Federal Reserve Bank of Dallas

The Rising Value of Time and the Origin of Urban Gentrification

Yichen Su

Working Paper 1913

Research Department

https://doi.org/10.24149/wp1913

The Rising Value of Time and the Origin of Urban Gentrification[*]

Yichen Su[†]

October 31, 2019

Abstract

In recent decades, gentrification has transformed American central city neighborhoods. I estimate a spatial equilibrium model to show that the rising value of high-skilled workers' time contributes to 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 the value of time has a modest direct effect on gentrification of central cities, the effect is substantially magnified by endogenous amenity improvement driven by the changes in local skill mix.

Keywords: Gentrification, Value of time, Amenities, Central cities

JEL codes: J22, R13, J24

[*] I thank my advisors Luigi Pistaferri, Caroline Hoxby, and Rebecca Diamond for their guidance and support. I also thank Ran Abramitzky, Yiwei Chen, Jessie Handbury, Eran Hoffmann, Sitian Liu, Isaac Sorkin, Yu Zheng, and participants at Stanford Labor and Public workshop, Stanford Applied Economics workshop, Stata Texas conference, UEA conference, AREUEA conference, System Applied-Micro conference at SF Fed. 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.

[†] Yichen Su, Federal Reserve Bank of Dallas, Yichen.Su@dal.frb.org.

1 Introduction

Since the 1990s, American cities have seen a wave of urban revival during which the growth of income and home value in 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 central cities and "flight" of affluent residents to the suburbs pre-1990s (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 (2019)). 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 this paper, I show that high-skilled workers' rising value of time is an exogenous force that has contributed to gentrification of central city neighborhoods. The rising value of time among high-skilled workers makes central cities more attractive to them due to shorter commuting time to work. As high-skilled "pioneers" move into the central cities, amenity conditions endogenously improve, and rents increase due to increased demand for central city housing. The improved amenities make central cities increasingly attractive to high-skilled workers, despite rising rents. On the other hand, low-skilled workers, while facing the same rising rents, demand these amenities much less than high-skilled workers do. As a result, low-skilled workers increasingly relocate to the suburbs while high-skilled workers increasingly sort into central city neighborhoods. I show that while the rising value of time among high-skilled workers has a modest direct effect on the gentrification of central cities, the resulting endogenous amenity change substantially amplifies its direct effect.

I motivate my analysis by documenting 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. Interestingly, 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 that the spatial relocation of high-skilled workers into the central cities is likely related to their changing value of time and increasing desire for shorter commute time.

To evaluate how value of time, commuting and amenities affect neighborhood sorting, I present and estimate a spatial equilibrium model of neighborhood choice. In my model, I allow workers to choose which neighborhood to live in based on their value of time, the commute time to their jobs, local amenities and rents. My model allows the changing value of time to exogenously affect workers' demand for locations with shorter commuting time. My model also allows local amenities and local rents to change endogenously as the local population mix changes. In the model, the mechanism of workers' spatial sorting is governed by how much workers' value of time, neighborhood amenities and rents each affect their demand for locations.

I estimate workers' location demand using a novel empirical strategy. The first and the most important parameter in location demand is the effect of value of time on the demand for shorter commute. The size of this parameter determines the direct effect of value of time on gentrification. To identify this parameter, I exploit the fact that job locations in different occupations are distributed differentially across space. And if rising value of time makes workers want to move closer to work, I should see them move toward their *occupation-specific* job locations, holding all else equal. By observing differential migration patterns by occupation and observing how much workers in occupations with increasing value of time migrate closer to occupation-specific job locations, I can identify how value of time affects workers' demand for shorter commuting time. To think about the strategy more intuitively, consider financial workers and physicians in the New York MSA. Financial jobs are very concentrated in downtown Manhattan while clinics and hospitals are spread throughout the metropolitan area. Therefore, to test whether financial workers and physicians migrate to reduce commuting time, I need to observe them *differentially* sort into locations closer to their respective job locations, controlling for other neighborhood characteristics that are occupation-invariant.

The other important parameter to identify is how much workers prefer local amenities. This parameter governs how much the endogenously changing amenities could amplify the exogenous neighborhood change induced by the changing value of time. To identify this parameter, I use the idea that locations of jobs that are unrelated to a worker may indirectly affect that worker's migration choice by changing local amenity levels through influencing *other* workers' migration choices. Continuing with the same examples of financial workers and physicians, downtown Manhattan has high concentration of financial firms but has less concentration of clinics and hospitals. The rising value of time of financial workers would induce inflow of high-skilled financial workers and thus higher levels of amenities. If I observe that physicians increasingly migrate into downtown Manhattan, even though physicians do not typically *work* there, such patterns would reveal their preference for amenities.

To implement this empirical strategy, which exploits variation in value of time by occupation, I first measure workers' value of time by estimating the "long-hour premium" for each detailed occupation, using repeated cross-sections from the Census data in 1990 and 2010. Using the differential changes in long-hour premiums in different occupations, I examine how much value of time affects workers' migration choice regarding commute time.

I measure the distance to job locations in terms of an "expected commute time", which is commute time weighted by the spatial distribution of jobs. To measure commute time, I use a travel time matrix (by driving) generated by Google Distance Matrix API and National Household Travel Survey data. This allows me to compute commuting time between all neighborhoods in all

MSAs in the U.S.. I then combine the travel time matrix with data on occupation-specific job location to measure the expected commute time for each residential neighborhood. Variation in expected commute time by occupation and neighborhood is a crucial ingredient for implementing the empirical strategy.

Using the estimated model, I show that a little more than 10% of gentrification of central cities is driven by the direct effect of the shock to the value of time, holding amenities and rents constant as the initial levels. I further show that additional 40% of gentrification of central cities driven by the indirect effect of endogenous amenity change and rent change. This means that the rising value of time is likely a contributing force behind gentrification, but its effect is greatly magnified by the effects of endogenous amenity improvement. The results also suggest that the changing value of time and the endogenous amenity change have limited ability to explain the full magnitude of central city gentrification. Other factors unrelated to the mechanisms described in this paper are likely to also have played a crucial role in gentrifying central cities.

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 pin 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 (2019) 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 (2019) 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 (2019) examine the role of crime reduction, which is another important exogenous force that generates inflow in high-skilled residents.

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 (2019), Davis, Dingel, Monras, and Morales (2019), 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 (2013) 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 the city level in Diamond's paper whereas they are modeled at the neighborhood level in this paper.

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 (2018)). My paper extends the analysis by investigating how the opportunity cost of time affects location choice and the housing market. My paper is particularly linked to and dependent on the work by Kuhn and Lozano (2008) who document the changing working-hour pattern among high and low-income workers in the U.S..

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 presents the conclusion.

2 Data

The main datasets I use are the 1990 U.S. Decennial Census data and the 2007-2011 American Community Survey (ACS). 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 at the census tract level for 1950, 1960, 1970, 1980, 1990, 2000, and also of the ACS for 2007-2011. This dataset enables me to analyze post-war trends of suburbanization and subsequent gentrification at the census tract level. NHGIS data also enable me to track workers' occupation affiliations at the census tract level, which I use to construct location choice probabilities for each census tract by occupation.[1]

I use Zip Code Business Patterns (ZCBP) data provided by the U.S. Census Bureau to measure the spatial distribution of jobs for each occupation in 1994 and 2010.[2] 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 within any given MSA. First, I use Google API to compute travel time[3] and travel distance from

[1] 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.

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

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

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).[4]

3 Descriptive patterns

To motivate the linkage between gentrification, rising value of time and amenity change, I document a few stylized facts that describe the gentrification patterns and time-use patterns observed since 1990.

3.1 Gentrification

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 (Holian and Kahn (2015))) 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.[5] In Figure 2, I plot the census tract level change in log skill ratio between 1990 and 2010 by distance to downtown (skill ratio is defined as ratio between number of residents of high-skilled occupations and residents of low-skilled occupations. High-skilled occupations are defined as occupations with $\geq 40\%$ of college graduates in the 1990 Census). The plot shows a dramatic change in presence of high-skilled residents near downtown locations.

High-skilled residents are increasingly living in central city neighborhoods, even though the locations of high-skilled jobs have not been centralizing. Figure 3a is a binscatter plot between the share of residents living in central city neighborhoods in 1990 and 2010. The plot shows that residential concentration in central cities rose significantly for workers in high-skilled occupations, while the residential concentration generally declined for low-skilled occupations. However, the binscatter plot in Figure 3b shows that the 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

[4] 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.

[5] 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.

neighborhoods is unlikely to be driven by concurrent sorting of jobs.[6]

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 increased the appeal of central city locations for high-skilled workers. Diamond (2016) and Couture and Handbury (2019) 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 increased 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 columns (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 skill ratio tend to also experience stronger growth in the abundance of 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.

3.2 Prevalence of working long hours

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. Interestingly the change in central cities around 1990-2010 is accompanied by the reversal of work-hour patterns in both the high-wage population and low-wage population. 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 5a 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.

In addition to using Census/ACS data, I use the CPS 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 4, I plot the probability of working long hours for workers in the top and bottom wage deciles respectively. Consistent with the Census data, low-

[6] 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).

[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.

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 the 1980s. But between the mid-1980 and late 1990, high-wage workers' probability of working long hours increased dramatically.

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, suggests that the rising cost of time among high-skilled workers could 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 5b that while commute time in all wage groups has increased between 1980 and 2010, the growth in higher wage groups is considerably smaller. This suggests that a substantial portion of higher-wage workers have re-optimized their locations in favor of shorter commute time.[8]

3.3 Correlation between long hours and central city location choice

To see whether workers who become more prone to work long hours are more likely to move into the central cities, I first conduct a simple regression analysis. I cut worker samples into detailed occupations, and calculate the percentages of workers working long hours in 1990 and 2010. Then, I examine whether workers in occupations with rising share of long-hour workers are increasingly likely to live in central cities.

Table 2 columns 1-3 show the results of the occupation-MSA level first-difference regressions. I regress the change in log share of residents living in central cities on change in percentage of those working long hours. The results show that workers in increasingly long-hour occupations are more likely to live in the central cities. In column 4-6, I show coefficients of regressions in which I use the change in log commute time (observed directly in the Census/ACS data) by occupation. The results show the workers in increasingly long-hour occupations tend to report shorter commuting time, although the elasticities are much smaller.

While these results suggest that increasing long-hour work may have contributed to some of the spatial sorting, there are two reasons that it is difficult to interpret the implication of these coefficients clearly or in a causal way. First, workers increasingly working long hours are disproportionately high-skilled. If high-skilled workers have time-varying taste for central city amenities, it could spuriously drive the coefficients. Second, if amenities are endogenous to the in-flow of high-skilled residents, the initial sorting of high-skilled could endogenously lead to further sorting

[8] 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 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.

of high-skilled residents. It is difficult to disentangle that with the simple regressions above.

For these reasons, I next perform an analysis using a spatial equilibrium model at census tract level to *separately* identify the effect of the rising value of time and endogenous amenity change on location choice. I then use the model to dissect how much gentrification of central cities is driven directly by the changing value of time and how much by the endogenous change in amenities.

4 Spatial equilibrium model of residential choice

To unpack the mechanism of gentrification, I use 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. Instead of modeling locations as central cities or suburbs, I treat each census tract as a distinct location in the model.

4.1 Worker's problem

Given the worker's occupation k and city m where she lives and works, a worker who chooses to live in neighborhood j and works in neighborhood n in year t enjoys utility:[9]

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

subject to budget constraint

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

C is consumption; H is the housing service; A_{jmt} is the amenity level for neighborhood j at time t;[10] $\tilde{\gamma}_k$ is the taste parameter for local amenities, which may differ by worker type; c_{jnmt} is the weekly commute time between residential location j and work location n. $\tilde{\omega}_t$ is a time-variant aversion parameter for commute time. $\varepsilon_{i,jmt}$ is the idiosyncratic preference for individual i, distributed as Type I Extreme Value, and σ is 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.

[9] 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.

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

4.1.1 Long-hour premium

Worker's weekly log earnings is a linear function of y_{0kt}, which is the basic log income the worker would receive if she were to supply only the minimum 40 hours of work or less. $v_{kt}\,(T - c_{jnmt})$ is the extra log income she receives if she works more than 40 hours a week; v_{kt} 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. The negative impact of commute time c_{jnmt} on log earnings (or utility) is larger if long-hour premium v_{kt} is larger.[11] I use the long-hour premium to approximate workers' value of time.

This way of measuring value of time differs from the traditional way of using hourly earnings or wage. Using wage as the value of time may work for workers who are paid by the hour. 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)).

4.1.2 Location demand

Each worker solves the utility maximization problem by choosing C and H conditional on her occupation and the locations she lives and works in. Derivation of the indirect utility is detailed in Appendix A1. I normalize the indirect utility function by σ, the standard deviation of the idiosyncratic preference. The indirect utility becomes

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

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)).

[11] 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.

