The Rising Value of Time and the Origin of Urban Gentrification

Yichen Su*

Federal Reserve Bank of Dallas

[Click here for the latest version]

December 12, 2018

Abstract

I estimate a spatial equilibrium model to show that the rising value of high-skilled workers' time is an important driving force behind the gentrification of American central cities. I show that the increasing value of time raises the cost of commuting and exogenously increases the demand for central locations by high-skilled workers. While change in value of time is an initial force behind gentrification, its effect is substantially magnified by endogenous amenity improvement. The model implies that welfare inequality in the recent decades increases by more than the rise in earnings inequality if the forces behind gentrification are considered.

*I thank my advisors Luigi Pistaferri, Caroline Hoxby, and Rebecca Diamond for their guidance and support. I also thank Ran Abramitzky, Yiwei Chen, Eran Hoffmann, Sitian Liu, Isaac Sorkin, Yu Zheng, and participants at Stanford Labor and Public workshop, Stanford Applied Economics workshop, Stata Texas conference, and UEA conference. I gratefully acknowledge that this research was supported by the Leonard W. Ely and Shirley R. Ely Graduate Student Fellowship through a grant to the Stanford Institute for Economic Policy Research. The views in this paper are those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System.

1 Introduction

Since the 1990s, American cities have seen a wave of urban revival during which the growth of income and home value in the central city neighborhoods far outpaces that in the suburbs. This process, often called gentrification, is characterized by an influx of affluent, educated residents, as well as improving amenities and rising housing cost. This recent prosperity contrasts sharply with the long period of decline of the central cities and "flight" of affluent residents to the suburbs in the earlier time (Baum-Snow (2007), Boustan (2010)).

Prior papers have shown that the increasing valuation of central city amenities explains the rising demand for central cities by college educated residents (Baum-Snow and Hartley (2017), Couture and Handbury (2017)). However, since local amenity change is likely an endogenous process (Guerrieri, Hartley, and Hurst (2013)), the increasing amenity value of central cities could be both a cause and a consequence of the inflow of high-skilled residents. To trace the causal origin of the central city revival after decades of persistent decline, one ought to identify the *exogenous* forces that push high-skilled urban "pioneers" back to the central cities despite the initially low level of amenities in these locations prior to gentrification, and understand how these forces bring about the endogenous amenity change in the central cities.

In this paper, I show that the demand for shorter commute time by the high-skilled workers due to their rising value of time is an important exogenous force that pushes high-skilled workers into the central city neighborhoods. I further show that the inflow of high-skilled workers leads to an endogenous process of amenity change. In the endogenous process, both the levels of amenities and rents in the central cities increase in response to the inflow of the high-skilled urban "pioneers." While all workers dislike higher rents, high-skilled workers value the improved amenities more than low-skilled workers do. As a result, more high-skilled workers sort into the central city neighborhoods while low-skilled workers sort out to the suburbs, which endogenously leads to further improvement in amenities in the central cities. I show that while the rising value of time among high-skilled workers is an important underlying driving force behind gentrification, the resulting endogenous amenity change greatly magnifies its direct effect. Finally, I find that the rising value of time, amenity and rent change lead to an increase in welfare gap larger than the increase in income gap between high- and low-skilled workers.

To motivate my analysis, I document that the time period of gentrification coincides with a period in which working long hours became more prevalent among high-wage earners. Evidence (Kuhn and Lozano (2008)) suggests that, before 1980, low-wage workers tended to work longer hours than high-wage workers. However, since mid-1980, this pattern has reversed itself. In recent years, high-wage workers have been much more likely to work long hours than their low-wage counterparts. Moreover, I show that, since 1980, the growth of reported commute time is much slower among the workers in the top wage deciles than workers in lower wage deciles, suggesting some of the spatial relocation of high-skilled workers into the central cities is likely attributed to their changing value of time and increasing desire for shorter commute time.

To capture the forces that drive changes in neighborhoods more generally, I present and estimate a spatial equilibrium model of neighborhood choice. In my model, I allow workers to choose which neighborhood to live based on their value of time, the commute time to their jobs, rents and local amenities. My model allows the changing value of time to generate exogenous shock to location demand for neighborhoods with short commute time. My model also allows changing local population mix to generate changes in local amenity levels and local rents. While all workers find high rents undesirable, my model allows high-skilled workers to potentially value the improved amenities differentially than low-skilled workers do, which is important in capturing the mechanism of endogenous amenity change.

In the model, the mechanism of workers' spatial sorting is governed by how much workers' value of time, neighborhood amenities and rents each affects their preference for locations. I parameterize the model such that each of these components in workers' preference is characterized by a set of parameters. Credible estimation of these parameters are thus important in determining the relative importance of each channel through which gentrification occurs.

To estimate how much the value of time affects workers' demand for shorter commute time, I first need a measure of workers' value of time. I measure the value of time by estimating the "long-hour premium" for each detailed occupation, using repeated cross-sections from the Census data. Using the differential changes in the long-hour premium in different occupations, I examine how much value of time affect workers' migration elasticity with respect to commute time to jobs.

The main identification challenge is that workers with rising value of time moving to locations with shorter commute time may not only be driven by their desire to reduce commute time but also could be driven by an increase in their desire for city-type amenities. While central locations typically provide shorter commute time to work, central locations also provide access to more variety of consumption venues such as restaurants and stores than living in suburban neighborhoods does (Couture and Handbury 2017). So when one observes workers' growing demand for central locations with shorter commute time, it is difficult to identify whether the desire to cut commute time is the factor behind their changing location demand or their growing affinity for city amenities is the true driver of such change. If workers with rising value of time also tend to have growing affinity for city amenities, my estimate would be biased.

To address this challenge, I exploit the fact that even within the same residential neighborhood, commute time can be dramatically different across different occupations. This is because jobs in different occupations are located differentially across space. For example, financial managers and medical doctors are both highly skilled workers. However, financial jobs are very concentrated in lower end of Manhattan while doctors in New York City area can work in any number of locations because clinics and hospitals are spread throughout the metropolitan area. I can therefore use the variation in expected commute time across occupations to identify the effect of commute cost on workers' valuation of neighborhoods, controlling for other neighborhood characteristics that are occupation-invariant.

To implement this neighborhood-by-occupation identification strategy, I exploit neighborhood-level, geocoded job location data for each occupation. I combine this dataset with within-city travel distance and a travel time matrix (by driving) generated by Google Distance Matrix API and National Household Travel Survey data. This allows me to compute every neighborhood's expected commute time for workers in each occupation. I find that workers with a rising value of time increasingly prefer neighborhoods with shorter commute time to the work locations of *their own* occupations, *holding constant* the commute time to work locations in occupations unrelated to their own.

The other empirical challenge I face is measuring the endogenous change in local amenities and estimating workers' preference for these amenities. While some local conditions (population and rent, for example) can be measured precisely, amenities are harder to measure. I can count the per capita number of restaurants, bars, gyms, parks, or measure local crime rates, but it is unclear how to aggregate up these numbers into a uni-dimensional measure of amenity level. On the other hand, I want to capture the information in the content of local amenities that can be affected by local residential mix. My approach is to use the ratio between local high- and low-skilled residents to approximate the level of amenities in a neighborhood. I use the local population change in high-skilled and low-skilled workers predicted by the shock to the value of time as the source of variation for the change in local skill ratio to identify workers' preference for local amenities.

After estimating the spatial equilibrium model of worker's location choice, I conduct a series of decomposition exercises to pin down the mechanisms of gentrification. Specifically, I decompose the spatial sorting of workers into components driven by (i) the direct effect of the shock to the value of time, and (ii) the indirect effect induced by the endogenous change in amenity levels. I find that the changing value of time alone can strongly predict the change in central city skill ratios, but underestimates the overall magnitude. Once the endogenous change in amenity levels is added, the predicted magnitude becomes much closer to data. This means that the rising value of time is likely an important driving force of gentrification, but its effect is greatly magnified by the effects of endogenous amenity improvement.

Finally, using the estimated model, I evaluate the welfare impact of the spatial sorting on high-skilled and low-skilled workers. I find that the change in the value of time, rents, and amenities leads to an increase in welfare gap equivalent to 8.41 percentage point in the earnings gap, which is 33% larger than the increase in the actual earnings gap.

This paper is related to several literatures. First, the paper contributes to the literature that examines the mechanisms behind the striking phenomenon of urban gentrification in the United States. Edlund, Machado, and Sviatchi (2015) is the first paper that examines how high-skilled workers' decreasing tolerance toward commuting induces them to move to the central cities, leading to gentrification. Inspired by their insights, my paper uses a spatial equilibrium model to demonstrate how the mechanisms plays out through rising value of time and endogenous amenity change, and I use a novel identification strategy to empirically pins down the each of the mechanisms. Many alternative hypotheses have been examined by prior papers. Brueckner and Rosenthal (2009) examine the role of the aging cycle of housing stock in urban gentrification. Baum-Snow and Hartley (2017) and Couture and Handbury (2017) both find that amenity change and high-skilled workers' valuation in amenities are important in explaining the recent changes in central cities. Couture, Gaubert, Handbury and Hurst (2018) demonstrate that income growth of the high-income workers and their non-homothetic preference for luxury urban amenities are significant forces that gentrify the city centers. Ellen, Horn, and Reed (2015) examine the role of crime reduction, which is another important exogenous force that generates inflow in high-skilled residents.

This paper also provides insight on the welfare consequence of spatial sorting caused by gentrification. I find that the welfare gap between high- and low-skilled workers increases by more than the income gap alone if I account for the gentrifying forces in my spatial equilibrium model. Couture, Gaubert, Handbury and Hurst (2018) similarly show that the increase in welfare inequality is larger once spatial sorting is accounted for, despite using a different model and method. In both our models, spatial sorting resulted from differential preference for central city amenities by skill or income is the key mechanism that generates widening welfare inequality. Other studies like Vigdor (2002), Ellen and O'Regan (2011) and Brummet and Reed (2018) find no evidence that the incumbent poor are displaced by gentrification. Using longitudinal data, Brummet and Reed find that gentrification raises the welfare of low-income homeowners but have little impact on the welfare low-income renters. Vigdor argues that gentrification brings various forms of benefits to the original residents and finds no evidence that gentrification negatively impacts the welfare of original residents. My paper differs from these papers because my focus is the overall welfare consequence for all residents, and not particularly the incumbent residents living in gentrifying neighborhoods.[1]

This paper also contributes to how neighborhood amenities change in response to changes in location demand and how, conversely, these neighborhood amenities affect how residents choose locations. Many papers highlight the role of amenities in the spatial economy (Glaeser, Kolko, and Saiz (2001), Bayer, Ferreira, McMillan (2007), Guerrieri, Hurst and Hartley (2011), Diamond (2016), Handbury (2013), Couture (2016), Couture and Handbury (2017), Davis, Dingel, Monras, and Morales (2017), Autor, Palmer, and Pathak (2017)). Glaeser, Kolko, and Saiz (2001) argue that cities are attractive to workers not only because they offer higher wages but also because their consumption amenities are greater. Another example is Guerrieri, Hurst, and Hartley (2011) who show that when cities experience positive labor demand shocks, incoming residents tend to demand housing near areas that were initially wealthy. In this paper, I use a method of modeling neighborhood amenities and identifying a worker's preference for amenities that is similar to the method used by Diamond (2016). However, amenities are modeled at city level in Diamond's paper whereas they are modeled at the neighborhood level in this paper.

This paper is also related to the study of equilibrium models of urban structure developed in the urban economics literature (Mills (1967), Brueckner, Thisse, and Zenou (1999), Lucas (2001), Lucas and Rossi-Hansberg (2002), Rossi-Hansberg (2004), Monte, Redding, and Rossi-Hansberg (2017), Ahlfeldt, Redding, Sturm, and Wolf (2015), Albouy and Lue (2015)). I contribute to this literature

[1] Even though there is weak evidence that incumbent residents in the gentrifying neighborhoods are displaced, the overall share of low-skilled residents living in the central cities has declined dramatically. This means many low-skilled residents choose to live in the suburbs. Such choice is affected various factors and how their choices are affected can have welfare consequences. My paper focuses on the overall welfare effect of gentrification, whereas the longitudinal studies focus more on the welfare effect on the incumbent residents living in the gentrifying neighborhoods.

by showing how the change in the value of time leads to the sorting of residents within cities.

Finally, this paper is closely linked to the literature on time-use. A number of papers have studied the effect of workers' opportunity cost of time on intra-household or intra-personal time allocation between market work time and home production (Aguiar and Hurst (2007), Becker (1965), Benhabib, Rogerson, and Wright (1991), Goldin (2014), Nevo and Wong (2017)). My paper extends the analysis by investigating how the opportunity cost of time affects location choice and the housing market. I am particularly grateful for the contributions of Kuhn and Lozano (2008) who document the changing working-hour pattern among high and low-income workers in the U.S.. My paper, which shows that the rising value of time has an effect on workers' location choice and urban structure, clearly depends on their work.

The remainder of the paper proceeds as follows. Section 2 describes the data. Section 3 presents descriptive patterns from the data. Section 4 describes the spatial equilibrium model. Section 5 discusses the estimation methodology. Section 6 presents the results. Section 7 analyzes the determinants of gentrification. Section 8 discusses the welfare impact of gentrification. Section 9 presents the conclusion.

2 Data

The main datasets I use are the U.S. Decennial Census data for year 1990 and the American Community Survey (ACS) of 2007-2011. The 5% Integrated Public Use Microdata Series (IPUMS) dataset provides Census and ACS microdata at the individual level for a large variety of demographic and economic variables, such as income and occupation (Ruggles et al. (2017)). IPUMS also provides geocoded microdata down to the level of Public Use Microdata Areas (PUMA), which is useful for computing changing location demand for central cities in various demographic subgroups. I also use IPUMS' national sample to estimate the value of time for each occupation.

Another data source for Census and ACS data is the National Historical Geographic Information System (Manson et al. (2017)). The NHGIS provides summary files of the Decennial Census and the ACS at the census tract level for 1950, 1960, 1970, 1980, 1990, 2000, and 2007-2011. This dataset enables me to analyze post-war trends in suburbanization and subsequent gentrification at the neighborhood level. NHGIS data also enable me to track workers' occupations affiliations at the census tract level, which I use to construct location choice probabilities for each census tract by workers in each occupation.[2]

I use Zip Code Business Patterns (ZCBP) data provided by the U.S. Census Bureau to measure spatial distribution of jobs in each occupation in 1994 and 2010.[3] The ZCBP is a comprehensive dataset at Zip Code Tabulation Area (ZCTA) level, developed from the Census's Business Register.

I measure commute time between each residential location and each potential work location

[2] Specifically, I use 1990 and 2007-2011 summary file data which provide the count of people in each occupation group at census tract level, and impute a detailed occupation count at census tract level in combination with IPUMS microdata at PUMA level. The imputation procedure is detailed in Appendix section B1.

[3] The employment location imputation procedure is described in Appendix section B2.

within any given MSA. First, I use Google API to compute travel time[4] and travel distance from every census tract to every ZCTA (Zip code) centroid within each MSA. I adjust for historical traffic conditions using an auxiliary dataset, the 1995 National Household Travel Survey (NHTS).[5]

3 Descriptive patterns

In this section, I document a few stylized facts that describe the gentrification patterns and time use patterns observed since 1990. These facts motivate the setup of the empirical spatial equilibrium model used in Section 4.

The growth of household income and home value in central city neighborhoods far outpaces that in suburban neighborhoods in the past three decades, which reverses decades of declining trends in central city neighborhoods. As shown in Figure 1, the ratio between average household income in central city neighborhoods (within 5 miles of the geographic pin of downtown by Google Map for the top 25 most populous MSAs) and suburban neighborhoods drops to its lowest value in 1980, and the home value ratio between central city and suburban neighborhoods drops to its lowest value in 1970 and remains relatively low, until both income ratio and home value ratio shoot up after 1990.[6]

High-skilled residents are increasingly living in central city neighborhoods, even though the locations of high-skilled jobs have not been centralizing. Figure 2a is a binscatter plot between the share of residents living in central city neighborhoods in 1990 and 2010. The plot shows that degree of residential concentration rose significantly for workers in high-skilled occupations (occupations with >=40% of college graduates in the 1990 Census), while the degrees of residential concentration generally declines for low-skilled occupations. But the binscatter plot in Figure 2b shows that the degree of concentration of job locations is slowly decreasing over time, and high-skilled jobs do not exhibit particularly different sorting patterns than do low-skilled jobs. These observations show that the increasing residential demand for central city neighborhoods is unlikely to be driven by sorting of jobs.

High-wage workers are increasingly likely to work long hours, while working long hours become less common for low-wage workers. Meanwhile, high-wage workers experience much slower growth in commute time than lower-wage workers. The change in central cities around 1990-2010 is accompanied by the reversal of work-hour patterns in both

[4] Travel time is computed with the traffic feature turned off.

[5] I do so by estimating a travel-speed model based on route distances and location characteristics of each trip's origin/destination, with trip samples that take place at rush hour during weekdays in 1995 (U.S. Department of Transportation (2009), Couture (2016)). A detailed description of how I generate the travel matrix is included in the Appendices B3 and B4.

[6] The terms "gentrification" or "urban revival" may give the impression that central neighborhoods are now seeing faster overall population growth than the suburbs. However, while central neighborhoods may be gaining in terms of absolute population, they have not gained in terms of shares of overall MSA population, since population growth in the suburbs continues to outpace that in central cities. American cities overall were still suburbanizing as recent as from 2000 to 2010, but at a much slower pace. Figure A6 in the appendix shows the share of central neighborhoods' population as a percentage of total metropolitan population in the 25 most populous MSAs. The revived demand for central neighborhoods comes primarily from high-income workers and not all workers.

high-wage population and low-wage populations. Before 1990, high-wage workers in general were less likely to work long hours than low-wage workers (Kuhn and Lozano (2008)). However, by 2010, high-wage workers are more likely to work long hours than low-wage workers, reversing the relationship between wage and work hours observed before 1990. Figure 4a shows the relation between (average hourly earnings) wage decile and percentage of workers working at least 50 hours a week in 1980 and 2010,[7] using Census data (Kuhn and Lozano (2008) document similar patterns from 1979 to 2006 using CPS data).

In addition to using Census/ACS data, I use the Current Population Survey to show this dramatic reversal in the context of a long-run trend. For each year, I compute the probability of working long-hours by using a three-year moving sample. I restrict the sample to male workers aged 25-65 working at least 30 hours a week (full time workers). In Figure 3, I plot the probability of working long hours for workers in the top and bottom wage deciles respectively. Consistent with the Census data, low-wage workers were more likely to work long hours prior to 1980. Since then, low-wage workers are increasingly less likely to work long-hours. In contrast, high-wage workers' probability of working long hours remains stable before the early part of 1980s. Between the mid-1980 and late 1990, high-wage workers' probability of working long hours increased dramatically. After 2000, the trend for high-wage stablized. This is consistent with the number documented in the Census/ACS data.

Notably, by the early 1990s, job locations for most occupations are much more concentrated near CBD areas than are residential locations, despite decades of the suburbanization of *residential* locations.[8] The increasing prevalence of working long hours among high-skilled workers since the 1980s, coupled with the fact that job locations are highly concentrated in central city locations, raises the possibility that the rising cost of time among high-skilled workers may have driven up their demand for housing in central city neighborhoods, due to the shorter expected commute time to work at central city locations.

Consistent with this conjecture, I show in Figure 4b that while commute time in all wage group has increased between 1980 and 2010, the growth in higher wage groups is considerably slower. This suggests that a substantial portion of higher-wage workers have re-optimized their location in favor of shorter commute time.[9]

[7] I restrict my sample to workers who report working no less than 30 hours (to avoid overestimating wage due to measurement error in reported work hours). In 1980 and 2010 respectively, I put each worker's wage into wage decile bins, and for each bin, I compute the percentage of workers who report working more than 50 hours a week. In calculating the percentage of long hour workers, I restrict the sample to males, aged 25-65, who work at least 30 hours per week. I exclude females from this calculation because I want to avoid the increase in female labor participation, which could confound the statistics.

[8] Figure A7 in the appendix shows the degree of job and residential concentration in central cities by occupation. Job locations by industry and occupation can be highly clustered and sticky to locations due to agglomeration and coagglomeration effects, as demonstrated by Ellison and Glaeser (1997), Rosenthal and Strange (2004), and Ellison, Glaeser and Kerr (2010).

[9] Interestingly, in the two decades between 1980 and 2000, the negative relation between growth of commute time and wage decile is very strong, while the relation is weakly positive between 2000 and 2010. This further suggests that the incentive to reduce commute time is likely an important initial reason why central cities became desirable among the skilled workers. Once the amenities started to improve and the feedback mechanism kicks in, the role of improving amenities in the central cities becomes gradually more important in attracting high-skilled workers than

Local amenities tend to improve in neighborhoods with a rising share of high-skilled residents. Furthermore, the change in skill mix of central city residents could have further increased the appeal of central city locations for high-skilled workers. Diamond (2016) and Couture and Handbury (2017) show that the share of educated residents in a city and neighborhood is correlated with the level of local amenities. Similarly, I find that an increasing presence of high-skilled workers is accompanied by improvement in local amenities, such as the quality of law enforcement and variety of consumption venues such as restaurants.

In Table 1 column (1)- (4), I show the relationship between log per-capita counts of four types of consumption establishments (restaurants, grocery stores, gyms, and personal services) and the changes in log skill ratios at census tract level. Results show that the census tracts that see stronger growth in local skill ratio tend to also experience stronger growth in the variety of consumption amenities. In columns (5) and (6), I show the relationship between changes in log crime rates and changes in log skill ratios at the municipal level, and find that stronger growth in skill ratios is associated with declining crime rates.

4 Spatial equilibrium model of residential choice

To understand the mechanism of workers' location sorting generally, I build a spatial equilibrium model in which I model workers' neighborhood choice as a function of their value of time, commute time, neighborhood amenity and rent, where amenity and rent can endogenously adjust in equilibrium. The model is intended to parsimoniously capture the effect of the changing value of time and the endogenous change in local amenity and rent on spatial sorting.

4.1 Value of time

The key mechanism in the model is that a rising value of time may lead to increasing demand for housing in the central city due to the rising cost of commute. To start, I discuss the concept of value of time in the context of commuting decision and then how I model it.

When workers' residential locations and work locations differ, they typically have to allocate a certain fraction of their time to commuting activities. Time spent on their daily commute is costly to workers, because that time could have been used as productive work hours or leisure time. The opportunity cost of commuting is often featured in urban models, such as Alonso-Muth-Mills model, as a force that counteracts people's desire for more spacious housing in the outer suburbs (Brueckner (1987)) (Glaeser, Kolko, and Saiz (2001)).

The best way to capture the cost of time spent on commuting is to look at the value of time a worker could have used in other activities. I assume that the fundamental source of a worker's value of time comes from the marginal earnings that the worker would realize by spending more

shorter commute time. In fact, the self-sustaining endogenous improvement in amenities in the central cities would lead to the rising prevalence of reverse-commute, which explains the slight positive relationship between growth in commute time and wage decile between 2000 and 2010. I discuss more supporting evidence in appendix section C4.

hours working.[10] In case of wage workers, the extra wage earnings the workers would receive if they devoted more hours to work would coincide with his value of time.

However, for non-wage workers, who are paid with fixed salaries, commissions, or more complex forms of compensation schedules, their pay may not be a linear function of their hours worked. Consider a teacher who works in a K-12 school and receives a fixed salary for 30 hours of weekly teaching obligations. Working more hours than 30 hours (e.g., spending extra time helping students with homework) would not necessarily increase earnings. In contrast, for a financial manager, receiving a bonus and or getting a promotion may depend crucially on the hours and effort devoted to the job. As a result, the financial manager's marginal incentive of hours supply may even exceed the average hourly earning and may be compensated disproportionately if she works longer hours (Goldin (2014)).

To capture such differential incentives to supply hours at the intensive margin, I use the concept of long-hour premium, which measures the percentage *return* of working extra hours (Kuhn and Lozano (2008)). In the model description, I discuss how the long-hour premium fits into workers' location demand.

4.2 Location demand

Given the worker's occupation k and city m where she lives and works, a worker who choose to live in neighborhood j and works in neighborhood n in year t enjoys utility:[11]

$$U\left(C,H,L,A_{jmt}\right)=C^{\theta_C}H^{\theta_H}L^{1-\theta_C-\theta_H}A_{jmt}^{\tilde{\gamma}_k}\exp\left(-\tilde{\omega}_t c_{jnmt}\right)\exp\left(\sigma\varepsilon_{i,jmt}\right) \tag{1}$$

subject to budget constraint

$$C+R_{jmt}H=\exp(y_{0mkt}+v_{mkt}\left(T-L-c_{jnmt}\right)).$$

C is consumption; H is the housing service; L is weekly hours of leisure time; A_{jmt} is the amenity level for neighborhood j at time t;[12] $\tilde{\gamma}_k$ is the taste parameter for local amenities, which may differ by worker type; c_{jnmt} is the weekly commute time between residential location in j and work location in n. $\tilde{\omega}_t$ is a time-variant aversion parameter for commute time. $\varepsilon_{i,jmt}$ is the idiosyncratic preference component for individual i, which I assume to be distributed as Type I Extreme Value, and σ is

[10] Under the assumption that work hours and leisure hours can be easily reallocated within a worker, the marginal value of leisure hours would equal the marginal value of work hours. In that case, a rise in the value of work hours would imply the value of leisure hours rises at the same rate.

[11] I use MSAs to represent cities. Given the choice of an MSA, a worker can choose which neighborhood to live in within that MSA. The reason I use MSA as a city unit for the analysis instead of commuting zones (CZs) is that CZs are constructed at a lower geographic level. For example, Jersey City, NJ belongs to the Newark CZ, which is different from the New York CZ, even though commute time from Jersey City to downtown New York is around 10 minutes. The New York MSA, on the other hand, covers both Newark and New York CZs. In this model, I would want workers who work in downtown New York to have the choice to live in Jersey City, NJ. Therefore, in the context of this analysis, MSA is a more natural choice.

[12] I allow the log transformed amenity level to be decomposed into a uni-dimensional observable amenity level and an unobservable component: $\log(A_{jmt})=a_{jmt}+\zeta_{jmt}$.

its standard deviation. I normalize the price of consumption good C to be 1, and I let R_{jmt} be the rent for housing services in j at time t. The worker's main source of income is labor earnings. Worker's weekly log earnings is a linear function of her base pay of working full time and the extra hours supplied in excess of the 40 hours full-time level. y_{0mkt} is the basic log income the worker would receive if she were to supply only the minimum 40 hours of work or less; v_{mkt} measures the log weekly earnings from each extra hour of work supplied in a week, or "long-hour premium". T is the worker's total possible hours supplied beyond 40 hours per week. If commute time c_{jnmt} is zero, all of T would be devoted to either work or leisure. As c_{jnmt} get large, a lower number of hours would be available, which would lower weekly earnings or leisure.[13] The negative impact of commute time c_{jnmt} on log earnings (or utility) is larger if long-hour premium v_{mkt} is larger.

Each worker solves the utility maximization problem by choosing C, H, L, conditional on his/her occupation and the locations he/she lives and works in. Derivation of the indirect utility is detailed in Appendix section A.

In the indirect utility, leisure choice is substituted by city/occupation/time fixed effects.[14] In addition, I normalize the indirect utility function by dividing it by σ, the standard deviation of the idiosyncratic preference component. δ_{mkt} contains all of the fixed-effects that are city/occupation/time specific. The normalization of coefficients allows me to interpret the coefficients as migration elasticities. The indirect utility becomes

$$V_{i,jnmt} = \delta_{mkt} - \mu v_{mkt} c_{jnmt} - \omega_t c_{jnmt} - \beta r_{jmt} + \gamma_k a_{jmt} + \gamma_k \zeta_{jmt} + \varepsilon_{i,jmt}.$$

Worker i then chooses residential neighborhood j within MSA m to maximize indirect utility. Since $\varepsilon_{i,jmt}$ is distributed as Type I Extreme Value, the probability that worker i would choose neighborhood j is given by a multinomial logit function (McFadden (1973)). Given city m where a worker lives and works, the worker's occupation k, and the neighborhood n which the worker works in, the probability of that worker choosing to live in neighborhood j is given by

$$s_{j|nmkt} = \frac{\exp\left(\tilde{V}_{jnmkt}\right)}{\sum_{j' \in J_m} \exp\left(\tilde{V}_{j'nmkt}\right)}$$

where $\tilde{V}_{jnmkt} = \delta_{mkt} - \mu v_{mkt} c_{jnmt} - \omega_t c_{jnmt} - \beta r_{jmt} + \gamma_k a_{jmt} + \gamma_k \zeta_{jmt}$ is the mean utility of occupation k living in j and working in n.

I assume that during the time period of the analysis, the spatial location of employment of each occupation within cities is sticky, and the cross-sectional variation in job location is driven

[13] Long commute could dip into people's work hours, which lowers earnings. Another possible hypothesis is that long commute time may not necessarily dip into a worker's work hours directly, but may instead eat into the worker's leisure hours. The predicted effect of value of time on locational sorting is robust to this assumption. Under the assumption that work hours and leisure hours can be easily reallocated within a worker, the marginal value of leisure hours would equal the marginal value of work hours. In that case, a rise in the value of work hours (long-hour premium) would imply that the value of leisure hours rises at the same rate, which would generate the same effect on location choice, even if the worker decides to keep his/her work hours unchanged.

[14] Assuming non-corner solution, leisure is a function of the value of time, which is city/occupation/time-specific.

by factors such as path-dependent patterns of industry clustering and firm agglomeration (Ellison and Glaeser (1997), Rosenthal and Strange (2004), Ellison, Glaeser and Kerr (2010)).[15] I take the spatial distribution of jobs for each occupation as exogenous to the model within the time frame of this analysis.

If I observe the residential location choice conditional on work location in the data, I can back out $\tilde{V}_{jnmkt}$ directly from the data, and model the mean utility directly. Unfortunately, I only observe unconditional location demand s_{jmkt}. To proceed, I assume, in equilibrium, for workers in each occupation k, the unconditional expected utility of working in any neighborhood n within the MSA is identical and remains identical over time. Under this simplying assumption, I essentially take a partial equilibrium framework in which a firm's location choice would not be affected by the change in residential sorting over the period of the analysis. I denote the expected utility value of working in each neighborhood in MSA m as Λ_{mkt}. The worker's conditional residential choice probability is then given by the following equation:[16]

$$s_{j|nmkt} = \exp\left(\tilde{V}_{jnmkt} - \Lambda_{mkt}\right)$$

Given the residential choice probability conditional on working in n, the unconditional residential choice probability is computed by weighting these conditional probabilities with the unconditional probability of working in neighborhood n in MSA m, which I denote as π_{nmkt}. Thus, the residential choice probability is:[17]

$$s_{jmkt} = \sum_{n' \in J_m} \pi_{n'mkt} \cdot s_{j|n'mkt}$$

After log transformation, I write the log location choice probability as a linear function of various location preference components.

$$\log\left(s_{jmkt}\right) = \underbrace{\tilde{\delta}_{mkt}}_{\text{fixed effects}} + \underbrace{\log\left(\sum_{n' \in J_m} \pi_{n'mkt} \exp\left(-\left(\omega_t + \mu v_{mkt}\right) \cdot c_{jn'mt}\right)\right)}_{\text{valuation of proximity to employment}} \quad (2)$$
$$- \underbrace{\beta r_{jmt}}_{\text{valuation of rent}} + \underbrace{\gamma_k a_{jmt}}_{\text{valuation of amenities}} + \underbrace{\gamma_k \zeta_{jmt}}_{\text{valuation of unobserved amenity}}$$

s_{jmkt} is the probability of choosing neighborhood j by workers in occupation k living in MSA m in year t. As can be seen in the location demand equation, the worker places positive value on the proximity to job locations, positive value on neighborhood amenities, and negative value on

[15] One example to illustrate this point is the concentration of financial-industry jobs in Lower Manhattan. This area has a high presence of financial jobs because financial firms are historically clustered around the southern tip of Manhattan, not because the southern tip of Manhattan is an ex-ante desirable place for financial workers to live.

