Federal Reserve Bank of Dallas

The Geography of Jobs and the Gender Wage Gap

Sitian Liu and Yichen Su

Working Paper 2028 **October 2020**

Research Department

https://doi.org/10.24149/wp2028

The Geography of Jobs and the Gender Wage Gap*

Sitian Liu† and Yichen Su‡

September 8, 2020

Abstract

Prior studies have shown that women are more willing to trade off wages for short commutes than men. Given the gender difference in commuting preferences, we show that the wage return to commuting (i.e., the wage penalty for reducing commute time) that stems from the spatial distribution of jobs contributes to the gender wage gap. We propose a simple job choice model, which predicts that differential commuting preferences would lead to a larger gender wage gap for workers who face greater wage returns to commuting based on their locations of residence and occupations. We then show empirical evidence that validates the model's prediction. Moreover, we estimate the model components: (i) the indifference curves between wages and commutes by gender, and (ii) the wage return to commuting faced by each worker. Our model shows that differential commuting choices account for about 16-21% of the gender wage gap on average, but the contribution varies widely across residential locations. The model also shows that policies that increase commute speed or density in the central city neighborhoods could moderately lower the gender wage gap.

Keywords: Gender wage gap, commuting, spatial distribution of jobs

JEL Codes: J16, J22, J31, R12, R41

* The views expressed in this article are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Dallas or the Federal Reserve System. Please click here to download the latest version.

† Sitian Liu, Department of Economics, Queen's University. Email: sitianliu@econ.queensu.ca.

‡ Yichen Su, Federal Reserve Bank of Dallas. Email: yichensu@outlook.com.

1 Introduction

The U.S. has seen a significant convergence of the gender wage gap since the 1960s (Blau and Kahn, 2017).[1] Despite the immense progress, a significant gender wage gap still remains today. By 2010, conventional human-capital factors can explain little of the gender wage gap (Blau and Kahn, 2017).

An increasing number of studies have examined the role of non-wage amenities, such as temporal flexibility, in explaining the remaining gender wage gap (Bertrand et al., 2010; Goldin and Katz, 2011; Mas and Pallais, 2017; Wiswall and Zafar, 2017; Wasserman, 2019). The literature has shown that female workers are more willing to give up higher wages in exchange for jobs that offer more non-wage amenities, such as flexible and predictable hours, compared with male workers. Because these amenities often come with wage penalties, the gender difference in preferences for non-wage amenities could lead to a gender wage gap (Goldin, 2014).[2]

This paper studies how commute time, as another source of non-wage amenities, contributes to the gender wage gap. A recent study by Le Barbanchon et al. (2019) shows that compared with men, women are more likely to give up higher wages in exchange for a shorter commute. This finding highlights commute time as a potential factor that could lead to a gender wage gap. However, we argue that the gender difference in commuting preferences alone is not sufficient to create a gender wage gap. In particular, if the wage penalty for choosing a shorter commute is negligible, then no significant gender wage gap would arise even if commuting preferences are different by gender. In contrast, if reducing commute time is associated with with a sizable wage penalty, while male workers may choose to bear a long commute to avoid the wage penalty, female workers may choose to give up a high-wage job for a job with shorter commute time.[3] In our paper, we show that the size of the wage penalty for reducing commute time is a determinant of the gender wage gap.

While the size of the wage penalty for typical non-wage amenities such as temporal flexibility varies by job, the wage penalty for reducing commute time is determined by the *geography* of job

[1] After the 1990s, the convergence continued but was much slower.

[2] Goldin (2014) shows that many job positions are rewarded disproportionately for the extremely long or unpredictable working hours. The disproportionate reward can be viewed as a form of compensating differential. This means that one has to take a wage penalty to take job positions that have more moderate working hours (as a form of non-wage amenities).

[3] Similarly in Goldin (2014), the size of the wage penalty for offering job flexibility is a key determinant of the gender wage gap across occupations. For lawyers, the wage penalty for flexibility is high, and thus female lawyers have to choose between lower wages or inflexible hours. Since female lawyers are more likely to give up inflexible hours, this results in a significant gender wage gap. In contrast, for pharmacists, the wage penalty for flexibility is negligible. Although female pharmacists may also be willing to give up higher wages for more job flexibility, they do not face such a wage/flexibility trade-off. Thus, the pharmacist profession sees a significantly smaller gender wage gap.

locations and workers' residential locations, and can vary by city and the neighborhood in which workers live.[4] If high-wage jobs are located near workers' residential locations, workers are not likely to face a large wage penalty for choosing a nearby job. Therefore, differential commuting preferences are not likely to lead to a gender wage gap. In contrast, if high-wage jobs are located far from workers' residential locations, workers are likely to face a large wage penalty for choosing a shorter commute. Since female workers could be more likely to give up far-away high-wage jobs for shorter commutes, the gender difference in commuting preferences could lead to a gender wage gap. Therefore, the geography of jobs and workers is an essential ingredient in understanding the size of wage penalties faced by workers and the gender wage gap.

To illustrate how commuting preferences and the geography of jobs can affect the gender wage gap, we build a job choice model in which workers trade off between wages and commute time. In the model, workers living in different residential locations face different job choice sets, defined with a two-dimensional wage-commute space. The wage return to commuting (equivalent to the wage penalty for reducing commute time) is higher for workers living farther from high-wage jobs, and lower (or close to zero) for workers living closer to high-wage jobs.[5] The model shows that, given the gender difference in commuting preferences, higher wage returns to commuting lead to a greater gender wage gap.[6]

Consistent with the prediction of the model, we document that there is substantial variation in the gender commute and wage gaps across locations within cities. In particular, the gaps are typically smaller for workers living in the central city, where many high-wage jobs are located. In contrast, the gaps are larger on average for workers living in the suburbs.[7] Alternatively, instead of considering the distance to the city center, we show that both commute and wage gaps are larger in neighborhoods where the average commute time is longer, suggesting that the gaps are likely to be greater in places with larger wage returns to commuting. To further validate the model, we show that for workers in occupations with a strong geographic concentration of high-wage jobs in a central city (e.g., management and business), the gender commute gap increases substantially with

[4]More precisely, the strength of the transportation infrastructure and the degree of congestion are also considered in the geography of jobs and workers.

[5]The wage return to commuting measures how much the highest wage faced by a worker can increase if the worker chooses to commute longer. It is equivalent to the wage penalty for choosing a shorter commute.

[6]Intuitively, if men have lower disutility of commuting, they are more likely to commute longer than women to reap the wage benefit, in response to a higher wage return to commuting.

[7]Later we confirm that the wage return to commuting is on average smaller for workers living closer to the city center and greater for workers living farther from the city center.

the distance to the central city. In contrast, for workers in occupations in which high-wage jobs are evenly distributed across locations (e.g., healthcare practitioners), we do not find such a positive correlation. We also find that the gender wage gap is highly correlated with the gender commute gap, controlling for observable characteristics.

Motivated by the descriptive evidence, we quantify the job choice model by estimating the model parameters: (i) the differential willingness to trade off between wages and commute time by gender, and (ii) the wage return to commuting by occupation and residential location. Using the quantitatively estimated model, we evaluate how much the spatial distribution of jobs and the gender difference in commuting preferences contribute to the observed gender gaps in commutes and wages.

We estimate men and women's willingness to trade off between wages and commute time by tracing the slope of indifference curves between wage and commute time using micro-data from the American Community Survey (ACS). Exploiting the theoretical prediction that all of the observed job bundles chosen by workers should locate above the indifference curve associated with workers' reservation utility (hereafter, reservation curve), we use quantile regressions to estimate the slope of the reservation curves.

Consistent with prior studies, we find that women have stronger disutility for commuting. On average, women are willing to give up 8.4% of wages in exchange for a 10-minute shorter one-way commute, compared with 6.31% for men. The gender difference in commuting preferences is greatest in married workers with children. In robustness tests, we also find such gender differences in commuting preferences for workers of different education levels and in different metropolitan statistical areas (MSAs).

We measure the wage return to commuting faced by each worker in a given residential location and occupation using data on the spatial distribution of jobs and wages from the ACS and ZIP Codes Business Patterns. Due to data limitations, we only focus on the New York MSA. For each residential Public Use Microdata Area (PUMA) and occupation, we simulate available jobs and wages according to the implied spatial distributions estimated from the data. Based on the simulated job choice set for each residential location and occupation, we measure the wage return to commuting using a nonparametric frontier estimator proposed by Cazals et al. (2002). We find substantial variation in the wage return to commuting across occupations and residential locations in the New York MSA. On average, increasing commute time by 10% is associated with an increase in the highest expected wage offer by 5.2%.

The estimated model predicts a gender commute gap of 0.0368 log points in the New York MSA, which explains 77% of the observed gender commute gap (0.0481 log points). The model-predicted gender wage gap is 0.0273 log points, which is 20.78% of the observed gender wage gap (0.1314 log points). In other words, differential commuting choices by gender could explain around 21% of the gender wage gap in the New York MSA. We show that the geography of jobs and the size of wage returns to commuting do matter. First, by equalizing wage returns to commuting across locations to the median (or mean) wage return, the model implies a substantially smaller gender wage gap. Second, we show that by reducing the wage return to commuting, we can further reduce the gender wage gap, even holding the gender difference in commuting preferences unchanged.

Lastly, we conduct several counterfactual analyses in which we evaluate various real-world changes that could alleviate workers' commuting burden, and assess the impact of such changes on the gender wage gap. We show that lowering commute time (e.g., through increasing travel speed) by up to 80% could only moderately reduce the gender wage gap, by 0.0084 log points, which is 6.4% of the observed wage gap. This is mainly because the wage return of commuting (i.e., the wage penalty for reducing commute time) does not fall proportionally with the reduction of the overall commute time. Moreover, we show that by reassigning 20% of the MSA's population from the suburbs (outside the 10-mile radius of city center) into central city neighborhoods (within the 10-mile radius), which is equivalent to increasing central city density by 56.15%, we can reduce the gender wage gap by 0.0041 log points, which is 3.1% of the observed wage gap.

These moderate effects suggest that the trade-off between wages and commutes is likely a lasting feature in the labor market, even as transportation infrastructure improves and housing supply increases in central cities.[8] As society increasingly smooths out wage penalties for other non-wage attributes in the labor market, such as temporal flexibility, the trade-off between wages and commutes may account for an increasingly larger portion of the gender wage gap going forward. An alternative way to reduce the gender wage gap could be reducing female workers' disutility of commuting by providing better childcare programs or making remote working options more widely available.

A large body of literature has assessed the contributing factors behind the gender wage gap (Blau and Kahn, 2017). Since the 1980s, traditional human capital factors have become less important

[8] Dispersing high-paying jobs spatially may also reduce pay gap, but it may have a negative productivity consequences due to the loss of the agglomeration effects.

in explaining the gender wage gap, due to the reversal of the educational gap and narrowing of the experience gap between men and women. In particular, our work is built upon the studies that investigate the gender difference in the willingness to trade off wages for non-wage job attributes and how the differential preferences affect equilibrium job outcomes by gender (Becker, 1985; Cha and Weeden, 2014; Goldin, 2014; Mas and Pallais, 2017; Wiswall and Zafar, 2017; Cubas et al., 2019).

Within this strand of literature, Gutierrez (2018) and Le Barbanchon et al. (2019) are the most closely related to our paper. Gutierrez (2018) is the first to document a gender difference in commuting patterns in the U.S. using the ACS, and relate such difference to intra-household earning disparities between husbands and wives. Le Barbanchon et al. (2019) use French administrative data on job search criteria for unemployed workers to show that female workers are more likely to trade off wages for shorter commute time.[9]

Our paper contributes to the literature by showing that differential gender preferences for commutes do not necessarily lead to a gender wage gap if the wage return to commuting (i.e., the wage penalty for reducing commute time) faced by workers is low. We highlight that size of the wage return to commuting is determined by the geography of jobs and workers. It is the *combination* of commuting preferences and the spatial distribution of jobs that determines the gender wage gap. The geography of jobs and commuting patterns is a central feature of our analysis, whereas previous studies—e.g., Le Barbanchon et al. (2019)—do not feature a spatial analysis explicitly.

Lastly, we contribute to the understanding of the linkages between spatial activities, transportation, and labor market outcomes (Becker, 1965; Small et al., 2005; Small and Verhoef, 2007; Ahlfeldt et al., 2015; Kreindler, 2018; Monte et al., 2018; Severen, 2019; Tsivanidis, 2019). The urban economics literature has examined the cost of spatial distance between jobs and workers as a form of trade cost/friction. Our paper highlights another consequence of such cost: the gender wage gap. Our results suggest that besides lowering the overall efficiency and welfare of the economy, the wage-commute trade-off created by the geography of jobs and workers also matters for gender equity.

The rest of the paper is organized as follows. Section 2 describes the model. Section 3 describes the data and provides descriptive facts. The estimation strategy of the model parameters is outlined in Section 4. Results are presented and discussed in Section 5. Section 6 analyzes the role of

[9]On the extensive margin, Black et al. (2014) show that commute time faced by workers contributes to the gap in the labor force participation rate between male and female workers. Another relevant study is by Lundborg et al. (2017), who find that women, after having children, work much closer to home.

commuting choices in the gender wage gap using the estimated model. Section 7 conducts two counterfactual analyses, and Section 8 concludes.

2 Model

We present a simple static job choice model to illustrate the mechanism through which commuting preferences and the geography of jobs can affect the gender wage gap. Workers have preferences over a residential location-specific choice set of jobs, given their residential location. A job is characterized by the wage and commute time, and the commute time is determined by the residential location of a worker and the location of the job.[10]

2.1 Job Choice Set

Given the residential location of a worker, an available job consists of two components: commute time and wage, (τ, w). τ is determined by the relative locations of the worker's home and the job. Each worker chooses one job out of the job choice set available to her.

To highlight the trade-off between wages and commute time, we assume that each worker i faces a log linear relationship between wages and commute time of jobs on the job choice set frontier:

$$\ln(w) = \xi_i + \beta_i \ln(\tau - \tau^{min}),\ \tau \geq \tau^{min},$$

where τ^{min} is the minimum commute time;[11] ξ_i captures individual productivity or ability, as well as i's best wage offer close to her residential location. This specification allows us to summarize the marginal wage return to commuting with a single parameter β_i, which is equivalent to the wage penalty for reducing commute time. Different workers face different wage returns to commuting, β_i, which is determined by the geography of jobs relative to the residential location of worker i. More specifically, if i lives in the suburbs and high-wage jobs are concentrated in the central city, then β_i

[10]We do not explicitly feature residential location choice in our model, mainly for simplicity. To empirically estimate each model component, we always include location fixed effects interacted with personal characteristics to account for location sorting. In Appendix A4, we write down an alternative job choice model in which spatial sorting is explicitly featured. We do demonstrate that spatial sorting could become important under some conditions.

[11]τ^{min} captures the fixed time cost of commuting faced by each commuter. Unless a person works from home, commute time can rarely be reduced to zero. Activities such as leaving home, going to the garage, starting up and heating up the car, and driving off the parking lot often take some time. In the model, a worker would optimally commute τ^{min} when β is zero. In our empirical specification, we set τ^{min} to 5 minutes.

is likely to be large. Conversely, if i lives in the central city, then β_i can be close to zero.

2.2 Worker

Each worker i has an additive utility function:

$$U(\tau, w) = \ln(w) - \lambda_i \tau,$$

where (τ, w) represent the commute time and wage of a job. Workers value higher wages but dislike longer commutes.[12] λ_i measures the degree to which worker i dislikes commuting relative to lower wages. We allow λ to differ by gender, $\lambda \in \{\lambda_m, \lambda_f\}$.[13] Let $U_i^R \in \mathbb{R}$ be the reservation utility: A worker chooses to work if and only if the corresponding level of utility is greater than U_i^R.

2.3 Equilibrium Job Choice

Each worker i chooses a job to maximize utility subject to the constraint of job availability:

$$\begin{aligned}
&\max_{\tau,w} \quad U(\tau, w) \\
\text{s.t.}\quad &\ln(w) = \xi_i + \beta_i \ln(\tau - \tau^{min}) \text{ if } \tau > \tau^{min}, \\
&\ln(w) = \xi_i \qquad\qquad\qquad\quad \text{ if } \tau = \tau^{min}.
\end{aligned}$$

The first-order condition yields

$$\tau_i^* = \begin{cases} \frac{\beta_i}{\lambda_i} + \tau^{min}, & \text{if } \beta_i > 0 \\ \tau^{min}, & \text{if } \beta_i = 0. \end{cases}$$

[12]Even though we allow λ to differ by gender, we do not imply that the gender difference in commuting preferences is necessarily rooted in any innate difference in the taste for commuting between men and women. Such a difference in preferences could be caused by a number of factors, such as social norms in the division of labor within households, the availability of childcare services, etc. Here, we do not attempt to dissect the reasons behind the gender difference in commuting preferences. Rather, we treat the differential preferences as "revealed preferences" and focus on the labor market consequence of the difference in the "revealed preferences."

[13]Gender differences in the preference for job amenity do not affect the prediction of the model as long as job amenity is not correlated with commutes. Specifically, suppose men and women trade off between wages and job amenity (A) differently: $A = a_i w$, $a_i \in \{a_m, a_f\}$. Suppose the utility function is $U(\tau, w, a) = \ln(w + A) - \lambda_i \tau$. Then the gender difference in the preference for job amenity only affects the level of the indifference curves, not the slope.

$$\ln(w_i^*) = \begin{cases} \xi_i + \beta_i \ln\left(\frac{\beta_i}{\lambda_i}\right), & \text{if } \beta_i > 0 \\ \xi_i, & \text{if } \beta_i = 0. \end{cases}$$

When $\beta_i > 0$ (i.e., there is a trade-off between wages and commutes), workers commute less if they dislike commuting more and commute more if the wage return to commuting is higher.