After derivation written in detail in Appendix A2, I write the log location choice probability as a linear function of various location preference components.

$$\log\left(s_{jmkt}\right) = \underbrace{\tilde{\delta}_{mkt}}_{\text{fixed effects}} + \underbrace{\log\left(\sum_{n'\in J_m} \pi_{n'mkt} \exp\left(-\left(\omega_t + \mu v_{kt}\right)\cdot c_{jn'mt}\right)\right)}_{\text{valuation of proximity to employment}} \tag{2}$$

$$- \underbrace{\beta r_{jmt}}_{\text{valuation of rent}} + \underbrace{\gamma_k a_{jmt}}_{\text{valuation of amenities}} + \underbrace{\gamma_k \zeta_{jmt}}_{\text{valuation of unobserved amenity}}$$

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 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 $\log\left(s_{jmkt}\right)$ with respect to the value of time:

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

where $\widetilde{\mathrm{E}}_t\left(c_{jmkt}\right) = \sum_{n\in J_m} \tilde{\pi}_{jnmkt} c_{jnmt}$, and $\tilde{\pi}_{jn'mkt}$ is an adjusted probability measure as: $\tilde{\pi}_{jnmkt} = \frac{\pi_{nmkt}\exp(-(\omega_t+\mu v_{kt})c_{jnmt})}{\sum_{n'\in J_m}\pi_{n'mkt}\exp\left(-(\omega_t+\mu v_{kt})c_{jn'mt}\right)}$. π_{nmkt} is the probability that a job in occupation k in MSA m is located in location n.

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 the fact that workers are less likely to work at locations too far away from home. μ governs how much value of time affects workers' sensitivity to commuting time. Therefore, estimating μ is an essential part of the paper, because the size of it determines the effect of the value of time on gentrification.

4.1.3 Endogenous amenity supply

I assume that the level of amenities can respond to the ratio of local high-skilled and low-skilled residents, similar to the method used by Diamond (2016).[12] Under this assumption, a rising share of high-skilled residents in a neighborhood would lead to the entries of suppliers of local goods and

[12] Some amenities 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)).

services and better funding for local public goods, such as effective local law enforcement.[13] I model amenity supply as follows:

$$a_{jmt} = \eta \ln \left(\frac{N^H_{jmt}}{N^L_{jmt}} \right) + \tilde{\theta}_t X_{jmt} + \delta_{jm} + \delta_{mt} + \xi^a_{jmt} \tag{4}$$

N^H_{jmt} and N^L_{jmt} 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^a_{jmt} represents the component of amenity supply that is unobservable and cannot be accounted for by the local skill ratio.[14]

Since a key driver of amenity supply is the ratio of high- to low-skilled residents, I endogenize amenity levels into workers' location demand by directly modeling location demand as an iso-elastic function of local skill ratios, governed by a reduced-form migration elasticity γ_k. Ideally, I would like to model neighborhood amenity directly. However, neighborhood amenities are multi-dimensional, and it is unclear how to aggregate various amenity variables. Local skill ratio captures the content of amenity that is driven by the changing local population.[15] Instead of modeling amenities directly into the equilibrium framework, I create measurements of crime and consumption amenities later in the paper,[16] 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:

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

The reduced-form migration elasticity γ_k is a combination of demand side elasticity and supply side elasticity, namely $\gamma_k \eta$; this is a sufficient statistic that can pin down the mechanism of

[13] 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 (2019) 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.

[14] 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.

[15] 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. 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.

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

the endogenous amenity change. $\theta_{kt}X_{jmt}$ is the component of amenities that is observable and exogenous.[17] $\gamma_k\xi^a_{jmt}$ 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}.

4.2 Housing supply

I assume that log rent is a reduced-form function of local demand for housing, and its interaction with the existing housing stock density (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 there is 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.

4.3 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 census tract:

1. **Amenity market clears in each census tract:** - The number of high- and low-skilled workers living in a census tract is determined by workers' location choice. 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.

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

5 Estimation

5.1 The long-hour premium

I first estimate long-hour premium for each occupation. I take the exact labor earnings function introduced in the model to the Census microdata to estimate the 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 earnings the worker would earn if she worked 40 hours/week. v_{kt} is the long-hour premium to be estimated. $\boldsymbol{\delta}_{demo,it}$ is the vector of demographic dummies. u_{ikt} is the error term. 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 the workers are paid a constant proportion of the hours worked. If v_{kt} rises over time, it means that workers are increasingly paid disproportionately more than before.

I estimate the long-hour premium v_{kt} using cross-sectional data on log earnings and hours within each occupation, controlling for individual workers' characteristics. Since hours worked is a labor supply choice variable, estimates of the long-hour premium may be driven by a selection effect related to unobserved worker ability. I describe in detail how I address endogeneity concerns in Appendix section D1.

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 6 the plots between residual log weekly earnings and hours worked for four occupations. For financial workers and lawyers, the slope rises dramatically from 1990 to 2010. In contrast, for office secretaries and teachers, the slope of the plot remains largely unchanged, despite increases in average hourly earnings over time. The variation in change in long-hour premiums (slopes of the curves) will be used in later analysis. The computed long-hour premiums are shown in Table A3 in the appendix.

Appendix section D3 and D4 talks about various validation tests performed for the long-hour premium as a measurement of the value of time.

5.2 Reduced-form relation between long-hour premium, long-hour work and central city location choice

To further validate the long-hour premium measurements, I regress change in probability of working long hours on change in long-hour premium. Table 3 shows the results, which shows that workers in occupations with rising long-hour premiums tend to increasingly work long hours. Table 3 also shows that workers in occupations with rising long-hour premiums are increasingly likely to live in the central cities. The change in log reported commute time is also negatively correlated with long-hour premium. Similar to the results in Table 2, the magnitude between the commuting time elasticities is smaller than central city sorting elasticities. Next, I describe my empirical strategy to identify the model parameters using location choice data and long-hour premiums.

5.3 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 also allow μ, β to differ by skill - μ_z, β_z.

Note that in specification (5), μ_z 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_z. Also, since I only have a static travel time matrix, the commute time is set to be time-invariant. Derivation of the linear approximation is included in Appendix D5. The following is the linearized location demand:

$$\log\left(s_{jmkt}\right) = \delta_{jmk} + \tilde{\delta}_{mkt} - \left(\phi + \omega_t\right)\widetilde{\mathrm{E}}_t\left(c_{jmk}\right) - \mu_z v_{kt}\widetilde{\mathrm{E}}_t\left(c_{jmk}\right) - \beta_z r_{jmt} + \gamma_z \log\left(\frac{N^H_{jmt}}{N^L_{jmt}}\right) + \theta_{kt}X_{jmt} + \xi_{jmkt} \tag{8}$$

As a result of the Taylor approximation, $\widetilde{\mathrm{E}}_t\left(c_{jmk}\right)$ is evaluated with a transformed probability measure, $\tilde{\pi}_{jnmk,t} = \frac{\pi_{nmkt}\exp(-\phi c_{jnm})}{\sum_{n' \in J_m}\pi_{n'mkt}\exp\left(-\phi c_{jn'm}\right)}$, where I calibrate ϕ such that the mean commute time matches the value reported in the 1990 Census data.[18] π_{nmkt} is the job distribution of occupation k at time t. $\widetilde{\mathrm{E}}_t\left(c_{jmk}\right)$ is the expected commute time weighted by job distribution of time t.

After linearization, μ_z 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.

$$\begin{aligned} \Delta\log\left(s_{jmkt}\right) \quad = \quad & \Delta\tilde{\delta}'_{mkt} - \Delta\omega_t\widetilde{\mathrm{E}}_{t-1}\left(c_{jmk}\right) - \mu_z\Delta\hat{v}_{kt}\widetilde{\mathrm{E}}_{t-1}\left(c_{jmk}\right) - \beta_z\Delta r_{jmt} \\ & + \gamma_z\Delta\log\left(\frac{N^H_{jmt}}{N^L_{jmt}}\right) + \varphi_{kt}\Delta\widetilde{\mathrm{E}}_t\left(c_{jmk}\right) + \Delta\theta_{kt}X_{jm} + \Delta\xi_{jmkt} \end{aligned} \tag{9}$$

where $\varphi_{kt} = -\phi - \omega_t - \mu_z\hat{v}_{kt}$. For estimation, I do not impose any restriction on the structure of φ_{kt} and allow it to freely vary by occupation. I set X_{jmt} to be time-invariant X_{jm}, which is the commute time weighted by initial locations of *all* jobs (excluding occupations similar to workers' own occupation) as a measure of *location centrality.* I allow each occupation to have an arbitrarily changing preference $\Delta\theta_{kt}$ for such location centrality measure, so that spatial sorting due to changing preference for location centrality is controlled for.

With this setup, I now discuss the identification of the three sets of parameters: μ_z, γ_z, β_z.

[18] ϕ is calibrated to be 0.3425.

The identification of μ_z - To identify μ_z, I exploit the fact that job locations are distributed differentially for different occupations. My identifying assumption is that while job locations are occupation-specific, amenities are occupation-invariant.[19] Spatial variation in job locations is captured in $\widetilde{\mathrm{E}}_{t-1}(c_{jmk})$, which is the expected commute time weighted by spatial distribution of jobs in occupation k at the initial time period. If location j is near a large concentration of jobs in occupation k, $\widetilde{\mathrm{E}}_{t-1}(c_{jmk})$ would be small, because short commute times would receive large weights. Therefore, if rising value of time makes workers want to move closer to work, I should see them move toward their *occupation-specific* job locations, which are locations with shorter $\widetilde{\mathrm{E}}_{t-1}(c_{jmk})$. By observing differential migration patterns by occupation and observing how much workers in occupations with increasing value of time migrate to locations with shorter $\widetilde{\mathrm{E}}_{t-1}(c_{jmk})$, I can identify μ_z. One usual worry for identification is that location with small $\widetilde{\mathrm{E}}_{t-1}(c_{jmk})$ tend to be in central cities, which have other amenity features for which workers may have time-varying preferences. To ensure identification, I include an occupation-invariant centrality measure X_{jm} in the demand equation, and allow workers to have differentially (by occupation) changing taste for such a measure. I exploit only the residual variation in occupation-specific job locations to identify μ_z.

The identification of γ_z - To identify γ_z, I exploit the idea that locations of jobs that are unrelated to a worker may indirectly affect that worker's migration choice by changing local amenity levels, through infuencing *other* workers' migration choices. Based on this idea, I construct instrumental variables for change in log skill ratio $\Delta \log \left(\frac{N^H_{jmt}}{N^L_{jmt}}\right)$, by computing the predicted log change in census tract populations of high- and low-skilled workers, driven purely by differential changes in *value of time* and *job locations* by occupation. While constructing instruments for workers in occupation k, I exclude occupations similar to occupation k[20], under the assumption that value of time and job locations of other occupations do not directly affect workers' location preference, but may indirectly affect workers' location choice through endogenously changing the census tract's population mix.

I compute the predicted population of each occupation in each neighborhood in 2010 using only variation in the value of time and expected commute time:

$$\hat{N}_{jmk,2010} = N_{mk,1990} \cdot \frac{\exp\left(\log\left(s_{jmk,1990}\right) - \hat{\mu}\Delta\hat{v}_{k,2010}\widetilde{\mathrm{E}}_{t-1}\left(c_{jmk}\right)\right)}{\sum_{j'\in J_m} \exp\left(\log\left(s_{j'mk,1990}\right) - \hat{\mu}\Delta\hat{v}_{k2010}\widetilde{\mathrm{E}}_{t-1}\left(c_{j'mk}\right)\right)}$$

where $\hat{\mu}$ is the preliminary parameter estimate from estimating the unconditional location demand equation without including amenities or rent.[21] The predicted log population changes of high- and low-skilled workers, respectively, are then $\Delta \log \hat{N}^H_{jm2010,-k} = \log\left(\sum_{\substack{k'\in K_H \\ k'\nsim k}} \hat{N}_{jmk'2010}\right) -$

[19] Preferences for amenities could be changing and occupation-specific, but the amenities themselves do not vary systematically by occupation. A counter-example would be that financial workers make prefer the amenities at location A to those at B, while doctors may prefer amenities at location B to those at A. If location A has high concentration of financial jobs and location B has high concentration of hospitals, it could undermine my identification. My identifying assumption preclude such occupation-specific amenities.

[20] I define similar occupations as occupations that belong to the same occupation group in the IPUMS Census/ACS data. There are 25 occupation groups.

[21] The size of $\hat{\mu}$ does not matter in the estimation.

$\log\left(\sum_{\substack{k'\in K_H\\ k'\nsim k}} N_{jmk'1990}\right)$, $\Delta\log\hat{N}^{L}_{jm2010,-k} = \log\left(\sum_{\substack{k'\in K_L\\ k'\nsim k}} \hat{N}_{jmk'2010}\right) - \log\left(\sum_{\substack{k'\in K_L\\ k'\nsim k}} N_{jmk'1990}\right)$. I use the $\Delta\log\hat{N}^{H}_{jmt,-k}$ and $\Delta\log\hat{N}^{L}_{jmt,-k}$ as instruments for the actual change in local skill ratio.

The identification of β_z - To identify β_z, the preference for 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.

5.3.1 Robustness of identification

Changing taste for central city amenities - One may worry that workers with rising value of time could also have increasing preference for central city amenities, which may lead to spurious relation between changing value of time and demand for central city locations. To deal with this concern, as mentioned previously, my demand equation includes a time-invariant component X_{jmk}, which measures the centrality of location j. I allow workers to have occupation-specific change in taste $\Delta\theta_{kt}$ for such centrality. This accounts for differentially changing tastes for central city amenities for any reasons, including increasing taste for urban-type amenities or decreasing crime in the city. What identifies μ_z and γ_z is the variation in occupation-*specific* job locations and the differentially changing value of time.