[16] $\Lambda_{nmkt} = \log\left(\sum_{j' \in J_m} \exp\left(\tilde{V}_{j'nmkt}\right)\right)$, which is the expected utility for worker in occupation k working in neighborhood $n(m)$. Under the assumption that the expected utility of working in each location is identical, I set $\Lambda_{nmkt} = \Lambda_{mkt}$. Due to the limitation of the unconditional location choice data, this simplifying assumption is necessary.

[17] $\tilde{\delta}_{mkt} = \delta_{mkt} - \Lambda_{mkt}$

neighborhood rents. The value of proximity to employment is particularly important, because it captures the key sorting mechanism by which workers with higher value of time choose locations closer to their workplace in terms of travel cost.

The specification of the valuation from proximity to employment is nonlinear with respect to value of time. To illustrate the marginal effect of value of time v_{mkt} on the demand for neighborhoods, I take the derivative for the $\log(s_{jmkt})$ with respect to the value of time:

$$\frac{\partial \log(s_{jmkt})}{\partial v_{mkt}} = \tilde{\delta}'_{mkt} - \mu \sum_{n'} \tilde{\pi}_{jn'mkt} \cdot c_{jn'mt}$$

where $\tilde{\pi}_{jn'mkt}$ is an adjusted probability measure as: $\tilde{\pi}_{jnmkt} = \frac{\pi_{nmkt} \exp(-(\omega_t + \mu v_{mkt}) c_{jnmt})}{\sum_{n' \in J_m} \pi_{n'mkt} \exp\left(-(\omega_t + \mu v_{mkt}) c_{jn'mt}\right)}$.

The adjusted $\tilde{\pi}_{jnmkt}$ is the probability of working in neighborhood n by worker of occupation k who lives in neighborhood j. The adjustment takes into account of the fact that workers are less likely to work at locations too far away from home. Factoring in the adjusted probability measure, I can rewrite the marginal value of living in neighborhood j is:

$$\frac{\partial \log(s_{jmkt})}{\partial v_{mkt}} = \underbrace{\tilde{\delta}'_{mkt}}_{\text{invariant across neighborhood}} - \mu \underbrace{\widetilde{\mathrm{E}}(c_{jmkt})}_{\text{expected commute time}} \tag{3}$$

The key insight from the derivative is that higher value of time leads the average worker to be less willing to live in neighorhoods with *longer expected commute time.* Living in a remotely located neighborhood would reduce the number of hours available for working due to high expected commute time. μ governs the sensitivity of a worker's location choice to commute cost. Therefore, evaluating whether commuting cost affects location choice boils down to testing whether $\mu > 0$.

4.2.1 Endogenous amenity supply

Workers also care about neighborhood amenities. Some are in the form of natural amenities such as parks and natural sceneries (Lee and Lin (2017)); some are in the form of public goods (e.g., crime and law enforcement), and others are in the form of consumption venues such as restaurants, retail stores, fitness facilities, etc. (Couture and Handbury (2017)).

In this model, I assume that the level of consumption amenities can respond to ratio between the number of local residents who are high-skilled and low-skilled, similar to the method use by Diamond (2016). Under this assumption, a rising share of high-skilled residents in a neighborhood would lead to the entries of suppliers of local goods and services and better funding for effective local law enforcement.[18] I model amenity supply as follows:

$$a_{jmt} = \eta \ln \left(\frac{N_{jmt}^{H}}{N_{jmt}^{L}} \right) + \tilde{\theta}_t X_{jmt} + \delta_{jm} + \delta_{mt} + \xi_{jmt}^{a} \tag{4}$$

N_{jmt}^{H} and N_{jmt}^{L} are the counts of high- and low-skilled workers living in neighborhood j. η represents the amenity supply elasticity with respect to the local skill ratio. X_{jmt} represents other observable neighborhood characteristic that workers may value, and I allow them to contribute to the amenity level at rate $\tilde{\theta}_t$. δ_{jm} represents census tract fixed-effects, and δ_{mt} represents MSA/time fixed-effects. ξ_{jmt}^{a} represents the component of amenity supply that is unobservable and cannot be accounted for by the local skill ratio. This may include amenities of a cultural and/or historical nature, which would affect neighborhood amenities regardless of the inflow and outflow of local residents.

Since a key driver of amenity supply is the ratio of high-skilled residents and low-skilled residents, I endogenize amenity levels into workers' location demand function by directly modeling location demand as an iso-elastic function of local skill ratios, governed by a reduced-form migration elasticity parameter γ_k. Ideally, I would like to model neighborhood amenity directly. However, neighborhood amenities are inherently multi-dimensional, and it is unclear how to aggregate various amenity variables into a uni-dimensional one. Local skill ratio provides a uni-dimensional measure of local amenity while capturing the content of amenity that is driven by changing local population. Alternatively, I could model each type of amenity in the model. Using that approach, I would face identification challenge. To separately identify preference parameters for different types of amenities (law enforcement, consumption venues, and public infrastructure), I need identifying variations for each one of these amenities.[19]

Instead of modeling amenities directly into the equilibrium framework, I create measurements of crime and consumption venues later in the paper,[20] and provide evidence that these amenity levels do respond to shocks to local skill ratios.

By plugging the amenity supply function into location demand, I get the following equation:

$$\begin{aligned} \log\left(s_{jmkt}\right) \quad = \quad & \tilde{\delta}_{mkt} + \log\left(\sum_{n' \in J_m} \pi_{n'mkt} \exp\left(-\left(\omega_t + \mu v_{mkt}\right) \cdot c_{jn'mt} \right) \right) - \beta r_{jmt} \\ & + \gamma_k \log\left(\frac{N_{jmt}^{H}}{N_{jmt}^{L}} \right) + \theta_{kt} X_{jmt} + \gamma_k \xi_{jmt}^{a} + \gamma_k \zeta_{jmt} \end{aligned} \tag{5}$$

The reduced-form migration elasticity γ_k is a combination of demand side elasticity and sup-

[18] The assumption is also consistent with what Guerrieri, Hurst and Hartley (2013) find: that at neighborhood level, people like to live close to a wealthy neighborhood, and therefore it is possible that local residential composition may directly influence people's location preference as well. Furthermore, Couture and Handbury (2017) show that the initial distribution of consumption amenity venues has some effect on highly educated people's preference for neighborhoods. It is also possible that a higher share of college graduates in a neighborhood is desirable in itself.

[19] If I create an amenity index that measures overall local amenity level a_{jmt}, I would still have to take a stance on how different measures of amenities ought to be aggregated, and it is difficult to favor one method over another.

[20] I will show amenity response elasticities for different amenities separate from the model to provide a full picture of the nature of amenity response.

ply side elasticity, namely $\gamma_k \eta$; this is a sufficient statistic that can pin down the mechanism of the endogenous amenity change. $\theta_{kt} X_{jmt}$ is the component of amenities that is observable and exogenous.[21] $\gamma_k \xi^a_{jmt}$ is the component of amenities that does not covary with local residential composition. Since $\gamma_k \xi^a_{jmt}$ and $\gamma_k \zeta_{jmt}$ are both unobservable, I denote the sum of the two terms as ξ_{jmkt}.

4.3 Housing supply

Housing services in each neighborhood are provided through the housing rental market. I assume the housing market is competitive, and home prices equal to the marginal cost of new construction, and rents are proportional to house prices.

I assume that log rent is a reduced-form function of local demand for housing, and the existing housing stock density within one mile of the neighborhood of interest (approximate the cost of construction). I approximate local housing demand by the aggregate income of residents in the neigborhood, which is $\sum_k \bar{Y}_{mkt} N_{jmkt}$. Additionally, I assume a national housing demand shock, captured by ι_t. I allow rent to respond positively to changes in location demand. I further assume that the density of housing stock would sharpen the rent response. For that purpose, I assume that the rent elasticity with respect to housing demand is a function of initial housing density. Therefore, the size of rent elasticity depends on housing stock density den_{jm} around neighorhood j. The following is the housing supply equation:

$$r_{jmt} = \pi den_{jm} \log \underbrace{(D_{jmt})}_{\text{housing demand}} + \xi^r_{jmt} \tag{6}$$

$$D_{jmt} = \exp(\iota_t) \sum_k \bar{Y}_{mkt} N_{jmkt}$$

πden_{jm} represents the inverse elasticity of housing supply at local level. I standardize den_{jm} with mean and standard deviation of housing stock densities across neighborhoods. ξ^r_{jmt} represents unobserved housing supply components, such as change in construction costs specific to neighborhood j but unrelated to initial housing stock density.

4.4 Equilibrium

Equilibrium is defined as the residential demand for each neighborhood by workers in each occupation k in each city m in each year t, s_{jmkt}, as well as rent r_{jmt}, such that the amenity market and housing market clear in each neighborhood (census tract):

1. Amenity market clears in each census tract:

The number of high-skilled workers and low-skilled workers living in a neighborhood is determined by the location choice of workers of different occupations. For notation purposes, K_H is the

[21] The parameter θ_{kt} is a reduced-form combination of demand side and supply side parameters. $\theta_{kt} = \gamma_k \tilde{\theta}_t$.

set of "high-skilled" occupations, which I define as occupations in which more than 40% of workers are college graduates in 1990 Census. K_L is the set of "low-skilled" occupations, which are occupations that are not in K_H. The number of high-skilled and low-skilled workers living in j is driven by the locational sorting of each occupation. The following equations describe how the high- and low-skilled populations are defined.

$$N_{jmt}^{H} = \sum_{k \in K_H} N_{jmkt}$$

$$N_{jmt}^{L} = \sum_{k \in K_L} N_{jmkt}$$

The amenities market clears if local skill ratios (amenity supply) lead to location choices such that the resulting local skill ratios (amenity demand) are identical.

2. Housing market clears in each census tract.

I cannot solve the system of equations analytically. Even so, the equilibrium framework is useful when I estimate the model parameters. In the estimation section, instrumental variables will be constructed using the framework from the model.

5 Estimation

5.1 The long-hour premium

In the model, I use long-hour premium to capture workers' value of time. The goal is to measure the differential changes in long-hour premium for workers in different occupations. With the estimated long-hour premium as a measure for the value of time for each occupation, I would then estimate how changing value of time affect workers' location choice. I take the exact labor earnings function introduced in the model to the Census data to estimate long-hour premium for each occupation in 1990 and 2010.

$$\log Y_{ikt} = y_{0kt} + v_{kt} hour_{ikt} + \boldsymbol{\delta}_{demo,it} + u_{ikt}. \tag{7}$$

I denote the variable *hour* as weekly work hours in excess of 40 hours. y_{0kt} is the log weekly earning the worker would earn if she worked 40 hours/week. v_{kt} is the marginal log weekly earnings on any hour worked in excess of 40 hours per week. $\boldsymbol{\delta}_{demo,it}$ is the vector of demographic dummies. u_{ikt} is the idiosyncratic earnings component. v_{kt} can be interpreted as the *percentage* of extra earnings that workers in occupation k can receive if he/she works one extra hour beyond 40 hours/week, and thus captures the value of time. Note that for workers who are paid a standard hourly wage, v_{kt} should remain roughly constant even if their wage increases, because wage workers are paid at a constant proportion of the hours worked. If v_{kt} rises over time, it means that workers are increasingly paid disproportionately more than before.

Similar to Kuhn and Lozano (2008), I estimate long-hour premium v_{kt} using cross-sectional data on log earnings and hours within each occupation, controlling for individual workers' characteristics in the Census microdata. Since hours worked is a labor supply choice variable, estimates of the long-hour premium may be driven by a selection effect by unobserved workers' ability. I describe in detail how I address endogeneity concerns in Appendix section D2.

To establish intuition for how the cross-sectional relationship between log earnings and hours worked can pin down the long-hour premium, I show in Figure 5 the plots between residual log weekly earnings and hours worked (the level of earnings is normalized such that log earnings at standard full time 40 hours is zero in each year) for four occupations. For financial workers and lawyers, the slope rises dramatically from 1990 to 2010, which means that workers in these occupations are increasingly disproportionately compensated for working longer hours. In contrast, for other occupations such as office secretaries and teachers the slope of the plot remains largely unchanged, despite increases in average hourly earnings over time. The computed long-hour premium is shown in Table A9 in the appendix.

Appendix section D4 talks about various validation tests performed for the long-hour premium as a measurement of the value of time.

5.2 Location demand

The key parameters to identify in the location demand function are μ, β, γ_k. I simplify γ_k to differ only by skills (high or low), or γ_z, $z \in \{high, low\}$. I therefore assume that workers in high-skill occupations potentially have different tastes for local amenities than workers in low-skill occupations. I also simplify the estimation by assuming that commute time and employment location remain unchanged over time, and assume the travel time matrix and employment locations to be fixed at the 1990/1994 value.[22]

Note that in specification (5), μ enters the equation nonlinearly. To simplify specification of the average worker's valuation of neighhorhoods, I use a Taylor approximation so that the location demand equation is a linear function of μ.[23] In addition, I estimate the long-hour premium v_{mkt} for each occupation in each MSA in each year, using microdata that *exclude* the state MSA m is in and the neighboring states. The following is the linearized location demand:

$$\log\left(s_{j(m)kt}\right) = \delta_{jmk} + \tilde{\delta}'_{mkt} - \omega_t \widetilde{\mathrm{E}}\left(c_{jmk}\right) - \mu \hat{v}_{mkt} \widetilde{\mathrm{E}}\left(c_{jmk}\right) - \beta r_{jmt} + \gamma_z \log\left(\frac{N^H_{jmt}}{N^L_{jmt}}\right) + \theta_{kt} X_{jmt} + \xi_{jmkt} \tag{8}$$

As a result of the Taylor approximiation, $\widetilde{\mathrm{E}}\left(c_{jmk}\right)$ is evaluated with a transformed probability measure, $\tilde{\pi}_{jnmk} = \frac{\pi_{nmk}\exp(-\phi c_{jnm})}{\sum_{n' \in J_m} \pi_{n'mk} \exp\left(-\phi c_{jn'm}\right)}$, where I calibrate ϕ such that the mean commute time matches the value reported in the 1990 Census data.[24]

[22] Travel time is partially computed using the NHTS in 1995. Employment location data are imputed from the ZCBP in 1994.

[23] Derivation of the linear approximation is included in Appendix section D6.

[24] ϕ is calibrated to be 0.3425

After linearization, μ is the migration elasticity with respect to expected commute cost measured in unit of log income, which is analogous to the interpretation of the parameter in the individual worker's indirect utility. $\tilde{\delta}'_{mkt}$ is the sum of all city/occupation/time specific fixed-effects, and δ_{jmk} is the sum of all neighborhood/occupation fixed-effects, which contain the constant terms from the Taylor approximation.

I take the first-difference for the location demand equation. To avoid confounding effects of potentially endogenous changes in neighborhood characteristics, I set X_{jmt} to be the time-invariant variable in the initial year, and I set it to be occupation-specific, and the reason is explained in the next paragraph.

$$\begin{aligned} \Delta \log \left(s_{jmkt}\right) \quad = \quad & \Delta\tilde{\delta}'_{mkt} - \Delta\omega_t \widetilde{\mathrm{E}}\left(c_{jmk}\right) - \mu\Delta\hat{v}_{mkt}\widetilde{\mathrm{E}}\left(c_{jmk}\right) - \beta\Delta r_{jmt} \\ & + \gamma_z \Delta \log\left(\frac{N^H_{jmt}}{N^L_{jmt}}\right) + \Delta\theta_{kt}X_{jmk} + \Delta\xi_{jmkt} \end{aligned} \tag{9}$$

The identification of μ - To identify μ, one examines the how much workers choose locations with shorter commute time in face of shocks to the value of time. One concern for its validity is that neighborhoods with short commute time $\widetilde{\mathrm{E}}\left(c_{jmk}\right)$ tend to be close to downtown locations, and downtown locations may have certain local amenities that attract workers with strong growth in the value of time $\Delta\hat{v}_{mkt}$. To account for such unobservable city amenities, for a worker of each occupation k in each MSA, I compute each neighborhood's average travel time to job locations for occupations *unrelated to* occupation k, and I include this variable as the control variable X_{jmk}. The commute time to jobs unrelated to the occupation of the worker of interest should be irrelevant to the worker's commute cost, because the commute time for the worker of interest should depend on the job locations of her occupation, not the job locations of other occupations. But the overall travel time to job location approximate the location's general geographic centrality to the city.

I let the preference coefficient $\Delta\theta_{kt}$ on the control variable X_{jmk} be *occupation-specific* so that the control variable should absorb the sorting effect between unobserved location amenities (which covary with the control variable) and unobserved workers' characteristics. Holding X_{jmk} constant, if workers exhibit increasing preference for neighborhoods with shorter commute time to jobs of their own occupations, this should mean that workers have increasing preference for shorter commute time.

To further build intuition, a thought experiment illustrates this idea. Consider financial workers who live and work in the New York metropolitan area. If they want to shorten commute time, many of them are likely to move closer to the financial district in Lower Manhattan, which has large concentration of financial jobs. On the other hand, if doctors in the New York area want to shorten commute time, they are unlikely to move toward Lower Manhattan necessarily, because hospitals and medical clinics are much more decentralized, and Lower Manhattan does not have a similar degree of concentration for hospitals as for financial firms. My empirical method is a generalization of such thought experiment.[25]

[25] It is possible that employers' past location decision are driven by some neighborhood amenity that their potential

The identification of γ_z - Next, I describe the identification of γ_z. The key endogeneity problem of naively regressing change in location demand on change in local skill ratio is that exogenous shocks to demand for neighborhoods by high-skilled workers would change the neighborhoods' skill ratios, leading to severe reverse causality problem. Therefore, I must find instrument for skill ratio to identify γ_z.

To instrument for $\Delta \log\left(\frac{N^H_{jmt}}{N^L_{jmt}}\right)$, I create the predicted log change in neighborhood populations of high-skilled and low-skilled workers, driven purely by change in value of time and differential commute time. While constructing these instruments, I exclude populations of workers in occupations similar to the occupation of workers in question.[26] I assume that other workers' value of time and neighborhoods' commute time to unrelated jobs do not directly affect workers' location preference, but may indirectly affect workers' location choice through endogenously changing the neighborhood's residential composition.

I compute the predicted population of each occupation in each neighborhood in 2010 using only information on value of time and commute time: $\hat{N}_{jmk,2010} = N_{mk,1990} \cdot \frac{\exp\left(\log\left(s_{jmk,1990}\right) - \hat{\mu}\Delta\hat{v}_{mk,2010}\widetilde{\mathrm{E}}\left(c_{jmk}\right)\right)}{\sum_{j' \in J_m} \exp\left(\log\left(s_{j'mk,1990}\right) - \hat{\mu}\Delta\hat{v}_{mk2010}\widetilde{\mathrm{E}}\left(c_{j'mk}\right)\right)}$. where $\hat{\mu}$ is the preliminary parameter estimate from estimating the unconditional location demand equation without including amenities or rent. The predicted log population changes of high-skilled and low-skilled workers, respectively, are then $\Delta \log \hat{N}^H_{jm2010,-k} = \log\left(\sum_{\substack{k' \in K_H \\ k' \nsim k}} \hat{N}_{jmk'2010}\right) - \log\left(\sum_{\substack{k' \in K_H \\ k' \nsim k}} N_{jmk'1990}\right)$, $\Delta \log \hat{N}^L_{jm2010,-k} = \log\left(\sum_{\substack{k' \in K_L \\ k' \nsim k}} \hat{N}_{jmk'2010}\right) - \log\left(\sum_{\substack{k' \in K_L \\ k' \nsim k}} N_{jmk'1990}\right)$. I use the $\Delta \log \hat{N}^H_{jmt,-k}$ and $\Delta \log \hat{N}^L_{jmt,-k}$ as instruments to provide identifying variation for the actual change in local skill ratio. [27]

I come back to the example of financial workers and doctors in the New York area for the intuition of identifying γ_z. The cluster of financial firms in Lower Manhattan may attract financial workers to settle nearby due to their rising value of time. The increased location demand by financial workers raises the level of amenities (skill ratio) in neighborhoods in Lower Manhattan.[28] If doctors

employees may desire. If this is true, then the workers' commute time (proximity) to jobs may reflect some amenity value of the neighborhoods that appeals specifically to workers in certain occupations, making the commute time to these jobs shorter. Variation in commute time may pick up people's preference for occupation-specific amenities. This concern may be mitigated by the argument that if job locations were strategically chosen to accommodate employees' amenity preference, workers who value these amenities should have already been living in locations close to their job locations as early as 1990. The fact that the downtown area in which jobs are highly concentrated tended to be near low-amenity areas before 1990 suggests that strategic firm location decisions motivated by employees' amenity demand are unlikely to have driven variation in the initial spatial job allocation.

[26] More precisely, I consider two occupations to be similar if they belong to an occupational group defined in the "occ2010" variable in the IPUMS data.

[27] I show in Table A1 that predicted changes in log populations have strong statistical power in predicting change in log skill ratio. As expected, change in a log high-skilled population carries a positive sign, and change in a log low-skilled population carries a negative sign.

[28] The main concern in identifying for preference for amenity is that the variation in the predicted change in population skill mix may pick up the variation of proximity to job locations for workers of interest. If high-skilled jobs with high growth in value of time tend to cluster together geographically, then a high-skilled worker with rising value of time may increasingly demand locations near these clusters both to cut commute time and to take advantage of the endogenously improving amenity that results from the influx of other high-skilled workers with rising value of time. In this case, the parameter γ_z would be confounded by the changing preference for value of time and preference for amenity. To ensure that γ_z identifies only preference for amenity, I exclude data for occupations similar to that of the

systematically increase their demand for a neighborhood near a large cluster of financial firms in Lower Manhattan, such a demand response can be attributed to the increased amenity levels.

The identification of β - To identify β, the migration elasticity with respect to rent, I use the setup described in the housing supply equation (6), in which Δr_{jmt} is driven by growth in local residents interacted with existing housing stock in neighborhood j. I construct instruments for Δr_{jmt} by interacting $\Delta \log \hat{N}^{H}_{jmt,-k}$, $\Delta \log \hat{N}^{L}_{jmt,-k}$ and $\Delta \log \hat{N}^{All}_{jmt,-k}$ with initial housing stock density den_{jm} in the 1980 Census to identify preference for rent.

5.3 Housing supply

To estimate elasticities in the housing supply equation, I take the first difference.

$$\Delta r_{jmt} = \pi_1 den_{jm} \Delta \log \left(\sum_k \bar{Y}_{mkt} N_{jmkt} \right) + \pi_2 den_{jm} + \delta_m + \Delta \xi^r_{jmt} \tag{10}$$

where $\pi_1 = \pi$ and $\pi_2 = \pi \Delta \iota_t$. δ_m is the MSA fixed-effects, after differencing the MSA/time fixed effects. For the identification of inverse housing supply elasticities, I need variation that drives the change in local aggregate income that is not correlated with $\Delta \xi^r_{jmt}$, which is neighborhood-level local housing supply shock (e.g. shock to local construction cost). I use the predicted log change in population of high-skilled, low-skilled and all workers $\Delta \log \hat{N}^{H}_{jmt}$, $\Delta \log \hat{N}^{L}_{jmt}$ and $\Delta \log \hat{N}^{All}_{jmt}$ to instrument for $\Delta \log \left(\sum_k \bar{Y}_{mkt} N_{jmkt} \right)$. Note that to identify the housing supply equation, instruments do not have to exclude data from workers in the occupation of interest as in the location demand equation, because the exclusion restriction only requires that instruments are uncorrelated with $\Delta \xi^r_{jmt}$.

To separately identify other parameters, I interact the instrument with housing stock density den_{jm}.

5.4 Linear GMM estimator

I jointly estimate the location demand and housing supply equations using an efficient GMM estimator.[29] I compute Heteroskedasticity and Autocorrelation Consistent (HAC) standard errors using Conley's (1999) method to account for spatial dependence of the unobserved error terms in both equations. Due to the fact that value of time is estimated from national data, I also correct the standard errors for the generated regressor following Murphy and Topel (1985).[30]

workers of interest when constructing instruments for changes in population skill mix. The idea is that rising value of time for unrelated workers only affects the migration decision of the worker of interest through amenity change. For example, to identify preference for amenity by doctors, I construct population shocks to non-doctors, and the migration response by doctors to such shocks must be due to the impact of the amenity change brought about by the migration decisions of non-doctors, not doctors' preference for shorter commute time.

[29] The estimation procedure is described in detail in Appendix section D7.

[30] See Appendix section D7.2 for the construction of Conley standard errors and correction for the generated regressor.

6 Estimation results

6.1 Location demand

Table 2 shows the estimates of the model using data from all non-rural census tracts. The estimate for μ is positive and significant, which shows that workers with higher value of time prefer neighborhoods with shorter expected commute time.[31] The estimate for μ is 3.166 (0.403), which means that a one percentage point rise in long-hour premium would lead to workers having 31.6% higher demand for saving one hour of daily commute time.[32]

Migration elasticity for local skill ratio γ_z is 1.851 (0.164) for high-skilled workers and 0.449 (0.142) for low-skilled workers, which means that census tracts with 1% higher skill ratio would raise demand from high-skilled workers 1.402 percentage point more than from low-skilled workers. Therefore, it can be easily seen that an exogenous shock that generates a rise in local skill ratio in a neighborhood could trigger an endogenous demand response from high-skilled workers that is much larger than that from low-skilled workers, and thus further raise the local skill ratio for this neighborhood. An important implication of this estimate is that some high-skilled workers may sort into the central city neighborhood even without experiencing a value of time shock themselves, so long as other high-skilled workers who work in the central city do.

The elasticity with respect to rent β is estimated to be 0.5704 (0.190). This estimate is quite small, because the implied housing expenditure share is 0.18, which is much lower than the number reported in the Consumer Expenditure Survey (Diamond (2016)). One possible reason for this is that many households may be homeowners, who may be much less sensitive to rising rents per se, which would drive down the estimate. If all households are indeed renters, the estimate would probably be much higher. The relatively small size of β would lead to weaker response to rent hike in my later analysis using the model, but the spatial sorting mechanism is generated mainly by the differential sizes of γ_z between high- and low-skilled workers.

6.2 Housing supply

In Table 2, I find that in the face of a rise in housing demand, rent responds positively. The elasticity of rent with respect to housing demand shock is higher in neighborhoods with higher density of housing stock. The increase in elasticity with each standard deviation of housing density is 0.167 (0.0262).[33] In addition, one standard deviation higher housing density also leads to 1.23% growth in rents.

[31] The linear coefficient for $\Delta v_{mkt}\,\mathrm{E}\left(c_{j(m)kt}\right)$ is $-\mu$. A positive estimate for μ means that the linear coefficient is negative.

[32] Recall that the long-hour premium is the marginal log weekly income gained from working an extra hour beyond a 40 hours/week threshold, and the expected commute time is scaled as the total commuting hours in a week. Assuming the average commuter goes to work 5 days a week, the weekly commute time should be 10 times the one-way commute time.

[33] I standardize housing density before using it in the estimation.

6.3 Amenity supply

I model the preference for amenities as reduced-form elasticities to local skill ratio, because of the lack of identifying variation for the many dimensions of local amenities. Nonetheless, it is feasible to estimate amenity *supply* elasticities with respect to local skill ratio separately for each dimension of amenities, and show that local skill ratio is a driving force of various types of local amenities. In Table 3, I estimate elasticities for the per-capita number of various local business establishments with respect to changes to local skill ratio.[34] I instrument the local skill ratio using the same instruments used to identify the reduced-form preference parameters for local skill ratio. I find that the per capita count of local businesses is positively responsive to exogenous shock to local skill ratio. The exception is the number of grocery stores in column (2). The lack of significant response in the number of grocery stores is consistent with the finding in the cross-MSA analysis of amenity response to MSA-level of college ratio in Diamond (2016).[35] In addition, municipality-level violent crime rate is negatively affected by the rise in local skill ratio, though no significant result is found to be associated with the property crime rate. Since crime is commonly regarded as a disamenity, the result shows some evidence that a rising skill ratio improves amenities by reducing the crime rate.

6.4 Robustness

I conduct a number of robustness checks for my estimation. First, to further validate the role of value of time in shifting location demand, I estimate another specifcation in which I also model workers' migration elasticity with respect to proximity to public transit stations, and I find that workers with rising value of time also sort into neighborhoods near transit stations. Since a fraction of commuter do use public transit, such finding further validates the idea that rising value of time affect workers' location choice through the reduction in commute time. (See Appendix section E.1 for detail)

I also estimate the model using alternative specification assuming μ and β to be heterogeneous across skills. I find that each parameter is positive and significant for both high- and low-skilled workers. (See Appendix section E.2 for detail).

Then, to ensure that my estimation is not driven by the particular choice of "long-hour premium" as a measurement of the value of time, I also test for spatial sorting using altnerative measurements of the value of time, and the results still show that workers with rising value of time are more likely to sort into neighborhoods closer to jobs, even though I use alternative measurement for the value of time. (See Appendix section E.3 for detail)

[34] I construct the analysis at census tract level. For each census tract, I compute the count of business establishments located within 1 mile of the census tract of interest. Meanwhile, I compute the total population within 1 mile of the census tract of interest. I then compute the per-capita count of business establishments by dividing the total count by population.

[35] I conjecture that the lack of response may be because the raw count of grocery stores is a mismeasurement of the true amenity level of grocery services, as small neighborhood stores may be replaced by large chain stores in the event of an amenity upgrade.

