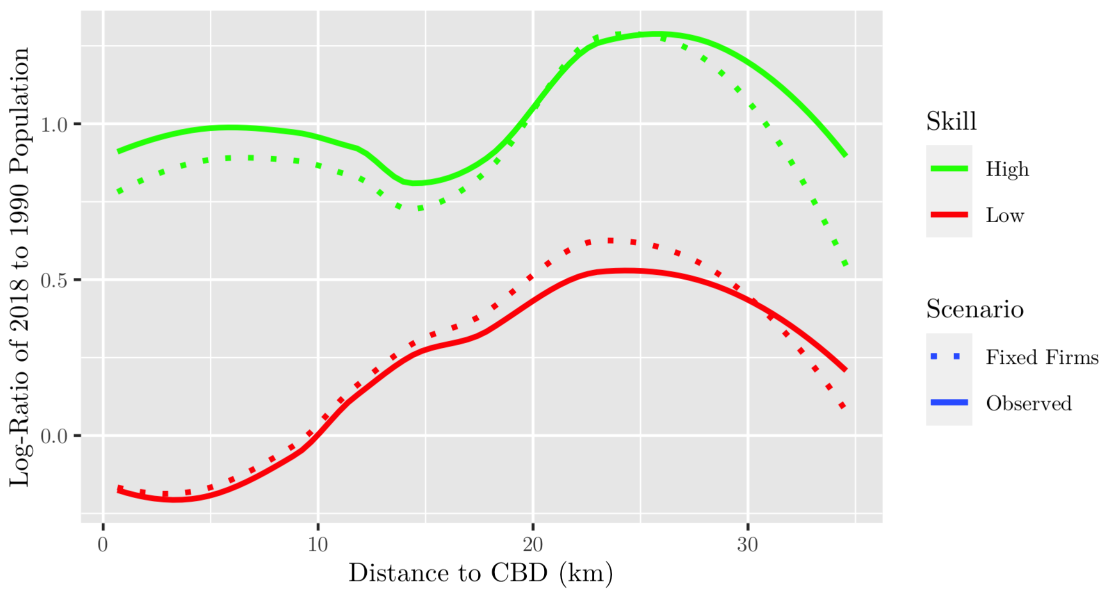
Figure 20.: Change in Log-Skill Ratio Holding Firm Distributions Constant



Observed and simulated change in log-skill ratio when firm sorting is shut down. Curves fit using a LOESS process.

log-skill ratio that is approximately 17% lower than the observed change in the log-skill ratio. I take this to be the average reduction in the intensive margin of gentrification. Overall, of the 78 zip codes included in the study, only four saw an increase in the log-skill ratio when firm sorting is fixed. Figure 21 plots the log of the ratio between 2018 and 1990 populations under the observed and fixed firm equilibrium. Here we see that high-skill individuals are generally less likely to sort into any given central city location when firms are fixed as they are under the observed equilibrium. Conversely, low-skill individuals are more likely to sort centrally under the fixed firm counterfactual. These two simultaneous patterns certainly explain the general decrease in the log-skill ratio when firm sorting is shut down. As in Couture (2013), it appears that the resorting of firms makes living downtown more attractive for high-skill households.

Figure 21.: Change in Log-Population Ratio Under Observed and Counterfactual Equilibria

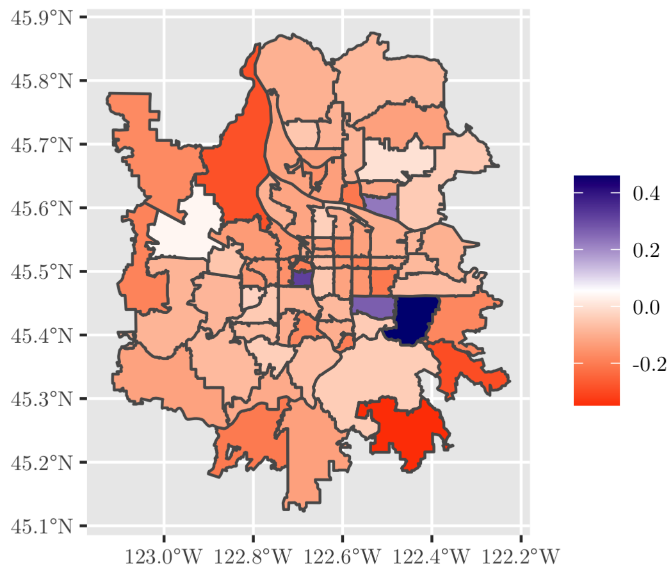


Observed and simulated log-ratio of population between 2018 and 1990 when firm sorting is shut down. Curves fit using a LOESS process.

Figure 22 maps the change in the change of the log-skill ratio under the fixed firm counterfactual. Reductions in gentrification are most intense in several zip codes around the edges of the city as defined, some of those areas in figure 15 that experienced gentrification. Interestingly, the few neighborhoods that experienced more gentrification under the counterfactual tend to be more centrally located; however, under the Freeman (2005) definition, none of these zip codes had low enough income in 1990 to be eligible for gentrification. These few zip codes have far greater populations of high-skill individuals under the counterfactual than they do in the observed equilibrium, suggesting that high-skill individuals are pooling into these limited locations.

Figure 23 presents gentrification results using the Freeman (2005) definition for the counterfactual. While nearly all neighborhoods saw a decrease in the

Figure 22.: Change in Change of Log-Skill Ratio Holding Firm Distributions Constant

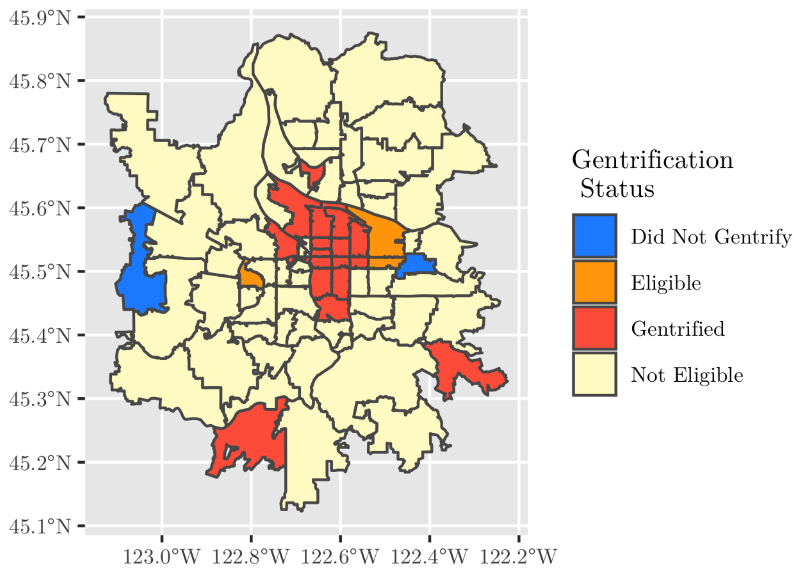


Difference in the change of the Log-skill ratio from the observed equilibrium to the no firm sorting counterfactual.

appreciation of the log-skill ratio, most neighborhoods that gentrified under the observed equilibrium still gentrified under the fixed firm counterfactual. Only two neighborhoods were actually prevented from gentrifying by shutting down firm sorting, or a 9.5% reduction in the extensive margin of gentrification.

3.5.3 Comparison of Firm Sorting Impacts in Portland and Seattle. In many ways, my analysis paints a very similar picture of gentrification and firm-sorting in both Seattle and Portland. Both cities exhibit minimal changes in the spatial sorting of households when firm-side production parameters are held constant, but drastically less sorting of high-skill households into the central city when amenities are held constant at 1990 levels. This is in line with the literature on amenities and gentrification (Couture, 2013; Couture & Handbury, 2017; Su, 2020) which establishes them as a significant source of gentrification over the last

Figure 23.: Gentrification Status of zip Codes with Firm Distributions Held Constant



Eligibility and gentrification status using the definitions laid out in Freeman (2005). "Did not gentrify" denotes neighborhoods that gentrified in the observed equilibrium that do not when firm sorting is held constant.

30 years. The main simulation in each, holding firm distributions fixed at 1990 levels, also result in similar outcomes; more low-skill individuals and fewer high-skill individuals sort centrally in the absence of firm resorting in both cities. Overall, this leads to a decrease in both the intensive margin (measured as appreciation in the log-skill ratio) and extensive margin (using the Freeman (2005) definitions) of gentrification in each city. The impacts of firm sorting on gentrification appear to be slightly larger in Seattle along both margins; the appreciation of the log-skill ratio declines by an average of 28% in Seattle and 21% fewer neighborhoods gentrify, as compared to 17% and 9.5% respectively in Portland. It should be expected that different cities, even in similar or common regions, will experience

gentrification differently. In both, however, I conclude that firm sorting had a significant positive impact on gentrification.

Some other qualitative similarities jump out as well. When looking at the figures mapping the change in the log-skill ratio, both cities exhibit mostly reductions in gentrification (red) with a few scattered neighborhoods in which gentrification increased when firm sorting is shut down. In both Seattle and Portland, most of these neighborhoods are not in the very center of the city, where gentrification is most peaked under the observed equilibrium. These pockets are areas that would have become centers of affluence sans firm sorting, as richer, more educated households moved in. The prominence of these pockets shows in particular that gentrification is not uniform, and not as simple as turning firm sorting "up or down." Rather, it is a spatial phenomenon with multiple overlapping causes; where firm growth may cause gentrification in one neighborhood, it may discourage it in another, depending on rents, wages, and amenities.

3.6 Conclusion

The application of the model I developed in chapter two of this dissertation to Portland has revealed many similarities in the patterns of gentrification between the two cities, even when many city specific factors such as neighborhood productivities and amenities, the relative masses of households and firms, and patterns of employment are recalculated. Portland shows less responsiveness to holding firm sorting constant, but the impact to gentrification is still significantly negative under this counterfactual: the increase in the log-skill ratio, the intensive margin of gentrification, declines an average of 17% from the observed outcome, and under common discrete definitions, 9.5% fewer tracks become gentrified.

It should be expected that the same model would find similar results between two cities that are in many ways common: they are the first and second largest in the Pacific Northwest of United States, separated by only 233 kilometers, experienced similar technological and sectoral growth during the period of study, and both experienced high levels of gentrification in the last 30 years. That the model does find differing results between them is evidence of the highly idiosyncratic process that gentrification is, not only in where and how intensely it occurs but also in the various factors that contribute to it in each location. An interesting application of this model would apply it to a novel qualitative context, such as a city outside the Pacific Northwest that has not yet experienced significant gentrification; several Mid-Western cities come to mind. The literature would also benefit from a more complete decomposition of gentrification in general, including not only firm sorting but also changes in factors like crime, housing age, and improving school quality. While this is among the first models to account for firm sorting in gentrification, it is designed to be open to future innovations and permutations, and I hope that it can be easily adapted to answer further questions on the process of gentrification.