Occupation choice and switching - One could also argue that increasing preference for the central cities may encourage people to choose careers for occupations of which work locations tend to be around central cities. By observing differential migration patterns by occupation, I may actually be capturing the effect of increasing preference for central cities through occupation choice and switching. My identifying assumption is that while job locations are occupation-specific, amenities or other location characteristics are occupation-invariant. If amenities are attractive, they are attractive to workers in all occupations, though with possibily differential degrees. Therefore, changes in preference for amenities can be accounted for by controlling for differential tastes for occupation-invariant location characteristics. In my estimation, I include location centrality as the location characteristic. In the appendix Table A1, I also show results controlling for initial skill ratio in addition to location centrality.[22]

5.4 Housing supply

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

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

[22] One may surmise that workers work long hours because they live in the central city and have less time spent on commuting. While this may be true, my estimates are not likely driven by this mechanism. This is because high-skilled workers' rising prevalence of working long hours is not specific to workers in the central cities, nor is rising long-hour premium specific to workers living in the central cities.

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 instruments with housing stock density den_{jm}.

5.5 Linear GMM estimator

I jointly estimate the location demand and housing supply equations using an efficient GMM estimator.[23] 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.[24]

6 Model estimates

The model is estimated with data from all census tracts in all MSAs in the United States in the 1990 Census and 2007-2011 ACS data.[25] Table 4 presents the summary statistics of the data I use.

6.1 First stage of IVs for local skill ratios

To separately estimate γ_z, I construct the instrumental variables based on predicted log population changes of high- and low-skilled workers. The first stage of these instrumental variables performs quite well. Table 5 presents the results from regressing actual change in log skill ratio on predicted change in log skill ratio and change in log population of high- and low-skilled workers at census tract level.[26] The instruments can predict the change in log skill ratio with very strong F-stats.

[23] The estimation procedure is described in detail in Appendix section D6.

[24] See Appendix section D7 for the construction of Conley standard errors.

[25] The estimation is done with long-hour premiums estimated using national (minus one) data. Long-hour premiums in each MSA are estimated using national data excluding own MSA.

[26] The predicted change in this regression are generated for each census tract. The populations are calculated by summing over predicted populations over all occupations. In the actual estimation, each observation is at occupation/census tract level. The instruments used in the estimation is created by excluding the populations of occupations in the occupation-group of the occupations in question.

6.2 Estimates

Table 6 shows the estimate for μ_z are positive and significant for both high- and low-skilled workers, which shows that workers with higher value of time prefer neighborhoods with shorter expected commute time. The estimate for high-skilled is 8.953, which means that one standard deviation rise in long-hour premium would lead to high-skilled workers having 44% higher demand residential location that can save one hour of daily commute time.[27] The estimate for low-skilled is 2.1035, which means that one standard deviation rise in long-hour premium would lead to low-skilled workers having 11.9% higher demand residential location that can save one hour of daily commute time.

Preference for endogenous amenities γ_z is 2.2193 for high-skilled workers and 0.6873 for low-skilled workers, which means that census tracts with 1% higher skill ratio would raise demand from high-skilled workers 1.432 percentage point more than from low-skilled workers. Therefore, it can be easily seen that an exogenous shock that generates a rise in the local skill ratio in a neighborhood could lead to 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. This implies that some high-skilled workers may sort into the central city neighborhood even without experiencing a value of time shock themselves, so long as the amenity level in the central cities increases.

The preference elasticity with respect to rent β_z for high-skilled is estimated to be 0.7950 and 0.4593 for low-skilled. Percentage-wise, high-skilled are moderately more elasticity with respect to rents. This is likely due to the fact that high-skilled workers are more mobile, and hence smaller σ. Recall from the model setup, the migration elasticities are inversely related to the standard deviation of logit component σ. While each of the preference parameters are larger in magnitude for high-skilled workers, the difference in β_z is relatively moderate.

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 0.9284.[28]

6.3 Amenity supply

In addition to the model estimates, I also demonstrate that local skill ratio is a driving force of various types of local amenities. Table 7 presents the elasticities for the per-capita number of various local business establishments with respect to changes to local skill ratio.[29] I use the same instrument for local skill ratio for identification. I find that the per-capita count of local businesses

[27] 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.

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

[29] 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.

is positively responsive to the 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).[30] 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

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 whether there is similar spatial sorting using alternative 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. (See Table A1 in the appendix) Moreover, I estimate the model using work location in the 2010 Zip-Code Business Pattern to make sure the results are not driven by particular year of job location data. I also re-defined high-skilled occupation using alternative definitions, and estimate the model again using the new definitions. The results remain quite robust. (See Table A1 in the appendix)

7 Determinants of gentrification

Having estimated the model parameters, I now evaluate how much gentrification is caused by a direct effect of a rising value of time, and how much gentrification is caused by the indirect effect of endogenously improved amenities and rent change.[31] I evaluate gentrification by examining the changes in neighborhoods' skill ratios predicted by the model and compare them with the data.

7.1 Direct effect of changing value of time

In the first exercise, I allow only workers' value of time to change from 1990 to 2010, holding neighborhood amenities, rent, all other components, constant at their 1990 levels. I use the model to generate the location choice in 2010 that would have been made if only the value of time had changed. The following equation is the predicted location demand:

$$\begin{aligned} \log\left(\widehat{s_{jmk,2010}}\right) \quad = \quad & \delta_{jmk} + \tilde{\delta}'_{mk,2010} + \log\left(\sum_{n'\in J_m} \pi_{n'mk,t-1} \exp\left(-\mu_z v_{k,2010} \cdot c_{jn'm}\right)\right) \\ & -\beta_z r_{jm,1990} + \gamma_z \log\left(\frac{N^H_{jm,1990}}{N^L_{jm,1990}}\right) + \theta_{k,1990} X_{jm} + \xi_{jmk,1990} \end{aligned} \tag{11}$$

[30] 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.

[31] 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.

To ensure that the predicted location choice probabilities add up to one for each occupation and each MSA, I adjust them accordingly with a normalizing constant $\tilde{\delta}'_{mk,2010}$.[32] All other components of the initial location demand are held fixed.

After computing the predicted location choices $\widehat{s_{jmk,2010}}$, I compute the predicted change in skill ratios (skill ratio in the central cities/skill ratio in the suburbs) and compare them to those observed in the data. The comparisons help assess how much of the neighborhood change is driven by the direct effect of the rising cost of commute time.

To quantify the amount that the model can explain gentrification of the central cities, I compute the model-predicted *relative* log skill ratio in the central cities and compare it with observed change in log skill ratio in the central cities. Relative log skill ratio is defined as the log skill ratio in the central cities *minus* the log skill ratio in the suburbs. I conduct this adjustment because skill ratios have increased in both central cities and suburbs, and relative change in skill ratio captures the degree of spatial sorting between central cities and suburbs.

I define central cities to be census tracts that are located within 3 miles or 5 miles of downtowns. Table 8 shows a little more than 10% of the change in the relative skill ratio in the central city can be explained directly by the changing value of time, with either definitions of central cities. This means that the rising value of time does contribute to gentrification of the central cities, but the direct effect is small. In other words, if we were to adjust workers' value of time in 1990 to 2010 levels, but hold everything else constant, we would only see a little more than 10% of the gentrification of central cities that we actually see.

7.2 Indirect effect through endogenous amenity change

One important reason for the small magnitude of model-predicted change in the last exercise is that the exercise mutes the channel of endogenous amenity changes. Next, I allow the indirect effect of the changing value of time on urban change to operate through endogenous amenity changes. In other words, I evaluate how much gentrification of central cities be explained by the migration of workers who move as a result of amenities changes *brought about* by the movers, in addition to the migration directly due to value of time change. For that purpose, I allow local skill ratio, which approximates endogenous levels of amenities, to vary endogenously and change by the amount predicted by the shock to the value of time. Then, I compute a new set of predicted skill ratios by census tract according to the following location demand equation:[33]

[32] This approach is described in the following equation,
$$\widehat{s_{jmk,2010}} = \frac{\exp\left(\log\left(s_{jmk,1990}\right) - \log\left(\sum_{n'} \pi_{n'mk} \exp\left(-\mu v_{k,1990} \cdot c_{jn'm,t-1}\right)\right) + \log\left(\sum_{n'} \pi_{n'mk} \exp\left(-\mu v_{k,2010} \cdot c_{jn'm,t-1}\right)\right)\right)}{\sum_{j'} \exp\left(\log\left(s_{j'mk,1990}\right) - \log\left(\sum_{n'} \pi_{n'mk} \exp\left(-\mu v_{k,1990} \cdot c_{j'n'm,t-1}\right)\right) + \log\left(\sum_{n'} \pi_{n'mk} \exp\left(-\mu v_{k,2010} \cdot c_{j'n'm,t-1}\right)\right)\right)}.$$

[33] I regress the actual change in log skill ratio on the change in log skill ratio 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.

$$\log\left(\widehat{s_{jmk,2010}}\right) = \delta_{jmk} + \tilde{\delta}'_{mk,2010} + \log\left(\sum_{n' \in J_m} \pi_{n'mk,t-1} \exp\left(-\mu_z v_{k,2010} \cdot c_{jn'm}\right)\right) \qquad (12)$$

$$-\beta_z r_{jm,2010} + \gamma_z \log\left(\widehat{\frac{N^H_{jm,2010}}{N^L_{jm,2010}}}\right) + \theta_{k,1990} X_{jm} + \xi_{jmk,1990}.$$

To obtain the predicted amenity level, I first regress the observed census tract-level change in log skill ratio on the change in log skill ratio predicted by the first exercise. The following is the fitted change in log skill ratio:

$$\Delta\log\left(\widehat{\frac{N^H_{jm,2010}}{N^L_{jm,2010}}}\right) = \hat{\alpha}_{m0} + \hat{\alpha}_1 \Delta\log\left(\frac{\hat{N}^H_{jm,2010}}{\hat{N}^L_{jm,2010}}\right)$$

where $\hat{\alpha}_1 = 3.025$ and $\hat{N}^H_{jm,2010}$ and $\hat{N}^L_{jm,2010}$ are the census tract population predicted by the first exercise.

I then take the fitted change in log skill ratio and add it to the log skill ratio observed in 1990 to compute the log skill ratio/amenity level in 2010.

With this new set of predicted location demand changes, the change in local skill mix is further affected by endogenous amenity changes and rent changes due to sorting of 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 the migration elasticity γ_z to amenity is much higher for high-skilled workers, adding the effect of endogenous amenity changes would draw in even more high-skilled workers in neighborhoods in which changes in skill ratios are already positively affected by the shock to the value of time.

Table 8 shows that once endogenous amenity and rents are adjusted, the predicted gentrification is much larger. This shows a large share of the gentrifiers who migrated to the central cities are attracted by the endogenously improved amenities, rather than an increased cost of commuting due to a rising value of time per se. However, the magnitude still only accounts for half of the full change of the central cities. This means that the gentrification of central cities cannot be entirely explained by the channels described in my paper. Other factors must also have played an important role in gentrifying the central cities.

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

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 quantify 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 show that the rising value of time has a modest direct role in gentrifying the central cities. However, the effect is substantially amplified by endogenous amenity change.

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 preferences for local amenities, but such preferences are estimated as a reduced-form elasticity with respect to the 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. https://doi.org/10.1086/491590

[2] Aguiar, Mark and Erik Hurst, 2007 "Measuring Leisure: The Allocation of Time over Five Decades." *Quarterly Journal of Economics* 122(3): 969-1006. https://doi.org/10.1162/qjec.122.3.969

[3] Ahlfeldt, Gabriel, Stephen Redding, Daniel Sturm, and Nikolaus Wolf, 2015. "The Economics of Density: Evidence From the Berlin Wall," *Econometrica* 83: 2127-2189. https://doi.org/10.3982/ecta10876

[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. https://doi.org/10.1016/j.jue.2015.03.003

[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. https://doi.org/10.1086/426036

[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. https://doi.org/10.1162/003355303322552801

[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. https://doi.org/10.1086/675536

[8] Autor, David, Christopher Palmer, and Parag Pathak, 2017. "Gentrification and the Amenity Value of Crime Reductions: Evidence from Rent Deregulation." *AEA Papers and Proceedings*, 109 : 381-84. https://doi.org/10.1257/pandp.20191022

[9] Baum-Snow, Nathaniel, 2007. "Did Highways Cause Suburbanization?" *Quarterly Journal of Economics*, 122(2): 775-805. https://doi.org/10.1162/qjec.122.2.775

[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. https://doi.org/10.1086/522381

[12] Becker, Gary. 1965. "A Theory of the Allocation of Time," *Economic Journal*, 75(299): 493-517. https://doi.org/10.2307/2228949

[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. https://doi.org/10.1086/261796

[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. https://doi.org/10.1162/qjec.2010.125.1.417

[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. https://doi.org/10.1016/s1574-0080(87)80006-8

[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. https://doi.org/10.1162/rest.91.4.725

[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. https://doi.org/10.1016/s0014-2921(98)00019-1

[19] Card, David, Alexandre Mas, and Jesse Rothstein, 2008. "Tipping and the Dynamics of Segregation," *Quarterly Journal of Economics* 123(1): 177-218. https://doi.org/10.1162/qjec.2008.123.1.177

[20] Conley, Timothy, 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics*, 92(1): 1-45. https://doi.org/10.1016/s0304-4076(98)00084-0