Furthermore, I use O*NET characteristics selected by a lasso procedure and see how workers sort into neighborhoods closer to jobs by these occupational characteristics. (See Appendix section D.4 and D.5 for description of the lasso procedure) Interestingly, I find that workers in occupation with higher scores of "time-pressure" and "duration of work week" are more likely to sort into neighborhood closer to jobs. (See Appendix section E.4 for detail)

Finally, I estimate the model under different calibrations of ϕ. Moreover, I estimate the model using work location in the 2010 Zip-Code Business Pattern to make sure the results are not driven by different job datasets. I also re-defined high-skilled occupation using alternative definitions, and estimate the model again using the new definitions. The results remains quite robust. (See Appendix section E.5 for detail)

7 Determinants of gentrification

The estimated model shows that the changing value of time affects the locational preference of workers in different occupations, both directly through incentivizing workers to reduce commute time by locating closer to work and indirectly through changing the local skill mix in the affected neighborhoods. In this section, I conduct a decomposition exercise to evaluate how much gentrification is explained directly by the migration of workers with a rising value of time, and how much gentrification is explained by the migration of workers who move to the central cities due to the endogenously improved amenities.[36]

7.1 Direct effect of changing commute cost

First, I evaluate the direct effect of the changes in the value of time on the re-sorting of residential locations within cities. In this exercise, I allow only workers' value of time to change from 1990 to 2010, holding neighborhood amenities, rent, and all other components in the location demand equation constant at their 1990 level. I use the location demand equation to generate the location choice in 2010 that would have been made if only the value of time had changed. After calculating these location choice probabilities by each occupation, I multiply them by MSA-level population in each occupation in 1990 to generate the predicted population for each occupation in each neighborhood. The following equation is the predicted location demand:

$$\log\left(\widehat{s_{jmk,2010}}\right) = \delta_{jmk} + \underbrace{\tilde{\delta}'_{mk,2010}}_{\text{Normalizing constant}} + \underbrace{\log\left(\sum_{n' \in J_m} \pi_{n'mk} \exp\left(-\mu v_{mk,2010} \cdot c_{jn'm}\right)\right)}_{\text{Direct effect of value of time}} \quad (11)$$

$$-\beta r_{jm,1990} + \gamma_z \log\left(\frac{N^H_{jm,1990}}{N^L_{jm,1990}}\right) + \theta_{k,1990} X_{jm} + \xi_{jmk,1990}$$

[36] Since the model predictions are in terms of occupation-specific location demand, the way the model replicates neighborhood changes is by predicting population changes for different occupations.

Since I observe the level of $s_{jmk,1990}$, I compute the predicted location demand equation by bringing the value of time to the 2010 level. To ensure that the predicted location choice probabilities add up to one for each occupation and each MSA, I adjust them accordingly with normalizing constant $\tilde{\delta}'_{mk,2010}$.[37] All other components of the initial location demand are held fixed. The components I hold fixed are the occupation-specific taste for each neighborhood revealed by the initial choice in 1990 that cannot be accounted for by value of time and commute time.[38] I assume that these idiosyncratic components do not change over time, and I save the preference components recovered in the data for 1990 and apply them to the demand equation for 2010.

After computing the predicted location choices $\widehat{s_{jmk,2010}}$, I compute the change in local skill ratios using the predicted location choices and compare them to those observed in the data. The comparison between the change in local skill ratios in the counterfactual neighborhoods and the change observed in the data help assess how much of the neighborhood change is driven by the direct effect of the rising cost of commute time.

I compute overall skill ratios for central city neighborhoods for each MSA in the US. To highlight the within-MSA changes in skill ratios in central cities relative to suburbs, I make sure that the population in each occupation and each MSA is held fixed at the 1990 level. As a result, all of the predicted skill ratio change in central cities is the result of the re-sorting of residents within an MSA, instead of cross-MSA variation in skill ratio change. I then compute the change in log skill ratio for each MSA observed in the data, which is plotted against the predicted change in skill ratio in Figure 6(a). In the plot, the change in skill ratio is highly correlated with those predicted by the model. R-squared is 0.3055 , which is very large given that the model is only predicted by change in value of time and commute time. It shows that the rising value of time and the resulting rise in cost of commute should be an important determinant of gentrification, because variation in the predicted changes in skill ratios are generated solely by the differential value of time growth across occupations (determined outside of the model) and the commute time that workers in each MSA can save by moving close to work (computed from the 1995 travel matrix and 1994 initial job locations).

However, the magnitude of the prediction of change in the central city skill ratio is substantial smaller than the actual change, which means that the predicted change in skill ratio grossly underestimates the actual change. The key reason for the underestimation is that the model mutes the channel of endogenous amenity changes. This implies that a rising value of time is indeed an important determinant of gentrification (cities predicted to gentrify do tend to gentrify), but the direct effect of rising value of time is too small to account for the full magnitude of neighborhood change.

---

[37] This approach is described in the following equation,

$$\widehat{s_{jmk,2010}} = \frac{\exp\Big(\log\big(s_{jmk,1990}\big)-\log\big(\sum_{n'}\pi_{n'mk}\exp\big(-\mu v_{mk,1990}\cdot c_{jn'm}\big)\big)+\log\big(\sum_{n'}\pi_{n'mk}\exp\big(-\mu v_{mk,2010}\cdot c_{jn'm}\big)\big)\Big)}{\sum_{j'}\exp\Big(\log\big(s_{j'mk,1990}\big)-\log\big(\sum_{n'}\pi_{n'mk}\exp\big(-\mu v_{mk,1990}\cdot c_{j'n'm}\big)\big)+\log\big(\sum_{n'}\pi_{n'mk}\exp\big(-\mu v_{mk,2010}\cdot c_{j'n'm}\big)\big)\Big)}.$$

[38] In particular, $\xi_{jmk,1990}$ is held fixed at the level computed in 1990. For workers in certain high-paying occupations, due to unobserved reasons unrelated to commuting, rent or endogenous amenities, $\xi_{jmk,1990}$ may be relatively large in the suburbs. And $\xi_{jmk,1990}$ for low-paying occupations may be relatively large in the central cities due to proximity to public transportation, social services, etc. I let the data determine the initial size of $\xi_{jmk,1990}$, and hold it fixed over time.

7.2 Indirect effect of endogenous amenity change

In the second part of this section, I evaluate the indirect effect of the changing value of time on urban change through endogenous amenity changes at neighborhood level. In other words, how much can gentrification be explained by the migration of workers who move as a result of amenities changes brought about by the initial movers. For that purpose, I allow local skill ratio to vary endogenously and change by the amount predicted by the shock to the value of time, and then compute a new set of predicted neighborhood residential mixes.[39]

With this new set of predicted location demand changes, the change in local skill mix is further affected by the differential taste for endogenous neighborhood amenity changes by high-skilled workers and low-skilled workers. Therefore, the neighborhoods in which there is an influx of high-skilled workers induced by rising value of time would become more attractive to other high-skilled workers. Since high-skilled workers like amenities more, adding the effect of endogenous amenity changes would raise the predicted change in local skill ratio even higher in neighborhoods in which changes in skill ratios are already positively affected by the shock to the value of time. The following equation is the predicted location demand equation:

$$
\begin{aligned}
\log\left(\widehat{s_{jmk,2010}}\right) \quad = \quad & \delta_{jmk} + \underbrace{\tilde{\delta}'_{mk,2010}}_{\text{Normalizing constant}} + \underbrace{\log\left(\sum_{n' \in J_m} \pi_{n'mk} \exp\left(-\mu v_{mk,2010} \cdot c_{jn'm}\right)\right)}_{\text{Direct effect of value of time}} \qquad (12) \\
& -\beta r_{jm,1990} + \underbrace{\gamma_z \log\widehat{\left(\frac{N^H_{jm,2010}}{N^L_{jm,2010}}\right)}}_{\text{endogenous amenity change}} + \theta_{k,1990} X_{jm} + \xi_{jmk,1990}.
\end{aligned}
$$

From equation (12), one can see that γ_z is different for high-skilled workers and low-skilled workers. When $\log\left(\frac{N^H_{jm,2010}}{N^L_{jm,2010}}\right)$ rises due to shocks to other workers' location demand, high-skilled workers increase their demand for these neighborhoods more than low-skilled workers do, raising the skill ratio in these neighborhoods even higher. By plugging in the predicted change in log skill ratio, I allow the predicted workers' location choices to take into account the endogenous change in skill ratio.

In Figure 6(b), I show the plot between actual change and predicted change in log skill ratio in central city relative to suburbs. The R-squared is now 0.35, which is higher than the R-squared in the previous exercise, 0.3055, in which the channel of endogenous amenity change is muted. Importantly, the magnitude of the predicted change catches up substantially to the observed change. The predicted change in skill ratio still underestimates actual change, but it greatly improves with

[39] I regress the actual change in log local skill ratio on the change in log population of high and low-skilled workers from the first exercise (allowing only the value of time to change). I then use the predicted value from the regression for the predicted local skill ratio in 2010. If a neighborhood's observed skill ratio has risen, but the changing value of time does not predict any change, then the $\Delta\log\widehat{\left(\frac{N^H_{jmt}}{N^L_{jmt}}\right)}$ would be zero. $\Delta\log\widehat{\left(\frac{N^H_{jmt}}{N^L_{jmt}}\right)}$ only picks up variation in changes predicted by the shock to the value of time.

the added channel of amenity change. This shows a large share of the gentrifiers who migrated to the central cities are attracted by the endogenously improved amenities.

Finally, in Figure 6(c), I plot the actual change in log skill ratio and the change in log skill ratio predicted by the full model using the observed change in skill ratio as input into the model; I also adjust rent changes. The graph shows that predicted changes in log skill ratio have extremely high correlation with actual changes in log skill ratio (R-squared is now 0.7454). The magnitude of the predicted changes also matches the observed magnitude quite well. This shows that some of the amenity changes that are also driving the neighborhood changes are not captured by the change in the cost of commute (value of time). This finding motivates further research on other factors that influence the changes in amenities and skill ratio. The changing taste for density and quality of existing consumption and cultural amenities documented by Couture and Handbury (2017), for example, would be an important determinant for some of the residual change in local skill ratio.

7.3 Gentrification mechanisms

From these exercises, I find that the change in housing demand induced by the shock to the value of time produces the exogenous momentum for gentrification. The exogenous shock to the demand for central city housing changes the local skill mix, which improves local amenities in the affected neighborhoods. Since high-skilled workers are shown to exhibit stronger preference for neighborhood amenities than low-skilled workers, improved amenities in central cities further attract high-skilled workers, leading to a "loop" effect for further gentrification.

Local rent also increases as a result of the increase in high-skilled workers. The increase in rent reduces the desirability of central cities, but at the same time the increase in amenity levels is highly valued by high-skilled workers, which compensates for the rent hikes. While low-skilled workers also value neighborhood amenities, they do not place as much value on amenities as high-skilled workers do. Therefore, in response to the rent hike in the central cities, low-skilled workers out-migrate in search of cheaper locations.

8 Welfare implications

Both researchers and the public are concerned that rising rents and out-migration of the low-skilled from the central cities may exacerbate the widening welfare inequality between the high- and low-skilled residents within cities. The challenge of evaluating whether gentrification widens or narrows welfare inequality is that the welfare impact of spatial sorting is theoretically ambiguous (Moretti (2011), Diamond (2016)).

If high-skilled workers move to the central cities entirely due to their rising cost of commute, the rising rent driven by their move may lower their welfare in equilibrium. In contrast, the out-migration of low-skilled workers means that the rising rent in the central cities may not impact them as much as it impacts high-skilled workers. In this case, the spatial sorting could potentially reduce welfare gap. Alternatively, if much of the inflow of high-skilled workers is also driven by endogenous amenity improvement in the central cities, high-skilled may benefit from the improving amenity in the central cities more than rising rent harms them. In this case, welfare gap could widen.

Fortunately, the estimated spatial equilibrium model provides a quantitative welfare framework to study welfare impact of the neighborhood sorting. In this section, I quantitatively unpack how changes in the value of time, amenities, and rent differentially affect the welfare of high- and low-skilled workers, and I determine whether the welfare gap between them increases or decreases due to these changes.

First, I compute workers' expected utility in 1990 using the estimated model. Then, I change the value of the observables (value of time, amenities, rent) to the 2010 level to identify how much each of these changes impacts workers' expected utility. The utility of a worker in occupation k who lives in neighborhood j and works in neighborhood n receives utility:

$$V_{i,jnmt} = \delta_{mkt} - \mu v_{mkt} c_{jnm} - \beta r_{jmt} + \gamma_k a_{jmt} + \gamma_k \zeta_{jmt} + \varepsilon_{i,jmt}.$$

δ_{mkt} is the occupation/MSA fixed-effects, which represents the utility component invariant to neighborhood choice. In the model, it is the sum of the utility derived from gross log earnings unadjusted for the cost of commute time and the utility derived from leisure. For this exercise, I abstract from the welfare effect of changing leisure inequality, because doing so requires the marginal utility of leisure, which is difficult to estimate from a location choice model.[40] Since μ is the marginal utility of log income. I approximate δ_{mkt} using the product of the marginal utility μ and the nationwide level of log earnings in occupation k. The observed and unobserved amenities will be replaced with the terms used in the estimation and evaluation, $\gamma_z \log\left(\frac{N^H_{jmt}}{N^L_{jmt}}\right) + \xi_{jmkt}$. To do the exercise, I use utility function:

$$V_{i,jnmt} = \mu y_{kt} - \mu v_{mkt} c_{jnm} - \beta r_{jmt} + \gamma_z \log\left(\frac{N^H_{jmt}}{N^L_{jmt}}\right) + \xi_{jmkt} + \varepsilon_{i,jmt}.$$

The expected utility for workers in occupation k working in neighborhood n is $\mathrm{E}\left(\max_j V_{i,jnmt}\right)$. The expected utility can be understood as the expected utility of the top choice neighborhoods of each worker within occupation k. Since the idiosyncratic component is distributed as Type I

[40] If leisure inequality is considered in the model, it could be a force that narrows the welfare gap between high- and low-skilled workers due to the divergence of leisure times between them, which is documented by Aguiar and Hurst (2007). But to make a precise statement on this matter, one must estimate how high-skilled and low-skilled workers value leisure differentially. If high-skilled workers do not value leisure as much as low-skilled workers, then the welfare implication due to leisure inequality may be small. Conversely, if high-skilled workers value leisure more than low-skilled workers do, welfare inequality may increase more if we take leisure into account.

Extreme Value, the expected utility becomes:

$$\mathrm{E}\left(U_{nmkt}\right)=\ln\left(\sum_{j\in J_m}\exp\left(V_{jnmkt}\right)\right)$$

$$\text{where the mean utility } V_{jnmkt}=\mu y_{kt}-\mu v_{mkt}c_{jnm}-\beta r_{jmt}+\gamma_z\log\left(\frac{N_{jmt}^H}{N_{jmt}^L}\right)+\xi_{jmkt}$$

Given the neighborhood n's share of jobs in occupation k in MSA m is π_{nmk}, the overall expected utility in an MSA is given by a weighted average of the expected utilities conditional on work locations:

$$\mathrm{E}\left(U_{mkt}\right)=\textstyle\sum_{n\in J_m}\pi_{nmk}\ln\left(\sum_{j\in J_m}\exp\left(V_{jnmkt}\right)\right).$$

To start the exercise, I compute the expected utility in the initial year 1990 by using the observables from the 1990 data.[41] The expected utility is given by:

$$\mathrm{E}\left(U_{mk,1990}\right)=\textstyle\sum_{n\in J_m}\pi_{nmk}\ln\left(\sum_{j\in J_m}\exp\left(V_{jnmk,1990}\right)\right)$$
$$\text{where } V_{jnmk,1990}=\mu y_{k1990}-\mu v_{mk1990}c_{jnm}-\beta r_{jm1990}+\gamma_z\log\left(\tfrac{N_{jm1990}^H}{N_{jm1990}^L}\right)+\xi_{jmk1990}.$$

Then, I proceed to compute hypothetical utility level in 2010 by adjusting earnings, value of time, rent, and amenities to 2010 levels. I normalize the expected utility measure by dividing it by the estimate of μ, which is the marginal utility of each unit of log earnings. The change in normalized expected utility is thus measured in the unit of log earnings. In other words, each of the change in welfare levels is measured as the equivalent change in earnings.[42]

8.1 Earnings

First, I compute the expected utility if workers' log earnings are changed to 2010 levels:

$$\mathrm{E}\left(\hat{U}_{mk,2010}^y\right)=\textstyle\sum_{n\in J_m}\pi_{n(m)k}\ln\left(\sum_{j\in J_m}\exp\left(V_{jnmk,2010}\right)\right)$$
$$\text{where } V_{jnmk,2010}=\underbrace{\mu y_{k2010}}_{\text{earnings change}}-\mu v_{mk1990}c_{jnm}-\beta r_{jm2010}+\gamma_z\log\left(\tfrac{N_{jm1990}^H}{N_{jm1990}^L}\right)+\xi_{jmk1990}.$$

Since log earnings are computed at occupation level, the change in expected utility for workers in each occupation is simply identical to the change in log earnings.[43] Table 4 shows that the welfare gap due to a change in earnings increases from 0.7368 to 0.8, which is 0.0632 log point.[44] I use the

[41] $\xi_{jmk1990}$ can be computed from the location demand equation. It is the residual term in year 1990.

[42] For example, if the normalized expected utility increases by 0.1 log point due to changing amenities, then the welfare impact of the changing amenities is equivalent to an increase of 0.1 log point in earnings.

[43] The willingness to pay to live in a world in which one gets paid by x unit more in log earnings would be exactly x unit of log earnings.

[44] The magnitude of the increase in the earnings gap between high-skilled workers and low-skilled workers depends on the definition of what constitutes high-skilled.

increase in welfare due to change in earnings as the benchmark for the welfare analysis. I evaluate whether the impact of value of time, rent, and amenities offsets or further increases the welfare gap caused by a change in earnings.

8.2 Value of time

To evaluate the impact of changing value of time on welfare, I let value of time adjust to 2010 levels. The expected utility is now the following:

$$\mathrm{E}\left(\hat{U}^{yv}_{mk,2010}\right) = \sum_{n\in J_m} \pi_{nmk} \ln\left(\sum_{j\in J_m} \exp\left(\hat{V}_{jnmk,2010}\right)\right)$$
$$\text{where } \hat{V}_{jnmk,2010} = \underbrace{\mu y_{k2010}}_{\text{earnings change}} - \underbrace{\mu v_{mk2010} c_{jnm}}_{\text{value of time change}} - \beta r_{jm1990} + \gamma_z \log\left(\frac{N^H_{jm1990}}{N^L_{jm1990}}\right) + \xi_{jmk1990}.$$

This is expected to reduce welfare for the workers whose value of time rises over time, especially in cities where their job locations are highly concentrated. Table 4 shows that the welfare gap due to earnings change and value of time change increases by 0.0564 log points, which is 10.76% smaller than the increase in the welfare gap due to earnings change alone. This is an intuitive result, given that high-skilled workers experience a larger increase in the value of time.

8.3 Rents

Next, I let rent also adjust to 2010 levels. The expected utility is now the following:

$$\mathrm{E}\left(\hat{U}^{yvr}_{mk,2010}\right) = \sum_{n\in J_m} \pi_{nmk} \ln\left(\sum_{j\in J_m} \exp\left(\hat{V}_{jnmk,2010}\right)\right)$$
$$\text{where } \hat{V}_{jnmk,2010} = \underbrace{\mu y_{k2010}}_{\text{earnings change}} - \underbrace{\mu v_{mk2010} c_{jnm}}_{\text{value of time change}} - \underbrace{\beta r_{jm2010}}_{\text{rent change}} + \gamma_z \log\left(\frac{N^H_{jm1990}}{N^L_{jm1990}}\right) + \xi_{jmk1990}.$$

Table 4 shows that the welfare gap increases by 0.0612 log points, which is 8.51% larger than the increase in the welfare gap if rent is not considered and 3.09% smaller than the increase in earnings gap alone. This means that the overall impact of rent hike hurts the low-skilled more. It may be surprising that a rent hike in the central city does not reduce the welfare gap, since high-skilled workers have increasing demand for central city neighborhoods. Diamond (2016) shows that the change in rents narrows the welfare gap. However, when one closely examines cross-MSA analysis, such as in Diamond (2016), the cities that experienced further improvement in amenities were already high-amenity cities in the initial year of analysis. So the rise in rents in high-amenity cities does not have as much of a negative impact on non-college-educated workers who disproportionately reside in low-amenity cities throughout the analyzed period.

On the other hand, in the context of within-city gentrification, low-skilled workers started out disproportionately residing in central cities. By revealed preference, the model infers that low-skilled workers have large unobserved preference for central locations. This could be due to cultural and historical reasons, or because central cities provide better access to social services and public transportation, etc. With large unobserved preference for central cities, low-skilled workers' welfare is negatively impacted by the rising rents in the central cities.[45]

8.4 Amenities

Finally, I let the skill ratio also adjust to 2010 levels. As Diamond (2016) points out, the change in skill ratio must be treated with caution when evaluating welfare, because variation in skill ratio only measures amenity levels *relative* to each other. In other words, a neighborhood with a relatively higher skill ratio compared to other neighborhoods has better amenities than other neighborhoods. The preference parameter for amenities γ_z is also identified off the relative change in skill ratios over time. However, in the data, the skill ratio increases across the nation, but the national increase in skill ratio does not necessarily mean that amenities have improved nationwide. Therefore, to evaluate the impact of amenity change due to change in skill ratio, I hold the MSA-level skill ratio fixed at the 1990 level, and only allow the neighborhood level skill ratio to change due to geographic re-sorting of high-skilled and low-skilled populations.

$$\mathrm{E}\left(\hat{U}^{yvra}_{mk,2010}\right) = \sum\nolimits_{n \in J_m} \pi_{nmk} \ln\left(\sum\nolimits_{j \in J_m} \exp\left(\hat{V}_{jnmk,2010}\right)\right)$$

$$\text{where } \hat{V}_{jnmk,2010} = \underbrace{\mu y_{k2010}}_{\text{earnings change}} - \underbrace{\mu v_{mk2010} c_{jnm}}_{\text{value of time change}} - \underbrace{\beta r_{jm2010}}_{\text{rent change}} + \underbrace{\gamma_z \log\left(\frac{\hat{N}^H_{jm2010}}{\hat{N}^L_{jm2010}}\right)}_{\text{amenity change}} + \xi_{jmk1990}$$

$$\text{and } \hat{N}^H_{jm2010} = \frac{N^H_{jm2010}}{N^H_{m2010}} N^H_{m1990},\ \hat{N}^L_{jm2010} = \frac{N^L_{jm2010}}{N^L_{m2010}} N^L_{m1990}$$

$\hat{N}^H_{jm2010}$ and $\hat{N}^L_{jm2010}$ are the number of high-skilled and low-skilled workers living in neighborhood j rescaled by the MSA-level population of high-skilled and low-skilled workers.

Table 4 shows that after taking into account the amenity change, the welfare gap increases by 0.0841 log points, which is 0.0209 log points larger than the welfare gap caused by changes in earnings alone. The overall change in the welfare gap after considering changes in commute cost, rents, and amenities is 33.07% larger than the earnings gap alone. One can easily see that most of the increase in the welfare gap is driven by the change in neighborhood amenities.

9 Conclusion

Central city neighborhoods experienced a dramatic reversal of fortune in the past few decades. High-skilled workers increased their demand for housing in central city neighborhoods, which raised rents and amenity levels in these neighborhoods. I show that the rise in the value of time among high-skilled workers leads them to increasingly prefer living in central city neighborhoods to avoid long and costly commute time. These changing location preferences contribute to the rising demand for housing in central city locations. In addition, the effect of the rising value of time on housing

[45] Since the expected utility can be understood as the expected utility derived from the top choice neighborhood. If low-skilled workers have high idiosyncratic preference for central cities, their top choices are more likely to be central city neighborhoods, which makes them vulnerable to rent hikes in central cities.

demand is magnified by endogenous improvements in amenities, which leads to further sorting of high-skilled workers, who tend to have stronger preferences for amenities than low-skilled workers do, into the central cities.

I estimate value of time for workers in each occupation and show that workers in occupations with rising value of time increasingly live in central city neighborhoods. I then estimate a spatial equilibrium model of residential choice to quantity the relative importance of the direct effect of rising value of time on gentrification and the indirect effect of endogenous amenity change on gentrification. I find that the surge in demand driven by rising value of time provides the initial momentum for gentrification. The initial in-flow of high-skilled workers into central cities leads to endogenous change in amenity levels in the central cities, and the endogenous change in amenities is an essential mechanism that magnifies the effect of initial shock on locational sorting.

Finally, I use the estimated model to quantitatively evaluate the impact of gentrification on the welfare gap between high-skilled and low-skilled workers. I find that the combined effect of value of time, rents, and amenities leads to an increase in the welfare gap equivalent to 8.41 percentage point in the earnings gap, which is 33% larger than the increase in the actual earnings gap.

While this paper shows that the rise in value of time contributes to gentrification, the mechanisms that cause the value of time to change remain unclear and open to future research. In addition, this paper shows that high-skilled workers have strong preference for local amenities, but such preference is estimated as a reduced-form elasticity to local skill ratio. Future research could focus on unpacking the mechanisms between the demand for and supply of local amenities. Moreover, firm locations are taken as given in this paper, and future research could examine how firms' location decisions may respond to workers' geographic re-sorting.

References

[1] Aguiar, Mark and Erik Hurst, 2005."Consumption vs Expenditure." *Journal of Political Economy* 113(5): 919-948.

[2] Aguiar, Mark and Erik Hurst, 2007 "Measuring Leisure: The Allocation of Time over Five Decades." *Quarterly Journal of Economics* 122(3): 969-1006.

[3] Ahlfeldt, Gabriel, Stephen Redding, Daniel Sturm, and Nikolaus Wolf, 2015. "The Economics of Density: Evidence From the Berlin Wall," *Econometrica* 83: 2127-2189.

[4] Albouy, David and Bert Lue, 2015. "Driving to Opportunity: Local Rents, Wages, Commuting, and Sub-Metropolitan Quality-of-Life Measures." *Journal of Urban Economics*, 89: 74–92

[5] Altonji, Joseph, Todd Elder, and Christopher Taber, 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy* 113(1): 151-184.

[6] Autor, David, Frank Levy, and Richard Murnane, 2003 "The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, 118(4): 1279-1333.

[7] Autor, David, Christopher Palmer, and Parag Pathak, 2014. "Housing Market Spillovers: Evidence from the End of Rent Control in Cambridge, Massachusetts," *Journal of Political Economy*, 122(3): 661-717

[8] Autor, David, Christopher Palmer, and Parag Pathak, 2017. "Gentrification and the Amenity Value of Crime Reductions: Evidence from Rent Deregulation." Working Paper.

[9] Baum-Snow, Nathaniel, 2007. "Did Highways Cause Suburbanization?" *Quarterly Journal of Economics*, 122(2): 775-805.

[10] Baum-Snow, Nathaniel and Daniel Hartley, 2017. "Accounting for Central Neighborhood Change, 1980-2010," Working paper.

[11] Bayer, Patrick, Fernando Ferreira, Robert McMillan, 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *Journal of Political Economy*, 115(4):588-638

[12] Becker, Gary. 1965. "A Theory of the Allocation of Time," *Economic Journal*, 75(299): 493-517

[13] Benhabib, Jess, Richard Rogerson, and Randall Wright, 1991. "Homework in Macroeconomics: Household Production and Aggregate Fluctuations," *Journal of Political Economy*, 99(6): 1166-1187.

[14] Berry, Steven, 1994. "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25(2), 242-262.

[15] Boustan, Leah, 2010. "Was Postwar Suburbanization "White Flight"? Evidence from the Black Migration," *Quarterly Journal of Economics*, 125(1): 417—43.

[16] Brueckner, Jan,1987. "The Structure of Urban Equilibria: A Unified Treatment of the Muth-Mills model," *Handbook of Regional and Urban Economics*, 2(20), 821-845.

[17] Brueckner, Jan and Stuart Rosenthal, 2009. "Gentrification and Neighborhood Housing Cycles: Will America's Future Downtowns Be Rich?" *Review of Economics and Statistics*, 91(4), 725-743.

[18] Brueckner, Jan, Jacques-Francois Thisse, and Yves Zenou, 1999. "Why is Central Paris Rich and Downtown Detroit Poor? An Amenity-based theory." *European Economic Review*, 43: 91-107.

[19] Brummet, Quentin and Davin Reed, 2018. "Gentrification and the Well-Being of Original Neighborhood Residents: Evidence from Longitudinal Census Microdata" Working Paper.

[20] Card, David, Alexandre Mas, and Jesse Rothstein, 2008. "Tipping and the Dynamics of Segregation," *Quarterly Journal of Economics* 123(1): 177-218.

[21] Conley, Timothy, 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics*, 92(1): 1-45.

[22] Couture, Victor, 2016. "Valuing the Consumption Benefits of Urban Density" Working Paper.

[23] Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst, 2018. "Income Growth and the Distributional Effects of Urban Spatial Sorting," Working Paper.

[24] Couture, Victor and Jessie Handbury, 2017. "Urban Revival in America, 2000 and 2010." Working Paper.

[25] Davis, Donald, Jonathan Dingel, Joan Monras, and Eduardo Morales, 2017 "How Segregated is Urban Consumption?" Working Paper.

[26] Diamond, Rebecca, 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000." *American Economic Review*, 106(3): 479-524.

[27] Edlund, Lena, Cecilia Machado, and Michaela Sviatchi, 2015 "Bright Minds, Big Rent: Gentrification and the Rising Returns to Skill." NBER Working Paper w21729.

[28] Ellen, Ingrid Gould, Keren Mertens Horn, and Davin Reed, 2016. "Has Falling Crime Invited Gentrification?" US Census Bureau Center for Economic Studies Paper No. CES-WP-17-27

[29] Ellen, Ingrid and Katherine O'Regan, 2011. "How Low Income Neighborhoods Change: Entry, Exit, and Enhancement," *Regional Science and Urban Economics*, 41(2): 89-97.

[30] Ellison, Glenn, Edward Glaeser, 1997. "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach." *Journal of Political Economy*, 105(5): 889-927.

[31] Ellison, Glenn, Edward Glaeser, and William Kerr, 2010. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns." *American Economic Review*, 100(3): 1195-1213.

[32] Glaeser, Edward, Jed Kolko, and Albert Saiz, 2001. "Consumer City," *Journal of Economic Geography* 1(1): 27-50.

[33] Goldin, Claudia, 2014 "A Grand Gender Convergence: Its Last Chapter." *American Economic Review*, 104(1): 1091-1119.

[34] Guerrieri, Veronica, Daniel Hartley, and Erik Hurst, 2013. "Endogenous Gentrification and Housing Price Dynamics." *Journal of Public Economics*,100(C): 45-60.

[35] Gyourko, J., A. Saiz, and A. Summers, 2008. "A New Measure of the Local Regulatory Environment for Housing Markets: The wharton residential land use regulatory index," *Urban Studies* 45(3), 693—729.

[36] Kneebone, Elizabeth and Natalie Holmes, 2015. "The Growing Distance Between People and Jobs in Metropolitan America," Brookings Metropolitan Policy Program.

[37] Handbury, Jessie, 2013. "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities." Working Paper.

[38] Holian, Matthew and Matthew Kahn, 2015. "Household Carbon Emissions from Driving and Center City Quality of Life." *Ecological Economics* 116: 362-368.

[39] Kuhn, Peter, and Fernando Lozano, 2008. "The Expanding Workweek? Understanding Trends in Long Work Hours among U.S. Men, 1979-2006," *Journal of Labor Economics*, 26(2): 311-343.