To understand how gender differences in λ_i contribute to the gender commute and wage gaps, we first consider what happens to the optimal commute time when a worker's disutility of commuting increases. We focus on the case in which $\beta_i > 0$:

$$\frac{\partial \ln(\tau_i^*)}{\partial \lambda_i} = -\frac{\beta_i}{\lambda_i (\beta_i + \lambda_i \tau^{min})} < 0.$$

If women have higher disutility of commute time than men, the optimal commute time for women is shorter than that of men. Similarly, the effect of λ on the equilibrium wage is

$$\frac{\partial \ln(w_i^*)}{\partial \lambda_i} = -\frac{\beta_i}{\lambda_i} < 0.$$

In addition, the magnitude of the effects depends on the wage return to commuting (β_i):

$$\frac{\partial^2 \ln(\tau_i^*)}{\partial \lambda_i \partial \beta_i} = -\frac{\tau^{min}}{(\beta_i + \lambda_i \tau^{min})^2} < 0;$$
$$\frac{\partial^2 \ln(w_i^*)}{\partial \lambda_i \partial \beta_i} = -\frac{1}{\lambda_i} < 0.$$

The impact of β_i on $\ln(\tau_i^*)$ is greater if λ_i is smaller. Intuitively, if the utility cost of long commutes is relatively mild, then workers will commute longer to reap the wage benefit should β_i be large. However, if the utility cost of long commutes is high, workers must weigh the benefit of long commutes against the utility cost. Therefore, a higher λ makes τ_i^* less responsive to β_i.[14]

The gender gaps in commute time and wages due to preference differentials can also be expressed

[14] Alternatively, we can consider what happens to the optimal commute time when the wage return to commute time increases, holding λ constant:

$$\frac{\partial \ln(\tau_i^*)}{\partial \beta_i} = \frac{1}{\beta_i + \lambda_i \tau^{min}} > 0 \text{ if } \beta_i > 0.$$

The optimal commute time is longer if the wage return to commuting is higher. However, larger λ_i dampens the effect of β_i on commute time.

with a first-order approximation:

$$\ln(\tau_m^*) - \ln(\tau_f^*) = \frac{\beta_i(\lambda_f - \lambda_m)}{\lambda_m\left(\beta_i + \lambda_m \tau^{min}\right)}; \tag{1}$$
$$\ln(w_m^*) - \ln(w_f^*) = \frac{\beta_i(\lambda_f - \lambda_m)}{\lambda_m}.$$

Crucially, a higher β_i increases gender gaps in commutes and wages. This is because if $\lambda_f > \lambda_m$, men's commute time is more responsive to β_i, and therefore the gender commute gap is greater in locations where workers face larger β_i.

2.4 Graphical Analysis

We use a series of graphical illustrations to further demonstrate the intuition of the model. In Figure 1, the upward-sloping curves represent indifference curves for a representative male worker and a representative female worker. The indifference curves are upward sloping because workers prefer higher wages and shorter commutes. If we assume $\lambda_f > \lambda_m$, then the woman's indifference curves are steeper than the man's indifference curves. Suppose they face the same job choice set. The slope of the job choice set frontier captures the wage return to commuting (i.e., the wage penalty for reducing commute time). The points below the job choice set frontier represent jobs available to the workers. The optimal job choice of a worker is given by the tangent point of her indifference curve and the frontier. Different slopes of the indifference curves between the man and the woman lead to different optimal job choices.

Figure 2 illustrates how the steepness of the job choice set frontier (i.e., the wage return to commuting or the wage penalty for reducing commute time) affects the gender gap in commute time and wages, holding the preferences fixed. Figure 2 shows a job choice set with a lower wage return to commuting compared with Figure 1. Figure 2 shows that the lower wage return to commuting leads to a smaller gender commute gap, and therefore a smaller gender wage gap.

2.5 Model Implications

The model suggests that the equilibrium gender commute and wage gaps are jointly determined by gender-specific preferences for commuting (λ_m and λ_f) and the wage return to commuting faced by each worker (β_i). The magnitude of β_i depends on both where the worker lives and the spatial distribution of jobs and wages of her occupation.

Below are some simple predictions of the model that can be tested in the data:

1. Gender wage and commute gaps should both be smaller near city centers, where high-wage jobs are disproportionately located.

2. Gender wage and commute gaps should be smaller near city centers for workers whose occupations have a high concentration of high-wage jobs in the city centers (e.g., financial jobs), while these gaps should be unrelated to distance to city centers for workers whose occupations do not exhibit such a geographic pattern.

3. Gender wage gaps should be spatially correlated with commute gaps. In other words, locations with larger commute gaps should be likely to see larger wage gaps.

In the next section, we show that the model's predictions are borne out by the data.

3 Data and Empirical Evidence

3.1 Data

We use data from two sources. First, we use the American Community Survey (ACS) 2013–2017 5-year sample from the Integrated Public Use Microdata Series (IPUMS) (Ruggles et al., 2019). The data provide individual- and household-level information on various demographic and socioeconomic characteristics, such as gender, race, marital status, education attainment, employment status, working hours, occupation, and income. Importantly, the ACS provides workers' self-reported commute time. The ACS also provides residential location at the level of Public Use Microdata Area (PUMA), and working location at the level of Place-of-Work PUMA (PWPUMA).[15] The PWPUMA variable is crucial for measuring the spatial wage distribution by work locations.

Second, we use the ZIP Codes Business Patterns (ZCBP) provided by the U.S. Census Bureau. The ZCBP provides the count of business establishments and employment sizes for each NAICS industry code at ZIP-Code level. We create a crosswalk between industry and occupation using the ACS, and impute the count of jobs by occupation and ZIP Code using the crosswalk.[16]

[15] PWPUMA roughly coincides with county.

[16] Appendix A2 includes more details on construction of the crosswalk and the imputation.

Table 1 provides summary statistics from the ACS. We restrict our sample to full-time (35+ hours) workers, aged between 25 and 69. Women in the sample are less likely to be married with spouse present and less likely to have a young child compared with male workers. Moreover, on average, women are better educated than men. The average *one-way* commute time is 26 minutes for men and 23 minutes for women. The average hourly wage is $30 for men and $24 for women. Gender differences in the average commute time and the average hourly wage are statistically significant.

A potential concern about using the self-reported commute time is that it may not reflect the actual amount of time people spend on their way between between the residence and the workplace. For instance, women might be more likely to stop at a grocery store than men on their way home from work, which could complicate the estimate of commute time. In Appendix A1, we use the 2017 National Household Travel Survey to show that including the amount of time spent on such chained trips does not affect gender difference in the log commute time.

Figure 3 presents scatter plots of gender differences in log commutes and log hourly wages without and with controls by age for full-time workers. In particular, Figure 3b shows that the life-cycle patterns between the gender commute and wage gaps are very similar, after controlling for demographic characteristics, occupation indicator, year fixed effects, MSA fixed effects, and the interaction between MSA and occupation indicators. The difference between the residual commute gap and the residual wage gap is around 0.189 log points at all ages from 25 to 69.

3.2 The Geography of Gender Commute and Wage Gaps

The model predicts that the gender commute gap is greater for workers who live farther from the location of high-wage jobs (Equation 1). Figures 4 and 5 map gender differences in log commute time and log hourly wage within four major MSAs by PUMA.[17] We find substantial variation in the gender commute and wage gaps across locations within MSAs. In particular, the gender gaps in both measures are typically smaller in the central city, where high-wage jobs tend to locate.

To show this geographical pattern of the gender commute and wage gaps more clearly, we use binned scatter plots of gender differences in log commute and log hourly wage by residential PUMA based on the nationwide sample (Figures 6a and 6b). To construct the figures, we bin residential

[17] We focus on Chicago, San Francisco, Boston, and New York MSAs because there are a relatively large number of PUMAs within these metropolitan areas. The patterns remain similar by using log commute time residual and log hourly wage residual. They are not presented in the paper due to space constraints.

PUMAs into 20 equal groups based on the distance to the city center of the MSA and control for MSA fixed effects.[18] The results suggest that the gender commute and wage gaps are larger for workers who live farther from the city center, where high-wage jobs tend to locate. More specifically, a 10-mile increase in the distance to the city center is associated with an increase in the gender commute (wage) gap by 0.028 (0.026) log points. Given that the average commute gap is 0.08 log points and the average gender wage gap is roughly 0.19 log points, the spatial variation is substantial.

According to the model, the gender wage gap results from the difference in commuting choices between men and women. Therefore, a place with a larger gender commute gap should also see a larger gender wage gap. Figure 7a presents a binned scatter plot of the log commute gap versus the log wage gap by residential PUMA. The figure suggests that a 1% increase in the gender commute gap is associated with a 0.38% increase in the gender wage gap.

However, the positive association between the gender commute (or wage) gap and the distance to the city center shown in Figures 6a and 6b, and the positive association between the gender commute gap and gender wage gap shown in Figure 7a, could be driven by spatial sorting of household characteristics unrelated to commuting preferences or the location of jobs. For instance, if married households are more likely to live in the suburbs than unmarried households, and married women are more likely choose lower-paying jobs and cut back on commuting time to enjoy greater job flexibility for household tasks and childcare, then we would obtain a positive correlation between the gender wage gap and distance to the city center. However, the positive correlation may not necessarily be driven by the geography of high-wage jobs. Therefore, the positive correlation between the gender commute gap and gender wage gap shown in Figure 7a could be driven by variation in the demand for job flexibility, not variation in the trade-off between commutes and wages.

3.2.1 Spatial Sorting by Observable Characteristics

To test whether the positive correlations found in Figures 6a, 6b, and 7a are driven by spatial sorting, we calculate the residual gender gaps in log commute time and log hourly wages for each PUMA location, controlling for observable characteristics and various fixed effects. More specifically, we

[18] To construct the figures, we restrict the sample to PUMAs of which the distance to downtown is below the 95th percentile, which is around 52 miles. The median distance to downtown for all PUMAs is 15 miles, and the mean is 20 miles.

estimate following regression equation:

$$y_{ijt} = \delta_j^m + \delta_j^f + \gamma \mathbf{X}_{ijt} + \varepsilon_{ijt},$$

where y_{ijt} is the log commute time or log hourly wage of worker i living in PUMA j in year t; δ_j^m is male-specific residential-location fixed effect; δ_j^f is female-specific residential-location fixed effect; $\mathbf{X}_{ijt}$ is a vector of of observable characteristics, including dummies for age, marital status, whether having a child younger than 18, race, Hispanic origin, education, occupation, year, and MSA. Therefore, $\hat{\delta}_j^m - \hat{\delta}_j^f$ is the estimated gender difference in commute residuals or wage residuals in residential PUMA j. We use $\hat{\delta}_j^m - \hat{\delta}_j^f$ to mitigate the concern about spatial sorting by observable characteristics.

Figures 6c and 6d present binned scatter plots of gender differences in log commute residual and log hourly wage residual by distance to the city center. Estimates of the slope of the fitted regression lines are still positive and statistically significant. The results suggest that increasing the distance to the city center by 10 miles is associated with an increase of 0.023 log points in both the residual commute gap and the residual wage gap, which again confirms the model's prediction.

One might still wonder whether the distance to city center is truly a good measure of the return to commuting. To address this concern, we alternatively use the average commute time of male workers for each residential PUMA to approximate the return to commuting faced by workers.[19] Figure 8 presents binned scatter plots of gender differences in residual log commute time and residual log wages by residential PUMA, in which residential PUMAs are binned into 20 equal groups based on the average commute time. The figures suggest that the gender commute and wage gaps are larger in places where workers' average commute time is longer.

Figure 7b presents a binned scatter plot of the residual log commute gap versus the residual log wage gap by residential PUMA, controlling for distance to the city center. We still find a positive correlation between the two gaps—a 1% increase in the gender commute gap is associated with a 0.24% increase in the gender wage gap. This shows that the commute gap and wage gap are spatially correlated, which validates another of our predictions. It is noteworthy that there is still a gender wage gap of 0.16 log points in places where the gender commute gap is zero, which is 16% lower than

[19]This idea comes from the fact that a worker would choose a long commute only if the return to commuting is high. Therefore, the average commute time for male workers should be indicative of such return.

the average residual gender wage gap. This simple calculation based on the descriptive statistics suggests that different commuting choices between men and women could account for approximately 16% of the gender wage gap.

3.2.2 The Gender Commute Gap by Occupation

One may still wonder whether controlling for the observables can eliminate all of the confounding effects of spatial sorting. To further address the concern, we examine the spatial patterns of the gender commute gap by occupation group, which differs in the spatial distribution of jobs. Different spatial patterns of the gender commute gap by occupation could provide further evidence that spatial variation in gender differences in commuting choices is likely a result of the geography of jobs, not spatial sorting of household types.

For some occupations, such as management and business, high-wage jobs are disproportionately concentrated in the central business districts of cities; for other occupations, such as healthcare practitioners, high-wage jobs are more evenly distributed across locations. The model predicts that for workers in occupations in which jobs are geographically concentrated in city centers, the gender commute gap should increase with distance to city centers. In contrast, for workers in occupations in which jobs are geographically evenly distributed, there should not be such a positive correlation.

Figure 9 presents binned scatter plots of gender differences in log commute residual by residential PUMA for managers and healthcare practitioners, respectively.[20] Consistent with the model prediction, Figure 9a shows a positive correlation between the gender commute gap and distance to city centers for managers, while Figure 9b shows a negative correlation for healthcare practitioners, such as dentists. The results suggest that the spatial patterns of the gender commute gap are not likely to be solely driven by spatial sorting, because if so, we would not have found different spatial patterns across occupations, given that workers in the selected occupation groups are likely to have similar skill and income levels, and therefore similar tastes for locations on average.

[20] "Management and business" refers to occupations whose codes are between 10 and 950 based on the ACS 2010 occupation classification. "Healthcare practitioners" refers to occupations whose codes are between 3000 and 3540.

4 Estimating the Job Choice Model

After validating the model with empirical evidence, we use the model to quantify the effect of the geography of jobs on the gender wage gap. To evaluate the model quantitatively, we need to estimate two sets of parameters: (i) gender-specific preferences for commuting, λ, and (ii) the wage return of commuting faced by each worker, β. Graphically, λ represents the slope of an indifference curve, and β represents the slope of the job choice set frontier.

4.1 Estimation of Commuting Preference λ

We first estimate the slope of workers' indifference curves corresponding to the utility function $U = \ln(w) - \lambda_s \tau,\ s \in \{m, f\}$. The goal is to trace workers' trade-off between commute time and wages at the same utility level. If workers are identical and can freely move across locations, the level of utility achieved in optimum should be equalized across locations for every worker. Under this assumption, the observed job bundles $(\tau, \ln(w))$ should locate on the same indifference curve. If so, regressing log wages on commute time should yield an unbiased estimate of λ_s.

However, the assumption of equalized utility is highly simplified and unlikely to hold in the data. If workers are heterogeneous or if job opportunities that are available for identical workers differ for idiosyncratic reasons, a simple regression would yield biased estimates. Below, we discuss two reasons that could lead to biased estimates of λ_s from simple regressions.

Location sorting by ability If workers with different levels of ability sort into neighborhoods with different distances to high-wage jobs, the observed commute-wage bundles may locate on different indifference curves and a simple regression would yield a biased estimate of the slope of indifference curves. For example, if workers with higher ability choose to live in the suburbs far from high-wage jobs for some unobserved reasons, then higher wages are associated longer commute time even in the absence of the commute-wage trade-off. Thus, a simple regression could overestimate λ_s.[21] Alternatively, if high-ability workers sort into neighborhoods close to job centers, a regression estimator may suffer from downward bias. Figure 10 provides a graphical illustration of the first case.

[21]High-ability workers may sort into suburban neighborhoods for their high amenities—e.g., school, law enforcement, larger houses, etc.

Random job arrival Another concern is that the observed commute-wage bundles may belong to different indifference curves due to the randomness of job arrivals. To illustrate, consider two workers with identical ability, living in the same location. Figure 11 provides a graphical illustration of the example. Worker i chooses the best available job A on the job choice set frontier. Although two workers have the same ability level, i' receives an additional job B idiosyncratically, which pays as much as A does but is closer to their residential location. In this case, worker i' has a higher utility level than worker i, and therefore these two observations belong to two different indifference curves. Tracing these two bundles together will yield an underestimation of the slope of their indifference curves. Even after controlling for workers' earning ability and location sorting, two observationally equivalent workers may still be on different indifference curves due to the randomness of job arrivals.

4.1.1 The Reservation Utility Indifference Curve

To overcome the difficulty of location sorting by ability or the randomness of wage realization, we estimate the slope of the indifference curve that corresponds to workers' reservation utility (hereafter *reservation curve*). The assumption of this approach is that the slopes of the indifference curves at different wage levels can be approximated by the slope of the indifference curve associated with reservation utility.