CHAPTER IV

RENT CONTROL AND HOUSING MARKET TURNOVER IN SAN FRANCISCO

On-going housing affordability crises in many major US cities have renewed populist cries for rent control in recent years. In 2019, the entire state of Oregon implemented rent control, capping rent increases to 7% plus inflation statewide (Zaveri, 2019). California followed suit in 2020. Seattle, Washington, is another such hotbed, where rent control was polled at 71% support by residents in 2020 (Rosner, 2020). Rising rents in recent decades have made many such large cities unaffordable to lower-income residents. Meanwhile, the growing attractiveness of downtowns to young, college-educated professionals has made affordable housing especially scarce in central cities, raising fears of gentrification and residential displacement. Advocates of rent control hold that limiting the amount rents are allowed to rise each year would make housing more affordable to more people, allowing those otherwise displaced by gentrification to stay on in a community. However, some recent studies have found that rent control can actually lead to less affordable downtowns and increased gentrification (Diamond et al., 2019a, 2019b; Kholodilin et al., 2016) as landowners withdraw their properties from the market in the face of artificially low rents. Specifically, landowners may sell their rent-controlled structures as condominiums or owner occupied housing; where rent control ordinances exempt newly built structures, landowners may also choose to renovate their properties extensively such that they qualify as new housing. While these papers have modeled spatial sorting by households and have noted changes in the supply of rent controlled housing, an equilibrium model of rent control that includes landowner decisions has yet to be formulated in the literature. The durable nature of housing also presents an irreversibility problem; namely, older, more affordable housing units that are renovated to newer, less-affordable units cannot be converted back if rent control is repealed. Rent control thus permanently alters the state of a city's housing to be less affordable for unregulated structures. A complete understanding of landowner reactions to such regulation is therefore critical to achieving desirable long-term outcomes.

In this paper, I investigate the impact of the San Francisco, California 1979 rent control ordinance on housing market turnover in the decade between 2010 and 2019. I focus specifically on the conversion of rent controlled structures to condominiums, a process that is heavily regulated in the city. To do this, I outline a simple, general equilibrium model of household and landowner decision making under the ordinance. Households choose which neighborhood and what type of housing to occupy conditional on housing costs (rents and mortgages) and neighborhood-housing type amenities. Landowners are endowed with an initial structure with a size in terms of units and decide whether or not to convert to condominiums. Condominiums earn the landowners more income than rent controlled structures but are subject to conversion and tenant buy out costs. I calibrate the model using techniques similar to S. T. Berry (1994), S. Berry et al. (1995), and Grigolon and Verboven (2014), parameter values from the literature, and moment matching.

I use the calibrated model to simulate several policy counterfactuals. First, I simulate the repeal regulations of the city's 1983-2013 condominium conversion lottery, which limits the size of structures than can be converted to structures with only two units. I find that conversions increase by 80% and that rents appreciate on both unoccupied controlled and uncontrolled housing units; this latter result occurs because of the contraction in the supply of vacant rent controlled housing from the increased number of conversions. Next, I repeal the city's rent control ordinance in two separate simulations; the first restricts the size of structures that may be converted to be the same as under the condominium lottery; the second allows all structures of up to 6 housing units to convert to condominiums. The former exercise leads to a 330% increase in conversions and the latter a 647% increase. Both results stem from the flood of decontrolled housing units able to be priced at market rates. In line with other papers in the literature, I find that decontrol leads to significant rent depreciation among never controlled units despite the high turnover.

As a final exercise, I freeze rents exogenously to levels observed in the data and repeat the same battery of simulations. Here, I once again find the lottery restrictions are effective at disincentivizing conversion to condominiums. However, under both decontrol scenarios, conversions decrease with fixed prices, confirming that rent control does encourage housing market turnover in principle once price effects are nullified. I conclude that the impact of rent control on turnover and affordability in the real world depends on the opposing forces of increased supply of unoccupied housing and depression of rents by the policy, and that the impacts of such ordinances is likely idiosyncratic to each geography implementing one.

The consideration of how rent control incentivizes housing turnover is important to policy makers. Economic theory has long posited that artificial price ceilings (such as rent control) for goods will result in constrained supply if producers are not subsidized or coerced to continue producing; regardless, efficiency is decreased under such distortions. In the housing market, the contraction in supply comes from turnover. As previously noted, if rents are sufficiently depressed, rent control can provide incentives for landowners of controlled structures to sell their properties as privately owned condos, effectively reducing the rental housing stock in an urban center. Where regulations exempt new or refurbished structures, landowners may also renovate their structure so as to avoid depressed profits from rent control; this latter action not only reduces the stock of controlled housing but also probably increases rents since newer units tend to attract higher rent. Higher rent on uncontrolled structures can also occur as the supply of controlled structures contracts, causing more households to seek alternative (remaining) housing (Mense, Michelsen, & Kholodilin, 2019). Ostensibly, rent control should reduce housing costs, but here we see that there are many mechanisms resulting from housing turnover that can cause affordability to decrease.

Even outside of turnover, rent control can also cause rising rents and efficiency losses among even controlled units. Many modern rent control measures, such as in San Francisco and Oregon, also feature "vacancy-decontrol," which allows landowners to reset rents to market rates on vacated units. Policies with this feature effectively divide the supply of controlled housing into occupied and unoccupied units, though equilibrium price for new tenants is dictated by the relative supply and demand of only the unoccupied units. Since the supply of unoccupied, controlled units is smaller than the total number of controlled units, vacancy-decontrol can cause the rent on unoccupied units to appreciated faster that it would sans-rent control. Additionally, vacancy-decontrol can lead to misallocation as tenants stay inefficiently long in their units to avoid losing reduced rents (Glaeser & Luttmer, 2003).

This paper contributes to the literature on rent control and affordability, which has yet to structurally model landowner decision making. The question of whether rent control worsens or alleviates affordability issues and gentrification in a particular city critically depends on the reaction of housing suppliers. Most rent control policies feature exemptions for new housing (so as not to discourage construction) and some other structures (such as condominiums). Rent control can then cause those landowners with controlled structures to pull their property from the market, selling it as owner occupied housing, or renovate significantly enough that the structure becomes rent control exempt. These supply contractions can put upward pressure on the rents for other units, worsening affordability. Or, conversely, the contractions in supply of controlled housing may not be severe enough to worsen affordability. reduced-form studies return different results from different case studies: Sims (2007) and D. H. Autor, Palmer, and Pathak (2014) find that rent control succeeded in keeping rents low in Massachusetts, where as Kholodilin et al. (2016) find evidence of the opposite in Berlin. Thomschke (2016) also investigates the rent control policy in Berlin, and finds that not only does rent control not produce any meaningful long-term attenuation of rental prices, but that in fact price appreciation in newly offered units outpaces pre-reform rates. Mense et al. (2019) note that rent control can cause price appreciation amongst uncontrolled units, worsening affordability for those unable to obtain controlled housing. Like this second set of papers, I find that the repeal of rent control actually improves affordability both among formerly controlled and never controlled units.

My paper is most closely related to Diamond et al. (2019a), who estimate a structural model of household (though not landowner) decisions. They find that that the expansion of rent control in San Francisco in 1995 led to an increase in average rents of about 7% as landowners pulled 30% of eligible inventory from the market. Like this work, I find that rent control provides substantial incentive for landowners to convert their properties away from rent-controlled housing once I repeal restrictions on conversion provided by the condominium lottery. However, in contrast, I find that rent control actually discourages conversion by inflating rents on vacant controlled units. This work is also closely related to Murphy (2018), who posits a dynamic structural model of housing supply in San Francisco that similarly uses idiosyncratic profit shocks to landowners to model construction decisions. That paper finds that geographical and idiosyncratic costs are key to understanding the provision of housing across time, and that construction and conversion can be thought of as an optimal stopping problem.

The rest of this paper is organized as follows. Section 4.1 details the San Francisco rent control ordinance and documents the extent of condominium conversions in the city; Section 4.2 presents the model; Section 4.3 describes the data sources I use to calibrate the model in section 4.4; the results of each counterfactual simulation are reported and analyzed in section 4.5; Section 4.6 concludes.

4.1 Background: San Francisco, Housing Turnover, and Rent Control

4.1.1 San Francisco's Rent Control Ordinance. Rent control in San Francisco began in 1979. The measure was in response to the high inflation endemic to the 1970's and 1980's in the broader US, which was causing rents in the city to skyrocket monthly; annual housing cost inflation in the Bay Area hovered around 8% for much of this period. The recent passage of Proposition 13, which dramatically limited soaring property taxes, had brought with it hope that landlords would pass the savings on to tenants. But in large part this did not happen (Forbes & Sheridan, 1999). By June of 1979, the public of California, and especially San Francisco were hungry for rent control. On June 13th, the acting mayor signed the first rent control bill into law in the Bay Area in 40 years.

The original rent control law had several facets worth noting. First, the law prevented month-to-month rent increases, mandating that rent be set on one-year contracts. Secondly, for continuing tenancies (i.e., a tenant choosing to occupy a structure for an additional year), the law caps rent increases to fixed percentages, set annually for the Rent Control Board (Rent Board). These increases accounted for inflation, capping nominal rent increases at 7%. The original law exempted structures built after 1979, and structures of less than 5 units; this was intended to exempt so called "mom and pop" landowners. However, it provided incentive for large national and international property management companies to begin buying up many of these small properties to get around rent control. In response, the law was extended in 1994 to all structures built before June 13, 1979 with two or more units.

In response to the original law passed in 1979, many landowners and property managers began converting their structures to condominiums. Condominiums (condos) are owner occupied housing units that feature shared common spaces or amenities for monthly "condo fees." Being owner occupied, they are exempt from rent control. For these landowners and managers, conversion of units to condos allowed them to obtain market prices for their structures, effectively circumnavigating the rent control ordinance. The prevalence of these conversions led to a significant reduction in the city's rent-controlled housing stock in the years following the ordinance's passage. In response, San Francisco implemented a condominium conversion lottery program (condo lottery) in 1983 that limited the conversion of rent-controlled housing units; those property owners who wished to convert their structures to condos were required pay a fee and be entered in an annual moderated lottery for the right to convert. To be eligible to apply for the lottery, a the landowner must have lived in their structure for at least 3 years; only landowners with structures of six or less units were allowed to apply. Approximately 200 units were approved for conversion per year (San Francisco Planning Department, 2020). Those who applied in previous years but did not win were given priority in following years. In 2013, the city government suspended the condo lottery until 2024 due to a tremendous backlog of applications. The program was replaced with the Expedited Conversion Program (ECP), which allowed owners of two-unit Tenancy in Common (TIC) structures to apply for a quicker conversion process. However, in 2017, the ECP was found to be unconstitutional and was ordered to cease. Throughout the entire history of the ordinance, certain two-unit structures were eligible to convert without a lottery win or ECP allowance so long as the building was 25% owned by two non-married individuals; this has been termed the "two unit bypass." After the lottery was suspended and the ECP shut down, two unit bypasses were the only way to convert a controlled structure into condos, and remain so to this day.