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

[22] Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst, 2019. "Income Growth and the Distributional Effects of Urban Spatial Sorting," Working Paper. https://doi.org/10.3386/w26142

[23] Couture, Victor and Jessie Handbury, 2019. "Urban Revival in America." Working Paper. https://doi.org/10.3386/w24084

[24] Davis, Donald, Jonathan Dingel, Joan Monras, and Eduardo Morales, 2019 "How Segregated is Urban Consumption?" *Journal of Political Economy*, 127(4): 1684-1738. https://doi.org/10.1086/701680

[25] 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. https://doi.org/10.1257/aer.20131706

[26] Edlund, Lena, Cecilia Machado, and Michaela Sviatchi, 2015 "Bright Minds, Big Rent: Gentrification and the Rising Returns to Skill." NBER Working Paper w21729. https://doi.org/10.3386/w21729

[27] Ellen, Ingrid Gould, Keren Mertens Horn, and Davin Reed, 2019. "Has Falling Crime Invited Gentrification?" *Journal of Housing Economics*, forthcoming. https://doi.org/10.1016/j.jhe.2019.101636

[28] Ellison, Glenn, Edward Glaeser, 1997. "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach." *Journal of Political Economy*, 105(5): 889-927. https://doi.org/10.1086/262098

[29] Ellison, Glenn, Edward Glaeser, and William Kerr, 2010. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns." *American Economic Review*, 100(3): 1195-1213. https://doi.org/10.1257/aer.100.3.1195

[30] Glaeser, Edward, Jed Kolko, and Albert Saiz, 2001. "Consumer City," *Journal of Economic Geography* 1(1): 27-50. https://doi.org/10.1093/jeg/1.1.27

[31] Goldin, Claudia, 2014 "A Grand Gender Convergence: Its Last Chapter." *American Economic Review*, 104(1): 1091-1119. https://doi.org/10.1257/aer.104.4.1091

[32] Guerrieri, Veronica, Daniel Hartley, and Erik Hurst, 2013. "Endogenous Gentrification and Housing Price Dynamics." *Journal of Public Economics*,100(C): 45-60. https://doi.org/10.1016/j.jpubeco.2013.02.001

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

[34] Holian, Matthew and Matthew Kahn, 2015. "Household Carbon Emissions from Driving and Center City Quality of Life." *Ecological Economics* 116: 362-368. https://doi.org/10.1016/j.ecolecon.2015.05.012

[35] 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. https://doi.org/10.1086/533618

[36] Lazear, Edward and Sherwin Rosen, 1981 "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy*, 89(5):841-864. https://doi.org/10.1086/261010

[37] Lee, Sanghoon and Jeffrey Lin, 2018. "Natural Amenities, Neighborhood Dynamics, and Persistence in the Spatial Distribution of Income," *Review of Economic Studies,* 85(1): 663-94. https://doi.org/10.1093/restud/rdx018

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

[39] 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.

[40] Nevo, Aviv and Arlene Wong, 2018. "The Elasticity of Substitution Between Time and Market Goods: Evidence from the Great Recession," International Economic Review, 60(1): 25-51. https://doi.org/10.1111/iere.12343

[41] Saiz, Albert, 2010. "The Geographic Determinants of Housing Supply," *Quarterly Journal of Economics*, 125(3):1253-1296. https://doi.org/10.1162/qjec.2010.125.3.1253

[42] Roback, Jennifer, 1982. Wages, Rents, and the Quality of Life. *Journal of Political Economy*, 90(6): 1257-1278. https://doi.org/10.1086/261120

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

[44] 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. https://doi.org/10.1016/s1574-0080(04)80006-3

[45] 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.

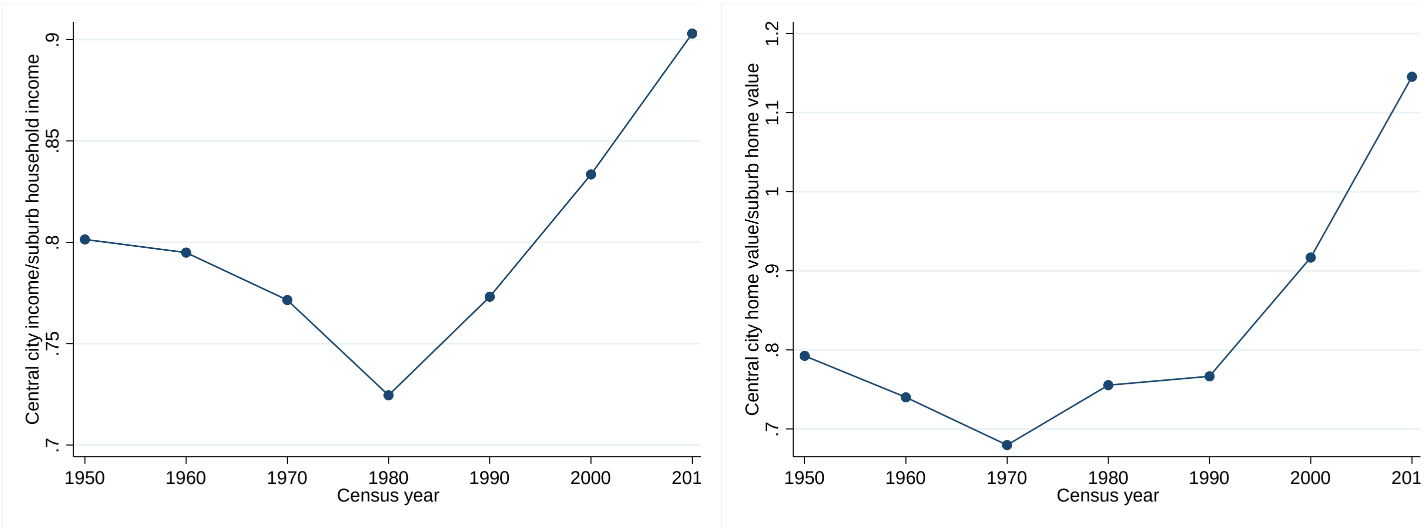
[46] U.S. Department of Transportation, Federal Highway Administration, 2009 National Household Travel Survey. URL: http://nhts.ornl.gov.

[47] 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

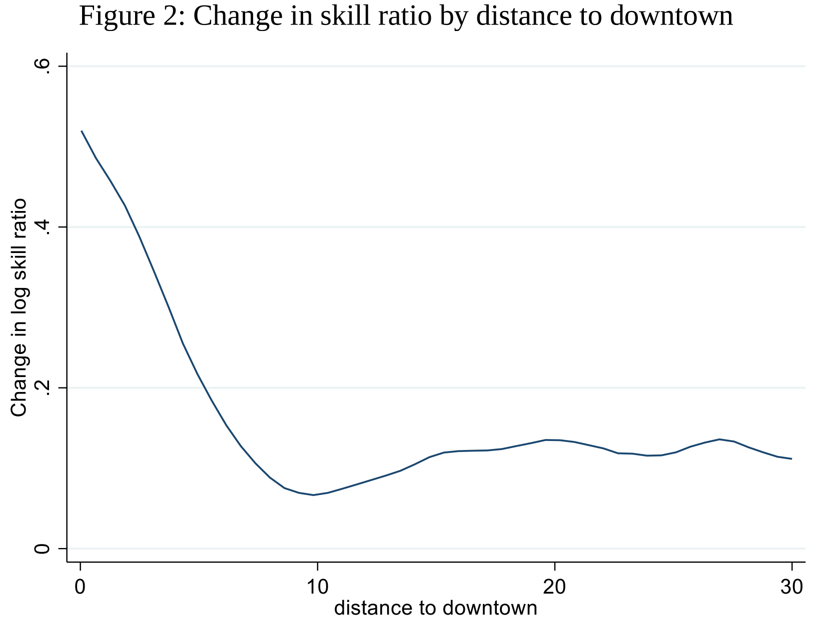
(a) Income ratio

(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.

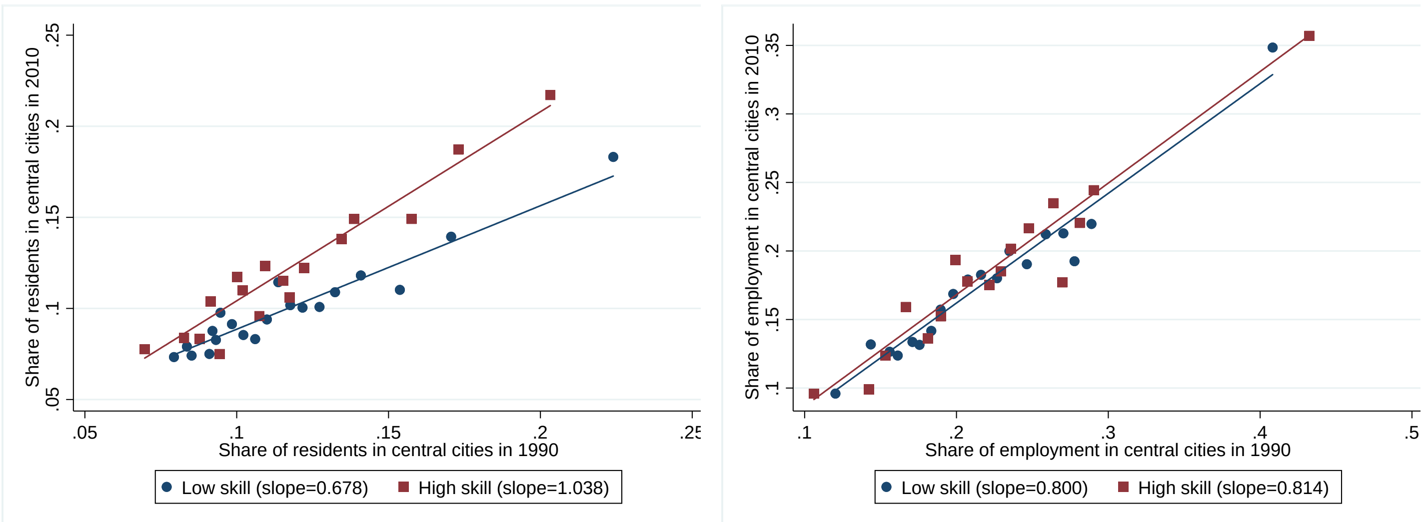
Figure 2: Change in skill ratio by distance to downtown



Notes: The graph shows a non-parametric plot between the change in log skill ratio and census tracts' distance to downtowns for the top 25 MSAs (defined by population ranking in 1990). The source of the data is Census and ACS provided by NHGIS.

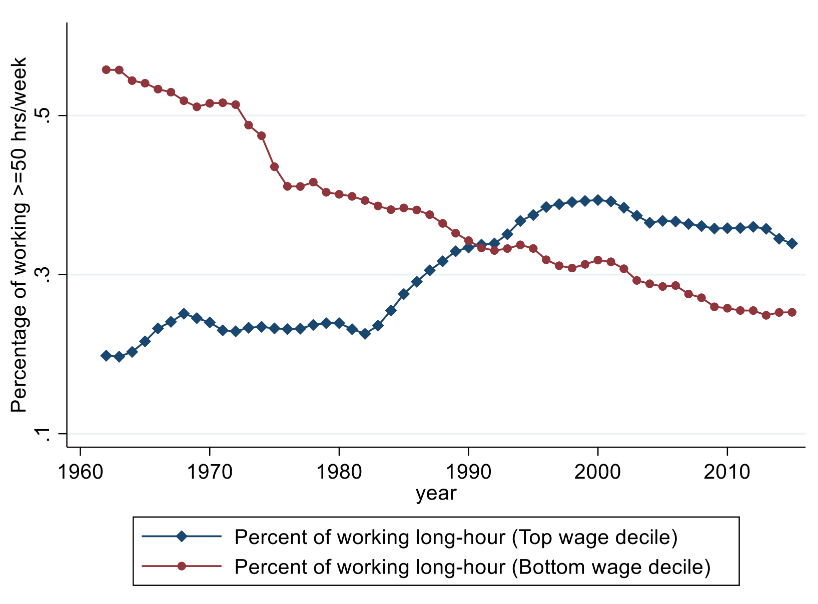
Figure 3: Residential and work location sorting by skill

(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 MSAs to produce these graphs.