We first residualize log hourly wage and commute time for each gender by controlling for various demographic and location fixed effects. In particular, we control for dummies for age, marital status, whether having a child younger than 18, race, Hispanic origin, education, residual PUMA fixed effects, and their interactions with occupation indicators to account for potential location sorting by marriage, ability, or preferences for local amenities. We then estimate the slope of indifference curves, exploiting the theoretical prediction that the observed job choice bundles should locate on or above the reservation curve. We argue that gender differences in preferences for job amenity are not likely to bias the estimate of λ if preferences for amenity are not correlated with commutes (see footnote 13 for more details).

For the moment, we assume that holding demographic and location controls constant, workers have the same reservation utility U^R within the same gender. If there is no measurement error in wages, the observed log wages must lie above the indifference curve corresponding to U^R. Therefore, taking the minimum log wages conditional on commute time should recover the slope of the reservation curve. However, this method is highly sensitive to extreme values or measurement error. For a given level of commute, if an individual reports extremely low wage due to measurement error, the estimated lower boundary will deviate from the true value. To overcome this problem, we use the quantile regression estimator.

4.1.2 Quantile Regression Estimator

Instead of analyzing the conditional mean of log wages as a function of commute time, we use the conditional quantile of log wages with respect to commute time. First, we use a simulation exercise to show that the slope of the reservation curve can be reasonably approximated using quantile regressions at a low percentile, even in the presence of measurement error in wages.

Simulation Exercise We simulate job offers by assigning commute (τ) to log wage ($\ln(w)$) observed from the data, where τ is randomly drawn from a uniform distribution. Assume that there is measurement error in the observed wage: $\ln(w^*) + \eta = \ln(w)$, where w^* is the true wage, η is a normally distributed noise term, and w is the observed wage by researchers. Suppose we know the "true" utility function $U(.)$ and the reservation utility U^R. The "true" reservation curve is shown in Figure 12. A job (w, τ) that can be accepted by workers must satisfy $U(\tau, w^*) \geq U^R$.

Figure 12 also shows binned scatter plots of various percentiles of log wages with respect to commute time, based on acceptable jobs. The reservation curve may lie above some lower quantile curves due to measurement error, and the quantiles depend on the standard deviation of η. However, Figure 12 suggests that the slope of the reservation curve can be reasonably approximated by the slope of a quantile curve at a relatively low percentile. The downside of using quantile regressions with a very low quantile is the lack of statistical power. In our preferred specification, we use the 1st percentile quantile regression estimator, because it is reasonably close to the slope of the true reservation curve in the simulation example, with enough power. We also use other percentiles for robustness checks.

Next, instead of using the simulated data, we present binned scatter plots of various percentiles of log wage residuals with respect to commute residuals using the real data in Figure 13. We show that the quantile curves with lower percentiles are considerably steeper than the conditional mean curve, which is consistent with the findings with the simulated data. This suggests that simply regressing

$\ln(w)$ on τ is likely to underestimate the slope of the indifference curves.

A sufficient assumption for the identification of λ is that holding demographic and location controls constant, workers have the same reservation wage conditional on commute time. Under this assumption, given the observables, variation in wages and commute time represents movement along the reservation curve.

More generally, our identification assumption relies on that (i) the unobserved ability is not correlated with the slope of the reservation utility curve and (ii) the wage return to commuting within each residential location and occupation is not correlated with the slope of the reservation utility curve. We can alleviate both of these concerns by controlling for the observable characteristics described previously, residential location fixed effects, and their interactions with occupation indicators.[22]

4.2 Measurement of the Wage Return to Commuting β

We measure β_i faced by each worker based on the worker's PUMA residential location j and occupation o. We use β_{jo} to denote the wage return to commuting for workers living in PUMA j in occupation o. It is measured based on the upper frontier of the simulated spatial distribution of jobs and their wages by occupation.

4.2.1 Construction of Spatial Wage and Commute Data

We construct a simulated job choice set for workers of each residential PUMA and occupation. To do so, we combine wage data from the ACS with job counts from the ZCBP following three steps:

1. We use the ACS to estimate the means and standard deviations of residual log wages for each occupation and PWPUMA. We assume that wage is multiplicative to workers' ability. The assumption ensures that the spatial distribution of residual log hourly wages is separable from workers' abilities. With this assumption, we can estimate the slope of a job choice set frontier holding workers' ability constant.[23]

[22] Another potential confounding factor in our identification is that intra-household bargaining may affect the location choice of the household or the allocation of housework. Often, workers' labor supply and commuting decisions are made jointly with their spouses. For example, if a man is working full time, the spouse may have a larger λ. Therefore, the labor supply decisions of the spouse may affect a worker' job choices and confound the estimate of λ. To deal with this concern, we control for the spouse's characteristics for married workers. We also estimate λ separately for workers with different types of spouses.

[23] Residual log wages within an occupation and PWPUMA are obtained by controlling for dummies for age, marital status, whether having a child younger than 18, race, Hispanic origin, and education, year fixed effects, and residential PUMA fixed effects.

2. We combine the empirical distribution of log wages for each occupation at the PWPUMA level with the number of jobs at ZIP-Code level from the ZCBP. For each ZIP Code and occupation, we calculate a job share—i.e., the number of jobs in the ZIP Code and occupation divided the total number of jobs of the occupation in the MSA. Then we simulate the job offers at each ZIP-code location for each occupation based on the empirical distribution of the residual log wages for each occupation.

3. We use Google API to compute the commute time between every residential PUMA and every job at ZIP-Code level within the same MSA, assuming that workers only consider jobs within the same MSA.

For workers in each residential PUMA j and occupation o, we construct a job choice set—i.e., commute-wage bundles $\{(\ln(\tau_{jko}), \ln(w_{ko}))\}$, where τ_{jko} is the commute time from PUMA j to job k in occupation o estimated in Step 3, and w_{ko} is the wage of job k drawn from the distribution estimated in Step 1. The simulated job choice set consists of n_{zo} randomly drawn jobs for each ZIP Code z and occupation o based on the above steps, where n_{zo} is equal to 10,000 multiplied by the job share described in Step 2. After constructing the job choice set, we measure the wage return to commuting β_{jo} using a frontier estimator.

Since the job choice set is simulated, we conduct the simulation 20 times for each occupation and each residential PUMA. In each simulation, we randomly draw n_{zo} jobs for each ZIP Code z and occupation o, and estimate β_{jo} based on the simulated data. We finally obtain the measurement for $\hat{\beta}_{jo}$ by averaging over the estimated β_{jo} in each simulation.

4.2.2 Frontier Estimator

We measure β_{jo} by estimating the slope of the frontier of the job choice set faced by workers in residential location j and occupation o.[24] Intuitively, β_{jo} represents the marginal log wage that a worker can gain by increasing her commute time. Therefore, the slope of the job choice set frontier captures the marginal return to commuting. We estimate the frontier of the job choice set simulated for each worker using a nonparametric frontier estimator proposed by Cazals et al. (2002). The

[24] Note that β_{jo} is determined by the location of jobs and the spatial wage distribution of jobs in each occupation. β_{jo} is not influenced by any random components faced by individual workers.

estimator can be robust to extreme values or outliers.[25]

From each worker's perspective, the set of *all available* job bundles is $\Psi = \{(\ln(\tau - \tau_{min}), \ln(w))\} \subseteq \mathbb{R}^2$. The upper boundary of the job choice set is given by

$$\phi\left(\ln(\tau - \tau_{min})\right) = \sup\{\ln(w) | (\ln(\tau - \tau_{min}), \ln(w)) \in \Psi\}.$$

Let $(\ln(w^1), ..., \ln(w^m))$ be m independent identically distributed random variables generated by the wage distribution given $\ln(\tau - \tau_{min}) \leq \ln(\tilde{\tau} - \tau_{min})$. Define the expected maximum wage function of order m denoted by $\phi_m\left(\ln(\tilde{\tau} - \tau_{min})\right)$ as

$$\phi_m\left(\ln(\tilde{\tau} - \tau_{min})\right) = E[\max(\ln(w^1), ..., \ln(w^m)) |\ \ln(\tau - \tau_{min}) \leq \ln(\tilde{\tau} - \tau_{min})],$$

assuming the expectation exists. Intuitively, $\phi_m\left(\ln(\tilde{\tau} - \tau_{min})\right)$ gives the expected highest log wage for m potential jobs in which the worker's log commute time is at least $\ln(\tilde{\tau} - \tau_{min})$. Cazals et al. (2002) show that for any fixed value $\tilde{\tau}$, $\lim_{m \to \infty} \phi_m\left(\ln(\tilde{\tau} - \tau_{min})\right) = \phi\left(\ln(\tilde{\tau} - \tau_{min})\right)$.

To estimate $\phi_m\left(\ln(\tilde{\tau} - \tau_{min})\right)$, consider an i.i.d. sample $\left(\ln\left(\tau_i - \tau_{min}\right), \ln\left(w_i\right)\right)$, $i = 1, ..., n$. The estimator of the expected minimum wage function of order m is defined by

$$\hat{\phi}_{m,n}\left(\ln(\tilde{\tau} - \tau_{min})\right) = \hat{E}[\max(\ln(w^1), ..., \ln(w^m)) |\ \ln(\tau - \tau_{min}) \leq \ln(\tilde{\tau} - \tau_{min})],$$

which can be easily computed in practice. More specifically, let $n(\tilde{\tau})$ be the number of observations of $\ln(\tau_i - \tau_{min})$ less than or equal to $\ln(\tilde{\tau} - \tau_{min})$. For $j = 1, ..., n(\tilde{\tau})$, let $\ln\left(w^{\tilde{\tau}}_{(j)}\right)$ be the jth-order statistic of the observations $\ln\left(w_i\right)$ such that $\ln\left(\tau_i - \tau_{min}\right) \leq \ln\left(\tilde{\tau} - \tau_{min}\right)$:
$\ln\left(w^{\tilde{\tau}}_{(1)}\right) > \ln\left(w^{\tilde{\tau}}_{(2)}\right) > ... > \ln\left(w^{\tilde{\tau}}_{n(\tilde{\tau})}\right)$. Then,

$$\hat{\phi}_{m,n}\left(\ln(\tilde{\tau} - \tau_{min})\right) = \ln\left(w^{\tilde{\tau}}_{(1)}\right) + \sum_{j=1}^{n(\tilde{\tau})-1} \left[\frac{n(\tilde{\tau}) - j}{n(\tilde{\tau})}\right]^m \left(\ln\left(w^{\tilde{\tau}}_{(j+1)}\right) - \ln\left(w^{\tilde{\tau}}_{(j)}\right)\right). \tag{2}$$

Cazals et al. (2002) establish the asymptotic properties of the estimator. In particular, they show

[25] Theoretically, we could employ quantile regressions to recover β. However, quantile regressions require very large amount of data observations. The required number of simulated job offers for the quantile regression to work would have to be unrealistically large. Therefore, instead of using quantile regressions, we employ the method proposed by Cazals et al. (2002), partly because it can produce a frontier estimate using a realistic number of job offers. Another reason is that the frontier estimator is originally proposed to approximate the production frontier, which is conceptually similar to the job choice frontier.

that as m and n grow larger, $\hat{\phi}_{m,n}(\ln(\tilde{\tau} - \tau_{min}))$ approaches $\phi(\ln(\tilde{\tau} - \tau_{min}))$. Choosing a finite m makes the estimator more robust to outliers and measurement error. Therefore, $\hat{\phi}_{m,n}(\ln(\tilde{\tau} - \tau_{min}))$ is a robust and consistent estimator of the maximum log wage available to workers within the log commute time $\ln(\tilde{\tau} - \tau_{min})$. We set $m = 10,000$.[26] Once we obtain the upper boundary of the job choice set using $\hat{\phi}_{m,n}(\ln(\tilde{\tau} - \tau_{min}))$, we use a linear regression to estimate the average slope of the estimated frontier. We measure the wage return to commuting $\hat{\beta}_{jo}$ using the estimated slope.

5 Estimation Results

5.1 OLS Regressions

We first provide OLS reduced-form evidence on gender differences for commuting preference. Table 2 presents OLS estimates of the following equation:

$$\begin{aligned} \ln(w_{ijt}) = & \alpha + \theta Fem_{ijt} + \gamma \tau_{ijt} + \delta Fem_{ijt} \cdot \tau_{ijt} \\ & + X_{ijt}\pi + d_t + d_j + d_o + d_{jo} + \epsilon_{ijt}, \end{aligned}$$

where $\ln(w_{ijt})$ is log hourly wage of individual i living in PUMA j in year t; Fem_{ijt} is an indicator for being female; τ_{ijt} is one-way commute time in hours; X_{ijt} is a vector of individual characteristics, including dummy variables for age, race, Hispanic origin, marital status, whether having a child younger than 18, and education; d_t is year fixed effects; d_j is residential location fixed effects; d_o is occupation fixed effects; and d_{jo} is the interaction between location and occupation indicators.

The results are reported in Table 2. The reduced-form evidence suggests that commute time is positively correlated with wages, and one hour of additional commute time is associated with a greater additional wage for women than for men. Moreover, we find such a gender difference in the relationship between commute time and wages for workers of different education groups and marital status, controlling for observable characteristics and various fixed effects.

[26] We set $m = 10,000$ because the estimate of $\hat{\phi}_{m,n}(\ln(\tilde{\tau} - \tau_{min}))$ does not change significantly if we further increase m.

5.2 Commuting Preferences

Table 3 shows the estimates of λ_s, $s \in \{m, f\}$, using the quantile regression estimator. Column 1 shows that the estimated disutility of commuting is 0.379 for men and 0.506 for women. These results suggest that in order to increase one-way commute time by 1 hour and to maintain the same level of utility, men must be compensated by 37.9% higher wages, whereas women must be compensated by 50.6% higher wages. On average, women's disutility of commuting is roughly 33.5% higher than that of men.

In Columns 2–5 of Table 3, we show the estimates of λ_s by workers' marital status and whether having a child. Column 2 shows that for single workers without children, there is not a significant gender gap in commuting preferences. The estimate of λ_m is slightly greater than the estimate of λ_f, but the difference is not statistically significant. Column 3 shows that for single workers with children, women dislike commuting more than men, but again the difference is not statistically significant. Columns 4–5 show that for married workers, women exhibit stronger disutility of commuting than men. The difference is greatest for married workers with children. In terms of the magnitude, λ_m for married workers is slightly lower than that for single workers, but λ_f for married workers is greater than that for single workers.

Columns 6 and 7 show the estimates of λ_s for workers with college degrees and workers without college degrees, respectively. The magnitude of λ_s is slightly higher for less educated workers, but the gender difference is slightly higher for more educated workers. Overall, the results suggest sizable gender differences in the disutility of commuting for workers in both educational groups.

Table 4 shows the estimates of λ_s for married workers. The estimation is based on the residualized log hourly wage and commute time, while further controlling for the spouse's characteristics, including the spouse's age, race, education, employment status, and occupation. Columns 1 and 2 show a gender difference in λ for married workers either with or without children. The results after controlling for the spouse's characteristics do not change significantly compared with the results in Columns 4–5 of Table 3. The results in Columns 3–4 of Table 4 show that the gender difference in λ is greater for workers with a more educated spouse than workers with a less educated spouse. The results in Columns 5–6 suggest a large gender difference in λ for both workers whose spouse is employed and workers whose spouse is unemployed or not in the labor force. The difference for workers whose spouse is unemployed is not statistically significant, which could partially be because of the small number of observations for this type of couple. Lastly, Columns 7–8 show estimates for workers whose spouses usually work more than 40 hours and workers whose spouses usually work less than or equal to 40 hours per week or do not work. The results suggest sizable and significant gender differences in λ for both types of workers.

We also explore heterogeneity in λ_s across residential locations. This is because workers who dislike commuting more may prefer living in the central city, where more high-wage jobs are concentrated. If women tend to have high disutility of commuting, this location sorting could have a greater effect for them. Table 5 shows the estimates of λ_s by marital status and distance to the central city. For single workers (Panel A), we do not find a discernible gender difference in λ_s, regardless of the residence location. For married workers (Panel B), we find a large and significant gender difference in λ_s. However, either workers' disutility of commuting or the gender difference varies with the distance to the central city.

Lastly, Table 6 shows the estimates of λ_s for workers living in four major MSAs: New York, Chicago, San Francisco, and Boston. The results suggest nonnegligible gender gaps in commuting preferences in all of the selected MSAs, although the statistical power is weaker within each subsample compared with the national sample.

5.3 The Wage Return to Commuting

We estimate β_{jo} for the New York MSA. This is because estimating β_{jo} requires wage data based on job locations, and the ACS provides information on wages and job locations at the PWPUMA level. We focus on the New York MSA because PWPUMAs are typically small and mostly resemble counties in the New York MSA. In contrast, other MSAs, such as Chicago, consist of extremely large PWPUMAs, which prohibit us from detecting accurate variation in spatial wage distributions.

Figure 14 presents the simulated job choice sets faced by workers in different occupations living in different areas of the New York MSA. Figures 14a and 14b show that financial workers living in the northern suburbs face an upward-sloping job choice set frontier, while financial workers living in Midtown Manhattan face a relatively flat job choice set frontier. This is not surprising, since high-wage jobs for financial workers are concentrated in Manhattan. In Contrast, Figures 14c and 14d show that physicians and surgeons living in both the northern suburbs and Midtown Manhattan face relatively flat job choice set frontiers. This is also reasonable since hospitals and clinics tend to be scattered within the MSA.