The law also has what is called "vacancy decontrol"; this means that when tenants move out, landlords are allowed to set rent on vacant units to any level, regardless of the Rent Board's mandated maximum percentage. Finally, the law contained language restricting the reasons a tenant could be evicted to so-called "Just Cause" evictions, limiting the ability of landowners to evict tenants just to increase rents. Additionally, under this clause, landowners are required to supply monetary compensation to tenants during renovation or eviction for property conversion. The Just Cause evictions clause applies to all rental structures in San Francisco, not just those that fall under rent control.

Landowners who wish to evict tenants either to raise rents or convert their structure have historically had the option of providing "buyout" payments to tenants. Here, a landowner offers a sum of money for a tenant to vacate a unit. While a 1996 ordinance barred landowners from converting or raising rents following a no-fault eviction, such actions are not prohibited following buyouts. The buyout negotiation process, being unregulated, has anecdotally involved coercive actions by landowners, threatening eviction or legal action if the tenant does not accept the buyout. Many tenants in these situations do not know the full extent of their rights to legal counsel or rejection of these offers. In 2015, the Rent Board passed an ordinance requiring landowners to, among other things, provide tenants with information about their rights during the buyout process, but the practice is still legal and widely implemented. This law also prevented the conversion of a unit to condominium following a buyout agreement, eviction of a senior, disabled, or catastrophically ill tenant.

In 2020, California succeeded in passing statewide rent control. Note however that this is outside the window of this study, and hence should not present confounding variation.

I choose to estimate the model for San Francisco for a number of reasons. First, San Francisco's model of rent control is fairly representative of modern rent control measures active in the US, featuring both exceptions for new housing and vacancy decontrol; results from my study may be externally valid in other regions with similar measures such as Oregon, Washington, DC, and Los Angeles, CA. Secondly, rent control and housing affordability in San Francisco has been previously studied (Diamond et al., 2019a, 2019b; Glaeser, 2002); since my approach is novel, it is desirable to have other studies of San Francisco to serve as guidelines for my results. Finally, the city of San Francisco publishes a complete data set of demolition/building permits dating back to 2004, which I need to identify property renovations and new housing construction.

4.1.2 Housing Affordability and Turnover. San Francisco ranks low among measures of housing affordability in general. The crisis, largely driven by restrictive zoning regulation (Hsieh & Moretti, 2017) and influxes of affluence in the tech sector, had been around for decades, but became particularly pronounced after the dot-com bubble of the 1990's and Great Recession of 2008-2009. In 2016, around half of the city's renters were considered "burdened" (paying in excess of 30% of their income for housing) compared to 32% nationally. That same year, the median home price sale was $1.29 million, which was affordable to only 13% of the city's population. Compare this to the 57% of Americans that could afford the median home price in their area nationally (Bellisario, Weinberg, Mena, & Yang, 2016).

San Francisco's housing cost appreciation during the study period dramatically outpaces national trends. Figure 24 maps percent changes in measures of housing cost by tract throughout the city. In many neighborhoods, the appreciation of rent for both controlled and exempt structures exceeds that of the national average for the same time period (32.2%). Particularly striking is the degree of appreciation even amongst rent controlled units, whose rents may only rise given a vacancy. Many tracts experienced appreciation in median rents of 50% or more over the study period. The appreciation of rents in structures exempt from rent control is even more drastic, with rents in a number of neighborhoods in

Figure 24.: Median Housing Cost Appreciation, 2010-2019

(a) Rent Controlled

(b) Rent Control Exempt

(c) Mortgages

(d) Home Values

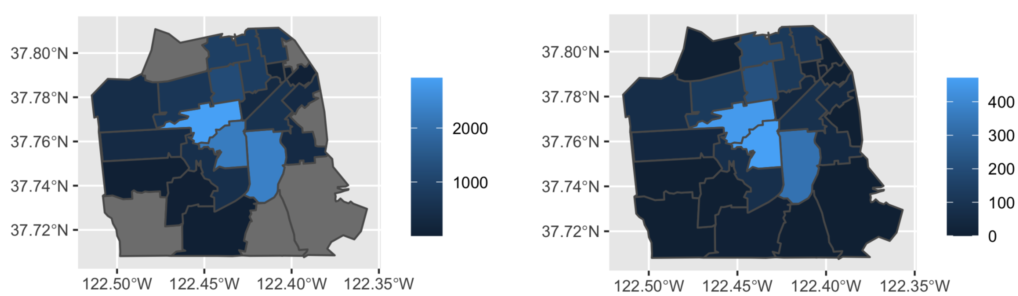
gentrifying areas appreciating by 70-80%.[1] Figure 24 depicts the appreciation in median housing costs in the city by census tract for visual reference.

This paper specifically focuses on turnover in the housing market in the form of conversions of rent controlled housing units into condominiums. Unfortunately, a comprehensive spatial census of these conversions does not exist. However, I do have some data which may help illustrate where conversions are most likely to occur in the city. Figure 25 maps the number of units in structures that applied for (a) and won (b) the condo lottery between the years 2007 and

[1]I determine gentrification status using the methodology in Freeman (2005); Specifically, to be considered gentrified during an intercensal period, a tract must: "1) Be located in the central city at the beginning of the intercensal period; 2) Have a median income less than the median (40th percentile) for that metropolitan area at the beginning of the intercensal period; 3) Have a proportion of housing built within the past 20 years lower than the proportion found at the median (40th percentile) for the respective metropolitan area; 4) Have a percentage increase in educational attainment greater than the median increase in educational attainment for that metropolitan area. 5) Have an increase in real housing prices during the intercensal period" (Freeman, 2005).

2013. Both maps tell roughly the same story: the majority of applicants are clustered in the geographical center of the city. Unsurprisingly, zip codes with the most applicants appear to be those with the most winners. Comparing figure 25 with figure 24(a), note that the most applicants and conversions loosely correspond to areas of the city experience the least appreciation in median rents among rent controlled structures. We can take this correlation as evidence that landowners are responding to depressed rents by seeking to convert their structures to condominiums. These results from the lottery include only 1,398 of the total 4,548 condo conversions reported in San Francisco Planning Department (2020) between 2010 and 2019; conversions via the two-unit bypass are not included in figure 25(b), meaning that the numbers shown are a partial sample of the total number of conversions.

Figure 25.: Units in Structures that Applied for and Won the Condominium Lottery



(a) Applicants (b) Winners

4.2 Model

I now outline the model I use to study the connection between rent control and housing turnover in the city. I assume that the city is partitioned into a set of discrete neighborhoods indexed by j. All action takes place in a single period. All economic actors are endowed with an initial state and decide whether and how to change their state during the action period. There are two types of economic actors: households and landowners. Households decide what neighborhood to live in and what type of housing to occupy, indexed by $\eta \in \{C, X, O\}$ for *Controlled*, *Exempt* from control, and *Owner* occupied. Landowners are endowed with a rent controlled structure of a certain size (measured in number of housing units) and decide whether to convert to condominiums or not. Condominiums are treated as O-type housing. I do not model the provision of other X- and O-type structures.

I assume that landowners contribute to a stock of housing in each neighborhood, the price of which is then set via market clearing.[2] Thus, I think of condo conversions as being a transfer of housing stock from $\eta = C$ to $\eta = O$. I also assume that if a landowner wishes to evict a tenant in order to convert, then they are able to do so, subject to a buyout cost.

I model three policy regimes, each of which has slightly different rules for landowners and households:

- **"Factual" scenario**: This regime takes place under both rent control and the condo lottery. Landowners may only convert their structure to condos if they have won the lottery or have a two-unit structure (and thus can use the two-unit bypass). I assume that any landowner that wins the condominium lottery converts their structure with probability 1; I justify this assumption with the reasoning that no landowner would enter the lottery unless they were already planning to convert should they win.[3] To convert, landowners pay a marginal conversion cost κ and neighborhood specific marginal buyout

[2]Modeling the search theoretical process of matching individual households to individual landowners is beyond the scope of this paper.

[3]Recall that there are fees, legal documents, and losses in rent associated with entering the lottery.

cost b_j paid to current tenants as compensation for their being removed from each unit. Landowners may only set rent to market rates on the fraction of units which are vacated by households in neighborhood j, the fraction $(1 - \varphi_j)$; the remaining fraction φ_j are set at the maximum allowable rent under the ordinance. As described below, φ_j is an endogenous object pinned down by the number of households that choose to move out of (C, j) housing.

– **"No-lotto" scenario**: This regime repeals just the condo lottery, allowing any landowner with a structure of six or less units to convert their structure to condos (I am unable to model condo-conversion costs for more than 6 units because the data does not exist since such conversions are not allowed). Landowners must still pay the marginal buyout and conversion costs in order to convert; they also still may only set a fraction $(1 - \varphi_j)$ to market rents.

– **"Decontrol" scenarios**: Under this regime, landowners are able to set market rents on their entire structure, regardless of household movement pattern. That is, $\varphi_j = 0$. Effectively, decontrol is the sans-rent control policy regime. I still assume, however, that residents must be bought out of their units before conversion to condos. Within this scenario, I actually model two different scenarios

  * **"Decontrol-2":** Here, I repeal rent control but only allow two-unit structures to convert. I conduct this exercise to more directly compare between the factual scenario (which only allows 2-unit conversions outside of the lottery) and the removal of rent control.

* **"Decontrol-6":** This scenario corresponds to the repeal of the entire rent control law, where I allow structures of up to 6 units to convert to condominiums.

I begin by describing the household sorting problem, then describe my handling of the rent-controlled landowners.