[40] Lazear, Edward and Sherwin Rosen, 1981 "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy*, 89(5):841-864.

[41] Lee, Sanghoon and Jeffrey Lin, 2017. "Natural Amenities, Neighborhood Dynamics, and Persistence in the Spatial Distribution of Income," *Review of Economic Studies,* forthcoming.

[42] McFadden, Daniel, 1973. "Conditional Logit Analyisis of Qualitative Choice Behavior," *Frontiers in Econometrics*, 105-142.

[43] Manson, Steven, Jonathan Schroeder, David Van Riper, and Steven Ruggles. IPUMS National Historical Geographic Information System: Version 12.0 [Database]. Minneapolis: University of Minnesota. 2017.

[44] Monte, Ferdinando, Stephen Redding, Esteban Rossi-Hansberg, 2015. "Commuting, Migration and Local Employment Elasticities," NBER Working Paper.

[45] Moretti, Enrico, 2011. "Real wage inequality," *University of California Berkeley Mimeo.*

[46] Nevo, Aviv and Arlene Wong, 2017. "The Elasticity of Substitution Between Time and Market Goods: Evidence from the Great Recession," NBER Working Paper.

[47] Saiz, Albert, 2010. "The Geographic Determinants of Housing Supply," *Quarterly Journal of Economics*, 125(3):1253-1296.

[48] Roback, Jennifer, 1982. Wages, Rents, and the Quality of Life. *Journal of Political Economy*, 90(6): 1257-1278.

[49] Rosen, Sherwin, 1979. "Wages-based Indexes of Urban Quality of Life," *Current Issues in Urban Economics.*

[50] Rosenthal, Stuart and William Strange, 2004. "Evidence on the Nature and Sources of Agglomeration Economies." *Handbook of Regional and Urban Economics*, 4(49): 2119-2171

[51] Rossi-Hansberg, Esteban, 2004. "Optimal Urban Land Use and Zoning," *Review of Economic Dynamics*, 7(1): 69-106.

[52] Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Raymond Owens, 2010. "Housing Externalities," *Journal of Political Economy*, 118(3): 485-535.

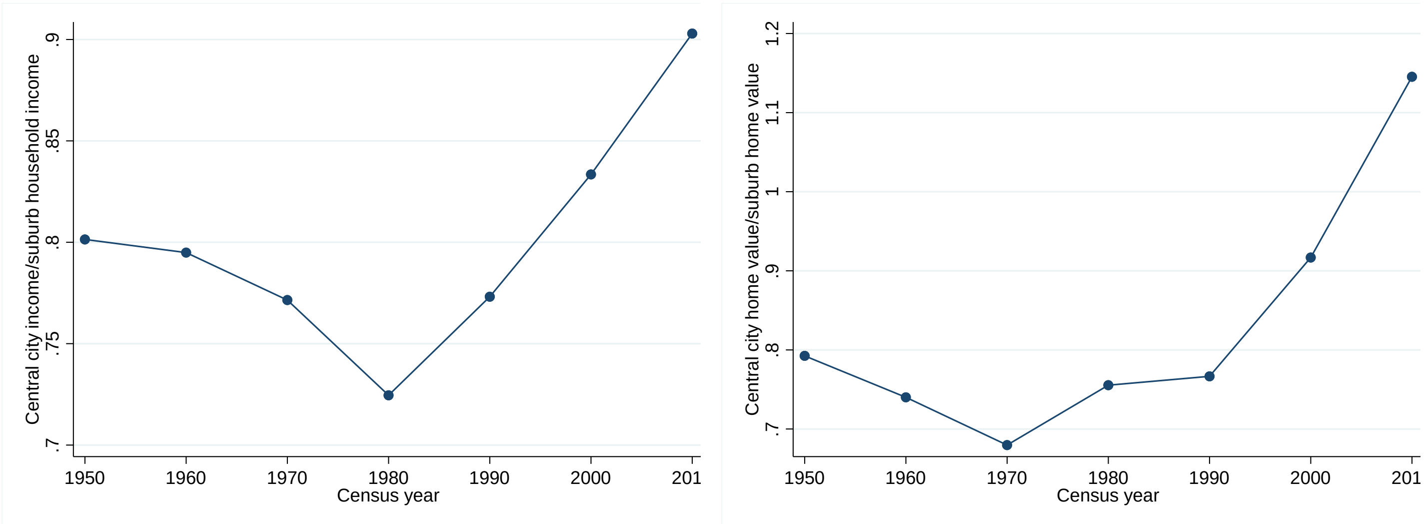
[53] Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek, 2015. Integrated Public Use Microdata Series: Version 7.0 [dataset]. Minneapolis: University of Minnesota, 2017.

[54] U.S. Department of Transportation, Federal Highway Administration, 2009 National Household Travel Survey. URL: http://nhts.ornl.gov.

[55] Vigdor, Jacob, 2002. "Does Gentrification Harm the Poor?" *Brookings- Wharton Papers on Urban Affairs*, 133-182.

Figure 1: Income and home value ratio between central city and suburban neighborhoods

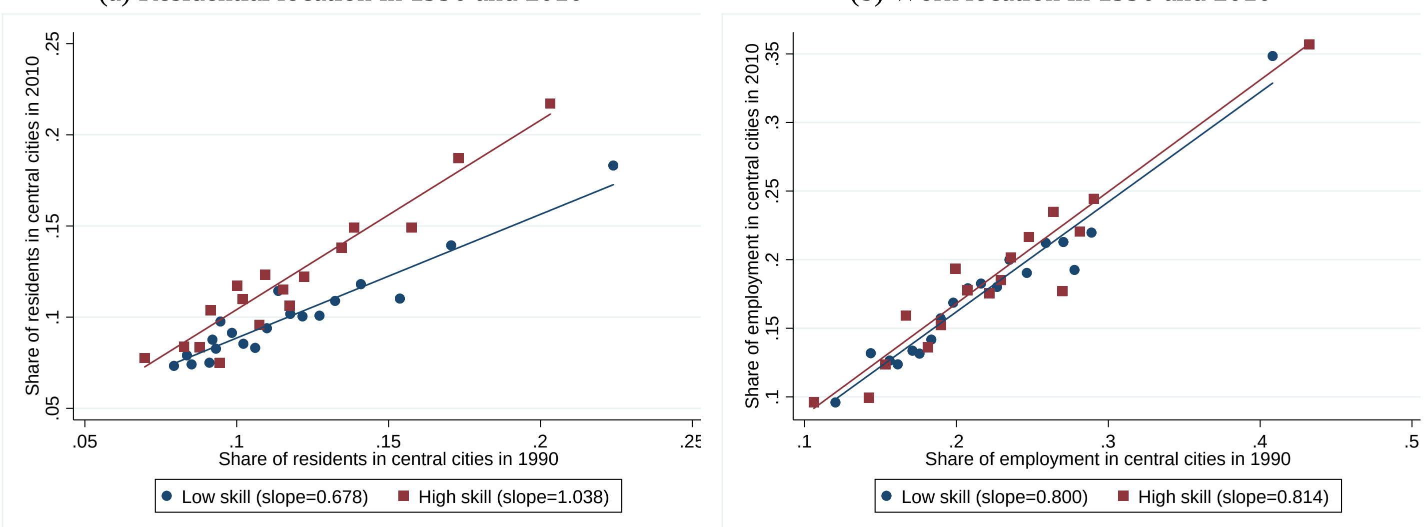
(a) Income ratio (b) Home value ratio



Notes: Central cities in this graph are census tracts that are located within 5 miles of the downtown pin on Google in the respective MSAs. The values plotted are the mean income and home value of the census tracts located in the central cities and the mean income and home value of non-central city census tracts in the top 25 MSAs (defined by population ranking in 1990). The source of the data is Census and ACS provided by NHGIS.

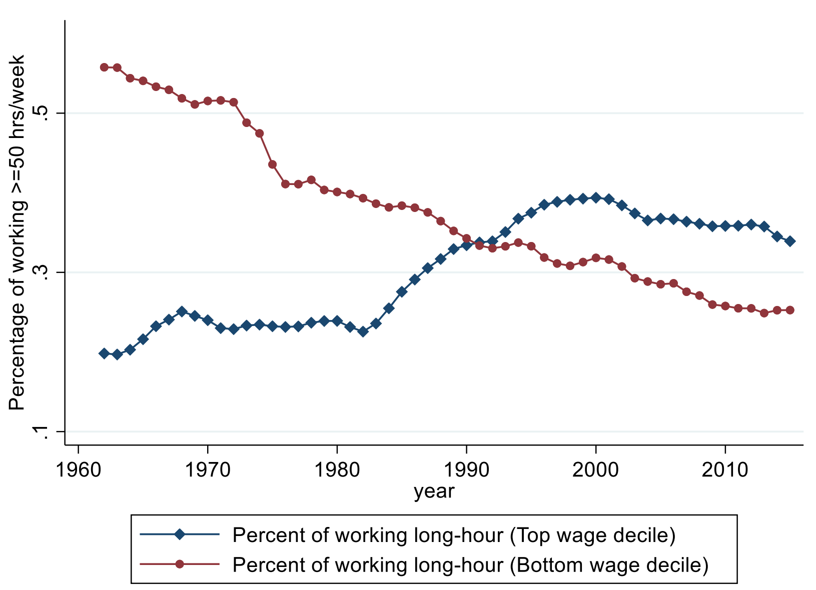
Figure 2: Residential and work location sorting by skill

(a) Residential location in 1990 and 2010 (b) Work location in 1990 and 2010



Notes: The Figure (a) is binscatter plot between each occupation's share of residents living in central city neighborhoods in 1990 and in 2010. Figure (b) is a binscatter plot between each occupation's share of job counts located in central city locations in 1994 and in 2010. Residential location data come from both IPUMS and NHGIS Census data. Details are described in the data section. Square dots represent binscatter plot of data in high-skilled occupations, and circle dots represent binscatter plot of data in low-skilled occupations. High-skilled occupations are defined as occupations in which more than 40% of the workers have college degrees in 1990. The employment data come from ZCBP at zip code level. Central cities are defined as census tracts and zip codes with centroids within 5-mile radius of the downtown pin. I use the sample from the largest 25 MSA to produce these graphs.

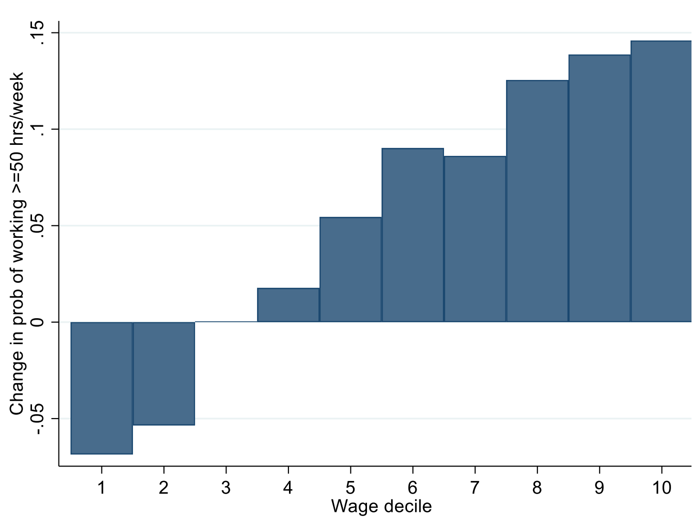
Figure 3: The evolution of long-hour working



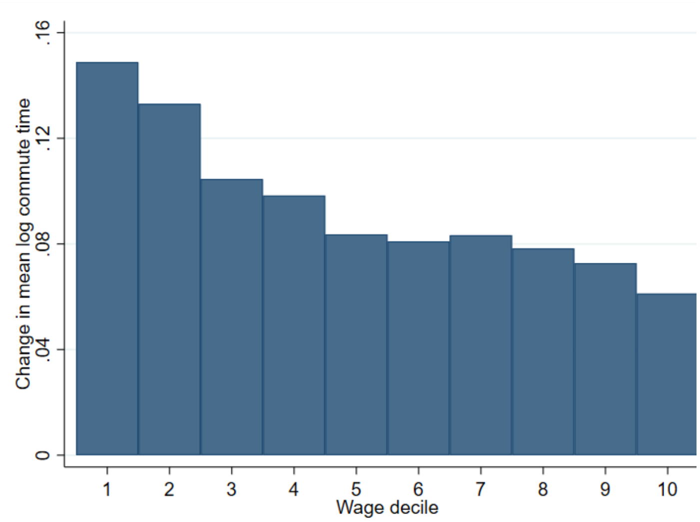
Notes: I plot the probability of working at least 50 hours a week using the CPS ASEC data from 1968 to 2016. The sample includes workers that are male, between age 25 and 65 and work at least 30 hours per week. I plot the probability of working long hours for workers in the top wage decile and the bottom wage decile over time. To smooth the plotted curve, each dot represents a three-year moving average.

Figure 4: Changing working hours and commute time by wage decile (1980-2010)

(a) Higher-wage workers more likely to work long hours



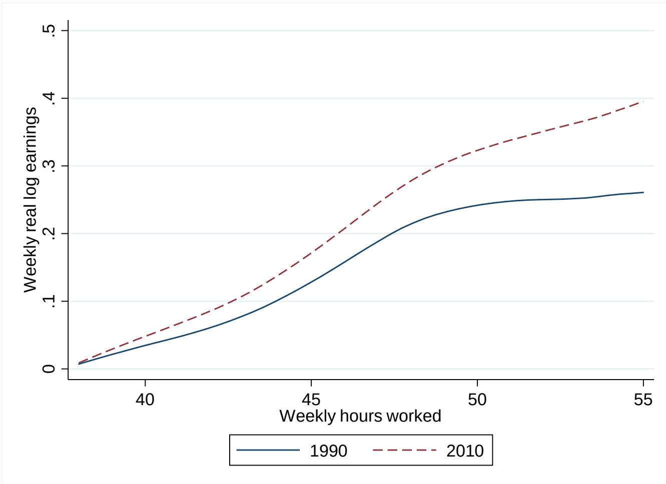
(b) Growth in commute time slower for higher-wage workers



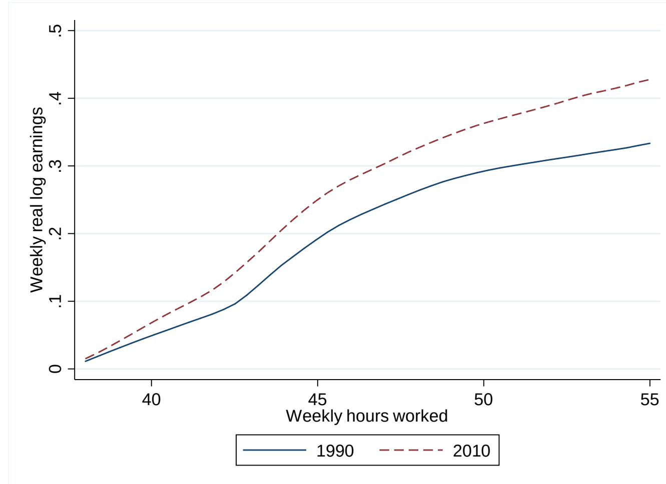
Notes: Data come from IPUMS census data in 1980 and 2010 (2007-2011 ACS). In a), I compute the change in probability of working at least 50 hours per week. The sample I use includes workers that are between 25 and 65 of age, males, and working at least 30 hours per week. I include only male in the sample to ensure that the changing female labor force participation does not distortion the statistics. In b), I compute the change in log commute time reported in the Census/ACS data. The sample includes workers that are between 25 and 65 of age, males, working at least 30 hours per week and living in the most populous 25 MSAs in the US.

Figure 5: Residual log weekly earnings against weekly hours worked

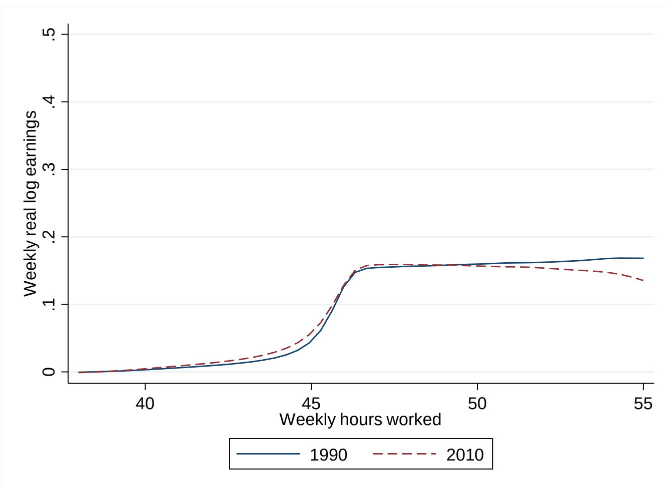
(a). Financial specialists



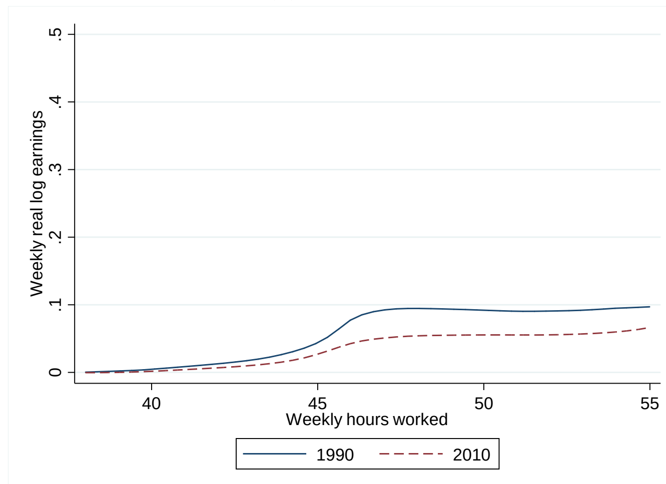
(b). Lawyers



(c). Secretaries and administrative assistants



(d). Teachers

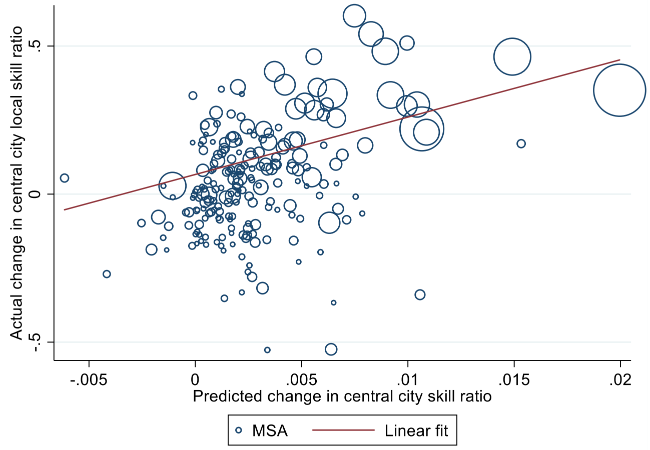


Notes: All samples come from Census data in IPUMS. ACS 2007-2011 is used for year 2010. The variables used in the plots are residual values after being regressed on individual level control variables (age, sex, race, education, industry code). The residual log earnings are normalized by constants such that the values in 1990 and 2010 start out from zero to help visual contrast. Financial specialists (a) include financial managers (occ2010- 120), accountants and auditors (occ2010- 800), and securities, commodities, and financial services sales agents (occ2010- 4820). Teachers include elementary school teachers (occ2010- 2310) and secondary school teachers (occ2010- 2320). The plot is the kernel-weighted local polynomial smoothing curve, with bandwidth equals 2.5, and Epanechnikov kernel function.

Figure 6: Predicted change in skill mix in central city vs actual change

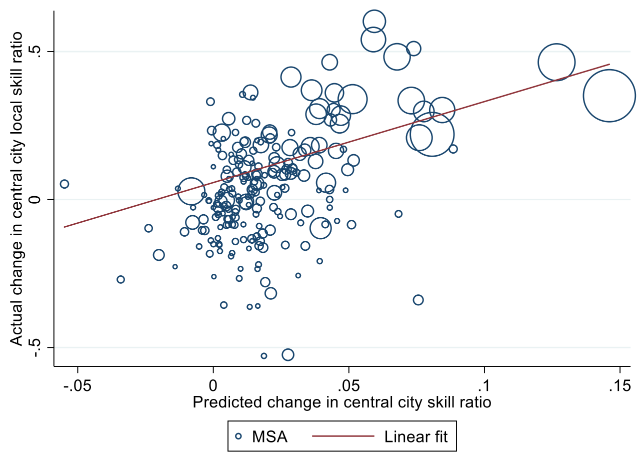
(a) Model specification: direct effect of commute cost only

(R-squared=0.3055)



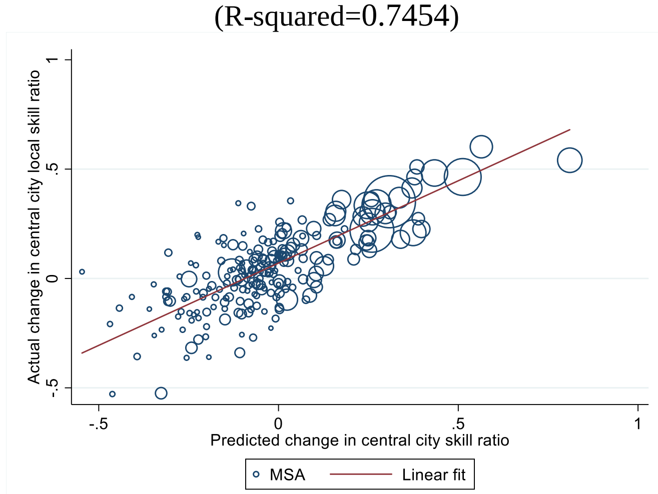
(b) Model specification: change in local skill ratio predicted by value of time shock

(R-squared=0.3500)



(c) Model specification: change in actual local skill ratio

(R-squared=0.7454)



Notes: These graphs are plots of actual log change in local skill ratio and the log change in local skill ratio predicted by the model in specification (a) (b) (c), respectively. Each circle represents an MSA. The local skill ratio is the computed as the change in local skill ratio in the central city. The values of local skill ratio are the ratio of the number of local residents of high-skilled occupations (more than 40% of workers have college degree in 1990 Census) and the number of local residents of low-skilled occupations.

Table 1: Relationship between local skill ratio and supply of local amenities

	Dependent variable: Δ ln (measurement of the selected amenity)					
	(1)	(2)	(3)	(4)	(5)	(6)
	Restaurants per 1000 residents	Grocery stores per 1000 residents	Gyms per 1000 residents	Personal serv. estab. per 1000 residents	Property crime per 1000 residents	Violent crime per 1000 residents
Δ ln (local skill ratio)	0.284***	0.013	0.454***	0.528***	-0.184**	-0.266***
	(0.036)	(0.033)	(0.034)	(0.038)	(0.072)	(0.047)
MSA fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	19,291	19,291	19,291	19,291	1,870	1,870
R-squared	0.1143	0.1246	0.1072	0.1751	0.3761	0.5415

Notes: Results shown above are OLS regressions, with sample from all MSAs. Each observation for column (1) – (4) are at census tract level. For each census tract, I sum up all the relevant business establishments (measured at zip code centroid) located within 1-mile radius, and I sum up the population in census tracts located within 1 miles, and compute the count of establishments per 1000 residents. The local skill ratio is computed as the ratio of the number of workers in high-skilled occupations and the number of workers in low-skilled occupations summed over all census tracts within 1 miles of each census tract. Conley (1999) HAC standard errors are reported with 1-mile threshold kernel function bandwidth. Each observation for column (5) – (6) are at municipality level. To compute local skill ratio for (5) – (6) I match census tracts to municipalities, and compute the overall local skill ratio using variables summed over across census tracts matched to municipalities.
*** p<0.01, ** p<0.05, * p<0.1

Table 2: Estimates of model parameters

Panel A: Worker's residential location demand			Panel B: Rent (housing supply equation)	
Commute cost (μ)		3.1661*** (0.4034)		
Amenity (γ)	High-skilled occupations	1.8512*** (0.1637)	Housing demand× housing stock density (π_1)	0.1669*** (0.0262)
	Low-skilled occupations	0.4494*** (0.1424)	Housing stock density (π_2)	0.0123*** (0.0045)
Rent (β)		0.5704*** (0.1895)		

Notes: Model estimated using occupation/census tract cell data from 1990 to 2010. Number of cells used is 8,755,411. The number of workers in each occupation/MSA in 1990 is used as analytical weight. I control for total expected commute (using expected commute time to jobs unrelated to workers' occupations), and I allow the coefficients on total expected commute to vary by occupation. Conley (1999) HAC standard errors are computed with 1-mile threshold for the kernel function. Since value of time is estimating using national (minus MSA in question) data, I adjust the standard error for the generated regressor using method introduced in Murphy and Topel (1985). Estimation detail can be found in the text.
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Estimates for amenity supply equations

	Dependent variable: Δ ln (measurement of the selected amenity)					
	(1)	(2)	(3)	(4)	(5)	(6)
	Restaurants per 1000 residents	Grocery stores per 1000 residents	Gyms per 1000 residents	Personal serv. estab. per 1000 residents	Property crime per 1000 residents	Violent crime per 1000 residents
Δ ln (local skill ratio)	0.514*** (0.165)	-0.165 (0.150)	1.058*** (0.136)	0.894*** (0.157)	0.739 (0.469)	-1.771*** (0.545)
MSA fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	19,291	19,291	19,291	19,291	1,870	1,870

Notes: Results shown above are GMM/IV regressions, with sample from all MSAs. I use the change in log number of high-skilled workers and change in log number of low-skilled workers predicted by expected commute time and change of value of time as instrumental variables for the change in local skill ratio. Each observation for column (1) – (4) are at census tract level. Each census tract, I sum up all the relevant business establishments located within 1-mile radius, and I sum up the population in census tracts located within 1 mile, and compute the count of establishments per 1000 residents. The local skill ratio is computed as the ratio of the number of workers in high-skilled occupations and the number of workers in low-skilled occupations summed over all census tracts within 1 miles of each census tract. Conley (1999) HAC standard errors are reported with 1-mile threshold kernel function bandwidth. Each observation for column (5) – (6) are at municipality level. To compute local skill ratio for (5) – (6) I match census tracts to municipalities, and compute the overall local skill ratio using variables summed over across census tracts matched to municipalities.
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Welfare gap between high and low-skilled workers

Year	Welfare gap due to **earnings** **(**earnings gap)	Welfare gap due to **earnings** & **commute cost**	Welfare gap due to **earnings** & **commute cost** & **rent**	Welfare gap due to **earnings** & **commute cost** & **rent & amenities**
1990	0.7368	0.7368	0.7368	0.7368
2010	0.8	0.7932	0.798	0.8209
Change	0.0632	0.0564	0.0612	0.0841
Effect on the change in welfare gap	---	-0.00678	-0.00195	0.0209

Notes: Welfare is computed as the mean expected utility of workers in each occupation in 1990 and 2010, weighted by the number of workers in each occupation in each MSA. Welfare gap is the difference between expected utility for high-skilled workers and low-skilled workers rescaled by the marginal utility of log earnings μ. I normalize the welfare gap in 1990 to its earnings gap. High-skilled workers are workers in high-skilled occupations, which are defined as occupations in which more than 40% of the sample hold college degree in 1990 Census. The last row "effect on the change in welfare gap" denotes the difference between the change in welfare gap and the change in earnings gap. This is shown to highlight how each factor adds or offsets the widening welfare gap.

Appendix for online publication

The appendix contains a few sections. Section A presents some extra details regarding my spatial equilibrium model. Section B contains the data appendix, in which I discuss how I obtain census tract residential locations, zip-code level job locations, and the procedure through which I obtain the travel time matrix. Section C describe some extra descriptive statistics that augment the analysis in the main text. Section D contains the estimation appendix, in which I discuss some reduced-form analysis, estimation of long-hour premium, and various validation test for long-hour premium, lasso analysis of occupation characteristics, alternative measurement of the value of time and some technical details of the estimation procedure. Section E contains detailed description of the robustness checks.

A Model

A.1 Worker's location choice problem

This section details the solution procedure that derives workers' indirect utility function given location characteristics (rents and amenities) and workers' occupation, from the basic assumption of Cobb-Douglas utility function. There is some additional complication in the solution procedure due to the fact leisure time enters the earnings function nonlinearly (I model *log* earnings linearly). Therefore, I solve the problem in two steps. In the first step, I solve the utility maximization problem holding leisure hours fixed, which makes the problem a standard utility maximization of Cobb-Douglas utility function. Once I obtain the partial indirect utility function given each level of leisure level, I then solve for optimal leisure hours, and in turn solve for the indirect utility function. Finally, I normalize the indirect utility function with the standard deviation of the idiosyncratic component of worker's preference. This section gives the complete micro-foundation of the indirect utility equation.

Step 1: Solve for partial indirect utility given leisure consumption.

Given the workers' utility function of C, H and L, I fix L first, and solve for C and H first. Let $\theta_L = 1 - \theta_C - \theta_H$.

$$\max_{C,H} U\left(C, H, L, A_{jmt}\right) = C^{\theta_C} H^{\theta_H} L^{\theta_L} A_{jmt}^{\tilde{\gamma}_k} \exp\left(-\tilde{\omega}_t c_{jnmt}\right) \exp\left(\sigma \varepsilon_{i,jmt}\right)$$

subject to budget constraint

$$C + R_{jmt} H = \exp(y_{0kt} + v_{mkt}\left(T - L - c_{jnmt}\right))$$

By Cobb-Douglas functional form, the demand for tradable consumption and housing services is (let I denote weekly earnings):

$$C^{*} = \frac{\theta_C}{\theta_C + \theta_H} I$$

$$H^{*} = \frac{\theta_H}{R_{jmt}\left(\theta_C + \theta_H\right)} I$$

The log-transformed partial indirect utility given leisure consumption L is then

$$V_{i,jnmt}\left(L\right) = \theta_C \log\left(\frac{\theta_C}{\theta_C + \theta_H} I\right) + \theta_H \log\left(\frac{\theta_H}{R_{jmt}\left(\theta_C + \theta_H\right)} I\right) + \theta_L \log\left(L\right) - \tilde{\omega}_t c_{jnmt} + \tilde{\gamma}_k a_{jmt} + \tilde{\gamma}_k \zeta_{jmt} + \sigma \varepsilon_{i,jmt}$$

The equation can be simplified with some algebra manipulation and by substitute I with the earnings equation.

$$\begin{aligned} V_{i,jnmt}\left(L\right) = \theta_C \log\left(\frac{\theta_C}{\theta_C + \theta_H}\right) + \theta_H \log\left(\frac{\theta_H}{\theta_C + \theta_H}\right) + \left(\theta_C + \theta_H\right)\left(y_{0mkt} + v_{mkt}\left(T - L - c_{jnmt}\right)\right) \\ -\theta_H \log\left(R_{jmt}\right) + \theta_L \log\left(L\right) - \tilde{\omega}_t c_{jnmt} + \tilde{\gamma}_k a_{jmt} + \tilde{\gamma}_k \zeta_{jmt} + \sigma \varepsilon_{i,jmt} \end{aligned}$$

Step 2: Solve for optimal leisure choice and the indirect utility function.