Figure 15 shows the distribution of the estimates of $\widehat{\beta_{jo}}$ for workers in the New York MSA. The mean is 0.523, the median is 0.489, and the standard deviation is 0.206. In other words, the spatial distribution of jobs implies that increasing commute time by 10% is associated with a 5.2% increase in the highest wage offer. It is noteworthy that $\widehat{\beta_{jo}}$ is positive for most workers. This is not surprising, because a longer commute radius is associated with a larger number of job openings, and therefore a higher probability of high wage realization. As a result, even if jobs are uniformly distributed across locations and the wage distributions of the jobs are identical, a worker would face a positive β_{jo}.

To further demonstrate the spatial pattern of the wage return to commuting, we present the relationship between $\widehat{\beta_{jo}}$ and distance to the city center in Figure 16. Figure 16a shows that on average, workers who live farther from the city center face a greater wage return to commuting. The model predicts that if women dislike commuting more than men do, there should be a higher gender commute gap for workers living farther from the city center. Indeed, we document the pattern in Figure 6. Figure 16b shows a positive correlation between $\widehat{\beta_{jo}}$ and distance to the city center for financial workers, but does not show a discernible correlation between $\widehat{\beta_{jo}}$ and distance to the city center for physicians and surgeons. The findings are consistent with the pattern presented in Figure 9: The gender commute gap is larger for managers living farther from the city center, but not for healthcare practitioners.

6 Decomposition of the Gender Wage Gap

Differential Commuting Preferences After estimating the model parameters, we compute how much the observed gender wage gap can be attributed to differential gender commuting preferences. To do so, we compare the model-predicted gaps in commutes and wages with the observed gaps.

For a given female worker, we can estimate the change in her log commute and log wage if her disutility of commuting λ_f is adjusted to λ_m, using Equation 1 in Section 2:

$$\ln(\tau_f^*) - \ln(\tau_m^*) = -\frac{\beta_i(\lambda_f - \lambda_m)}{\lambda_f\left(\beta_i + \lambda_f \tau^{min}\right)},$$
$$\ln(w_f^*) - \ln(w_m^*) = -\frac{\beta_i(\lambda_f - \lambda_m)}{\lambda_f}.$$

The differences estimated with the above equations can be seen as the gender gaps in log commutes and log wages due to differential commuting preferences. We refer to these estimated gender gaps as the model-predicted gender gaps.

We compute the model-predicted gender gaps for workers, and then collapse them by residential PUMA. Figure 17 presents the model-predicted gender gap in log commute residuals and log hourly wage residuals versus the corresponding observed gaps across residential PUMAs. Figure 17a shows that the model-predicted gender commute gap is highly correlated with the observed gender commute gap. The average model-predicted commute gap is 0.0368 log points, compared with the average observed gap of 0.0481 log points. The results suggest that the observed gender commute gap is largely driven by differential gender commuting preferences.

Figure 17b shows that the average model-predicted gender wage gap is 0.0273 log points, compared with the average observed gap of 0.1314 log points. The results suggest that commuting preferences could explain 20.78% of the observed gender wage gap. This is higher than the implication in Figure 7, which suggests that commuting preferences could explain 16% of the gender wage gap. The difference between these results may be attributed to two reasons. First, the reduced-form implication has to assume that the positive correlation shown in Figure 7 represents causality, which may not hold. Second, the decomposition result is based on the New York MSA only; high-wage jobs are highly concentrated in central business district in the New York MSA, which may lead to a greater β on average, compared with other MSAs.

Geographic Variation in the Role of Commuting Preferences The model predicts that the gender difference in commuting preferences leads to a larger gender wage gap when the wage return to commuting is larger. Figure 16 shows that the wage return to commuting is on average greater for workers living farther from the city center. Thus, we also compute the contribution of commuting preferences to the gender wage gap by distance to the city center.

Table 7 Columns 2–5 show that the contribution of commuting preferences to the gender wage gap differs across residential locations. In particular, for workers living far from the central city, gender differences in commuting preferences play a more important role, accounting for 19.18% of the gender wage gap for those living more than 25 miles from the city center. However, for workers living within 5 miles of the city center, the difference in commuting preferences actually reduces the gender wage gap. This is partially due to the higher percentage of single female workers with no children, who have higher λ than male workers, in the central city, and the relatively smaller wage return to commuting faced by workers living in the central city. These results suggest that gender differences in commuting preferences alone cannot explain the geographic variation in the gender wage gap.

Geography of Jobs and the Wage Return to Commuting To illustrate the role of the geography of jobs and the wage return to commuting in the gender wage gap, we conduct several exercises in which we change β faced by workers.

In the main exercise, we assume workers face the wage return to commuting β_{jo} assigned to them based on their location of residence j and occupation o. Now, we examine how removing spatial variation in β could affect the gender wage gap. Columns 1 and 2 in Table 8 show the wage gaps if we set every worker's β to the median and mean β, respectively. We find that the model-predicted wage gaps are reduced significantly. The large reduction in the gender wage gap is due to the fact that married households tend to live in the suburbs, where β tend to be larger. Since the gender difference in λ is higher for married households, a reduction of β to the median or mean would lead to a significant reduction in the wage gap. In contrast, the gender difference in λ is smaller for single households (men have larger λ than women for single workers without children). Thus, an increase of β to the median or mean would not lead to a large increase in the gender wage gap for the single workers. This exercise shows that spatial variation in the wage return to commuting is important for correctly understanding the contribution of differential commuting choices to the gender wage gap.

Next, we take the existing spatial variation in β and reduce each worker's β proportionally to see how the reductions affect the gender wage gap. Columns 3 and 4 show that the model-predicted wage gaps are significantly lower if β faced by each worker is reduced by 50% or 80%. The magnitude of the change in the gender wage gap is nearly proportional to the percentage reduction in β.

7 Counterfactual Analysis

Based on the estimated model, we evaluate by how much real-world technological or policy changes could affect the wage gap. We consider two scenarios: (i) improving the efficiency of transportation (e.g., increasing travel speed), and (ii) relocating workers' residential locations.

7.1 Increasing Travel Speed

We assess by how much increasing travel speed could affect the gender wage gap. More specifically, we hold workers' preference parameters and job locations constant, and examine what happens to workers' optimal job choices if the commute time to all available jobs is reduced.

We take the simulated spatial distributions of jobs as given, and reduce the commute time from each residential location j to all job locations by 20%, 50%, and 80%. Based on the new commute time for each job, we measure a new set of counterfactual wage returns to commuting $\{\beta'_{jo}\}$ and compute the implied gender wage gaps using $\{\beta'_{jo}\}$.[27]

Table 9 shows how much the model-predicted gender wage gap decreases as a result of reduced commuting time. The results suggest that the effect of reducing commute time to all jobs (e.g., through improving transportation or traffic conditions) is not linear. Specifically, comparing Column 2 with Column 1 shows that the effect of reducing commute time by 20% on the model-explained gap is very small. However, when commute time is further reduced, the contribution of differential commuting choices to the gender wage gap becomes much smaller. For instance, when commute time is reduced by 80%, the model-predicted gender wage gap decreases from 0.0273 log points to 0.0189 log points. As a result, the contribution of differential commuting choices is lowered from 20.78% to 14.38%.

The nonlinear effect of reducing commute time on the model-explained gender wage gap is due to its nonlinear effect on β_{jo}. Specifically, reducing commute time to all jobs has two effects on a worker's job choice set.

1. Reducing commute time raises the wage a worker could attain within a relatively short commute, and therefore may lower the return of long commutes. As a result, this effect could lower the gender wage gap, since workers now face a smaller trade-off between commutes and wages.

2. As commute time to all locations decreases, previously unattainable jobs that are very far away could become more attainable. As a result, this could drive up the wage return to commuting. For example, the construction of highways or high-speed trains could make extremely distant but high-wage jobs more reachable. Therefore, as travel speed increases, people may travel farther and still face a trade-off between commutes and wages.

[27] We do not change τ_{min} as we do the exercises. This is an important setup because if τ_{min} is reduced proportionally, the estimated βs should remain the same due to the log-log frontier specification.

7.2 Increasing Residential Density in Urban Cores

Alternatively, we evaluate how the gender wage gap might change if residential density in the urban core is increased. We do so by randomly selecting residents living outside the 10-mile radius of downtown New York and reassign them to locations within the 10-mile radius of downtown. The probability of being assigned to a PUMA location is based on the PUMA's population share out of all PUMAs within the 10-mile radius of downtown. Once the reassignment is complete, we calculate the implied gender wage gap based on workers' occupations and newly assigned residential locations. Since the reassignment is based on simulated reassignment, we repeat the exercise 100 times. The counterfactual gender wage gap is computed as the average of the implied wage gaps computed in each simulation.

Table 10 shows the results of the counterfactual exercises. We show that by reallocating 5%, 10%, and 20% of the MSA's population from suburban neighborhoods to central city neighborhoods, we can reduce the gender wage gap by 0.4, 0.77, and 1.5 percentage points, respectively.

7.3 Discussion

The counterfactual analyses show that reducing travel time or increasing central city density could reduce the gender wage gap, but only with a moderate magnitude. Unless travel cost is entirely eliminated, workers are sill likely to face a trade-off between wages and commutes for the reasons described above. A spatial concentration of population, though effective, can be very costly due to increased congestion and cost of construction.

The results of the counterfactual analyses suggest that the geography of jobs and workers will likely play a lasting role in the gender wage gap, especially since other contributing factors are being addressed and mitigated. It is possible, however, to reduce the effect of commuting choices on the gender wage gap if the following two changes occur:

1. The preference parameter for commuting (λ) decreases for female workers. Since λ is determined by how much female workers are willing to trade off wages for shorter commute time, if we can reduce the burden of long commute time, we can lower λ for female workers. One such change would be to improve the quality and convenience of childcare and after-school programs so that fewer female workers have to trade off taking a high wage job for childcare.

2. Remote work becomes a more common option for female workers. If remote work becomes increasingly common, the returns to commuting would be lowered dramatically without changing the geography of jobs and workers or improving transportation infrastructure. The effect of remote working options on the wage gap could be a subject for future research, particularly in light of the COVID-19 pandemic.

8 Conclusion

This paper studies how gender differences in commuting choices contribute to the gender wage gap in the United States. In particular, we emphasize the role of the wage penalty for reducing commute time that stems from the geography of jobs. We use a job choice model to show that if women value short commutes more than men do, such differential preferences could lead to a gender wage gap. However, the gender wage gap would occur only when workers face a trade-off between higher wages and shorter commutes. In other words, the gender wage gap caused by differential commuting preferences is increasing in the wage return to commuting.

Using ACS data, we present empirical evidence that supports the predictions of the model. First, we show that the gender commute and wage gaps are greater for workers living farther from the city center. The finding remains after controlling for workers' observable characteristics. This finding supports the model since many high-wage jobs are concentrated in the central city, and this geography of jobs leads to relatively greater wage returns to commuting for workers living farther from the city center. Moreover, we show that the residual gender wage gap is highly correlated with the residual gender commute gap. The positive correlation suggests that different commuting choices between men and women due to different commuting preferences are likely to contribute to their different wage outcomes.

We then estimate the model parameters: (i) the slope of indifference curves for men and women separately, and (ii) the wage return to commuting for each worker living in the New York MSA. To estimate the slope of indifference curves, we use a quantile regression model to estimate the slope of the indifference curve corresponding to the reservation utility, assuming that after controlling for observable characteristics and various fixed effects, workers of the same gender face the same reservation wage conditional on the commute time. To measure the wage return to commuting, we use wage and job location data to simulate spatial distributions of jobs and wages. We then apply a frontier estimator to the simulated data to recover the wage return to commuting for each occupation and residential location.

With the estimated model, we show that the model-predicted gender commute gap is very close to the observed commute gap. Moreover, differential commuting choices could account for about 21% of the observed gender wage gap. This decomposition result obtained through the lens of the model closely mirrors the reduced-form evidence.

Finally, we use the model to conduct several counterfactual analyses and evaluate how technological or policy changes may affect the gender wage gap through the mechanism described by the model. We show that reducing commute time to all jobs (e.g., through improving traffic conditions) and increasing density in central city neighborhoods could moderately lower the contribution of differential commuting choices to the gender wage gap. However, the effects are moderate, which suggests that even if transportation infrastructure is improved and housing supply in central cities increases, the geography of jobs is likely to play a lasting role in the gender wage gap, unless we reduce female workers' trade-off for shorter commutes (e.g., by improving childcare programs) or make remote work options more available for women.

References

Ahlfeldt, G., Redding, S., Sturm, D., and Wolk, N. (2015). The economics of density: Evidence from the berlin wall. *Econometrica*, 83(6):21272189. DOI: `https://doi.org/10.3982/ECTA10876`.

Becker, G. (1965). A theory of the allocation of time. *Economic Journal*, 75(299):493–517. DOI: `https://doi.org/10.2307/2228949`.

Becker, G. (1985). Human capital, effort, and the sexual division of labor. *Journal of Labor Economics*, 3(1):33–58. DOI: `https://doi.org/10.1086/298075`.

Bertrand, M., Goldin, C., and Katz, L. F. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2(3):228–55. DOI: `https://doi.org/10.1257/app.2.3.228`.

Black, D. A., Kolesnikova, N., and Taylor, L. J. (2014). Why do so few women work in new york (and so many in minneapolis)? labor supply of married women across us cities. *Journal of Urban Economics*, 79:59–71. DOI: `https://doi.org/10.1016/j.jue.2013.03.003`.

Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865. DOI: `https://doi.org/10.1257/jel.20160995`.

Cazals, C., Florens, J.-P., and Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of econometrics*, 106(1):1–25. DOI: `https://doi.org/10.1016/S0304-4076(01)00080-X`.

Cha, Y. and Weeden, K. (2014). Overwork and the slow convergence in the gender gap in wages. *American Sociological Review*, 79(3):457–484. DOI: `https://doi.org/10.1177/0003122414528936`.

Cubas, G., Juhn, C., and Silos, P. (2019). Coordinated work schedules and the gender wage gap. *Working paper*. DOI: .

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119. DOI: `https://dx.doi.org/10.1257/aer.104.4.1091`.

Goldin, C. and Katz, L. (2011). The cost of workplace flexibility for high-powered professionals. *The ANNALS of the American Academy of Political and Social Science*, 638(45):45–67. DOI: `https://doi.org/10.1177/0002716211414398`.

Gutierrez, F. (2018). Commuting patterns, the spatial distribution of jobs and the gender pay gap in the U.S. *Available at SSRN 3290650*. DOI: `http://dx.doi.org/10.2139/ssrn.3290650`.

Kreindler, G. (2018). The welfare effect of road congestion pricing: Experimental evidence and equilibrium implications. *Working paper*.

Le Barbanchon, T., Rathelot, R., and Roulet, A. (2019). Gender differences in job search: Trading off commute against wage. *Working Paper*.

Lundborg, P., Plug, E., and Rasmussen, A. W. (2017). Can women have children and a career? IV evidence from IVF treatments. *American Economic Review*, 107(6):1611–37.

Mas, A. and Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, 107(12):3722–3759. DOI: `https://doi.org/10.1257/aer.20161500`.

Monte, F., Redding, S., and Rossi-Hansberg, E. (2018). Commuting, migration, and local employment elasticities. *American Economic Review*, 108(12):38553890. DOI: `https://doi.org/10.1257/aer.20151507`.

Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J., and Sobek, M. (2019). IPUMS USA: Version 9.0 [dataset]. Minneapolis, MN: IPUMS.

Severen, C. (2019). Commuting, labor, and housing market effects of mass transportation: Welfare and identification. *Working paper*. DOI: .

Small, K. and Verhoef, E. (2007). *The Economics of Urban Transportation*. Routledge.

Small, K., Winston, C., and Yan, J. (2005). Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica*, 73(4):1367–1382. DOI: `https://doi.org/10.1111/j.1468-0262.2005.00619.x`.

Su, Y. (2019). The rising value of time and the origin of urban gentrification. *Working Paper*. DOI: `http://dx.doi.org/10.2139/ssrn.3216013`.

Tsivanidis, N. (2019). Evaluating the impact of urban transit infrastructure: Evidence from bogota's transmilenio. *Working paper*.

Wasserman, M. (2019). Hours constraints, occupational choice, and gender: Evidence from medical residents. *Working paper*. DOI: `http://dx.doi.org/10.2139/ssrn.3371100`.

Wiswall, M. and Zafar, B. (2017). Preference for the workplace, investment in human capital, and gender. *Quarterly Journal of Economics*, 133(1):457507. DOI: `https://doi.org/10.1093/qje/qjx035`.