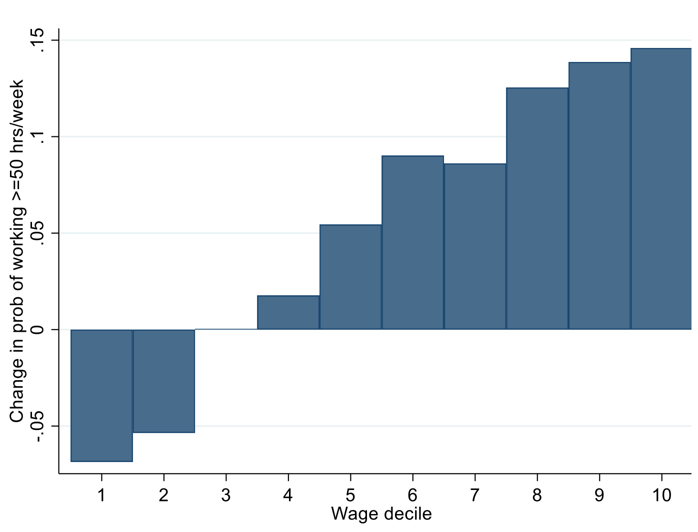
Figure 4: The evolution of long-hour working



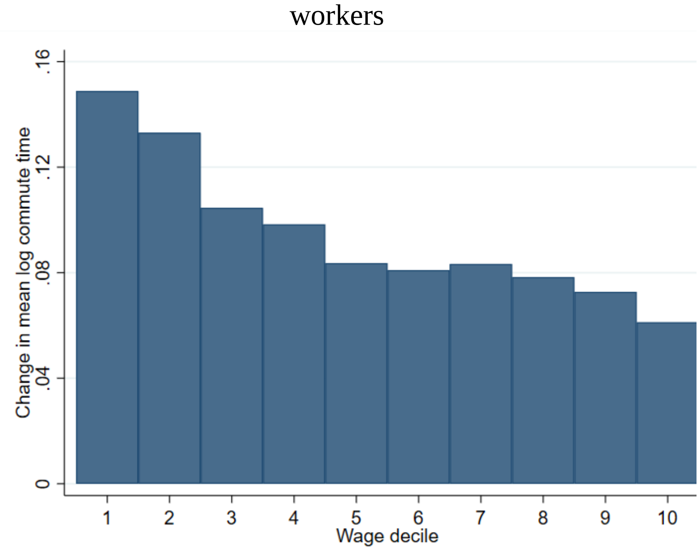
Notes: I plot the probability of working at least 50 hours a week using the CPS ASEC data from 1968 to 2016. The sample includes workers that are male, between age 25 and 65 and work at least 30 hours per week. I plot the probability of working long hours for workers in the top wage decile and the bottom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

Figure 5: Changing working hours and commute time by wage decile (1980-2010)

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

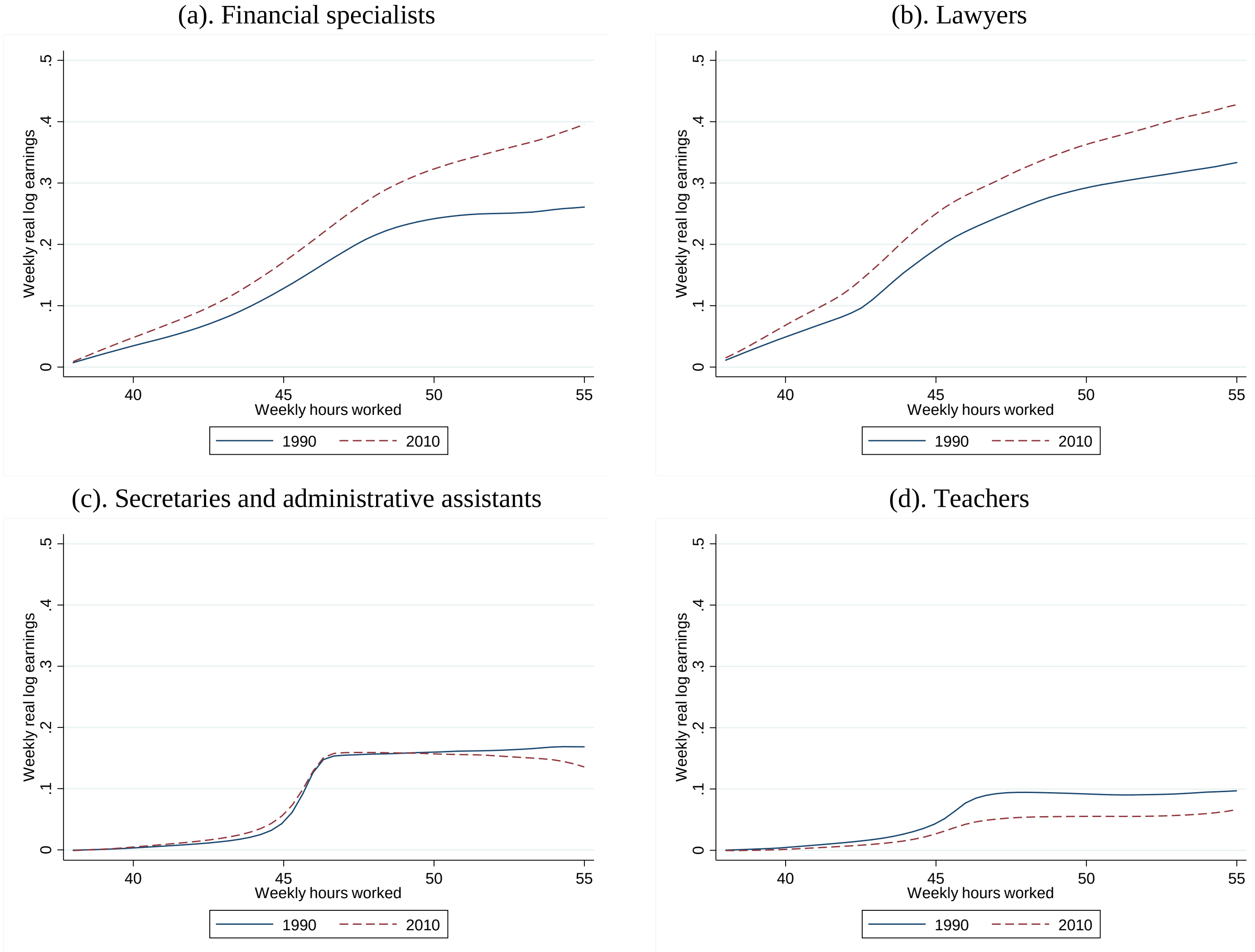


(b) Growth in commute time slower for higher-wage workers



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 6: Residual log weekly earnings against weekly hours worked



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.

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 (skill ratio)	0.284***	0.013	0.454***	0.528***	-0.495***	0.597***
	(0.036)	(0.033)	(0.034)	(0.038)	(0.112)	(0.0929)
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	0.2836	0.4849

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 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 skill ratio for (5) – (6) I match census tracts to municipalities, and compute the overall skill ratio using variables summed over across census tracts matched to municipalities.
*** p<0.01, ** p<0.05, * p<0.1

Table 2: Workers in increasingly long-hour occupations increasingly live the central cities and have shorter commuting time

	Δ ln (share in central city)			Δ ln (reported commuting time)		
	Largest 10 MSAs	Largest 25 MSAs	all MSAs	Largest 10 MSAs	Largest 25 MSAs	all MSAs
Δ ln (pct long-hour)	0.244**	0.208***	0.0984**	-0.0393*	-0.029*	-0.0176
	(0.081)	(0.057)	(0.040)	(0.018)	(0.015)	(0.011)
Observations	2,140	5,347	45,279	2,120	5,276	39,386
Fixed-Effects	MSA			MSA		
Tabulation	MSA/occupation			MSA/occupation		
S.E.	Cluster at MSA			Cluster at MSA		

Notes: Results shown above are OLS regressions, with tabulated cells by MSA and occupation. I compute the share in central city by computing the percentage of workers in each occupation in each MSA who live within 5-mile radius of downtown pin. The percentage of long-hour is defined as the share of workers within each occupation who work at least 50 hours a week. The regressions are conducted by first-difference between data in 2010 and 1990. MSA fixed effects are included. Standard errors are clustered at MSA level.
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Reduced-form relationship between long-hour premium and long-hour worked, central city sorting, and commuting time

	Δ ln (pct long-hour)	Δ ln (share in central city)			Δ ln (reported commuting time)		
		Largest 10 MSAs	Largest 25 MSAs	all MSAs	Largest 10 MSAs	Largest 25 MSAs	all MSAs
ΔLHP	14.38***	9.82**	8.05**	4.04*	-2.35***	-1.72***	-0.68
	(3.70)	(4.21)	(3.00)	(2.08)	(0.61)	(0.59)	(0.54)
Observations	214	2,140	5,347	45,279	2,120	5,276	39,386
Fixed-Effects	N/A	MSA			MSA		
Tabulation	Occupation	MSA/occupation			MSA/occupation		
S.E.	Robust	Cluster at MSA			Cluster at MSA		

Notes: Results shown above are OLS regressions, with tabulated cells by MSA and occupation. I compute the share in central city by computing the percentage of workers in each occupation in each MSA who live within 5-mile radius of downtown pin. The percentage of long-hour is defined as the share of workers within each occupation who work at least 50 hours a week. The regressions are conducted by first-difference between data in 2010 and 1990. LHP denotes long-hour premium. MSA fixed effects are included. Standard errors are clustered at MSA level.
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Summary statistics

		Obs	Mean	SD	min	Max
Long-hour premium	1990	214	0.0145	0.00455	0.00324	0.0365
	2010	214	0.0144	0.00579	-0.00597	0.0355
	Change	214	-0.000136	0.00548	-0.0238	0.0197
Long-hour premium (high-skilled)	1990	58	0.0135	0.00427	0.00475	0.0237
	2010	58	0.0145	0.00656	0.00286	0.0355
	Change	58	0.000955	0.00491	-0.0115	0.0197
Long-hour premium (low-skilled)	1990	156	0.0149	0.00461	0.00324	0.0365
	2010	156	0.0143	0.00551	-0.00597	0.0332
	Change	156	-0.000541	0.00564	-0.0238	0.0158
Skill ratio	1990	42,346	-1.112	0.577	-4.311	0.747
	2010	42,346	-0.982	0.626	-3.910	1.148
	Change	42,346	0.130	0.359	-1.651	2.367
Rent	1990	42,346	6.492	0.429	5.107	7.421
	2010	42,346	6.653	0.424	4.595	7.601
	Change	42,346	0.160	0.264	-2.404	2.479
Restaurants per 1000 residents	1990	19,291	-5.194	1.609	-9.861	4.956
	2010	19,291	-5.081	1.538	-10.815	5.187
	Change	19,291	0.114	0.931	-9.242	5.406
Grocery stores per 1000 residents	1990	19,291	-5.929	1.454	-10.077	3.871
	2010	19,291	-6.128	1.381	-10.815	3.730
	Change	19,291	-0.199	0.877	-8.756	4.491
Gyms per 1000 residents	1990	19,291	-8.435	1.729	-11.880	1.946
	2010	19,291	-7.984	1.645	-11.976	2.767
	Change	19,291	0.450	0.800	-8.973	4.725
Personal services per 1000 residents	1990	19,291	-6.673	1.749	-11.432	4.127
	2010	19,291	-6.789	1.611	-10.927	4.002
	Change	19,291	-0.116	0.965	-10.882	4.586
Violent crime rate	1990	1,870	1.257	1.023	-2.303	4.150
	2010	1,870	0.878	0.947	-2.303	4.177
	Change	1,870	-0.378	0.746	-5.133	3.204
Property crime rate	1990	1,870	3.797	0.598	-2.303	5.843
	2010	1,870	2.821	1.043	-2.303	6.763
	Change	1,870	-0.976	0.826	-6.367	4.209

Notes: The table shows the summary statistics for long-hour premium (high- and low-skilled), skill ratio, rent, and various amenities and crime. For long-hour premium, each observation is an occupation. High-skilled occupations are those with at least 40% college degree in 1990 Census, and low-skilled occupations are the rest of the occupations. Skill ratio and rents are panels by census tracts in 1990 and 2010. Skill ratio is defined Amenities (restaurant, grocery, gym, personal services) are panels by census tracts as well. For 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. Violent and property crime rates are panel at municipality level.

Table 5: First-stage between actual change in skill ratio and predicted change in skill ratio

	Dep variable: Actual Δ log skill ratio	
	(1)	(2)
Predicted Δ in log skill ratio	0.996***	-
	(0.10)	
Predicted Δ change in high-skilled workers	-	1.659***
		(0.14)
Predicted Δ change in low-skilled workers	-	-0.594***
		(0.076)
MSA fixed-effects	Yes	Yes
Observations	43,246	43,246
F-statistics	96.62	154.65

Notes: Results shown above are OLS regressions. Each observation is a census tract. The first-difference is between 1990 and 2010. I use 2007-2011 ACS for year 2010. Column (1) reports regression result when predicted change in log skill ratio is included as the regressor. Column (2) reports regression result when change in high- and low-skilled workers are included separately as the regressors. The predicted change in log skill ratio is generate by changing only the long-hour premium in the model, assuming $\hat{\mu} = 8.96$, which I obtain by estimating the location demand with only the term of commuting cost and no heterogeneity by skill. The model estimates do not respond to the value of $\hat{\mu}$. Conley (1999) HAC standard errors are computed with 1-mile threshold for the kernel function. F-stats are computed accounting for the spatial dependency. MSA fixed effects are used for all regressions.
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Estimates of model parameters

Panel A: Worker's residential location demand			Panel B: Rent	
Commute cost (μ)	High-skilled occupations	8.9532*** (1.0461)		
	Low-skilled occupations	2.1035*** (0.4679)		
Amenity (γ)	High-skilled occupations	2.2193*** (0.1890)	Housing demand$\times$ housing stock density (π_1)	0.9284*** (0.1317)
	Low-skilled occupations	0.6873*** (0.1388)	Housing stock density (π_2)	-0.0675*** (0.0197)
Rent (β)	High-skilled occupations	0.7950*** (0.2338)		
	Low-skilled occupations	0.4593*** (0.1853)		

Notes: Model estimated using occupation/census tract cell data from 1990 to 2010. Number of cells used is 8,755,373. 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. Estimation detail can be found in the text.
*** p<0.01, ** p<0.05, * p<0.1

Table 7: 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 (skill ratio)	0.514***	-0.165	1.058***	0.894***	2.659**	-2.710**
	(0.165)	(0.150)	(0.136)	(0.157)	(1.063)	(1.110)
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 skill ratio. Each observation for column (1) – (4) is at census tract level. For 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 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 skill ratio for (5) – (6), I match census tracts to municipalities, and compute the skill ratio using variables summed over across census tracts matched to municipalities.
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Gentrification decomposition – relative log skill ratio

		Largest 25 MSAs		Largest 50 MSAs	
Central city definition		Δ Value of time	Δ Value of time + Δ rent + endogenous Δ amenity	Δ Value of time	Δ Value of time + Δ rent + endogenous Δ amenity
3 miles within downtown	Actual change	0.305	0.305	0.269	0.269
	Model-predicted change	0.0340	0.131	0.0301	0.111
	% explained	11.1%	43.0%	11.2%	41.3%
5 miles within downtown	Actual change	0.235	0.235	0.201	0.201
	Model-predicted change	0.0299	0.120	0.0267	0.103
	% explained	12.7%	51.1%	13.3%	51.2%

Notes: The results shown in this table show the comparison between actual changes in relative skill ratio and model-predicted changes in relative skill ratio. Relative skill ratio is defined as ratio between skill ratio (residents in high-skilled occupations/residents in low-skilled occupations) in central cities and skill ratio in the suburbs. I use varying definition of central city, and sample from largest 25 MSAs and largest 50 MSAs. Actual changes in relative skill ratio are computed using observed spatial data by occupation. The values shown are the mean change in skill ratio weighted by MSAs' population.