The maximization of the second step is the following.

$$\max_{L} V_{i,jnmt}\left(L\right)$$

It can be seen that leisure consumption increases utility through the term $\theta_L \log\left(L\right)$. But higher leisure hours means lower working hours, which reduces log earnings. Optimal leisure is obtained by solving the tradeoff problem.

A simple first-order condition leads to:

$$L^{*} = \frac{1 - \theta_C - \theta_H}{\left(\theta_C + \theta_H\right) v_{mkt}}$$

Substituting optimal leisure back into the partial indirect utility, I obtain the indirect utility function.

$$\begin{aligned} V_{i,jnmt} = \theta_C \log\left(\frac{\theta_C}{\theta_C + \theta_H}\right) + \theta_H \log\left(\frac{\theta_H}{\theta_C + \theta_H}\right) \\ + \left(\theta_C + \theta_H\right)\left(y_{0mkt} + v_{mkt}\left(T - \frac{1 - \theta_C - \theta_H}{\left(\theta_C + \theta_H\right) v_{mkt}} - c_{jnmt}\right)\right) \\ -\theta_H \log\left(R_{jmt}\right) + \theta_L \log\left(\frac{1 - \theta_C - \theta_H}{\left(\theta_C + \theta_H\right) v_{mkt}}\right) - \tilde{\omega}_t c_{jnmt} + \tilde{\gamma}_k a_{jmt} + \tilde{\gamma}_k \zeta_{jmt} + \sigma \varepsilon_{i,jmt} \end{aligned}$$

I then re-normalize the indirect utility function by dividing the entire utility function by σ, so that I can interpret the all coefficients as migration elasticities.

$$
\begin{aligned}
V_{i,jnmt} = \frac{1}{\sigma}\left(\theta_C \log\left(\frac{\theta_C}{\theta_C+\theta_H}\right) + \theta_H \log\left(\frac{\theta_H}{\theta_C+\theta_H}\right)\right) \\
+\left(\frac{\theta_C+\theta_H}{\sigma}\right)\left(y_{0mkt} + v_{mkt}\left(T - \frac{1-\theta_C-\theta_H}{(\theta_C+\theta_H)\,v_{mkt}} - c_{jnmt}\right)\right) - \frac{\theta_H}{\sigma}\log\left(R_{jmt}\right) \\
+\frac{1-\theta_C-\theta_H}{\sigma}\log\left(\frac{1-\theta_C-\theta_H}{(\theta_C+\theta_H)\,v_{mkt}}\right) - \frac{\tilde{\omega}_t}{\sigma}c_{jnmt} + \frac{\tilde{\gamma}_k}{\sigma}a_{jmt} + \frac{\tilde{\gamma}_k}{\sigma}\zeta_{jmt} + \varepsilon_{i,jmt}
\end{aligned}
$$

I simplify the above equation by combining terms and normalize the constant term to zero. . By doing so, I arrive at the following equation which is the one presented in the main body of the paper.

$$
V_{i,jnmt} = \delta_{mkt} - \mu v_{mkt} c_{jnmt} - \omega_t c_{jnmt} - \beta r_{jmt} + \gamma_k a_{jmt} + \gamma_k \zeta_{jmt} + \varepsilon_{i,jmt}
$$

Each coefficient is written in terms of the underlying parameters:

$$
\delta_{mkt} = \left(\frac{\theta_C+\theta_H}{\sigma}\right)\left(y_{0mkt} + v_{mkt}\left(T - \frac{1-\theta_C-\theta_H}{(\theta_C+\theta_H)\,v_{mkt}}\right)\right) + \frac{1-\theta_C-\theta_H}{\sigma}\log\left(\frac{1-\theta_C-\theta_H}{(\theta_C+\theta_H)\,v_{mkt}}\right)
$$

$$
\mu = \frac{\theta_C+\theta_H}{\sigma}
$$

$$
\beta = \frac{\theta_H}{\sigma}
$$

$$
\gamma_k = \frac{\tilde{\gamma}_k}{\sigma}
$$

$$
\omega_k = \frac{\tilde{\omega}_k}{\sigma}
$$

B Data

B.1 Residential location imputation procedure

The key dependent variable in this research is the location choice of workers in different occupations and how their location choice changes over time. The choice set for workers is the set of neighborhoods given the city that the workers live in. The best geographic unit that captures the essence of neighborhood would be census tract. The boundary of census tracts is relatively stable over time, and census tracts are designed to be fairly homogeneous in terms of population characteristics and economic status. Therefore, census tract is the natural choice for the definition of neighborhood. Nevertheless, the lowest geographic identifier in the Census microdata released to the public in IPUMS is PUMA, which is a much more aggregate level than census tract. The data that I use for occupation-specific location data at census tract level are resident count by occupation group from each census tract, provided by the NHGIS. I then impute census tract level occupation-specific count of residents using census tract level summary statistics and PUMA level microdata. I document the imputation procedure below.

Since NHGIS only provide counts of residents at census tract level for at aggregate occupation level K, I would only observe n^j_K for each census tract j. My goal is to impute the count of residents by detailed occupation level k, namely n^j_k. I do so by first imputing $\hat{\theta}^j_{k|K}$, which is the conditional probability of being in occupation k given one is in occupation-group K. I compute $\hat{\theta}^j_{k|K}$ using IPUMS microdata at PUMA level, assuming that $\hat{\theta}^j_{k|K}$ is the same for every census tract j within the same PUMA area. Then, finally compute the census tract level count of residents in occupation k, by multiplying the count of residents in occupation group K with the imputed probability of a worker being occupation k given he/she is in occupation group K.

$$\hat{n}^j_k = \hat{\theta}^j_{k|K} \cdot n^j_K$$

Once I get $\hat{n}^j_k$, I generate the location choice probability for each occupation and in each city in each year s_{jmkt}, which is the probability of living in census tract j, conditional on living in MSA m, working in occupation k and at year t. The share of each neighborhood among each type of workers reveals information about the demand for the neighborhood, and it will be used in the location choice model to infer the mean indirect utility of the average workers in each occupation.

B.2 Employment location imputation procedure

The employment location information is derived from the ZCBP from U.S. Census Bureau, which provides establishment counts by the employment size of business establishments. The dataset comes at the level of detailed SIC and NAICS code for each zip code from 1994 on, annually. Unfortunately, the dataset does not go back farther than 1994. Therefore, I use the employment location data in 1994 to proxy those in 1990. The spatial distribution of employment changes fairly slowly over time, so I expect the four year difference in data is unlikely to bias the data significantly.

For each zip code z, I first impute the employment count n^z_h for each industry h using establishment count and establishment sizes. Establishment size data are in the form of tabulated count: count of establishments with 1-4 employees, 5-9 employees, etc. I sum up these establishment counts weighted by the mid-value of the employee counts, to impute the total employment count for each industry in each zip code. Then I use $\hat{\theta}_{k|h}$, which is conditional probability of working in occupation k, given he/she works in industry h, to impute the number of employment in occupation k at zip code z. $\hat{\theta}_{k|h}$ is computed using contemporaneous national microdata from IPUMS.

$$\hat{n}^z_k = \sum_h n^z_h \cdot \hat{\theta}_{k|h}$$

The set of $\hat{n}^z_k$ measured for each zip code and each occupation will form the basis of the spatial allocation of employment. I use these data and travel time matrix to compute expected commute time for each census tract and for worker of each occupation.

B.3 Data acquisition procedure for travel time matrix

I acquire the travel time and travel distance from the Google Distance Matrix API (Application Programming Interface). The number of entries in travel matrix from every census tract to every zip code within every MSA is more than 7 million (7,363,850), which is too large to extract from the API directly. One reason that such travel matrix suffers from the curse of dimensionality is that large metro areas such as New York contain very large number of entries connecting numerous locations that are very far apart. For example, from east Long Island to Manhattan, there are tens of thousands of entries connecting all zip codes to all census tracts in Manhattan and east Long Island, even though most of these entries have almost identical travel time and distances. Hence, it is in fact not necessary to compute distance and time for all entries between census tracts and zip codes. I can group various zip code destinations and compute travel distance and time from all census tracts to one destination per zip code group if the trip distance is very long, and thereby reducing the dimensionality of the data dimension.

An intuitive real-life example that demonstrates this logic would be the use of GPS navigation for a long trip. When taking a long trip by car (such as from Palo Alto to San Francisco), setting the GPS destination in whichever specific location near downtown San Francisco would not make much of a difference, because one has to get on the freeway and the exact location of the destination makes *relatively* little impact on the ETA. However, if one takes a trip that is around 3 to 4 miles that starts and ends within San Francisco, ETA would be sensitive to the exact location of the destination.

Motivated from this observation, I reduce dimension by only directly extracting travel distance and time information between census tracts and zip code for all pairs that are located within 5 miles Euclidean distance (centroids of census tracts and long/lat of zip code gazetteer). For the pairs that are farther than 5 miles apart, I proxy the location of each zip code with the closest PUMA centroid, and I extract the travel distance and time between each census tract to the assigned PUMA centroid. It significantly reduces the dimension required for the data extract.

B.4 Historical travel time

In this section, I describe how I generate the 1990 historical travel time matrix for each MSA. Why estimate historical travel speed? If Google map exists in 1990, I could easily compute the travel time matrix using the historical traffic data. Unfortunately, the Google traffic model is only applicable to today's traffic condition and can only generate reliable travel time matrix relevant for the present day. One obvious concern of using today's traffic condition is measurement error problem. But a much bigger concern is that traffic condition is a highly local variable and it is very likely to be endogenous to location demand. Here is an example of such endogeneity problem. An exogenous demand surge (e.g. amenity shock) for a certain neighborhood location X makes traffic around location X more congested, which prolongs travel time to and from location X. The long travel time into and out of location X coupled with the observation of a demand surge for location X would lead the model to interpret that the demand surge is caused by people's desire to save on commute time. Using today's traffic model could introduce this "self-fullfiling prophecy" that could introduce serious endogeneity problem into the estimation of the model. Hence, the historical travel time matrix needs to be traffic information from the past.

To that purpose, I use two sources of data, Google API and the 1995 National Household Travel Survey (NHTS), to impute the historical travel time matrix. I first impute the historical travel speed (using NHTS and Google) for all travel routes within MSAs in 1995 rush hour, and then multiply the historical travel speed with travel distance (from Google) for each route to get expected travel time.

First, I use Google Distance Matrix API to obtain travel time (with traffic model turned off) and travel distance from each census tract to each zip code within each MSA. I make sure that travel time from Google is derived under the condition that the trips take place at midnight, so that no traffic is expected. The traffic-free travel time gives me information on the route fixed-effects (such as slowing-down effect of crossing a bridge, windy road, or dense city blocks with traffic lights).

Second, I use the 1995 NHTS data to fit a simple traffic speed model (Couture 2016) so that I could take the parameters estimated in the model onto the observable neighborhood characteristics in the 1990 Census and predict historical travel speed. I model travel speed as following:

$$\log(\text{speed}_{jnt}) = \beta_{0,t} + \beta_{1,t}\log(\text{distance}_{jn}) + \beta_{2,t}\log(\text{distance}_{jn})^2 + \bar{\mathbf{X}}_{jn}\Gamma_t + d_{jn} + \epsilon_{jnt}$$

j is origin census tract; n is the destination zip code; t is the year in which the trip is taken. I assume log speed of the trip is a function of trip distance, because longer trips usually have higher speeds because people take freeway or use main thoroughfare when the distance is long enough. I assume travel speed is also a function of the average neighborhood characteristics (population density, median income, and percentage of population working) of the origin and destination. Travel speed heavily depends on the types of neighborhood on which the trips take place. A trip to or from densely populated neighborhoods are expected to experience heavier congestion than another trip taken place in the suburbs. Additionally, I assume each route admits a time-invariant fixed-effects component, which accounts for the road conditions other than traffic congestion, such as slowing-down effect of crossing a bridge, windy road, or dense city blocks with traffic lights. I assume these fixed-effects do not change over time. The parameters of the model $\beta_{0,t}$, $\beta_{1,t}$, $\beta_{2,t}$, Γ_t governs how location characteristics and trip distance are mapped into travel speed. Since traffic condition evolves over time, these parameters are assumed to be year-specific.

I use 1995 NHTS data to estimate these parameters to obtain parameters applicable to 1995 traffic condition. I restrict the trip samples to those take place Monday to Friday and with departure time between 6:30 to 10:30 am and between 4:30 to 8:30 pm. I also restrict the trips either originate from or destine toward respondents' location of residence. $\bar{\mathbf{X}}_{jn}$ takes the location characteristics of the census tract which respondent lives in (neighborhood characteristics for the other end of the trip is unavailable). Additionally, I use Google API travel time (with traffic model turned off) to estimate the fixed-effects d_{jn} for each route. I impute traffic speed using the following equation.

$$\log(\widehat{\text{speed}}_{jn,1995}) = \hat{\beta}_{0,1995} + \hat{\beta}_{1,1995}\log(\text{distance}_{jn}) + \hat{\beta}_{2,1995}\log(\text{distance}_{jn})^2 + \bar{\mathbf{X}}_{jn}\hat{\Gamma}_{1995} + \hat{d}_{jn}$$

The travel time is then obtained by multiplying imputed travel speed with travel distance

$$\text{time}_{jn,1995} = \exp\Big(\log(\widehat{\text{speed}}_{jn,1995})\Big) \cdot \text{distance}_{jn}$$

C Descriptive statistics

C.1 Definition of central city neigbborhoods

As described in the descriptive statistics section of the paper, central city neighborhoods are defined as census tracts that fall within the 5-mile pin of downtown (defined by Google search). In Figure A1, I show the maps of a few cities as examples. In the map, the pin is defined as the point of downtown. The smaller circle represents the 3-mile radius, and the larger circle represents the 5-mile radius. The definition of central city neighborhoods throughout the paper is given by the 5-mile radius of the downtown pin.

C.2 Neighborhood change on the map (Chicago and New York)

The first descriptive facts that I show in the paper (Figure 1) is that income ratio between central city and suburban neighborhoods declined precipitously and reversed dramatically after 1980. The reversal of fortune in the central cities after 1980 is the main subject of ths paper.

Therefore, to build intuition for such change after the 1980, I demonstrate the neighborhood changes on maps for two prominent cities in the United States: Chicago and New York. I rank census tracts by income quintile within Chicago's MSA and New York's MSA, then plot the income quintile by the census tract's distance to downtown for the three decades from 1980 to 2010. Figure A4 shows that central city neighborhoods in Chicago are overwhelmingly low-income relative to the overall MSA income level in 1980, but after several decades of increase, central city neighborhood income levels are well above the overall MSA income level. A similar pattern can be observed in New York's central city neighborhoods in Figure A4. To various degrees, most major MSAs in the U.S. exhibit a similar pattern of income reversal between central cities and suburbs.

Furthermore, in Figure A5, I plot the census tract income quintile by distance to downtown for Chicago and New York. One can clearly see that the census tracts near downtown experienced a dramatic increase in their rank since 1980.

C.3 Central city population

The terms "gentrification" or "urban revival" may give the impression that central neighborhoods are now seeing faster overall population growth than the suburbs. However, while central neighborhoods may be gaining in terms of absolute population, they have not gained in terms of shares of overall

MSA population, since population growth in the suburbs continues to outpace that in central cities. American cities overall were still suburbanizing as recent as from 2000 to 2010, but at a much slower pace. Figure A6 shows the share of central neighborhoods' population as a percentage of total metropolitan population in the 25 most populous MSAs. The revived demand for central neighborhoods comes primarily from high-income workers and not all workers.

C.4 Change in work hours and commute time by wage decile

I use this section to discuss the timing of the rising probability of working long hours as well as the timing of the growth of commute time. In the paper, I highlight the fact that high-wage workers experienced a rising probability of working long hours and a slower growth of commute time between 1980 and 2010, which coincides with episode of gentrification. If we zoom in, we find that the sharp increase in the probability of working long hours occurred mainly before 2000. Coincidentally, the strong negative relationship between the growth in commute time and wage decile also mainly occurred before 2000. After 2000, both high- and low-skilled workers actually were less likely to work long hours (although low-skilled workers' probability of working long hours decreased much more). Also, after 2000, the negative relationship between growth in commute time and wage decile disappears. In fact, the workers in the top wage decile actually experience a weakly stronger growth in commute time than workers in lower wage deciles do.

Figure A2 shows the changing probability of working long hours by wage decile for two different periods: 1. 1980 - 2000; 2. 2000 - 2010. Figure A3 shows the growth in commute time by wage decile for the same two periods. These facts are suggestive evidence that the rising value of time provided the *initial* force that attracted high-skilled workers into the central cities. Once the endogenous amenity process starts, many high-skilled workers started to move into the city due to amenity rather than shorter commute. As amenity change evolved, the role of amenities started to overwhelm the role of shorter commute time. In fact, many high-skilled workers live in the central cities for the amenities even though they work in the suburbs. This explains why the high-wage workers experience slightly higher growth in commute time between 2000 and 2010. This evidence is also consistent with Couture and Handbury (2017)'s results in which reverse commuting became more prevalent after 2002.

D Estimation

D.1 Reduced-form relationship between changing long-hour premium and central city sorting

With the value of time defined and measured as long-hour premium, I show that there is a positive relationship between changes in the long-hour premium and the increase in preference for living in central city neighborhoods.

I first assign the estimated long-hour premiums as the measurement for value of time to each occupation in each MSA for years 1990 and 2010 using the methodology described previously. I estimate the long-hour premium for each MSA, excluding information for that MSA. I take an additional precautionary step by also excluding data on all states which the MSA borders. I do this to ensure that the long-hour premium measurement assigned to workers in each MSA is not confounded by local labor supply factors in that MSA. In Figure A9a), I show the binscatter plot that shows a positive relationship between changes in the value of time (long-hour premium) and changes in the demand for central city locations. To corroborate that the relationship is indeed related to workers' intention to shorten their commute time, in Figure A9b) I show another binscatter plot that shows a negative relationship between changes in commute time and changes in the value of time. In other words, stronger growth in the value of time leads to *slower* growth in commute time.

While the relationship between the rising value of time and the increasing preference for central city housing may suggest that the value of time is a likely driver for the changing taste for central city locations among workers and a likely cause of gentrification, these relations could still be driven by a spurious correlation between a changing preference for urban amenities and the skill content of different occupations. Since high-skilled workers are more likely to face stronger labor demand shocks due to skill-biased technological change, and high-skilled workers generally have a stronger preference for urban amenities (Diamond (2016); Couture and Handbury (2017)), any correlation between demand-driven change in the value of time and rising preference for central city locations may be confounded by a shifting taste for city amenities among high-skilled workers. Therefore, in the paper, I conduct an empirical analysis using a spatial equilibrium framework, and specifically estimate the taste for commute time for workers in each occupation and show that workers' rising value of time leads to their lower tolerance for commute time. I then use the estimated preference parameters to assess the role of the value of time in bringing about gentrification.

The other problem with the reduced-form relationship between value of time and location preference is that it cannot help to answer the question of whether the shock to the value of time changes the amenity levels of central city neighborhoods, and how the endogenous change in amenity levels affects people's location demand for central cities. To fully recover the mechanism of worker's location choice preference, one ought to causally estimate workers' migration elasticity with respect to cost of commute, neighborhood amenity, and rent simultaneously, and use these elasticities to quantitatively gauge the relative importance of each channel.

D.2 Potential biases in estimating long-hour premium

The long-hour premium is measured off the cross-sectional relationship between weekly log earnings and weekly hours worked. One potential reason for biased estimate for LHP (long-hour premium) is that weekly hours worked is a result of workers' labor supply choice, and therefore the variable of hours worked may be endogenous.

In the context of my estimation, the variation that I use to identify the spatial equilibrium model is the differential *change* in the long-hour premium. While the endogeneity of the hours variable may overstate the size of the static estimate of long-hour premium, the real threat to identification is if the change in the estimated LHP within occupations is driven by *changing* degree of sorting on earnings and hours described above.

To fix idea, consider the case of financial workers. It is possible that over time, high-ability financial workers increasingly supply longer hours and receive higher earnings, relative to the low-ability counterparts. Their increasing supply of long-hour may simply due to a change in preference, and the fact that they receives higher earnings may not be *due to* their increasingly longer working hours, but instead *due to* their high-abilities. As a result of this increasingly selection on abilities, I would observe increasing association between high earnings and high work hours among financial workers, but such association may be not driven by the increasing payoff of working long hours.

If I observe workers' true abilities, I would re-estimate the long-hour premium controlling for the levels of ability and see whether controlling for abilities would change the estimate for LHP. The difference between LHP estimates with and without control for levels of ability indicates the degree of selection on workers' abilities. If the degree of selection on ability increases over time, it would raise suspicion that long-hour premium estimate may be driven by increasing selection effect.

Since I do not observe workers' unobservable abilities, I conduct a similar test on observable abilities: reported education levels. I assume that if there is increasing selection on the unobservable abilities, I should see the same increase in selection on the observable abilities, such as education levels (Altonji, Elder, and Taber (2005)).

To do that, I re-estimate the long-hour premium for several key occupations with and without controlling for the levels of education, and show the comparison of LHP estimates.

Figure A10 shows the degree of selection on the observable skills for estimates of LHP in 1990 and in 2010. The degree of selection is computed as the difference between the LHP estimates without education control and estimates with education control. There are two observations that can be made here: 1. there is selection effect on the observable skill levels for almost all occupations for the level estimates of LHP in both 1990 and 2010; 2. the selection effect is larger for occupations with more skill content. These observations suggest that the level estimates of LHP are likely partly driven by selection effect on the unobservable skill levels.

However, in Figure A11, I show the *change* in the degree of selection on observable skills. The selection effect on average is not increasing. In fact, occupations are equally likely to see increasing and decreasing selection effect. In addition, the change in selection effect is not correlated with skill content of the occupation at all. Since the estimates for LHP is not driven by selection by the observables, the *change* in LHP estimated from cross-sectional data is unlikely to be driven by the *changing* degree of selection by the unobservables.

Table A8 in the appendix reports some of these estimates. The level estimates are smaller with education control. But the change in estimated LHP does not seem to sustain a substantial effect from adding education control. For computer scientist profession, the negative bias in LHP is relatively large by adding education control, but even for this, the bias is around 11%. On the other hand, for lawyer profession, the change in estimated LHP is actually larger with education control.

D.3 Alternative measures of the value of time

Another variable that tracks the marginal earnings of hours supply could be constructed based on a "tournament scheme" of compensation, in which workers get paid with prizes from tournament competitions within firms or within labor markets (Lazear and Rosen (1981)). A "prize" such as a job promotion or securing lucrative projects is awarded to the workers who outperform their competitors. Under this scheme, increasing work effort can increase the chance of winning such prize. If the reward of a "prize" is very high, the payoff of effort is thus likely very high, since even narrowly losing the "tournament" means missing the prize entirely. Therefore, the effort level is an increasing function of the prize *spread* between winning and losing. Since work hour is a crucial input of worker's effort level, the marginal earnings of hours supply would rise if the reward spread between winning and losing the "tournaments" becomes higher (Bell and Freeman (2001)). A measurement of log earnings dispersion within the same occupation could track the size of the "spread" of the financial reward for workers in the occupation. Therefore, I use the Census data and compute the standard deviation of the residual log earnings for each occupation, after controling for the individual characteristics, and I use it as an alternative measurement for value of time for the purpose of checking robustness of the main results.

D.4 Validation tests for long-hour premium and earnings dispersion as the value of time

To validate the measurements of the long-hour premium, I show that the occupations with positive change in long-hour premium tend to have increasing prevalence of working long hours (working at least 50 hours per week). In Figure A8, I show the relationship between the change of log probability of working long hours on the change of long-hour premium by occupation in a binscatter plot, and the result shows that rising long-hour premium is significantly associated with rising incidence of working long hours.

Also, I find that rising earnings dispersion contribute to the rising prevalence of working long hour as well. Figure 9b show the relationship between the change of log probability of working long hours and the change of log earnings dispersion in a binscatter plot, and I find strong correlation between them, which suggests the rising earnings dispersion is likely to raise marginal earnings of hours as well.

To further validate the measurements of long-hour premium as the value of time, I use data on occupation characteristics from the O*NET website (developed under the sponsorship of the U.S. Department of Labor/Employment and Training Administration) to characterize occupations with rising long-hour premium. I conduct a lasso regression analysis for estimated long-hour premium. Lasso analysis can help select a subset of occupation characteristics that can best predict the variation in the changes in long-hour premium. I then assess whether the characteristics selected by the lasso regression make sense intuitively.

I extract 57 work-context variables from O*NET database. Each work-context variables tracks the score that workers and labor experts give for each occupation on a specific occupational characteristics (e.g. level of competition). These variables describe the importance of interpersonal relationships (14 variables), physical work conditions (30 variables), and structural job characteristics (13 variables). The scores are collected at various times before and after 2010. I use the mean score given by both workers and expert for each variable.

In the lasso analysis, I use the change of long-hour premium from 1990 to 2010 as the outcome variable and the work-context characteristics as covariates. I describe the detail of lasso regression in the next subsection.

Out of the 57 work characteristics, the selected characteristics which positively correlate with change in long-hour premium and remain in the lasso regression are "time pressure," "degree of automation," "frequency of decision making," and "importance of repeating the same tasks." It is remarkable that lasso analysis picks up "time pressure" as among the variables that effectively explain the variation in the change in long-hour premium. In Appendix Figure A12, I show the lasso trace plot of the regression. Since the idea of long-hour premium is the log return of working extra hours beyond standard full time, occupations with rising long-hour premium should be those with increasing demand for hours worked. The fact that degree of time pressure predicts the change in long-hour premium is an additional validation that the long-hour premium measure picks up information about value of time.

I also conduct lasso regression for change in earnings dispersion. Out of 57 work characteristics, the last five remaining characteristics from lasso regression that positively correlate with change in residual earnings dispersion are "level of competition," "duration of typical work week," "electronic mail," "outdoors, under cover," "Telephone." Interestingly, "level of competition" is among the variables in residual earnings dispersion. Since the measurement of earnings dispersion intend to capture the "spread" of prizes between winning and losing a "tournament" competition at workplace, the stake of winning should be higher if earnings dispersion is higher. The fact that stronger rise in earnings dispersion is well predicted higher level of workplace competition validates that very idea. In addition, the fact that "duration of typical work week" is picked up in the lasso further validates the idea that workers in occupations with stronger rise in earnings dispersion tends to work more and are more likely to be time-constrained.

D.5 Lasso regression using O*NET occupation characteristics

I project the change in long-hour premium and change in earnings dispersion of each occupation onto the 57 occupation characteristics from O*NET. I standardize the occupation characteristics by their respective mean and standard deviation, so that the variation in each variable is not confounded by the scale of each characteristic. I also standardize the outcome variables (change in long-hour premium and change in earnings dispersion).

The lasso coefficients is chosen by solving the following constrained minimization problem:

$$\min_{\beta_1 \ldots \beta_{57}} \frac{1}{N} \sum_{i=1}^{N} \left(y_i - \beta_1 x_{1i} - \ldots - \beta_{57} x_{57i}\right)^2$$
$$\text{subject to } \sum_{j=1}^{57} \left|\beta_j\right| \leq t$$

y_i is the outcome variable, which is the change in value of time (measured as change in long-hour premium or earnings dispersion). x_{ji} where $j = 1, ...57$ are the 57 standardized occupation characteristics from O*NET. t is some size constraint for the norm of the coefficients. There is no intercept in the regression because standardized variables are centered around zero.

One could rewrite the minimization problem with a minimization problme with a single equation and a Lagrange multiplier λ.

$$\min_{\beta_1 \ldots \beta_{57}} \frac{1}{N} \sum_{i=1}^{N} \left(y_i - \beta_1 x_{1i} - \ldots - \beta_{57} x_{57i}\right)^2 + \lambda \sum_{j=1}^{57} \left|\beta_j\right|$$

λ is the weight that the regression gives to the norm of all the regression coefficients. When λ is zero, the lasso regression coefficients are identical to those estimated from OLS regression. With large value of λ, I penalize large values of the coefficients on any of the explanatory variable, which forces the regression coefficients to drop out and become zero if the corresponding variables do not perform as well in predicting the variation in outcome variable and minimizing the mean squared residual. Therefore, with different level of λ, regression coefficients would be different. A useful exercise to do would be to raise the size of λ incrementally, and observe which explanatory variables drop out and which remain. Those that remain with large size of λ tend to be those with the best explanatory power.

Finally, I use the variable selection and coefficients that gives the minimum mean squared error under a 5-fold cross-validation.

D.6 Linearization of location demand

To facilitate the estimation procedure, I linearize the location demand equation by evaluating the equation with Taylor approximation around $\omega_t + \mu v_{mkt}$ at some constant. One can think of $\omega_t + \mu v_{mkt}$ as the marginal disutility of commute time. I let $\omega_t + \mu v_{mkt} = \phi$, so that commute time is discounted with a constant coefficient ϕ. Taking derivative for $\log\left(\sum_{n' \in J_m} \pi_{n'mk} \exp\left(-\left(\omega_t + \mu v_{mkt}\right) c_{jn'm}\right)\right)$ with respect to $\omega_t + \mu v_{mkt}$, leads to $-\widetilde{\mathrm{E}}\left(c_{jmk}\right)$ where $\widetilde{\mathrm{E}}\left(c_{jmk}\right)$ is the expected commute on an adjusted probability measure (the adjustment depends on the size of ϕ). Therefore, Taylor expansion around $\omega_t + \mu v_{mkt} = \phi$ equals the following equation.

$$\log\left(s_{jmkt}\right) \approx \log\left(\sum_{n'\in J_m} \pi_{n'mk} \exp\left(-\phi \cdot c_{jn'm}\right)\right) + \tilde{\delta}_{mkt} - \widetilde{\mathrm{E}}\left(c_{jmk}\right)\left(\omega_t + \mu v_{mkt} - \phi\right)$$
$$-\beta r_{jmt} + \gamma_z \log\left(\frac{N^H_{jmt}}{N^L_{jmt}}\right) + \mathbf{X}_{jmt}\Theta_{kt} + \xi_{jmkt}$$

The nonlinear term $\log\left(\sum_{n'\in J_m} \pi_{n'mk} \exp\left(-\phi \cdot c_{jn'm}\right)\right)$ becomes a constant that is j and k specific, which I write it as a fixed-effects term. $\phi\widetilde{\mathrm{E}}\left(c_{jmk}\right)$ can be pulled out and included in the fixed-effects term. After some algebraic arrangement, the location demand equation can be approximate as following

$$\log\left(s_{jmkt}\right) \approx \delta_{jmk} + \tilde{\delta}_{mkt} - \omega_t \widetilde{\mathrm{E}}\left(c_{jmk}\right) - \mu v_{mkt}\widetilde{\mathrm{E}}\left(c_{jmk}\right) - \beta r_{jmt} + \gamma_z \log\left(\frac{N^H_{jmt}}{N^L_{jmt}}\right) + \mathbf{X}_{jmt}\Theta_{kt} + \xi_{jmkt}$$

D.7 GMM estimation and standard errors

I use a iterative linear GMM estimation procedure to estimate the parameters of the model. The estimator and the standard errors need to address 1. correction for the statistical errors of the estimated value of time in the regressor; 2. potential spatial dependence of the error terms for observations that are physically in proximity to each other. In this section, I describe how I address these issues and derive the estimator and corresponding standard errors that are robust to these concerns.