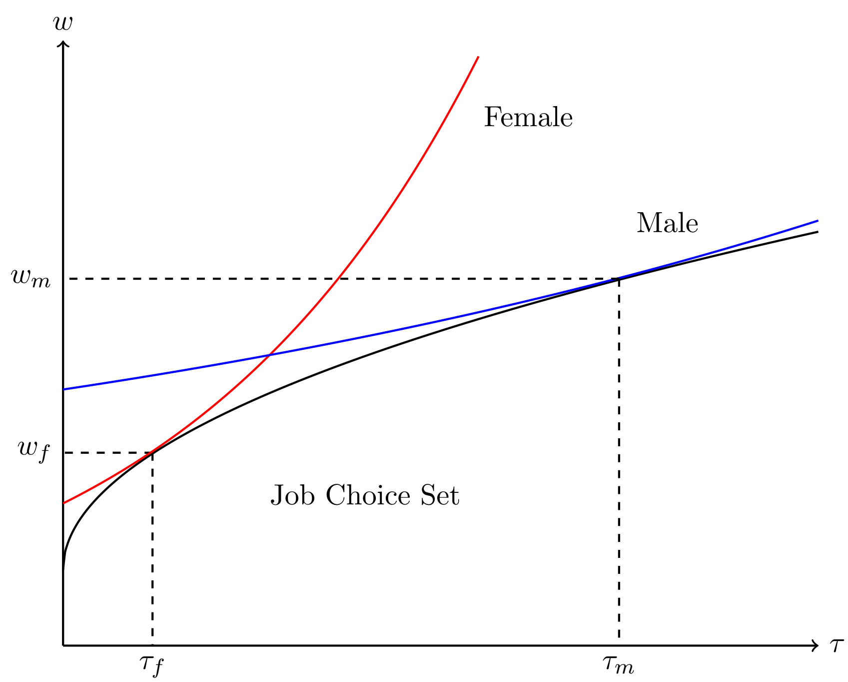


Figure 1: Indifference Curves and Job Choice Set Frontier

Note: The red and blue convex curves represent the indifference curves for a representative female and male worker, respectively. They face the same job choice set, and the concave curve represents the choice set frontier. Tangent points (τ_f, w_f) and (τ_m, w_m) represent the optimal job choice for the female and male worker, respectively.

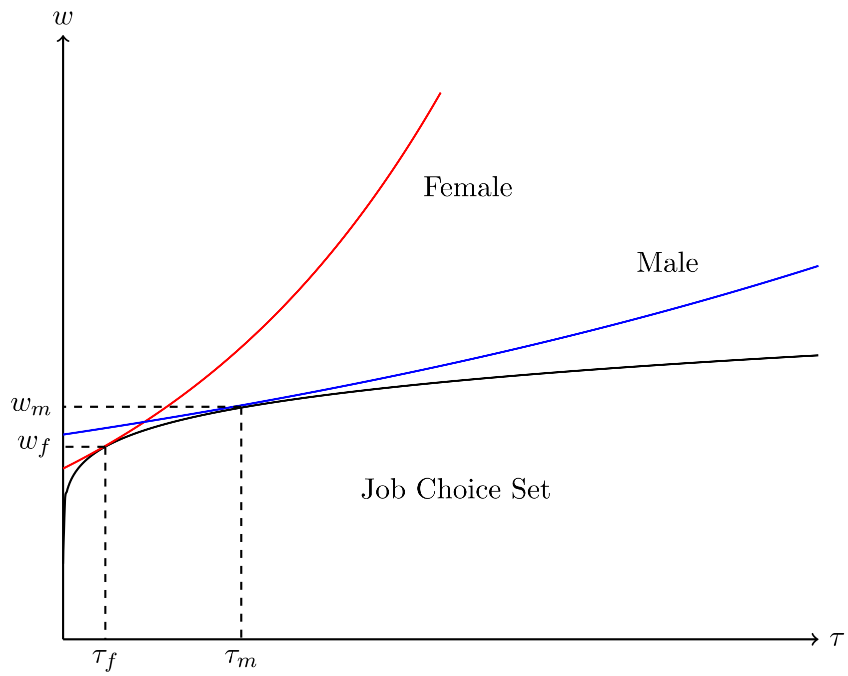
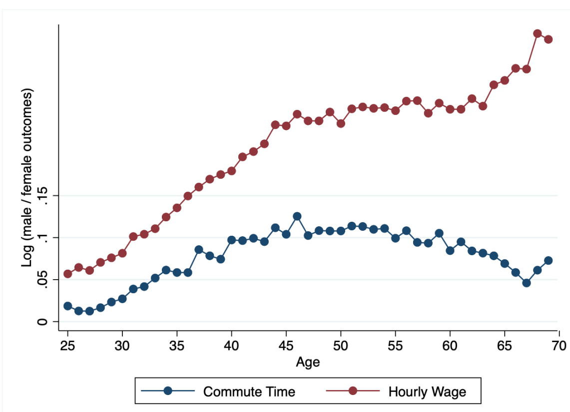
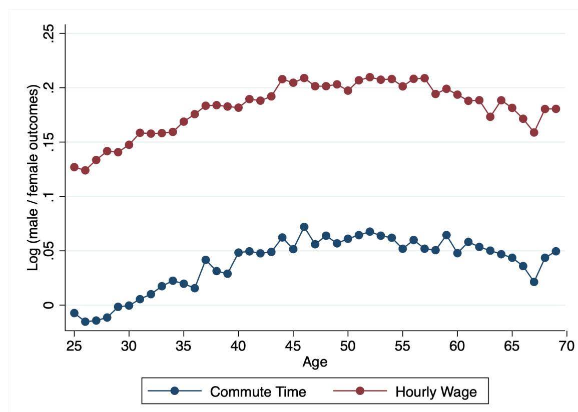


Figure 2: Indifference Curves and Job Choice Set Frontier: Low Wage Return to Commuting

Note: The red and blue convex curves represent the indifference curves for a representative female and male worker, respectively. They face the same job choice set, and the concave curve represents the choice set frontier. Tangent points (τ_f, w_f) and (τ_m, w_m) represent the optimal job choice for the female and male worker, respectively.



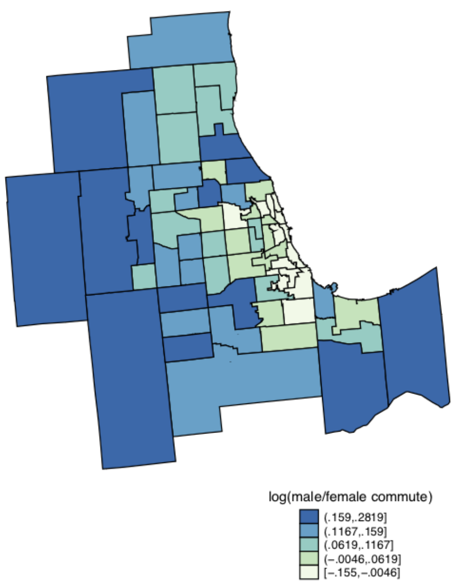
(a) Gender Commute and Wage Gaps



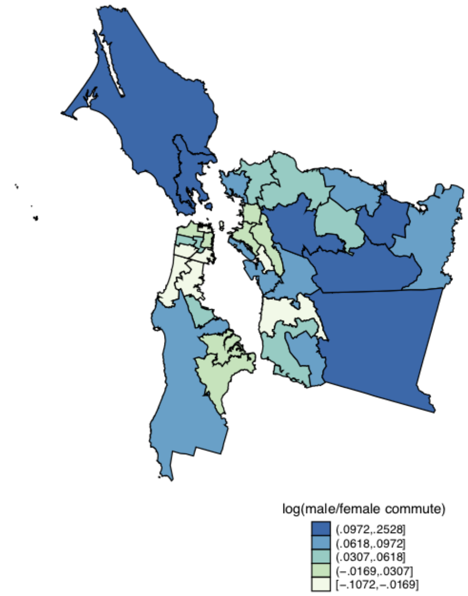
(b) Residual Gender Commute and Wage Gaps

Figure 3: Life-cycle Patterns in Gender Commute and Wage Gaps

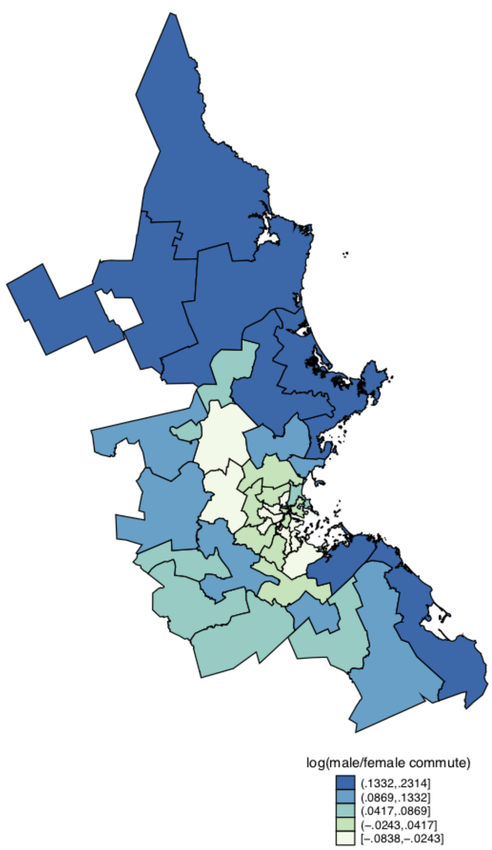
Note: The figures present scatter plots of gender differences in log commute (residuals) and log wage (residuals) by age. The sample comprises full-time workers aged between 25 and 69. Commute and wage residuals are obtained by controlling for dummies of age, marital status, whether having a child younger than 18, race, Hispanic origin, education level, year, occupation.



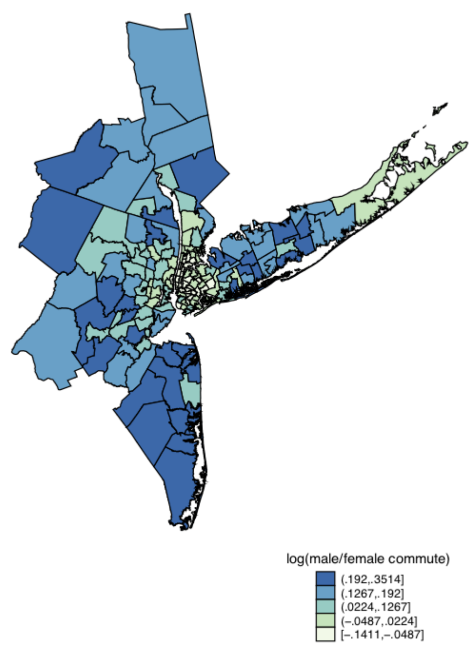
(a) Chicago



(b) San Francisco

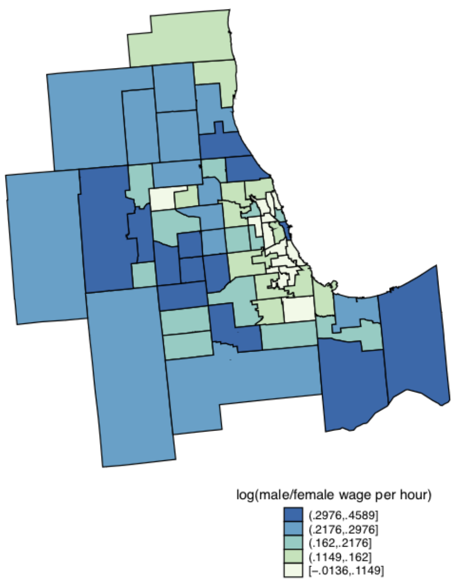


(c) Boston

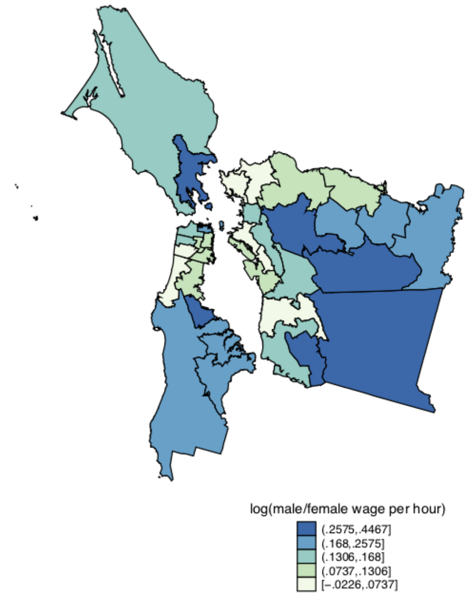


(d) New York

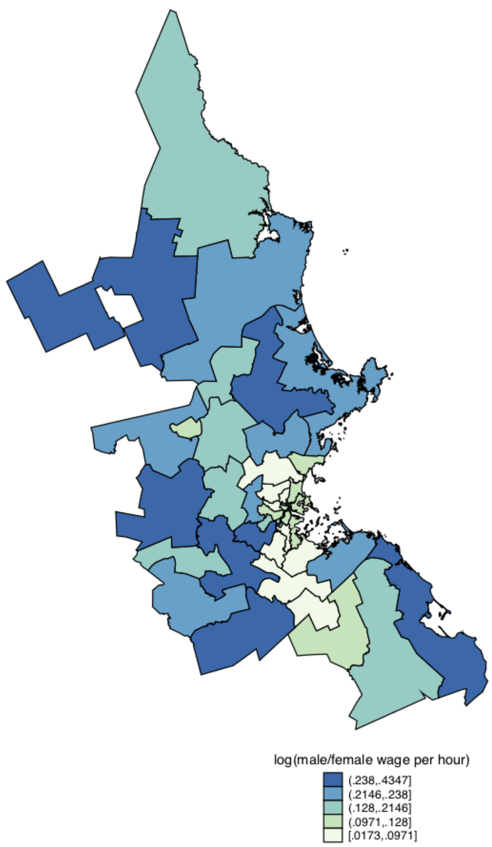
Figure 4: Gender Differences in Commute Time within Four Metropolitan Areas



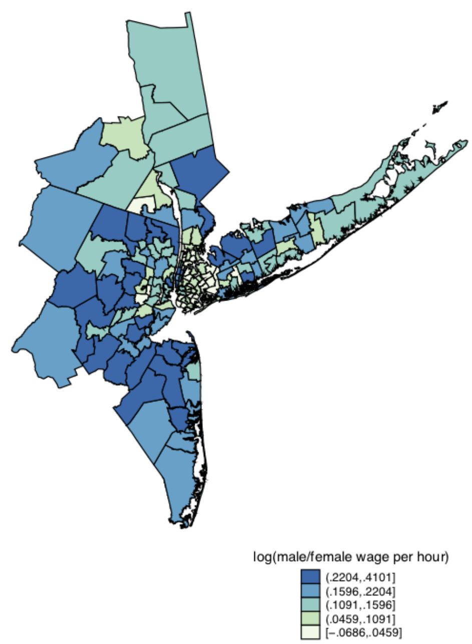
(a) Chicago



(b) San Francisco

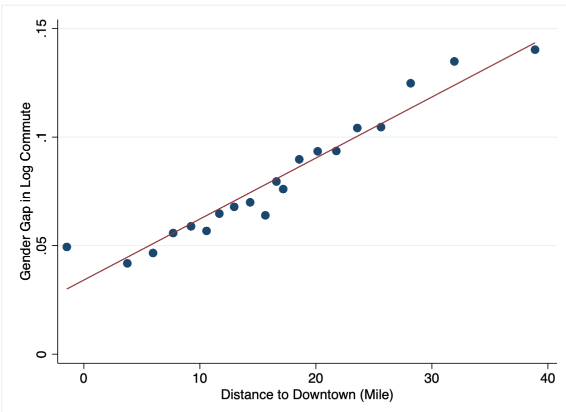


(c) Boston

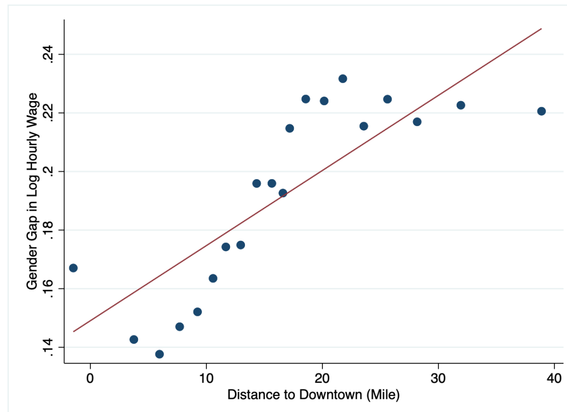


(d) New York

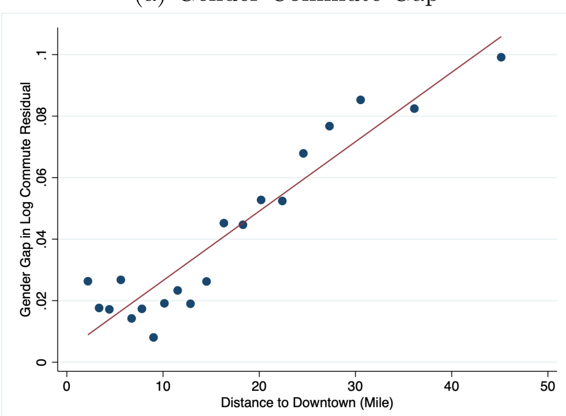
Figure 5: Gender Differences in Log Hourly Wage within Four Metropolitan Areas



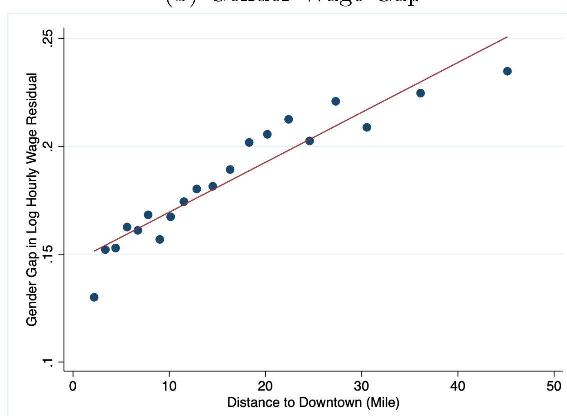
(a) Gender Commute Gap



(b) Gender Wage Gap



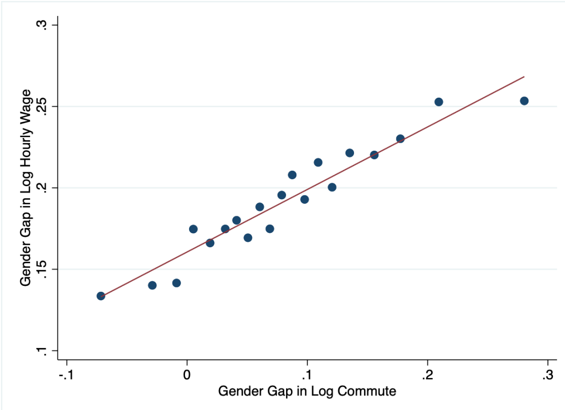
(c) Residual Gender Commute Gap



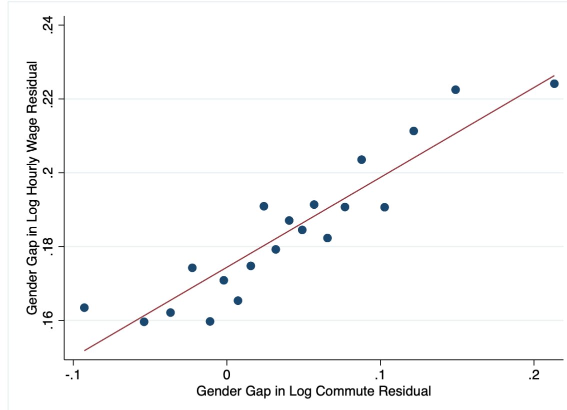
(d) Residual Gender Wage Gap

Figure 6: Gender Gaps and Distance to Downtown

Note: The figures present binned scatter plots of (residual) gender differences in log commutes and log hourly wages by residential PUMA. The residual gender gaps are obtained by controlling for dummies of age, marital status, whether having a child younger than 18, race, Hispanic origin, education level, occupation, year, and MSA. To construct the figures, we bin PUMAs into 20 equal groups based on the distance to the city center of the corresponding MSA.