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.

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.

Given the workers' utility function of C, H, the utility maximization problem is

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

subject to budget constraint

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

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

$$C^* = \theta I$$
$$H^* = \frac{1-\theta}{R_{jmt}} I$$

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

$$V_{i,jnmt}\left(L\right) = \theta \log\left(\theta I\right) + \left(1-\theta\right) \log\left(\frac{1-\theta}{R_{jmt}} I\right) - \tilde{\omega}_t c_{jnmt} + \tilde{\gamma}_k a_{jmt} + \tilde{\gamma}_k \zeta_{jmt} + \sigma \varepsilon_{i,jmt}$$

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

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

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.

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

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.

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

Each coefficient is written in terms of the underlying parameters:

$$
\delta_{mkt} = \frac{1}{\sigma}\left(y_{0kt} + v_{kt}T\right)
$$

$$
\mu = \frac{1}{\sigma}
$$

$$
\beta = \frac{1-\theta}{\sigma}
$$

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

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

A.2 Derivation of location demand equation

In the appendix, I start from the normalized indirect utility:

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

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 $\tilde{V}_{jnmkt} = \delta_{mkt} - \mu v_{kt}c_{jnmt} - \omega_t c_{jnmt} - \beta r_{jmt} + \gamma_k a_{jmt} + \gamma_k \zeta_{jmt}$ is the mean utility of occupation

k living in j and working in n.

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:

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

$\Lambda_{nmkt} = \log\left(\sum_{j' \in J_m} \exp\left(\tilde{V}_{j'nmkt}\right)\right)$, which is the expected utility for worker in occupation k working in neighborhood n. 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.

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:

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

I assume the spatial distribution of jobs for each occupation- $\pi_{n'mkt}$ as exogenous to the model within the time frame of this analysis, and the cross-sectional variation in job location is driven 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)). 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.

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

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

where $\tilde{\delta}_{mkt} = \delta_{mkt} - \Lambda_{mkt}$

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_K^j for each census tract j. My goal is to impute the count of residents by detailed occupation level k, namely n_k^j. I do so by first imputing $\hat{\theta}_{k|K}^j$, which is the conditional probability of being in occupation k given one is in occupation-group K. I compute $\hat{\theta}_{k|K}^j$ using IPUMS microdata at PUMA level, assuming that $\hat{\theta}_{k|K}^j$ 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.

$$\hat{n}_k^j = \hat{\theta}_{k|K}^j \cdot n_K^j$$

Once I get $\hat{n}_k^j$, 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_h^z 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.

$$\hat{n}_k^z = \sum_h n_h^z \cdot \hat{\theta}_{k|h}$$

The set of $\hat{n}_k^z$ 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

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

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 (2019)'s results in which reverse commuting became more prevalent after 2002.

D Estimation

D.1 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 A8 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 A9, 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 A2 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.2 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 (See Table A1).

D.3 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 A10a, 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 A10b 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 A11, 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.4 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:

$$\min_{\beta_1 \dots \beta_{57}} \frac{1}{N} \sum_{i=1}^{N} (y_i - \beta_1 x_{1i} - \dots - \beta_{57} x_{57i})^2$$

$$\text{subject to } \sum_{j=1}^{57} |\beta_j| \leq t$$

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 $j = 1, \dots 57$ 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.

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

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.5 Linearization of location demand

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

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

The nonlinear term $\log\left(\sum_{n'\in J_m}\pi_{n'mkt}\exp\left(-\phi\cdot c_{jn'm}\right)\right)$ can be approximated by

$$
\begin{aligned}
& \log\left(\sum_{n'\in J_m}\pi_{n'mkt}\exp\left(-\phi\cdot c_{jn'm}\right)\right)\\
\approx\ & \log\left(\sum_{n'\in J_m}\pi_{n'mk,t-1}\exp\left(-\phi\cdot c_{jn'm}\right)\right)\\
& +\frac{\sum_{n'\in J_m}\pi_{n'mkt}\exp\left(-\phi\cdot c_{jn'm}\right)-\sum_{n'\in J_m}\pi_{n'mk,t-1}\exp\left(-\phi\cdot c_{jn'm}\right)}{\sum_{n'\in J_m}\pi_{n'mk,t-1}\exp\left(-\phi\cdot c_{jn'm}\right)}\\
=\ & \log\left(\sum_{n'\in J_m}\pi_{n'mk,t-1}\exp\left(-\phi\cdot c_{jn'm}\right)\right)+\frac{\sum_{n'\in J_m}\pi_{n'mkt}\exp\left(-\phi\cdot c_{jn'm}\right)}{\sum_{n'\in J_m}\pi_{n'mk,t-1}\exp\left(-\phi\cdot c_{jn'm}\right)}-1\\
\approx\ & \delta_{jmk}
\end{aligned}
$$

The term itself varies by j,m,k,t. To simplify, I decompose the term into two parts. The first part is the term evaluated with initial job location. The second part is a the ratio between the expected utilities evaluated with job locations at $t-1$ and job locations at t, *holding the distaste for commuting time constant (ϕ)*. For feasibility reason, I assume that the ratio is constant across occupations. If jobs in occupations in which workers experience rising value of time are not becoming increasingly concentrated in the initial locations, this assumption would not affect my estimation. Unde this assumption, the first term becomes a constant that is j and k specific, which I write it as a fixed-effects term δ_{jmk}. After some algebraic arrangement, the location demand equation can be approximate as following

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

D.6 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.6.1 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 $\mathbf{\Sigma}_{\boldsymbol{\varepsilon}}$. 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.

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

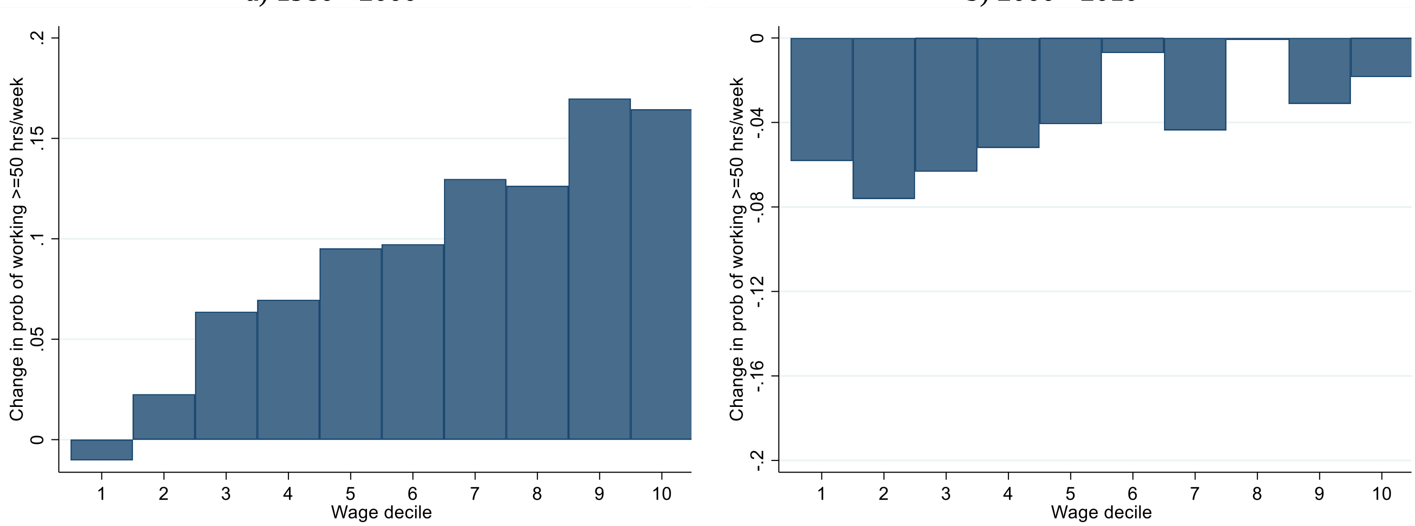
(c) San Francisco

(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)

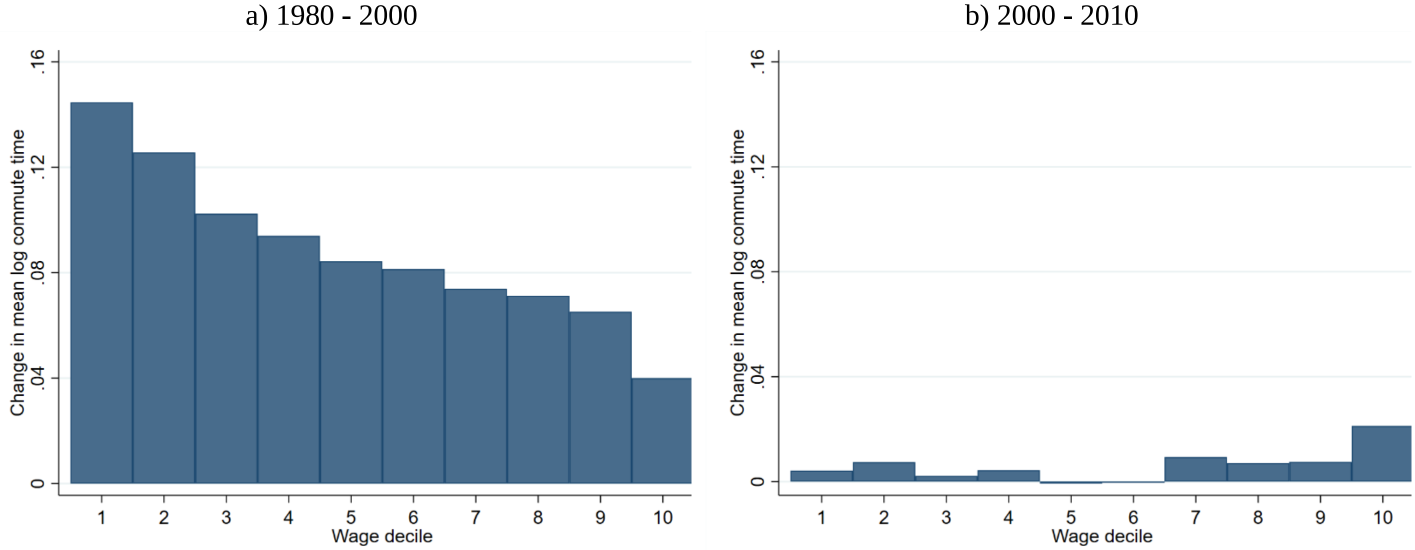
a) 1980 - 2000 b) 2000 - 2010



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.