4.2.1 Households. Households are rational utility maximizers; they are endowed with an initial state, (η_0, j_0), which is their initial housing type and neighborhood (respectively), and an education level $e \in \{H, L\}$ (high or low) which correspond to having a college degree or not. In the action period, households must decide what type of housing and what neighborhood to live in. Each household occupies a single unit of housing and sorts so as to maximize utility,

$$\begin{aligned} V_{ij}^e(\eta|\eta_0, j_0) &= \xi_j^e(\eta|\eta_0, j_0) - \alpha^e \log R_j(\eta|\eta_0, j_0) + \varepsilon_{ij}(\eta) \\ &= \delta_j^e(\eta|\eta_0, j_0) + \varepsilon_{ij}^e(\eta). \end{aligned} \tag{4.1}$$

$\xi_j^e(\eta|\eta_0, j_0)$ is an amenity term for type-e households living in type-η housing in neighborhood j, conditional on households' initial living situation (η_0, j_0). I allow amenities to be conditioned on initial living situation to implicitly model moving costs, many of which are unobservable (e.g., sentimental attachment to a given dwelling unit or neighborhood). I assume that $\xi_j^e(\eta|\eta_0, j_0)$ is the same for all individuals moving to a new neighborhood-housing type pair; i.e., abusing notation, $\xi_j^e(\eta|\eta_0, j_0) = \xi_j^e(\eta)$ if $(\eta, j) \neq (\eta_0, j_0)$. $R_j(\eta|\eta_0, j_0)$ is the rent faced by households for η-type housing in j conditional on their initial living situation. If individuals lived in $\eta = \eta_0 = C$, $j_0 = j$ in the initial period and were not pushed out by a condo conversion, then they are only required to pay the rent controlled rent, $\bar{R}_j(C)$. If $\eta \neq C$, $\eta_0 \neq C$, or $j \neq j_0$, then they must pay market rents, $R_j(C)$. $R_j(C)$ is determined endogenously by the supply and demand of vacant controlled housing in j, where as $\bar{R}_j(C)$ is pinned down by rent control ordinance. In excess:

$$R_j(\eta|\eta_0, j_0) = \begin{cases} \bar{R}_j(C) & (\eta_0, j_0) = (\eta, j) \\ R_j(C) & \text{else} \end{cases}.$$

α^e is an e-specific rent sensitivity parameter. Finally, $\varepsilon_{ij}(\eta)$ is a household specific idiosyncratic taste shock for housing type-neighborhood pairs that is drawn from a generalized extreme value (GEV) distribution with nesting parameter $\lambda^e(\eta)$, such that the cumulative distribution function for ε is

$$F^e(\varepsilon^e_{i1}(\eta), ..., \varepsilon^e_{iJ}(\eta)) = \exp\Bigg(-\sum_\eta \Bigg(\sum_j \exp\Big(-\varepsilon_{ij}(\eta)/\lambda^e(\eta)\Big)\Bigg)^{\lambda^e(\eta)}\Bigg).$$

In essence, the parameter $\lambda^e(\eta)$ captures the degree to which different housing types are substitutable within neighborhoods.

I assume that there is an outside location where households can choose to live if not in the city proper that affords utility $\delta^e_0 = 0$. Given the properties of the GEV distribution, the probability that a type e household will choose to live in housing-neighborhood combination (η, j) conditional on their initial living conditions (η_0, j_0) is

$$p^e_j(\eta|\eta_0, j_0) = \frac{\exp\left(\delta^e_j(\eta)/\lambda^e(\eta)\right)\Bigg(\sum_\ell \exp\left(\delta^e_\ell(\eta)/\lambda^e(\eta)\right)\Bigg)^{\lambda^e(\eta)-1}}{1 + \sum_\eta \Bigg(\sum_\ell \exp\left(\delta^e_\ell(\eta)/\lambda^e(\eta)\right)\Bigg)^{\lambda^e(\eta)}}\Bigg| \eta_0, j_0. \tag{4.2}$$

Equation (4.2) allows me to calculate the fraction of housing which is not vacated by residents during the action period, φ_j. Letting $\mathbf{\Lambda}_j$ be the total supply of $\eta = C$ housing in j,

$$\varphi_j = \frac{\sum_e N^e_j(\eta_0 = C) \times \sum_{\eta, \ell \neq j} p^e_\ell(\eta|C, j)}{\mathbf{\Lambda}_j} \tag{4.3}$$

where $N_j^e(\eta_0)$ is the number of skill e households in j endowed with housing type η_0.

4.2.2 Landowners. Landowners, indexed k, are endowed with an initial structure of type $\eta = C$ and size Λ_k in neighborhood j. Under the factual scenario, the only landowners of concern are non-lottery winners with $\Lambda_k = 2$, as they are able to use the 2-unit bypass in the ordinance to convert; all other landowners have either won the lottery, and are assumed to convert, or are locked into their structure by the rent control ordinance. However, under the No-lotto or Decontrol scenarios, structures of greater size become able to convert. Therefore, I describe the landowner problem in the most general sense so that it applies to all three scenarios. In all scenarios, I assume that landowners earn rent on their entire structure, 12 times per year, and that any conversions take place half way between 2010 and 2019 (5 years in). I make this assumption to recognize that conversion is expensive and the decision to do so is made in anticipation of years of earned rental or mortgage income.

Landowners seek to maximize their profit, which I assume is given by

$$\pi_j(\eta|\Lambda_k) = 12 \times 5 \times \Lambda_k \times \begin{cases} \varphi \bar{R}_j(C) + (1-\varphi_j)R_j(C) & \eta = C \\ \\ R_j(O) - (b_j + \kappa)/(12 \times 5) & \eta = O \end{cases} \tag{4.4}$$

by choosing their desired housing type η. The $12 \times 5 \times \Lambda_k$ reflects the fact that landowners receive rent 12 times a year for 5 years on their entire structure. The first case corresponds to a landowner choosing not to convert their structure; they earn controlled rent on the fraction of their structure not vacated and are able to charge market rent on the remaining fraction. The second case corresponds to the decision to convert to condominiums; landowners pay buyout cost b_j and conversion cost κ but earn rental income $R_j(O)$.[4]

I also assume that landowner profits are augmented by landowner-housing type specific idiosyncratic shocks drawn from a Gumbel distribution with scale parameter σ. Given the properties of the Gumbel distribution, this means that the probability that a landowner in j with structure size Λ_k will choose housing type η is given by

$$p_{kj}(\eta|\Lambda_k) = \frac{\exp\left(\pi_{kj}(\eta|\Lambda_k)/\sigma\right)}{\sum_{\eta'} \exp\left(\pi_{kj}(\eta'|\Lambda_k)/\sigma\right)} \tag{4.5}$$

where $\eta' \in \{C, O\}$. Equation 4.5 makes clear the importance of the scale parameter of the landlord's cost shocks, σ. Larger values of σ mean that the variance of the shocks is high, and thus the stochastic component of landlord profit dominates and the deterministic component is overwhelmed. As $\sigma \to \infty$, landlords will begin to pursue each action with equal probability. Conversely, as $\sigma \to 0$, the deterministic component dominates as the variance of the Gumbel distribution collapses. Essentially, σ captures the sensitivity of landlords to market size and costs. This is a critical parameter to calibrate in the model, because it so fundamentally controls the reactivity of the housing market.

4.2.3 Evictions from Conversions. As mentioned previously, an ideal model would individually match households to landowners and vice-versa using search theory. However, I am constrained by the availability of data, and the requirements of such a model would be rather exhaustive; hence I treat the housing stock in each neighborhood as a pool to which landowners contribute and from which households buy. This creates an issue, however, when it comes to considering

[4]Note that under my assumptions all landowners would receive the $\eta = C$ case from above multiplied by $12 \times 5 \times \Lambda_k$ for the first five years regardless of their conversion decision. Hence, I remove this term for sorting purposes.

the situation of households who are initially endowed with rent-controlled housing but are pushed out by condominium conversion. These households are not covered by the rent control ordinance, and hence are no longer able to obtain type-C housing at the lower rent-controlled rent. The question becomes: if a landowner removes a unit from the stock of type-C housing in a given neighborhood, how many households endowed with $(\eta_0, j_0) = (C, j)$ should become ineligible for controlled rents in the action period? Since all households buy from a pool of housing, how many will lose their rent controlled status when a unit is removed from that pool?

I model evictions using a simple reduced-form mechanism. Suppose that there are $N_j^e(C)$ type-e individuals endowed with rent controlled housing, and a total type-C housing stock of $\mathbf{\Lambda}_j$ in neighborhood j. Further, let $\tilde{\Lambda}_j$ be the number of units converted to condominiums during the action period. Then, I assume that the number of type-e households to be pushed out of rent-controlled housing by conversions to be

$$\tilde{N}_j^e(C) = \tilde{\Lambda}_j \times \frac{N_j^e(C)}{\mathbf{\Lambda}}. \tag{4.6}$$

This expression says that the number of households pushed out by evictions is equal to the average number of households per housing unit times the number of conversions. For example, if there are on average two high-skill households per housing unit in a neighborhood, then for each unit converted we should expect that 2 such households are evicted.

4.2.4 Rent Formation and Equilibrium.

Although briefly touched on throughout this section, I now more carefully characterize how the market adjusts $R_j(C)$ to equate the supply of and demand for vacant $\eta = C$ housing and define equilibrium both with and without rent control. First, it is worth discussing how the market for available rent controlled housing clears. From equation (4.5), I have the probability that a landowner will choose to maintain their structure as rent controlled, $p_{kj}(C|\Lambda_k)$. Given the share of households that remain in their units φ_j, the expected amount of vacant housing from k is simply $p_{kj}(C|\Lambda_k) \times (1-\varphi_j) \times \Lambda_k$. In order for supply and demand to equate, $R_j(C)$ must solve

$$(1-\varphi_j)\sum_{k\in j} p_{kj}(C|\Lambda_k)\Lambda_k = \sum_{e,\eta_0,\ell\neq j} N_\ell^e(\eta_0)p_j^e(C|\eta_0,\ell), \tag{4.7}$$

where $R_j(C)$ is buried in the sorting probability functions. As such, there is no explicit formula for $R_j(C)$, but it is implicitly defined by this equation.