Let $\mathbf{X}$ be the stacked matrix for model regressors of both equations; $\mathbf{Z}$ be the stacked matrix for instruments (both included and excluded) of both equations; $\mathbf{y}$ be the outcome variable vector. All variables are analytically weighted by the population of the cell data point. $\mathbf{W}$ be the optimal weighting matrix, and the estimating equations can be written as $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. $\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}$. $\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 & 0 \\ 0 & \mathbf{X}_2 \end{pmatrix}$, where $\mathbf{X}_1$ is the matrix of the location demand estimating equation and $\mathbf{X}_2$ is the matrix of the housing supply estimating equation. $\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_1 & 0 \\ 0 & \mathbf{Z}_2 \end{pmatrix}$, where $\mathbf{Z}_1$ is the matrix of the instruments for the first equation and $\mathbf{Z}_2$ is the matrix of the instruments for the second equation. Then the linear GMM is written as follows.

$$\hat{\boldsymbol{\beta}}_{GMM} = \left(\mathbf{X}'\mathbf{Z}\mathbf{W}\mathbf{Z}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{Z}\mathbf{W}\mathbf{Z}'\mathbf{y}$$

Obtaining the optimal weighting matrix $\mathbf{W}$ would depend on the standard error estimation, which I describe below.

D.7.1 Standard error correction for generated regressor (Murphy and Topel (1985))

Since value of time enters into one of the regressors of the estimating equation and value of time is estimated from national data (excluding the MSAs of interest), the statistical errors of the regressor should be taken into account in the standard errors. Otherwise, the standard errors may be underestimated. In the linear estimating equation, the only regressor that contain value of time is $\Delta\hat{v}_{mkt}\widetilde{\mathrm{E}}\left(c_{jmk}\right)$, in which $\Delta\hat{v}_{mkt}$ is estimated. For notation purpose, let $\hat{\mathbf{x}}$ be the vector for the term $\Delta\hat{v}_{mkt}\widetilde{\mathrm{E}}\left(c_{jmk}\right)$ and $\mathbf{x}$ be the vector for the term $\Delta v_{mkt}\widetilde{\mathrm{E}}\left(c_{jmk}\right)$, and μ is the coefficient on this term. Using similar method demonstrated in Murphy and Topel (1985), the estimating equation can be rewritten as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}+\mu\left(\mathbf{x}-\hat{\mathbf{x}}\right)+\boldsymbol{\varepsilon}$$

Hence, the GMM estimator becomes

$$\hat{\boldsymbol{\beta}}_{GMM} = \boldsymbol{\beta}+\left(\mathbf{X}'\mathbf{ZWZ}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{ZWZ}'\mu\left(\mathbf{x}-\hat{\mathbf{x}}\right)+\left(\mathbf{X}'\mathbf{ZWZ}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{ZWZ}'\boldsymbol{\varepsilon}$$

The variance-covariance matrix of the estimator can be writte in two terms, assuming that $\hat{\mathbf{x}}$ and $\boldsymbol{\varepsilon}$ are uncorrelated. It is a reasonable assumption, since the $\Delta\hat{v}_{mkt}$ is estimated using data *outside* of MSAs of interest.

$$\begin{aligned}\mathbf{V}_{\hat{\boldsymbol{\beta}}} &= \left(\mathbf{X}'\mathbf{ZWZ}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{ZWZ}'\mu^2\mathbf{V}_{\hat{\mathbf{x}}}\mathbf{ZWZ}'\mathbf{X}\left(\mathbf{X}'\mathbf{ZWZ}'\mathbf{X}\right)^{-1}\\ &\quad+\left(\mathbf{X}'\mathbf{ZWZ}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{ZWZ}'\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}\mathbf{ZWZ}'\mathbf{X}\left(\mathbf{X}'\mathbf{ZWZ}'\mathbf{X}\right)^{-1}\end{aligned}$$

To estimate the variance-covariance matrix, I need to account for the unknown parameters: μ, $\mathbf{V}_{\hat{\mathbf{x}}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$. Since μ is a parameter in the equation, I substitute it with the model estimate $\hat{\mu}$. $\mathbf{V}_{\hat{\mathbf{x}}}$ can be constructed with the estimated variance of $\Delta\hat{v}_{mkt}$ from the auxiliary value of time estimation.

D.7.2 Spatial Heteroskedasticity and Autocorrelation Consistent Standard Errors (Conley (1999))

I assume a non-parametric approach to account for the spatial dependence of $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$. Spatial dependence is a common issue in estimating models with highly localized spatial outcome variable. Among census tracts that are spatially close, the unobservable error term is very likely to be correlated. For example, the construction of a nice neighborhood park increases the attractiveness of all the census tracts that are located within a reasonable distance to the park. If the park construction is not included in the observable amenity shock variable, it would be included in the error term. Such error term would apply to all data points in proximity to the park. As a result, error terms are likely to be correlated across nearby census tracts. Even if they are clustered at census tract level, standard error could still be underestimated.

Conley (1999) standard error estimator is the standard procedure for adjustment for spatial dependence. I implement this estimator to obtain HAC standard errors robust to spatial dependence in $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$. The Conley estimator is spatial analogue of non-parametric estimator introduced in Newey and West (1984), in which the variance-covariance matrix of the moment restrictions $\mathbf{Z}'\mathbf{\Sigma}_{\boldsymbol{\varepsilon}}\mathbf{Z}$ is estimated from sample covariance using pairs of data points that are located within some distance from each other. The covariance is estimated with a Bartlett kernel weighting function, with bandwidth of 1 mile. For notation purpose, I set $\Omega = \mathbf{Z}'\mathbf{\Sigma}_{\boldsymbol{\varepsilon}}\mathbf{Z}$. I also set g_{jmk} to be the vector of a sample moment from the data, where $g_{jmkt} = Z_{jmkt}\hat{e}_{jmkt}$.

Using the Bartlett kernel weighting function, I assume that two observations may be spatially correlated if they are located within 1 mile to each other, and the weighting takes the Bartlett functional form:

$$K\left(d_{jj'm}\right) = \begin{cases} 1 - \frac{d_{jj'm}}{\bar{d}}, & \text{if } d_{jj'm} < \bar{d} \\ 0, & \text{if } d_{jj'm} \geq \bar{d} \end{cases}$$

$d_{jj'm}$ is the distance between census tract j and j', and $\bar{d}$ is the one mile threshold. The weights give more weight to pairs of observations that are closer to each other, and giving a weight of one for two observations from the same census tract. The weights decline from 1 to zero linearly, giving zero weight for observations that are farther than one mile apart. The estimate for Ω is constructed as following.

$$\hat{\Omega} = \frac{1}{N} \sum_{jmkt} \sum_{j'mk't'} K\left(d_{jj'm}\right) g_{jmkt} g'_{j'mk't'}$$

I implement the iterative estimation procedure for the GMM estimate, in which I first conduct estimation assuming the weighting matrix $W_0 = I$. Using the preliminary estimate $\hat{\boldsymbol{\beta}}^0_{GMM}$, I estimate $\hat{\Omega}_0$, and I let $W = \hat{\Omega}_0^{-1}$. Using the new weighting matrix, and then re-estimate the model parameters $\hat{\boldsymbol{\beta}}_{GMM}$ and $\hat{\Omega}$. I then repeat the process, until $\hat{\boldsymbol{\beta}}_{GMM}$ converges.

E Robustness

E.1 Spatial sorting to public transit stations

To further validate the role of value of time in shifting location demand, I estimate another specification of the location demand equation in which I also model workers' migration elasticity with respect to proximity to public transit stations. Figure A15 shows the maps of Chicago and New York as examples, in which the public transit stations are plotted. I assume that living close to public transit stations can save time for commuters who do not own automobiles or those who find driving to work less time-efficient. Table A2 shows the estimation results of the location demand equation in which I allow the value of time to impact workers' preference for a neighborhood's proximity to public transit station. I include interaction terms between value of time and an indicator variable of the neighborhood's being located within 1 mile of a transit station. The results show that workers with higher value of time have increasing demand for locations close to public transit stations.

E.2 Heterogeneous μ and β

I also estimate another version of the model in which high-skilled workers have potentially different values of μ or β, which are preference parameter for commuting cost and to rents. Note that in the main estimation, I assume that μ and β are homogeneous across high- and low-skilled workers. The only parameter that I allow for heterogeneity is γ, which is the preference parameter for amenities. Such setup is intended to keep the functional form the simplest and the easiest to interpret.

Table A3 shows the results from the specification in which I allow heterogeneity for μ or β. The estimates for μ are both positive and significant for high- and low-skilled workers, although the estimate is far larger for high-skilled workers than for low-skilled workers. This means that rising long-hour premium have stronger impact on the choice of commute time for high-skilled workers than it does for low-skilled workers. Interestingly, the preference elasticities to rents is higher for high-skilled workers, which means that high-skilled workers are more responsive to rent hike. If one look carefully, all preference elasticities are larger for high-skilled workers. This strongly suggests that the high-skilled workers are simply just more mobile (lower σ for their idiosyncratic component).

In the main specification of the model, I keep μ or β homogeneous for simplicity. Also, I keep μ homogeneous, because I want to avoid generating spatial sorting artificially through heterogeous μ. In the decomposition exercises, the rising value of time generates exogenous momentus for spatial sorting due to the *differential rise* in the value of time across occupation and the *differential spatial location* of jobs, not the differential value of μ.

E.3 Alternative measurements of value of time

In the model estimation, I use change in long-hour premium to proxy the change in the value of time for workers in each occupation. In this section, I use various alternative variables that may track the changing value of time, and test whether workers with rising value of these alternative measures increasingly choose locations with shorter commute time. If so, it would further validates my hypothesis. To do that, I estimate the following equation:

$$\Delta \log \left(s_{jmkt}\right) - \left(\hat{\gamma}_z \Delta \log \left(\frac{N^H_{jmt}}{N^L_{jmt}}\right) - \hat{\beta}\Delta r_{jmt}\right) =$$
$$\Delta\tilde{\delta}'_{mkt} + \Delta\omega_{t,alt}\widetilde{\mathrm{E}}\left(c_{jmk}\right) - \mu_{alt}\Delta v^{alt}_{mkt}\widetilde{\mathrm{E}}\left(c_{jmk}\right) + \Delta\theta_{kt}X_{jmk} + \Delta\xi_{jmkt}.$$

The equation is a rearrangement of the estimating equation of the full model. The outcome variable is the change in log demand *net* of amenity change and rent change. In other words, the outcome variable can be understood as the change in location demand with amenity levels and rents unchanged. v^{alt}_{mkt} is the alternative measurement of value of time (residual earnings dispersion in this case). $-\mu_{alt}$ is the migration elasticity with respect to commute cost under the alternative definition of the value of time. If the estimate for μ_{alt} is positive, that would validates the hypothesis that an increasing value of time leads to preference for a shorter commute.

Table A4 column (1) reports the estimate for μ_{alt} using residual earnings dispersion as a measure for value of time. I report results for the specification using gross change in log demand and net change in log demand (with and without holding amenity and rent values fixed). I find that μ_{alt} is positive, which means that rising earnings dispersion leads to preference for shorter commute time. To further corroborate the results, I replace Δv_{mkt}^{alt} with change in log probability of working at least 50 hours a week, and I show in column (2) that μ_{alt} is again positive, which means that workers in occupations with higher incidence of working long hours prefer location with shorter commute time.

E.4 Sorting by occupation characteristics

I present results from a lasso regression in which I select O*NET occupational characteristics that best explain the variation in the change in long-hour premium and earnings dispersion.

Here, I exploit the variation of these occupational characteristics directly, and examine whether workers in occupations with higher value of a certain characteristic exhibit increasing preference for shorter commute time. This exercise directly estimates how these occupational characteristics are correlated with the changing migration elasticity with respect to commute time. If workers who work in occupations that are more time pressured and more competitive become more averse to long commute time, it would provide further validation that value of time plays a role in people's changing location preferences. The location demand specification is the following:

$$\Delta \log\left(s_{jmkt}\right) - \left(\hat{\gamma}_z \Delta \log\left(\frac{N_{jmt}^H}{N_{jmt}^L}\right) - \hat{\beta}\Delta r_{jmt}\right) =$$
$$\Delta\tilde{\delta}'_{mkt} + \Delta\omega_{t,occ}\widetilde{\mathrm{E}}\left(c_{jmk}\right) - \sum_l \mu_{occ,l}x_l\widetilde{\mathrm{E}}\left(c_{jmk}\right) + \Delta\theta_{kt}X_{jmk} + \Delta\xi_{jmkt}.$$

I regress the change in location demand on expected commute time and the interaction terms between the selected occupation characteristics x_l and expected commute time. I include all of the characteristics selected by lasso and include two other characteristics that I suspect may be related to the incentive to supply hours: duration of work week and level of competition. A positive estimate for $\mu_{occ,l}$ means that workers with higher value of the occupation characteristic included (say, degree of time pressure) increasingly prefer locations with shorter commute time. I also estimate a specification of this equation in which I do not subtract the value of endogenous amenity and rent.

Table A5 reports estimation results. Interestingly, it shows that higher degree of time pressure and longer duration of work week are more likely to reduce commute time, while most other characteristics do not have positive impact on the incentive to shorten commute.

In addition, I use these occupation characteristics (time pressure, longer duration of work week, degree of automation, frequency of decision making, importance of repeating same tasks, time pressure, physical proximity, and spending time kneeling, crouching, stooping, or crawling) and jointly instrument for the change in long-hour premium in the location demand specification, then re-estimate μ to see whether the sign of the new estimate changes. I report results for the specification using gross change in log demand and net change in log demand (with and without holding amenity and rent values fixed). I use standardized variables of the occupation characteristics and their interactions with $\widetilde{\mathrm{E}}\left(c_{jmk}\right)$ as instrumental variables. Table A6 reports the estimate μ_{IV}, which has the same sign as the main estimates.

E.5 Other robustness tests

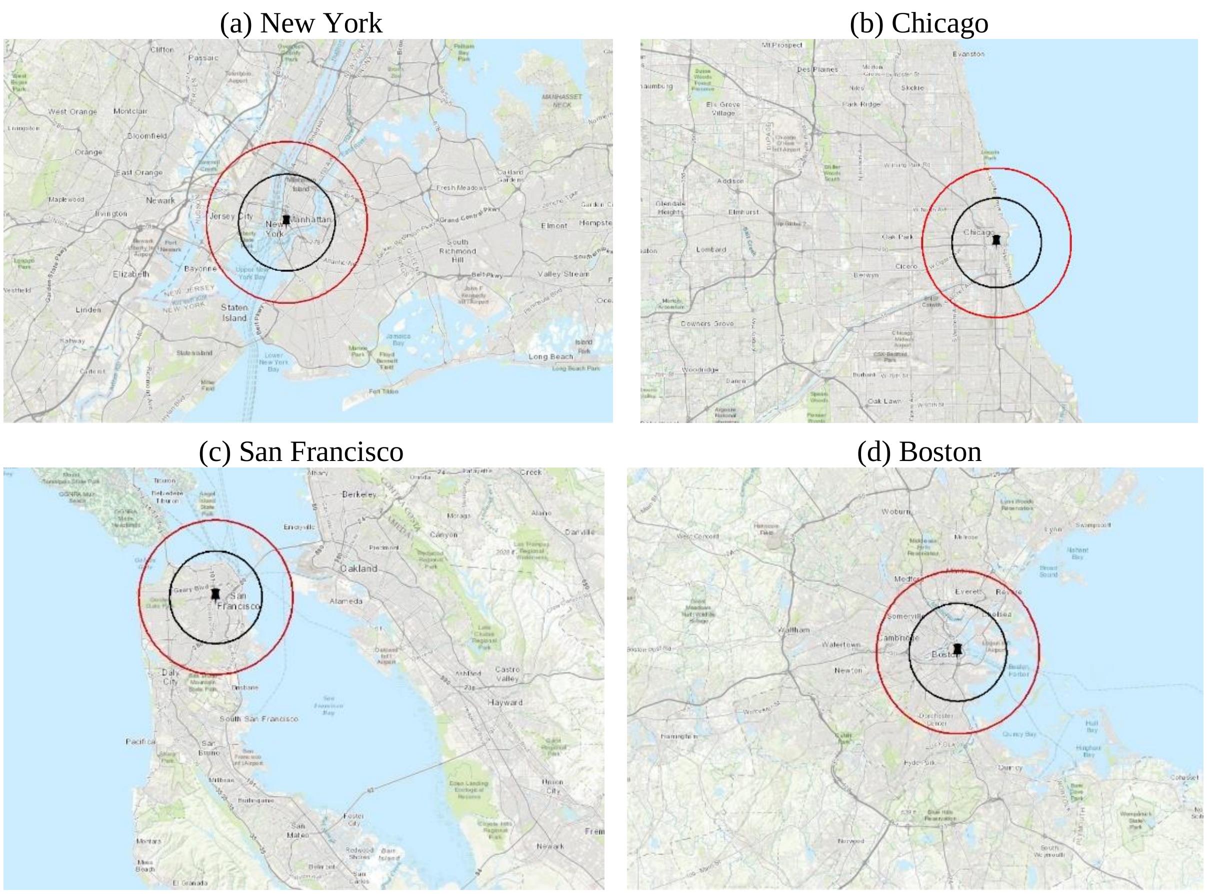
To check that the signs of parameters from the estimation results do not depend on the specification of the value of ϕ. I re-create the expected commute time matrices under various different specifications of ϕ, and redo the estimation under various specifications, which are shown in Table A7. I find that the scales of the parameter change, but the signs and comparative sizes of the parameters remain the same as in the main estimation, except the preference parameter for rents, which become insignificant under some values of ϕ.

I also re-create the expected commute time matrix using job locations in 2010 (rather than 1994) to determine whether the results are sensitive to such change. The results do not change by much.

Finally, I re-estimate the model using various definitions of high-skilled occupations (in the main estimation, occupations with more than 40% of college graduates in 1990 Census are designated as high-skilled). The results are largely similar.

Appendix: Figures and tables

Figure A1: Map of downtowns and the 3-mile and 5-mile ring in selected MSAs

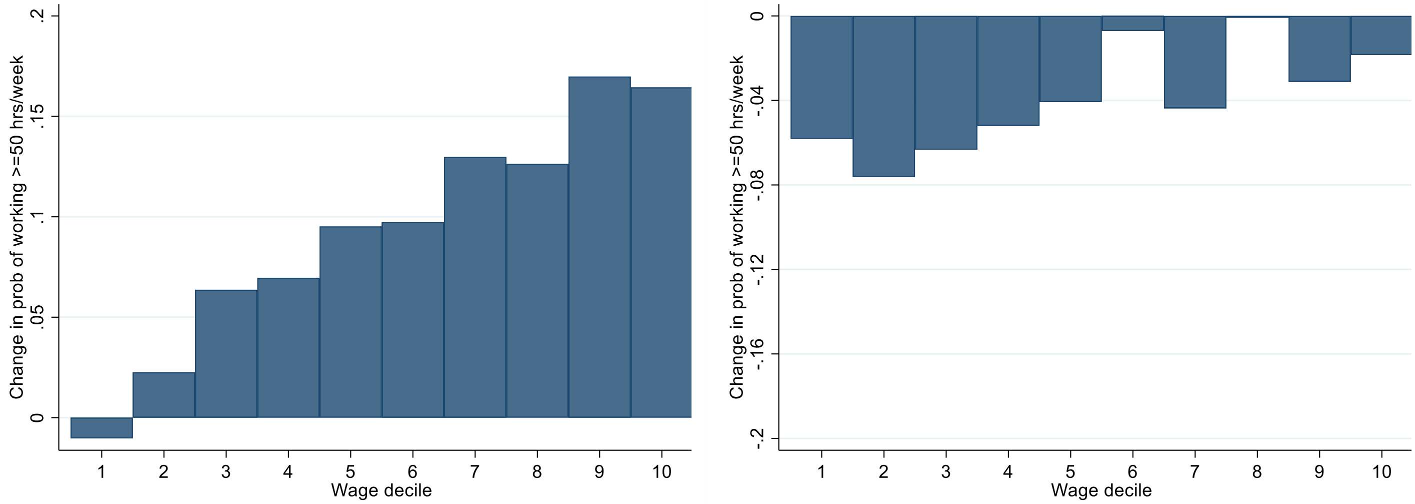


(a) New York
(b) Chicago
(c) San Francisco
(d) Boston

Notes: The longitudes and latitudes of the downtown pins are provided by Holian and Kahn (2015). The pins are the geographic location given by Google Map after searching for the respective cities. The smaller circles indicate the 3-mile (Euclidean distance) rings around the indicated downtown pins, and the larger circles indicate the 5-mile rings around the indicated downtown pins.

Figure A2: Changing probability of working long hours by wage decile (>=50 hours per week)

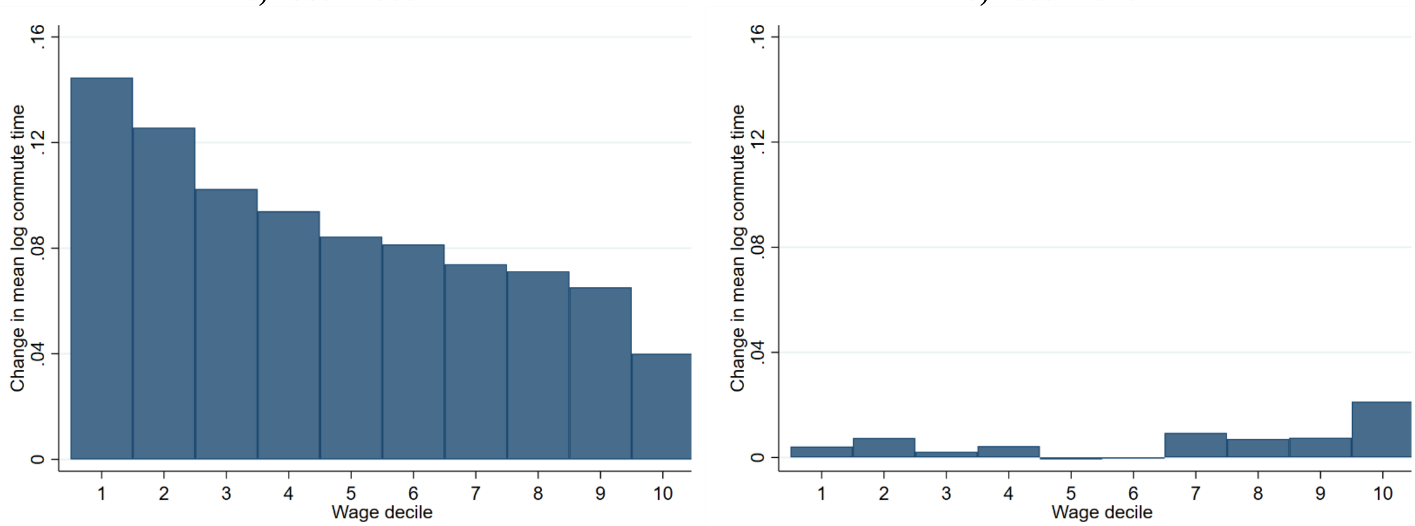
a) 1980 - 2000 b) 2000 - 2010



Notes: Data come from IPUMS census data in 1980, 2000, 2010 (2007-2011 ACS). To compute the probability of working at least 50 hours per week, the sample I use is workers that are between 25 and 65 of age, males, and working at least 30 hours per week. I include only male in the sample to ensure that the changing female labor force participation does not distortion the statistics. In a), I compute the change in probability of working long hours (>=50 hours per week) from 1980 to 2000. In b), I compute the change in probability of working long hours (>=50 hours per week) from 2000 to 2010.

Figure A3: Growth of commute time by wage decile

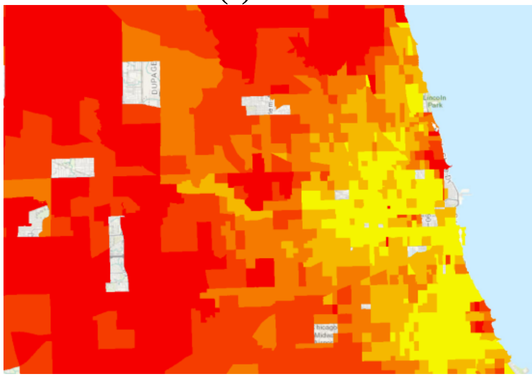
a) 1980 - 2000 b) 2000 - 2010



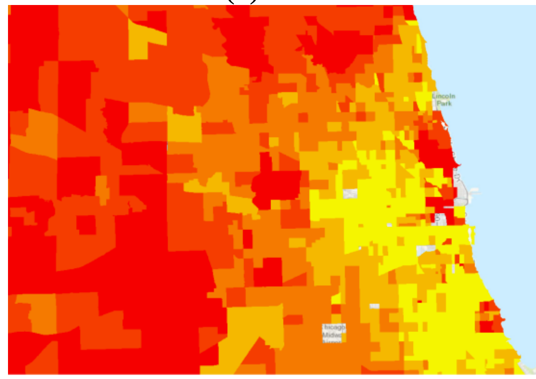
Notes: Data come from IPUMS census data in 1980, 2000, 2010 (2007-2011 ACS). I compute the change in log commute time reported in the Census/ACS data. The sample includes workers that are between 25 and 65 of age, males, working at least 30 hours per week and living in the most populous 25 MSAs in the US. In a), I plot the change in log commute time between year 2000 and 1980. In b), I plot the change in log commute time between year 2010 and 2000.

Figure A4: Income quintile by neighborhood within Chicago MSA (1980 – 2010)

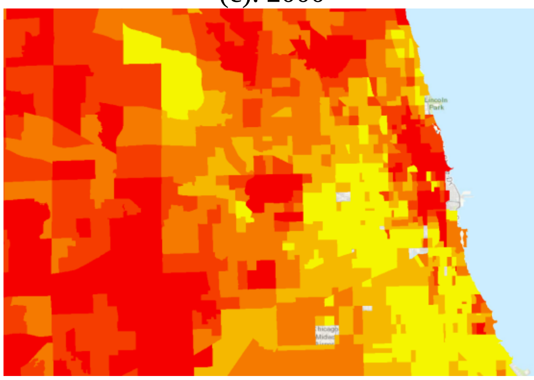
(a). 1980



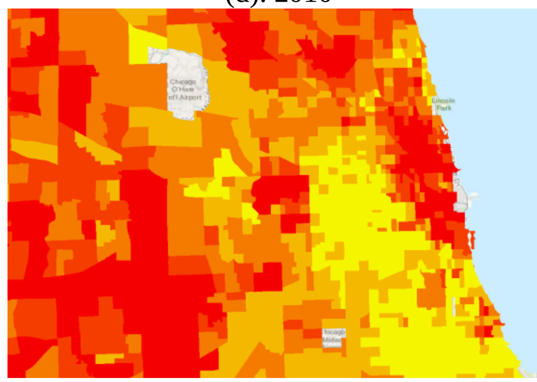
(b). 1990



(c). 2000

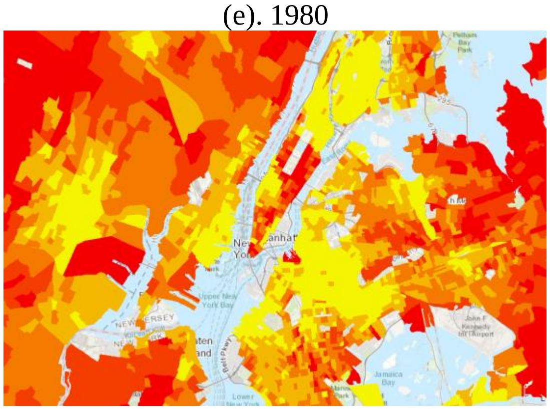


(d). 2010

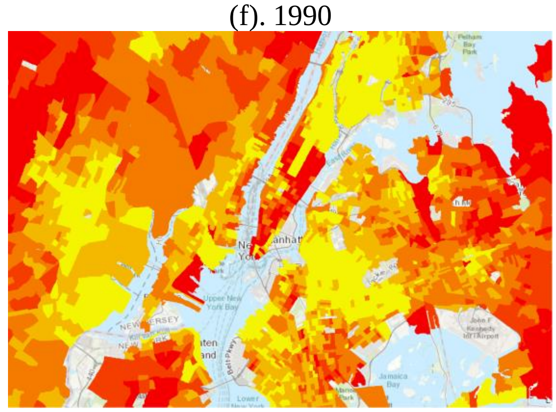


Income quintile by neighborhood within New York MSA (1980 – 2010)

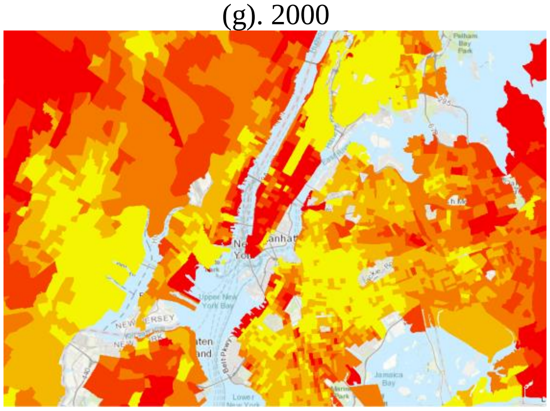
(e). 1980



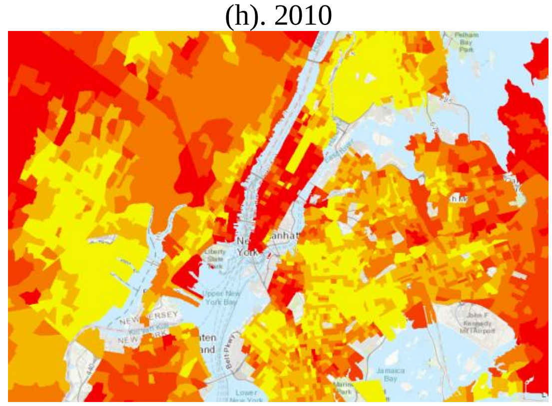
(f). 1990



(g). 2000

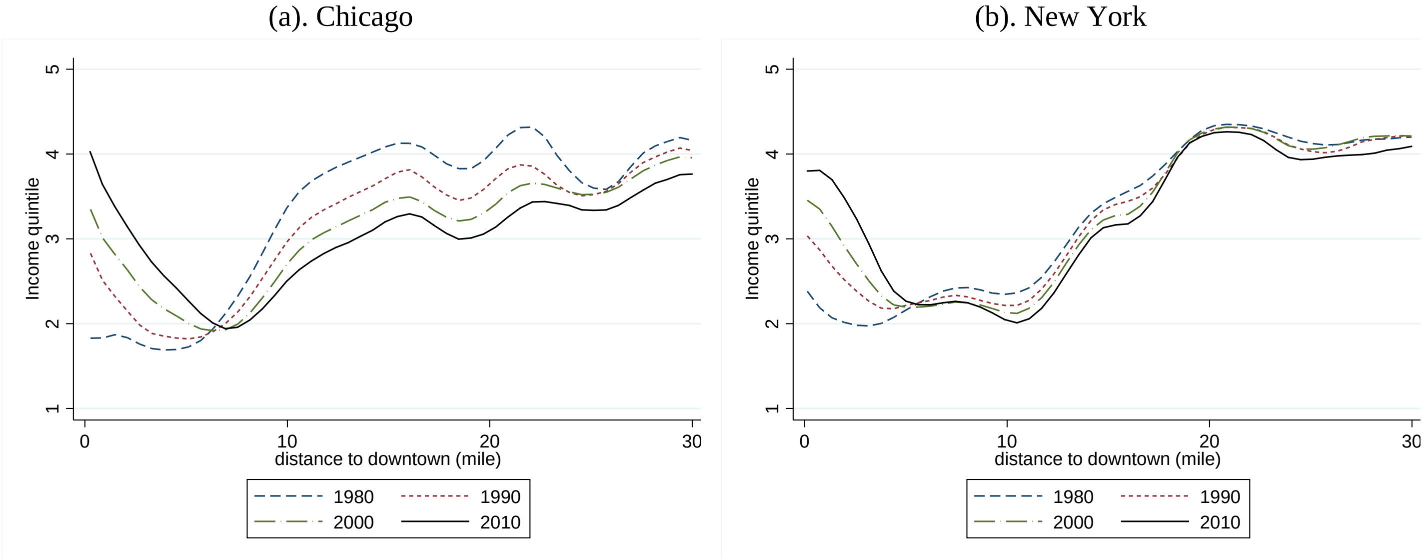


(h). 2010



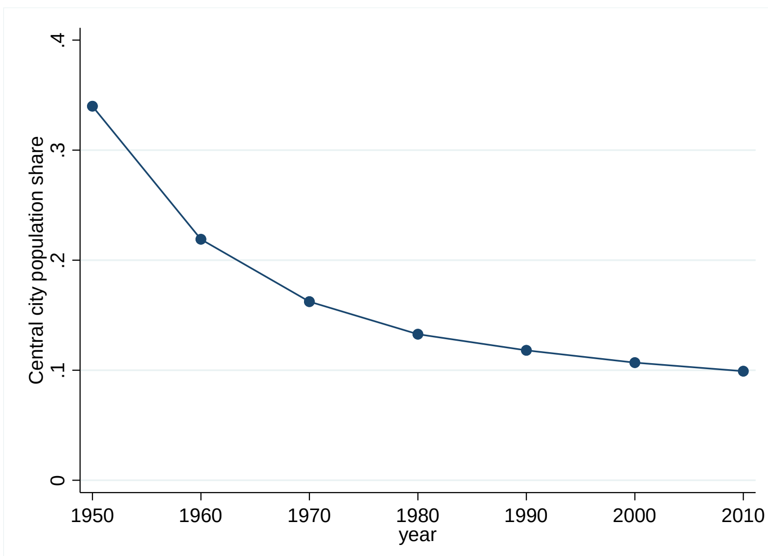
Notes: The plotted values are quintile ranking of census tract level income within the Chicago MSA and New York MSA respectively, from year 1990 to year 2010 using the Census summary statistics (NHGIS). The yellow color represents lowly ranked census tracts, and the darker red color represents more highly ranked census tracts in each contemporaneous year.