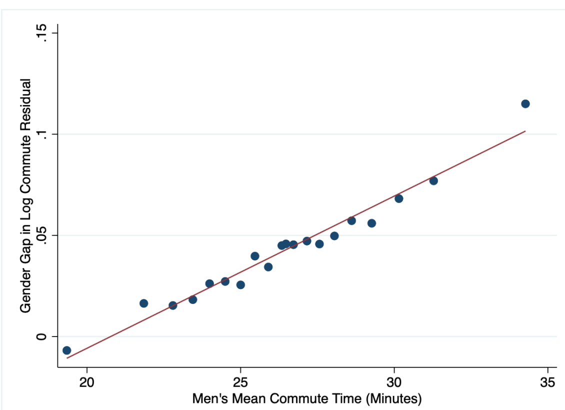
(a) Gender Commute and Wage Gaps



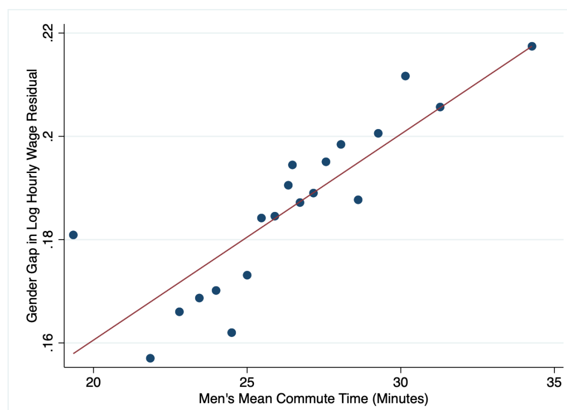
(b) Residual Gender Commute and Wage Gaps

Figure 7: Gender Commute Gap vs. Gender Wage Gap

Note: The figures present binned scatter plots of gender differences in log commute (residuals) vs. gender differences in log wage (residuals) by residential PUMA. The residual gaps are obtained by controlling for dummies of age, marital status, whether having a child younger than 18, race, Hispanic origin, education level, occupation, year, and MSA.



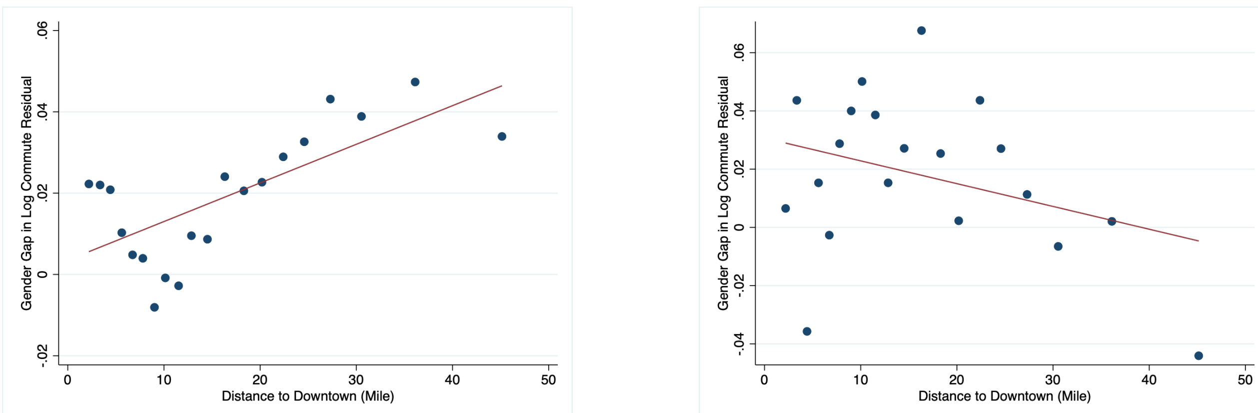
(a) Residual Commute Gaps and Average Commute Time



(b) Residual Wage Gaps and Average Commute Time

Figure 8: Gender Gaps and Average Commute Time

Note: The figures present binned scatter plots of gender differences in log commute and wage residuals by residential PUMA. We bin PUMAs into 20 equal groups based on the average commute time of male workers. The residual gaps are obtained by controlling for dummies of age, marital status, whether having a child younger than 18, race, Hispanic origin, education level, occupation, year



(a) Residual Gender Commute Gap:
Management and Business

(b) Residual Gender Commute Gap:
Healthcare Practitioners

Figure 9: Gender Commute Gaps: By Occupation Group

Note: The figures present binned scatter plots of residual gender differences in log commutes by residential PUMA for two occupation groups. "Management and Business" refers to occupations whose codes are between 10 and 950 (in ACS 2010 occupation classification). "Healthcare Practitioners" refers to occupations whose codes are between 3000 and 3540. The residual gaps are obtained by controlling for dummies of age, marital status, whether having a child younger than 18, race, Hispanic origin, education level, occupation, year, and MSA. To construct the figures, we bin PUMAs into 20 equal groups based on the distance to downtown.

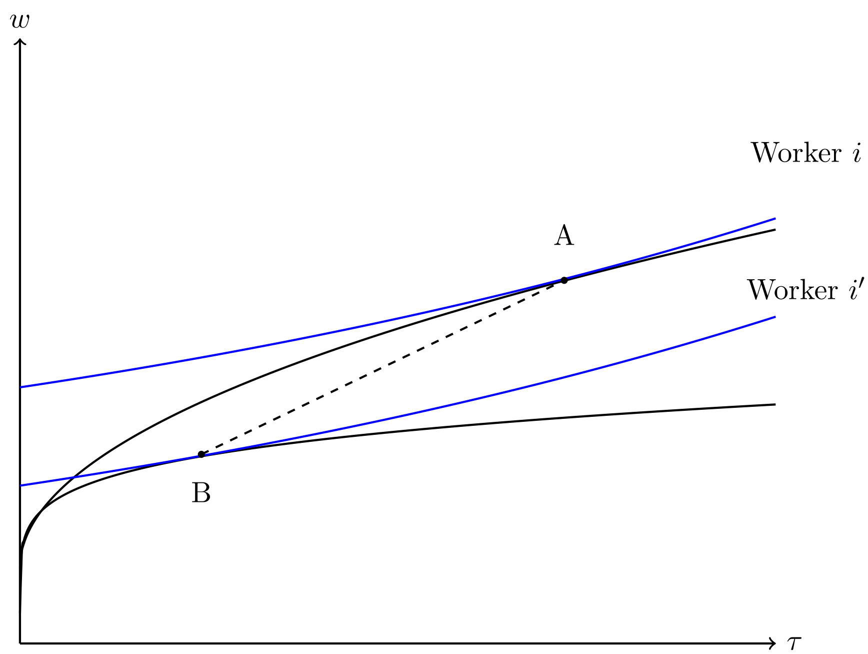


Figure 10: Location Sorting by Ability

Notes: The figure illustrates why a simple regression may lead to a biased estimate of the slope of difference curves due to location sorting by ability. Assume worker i has higher ability than worker i', and worker i prefers living in the suburbs while worker i' prefers living in the center city. Worker i faces a higher job choice set frontier than i'. Worker i chooses point A while worker i' chooses point B. Tracing out points A and B would lead to an upward-biased estimate of the slope of indifference curves.

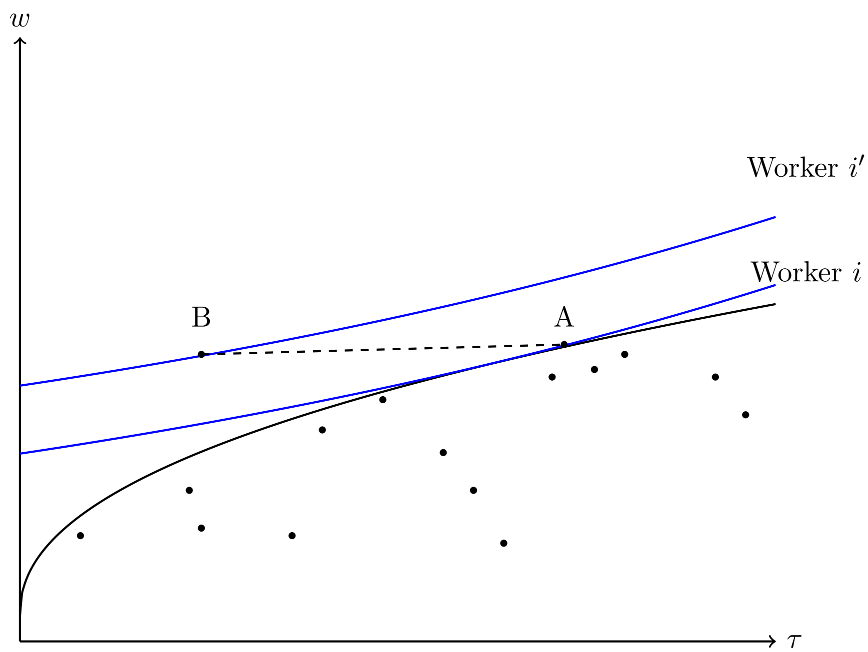


Figure 11: Random Job Arrival

Notes: The figure illustrates why a simple regression may lead to a biased estimate of the slope of difference curves due to the randomness of job arrival. If job openings are idiosyncratic across workers, then two identical workers may receive slightly different sets of jobs. Assume worker i only receives jobs within her job choice set and chooses point A along the frontier. Worker i' receives an additional job offer B idiosyncratically, which is closer and pays similar wage as A. Therefore, worker i' would accept B. In this case, tracing points A and B would underestimate the slope of the indifference curves.

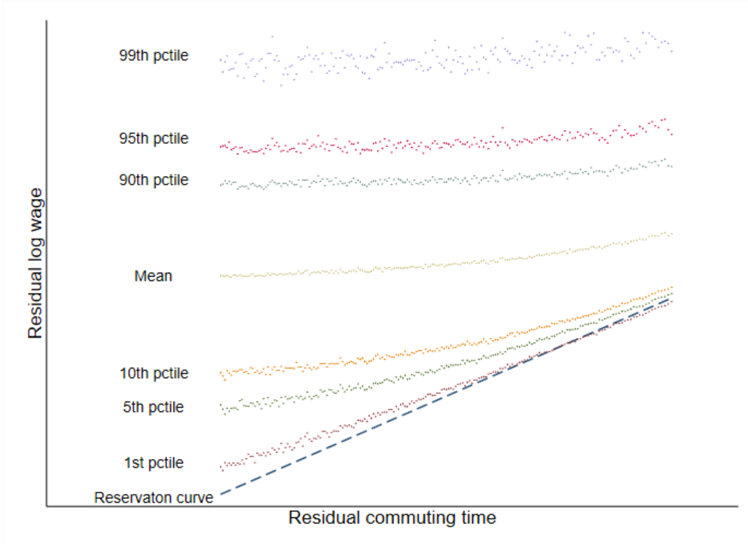


Figure 12: Log Wage Quantiles and Reservation Curve: Simulated Data

Note: The figure is based on simulated data, where commute residuals τ are assigned to log wage residuals $\ln(w)$ observed from the data, and τ is randomly drawn from a uniform distribution between -1.5 and 1.6. Assume that there is measurement error in the observed wage w: $\ln(w) = \ln(w^*) + \eta$, where w^* is the true wage and η is a normally distributed noise term with mean 0 and standard deviation 0.15. Assume the true utility function is $U(\tau, w^*) = \ln(w^*) - 0.5\tau$ and the reservation utility is -0.84. The figure presents the indifference curve corresponding to the reservation utility. The figure also presents binned scatter plots of various percentiles of log wage residuals with respect to commute residuals, based on accepted job bundles. In particular, a job is accepted if $U(\tau, w^*) \geq -0.84$.

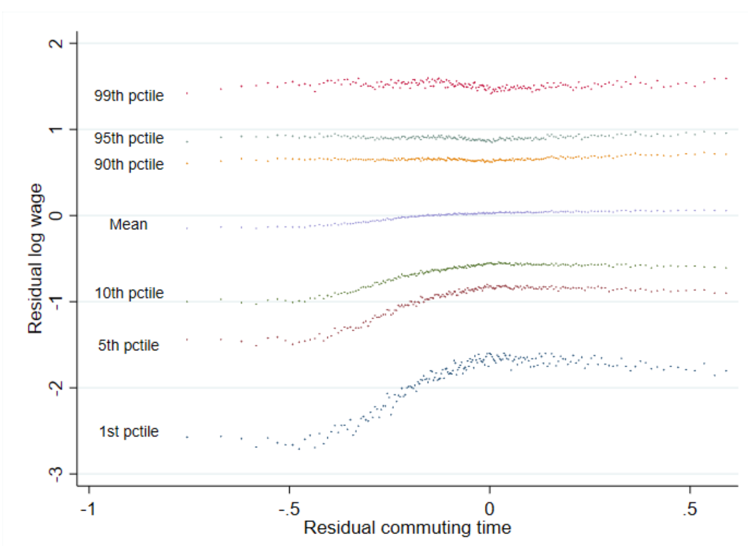
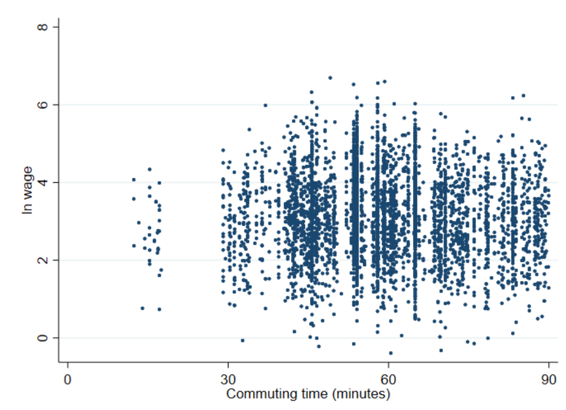
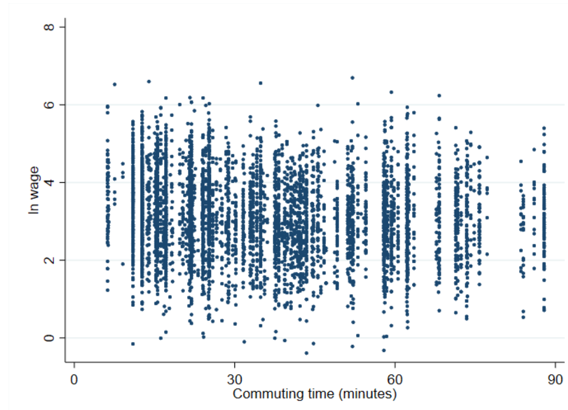


Figure 13: Log Wage Quantiles: Real Data

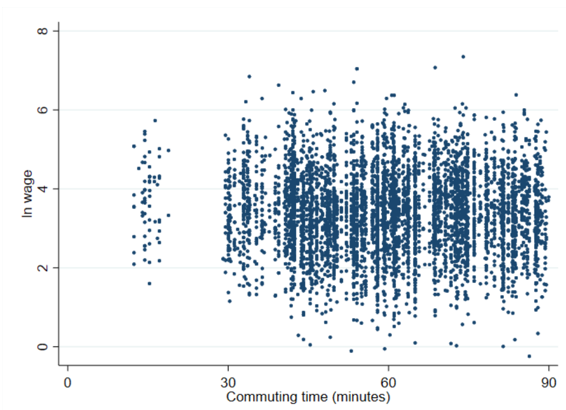
Note: The figure is based on real data of observed commute residuals τ and log hourly wage residuals $\ln(w)$. The figure presents binned scatter plots of various percentiles of log wage residuals with respect to commute residuals.