I now turn to equilibrium. Since I model the city with and without rent control, I will need two different definitions of equilibrium that are very similar in concept. I first start by defining a *rent control equilibrium.*

Definition 1. *Given a set of parameters* $\{\alpha^e, \lambda^e(\eta), \kappa, \sigma\}_{e,\eta}$*, neighborhood-housing amenities* $\xi_j^e(\eta|\eta_0, j_0)$*, initial endowments for households* $\{\eta_{i0}, j_{i0}\}_i$ *and* Λ_k *and exogenous amounts of other types of housing* $\{\mathbf{\Lambda}_j(\eta)\}_{\eta,j}$*, maximum allowable rents* $\bar{R}_j(C)$*, and buyout costs* $\{b_j\}_j$*, a* ***rent control equilibrium*** *is a set of rents, housing quantities, and distribution of households such that*

1. *Households sort themselves so as to maximize utility given by (4.1) and according to probabilities (4.2), and* φ_j *is formed endogenously by their actions;*

2. *Landowners choose whether or not to convert their structures to condominiums (given the ability to) so as to maximize profit (4.4) and according to probabilities (4.5);*

3. *Rents on vacated rent controlled housing units solve (4.7).*

I define a *decontrol equilibrium* identically, except that there is no maximum allowable rent and all $\eta = C$ structures are rented at the market clearing rent $R_j(C)$:

Definition 2. *Given a set of parameters* $\{\alpha^e, \lambda^e(\eta), \kappa, \sigma\}_{e,\eta}$*, neighborhood-housing amenities* $\xi_j^e(\eta|\eta_0, j_0)$*, initial endowments for households* $\{\eta_{i0}, j_{i0}\}_i$ *and* Λ_k *and exogenous amounts of other types of housing* $\{\mathbf{\Lambda}_j(\eta)\}_{\eta,j}$*, and buyout costs* $\{b_j\}_j$*, a* ***decontrol equilibrium*** *is a set of rents, housing quantities, and distribution of households such that*

1. *Households sort themselves so as to maximize utility given by (4.1) and according to probabilities (4.2), and* φ *is exogenously equal to 1;*
2. *Landowners choose whether or not to convert their structures to condominiums (given the ability to) so as to maximize profit (4.4) and according to probabilities (4.5);*
3. *Rents on all* $\eta = C$ *structures solve (4.7).*

4.3 Data

I combine a variety of publicly available data sets to estimate my model for the years 2010-2019. I take the city of San Francisco to be defined by the political boundary of the city as mapped by the state of California; I use official city definitions of neighborhoods, omitting Golden Gate Park and Treasure Island due to lack of data. Unfortunately, my data come only at the tract, public use microdata area (PUMA), or zip code. Therefore, when imputing data from one of these other geographic units into neighborhoods, I assume that each belongs to the neighborhood in which the majority of its area is located; few tracts or zip codes are located in more than two neighborhoods. In a similar manner, I assume that neighborhoods follow the trends of the PUMA to which the majority of their area belongs.

For the household side of the model, I use data collected by the US Census Bureau. To match populations of differently skilled individuals, I turn to the tidycensus package from R; this is also where I obtain shape files of the tracts in the city. I gather data on income from IPUMS through the University of Minnesota at the city by education level. From IPUMS I also obtain tenancy rates for C, X, and O type housing at the PUMA level by skill level. These data also allow me to see how long each sampled household has lived at their current address, allowing me to also estimate the share of residents in each neighborhood that remain in their housing throughout the 2010-2019 period; I assume that any household staying for more than 5 years stays through the entire period. Unfortunately these tenancy rates are not available at finer geographic detail that the PUMA; therefore, I assume that all neighborhoods within a single PUMA share the same tenancy rates for each housing type. Finally, to help calibrate household housing demand parameters, I turn to the Consumer Expenditure Survey published by the Bureau of Labor Statistics (BLS), which reports the budget share of housing expenditures for consumers of various education levels in the Bay Area. I make limited use of other Census data, such as the Bureau's Quick Fact Finder.

For data on housing units counts, median rents, and average mortgages, I turn to the National Historic GIS (NHGIS) data tables, which aggregate Census data to the tract-year level. The unit counts data reports the number of individual housing units existing in structures of size 1 unit, 2-4, 5-19, 20-49, and 50+ units by decade built and owned/rented status. I obtain an estimate of the number of structures of each binned size by taking the total unit count and dividing it by the midpoint of each bin (i.e., I divide the number of units in buildings of 2-4 units by 3, units in 5-19 unit buildings by 12, and so on). I then filter down to only those structures built pre-1980 and with more than 1 unit to obtain inventories of C-type housing by neighborhood in 2010. This not only gives me a structure count but also an idea of the distribution of structure sizes which I leverage in section 4.4.3 to help initialize the simulation exercises. Since two unit structures in particular are important for the model, I use building footprints from the San Francisco Planning Department to obtain the exact count of two unit structures in each neighborhood; I multiply this number by the share of 2-4 unit building units to more precisely estimate the number of rent controlled two-unit buildings. The data on rents comes binned by units in structure and decade, meaning I can once again separate out C- from X- type buildings; I construct neighborhood rents from an average of these data weighted by the number of units falling in each bin. I use average mortgages per tract as monthly housing costs ($R_j(O)$) for O-type buildings.

In order to properly initialize the simulations, I need to know the distribution of lottery winners across the city by structure size. The city publishes a list of the lottery winners by structure size and zip code that I use to find this. However, just because a certain landowner wins in a given year does not mean that they will convert immediately. This fact means that I cannot simply look at lottery winners in 2010-2013 (the overlap between my study and the lottery), because there may be more landowners eligible to convert left over from previous years. For instance, 4,548 residential units underwent condominium conversion between 2010 and 2019, despite only 799 winning (San Francisco Planning Department, 2020). Therefore, I include lottery winners from 2007-2009 in my sample of the distribution of lottery winners. This gives me a total number of lottery-winning units of 1,398, leaving 3,150 two-unit bypass conversions.

4.4 Calibration and Estimation

In this section, I describe how I put numerical values to the parameters in my model. Generally, I do this in two steps: first, I calibrate as many parameters as I am able to take from the literature to cut down on the dimensionality of the problem. Second, I use standard estimation techniques to retrieve the remaining parameters. I also briefly explain how I initialize each of the counterfactual simulations and obtain results via Monte-Carlo integration.

4.4.1 Calibration. I begin by calibrating the household rent sensitivity parameters, α^H and α^L. Estimating these parameters would be a challenging endeavour because rents are obviously endogenous to household demand. The availability of instruments at the intra-city level to properly identify these parameters is extremely limited; therefore, I choose to calibrate these values. To do this, I rely on two sources from the literature. Su (2020) also uses a Cobb-Douglas utility function and derives indirect utility similarly to my methodology, and finds a rent sensitivity parameter of 0.7950 high-skill households and 0.4593 for low skill households. That value represents the quotient of the household budget share for housing and the standard deviation of Gumbel shock idiosyncratic preferences for neighborhood choice. Thus, if I assume that budget shares for housing follow national patterns of consumption, I should be able to back out skill-level specific Gumbel scale parameters (call these ς^e) to use in calibration of my own model. Foster (2014) estimates that high- and low-skill households spend 33.7% and 32.1% percent of their budget on housing at the national level. Therefore, assuming these trends hold in Su (2020), I am able to obtain that paper's estimates of ς^H and ς^L to use in my own. This methodology gives me $\varsigma^H = 0.04038$ and $\varsigma^L = 0.6804$. Since $\alpha^e = \hat{\alpha}^e/\varsigma^e$, this gives me $\alpha^H = 0.906$ and $\alpha^L = 0.538$. The greater sensitivity of high-skill households to the rent index is indicative of the greater mobility ($\varsigma^H < \varsigma^L$) that is typically found to hold in papers that estimate these scale parameters (Su, 2020; Tsivanidis, 2019).

I also calibrate the number of households of each skill level. I treat each resident of the greater San Francisco metropolitan statistical area as an independent household, giving me populations as given in table 9.

Table 9.: Households by Year and Skill Level

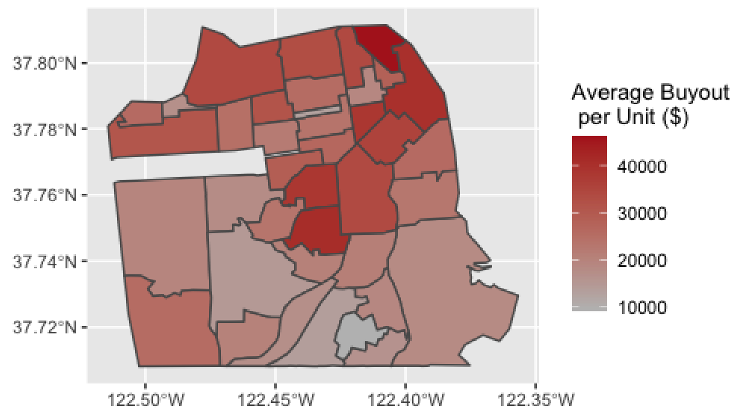
Year	High-skill	Low-skill
2010	1,286,088	1,699,387
2019	1,668,244	1,722,384

On the landowner side, I calibrate $\sigma = \$35,490$ as in Murphy (2018); that paper estimates Gumbel shocks for landowners in San Francisco over a similar time horizon, making this value very suitable for my own work. I also calibrate the buyout costs per unit, b_j, using San Francisco city data. Figure 26 presents average buyout costs by neighborhood; the mean buyout cost by neighborhood overall is $25,400 per unit, and ranges from a minimum of $9,233 to a max of $46,064 per unit. Buyouts generally appear to be higher in the Northwest of the city, close to the downtown financial district where the most affluent populations in the city tend to live.

4.4.2 Estimation.

Here, I calibrate the remaining parameters of the model, GEV nesting parameters $\lambda^e(\eta)$ and amenities $\xi_j^e(\eta|\eta_0, j_0)$ on the household side, and the marginal condo conversion cost κ on the landowner side. I start by estimating the remaining household parameters. If I could observe the

Figure 26.: Average Buyout By Neighborhood, 2010-2019



Data from city of San Francisco.

demographics and endowments of initial and final states for individual households, I could estimate household-side parameters using basic instrumental least squares on first differences of log-population shares; however, I observe only more aggregate data, so I must calibrate these parameters by matching more aggregated moments. To do this, I rely on techniques developed by S. T. Berry (1994) and S. Berry et al. (1995), modified for the GEV distribution by Grigolon and Verboven (2014). I do this in two steps: first I calculate mean utility for households occupying each neighborhood-housing type conditional on their initial location and hypothetical values of $\lambda^e(\eta)$; I then search for the set of $\lambda^e(\eta)$ which minimizes the share of mean utility unexplained by rents, or $\xi^e_j(\eta|\eta_0, j_0)$.