Figure A5: Income quintile by distance to downtown.



Notes: The plotted values are quintile ranking of census tract level income within the Chicago MSA and New York MSA respectively. I plot the census tract income ranking from year 1980 to year 2010 against the distance (in mile) to downtown. The plot is the kernel-weighted local polynomial smoothing curve, and Epanechnikov kernel function.

Figure A6: Central city population percentage among the largest 25 MSAs.



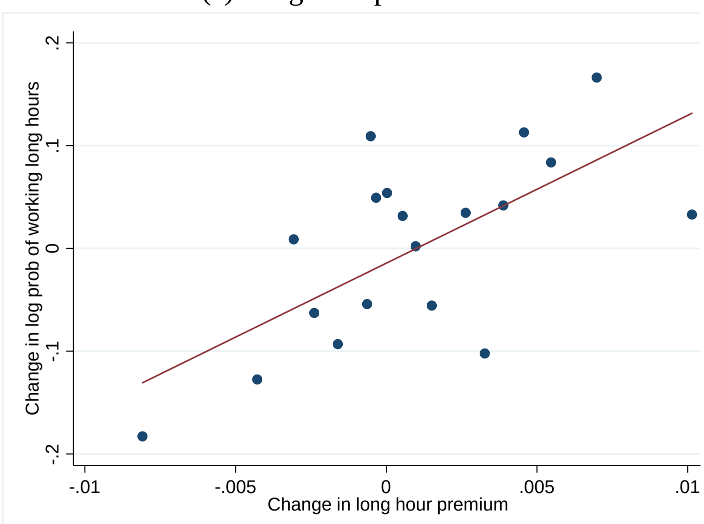
Notes: Central cities in this graph are defined as census tracts that are located within 5 miles of the downtown pin on Google in the respective MSAs. The value plotted in the graph are the population ratio between the population in the census tracts located in the central cities and the total population in the top 25 MSAs (defined by population ranking in 1990). The source of the data is Census and ACS provided by NHGIS.

Figure A7: Work and residential location in 1990

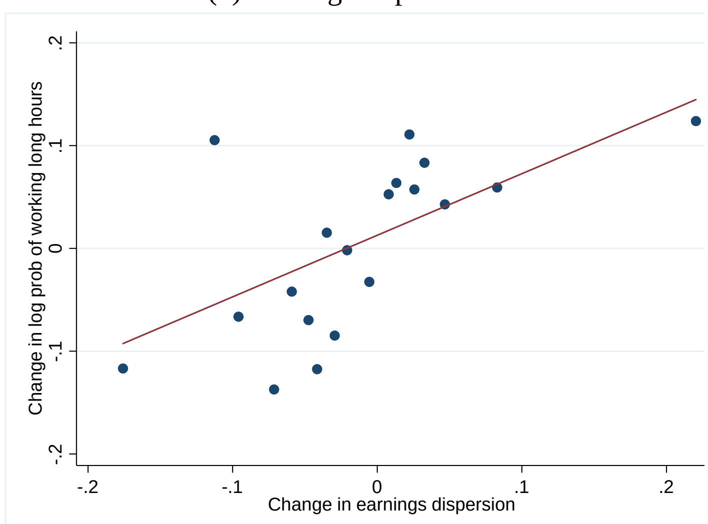
Notes: Residential location data come from both IPUMS and NHGIS Census data. Details are described in the data section. The employment data come from ZCBP at zip code level. Central cities are defined as census tracts and zip codes with centroids within 5 miles radius of the downtown pin. I use the sample from the largest 25 MSA to produce these graphs. The redline is the 45-degree line.

Figure A8: Change in the value of time and prevalence of working long hours (binscatter plots)

(a) Long-hour premium



(b) Earnings dispersion

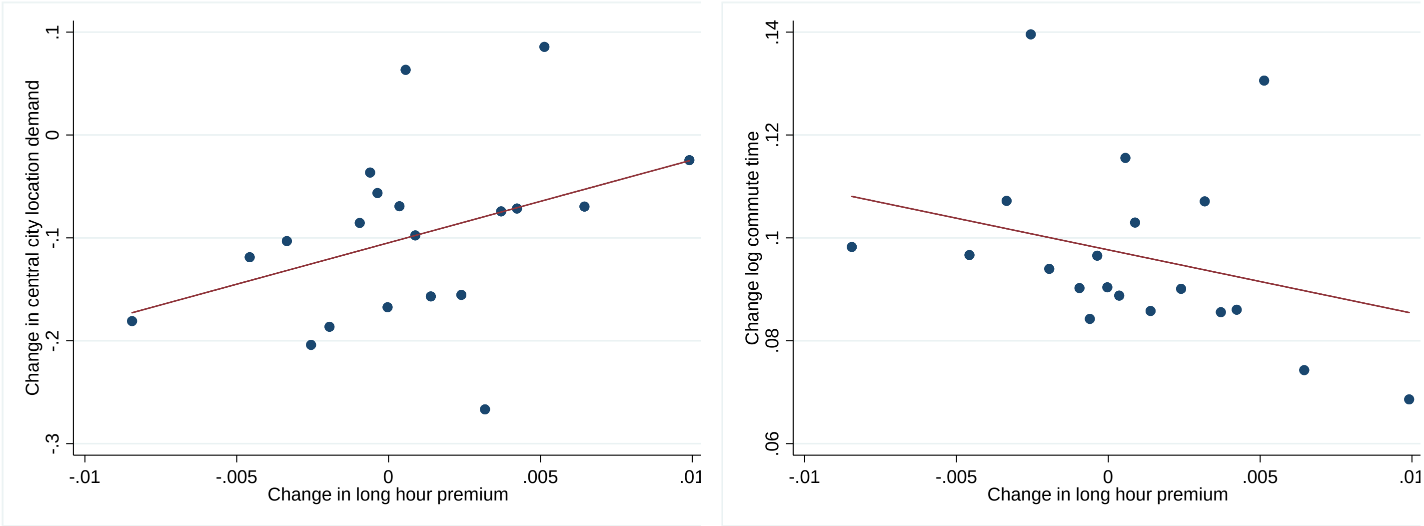


Notes: Prevalence of working long hours is computed with sample from male full-time workers between the age 25 and age 65 in the Census data. The first-difference is between 1990 and 2010. Data for year 2010 come from ACS 2007-2011. Each observation is an occupation-specific value. Binscatters are weighted by number of workers in each occupation in the Census data.

Figure A9: Change in long-hour premium, central city sorting and commute time (binscatter plots)

(a) Central city sorting

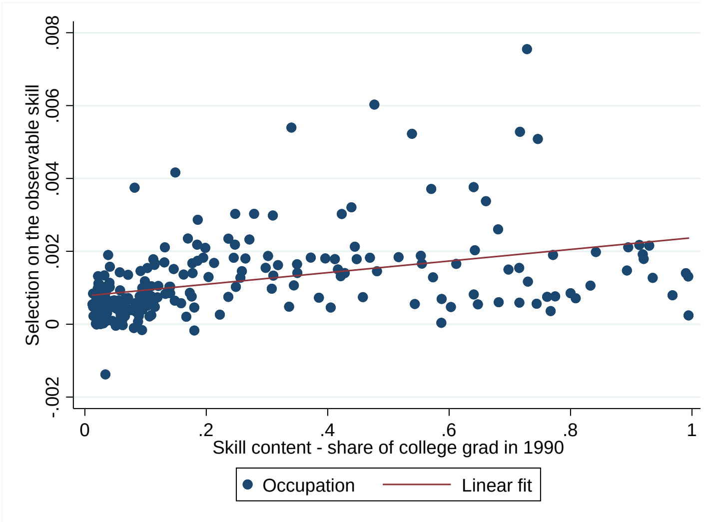
(b) Commute time



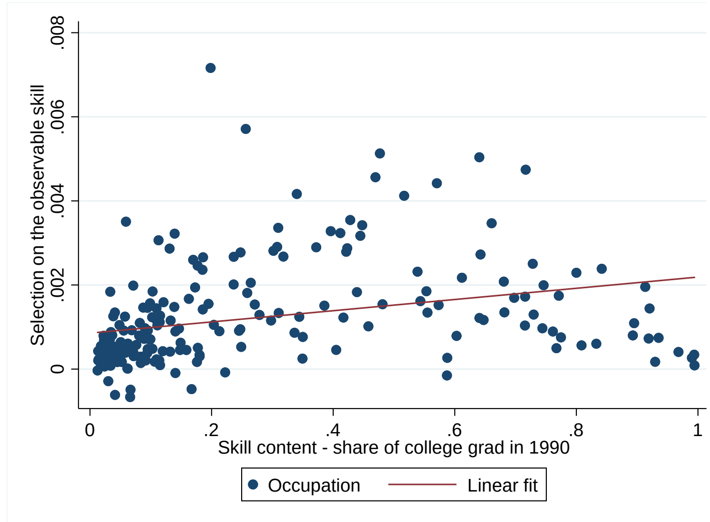
Notes: Each observation is a cell by occupation and MSA. I include MSA fixed-effects. The outcome variables (central city location choice and commute time) are constructed with sample from the most populous 25 MSAs. The first-difference is between 1990 and 2010. I use 2007-2011 ACS for year 2010. Binscatters are weighted by the occupation/MSA cell count.

Figure A10: Degree of selection for long-hour premium estimates on observable skills in 1990 and 2010

(a) Degree of selection in 1990

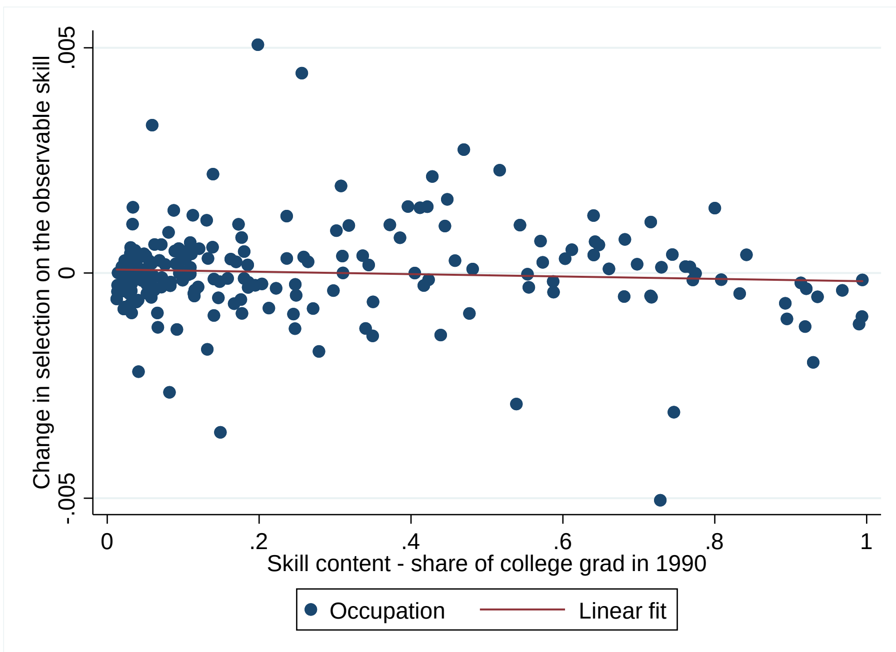


(b) Degree of selection in 2010



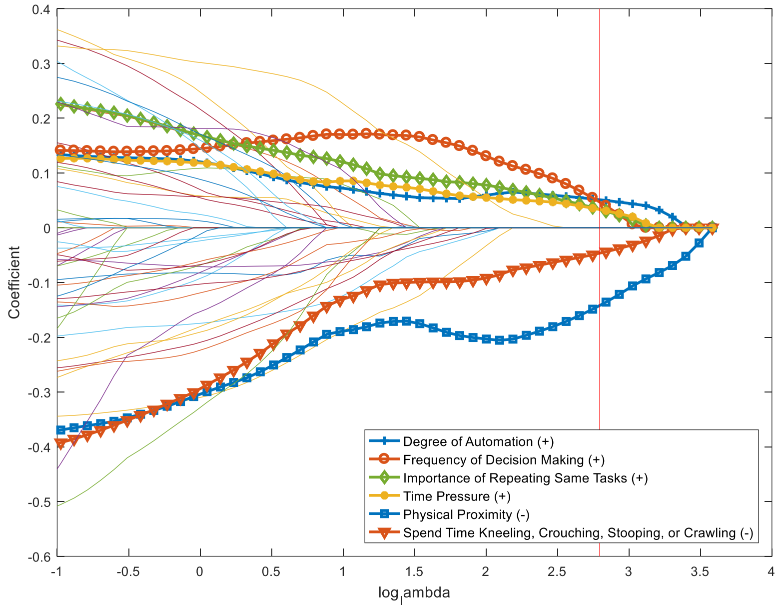
Notes: The y-axis is the difference between the estimates of long-hour premium without controlling for education levels and the estimates controlling for education levels. The difference between the two estimates indicates the degree of selection on the observable skill levels. X-axis is the skill content of each occupation, measured as the share of college graduates in 1990 Census.

Figure A11: Change in the degree of selection for long-hour premium estimates on observable skills



Notes: The y-axis is the difference between the change in long-hour premium estimated without education control and with education control. X-axis is the skill content of each occupation, measured as the share of college graduates in 1990 Census.

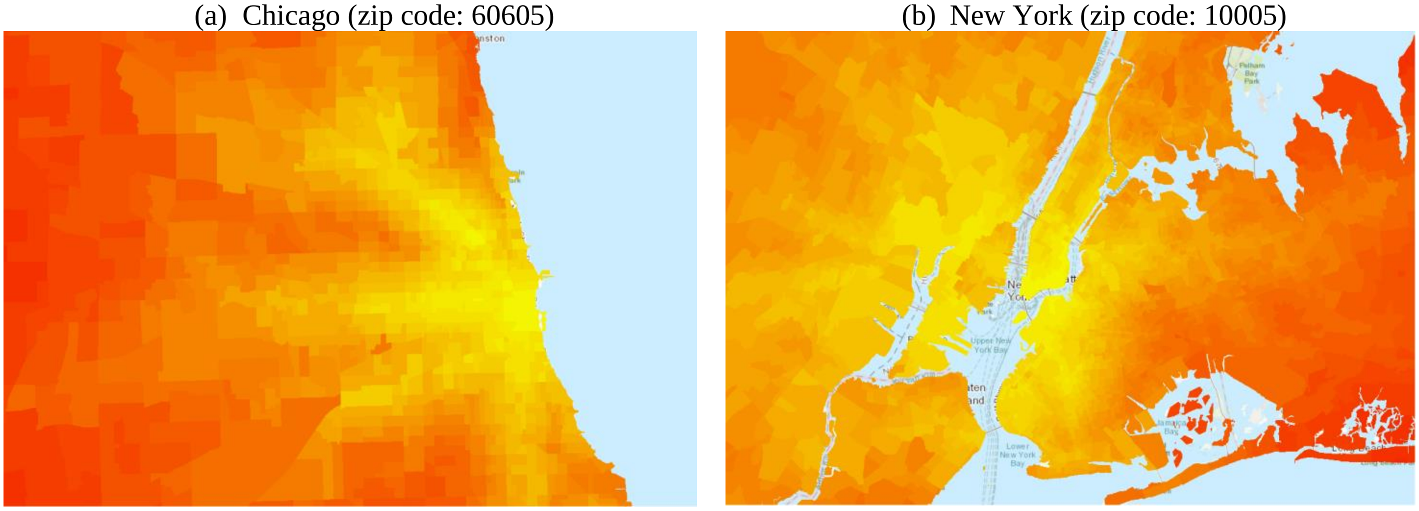
Figure A12: Lasso trace plot of the O*NET characteristics at predicting change in long-hour premium



Notes: I plot the coefficients on each of the 57 O*NET occupation characteristics for different levels of lambda (regularization penalty). The outcome variable is the change in long-hour premium. The red vertical line marks the lambda selected by 10-fold cross-validation. The characteristics that are non-zero at the red line are the non-redundant characteristics.

Figure A13: Imputed 1995 rush-hour driving time

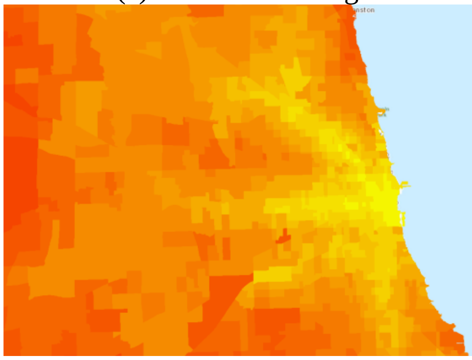
(a) Chicago (zip code: 60605) (b) New York (zip code: 10005)



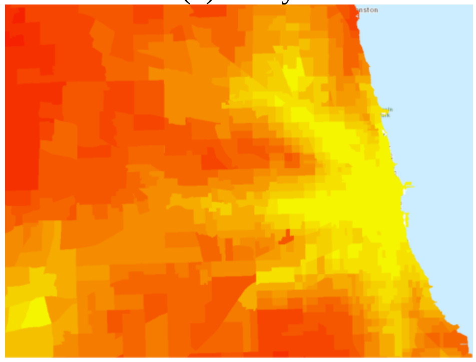
Notes: The above maps plot travel time from each census tract to the downtown of the MSA. I designate the destination for Chicago MSA as zip code 60605 (downtown Chicago) and destination for New York MSA as zip code 10005 (downtown Manhattan). The yellow color represents census tract with short travel time to the center of the city and red color represents long travel time. The maps are shown for the purpose of demonstration. To conduct the model estimation, I impute driving time to every zip code from every census tract in the non-rural counties of the US.

Figure A14: Expected commute time for selected occupations in Chicago MSA.

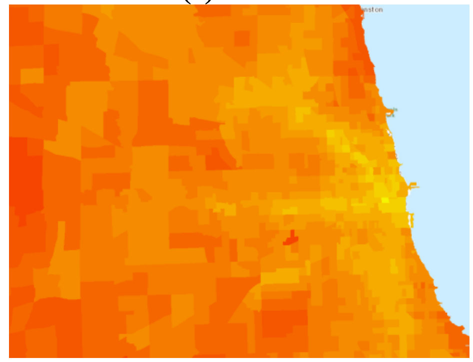
(a) Financial managers



(b) Lawyers

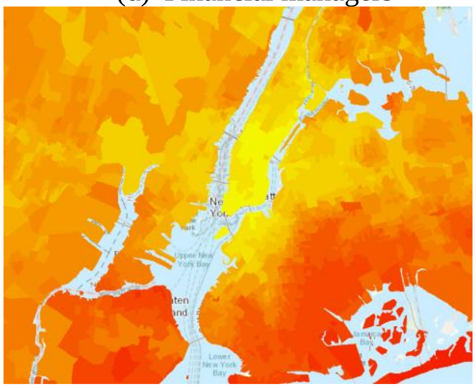


(c) Cashiers

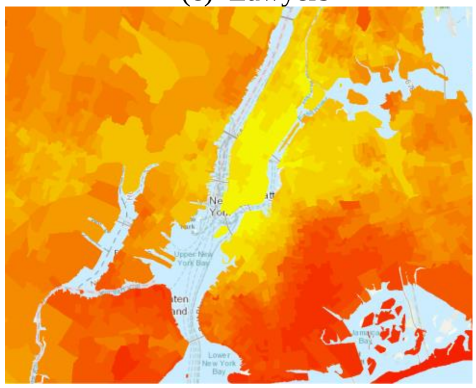


Expected commute time for selected occupations in New York MSA.

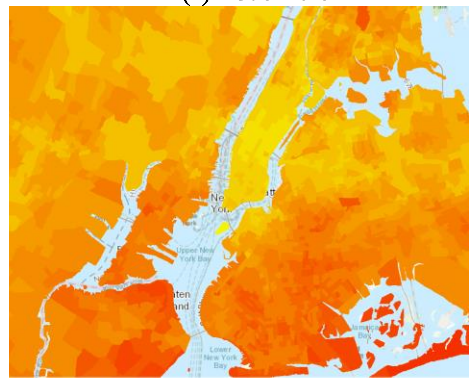
(d) Financial managers



(e) Lawyers



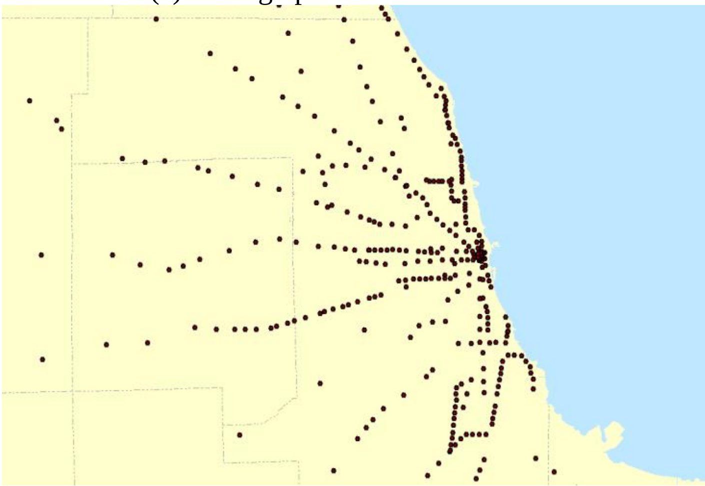
(f) Cashiers



Notes: The above maps are selected demonstrations of the expected commute time computed using employment allocation data (ZCBP data) and travel time matrix. The geographic unit displayed in the graphs is census tract. The color ranges from yellow to red. The yellow color represents short commute time, and red color represents long commute time. The color scale is consistent within respective MSA. The purpose of the maps is to show that the expected commute time by census tracts is quite different across different occupations, due to the differential allocation of job locations.

Figure A15: Public transit stations map for Chicago and New York

(a) Chicago public transit stations



(b) New York public transit stations

Notes: The dots plotted in the maps are public transit stations. The source of the data is Homeland Infrastructure Foundation-Level Data (HIFLD). https://hifld-geoplatform.opendata.arcgis.com/datasets/public-transit-stations/data?geometry=-93.974%2C41.551%2C-83.52%2C42.974

Table A1: First-stage regression between actual change in skill ratio and predicted change in skill ratio

	Dep variable: Δ actual change in log skill ratio	
	(1)	(2)
Δ predicted change in log skill ratio	1.215***	-
	(0.130)	
Δ log number of high-skilled workers	-	1.888***
		(0.170)
Δ log number of low-skilled workers	-	-0.609***
		(0.114)
MSA fixed-effects	Yes	Yes
Observations	42,346	42,346
R-squared	0.0791	0.0804

Notes: Results shown above are OLS regressions, with sample from the most populous 25 MSAs. The first-difference is between 1990 and 2010. I use 2007-2011 ACS for year 2010. Each observation is a cell by occupation and MSA. I include MSA fixed-effects. Regressions are weighted by the number of workers in each MSA/occupation cell included in the regression. A very few occupations are not present in the sample of every MSA in the IPUMS data, and thus those MSA/occupation cells are missing. There are two fewer cells in column (2) than in column (1) because workers in some cells do not report commute time in the Census data. Standard errors clustered at the MSA level. MSA fixed effects are used for all regressions.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A2: Estimates of alternative model specification with access to public mass transit stations

Panel A: Worker's residential location demand			Panel B: Rent (housing supply equation)	
Commute cost (μ)		3.024***		
		(0.4583)		
$\Delta\ v_{mkt}$ × 1{public transit < 1 mile}		1.4058***	Housing demand× housing stock density (π_1)	0.1664***
		(0.5451)		(0.0229)
Amenity (γ)	High-skilled occupations	1.8644***	Housing stock density (π_2)	0.0122***
		(0.2045)		(0.0044)
	Low-skilled occupations	0.4596***		
		(0.1918)		
Rent (β)		0.6534***		
		(0.2723)		

Notes: Model estimated using occupation/census tract cell data from 1990 to 2010. Number of cells used is 8,755,411. The number of workers in each occupation/MSA in 1990 is used as analytical weight. I control for total expected commute (using expected commute time to jobs unrelated to workers' occupations), and I allow the coefficients on total expected commute to vary by occupation. Conley (1999) HAC standard errors are computed with 1 miles threshold for the kernel function. Since value of time is estimating using national (minus MSA in question) data, I adjust the standard error for the generated regressor using method introduced in Murphy and Topel (1985). Same estimation procedure applies to this table as in Table 2, except that two additional terms are added in worker's location demand equation: $\Delta\ v_{mkt}$ × 1{public transit < 1 mile} and 1{public transit < 1 mile}.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A3: Estimates of model parameters – heterogeneous μ and β

Panel A: Worker's residential location demand			Panel B: Rent (housing supply equation)	
Commute cost (μ)	High-skilled occupations	6.3388*** (1.0316)		
	Low-skilled occupations	1.3245*** (0.4412)		
Amenity (γ)	High-skilled occupations	2.2361*** (0.1875)	Housing demand× housing stock density (π_1)	0.3410*** (0.0385)
	Low-skilled occupations	0.6910*** (0.1384)	Housing stock density (π_2)	-0.0002 (0.0058)
Rent (β)	High-skilled occupations	0.7100*** (0.2291)		
	Low-skilled occupations	0.3910*** (0.1825)		

Notes: Model estimated using occupation/census tract cell data from 1990 to 2010. Number of cells used is 8,755,411. The number of workers in each occupation/MSA in 1990 is used as analytical weight. I control for total expected commute (using expected commute time to jobs unrelated to workers' occupations), and I allow the coefficients on total expected commute to vary by occupation. Conley (1999) HAC standard errors are computed with 1-mile threshold for the kernel function. Since value of time is estimating using national (minus MSA in question) data, I adjust the standard error for the generated regressor using method introduced in Murphy and Topel (1985). Estimation detail can be found in the text.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A4: Robustness test: the effect of alternative measures of value of time on preferences for commute time

Alternative measures for the value of time	(1)	(2)
	Dependent variable: Δ ln (s_{jmkt})	
	Residual log earnings dispersion	Percentage of working long hour (weekly hours>=50)
Panel A: Unconditional Δ ln (s_{jmkt})		
μ_{alt}	1.455***	0.198***
	(0.0903)	(0.0149)
Panel B: Δ ln (s_{jmkt}) net of amenity and rent changes		
μ_{alt}	0.927***	0.149***
	(0.0984)	(0.0151)

Notes: Model estimated using occupation/census tract cell data of change from 1990 to 2010. Number of cells used is 8,755,411. The number of workers in each occupation/MSA in 1990 is used as analytical weight. For all estimations, I control for total expected commute (using expected commute time to jobs unrelated to workers' occupations), and I allow the coefficients on total expected commute to vary by occupation.

Table A5: Robustness test: association of O*NET occupation characteristics with the changes in preferences for commute time

O*NET characteristics (x_l):	Time pressure	Duration of work week	Level of competition	Degree of automation	Frequency of decision making	Importance of repeating the same tasks	Physical proximity	Spending time kneeling, crouching, stooping and crawling
Panel A: Unconditional estimates								
$\mu_{occ,l}$	0.0145***	0.027***	0.0024	0.0006	-0.0136***	0.0457***	-0.0911***	-0.0132***
	(0.0034)	(0.0044)	(0.0034)	(0.0027)	(0.0043)	(0.0045)	(0.0059)	(0.0051)
Panel B: Δ ln (s_{jmkt}) net of amenity and rent changes								
$\mu_{occ,l}$	0.0149***	0.0376***	0.0124***	0.0057**	-0.0211***	0.0421***	-0.0383***	0.0080
	(0.0035)	(0.0047)	(0.0038)	(0.0028)	(0.0047)	(0.0051)	(0.0065)	(0.0059)

Notes: Model estimated using occupation/census tract cell data of change from 1990 to 2010. Number of cells used is 8,083,276. The number of workers in each occupation/MSA in 1990 is used as analytical weight. For all estimations, I control for total expected commute (using expected commute time to jobs unrelated to workers' occupations), and I allow the coefficients on total expected commute to vary by occupation.