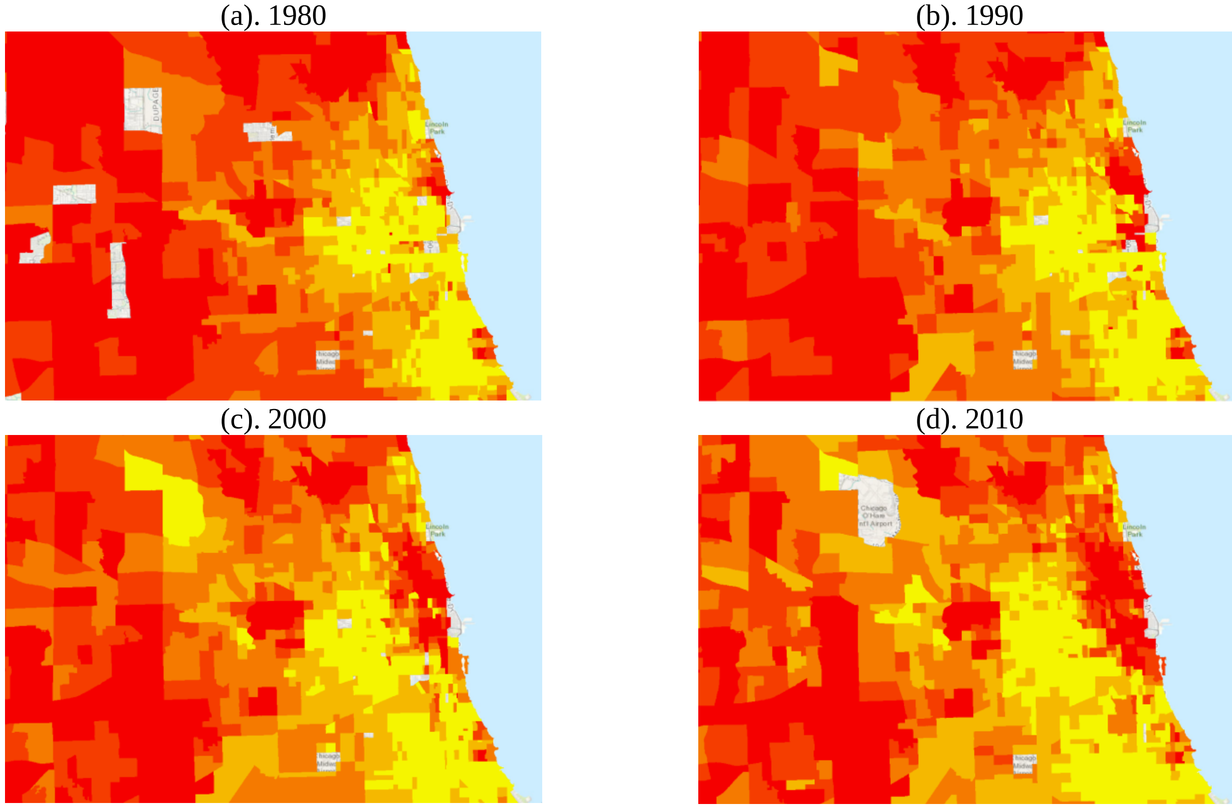
Figure A3: Growth of commute time by wage decile

a) 1980 - 2000 b) 2000 - 2010

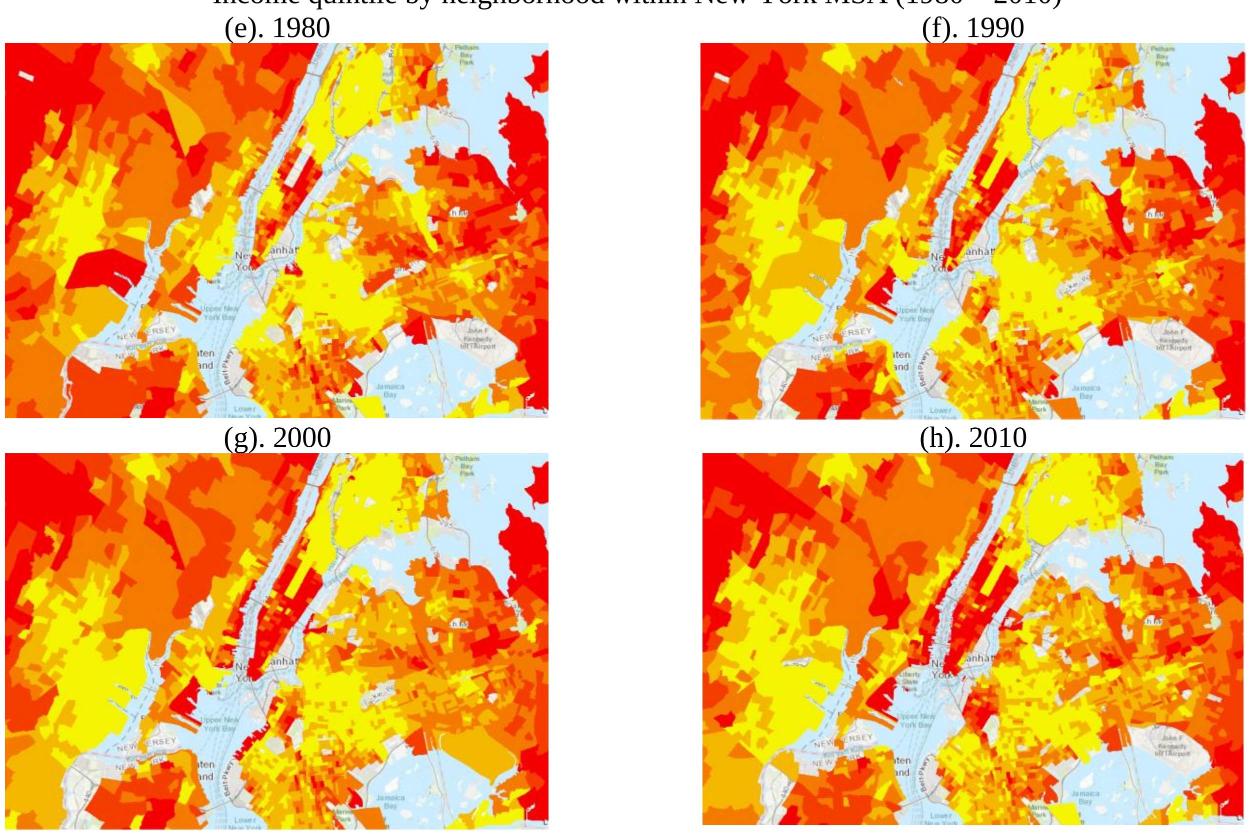


Notes: Data come from IPUMS census data in 1980, 2000, 2010 (2007-2011 ACS). I compute the change in log commute time reported in the Census/ACS data. The sample includes workers that are between 25 and 65 of age, males, working at least 30 hours per week and living in the most populous 25 MSAs in the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



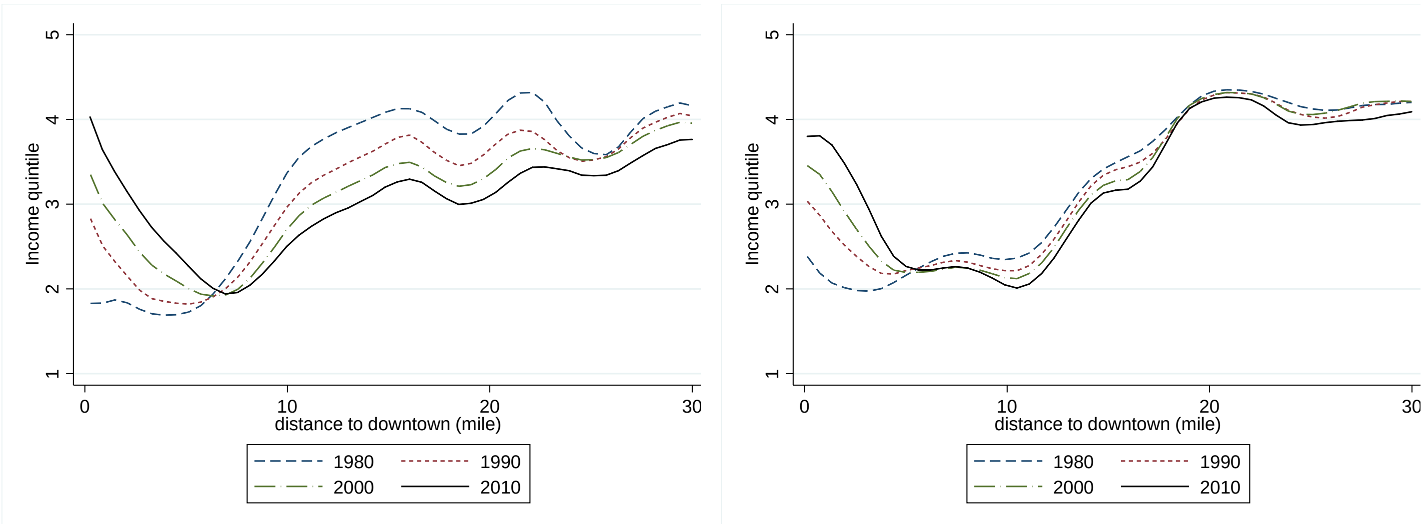
Income quintile by neighborhood within New York MSA (1980 – 2010)



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.

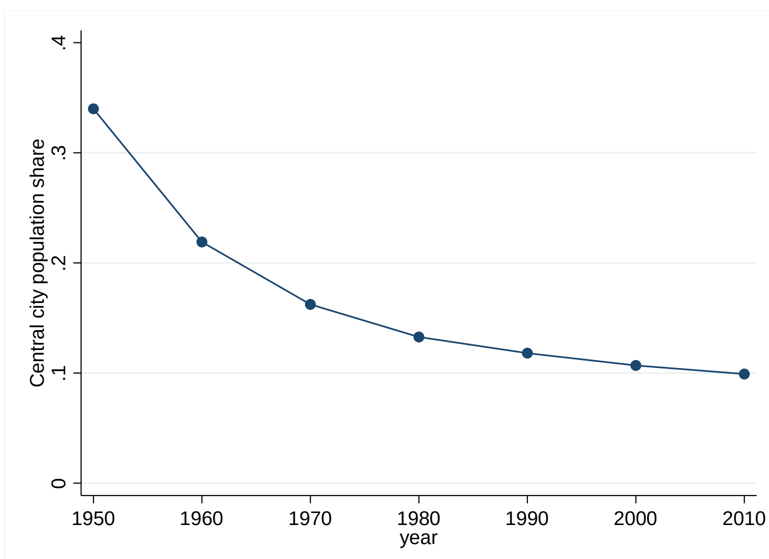
Figure A5: Income quintile by distance to downtown.

(a). Chicago (b). New York



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.



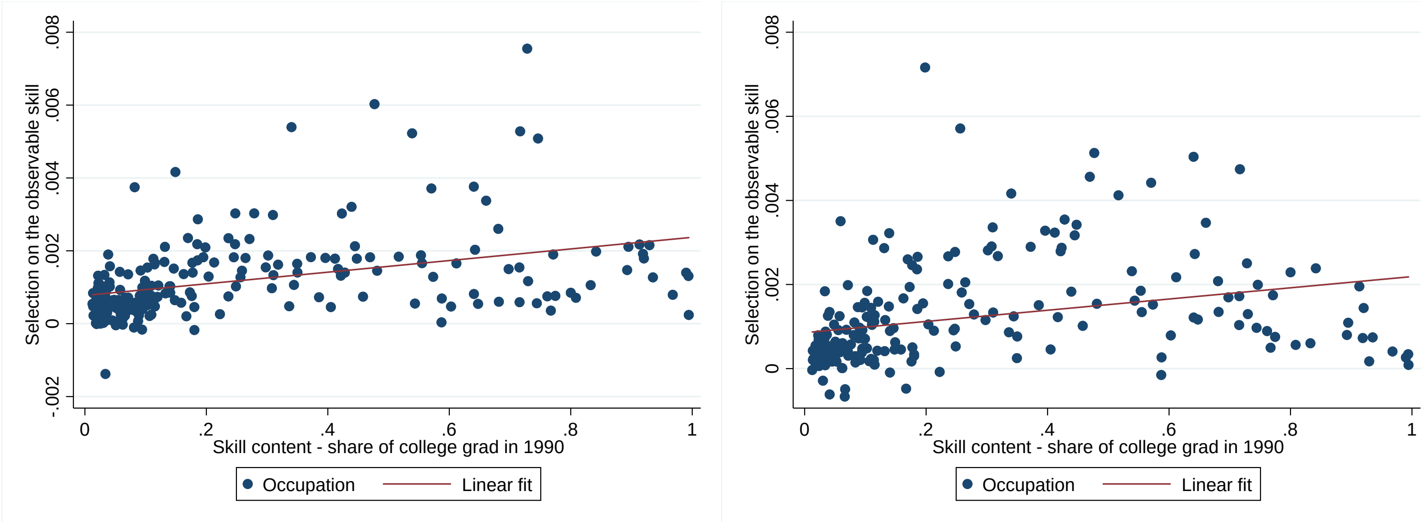
Notes: Central cities in this graph are defined as census tracts that are located within 5 miles of the downtown pin on Google in the respective MSAs. The value plotted in the graph are the population ratio between the population in the census tracts located in the central cities and the total population 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.

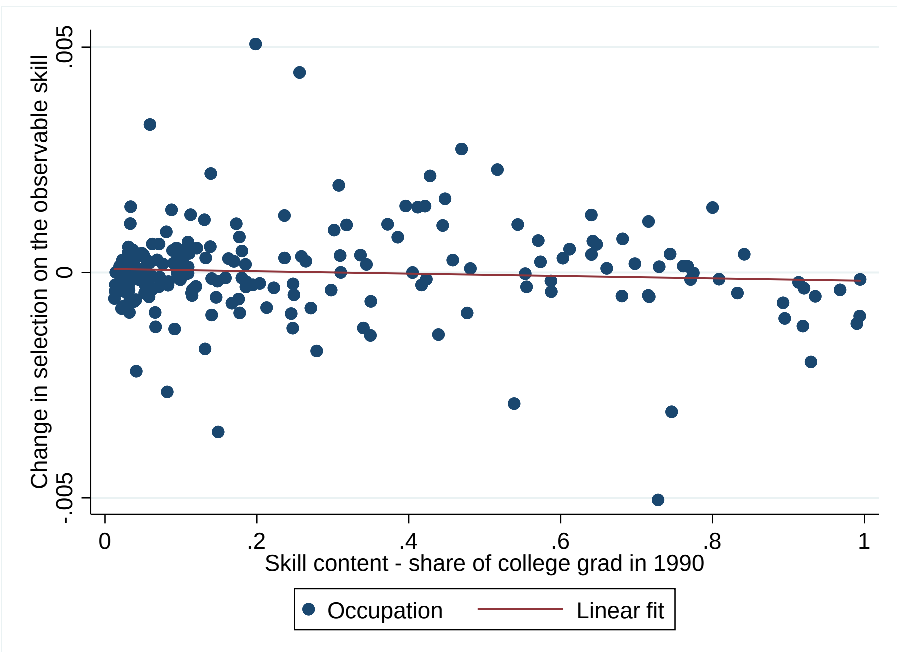
Figure A8: Degree of selection for long-hour premium estimates on observable skills in 1990 and 2010

(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.