In step one, I posit a set of values for $\lambda^e(\eta)$ and calculate mean utility for each skill-neighborhood-housing type-initial condition combination, which I denote as $\delta^e_j(\eta|\eta_0, j_0) = \xi^e_j(\eta|\eta_0, j_0) - \alpha^e \log R^e(\eta|\eta_0, j_0)$ as in section 4.2.1. S. T. Berry (1994) showed that for extreme value preferences there exists an iterative contraction mapping that is guaranteed to result in a fixed point for $\delta^e_j(\eta|\eta_0, j_0)$. Grigolon and Verboven (2014) modify this contraction mapping slightly for the usage of generalized extreme value shocks, which I reproduce here using my notation:

$$\delta^{\iota+1} = \delta^{\iota} + \lambda^{e}(\eta)\big(\log p^{e}(\eta) - \log p^{e}(\eta|\delta^{\iota})\big) \tag{4.8}$$

where ι is the iteration number, p^e is observed shares and $p(\eta|\delta)$ is model predicted shares based on the current guess of δ, both conditioned on initial conditions. This mapping is repeated until $\delta^{\iota+1} \approx \delta^{\iota}$. Here, as in Grigolon and Verboven (2014), the $\log p - \log p(\delta)$ term is dampened by the GEV nesting parameter $\lambda^e(\eta)$, as compared to BLP and S. T. Berry (1994) where it is not (this is necessary for convergence with GEV preferences). Unlike in those papers, however, not all households have access to the same set of "goods" (combinations of neighborhood, housing type, and initial endowment). I therefore restrict the number of households with access to $\delta^e_j(\eta\eta_0, j_0)$ to only those with initial endowment (η_0, j_0). As discussed in section 4.3, I am able to observe the rate which households stay in the same housing unit for the majority of the study period at the PUMA level, allowing me to estimate the moment $p^e_j(\eta|\eta, j)$.

In step two, I search for the optimal $\lambda^e(\eta)$'s. For every guess of parameters $\lambda^e(\eta)$, I estimate $\delta^e_j(\eta|\eta_0, j_0)$ and $\xi^e_j(\eta|\eta_0, j_0)$; I then use a Nelder-Mead simplex method to search for the parameters $\lambda^e(\eta)$ that minimize the sum of squared $\xi^e_j(\eta\eta_0, j_0)$ terms. Roughly speaking, this is akin to finding the set of nesting parameters that minimize the explanatory power of neighborhood-housing type amenities in generating the observed equilibrium. Table 10 presents the likelihood maximizing values of $\lambda^e(\eta)$ that I obtain from this exercise.

Finally, I obtain κ, the cost to landowners of converting each of their units to condominiums. It is important to note here that κ not only captures any construction or physical costs of conversion but also costs from redesign, time off

Table 10.: Estimated GEV Nesting Parameters, $\lambda^e(\eta)$

η	High-Skill	Low-Skill
C	0.842	0.188
X	0.934	0.122
O	1.000	0.406

of the market, and so on. Hence, while the city of San Francisco publishes data on the physical costs of condo conversion, directly using only that data would tend to underestimate the size of κ. Therefore, I calibrate κ "by hand" (making my own adjustments via trial and error). Using all the other parameters of the model, I simulate equilibrium and find the number of conversions over an average of several simulations;[5] I then adjust κ until the simulated number of condo conversions closely matches the observed total citywide, 3,150. This exercise gives me a value for κ of \$108,500.

4.4.3 Simulation Details. Unfortunately, my data does not allow me to observe the exact number of structures of each type and size, which the results of simulation critically depend on. Therefore, for each scenario I model, I conduct a number of simulations and take averages of outcome variables across them. To initialize each simulation, I begin by drawing initial landowner conditions for each neighborhood. From the data on units by year built by structure size, I estimate the number of structures of each binned size in each neighborhood as described in section 4.3. For each bin, I then draw structure sizes from a uniform distribution over the bin's respective range, rounded to the nearest whole unit. I draw as many such sizes as I estimate there are structures fitting within that bin.

[5] The variance across simulations that use the same value of κ is relatively low but each equilibrium is relatively time consuming to calculate; therefore, I only simulate each ten times. Section 4.4.3 details how I conduct each simulation.

I truncate the 50+ unit bin to 50 to 75 units; the building footprints data indicate that there are very few buildings of more than 75 units in any neighborhood in the city.[6] From the data on lottery winners I know how many winners there are in each neighborhood by size, allowing me to remove the proper number of landowners of the appropriate size from each neighborhood (since I assume that these landowners convert with probability 1; their unit counts are reflected in the observed number of owner-occupied units in 2019 from the NHGIS data). Hopefully, the repeated simulation and integration of results smooths out the uncertainty from not knowing precise structure counts.

At a high level of summary, my solution algorithm proceeds as follows: First, I draw initial structure sizes for landowners in each neighborhood and the share of households with given initial conditions according to the observed equilibrium; I initialize rents to the values from NHGIS. Next, I allow households to sort themselves. Then, I calculate the transition probabilities for each landowner and use these to calculate the resulting C-housing stock and number of condominium conversions. The number of conversions gives me the number of households of each skill level to evict as according to equation 4.6. I then calculate market clearing rent for $\eta \in \{C, X\}$ according to equation 4.7; I do not model the formation of $R_j(O)$, which I assume to be equal to the observed average monthly mortgage. I compare these newly generated rents to those used for household sorting and consider my solution algorithm to have converged if all rents on neighborhood-housing type combinations differ by less that \$50. If greater differences exist, I take the average of the old and new rents to form a new guess

[6]The building footprints data shows that each individual bin does not neatly follow a clear distribution, and there are not enough data to estimate distributions by bin by neighborhood in any event. I tried modeling the distribution of structures in each bin as normal, but received implausibly large variances from doing so.

of equilibrium rent, evict the proper number of households from rent controlled structures, and repeat the process until the difference in generated rents satisfies the stopping condition.

4.5 Results

The results of each simulation are presented in tables 11 and 12. Specifically, the first numerical column in table 11 reports the number of city-wide condominium conversions under each modeled scenario and table 12 presents the mean change in endogenous rents between the factual and other scenarios. Here we see that, under my calibration, my model has several distinct predictions. First, conversions increase by about 80% once the condo lottery restrictions are removed, allowing any structure of up to 6 units to be converted. Secondly, both Decontrol scenarios result in far more condo conversions that the Factual scenario; Decontrol-2 results in a 330% increase and Decontrol-6 results in a 647% increase in conversions. This second set of results runs counter to the predictions of Diamond et al. (2019a) and Diamond et al. (2019b), which suggest that rent control leads to increased turnover in the housing market.

Looking at table 12, we can see what is driving these massive increases in the number of conversions under the Decontrol scenarios. Across simulations and neighborhoods, rents under Decontrol-2 and Decontrol-6 on average are about 65% and 61% lower respectively as compared to rents on vacant units in the Factual scenario. Landowners, as a general rule, are more likely to convert when rents on C-type structures are lower, because it cuts into their profit margins. Sans rent control, the stock of available C-type housing dramatically increases, sending rents downward and thus *encouraging* condo conversions, even though the depression in rents on occupied units enforced by rent control is no longer in effect.

Table 11.: Number of Units Converted to Condos by Scenario

	Endogenous Prices	Fixed Prices
Factual	3,198	1,694
	(26.074)	(0.216)
No-lotto	5,752	1,973
	(57.800)	(1.815)
Decontrol-2	13,757	1,033
	(38.104)	(0.143)
Decontrol-6	23,881	1,153
	(84.076)	(1.153)

Average across 10 simulations, with standard deviation reported in parenthesis.

By restricting the quantity of housing that may be repriced over the decade, rent control leads to *more* conversions than its absence does. As a researcher, I have to question the validity of the magnitude of the decrease in rents, but this dynamic of rent control, to my reading, has yet to be highlighted in the literature. As found in Mense et al. (2019), however, both Decontrol scenarios also cause rents on never controlled units to decrease significantly and much more robustly. The increase in the quantity of available housing, despite the dramatic increases in turnover, take enough pressure off of the market for X-type housing that affordability actually improves in these units.

The repeal of the restrictions associated with the lottery actually lead to mean increases in rent across neighborhoods and simulations, in the markets for both controlled and uncontrolled housing units, relative to the Factual scenario. Evidently, this is because of the contraction in the supply of C-type housing units as structures of 3-6 units are now allowed to convert. The key difference between

Table 12.: Average Percent Changes in Endogenous Rents Compared to Factual Scenario

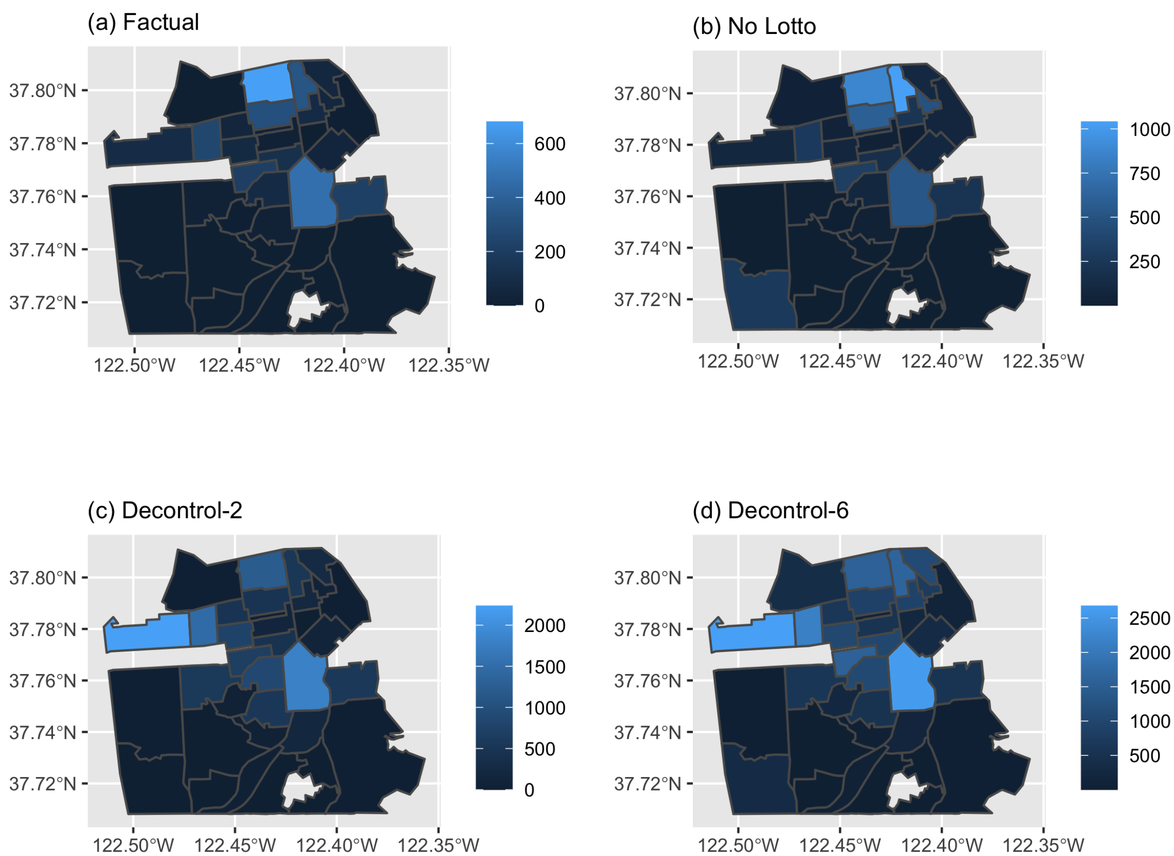
	$R(C)$	$R(X)$
No-lotto	2.29	0.12
	(17.168)	(0.002)
Decontrol-2	-65.01	-5.25
	(59.037)	(0.003)
Decontrol-6	-60.81	-4.63
	(10.204)	(0.060)

Average across 10 simulations, with standard deviations in parenthesis.

this scenario and the Decontrol scenarios is that rent control still mandates that rents can only be set to market rents on vacated units, restricting the supply of available housing for movers to each neighborhood and driving rents upwards.