Table A6: Robustness test: effect of value of time on preferences for commute time using O*NET characteristics as IVs

	Dependent variable: Δ ln (s_{jmkt})
	IV estimates
Panel A: Unconditional Δ ln (s_{jmkt})	
μ_{IV}	12.975***
	(1.144)
Panel B: Δ ln (s_{jmkt}) net of amenity and rent changes	
μ_{IV}	7.507***
	(1.184)

Notes: I use GMM/IV estimator in this regression. The instruments are the O*NET characteristics (Degree of automation, frequency of decision making, importance of repealing same tasks, time pressure, physical proximity, spend time kneeling, crouching, stooping, or crawling) and the interaction terms between each of them and E(c_{jmk}). Model estimated using occupation/census tract cell data of change from 1990 to 2010. I use the Number of cells used is 8,083,276. The number of workers in each occupation/MSA in 1990 is used as analytical weight. For all estimations, I control for total expected commute (using expected commute time to jobs unrelated to workers' occupations), and I allow the coefficients on total expected commute to vary by occupation.

Table A7: Robustness test with alternative model specifications

		Alternative ϕ (travel matrix)			2010 job location	Alternative definition for high-skilled workers	
		ϕ =0.1	ϕ =0.2	ϕ =0.4		College>30%	College>50%
Panel A: Worker's location demand							
Commute cost (μ)		2.283***	2.589***	3.734***	3.141***	2.455***	3.0817***
		(0.29)	(0.334)	(0.398)	(0.412)	(0.464)	(0.474)
Amenity (γ)	High-skilled occupations	1.490***	1.658***	1.780***	2.116***	2.384***	2.696***
		(0.120)	(0.132)	(0.162)	(0.182)	(0.226)	(0.312)
	Low-skilled occupations	0.166***	0.301***	0.362***	0.619***	0.916***	1.214***
		(0.109)	(0.116)	(0.139)	(0.155)	(0.193)	(0.236)
Rent (β)		-0.0351	0.0515	0.266	0.751***	1.003***	0.401*
		(0.1503)	(0.158)	(0.181)	(0.207)	(0.257)	(0.235)
Panel B: Rent (housing supply equation)							
Housing demand× housing stock density (π_1)		0.41***	0.31***	0.32***	0.171***	0.185***	0.163***
		(0.059)	(0.0438)	(0.043)	(0.0294)	(0.031)	(0.028)
Housing stock density (π_2)		-0.032***	-0.009	-0.010	0.0127***	0.0078***	0.0117***
		(0.012)	(0.0086)	(0.0078)	(0.0046)	(0.0055)	(0.0049)

Notes: Each model specification is estimated using occupation/census tract cell data from 1990 to 2010. The number of workers in each occupation/MSA in 1990 is used as analytical weight. I control for total expected commute (using expected commute time to jobs unrelated to workers' occupations), and I allow the coefficients on total expected commute to vary by occupation. Conley (1999) HAC standard errors are computed with 1-mile threshold for the kernel function. Since value of time is estimating using national (minus MSA in question) data, I adjust the standard error for the generated regressor using method introduced in Murphy and Topel (1985). Column 1 – 3 show results from estimation with various different calibrations of ϕ in computing expected commute time. Column 4 uses 2010 Zip-Code Business Pattern data to compute the expected commute time. Column 5 – 6 show results from estimation using alternative definitions of high-skilled occupations (using 30% or 50% as thresholds rather than 40%).

Table A8: Estimate of long-hour premium with or without controls for education

Occupation name	code	Without educ. control: LRP - 1990	LRP - 2010	Δ in LRP	With educ. control: LRP - 1990	LRP - 2010	Δ in LRP
Managers in Marketing, Advertising, and Public Relations	30	0.0177	0.0232	0.0055	0.0164	0.0217	0.0053
Financial Managers	120	0.0218	0.0304	0.0085	0.0181	0.026	0.0078
Accountants and Auditors	800	0.0231	0.0306	0.0075	0.0198	0.0272	0.0074
Computer Scientists and Systems Analysts/Network systems Analysts/Web Developers	1000	0.0106	0.0168	0.0061	0.0100	0.0154	0.0054
Lawyers, and judges, magistrates, and other judicial workers	2100	0.0208	0.0254	0.0046	0.0194	0.0252	0.0058
Secondary School Teachers	2320	0.0067	0.0053	-0.0014	0.0048	0.0045	-0.0002
Securities, Commodities, and Financial Services Sales Agents	4820	0.0195	0.0405	0.0210	0.0158	0.0355	0.0197
Secretaries and Administrative Assistants	5700	0.0152	0.0152	-0.00	0.015	0.0144	-0.0006

Notes: The measurements are computed with microdata from Census IPUMS data. To compute the long-hour premium, I restrict the sample to workers between age of 25 and 65, and work at least 40 hours per week but does not work more than 60 hours per week. The results on the left are estimates without education controls, whereas the results on the right are estimates with education controls.

Table A9: List of occupations included and the long-hour premium and associated residual log hourly earnings dispersion

Occupation name	code	LRP - 1990	LRP - 2010	Δ in LRP	Earnings disp - 1990	Earnings disp - 2010	Δ in Earnings disp
Managers in Marketing, Advertising, and Public Relations	30	0.0164	0.0217	0.0053	0.5992	0.6550	0.0559
Financial Managers	120	0.0181	0.0260	0.0078	0.5492	0.6321	0.0829
Human Resources Managers	130	0.0149	0.0188	0.0039	0.5842	0.5913	0.0071
Purchasing Managers	150	0.0146	0.0193	0.0047	0.4589	0.4898	0.0309
Farmers, Ranchers, and Other Agricultural Managers	205	0.0084	0.0071	-0.0013	1.0270	1.1527	0.1257
Education Administrators	230	0.0127	0.0127	-0.0001	0.5663	0.5490	-0.0173

Food Service and Lodging Managers	310	0.0169	0.0134	-0.0034	0.7515	0.8118	0.0603
Medical and Health Services Managers	350	0.0184	0.0188	0.0004	0.5803	0.5959	0.0156
Property, Real Estate, and Community Association Managers	410	0.0166	0.0148	-0.0018	0.7931	0.8156	0.0225
Managers, nec (including Postmasters)	430	0.0156	0.0162	0.0006	0.6646	0.7125	0.0479
Wholesale and Retail Buyers, Except Farm Products	520	0.0210	0.0226	0.0016	0.6675	0.6944	0.0269
Purchasing Agents, Except Wholesale, Retail, and Farm Products	530	0.0142	0.0177	0.0036	0.4703	0.4961	0.0257
Claims Adjusters, Appraisers, Examiners, and Investigators	540	0.0130	0.0117	-0.0012	0.5013	0.4836	-0.0178
Compliance Officers, Except Agriculture	560	0.0132	0.0226	0.0094	0.5266	0.5275	0.0009
Human Resources, Training, and Labor Relations Specialists	620	0.0160	0.0244	0.0083	0.6157	0.6105	-0.0052
Management Analysts	710	0.0171	0.0186	0.0015	0.7794	0.7638	-0.0155
Other Business Operations and Management Specialists	730	0.0147	0.0236	0.0089	0.5149	0.6492	0.1343
Accountants and Auditors	800	0.0198	0.0272	0.0074	0.5897	0.6378	0.0481
Insurance Underwriters	860	0.0184	0.0267	0.0083	0.4606	0.4933	0.0327
Computer Scientists and Systems Analysts/Network systems Analysts/Web Developers	1000	0.0101	0.0154	0.0054	0.4501	0.5992	0.1492
Computer Programmers	1010	0.0095	0.0123	0.0028	0.5224	0.5531	0.0307
Operations Research Analysts	1220	0.0113	0.0120	0.0007	0.4723	0.4492	-0.0230
Architects, Except Naval	1300	0.0129	0.0179	0.0050	0.6649	0.6914	0.0265
Aerospace Engineers	1320	0.0129	0.0123	-0.0006	0.3911	0.4001	0.0091
Chemical Engineers	1350	0.0076	0.0105	0.0029	0.3935	0.4494	0.0559
Civil Engineers	1360	0.0127	0.0132	0.0005	0.4830	0.5091	0.0261
Electrical and Electronics Engineers	1410	0.0103	0.0098	-0.0005	0.4487	0.4707	0.0219
Industrial Engineers, including Health and Safety	1430	0.0125	0.0124	-0.0001	0.4092	0.4665	0.0572
Mechanical Engineers	1460	0.0136	0.0126	-0.0011	0.4163	0.4548	0.0385
Engineers, nec	1530	0.0130	0.0122	-0.0008	0.4643	0.4964	0.0321
Drafters	1540	0.0210	0.0159	-0.0052	0.5690	0.5688	-0.0002
Engineering Technicians, Except Drafters	1550	0.0165	0.0142	-0.0023	0.5551	0.5365	-0.0186
Surveying and Mapping Technicians	1560	0.0156	0.0142	-0.0014	0.6023	0.6497	0.0474
Biological Scientists	1610	0.0065	0.0069	0.0004	0.5050	0.5076	0.0025
Chemists and Materials Scientists	1720	0.0115	0.0126	0.0011	0.5104	0.4903	-0.0202
Environmental Scientists and Geoscientists	1740	0.0092	0.0117	0.0025	0.5301	0.5565	0.0263
Psychologists	1820	0.0194	0.0102	-0.0092	0.6446	0.6107	-0.0340
Chemical Technicians	1920	0.0161	0.0196	0.0035	0.5021	0.5045	0.0024
Life, Physical, and Social Science Technicians, nec	1960	0.0125	0.0120	-0.0005	0.5689	0.6881	0.1193
Counselors	2000	0.0089	0.0122	0.0033	0.5605	0.5889	0.0284
Social Workers	2010	0.0092	0.0085	-0.0006	0.5675	0.5061	-0.0615
Clergy	2040	0.0086	0.0051	-0.0035	0.5990	0.5755	-0.0235
Religious Workers, nec	2060	0.0116	0.0096	-0.0020	0.7105	0.7220	0.0115
Lawyers, and judges, magistrates, and other judicial workers	2100	0.0194	0.0252	0.0058	0.7271	0.7976	0.0705
Paralegals and Legal Assistants	2140	0.0206	0.0171	-0.0035	0.6320	0.6220	-0.0100
Postsecondary Teachers	2200	0.0131	0.0128	-0.0003	0.6060	0.6308	0.0249
Preschool and Kindergarten Teachers	2300	0.0103	0.0111	0.0008	0.7732	0.6404	-0.1328
Elementary and Middle School Teachers	2310	0.0078	0.0046	-0.0032	0.6128	0.5436	-0.0692
Secondary School Teachers	2320	0.0048	0.0045	-0.0002	0.5419	0.4840	-0.0579
Other Teachers and Instructors	2340	0.0107	0.0129	0.0022	0.7830	0.7659	-0.0171
Librarians	2430	0.0114	0.0101	-0.0013	0.5611	0.4591	-0.1020
Teacher Assistants	2540	0.0065	0.0155	0.0090	0.8637	0.6381	-0.2256
Artists and Related Workers	2600	0.0124	0.0133	0.0009	0.9209	1.0133	0.0924
Designers	2630	0.0169	0.0138	-0.0031	0.8058	0.8054	-0.0004
Actors, Producers, and Directors	2700	0.0194	0.0190	-0.0004	0.9462	0.9079	-0.0383
Athletes, Coaches, Umpires, and Related Workers	2720	0.0193	0.0153	-0.0040	0.9593	0.9820	0.0227
Musicians, Singers, and Related Workers	2750	0.0137	0.0136	-0.0001	0.9611	0.9970	0.0359
Editors, News Analysts, Reporters, and Correspondents	2810	0.0181	0.0178	-0.0003	0.6815	0.6862	0.0047
Public Relations Specialists	2825	0.0223	0.0220	-0.0003	0.6664	0.6704	0.0040
Technical Writers	2840	0.0107	0.0139	0.0032	0.5196	0.5493	0.0297
Writers and Authors	2850	0.0186	0.0162	-0.0024	0.9536	0.9783	0.0247
Photographers	2910	0.0152	0.0070	-0.0083	0.8685	1.0318	0.1633
Dentists	3010	0.0077	0.0084	0.0007	0.7503	0.7733	0.0230
Dieticians and Nutritionists	3030	0.0131	0.0111	-0.0020	0.5921	0.5651	-0.0270
Pharmacists	3050	0.0100	0.0071	-0.0029	0.5412	0.5940	0.0527
Physicians and Surgeons	3060	0.0119	0.0109	-0.0010	0.6860	0.7366	0.0505
Registered Nurses	3130	0.0166	0.0141	-0.0025	0.5643	0.5381	-0.0262
Physical Therapists	3160	0.0237	0.0122	-0.0115	0.6423	0.5473	-0.0949
Respiratory Therapists	3220	0.0130	0.0125	-0.0005	0.4853	0.4563	-0.0290
Speech Language Pathologists	3230	0.0089	0.0051	-0.0039	0.4882	0.4794	-0.0087
Therapists, nec	3240	0.0170	0.0078	-0.0091	0.6281	0.6111	-0.0169
Clinical Laboratory Technologists and Technicians	3300	0.0153	0.0109	-0.0045	0.5448	0.5764	0.0316
Dental Hygienists	3310	0.0236	0.0029	-0.0207	0.5799	0.5507	-0.0292
Health Diagnosing and Treating Practitioner Support Technicians	3410	0.0147	0.0160	0.0013	0.5275	0.5986	0.0711
Licensed Practical and Licensed Vocational Nurses	3500	0.0182	0.0140	-0.0042	0.6373	0.5790	-0.0583
Health Technologists and Technicians, nec	3530	0.0185	0.0198	0.0012	0.6345	0.6633	0.0288
Dental Assistants	3640	0.0097	0.0090	-0.0006	0.6714	0.6037	-0.0677
Medical Assistants and Other Healthcare Support Occupations, nec	3650	0.0130	0.0163	0.0033	0.7748	0.6165	-0.1583
First-Line Supervisors of Police and Detectives	3710	0.0032	0.0078	0.0046	0.3749	0.4493	0.0744
Firefighters	3740	0.0039	0.0074	0.0035	0.4299	0.4944	0.0645

Sheriffs, Bailiffs, Correctional Officers, and Jailers	3800	0.0108	0.0093	-0.0015	0.5114	0.5166	0.0053
Police Officers and Detectives	3820	0.0113	0.0131	0.0017	0.4333	0.4681	0.0348
Security Guards and Gaming Surveillance Officers	3930	0.0184	0.0186	0.0003	0.7448	0.7197	-0.0250
Crossing Guards	3940	0.0365	0.0126	-0.0238	0.8201	0.7972	-0.0230
Law enforcement workers, nec	3950	0.0119	0.0162	0.0043	0.7481	0.7448	-0.0033
Chefs and Cooks	4000	0.0185	0.0185	0.0001	0.8288	0.7588	-0.0700
First-Line Supervisors of Food Preparation and Serving Workers	4010	0.0208	0.0208	0.0001	0.7263	0.7169	-0.0093
Food Preparation Workers	4030	0.0169	0.0116	-0.0053	0.8651	0.7822	-0.0829
Bartenders	4040	0.0126	0.0064	-0.0061	0.7358	0.6780	-0.0577
Counter Attendant, Cafeteria, Food Concession, and Coffee Shop	4060	0.0127	0.0089	-0.0038	0.9357	0.9060	-0.0297
Waiters and Waitresses	4110	0.0105	0.0079	-0.0025	0.8117	0.7822	-0.0295
Food preparation and serving related workers, nec	4130	0.0123	0.0101	-0.0023	0.8885	0.7531	-0.1354
First-Line Supervisors of Housekeeping and Janitorial Workers	4200	0.0136	0.0175	0.0039	0.5861	0.6775	0.0914
First-Line Supervisors of Landscaping, Lawn Service, and Groundskeeping Workers	4210	0.0195	0.0145	-0.0049	0.6607	0.7929	0.1322
Janitors and Building Cleaners	4220	0.0148	0.0126	-0.0022	0.7912	0.7226	-0.0686
Maids and Housekeeping Cleaners	4230	0.0047	0.0063	0.0016	0.8021	0.7552	-0.0469
Grounds Maintenance Workers	4250	0.0185	0.0147	-0.0038	0.8693	0.8563	-0.0130
First-Line Supervisors of Personal Service Workers	4320	0.0123	0.0087	-0.0036	0.7598	0.8326	0.0729
Nonfarm Animal Caretakers	4350	0.0122	0.0062	-0.0060	0.8443	0.8647	0.0204
Entertainment Attendants and Related Workers, nec	4430	0.0070	0.0135	0.0065	0.8259	0.8544	0.0285
Barbers	4500	0.0171	0.0102	-0.0069	0.7763	0.8134	0.0371
Hairdressers, Hairstylists, and Cosmetologists	4510	0.0193	0.0115	-0.0079	0.8156	0.8066	-0.0090
Childcare Workers	4600	0.0211	0.0132	-0.0079	1.0126	0.9613	-0.0513
Recreation and Fitness Workers	4620	0.0069	0.0071	0.0003	0.8252	0.7988	-0.0264
First-Line Supervisors of Sales Workers	4700	0.0147	0.0188	0.0040	0.6990	0.7456	0.0466
Cashiers	4720	0.0170	0.0163	-0.0006	0.8906	0.8425	-0.0481
Counter and Rental Clerks	4740	0.0160	0.0123	-0.0036	0.8446	0.7902	-0.0544
Parts Salespersons	4750	0.0122	0.0161	0.0039	0.5964	0.5629	-0.0335
Retail Salespersons	4760	0.0176	0.0186	0.0010	0.8094	0.8383	0.0289
Advertising Sales Agents	4800	0.0181	0.0244	0.0063	0.7962	0.8024	0.0062
Insurance Sales Agents	4810	0.0131	0.0189	0.0058	0.7523	0.8567	0.1043
Securities, Commodities, and Financial Services Sales Agents	4820	0.0158	0.0355	0.0197	0.8390	0.9476	0.1086
Sales Representatives, Services, All Other	4840	0.0174	0.0215	0.0041	0.7575	0.8069	0.0494
Sales Representatives, Wholesale and Manufacturing	4850	0.0144	0.0181	0.0037	0.6812	0.7116	0.0304
Models, Demonstrators, and Product Promoters	4900	0.0246	0.0041	-0.0205	1.1193	0.9631	-0.1563
Door-to-Door Sales Workers, News and Street Vendors, and Related Workers	4950	0.0150	0.0103	-0.0047	0.9567	1.0600	0.1033
Sales and Related Workers, All Other	4965	0.0161	0.0210	0.0050	0.8826	0.7361	-0.1464
First-Line Supervisors of Office and Administrative Support Workers	5000	0.0134	0.0185	0.0050	0.5162	0.5671	0.0509
Telephone Operators	5020	0.0162	0.0114	-0.0048	0.6405	0.6335	-0.0070
Bill and Account Collectors	5100	0.0170	0.0168	-0.0002	0.5962	0.6113	0.0151
Billing and Posting Clerks	5110	0.0151	0.0212	0.0061	0.5824	0.5308	-0.0516
Bookkeeping, Accounting, and Auditing Clerks	5120	0.0144	0.0176	0.0032	0.6428	0.5918	-0.0510
Payroll and Timekeeping Clerks	5140	0.0147	0.0138	-0.0010	0.5435	0.5081	-0.0354
Bank Tellers	5160	0.0177	0.0169	-0.0009	0.6490	0.5798	-0.0692
File Clerks	5260	0.0134	0.0208	0.0074	0.8376	0.7000	-0.1376
Hotel, Motel, and Resort Desk Clerks	5300	0.0155	0.0119	-0.0036	0.7328	0.7264	-0.0065
Interviewers, Except Eligibility and Loan	5310	0.0231	0.0263	0.0032	0.7927	0.7959	0.0032
Library Assistants, Clerical	5320	0.0118	-0.0060	-0.0177	0.6894	0.5876	-0.1018
Loan Interviewers and Clerks	5330	0.0155	0.0211	0.0056	0.5862	0.6110	0.0248
Correspondent clerks and order clerks	5350	0.0081	0.0167	0.0085	0.6100	0.6639	0.0539
Human Resources Assistants, Except Payroll and Timekeeping	5360	0.0215	0.0242	0.0027	0.5378	0.5813	0.0434
Receptionists and Information Clerks	5400	0.0143	0.0138	-0.0004	0.7687	0.7029	-0.0658
Reservation and Transportation Ticket Agents and Travel Clerks	5410	0.0137	0.0295	0.0158	0.6707	0.6554	-0.0153
Information and Record Clerks, All Other	5420	0.0175	0.0128	-0.0047	0.7371	0.6198	-0.1173
Couriers and Messengers	5510	0.0174	0.0189	0.0015	0.7354	0.7285	-0.0069
Dispatchers	5520	0.0131	0.0138	0.0007	0.5772	0.5747	-0.0025
Postal Service Clerks	5540	0.0104	0.0148	0.0045	0.5351	0.4676	-0.0675
Postal Service Mail Carriers	5550	0.0090	0.0137	0.0047	0.4418	0.4065	-0.0353
Production, Planning, and Expediting Clerks	5600	0.0185	0.0161	-0.0024	0.6153	0.5582	-0.0571
Shipping, Receiving, and Traffic Clerks	5610	0.0191	0.0176	-0.0015	0.6030	0.6173	0.0143
Stock Clerks and Order Fillers	5620	0.0155	0.0158	0.0003	0.7187	0.7519	0.0332
Weighers, Measurers, Checkers, and Samplers, Recordkeeping	5630	0.0197	0.0044	-0.0153	0.7341	0.6572	-0.0769
Secretaries and Administrative Assistants	5700	0.0150	0.0144	-0.0006	0.6354	0.5930	-0.0424
Computer Operators	5800	0.0190	0.0235	0.0045	0.5896	0.6069	0.0172
Data Entry Keyers	5810	0.0139	0.0224	0.0085	0.6526	0.6726	0.0199
Word Processors and Typists	5820	0.0235	0.0129	-0.0106	0.7070	0.6357	-0.0713
Mail Clerks and Mail Machine Operators, Except Postal Service	5850	0.0183	0.0234	0.0051	0.7562	0.6608	-0.0954
Office Clerks, General	5860	0.0169	0.0183	0.0014	0.7258	0.6729	-0.0529
Office Machine Operators, Except Computer	5900	0.0236	0.0332	0.0096	0.6777	0.6436	-0.0341
Office and administrative support workers, nec	5940	0.0160	0.0203	0.0043	0.6318	0.6346	0.0028
Agricultural workers, nec	6050	0.0142	0.0137	-0.0005	0.8886	0.7933	-0.0954
First-Line Supervisors of Construction Trades and Extraction Workers	6200	0.0128	0.0133	0.0005	0.6936	0.7418	0.0481
Brickmasons, Blockmasons, and Stonemasons	6220	0.0116	0.0100	-0.0016	0.7574	0.8133	0.0559

Carpenters	6230	0.0112	0.0068	-0.0043	0.7925	0.8327	0.0403
Carpet, Floor, and Tile Installers and Finishers	6240	0.0179	0.0108	-0.0071	0.8545	0.8690	0.0145
Cement Masons, Concrete Finishers, and Terrazzo Workers	6250	0.0033	0.0092	0.0059	0.7944	0.7786	-0.0158
Construction Laborers	6260	0.0188	0.0127	-0.0061	0.8776	0.8918	0.0142
Construction equipment operators except paving, surfacing, and tamping equipment operators	6320	0.0116	0.0091	-0.0025	0.6076	0.6690	0.0614
Drywall Installers, Ceiling Tile Installers, and Tapers	6330	0.0105	0.0051	-0.0055	0.7965	0.8227	0.0262
Electricians	6355	0.0130	0.0107	-0.0023	0.5866	0.6703	0.0837
Painters, Construction and Maintenance	6420	0.0169	0.0114	-0.0054	0.8842	0.8975	0.0133
Pipelayers, Plumbers, Pipefitters, and Steamfitters	6440	0.0073	0.0104	0.0031	0.6738	0.6968	0.0230
Roofers	6515	0.0152	0.0099	-0.0053	0.8914	0.9077	0.0163
Sheet Metal Workers, metal-working	6520	0.0130	0.0116	-0.0014	0.5752	0.6677	0.0926
Structural Iron and Steel Workers	6530	0.0064	0.0107	0.0042	0.6222	0.7322	0.1100
Helpers, Construction Trades	6600	0.0126	0.0114	-0.0011	0.9033	0.8613	-0.0420
Construction and Building Inspectors	6660	0.0083	0.0039	-0.0044	0.5031	0.5935	0.0904
First-Line Supervisors of Mechanics, Installers, and Repairers	7000	0.0115	0.0133	0.0018	0.4692	0.5058	0.0366
Computer, Automated Teller, and Office Machine Repairers	7010	0.0110	0.0125	0.0015	0.5377	0.6742	0.1365
Radio and Telecommunications Equipment Installers and Repairers	7020	0.0153	0.0125	-0.0028	0.4656	0.5510	0.0854
Aircraft Mechanics and Service Technicians	7140	0.0101	0.0089	-0.0012	0.5247	0.4916	-0.0331
Automotive Body and Related Repairers	7150	0.0100	0.0102	0.0003	0.7530	0.7501	-0.0029
Automotive Service Technicians and Mechanics	7200	0.0124	0.0119	-0.0005	0.7149	0.7262	0.0113
Bus and Truck Mechanics and Diesel Engine Specialists	7210	0.0118	0.0123	0.0005	0.5657	0.5589	-0.0068
Heavy Vehicle and Mobile Equipment Service Technicians and Mechanics	7220	0.0153	0.0128	-0.0025	0.6049	0.5897	-0.0152
Heating, Air Conditioning, and Refrigeration Mechanics and Installers	7315	0.0115	0.0088	-0.0027	0.6128	0.6794	0.0666
Industrial and Refractory Machinery Mechanics	7330	0.0173	0.0172	-0.0001	0.5266	0.5376	0.0110
Maintenance and Repair Workers, General	7340	0.0130	0.0152	0.0022	0.6240	0.5835	-0.0405
First-Line Supervisors of Production and Operating Workers	7700	0.0146	0.0157	0.0011	0.5060	0.5504	0.0444
Electrical, Electronics, and Electromechanical Assemblers	7720	0.0190	0.0212	0.0022	0.6710	0.6538	-0.0171
Assemblers and Fabricators, nec	7750	0.0171	0.0175	0.0004	0.7312	0.7135	-0.0177
Bakers	7800	0.0117	0.0136	0.0019	0.7285	0.7299	0.0014
Butchers and Other Meat, Poultry, and Fish Processing Workers	7810	0.0095	0.0191	0.0096	0.6554	0.6806	0.0252
Cutting, Punching, and Press Machine Setters, Operators, and Tenders, Metal and Plastic	7950	0.0179	0.0209	0.0030	0.6022	0.6758	0.0736
Machinists	8030	0.0191	0.0214	0.0023	0.5190	0.5551	0.0361
Tool and Die Makers	8130	0.0201	0.0188	-0.0013	0.4501	0.4680	0.0179
Welding, Soldering, and Brazing Workers	8140	0.0138	0.0171	0.0034	0.6522	0.6782	0.0260
Metal workers and plastic workers, nec	8220	0.0186	0.0209	0.0024	0.6189	0.6042	-0.0147
Bookbinders, Printing Machine Operators, and Job Printers	8230	0.0188	0.0144	-0.0045	0.6471	0.6149	-0.0323
Laundry and Dry-Cleaning Workers	8300	0.0162	0.0114	-0.0047	0.7909	0.7653	-0.0255
Sewing Machine Operators	8320	0.0138	0.0047	-0.0091	0.7212	0.6928	-0.0284
Tailors, Dressmakers, and Sewers	8350	0.0089	0.0044	-0.0046	0.7878	0.7692	-0.0186
Cabinetmakers and Bench Carpenters	8500	0.0129	0.0175	0.0046	0.7046	0.7263	0.0217
Stationary Engineers and Boiler Operators	8610	0.0133	0.0136	0.0003	0.5111	0.5014	-0.0096
Crushing, Grinding, Polishing, Mixing, and Blending Workers	8650	0.0229	0.0203	-0.0027	0.6061	0.6452	0.0391
Cutting Workers	8710	0.0189	0.0250	0.0061	0.7055	0.6515	-0.0540
Inspectors, Testers, Sorters, Samplers, and Weighers	8740	0.0133	0.0196	0.0063	0.6364	0.6475	0.0111
Medical, Dental, and Ophthalmic Laboratory Technicians	8760	0.0225	0.0186	-0.0038	0.6714	0.6275	-0.0439
Packaging and Filling Machine Operators and Tenders	8800	0.0123	0.0162	0.0039	0.7999	0.7525	-0.0474
Painting Workers and Dyers	8810	0.0149	0.0209	0.0061	0.6951	0.7312	0.0361
Photographic Process Workers and Processing Machine Operators	8830	0.0180	0.0196	0.0016	0.7614	0.6799	-0.0815
Other production workers including semiconductor processors and cooling and freezing equipment operators	8965	0.0182	0.0194	0.0011	0.6645	0.6924	0.0278
Supervisors of Transportation and Material Moving Workers	9000	0.0126	0.0148	0.0021	0.5437	0.5599	0.0162
Aircraft Pilots and Flight Engineers	9030	0.0060	0.0029	-0.0031	0.6009	0.5976	-0.0034
Flight Attendants and Transportation Workers and Attendants	9050	0.0076	0.0106	0.0030	0.6493	0.6064	-0.0429
Bus and Ambulance Drivers and Attendants	9100	0.0128	0.0120	-0.0008	0.6477	0.5953	-0.0524
Driver/Sales Workers and Truck Drivers	9130	0.0161	0.0160	-0.0001	0.7214	0.7067	-0.0147
Taxi Drivers and Chauffeurs	9140	0.0115	0.0070	-0.0045	0.8029	0.7732	-0.0297
Parking Lot Attendants	9350	0.0246	0.0088	-0.0158	0.8102	0.7369	-0.0733
Crane and Tower Operators	9510	0.0177	0.0153	-0.0024	0.5274	0.6335	0.1061
Industrial Truck and Tractor Operators	9600	0.0122	0.0151	0.0029	0.6227	0.6573	0.0346
Cleaners of Vehicles and Equipment	9610	0.0164	0.0149	-0.0015	0.8892	0.8991	0.0099
Laborers and Freight, Stock, and Material Movers, Hand	9620	0.0133	0.0163	0.0030	0.8231	0.8178	-0.0053
Packers and Packagers, Hand	9640	0.0139	0.0113	-0.0026	0.8429	0.8158	-0.0271

Notes: The measurements are computed with microdata from Census IPUMS data. To compute the long-hour premium, I restrict the sample to workers between age of 25 and 65, and work at least 40 hours per week but does not work more than 60 hours per week. To compute the residual log earnings dispersion, I regress log earnings on individual characteristics (age, sex, race, education, industry code), and compute the standard deviation of the residual log earning. The residual earnings dispersion is computed as the standard deviation of the residual log earnings.