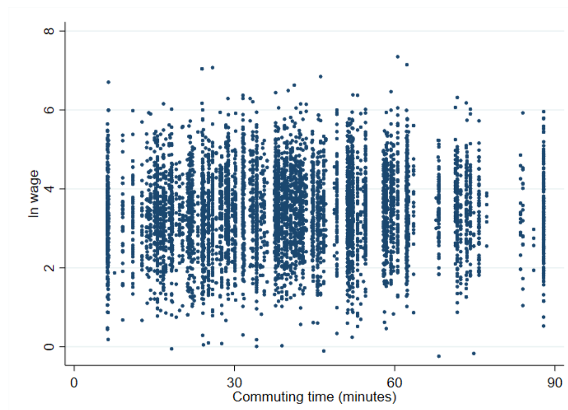
(a) Financial Managers Living in Northern Suburbs



(b) Financial Managers Living in Midtown Manhattan



(c) Physicians & Surgeons Living in Northern Suburbs



(d) Physicians & Surgeons Living in Midtown Manhattan

Figure 14: Simulated Job Choice Set by Occupation and Residential Location

Note: The figures present the job choice sets faced by workers in different occupations (financial workers vs. physicians and surgeons) living in different PUMAs in the New York MSA (northern suburbs vs. Midtown Manhattan). Each dot represents a job (i.e., a commute-wage bundle).

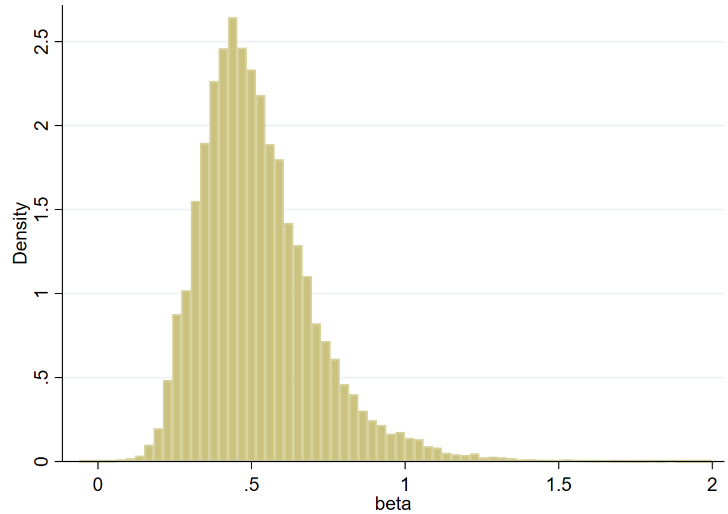
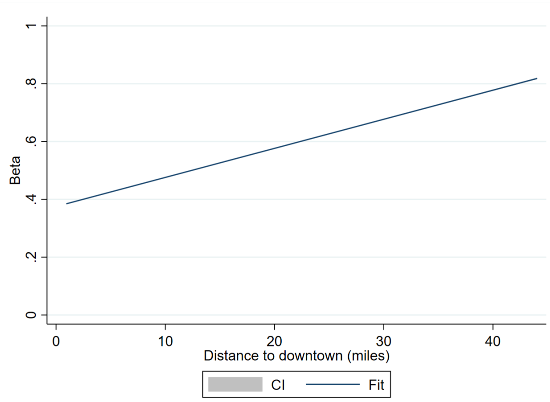
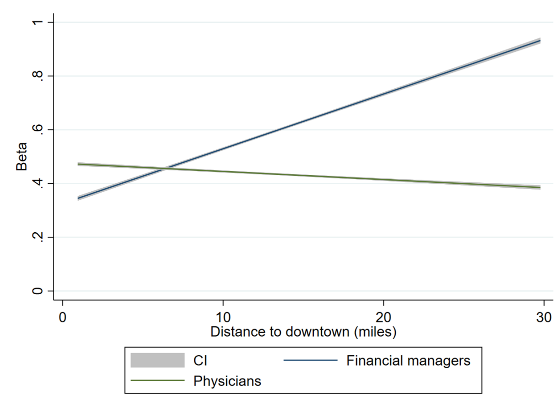


Figure 15: Distribution of estimated $\widehat{\beta_{jo}}$

Note: This figure shows the histogram of estimated $\widehat{\beta_{jo}}$. Each observation of the data is a worker observed in the ACS microdata. For each worker, we assign the estimated $\widehat{\beta_{jo}}$ based on his/her PUMA location of residence and occupation. The sample is weighted by personal sampling weight (perwt).



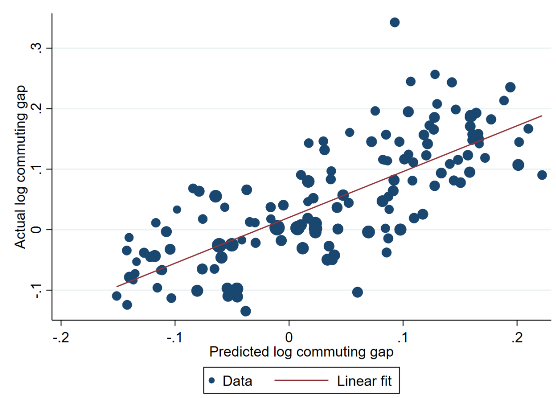
(a) All Occupations



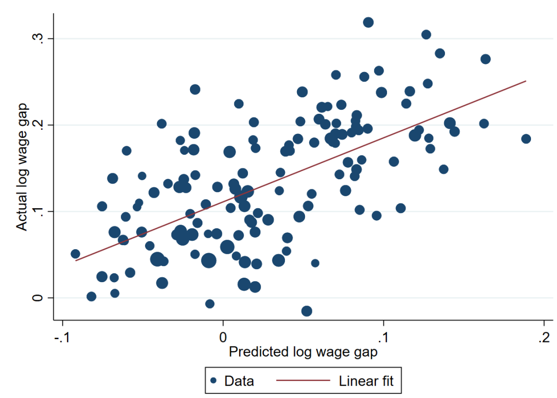
(b) Financial Managers vs. Physicians & Surgeons

Figure 16: Relationship between Wage Return to Commuting and Distance to the City Center

Note: The figures present the relationship between the wage return to commuting and distance to the city center for all occupations (in Figure a), and for financial workers and physicians and surgeons (in Figure b).



(a) Gender Commute Gap



(b) Gender Wage Gap

Figure 17: Model-predicted Gender Gaps vs. Observed Gender Gaps

Note: The figures present the model-predicted gaps in log commute residuals and log hourly wage residuals vs. the corresponding gaps observed in the data. The model-predicted gaps are computed using Equation 1. We compute the model-predicted gaps for each worker, then collapse them by residential PUMA. Each point in the figure represents a residential PUMA.

Table 1: Summary Statistics

Variables	Men			Women		
	Mean	SD	Observations	Mean	SD	Observations
Age	44.034	11.584	3,003,648	44.077	11.623	2,355,573
Married	0.613	0.487	3,003,648	0.525	0.499	2,355,573
Child (< 18)	0.393	0.488	3,003,648	0.384	0.486	2,355,573
High school or lower	0.420	0.494	3,003,648	0.327	0.469	2,355,573
Some college	0.231	0.421	3,003,648	0.262	0.440	2,355,573
College or higher	0.350	0.477	3,003,648	0.412	0.492	2,355,573
Commute time (in hour)	0.436	0.405	3,003,648	0.389	0.356	2,355,573
Hourly wage	30.471	31.348	3,003,648	23.621	21.174	2,355,573

Note: The sample comprises full-time workers (i.e., usually worked at least 35 hours per week) aged between 25 and 64 from the 2013–2017 ACS. *Married* is an indicator of being married and spouse being present. *Child (<18)* is an indicator for having a child younger than 18. Variables on education represent the highest educational level. *Commute time* represents one-way commute time measured in hours.

Table 2: Relationship Between Log Wages and Commute Time

	Log (Wage)						
	All	All	College	<College	Single	Married	Married
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.227***	-0.196***	-0.176***	-0.209***	-0.0973***	-0.267***	-0.205***
	(0.00256)	(0.00174)	(0.00239)	(0.00228)	(0.00244)	(0.00206)	(0.00253)
Commute	0.194***	0.115***	0.0871***	0.128***	0.131***	0.100***	0.112***
	(0.00449)	(0.00159)	(0.00289)	(0.00191)	(0.00290)	(0.00181)	(0.00238)
Commute × Female	0.111***	0.0561***	0.0554***	0.0607***	0.00929**	0.0930***	0.0951***
	(0.00373)	(0.00227)	(0.00356)	(0.00309)	(0.00367)	(0.00286)	(0.00354)
Commute (Spouse)							-0.0345***
							(0.00176)
Log Hours Worked							-0.231***
(Spouse)							(0.00476)
Observations	5,359,221	5,182,799	3,080,165	1,946,585	1,846,323	3,143,555	1,924,987
R-Squared	0.026	0.470	0.393	0.455	0.478	0.499	0.505
Controls (Own)	No	Yes	Yes	Yes	Yes	Yes	Yes
Controls (Spouse)	No	No	No	No	No	No	Yes

Note: The sample contains full-time workers (i.e., usually worked at least 35 hours per week) aged between 25 and 64 from the 2013–2017 ACS. We use 1st percentile quantile regressions to estimate λ_s. Columns 2–6 control dummy variables for age, race, Hispanic origin, marital status, whether having a child younger than 18, education, occupation, and residual PUMA, and the interactions of residential PUMA fixed effects and occupation indicators. Column 7 further controls for dummy variables for the spouse's age, race, education, employment status, and occupation. Robust standard errors in parentheses *** p<0.01, ** p<0.05, *p<0.1.

Table 3: Estimates of Disutility of Commuting (λ)

	All (1)	Single no children (2)	Single with children (3)	Married no children (4)	Married with children (5)	College (6)	< College (7)
λ_m	0.379***	0.408***	0.322***	0.467***	0.297***	0.330***	0.386***
	(0.0119)	(0.0266)	(0.064)	(0.0253)	(0.0174)	(0.0197)	(0.0159)
λ_f	0.506***	0.357***	0.382***	0.616***	0.588***	0.491***	0.501***
	(0.0144)	(0.0291)	(0.0377)	(0.0312)	(0.0267)	(0.0231)	(0.0199)
$\lambda_f - \lambda_m$	0.127***	-0.051	0.0598	0.149***	0.29***	0.161***	0.116***
	(0.0187)	(0.0395)	(0.0743)	(0.0401)	(0.0319)	(0.0303)	(0.0255)
Observations	5,022,015	1,167,374	386,044	983,149	1,699,917	1,844,513	2,949,877

Note: The sample contains full-time workers (i.e., usually worked at least 35 hours per week) aged between 25 and 64 from the 2013–2017 ACS. We use 1st percentile quantile regressions to estimate λ_s. Estimation is based on log wage residuals and commute time (one-way commute in hours) residuals, controlling for dummy variables for age, race, Hispanic origin, marital status, whether having a child younger than 18, education, occupation, and residual PUMA, and the interactions of residential PUMA fixed effects and occupation indicators. Robust standard errors in parentheses *** p<0.01, ** p<0.05, *p<0.1.

Table 4: Estimates of Disutility of Commuting (λ) for Married Workers

	No Children (1)	With Children (2)	College Spouse (3)	Non-College Spouse (4)	Employed Spouse (5)	Unemployed Spouse (6)	Spouse works $>$ 40 hours (7)	Spouse works $\leq$ 40 hours (8)
λ_m	0.323***	0.311***	0.341***	0.339***	0.352***	0.212*	0.367***	0.318***
	(0.0380)	(0.0258)	(0.0316)	(0.0298)	(0.0213)	(0.108)	(0.0395)	(0.0264)
λ_f	0.497***	0.571***	0.611***	0.508***	0.579***	0.406***	0.537***	0.513***
	(0.0378)	(0.0307)	(0.0373)	(0.0295)	(0.0225)	(0.116)	(0.0405)	(0.0297)
$\lambda_f - \lambda_m$	0.174***	0.260***	0.270***	0.169***	0.227***	0.194	0.170***	0.195***
	(0.0536)	(0.0401)	(0.0489)	(0.0419)	(0.0310)	(0.1585)	(0.0566)	(0.0397)
Observations	586,730	1,027,865	693,968	940,337	1,692,421	39,475	552,627	1,071,337

Note: The sample contains married full-time workers (i.e., usually worked at least 35 hours per week) aged between 25 and 64 from the 2013–2017 ACS. We use 1st percentile quantile regressions to estimate λ_s. Estimation is based on log wage residuals and commute time (one-way commute in hours) residuals, controlling for dummy variables for age, race, Hispanic origin, marital status, whether having a child younger than 18, education, occupation, and residual PUMA, the interactions of residential PUMA fixed effects and occupation indicators, as well as dummy variables for the spouse's age, race, education, education, employment status, and occupation. Robust standard errors in parentheses *** p<0.01, ** p<0.05, *p<0.1.

Table 5: Estimates of Disutility of Commuting (λ)
by Marital Status and Distance to Central Cities

Residence	Distance to Downtown (Miles)			
	(0,5]	(5,15]	(15,25]	15+
	(1)	(2)	(3)	(4)
Panel A: Single Workers				
λ_m	0.404***	0.417***	0.413***	0.390***
	(0.0668)	(0.0373)	(0.0494)	(0.0473)
λ_f	0.357***	0.372***	0.432***	0.396***
	(0.0612)	(0.0384)	(0.0523)	(0.0435)
$\lambda_f - \lambda_m$	-0.047	-0.045	0.019	0.006
	(0.0906)	(0.0535)	(0.0719)	(0.0643)
Observations	308,321	631,653	345,539	437,779
Panel B: Married Workers				
λ_m	0.361***	0.348***	0.387***	0.331***
	(0.0604)	(0.0340)	(0.0454)	(0.0299)
λ_f	0.570***	0.590***	0.602***	0.559***
	(0.0739)	(0.0425)	(0.0455)	(0.0379)
$\lambda_f - \lambda_m$	0.209**	0.242***	0.215***	0.228***
	(0.0954)	(0.0544)	(0.0643)	(0.0483)
Observations	177,857	610,520	443,747	577,195

Note: In Panel A, the sample contains single full-time workers (i.e., usually worked at least 35 hours per week) aged between 25 and 64 from the 2013–2017 ACS. We use 1st percentile quantile regressions to estimate λ_s. Estimation is based on log wage residuals and commute time (one-way commute in hours) residuals, controlling for dummy variables for age, race, Hispanic origin, marital status, whether having a child younger than 18, education, occupation, and residual PUMA, and the interactions of residential PUMA fixed effects and occupation indicators. In Panel B, the sample contains married full-time workers aged between 25 and 64 from the 2013-2017 ACS. Estimation is based on log wage residuals and commute time residuals, controlling for workers' characteristics as well as dummy variables for their spouse's age, race, education, education, employment status, and occupation. Robust standard errors in parentheses *** p<0.01, ** p<0.05, *p<0.1.

Table 6: Estimates of Disutility of Commuting (λ) for Selected MSAs

Residence	NY (1)	Chicago (2)	SF (3)	Boston (4)
λ_m	0.379***	0.347***	0.475***	0.464***
	(0.0439)	(0.0652)	(0.101)	(0.0704)
λ_f	0.429***	0.592***	0.660***	0.508***
	(0.0467)	(0.0807)	(0.0827)	(0.105)
$\lambda_f - \lambda_m$	0.050	0.245***	0.185	0.0444
	(0.0641)	(0.1037)	(0.131)	(0.126)
Observations	309,236	137,728	80,582	82,712

Note: The sample contains full-time workers (i.e., usually worked at least 35 hours per week) aged between 25 and 64 from the 2013–2017 ACS. We use 1st percentile quantile regressions to estimate λ_s. Estimation is based on log wage residuals and commute time (one-way commute in hours) residuals, controlling for dummy variables for age, race, Hispanic origin, marital status, whether having a child younger than 18, education, occupation, and residual PUMA, and the interactions of residential PUMA fixed effects and occupation indicators. Robust standard errors in parentheses *** p<0.01, ** p<0.05, *p<0.1.

Table 7: Contribution of Commuting Choices to the Gender Wage Gap
(New York MSA)

Residence	Overall	Distance to Downtown (Miles)			
		(0,5]	(5,10]	(10,25]	25+
Observed gap in log wage	0.1314	0.1007	0.085	0.1453	0.1811
Model-predicted gap	0.0273	-0.0284	-0.0083	0.0369	0.0932
Unexplained gap	0.1041	0.1291	0.0933	0.1084	0.0879
Fraction explained	20.78%	-28.2%	-2.35%	25.4%	51.46%

Note: This table shows results of the decomposition exercises. Column 1 shows the overall observed gender gap in log wage residuals, the log wage gap predicted by the model, the gap that cannot be explained the model, and the fraction of the observed gap that can be explained by the model. Columns 2–5 show the decomposition results for workers by their distance to downtown.

Table 8: Effect of Wage Returns to Commuting (β) on the Gender Wage Gap
(New York MSA)

	β			
	0.489 (Median)	0.523 (Mean)	Cut by 50%	Cut by 80%
Observed gap in log wage	0.1314	0.1314	0.1314	0.1314
Model-predicted gap	0.0175	0.0187	0.0136	0.0055
Fraction explained	13.32%	14.23%	10.35%	4.19%

Note: The table shows the results of a counterfactual analysis, in which the commute time of all available jobs faced by workers is reduced by 20%, 50%, and 80%. Row 1 shows the observed gender gap in log wage residuals. Row 2 shows the model-predicted gap corresponding to the new commute time. Row 3 shows the fraction of the observed gap that can be explained by the model.