Figure 27 depicts the simulated number of conversions by scenario. These maps of the city are not directly comparable to those in figure 25, as they use neighborhoods rather than zip codes. Still, the factual simulation (a) seems to find a hot bed of conversions near the geographical center of the city as in figure 25, albeit with a marked over-prediction in the central north neighborhood of Marina. The removal of the lottery results in a fairly similar picture, but with conversions taking place at a much greater scale (b). When rent control is removed in the Decontrol simulations however, an entirely new locus of conversions in the neighborhood of Outer Richmond appears (although the midtown hot-spot remains prominent). This neighborhood has a relatively high proportion of its controlled housing stock in buildings with three to 4 units, making them ineligible for conversion under the condo conversion lottery rules. Other neighborhoods, such

Figure 27.: Simulated Condo Conversions by Scenario



Average number of units converted to condominium by study neighborhood, average across 10 simulations.

as West of Twin Peaks and Portola, see no new conversions after the removal of rent control; these neighborhoods have relatively few rent controlled units overall, and hence are not impacted as much by decontrol. Thus the removal of rent control does not impact all areas of the city uniformly, and shifts the locus of major housing market turnover to neighborhoods with relatively more rent controlled units.

As a final exercise, I run another set of simulations in the model where I set rents on all housing, including $R(C)$ and $R(X)$, to values observed in the data; this corresponds to the second numerical column in table 11. I do this to examine landowner decision making absent of any endogenous price impacts, that is, to compare across regulatory scenarios under the observed prices. When prices are endogenous, the flood of $\eta = C$ housing from the removal of rent control in the Decontrol scenarios causes $R_j(C)$ to plummet compared to the factual scenario, a mechanism that we might expect to see in the real world following repeal of the ordinance. With this battery of simulations, however, I remove this confounding variation. These exercises also keep neighborhood populations of each skill level much closer to actual populations. Loosely speaking, these simulations show what landowners would have done if everything was kept identical to the Factual scenario except for the various regulations they remove. It is worth noting here that under fixed rents the number of conversions in the Factual scenario is roughly half of that observed in the data. This is because the condo conversion cost, κ, is calibrated to match the observed number of conversions under endogenous prices. Endogenous rents resulting from the Factual simulation, however, are almost uniformly much lower than observed rents, and hence there exists greater incentive to convert to condominiums. A greater conversion cost calibration is thus required to deter conversion. I choose not to recalibrate κ for the fixed rent simulations for internal consistency.

The results of the exogenous price simulations confirm the intuition of Diamond et al. (2019a) and Diamond et al. (2019b). The repeal of the condo lottery, represented by the No-lotto scenario, leads to a 16.5% increase in the number of condominium conversions, reflecting that landowners with up to 6 unit structures are newly able to convert. The most striking results however are from the Decontrol scenarios, where the removal of rent control leads to significantly fewer condo conversions. Keeping the pool of structures eligible for conversions static in Decontrol-2 absent rent control leads to about 39% fewer units being converted to condominiums. This is as expected; removing rent control allows these landowners to set rent on all of their units to market levels $R(C)$, regardless of their occupation status, reducing the profit differential between C and O structures. Decontrol-6, however, shows that the reduction of conversions remains significant even when all structures with 6 or fewer units become eligible for conversion. Compared to the Factual scenario, Decontrol-6 still leads to a 32% reduction in the number of conversions. Taken together, the exogenous price simulations demonstrate two key facts about rent control in the city of San Francisco: 1) the rent control ordinance provides landowners with greater incentive to convert to condominiums even under fixed prices; and 2) the condominium lottery is important to reduce total conversions.

The comparison of the endogenous and exogenous price simulations show that there are two main factors that would contribute to the number of condominium conversions were rent control repealed in San Francisco. First, rent control depresses rents, causing fewer landowners to convert to condos upon its removal. Second, rent control restricts the quantity of available rental units, causing rents on available units to be higher under the policy than without it; thus, the removal of rent control leads to lower rent and causes more landowners to convert. These two forces obviously oppose each other, and their relative sizes determine whether the repeal of rent control causes housing turnover to increase or decrease. My results from the endogenous price scenarios suggest that the second effect may dominate the first under repeal. However, the results of the exogenous price scenarios prove that both are occurring.

4.5.1 Welfare and Profit. I briefly document the winners and losers of each policy scenario in table 13, which records average percent changes in the welfare of households and profit per unit of landowners across neighborhoods in each scenario.[7] Unsurprisingly, the repeal of the condo lottery results in a slight net loss for both high- and low-skill households, reflecting the increase in rent outlined in table 12. The contraction in the supply caused from increased condo conversions under this counterfactual drastically limits household options, enough to negatively impact utility (albeit by less than 0.2% for both household types). Landowners are made slightly better off, however. Many landowners convert to condominiums to secure higher rent; those who do not convert are made better off by the resulting rising rents from the supply contraction.

Table 13.: Percent Change in Utility for Households and Profits per Unit for Landowners by Scenario

	H-skill Households	L-skill Households	Landowners
No-lotto	-0.13	-0.06	1.27
Decontrol-2	9.33	2.06	-57.64
Decontrol-6	8.13	1.84	-47.09

Percent changes calculated relative to the Factual scenario simulation, averaged across neighborhoods. Landowner profits are normalized by structure size to give profit per unit in the last column.

Given the housing cost results outlined in table 12, the changes in welfare and profit in table 13 for the Decontrol scenarios also make sense. As described

[7]Unfortunately, since I draw new structure sizes for each simulation run, profit outcomes are not directly comparable across simulations. Thus, I normalize landowner profit by the number of units in their structure, giving me profit per unit, which I can compare across simulations. Since landowner profit is essentially linear in structure size, there is no systematic relationship between changes in profit from the Factual simulations and structure size within each scenario.

previously, the introduction of the entire stock of formerly controlled housing under Decontrol during the action period (rather than just the vacant controlled housing) drives rents down significantly. Thus, households are made better off and landowners are made worse off relative to the factual scenario. High-skill households, in particular, seem to benefit from Decontrol. This is likely because high-skill households have a greater rent sensitivity parameter, α^e, which essentially governs how freely mobile each household is with respect to rents. The higher α means that high-skill households can move around the city to capitalize on lower rents more easily than low-skill households. Note that households of both types do slightly better under Decontrol-2 than Decontrol-6. This is because rents depreciate more under Decontrol-2 than Decontrol-6 relative to the Factual scenario. Landowners, conversely, do worse under Decontrol-2 than Decontrol-6, for the same reason.

4.6 Conclusion

This paper is among the first to structurally model the decision making process of landowners in a rent control environment. I find that the repeal of the condominium conversion lottery restrictions, which limit the number and size of structures that can be converted to condominiums, leads to an 80% increase in the number of conversions. However, I also find that the removal of the entire rent control ordinance leads to a 647% increase in the number conversions, because such a counterfactual floods the market with available formerly controlled units, eating into landowners' profits. Affordability increases significantly under this scenario. The results of this paper could be important to policy makers considering implementing rent stabilization measures in response to housing affordability crises and gentrification. While not all simulations return results of believable size, qualitatively I demonstrate a number of key forces at play in cities under a modern rent control measure.

While this paper and others (Diamond et al., 2019a, 2019b; Kholodilin et al., 2016) have commented on the potential for rent control to lead to increased gentrification, to my knowledge, no paper has yet established such a relationship robustly. Similarly, a truly structural and dynamic model of landowner reactions to a rent control policy has yet to be drafted and estimated. This paper therefore leaves a substantial mantle for future work to take up.

CHAPTER V

CONCLUSION

This dissertation has sought to structurally model a series of two-sided interactions that take place in cities in order to determine their contribution to emergent city-wide patterns of change. Chapters two and three treated the aggregate pattern known as gentrification as being the result, in part, of interactions between households and firms. In a spatial setting, both non-tradable firms and households are incentivized to locate near one another to mitigate travel costs. Firms in particular wish to locate near wealthier households to maximize their revenue, exactly the sorts of households that tend to flood into central city neighborhoods as gentrification begins. More of these firms move into these neighborhoods, prompting additional affluent households to move in, and so on in a feedback cycle. Fundamentally underlying this cycle are simple laws of economics and space: it is harder to travel longer distances, and having more money tends to increase consumption. This work has shown that these forces alone can explain about 28% of the intensity of gentrification in Seattle, WA, and 17% in Portland, OR, and about 21% and 9.5% of the extent of it respectively. Chapter four models households and landowners in San Francisco, CA, under a rent control ordinance and sheds light on the factors of that policy that lead to turnover in the housing market, and by extension likely other emergent patterns such a gentrification. Once again, the forces at play are simple: landowners seek to maximize profits in the presence of construction costs, and households react to maximize their happiness. The interaction of these forces can lead to massive changes in the stock of rental housing in the city despite their simplicity. If there is any over-arching theme to this dissertation, it is this: even very basic economic incentives and spatial considerations can lead to enormous aggregate patterns that entirely reshape urban economic outcomes simply by interacting in a strategic environment. This is a key concept in economics, of course, but is here once again demonstrated.

Since completing this dissertation, I have started to think more about the role that commuting might play in household and firm interactions. Specifically, other papers such as Su (2020) and Edlund et al. (2015) have studied the roll that the increasing value of time for high skilled workers, brought on by longer working hours, may have induced these workers to sort towards the locations of their employers in city centers. This theory undoubtedly has relevance in chapters two and three, where households sort themselves in relation to firms' sorting decisions. Were high-skill tradable jobs to appear disproportionately downtown (as they seem to in Portland over the study window), it could induce many high-skill households to sort toward the city center even though doing so bears no consumptive benefit from these firms; rather, they may so sort to decrease time spent commuting. The omission of a commuting framework from my model, under this hypothetical, will thus cause me to underestimate the impact of firm sorting on gentrification. Including commuting in the model would be a fairly natural extension of the existing framework, and would highlight all the more the careful consideration that is due to economic relationships in space.