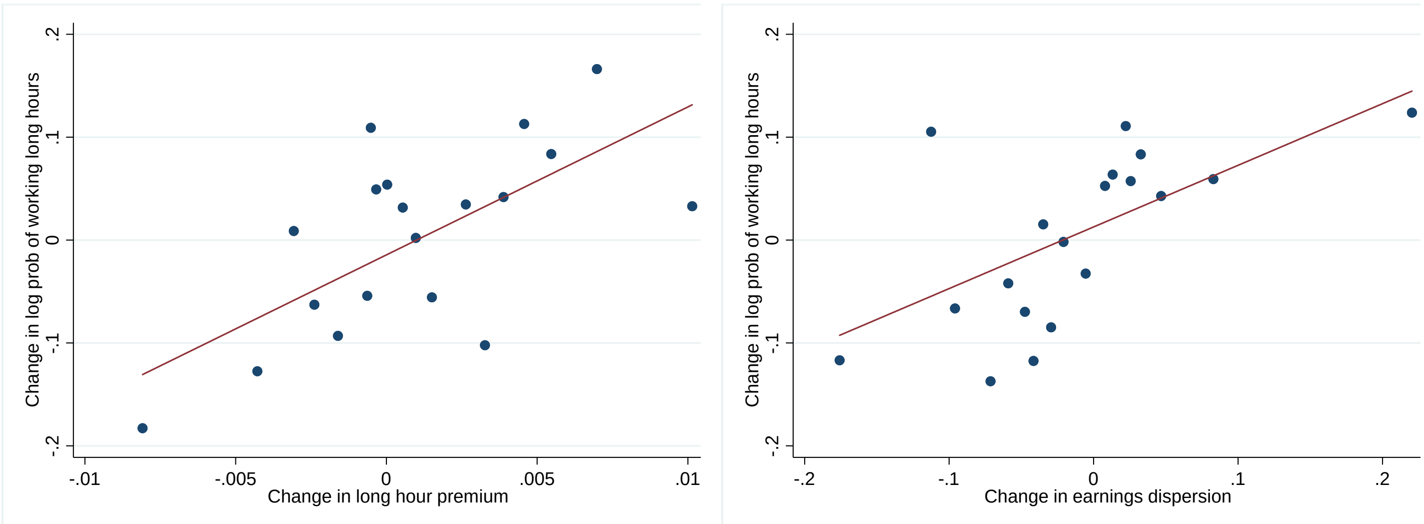
Figure A9: Change in the degree of selection for long-hour premium estimates on observable skills



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.

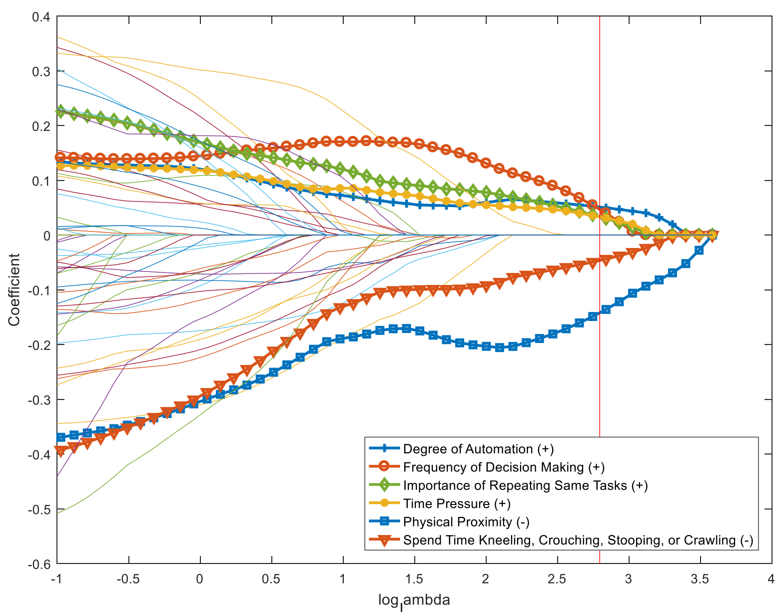
Figure A10: Change in the value of time and prevalence of working long hours (binscatter plots)

(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.

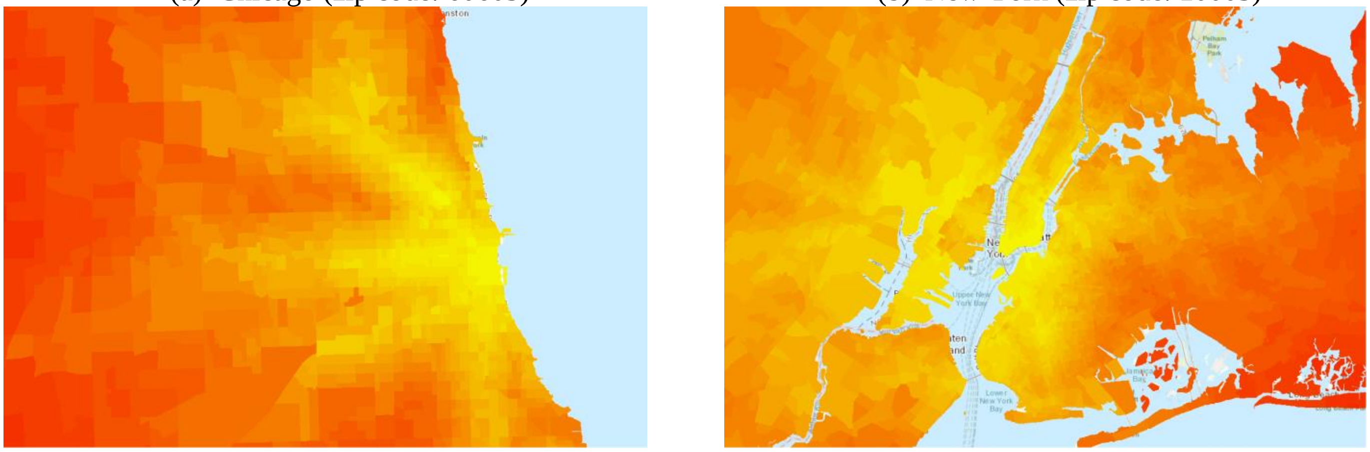
Figure A11: Lasso trace plot of the O*NET characteristics at predicting change in long-hour premium



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 A12: Imputed 1995 rush-hour driving time

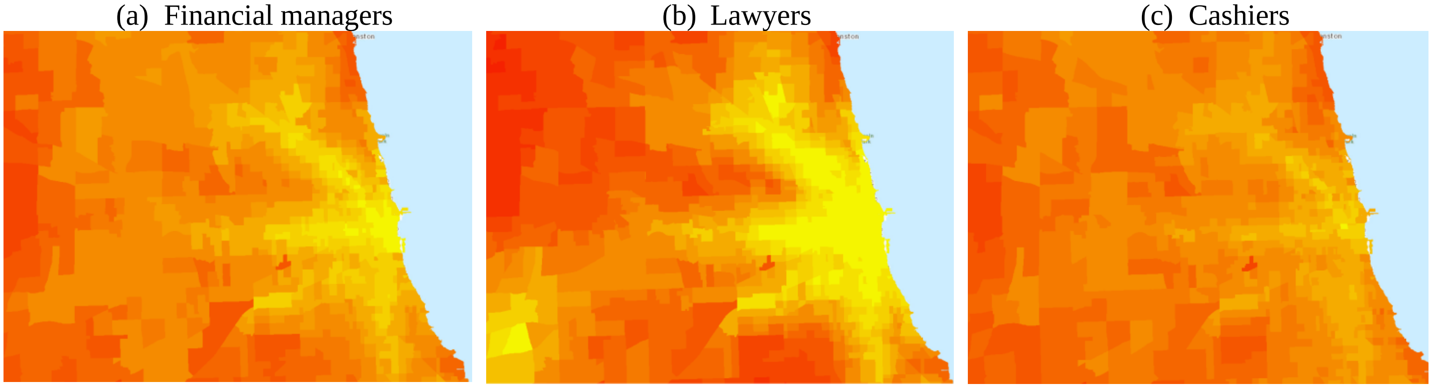
(a) Chicago (zip code: 60605) (b) New York (zip code: 10005)



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.

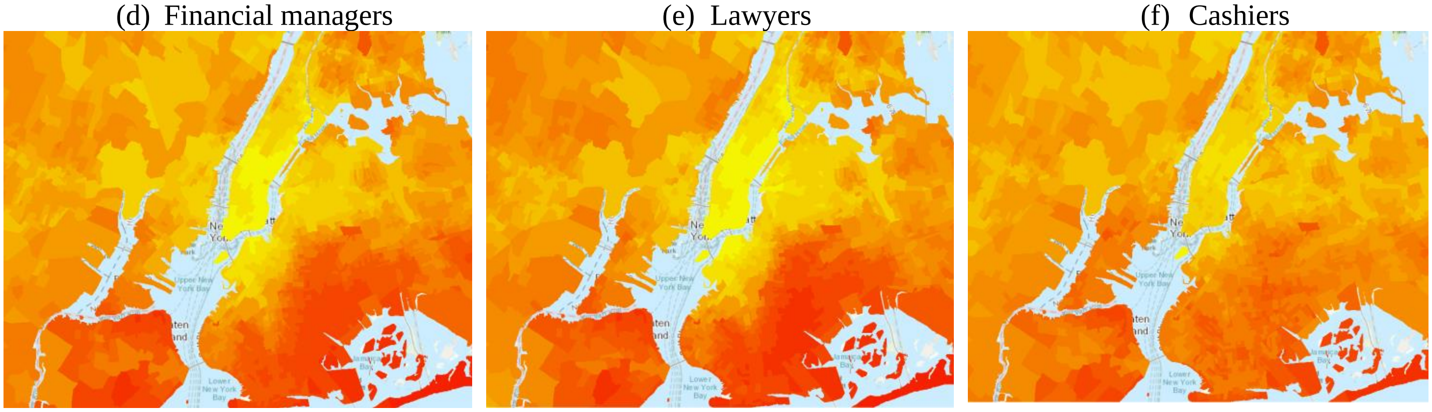
Figure A13: Expected commute time for selected occupations in Chicago MSA.

(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.

Table A1: Robustness tests with alternative model specifications

		Commute only	Alternative value of time measure		2010 job location	Alternative definition for high-skilled workers		Controls for initial skill ratio
			Residual log earnings dispersion	Pct of long hour		Col>30%	Col>50%	
Panel A: Worker's location demand								
Commute cost (μ)	High-skilled occupations	21.947***	0.175***	0.0451	7.119***	3.765***	5.727***	9.787***
		(1.248)	(0.0156)	(0.0352)	(1.088)	(1.171)	(0.860)	(0.924)
	Low-skilled occupations	2.777***	0.095***	0.215***	1.657***	3.002***	5.491***	2.228***
		(0.453)	(0.0124)	(0.0209)	(0.468)	(0.499)	(0.420)	(0.453)
Amenity (γ)	High-skilled occupations	-	-	-	2.347***	2.900***	1.469***	1.156***
					(0.209)	(0.274)	(0.178)	(0.069)
	Low-skilled occupations	-	-	-	0.747***	1.257***	0.521***	0.361***
					(0.148)	(0.201)	(0.158)	(0.085)
Rent (β)	High-skilled occupations	-	-	-	0.762***	1.255***	-0.482***	-0.0415
					(0.253)	(0.344)	(0.172)	(0.102)
	Low-skilled occupations	-	-	-	0.476***	0.780***	-0.341**	0.268**
					(0.196)	(0.267)	(0.164)	(0.1212)
Panel B: Rent								
Housing demand× housing stock density (π_1)		-	-	-	1.035***	0.295***	0.232***	0.0549***
					(0.128)	(0.0411)	(0.0314)	(0.0201)
Housing stock density (π_2)		-	-	-	-0.0954***	0.0022	-0.0065	0.0342***
					(0.0204)	(0.0069)	(0.0052)	(0.0041)

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 the change in occupation-specific expected commuting time, and I allow the coefficients on these controls to vary by occupation. Conley (1999) HAC standard errors are computed with 1-mile threshold for the kernel function. Column 1 shows results from estimation with only commuting cost (long-hour premium). Column 2 – 3 show results from estimation with various measures of value of 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%). Column 7 shows results from estimation using initial skill ratio as an additional control (in addition to the centrality measure – total expected commute and change in expected commute).

Table A2: Estimate of long-hour premium with or without controls for education

Occupation name	code	Without educ. control LRP - 1990	Without educ. control LRP - 2010	Without educ. control Δ in LRP	With educ. control LRP - 1990	With educ. control LRP - 2010	With educ. control 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 A3: 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.