Table 9: Effect of Reducing Commute Time on the Gender Wage Gap
(New York MSA)

	Reducing Commute Time by			
	0%	20%	50%	80%
Observed gap in log wage	0.1314	0.1314	0.1314	0.1314
Model-predicted gap	0.0273	0.0267	0.0250	0.0189
Fraction explained	20.78%	20.32%	19.03%	14.38%

Note: The table shows the results of a counterfactual analysis, in which the commute time of all available jobs faced by workers is reduced by 20%, 50%, and 80%. Row 1 shows the observed gender gap in log wage residuals. Row 2 shows the model-predicted gap corresponding to the new commute time. Row 3 shows the fraction of the observed gap that can be explained by the model.

Table 10: The Effect of Reallocating Population to Central Cities on the Gender Wage Gap
(New York MSA)

	Share of Population Reallocated			
	0%	5%	10%	20%
Observed gap in log wage	0.1314	0.1314	0.1314	0.1314
Model-predicted gap	0.0273	0.0262	0.0251	0.0232
Fraction explained	20.78%	19.94%	19.10%	17.66%

Note: The table shows the results of a counterfactual analysis, in which residents living in the suburbs (PUMAs outside the 10-mile radius circle of downtown New York) are randomly reallocated to live in the central city (PUMAs within the 10-mile radius circle of downtown). Column 1 shows the baseline decomposition results. Columns 2, 3, and 4 show the model-predicted gap (and the fraction of the observed gap that can be explained by the model) if 5%, 10%, and 20% of the MSA's population are reallocated from the suburbs into central city neighborhoods, respectively.

Appendix

A1 Trip Chaining

Gender differences in trip-chaining behavior could complicate estimates of commute time. For instance, women may be more likely to have non-work stops for personal or family-related business (e.g., going to a grocery store or day care) during their trips between the residence and the workplace. If so, the gender commute gap may be much smaller if the amount of time spent on these chained trips is included.

We use the 2017 National Household Travel Survey (NHTS) to study whether trip chaining affects estimates of the gender commute gap. The NHTS Person File collects information on each interviewed household member, such as age, race, Hispanic origin, education, employment status, trip time (in minutes) to work without traffic, and state and metropolitan statistical area (MSA) of residence.[28] The NHTS Travel Day Trip File collects information on each trip made by each household member during the household's travel day, such as the purpose of the trip, trip origin purpose, trip destination purpose, and trip duration.

To study whether women are more likely to have chained trips for personal or family-related business on their way to/from work, we focus on trips with the purpose of "to/from work." For example, suppose a person's trips are as follows:

$$\text{Home} \xrightarrow{\text{Trip 1}} \text{Dry Cleaner} \xrightarrow{\text{Trip 2}} \text{Work} \xrightarrow{\text{Trip 3}} \text{Movie Theater} \xrightarrow{\text{Trip 4}} \text{Home.}$$

Then the purposes of Trip 2 and Trip 3 are "to/from work." Although the data provide self-reported trip time to work with traffic, we also estimate the average trip time of trips to/from work. This is the average duration of Trip 2 and Trip 3 in the example. In addition, We say that a person has a chained trip for personal or family-related business on the way to/from work if the trip origin purpose or the trip destination purpose of the to/from work trip is personal or family-related business, including (i) drop off/pick up someone, (ii) attend child or adult care, (iii) buy goods (e.g., groceries), services (e.g., dry cleaners), or meals, and (iii) other general errands (e.g., post office). Then we estimate the total trip time to and from work with other chained trips. This is equal to the total duration of Trip

[28] The NHTS does not provide information on marital status, so we cannot study trip-chaining behaviors by gender and marital status.

1, Trip 2, and Trip 3 in the previous example. We exclude the duration of Trip 4 in the calculation because the destination purpose of Trip 3 is recreation, not family-related business.

Table A1 provides summary statistics for full-time workers aged between 25 and 64 whose travel day is a weekday from the 2017 NHTS. These workers are slightly younger than those from the 2013-2017 ACS, but much more educated. For example, more than 50% of the workers in the NHTS sample have college degrees or higher, compared with 35% of men and 41% of women in the ACS sample. However, the average gender difference in one-way commute time is comparable between two data sets. The gender commute gap is 2.82 minutes from the ACS; it is 2.9 minutes in terms of trip time to work without traffic, and 3.59 minutes in terms of estimated trip time to/from work from the NHTS. Lastly, the gender difference in the trip time to and from work with other chained trips is 4.12 minutes, which is around 57–71% of the gender difference in two-way commute time. This suggests that including chained trips might slightly lower—but is not likely to close—the gender commute gap.

Table A2 provides OLS regression results of the gender difference in (i) the likelihood of having a trip for personal or family-related business, (ii) the likelihood of linking such trips to trips to/from work, (iii) the log trip time to work without traffic, (iv) the log trip time to/from work, and (v) the log trip time to and from work with other chained trips for personal or family-related business. In particular, we control for dummy variables for age, race, Hispanic origin, education, and MSA of residence. The results suggest that women are 6.6% more likely to have a trip for personal or family-related business and 11% more likely to link such trips to their trips to/from work. However, adding the amount of time spent on such chained trips to time to work does not affect the gender commute gap. The gender gaps in time to work without traffic, time to/from work, and time to and from work with chained trips are all around 13 log points.

A2 Crosswalk between Industry and Occupation

Since we estimate β_{jo} by occupation, but our job location data come by industry code (NAICS), we need to construct a crosswalk between NAICS and occupation. The occupation code we use is the occupation code in 2010 (occ2010) defined in IPUMS. There is no direct crosswalk between NAICS and occ2010. Instead, we construct the crosswalk in two steps. In the first step, we construct the crosswalk between NAICS and ind1990 (Census Industrial Classification in 1990). Once we have created the crosswalk, we use the Census microdata to generate a relation file between occ2010 and ind1990. For each ind1990, we calculate the percentage of that working population working in each occ2010 occupation. This enables us to generate a proportional assignment between ind1990 and occ2010. Using the crosswalk between NAICS and ind1990, we are able to compute the probability of workers working in each occ2010 for each NAICS.

Using the ZCBP business count, we construct a job location count by averaging over establishment size (for example, establishments with 1 to 4 employees would be assumed to have 2.5 employees) by ZIP Code and NAICS. We then apply the crosswalk between NAICS and occ2010 to compute the expected number of jobs in each occupation (occ2010) in each ZIP Code.

A3 Commute Time Imputation

We use the travel time matrix between neighborhoods based on work by Su (2019), who acquires the travel time and travel distance from the Google Distance Matrix API (Application Programming Interface). He augments the data using the National Household Travel Survey to impute travel time to reflect traffic conditions using a simple travel speed model. We take the travel time matrix parameterized with the 2009 National Household Travel Survey.

First, we use the Google Distance Matrix API to obtain travel time (with the traffic model turned off) and travel distance from each census tract to each ZIP Code within each MSA. We make sure that the 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 us information on route fixed effects (such as a slowing-down effect of crossing a bridge, a winding road, or dense city blocks with traffic lights).

Second, we use the 2009 NHTS data to fit a simple traffic speed model so that we can map the parameters estimated in the model onto the observable neighborhood characteristics and predict historical travel speed. We model travel speed as follows:

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

where j is the origin census tract; n is the destination ZIP Code; t is the year in which the trip is taken.

We assume log speed of the trip is a function of trip distance, because longer trips usually have higher speeds because people take the freeway or use a main thoroughfare when the distance is long enough. We assume that 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 neighborhoods in which the trips take place. A trip to or from densely populated neighborhoods are expected to experience heavier congestion than another trip that takes place in the suburbs. Additionally, we assume that each route admits a time-invariant fixed-effects component, which accounts for road conditions other than traffic congestion, such as the slowing-down effects described above. We assume these fixed effects do not change over time. The parameters of the model, $\beta_{0,t}$, $\beta_{1,t}$, $\beta_{2,t}$, and Γ_t, govern how location characteristics and trip distance are mapped onto travel speed. Since traffic conditions evolve over time, these parameters are assumed to be year-specific.

We use 2009 NHTS data to estimate these parameters to obtain parameters applicable to 2009 traffic conditions. We restrict trip samples to those that take place Monday to Friday and with departure time between 6:30 and 10:30 am and between 4:30 and 8:30 pm. We also restrict trips that either originate from or are destined for the respondent's location of residence. $\bar{\mathbf{X}}_{jn}$ takes the location characteristics of the Census tract the respondent lives in (neighborhood characteristics for the other end of the trip are unavailable). Additionally, we use Google API travel time (with the traffic model turned off) to estimate the fixed effects d_{jn} for each route. We impute traffic speed using the following equation:

$$\log(\widehat{\text{speed}_{jn,2009}}) = \hat{\beta}_{0,2009} + \hat{\beta}_{1,2009} \log(\text{distance}_{jn}) + \hat{\beta}_{2,2009} \log(\text{distance}_{jn})^2 + \bar{\mathbf{X}}_{jn}\hat{\Gamma}_{2009} + \hat{d}_{jn}$$

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

$$\text{time}_{jn,2009} = \exp\left(\log(\widehat{\text{speed}_{jn,2009}})\right) \cdot \text{distance}_{jn}.$$

A4 Job Choice Model with Spatial Sorting

In the main text, the model we use does not explicitly factor in spatial sorting. We define spatial sorting as households with different preferences for commuting that have demand for different neighborhoods. Empirically, we address this by including detailed location fixed effects interacted with observable characteristics.

However, in the counterfactual exercises, spatial sorting may be a material concern, since female workers may decide to live in different neighborhoods when their taste for commuting β is adjusted. In our exercise, however, we hold workers' residential locations fixed. To assess the potential bias of this omission, we write down a job choice model with spatial sorting.

We assume that the structure of the job choice set remains the same as in the main model:

$$\ln(w) = \xi_i + \beta_{ij} \ln(\tau - \tau^{min}),\ \tau \geq \tau^{min},$$

where β_{ij} is the return to commuting faced by worker i in neighborhood j.

Here, we allow workers' utility function to be a function of housing consumption and amenities. Worker i has instead the following utility function:

$$U(\tau, w) = \ln(w) - \lambda_i \tau - \kappa_p \ln P_j + \kappa_a A_j,$$

where P_j is the cost of housing in neighborhood j; A_j is the amenity level in neighborhood j.

Worker i solves a two-step problem:

1. Conditional on living in neighborhood j (given β_{ij}, P_j, A_j), worker i chooses a job that maximizes utility. From this step, we get worker i's indirect utility of living in neighborhood j.

2. Once we get the indirect utility associated with each neighborhood for worker i, we then solve for the optimal location and obtain the final indirect utility. Such final indirect utility is a function of the fundamentals: λ_i, κ_p, κ_a, and the spatial distribution of P_j and A_j.

Step 1 Conditional on living in neighborhood j, the optimal job choice is:

$$\tau^*_{ij} = \begin{cases} \frac{\beta_{ij}}{\lambda_i} + \tau^{min}, & \text{if } \beta_{ij} > 0 \\ \tau^{min}, & \text{if } \beta_{ij} = 0. \end{cases}$$

$$\ln(w_{ij}^*) = \begin{cases} \xi_{ij} + \beta_i \ln\left(\frac{\beta_{ij}}{\lambda_i}\right), & \text{if } \beta_{ij} > 0 \\ \xi_{ij}, & \text{if } \beta_{ij} = 0. \end{cases}$$

The indirect utility is as follows:

$$V(\beta_{ij}, P_j, A_j) = \begin{cases} \xi_{ij} + \beta_i \ln\left(\frac{\beta_{ij}}{\lambda_i}\right) - \beta_{ij} - \lambda_i \tau^{min} - \kappa_p \ln P_j + \kappa_a A_j, & \text{if } \beta_{ij} > 0 \\ \xi_{ij} - \lambda_i \tau^{min} - \kappa_p \ln P_j + \kappa_a A_j, & \text{if } \beta_{ij} = 0. \end{cases}$$

From the indirect utility, we can see that if workers are allowed to factor in the spatial distribution of the cost of housing P_j and the levels of amenities A_j, workers may trade off locations with high cost or amenities for places with low wage penalty for short commute.

Step 2 Given the indirect utility of living in each neighborhood j, worker i chooses a neighborhood that maximizes utility. To fix this concept, we assume that locations are indexed as $x \in [0, 1]$, 0 is the central city, and 1 is the periphery of the city. We assume that all neighborhoods characteristics are functions of x:

$$\begin{aligned} \beta &= \eta x \\ \xi &= \xi(\beta) = \xi(\eta x) \\ \ln P &= p(x) \\ A &= A(x). \end{aligned}$$

By doing this, we can rewrite the indirect utility as a function of the one-dimensional distance to central city x:

$$V(x) = \xi(\eta x) + \eta x \ln\left(\frac{\eta x}{\lambda_i}\right) - \eta x - \lambda_i \tau^{min} - \kappa_p p(x) + \kappa_a A(x).$$

The first-order condition in the search for optimal β is

$$V'(x) = \eta \xi'(\eta x) + \eta \ln(\eta x) - \eta \ln \lambda_i - \kappa_p p'(x) + \kappa_a A'(x) = 0.$$

By implicit function theorem and the fact that β is assumed to be proportional to x, we can easily show that the optimal β^* is decreasing in λ_i:

$$\frac{\partial \beta^*}{\partial \lambda_i} = \frac{\frac{\eta}{\lambda_i}}{V''(x)}.$$

We assume that $V(x)$ is strictly concave so that a unique solution for x^* and β^* is guaranteed. Given this assumption, we know that $V''(x) < 0$, and therefore $\frac{\partial \beta^*}{\partial \lambda_i} < 0$. The result is intuitive, because if the distaste for commuting is larger, workers are more likely to choose a place with lower wage penalty for short commute. The size of $\frac{\partial \beta^*}{\partial \lambda_i}$ depends on the size of $V''(x)$:

$$V''(x) = \eta^2 \xi''(\eta x) + \frac{\eta}{x} - \kappa_p p''(x) + \kappa_a A''(x).$$

Intuitively, it is likely that $\xi''(\eta x) < 0$, because if one is too far away from the city, local wages likely go down to close to zero. Also, it is likely that $p''(x) < 0$ and $A''(x) < 0$, because the marginal benefit of saving on housing cost and amenities by moving away from the city center likely decreases as one moves too far from the city. The derivative also depends on the sizes of κ_p and κ_a.

When we increase travel speed, we can think of the change as reducing the size of η. In other words, β does not increase as much as when distance to downtown (job center) increases. By implicit function theorem, we can show that

$$\frac{\partial \beta^*}{\partial \eta} = -\frac{\eta \xi'(\eta x) + \eta^2 x \xi''(\eta x) + \ln(\eta x) + 1 - \eta \ln \lambda}{V''(x)}.$$

The numerator represents how much the marginal utility with respect to x changes with η. The numerator term is intuitively negative, because holding the spatial distribution of cost of housing and amenities constant, larger η increases the utility cost of living farther from the center center. Given that the numerator is negative, $\frac{\partial \beta^*}{\partial \eta} < 0$.

This means, when lowering η (increasing travel speed), that it is possible that workers would sort into neighborhoods in which β is larger, offsetting some of the reduction in the wage gap. The strength of the offsetting effect depends on the size of $V''(x)$, and ultimately the sizes of κ_p and κ_a and $P''(x)$ and $A''(x)$. The intuition is that the reduction in η reduces the marginal cost of moving away from the city center, and thus could encourage workers to sort away from the center. But if the marginal benefit of moving away from the center decreases rapidly or the marginal cost of moving away from the center increases rapidly, the reduction η would not be strong enough to generate enough incentive for workers to move away from the center.

Table A1: Summary Statistics: NHTS

Variables	Men			Women		
	Mean	SD	Obs.	Mean	SD	Obs.
Age	43.54	10.83	34,155	43.19	10.99	28,882
High school or lower	0.198	0.398	34,155	0.137	0.343	28,882
Some college	0.251	0.434	34,155	0.289	0.454	28,882
College or higher	0.549	0.498	34,155	0.573	0.495	28,882
Trip time to work w/o traffic	20.67	16.65	28,593	17.77	12.79	24,798
Trip time to/from work	29.07	28.1	34,155	25.48	22.49	28,882
Trip time to and from work with other chained trips	50.71	48.39	34,155	46.59	42.95	28,882

Note: The sample comprises full-time workers aged between 25 and 64 whose travel day is a weekday from the 2017 NHTS. Travel times are measured in minutes.

Table A2: Gender Differences in Trip-chaining Behaviors and Commute Time

Variables	Have a trip for personal/family business		Link personal trips to trips to/from work		Log time to work without traffic		Log time to/from work		Log time to and from work with chained trips	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female	0.072***	0.066***	0.116***	0.109***	-0.124***	-0.131***	-0.134***	-0.130***	-0.122***	-0.127***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	63,037	62,704	63,037	62,704	99,882	99,368	53,204	52,934	50,643	50,373
R-squared	0.006	0.027	0.013	0.045	0.005	0.090	0.008	0.075	0.007	0.096
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: The sample comprises full-time workers aged between 25 and 64 whose travel day is a weekday from the 2017 NHTS. The control variable includes dummies for age, race, Hispanic origin, education, and MSA of residence.