I can also think of several improvements to chapter four as well, though they would add serious complexity. A more complete model of household and landowner interactions in a rental market necessitates a two-sided search theoretical framework, where households search for differentiated housing and landowners for differentiated tenants. To my knowledge, no such model has yet been postulated because of the exorbitant difficulty in obtaining relevant data and estimating such a model. Still, a thoughtful implementation of a monopolistically competitive housing market could be a substantial improvement over my competitively market-clearing model. Additionally, the fit and performance of the model could be improved by the inclusion of a market clearing process within the paper's BLP algorithm, which would generate prices that rationalize the observed equilibrium under the model's assumptions. This may lead it to more accurate and less drastic counterfactual predictions.

Writing this dissertation has equipped me with the structural modeling skills that I have desired since first entering the discipline of economics, and they in turn have allowed me to ask questions of cities and economies that I hope will prove useful to policy makers. In future work, I hope to continue exploring how cities grow and change with respect to income, education, and spatial landscape, with a particular eye towards concerns of equity and inclusion in socioeconomic processes. I hold no doubt that the tools I have acquired throughout this process will continue to serve me well in my career henceforth, and I am thankful for the opportunity to produce this work.

REFERENCES CITED

Atkinson, R., Wulff, M., Reynolds, M., & Spinney, A. (2011). Gentrification and displacement: the household impacts of neighbourhood change. *AHURI Final Report*, *160*, 1–89.

Autor, D., Palmer, C. J., & Pathak, P. A. (2017). *Gentrification and the amenity value of crime reductions: Evidence from rent deregulation* (Tech. Rep.). National Bureau of Economic Research.

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

Baum-Snow, N. (2007). Did highways cause suburbanization? *The quarterly journal of economics*, *122*(2), 775–805.

Baum-Snow, N., & Hartley, D. (2019). Accounting for central neighborhood change, 1980-2010. *Journal of Urban Economics*, 103228.

Bayer, P., McMillan, R., & Rueben, K. (2004). *An equilibrium model of sorting in an urban housing market* (Tech. Rep.). National Bureau of Economic Research.

Bellisario, J., Weinberg, M., Mena, C., & Yang, L. (2016). *Solving the housing affordability crisis* (Tech. Rep.). Bay Area Council Economic Institute.

Bems, R. (2008). Aggregate investment expenditures on tradable and nontradable goods. *Review of Economic Dynamics*, *11*(4), 852–883.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841–890.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 242–262.

Boustan, L. (2010). Was postwar suburbanization "white flight"? evidence from the black migration. *The Quarterly Journal of Economics*, *125*(1), 417-443. Retrieved from `https://EconPapers.repec.org/RePEc:oup:qjecon:v:125:y:2010:i:1:p:417-443.`

Card, D. (2009). Immigration and inequality. *American Economic Review*, *99*(2), 1–21.

Card, D., Mas, A., & Rothstein, J. (2008). Tipping and the dynamics of segregation. *The Quarterly Journal of Economics*, *123*(1), 177–218.

City of San Francisco. (2022). *City of san francisco open data portal.* Retrieved from `https://data.sfgov.org`

Costa, D. L. (2000). The wage and the length of the work day: From the 1890s to 1991. *Journal of Labor Economics*, *18*(1), 156–181.

Couture, V. (2013). Valuing the consumption benefits of urban density. *University of California, Berkeley. Processed*.

Couture, V., Gaubert, C., Handbury, J., & Hurst, E. (2019). *Income growth and the distributional effects of urban spatial sorting* (Tech. Rep.). National Bureau of Economic Research.

Couture, V., & Handbury, J. (2017). *Urban revival in america, 2000 to 2010* (Tech. Rep.). National Bureau of Economic Research.

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

Diamond, R., McQuade, T., & Qian, F. (2019a). The effects of rent control expansion on tenants, landlords, and inequality: Evidence from san francisco. *American Economic Review*, *109*(9), 3365–94.

Diamond, R., McQuade, T., & Qian, F. (2019b). Who pays for rent control? heterogeneous landlord response to san francisco's rent control expansion. In *Aea papers and proceedings* (Vol. 109, pp. 377–80).

Dixit, A. K., & Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *The American economic review*, *67*(3), 297–308.

Doucet, B. (2009). Living through gentrification: subjective experiences of local, non-gentrifying residents in leith, edinburgh. *Journal of Housing and the Built Environment*, *24*(3), 299–315.

Eckert, F., Ganapati, S., & Walsh, C. (2019). Skilled tradable services: The transformation of us high-skill labor markets. *Available at SSRN 3439118*.

Edlund, L., Machado, C., & Sviatschi, M. M. (2015). *Gentrification and the rising returns to skill* (Tech. Rep.). National Bureau of Economic Research.

Ellen, I. G., Horn, K. M., & Reed, D. (2019). Has falling crime invited gentrification? *Journal of Housing Economics*, *46*, 101636.

Forbes, J., & Sheridan, M. C. (1999). The birth of rent control in san francisco. *San Francisco Apartment Magazine*. Retrieved from `https://web.archive.org/web/20080720034219/http://www.sfaa.org/0406forbes.html`

Foster, A. C. (2014). New education classification better reflects income and spending patterns in the consumer expenditure survey.

Freeman, L. (2005). Displacement or succession? residential mobility in gentrifying neighborhoods. *Urban Affairs Review*, *40*(4), 463–491.

Glaeser, E. L. (2002). Does rent control reduce segregation? *Available at SSRN 348084*.

Glaeser, E. L., & Luttmer, E. F. (2003). The misallocation of housing under rent control. *American Economic Review*, *93*(4), 1027–1046.

Grigolon, L., & Verboven, F. (2014). Nested logit or random coefficients logit? a comparison of alternative discrete choice models of product differentiation. *Review of Economics and Statistics*, *96*(5), 916-935.

Guerrieri, V., Hartley, D., & Hurst, E. (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics*, *100*, 45–60.

Handbury, J. (2019). *Are poor cities cheap for everyone? non-homotheticity and the cost of living across us cities* (Tech. Rep.). National Bureau of Economic Research.

Head, K., & Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of international economics* (Vol. 4, pp. 131–195). Elsevier.

Hoelzlein, M. (2019). Two-sided sorting and spatial inequality in cities.

Hsieh, C.-T., & Moretti, E. (2017). How local housing regulations smother the us economy. *The New York Times*.

Huante, A. (2021). A lighter shade of brown? racial formation and gentrification in latino los angeles. *Social Problems*, *68*(1), 63–79.

Kholodilin, K. A., Mense, A., & Michelsen, C. (2016). Market break or simply fake? empirics on the causal effects of rent controls in germany.

Lester, T. W., & Hartley, D. A. (2014). The long term employment impacts of gentrification in the 1990s. *Regional Science and Urban Economics*, *45*, 80–89.

Macaig, M. (2015). *Gentrification in america report* (Tech. Rep.). Governing.

Manson, S., Schroeder, J., Van Riper, D., Kugler, T., & Ruggles, S. (2021). *Ipums national historical geographic information system: Version 16.0 [dataset].* IPUMS. doi: http://doi.org/10.18128/D050.V16.0

Margo, R. A. (1992). Explaining the postwar suburbanization of population in the united states: The role of income. *Journal of urban economics*, *31*(3), 301–310.

Mayer, H. (2005). Taking root in the silicon forest: High-technology firms as surrogate universities in portland, oregon. *Journal of the American Planning Association*, *71*(3), 318–333.

Mayer, H. (2006). What is the role of universities in high-tech economic development? the case of portland, oregon, and washington, dc. *Local Economy*, *21*(3), 292–315.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.

Mense, A., Michelsen, C., & Kholodilin, K. A. (2019). The effects of second-generation rent control on land values. In *Aea papers and proceedings* (Vol. 109, pp. 385–88).

Murdie, R., & Teixeira, C. (2011). The impact of gentrification on ethnic neighbourhoods in toronto: A case study of little portugal. *Urban Studies*, *48*(1), 61–83.

Murphy, A. (2018). A dynamic model of housing supply. *American economic journal: economic policy*, *10*(4), 243–67.

Myers, D., & Pitkin, J. (2009). Demographic forces and turning points in the american city, 1950-2040. *The Annals of the American Academy of Political and Social Science*, *626*(1), 91–111.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of political Economy*, *90*(6), 1257–1278.

Rosen, S. (1979). Wage-based indexes of urban quality of life. *Current issues in urban economics*, 74–104.

Rosner, G. Q. (2020). *Strong support for reducing rent and creating good cause eviction protections in washington* (Tech. Rep.). Washington Community Action Network and Greenberg Quinlan Rosner.

Ruggles, S., Flood, S., Ronald, G., Grover, J., Meyer, E., Pacas, J., & Sobek, M. (2020). *Ipums usa: Version 10.0 [dataset].* Retrieved from `https://doi.org/10.18128/D010.V10.0`

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, *125*(3), 1253–1296.

San Francisco Planning Department, T. (2020). *2019 san francisco housing inventory* (Tech. Rep.).

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of mathematical sociology*, *1*(2), 143–186.

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

Smith, G. S., Breakstone, H., Dean, L. T., & Thorpe, R. J. (2020). Impacts of gentrification on health in the us: A systematic review of the literature. *Journal of Urban Health*, *97*(6), 845–856.

Su, Y. (2020). The rising value of time and the origin of urban gentrification. *Available at SSRN 3216013*.

Suárez Serrato, J. C., & Zidar, O. (2016). Who benefits from state corporate tax cuts? a local labor markets approach with heterogeneous firms. *American Economic Review*, *106*(9), 2582–2624.

Thomschke, L. (2016). *Distributional price effects of rent controls in berlin: When expectation meets reality* (Tech. Rep.). CAWM Discussion Paper.

Tsivanidis, N. (2019). Evaluating the impact of urban transit infrastructure: Evidence from bogotá's transmilenio. *UC Berkeley, memo*.

US Bureau of Economic Analysis. (2020). *Gross value added by industry.* Retrieved from `https://apps.bea.gov/iTable/iTable.cfm?reqid=150&step=2&isuri=1&categories=gdpxind`

US Census Bureau. (2022). *Zip code business patterns survey.* Retrieved from `https://www.census.gov/data/developers/data-sets/cbp-nonemp-zbp/zbp-api.html`

World Bank (Ed.). (2020, April). *Urban development overview.* Retrieved from `https://www.worldbank.org/en/topic/urbandevelopment/overview#1`

Zaveri, M. (2019, Feb). Oregon to become first state to impose statewide rent control. *New York Times*. Retrieved from `https://www.nytimes.com/2019/02/26/us/oregon-rent-control.html`