Re-assessing the Spatial Mismatch Hypothesis*

David Card — UC Berkeley
Jesse Rothstein — UC Berkeley
Moises Yi — US Census Bureau

March 2024

ABSTRACT

We use detailed location information from the Longitudinal Employer-Household Dynamics (LEHD) database to develop new evidence on the effects of spatial mismatch on the relative earnings of Black workers in large US cities. We classify workplaces by the size of the pay premiums they offer in a two-way fixed effects model, providing a simple metric for defining "good" jobs. We show that: (a) Black workers earn nearly the same average wage premiums as whites; (b) in most cities Black workers live closer to jobs, and closer to good jobs, than do whites; (c) Black workers typically commute shorter distances than whites; and (d) people who commute further earn higher average pay premiums, but the elasticity with respect to distance traveled is slightly lower for Black workers. We conclude that geographic proximity to good jobs is unlikely to be a major source of the racial earnings gaps in major U.S. cities today.

*We thank Richard Jin and Nicole Gandre for excellent research assistance and Ellora Derenoncourt, Patrick Kline, and Stephen Ross for valuable comments. Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau. The Census Bureau has ensured appropriate access and use of confidential data and has reviewed these results for disclosure avoidance (Project 6000266: CBDRB-FY24-CES024-001/002).

In a pair of influential papers written in the early 1960s, Kain (1964, 1968) proposed the *spatial mismatch* hypothesis: that the combination of residential segregation and the suburbanization of jobs was limiting the employment opportunities of African Americans and contributing to the wave of civil unrest in many large US cities.[1] Kain's hypothesis attracted much attention at the time, and a large body of subsequent work has focused on whether the different residential locations of Blacks and whites can explain part of the persistent racial gap in labor market outcomes (see Holzer, 1991; Ihlanfeldt and Sjoquist, 1998; Gobillon, Selod and Zenou, 2007; and Wang, Wu, and Zhao, 2022 for reviews).

We use data from the 2010-2018 Longitudinal Employer-Household Dynamics (LEHD) program to evaluate the impact of spatial mismatch (SM) on the racial earnings gap in contemporary labor markets. LEHD has detailed information on the place of residence and place of work for of nearly all wage and salary earners in the U.S., offering a significant advance in the quality and quantity of data available for assessing SM. We also introduce a new measure of job quality to the SM literature: the establishment earnings premium, estimated from an Abowd-Kramarz-Margolis (1999) (hereafter AKM) earnings decomposition. The combination of LEHD data and the AKM framework allows us to measure both the local availability of all jobs for any given worker and the availability of "good" jobs, as defined by pay premiums they offer relative to a typical job in the same market.

We focus on racial differences in job accessibility in two major groups of cities: older industrial cities in the Northeast and Midwest; and newer Sunbelt cities in the South and Southwest. We also examine racial gaps in realized commute distances and in the wage premiums that workers receive when they commute further. These comparisons parallel the outcome tests widely used in the literature on discriminatory policing (e.g., Knowles, Persico and Todd, 2001) and help to address the concern that job *proximity* may give a misleading picture of job *opportunities* if Black applicants are less likely to be hired at higher-quality jobs.

Our first key finding is that virtually none of the 30+ percentage point Black-white earnings gap among full-time workers reflects differences in average establishment pay premiums. Rather, the entire premium is attributable to the "worker effect" in the AKM

[1] Kain (2004) presents a history of his work on this topic and discusses some of its impact.

decomposition. This result is the opposite of what we would expect if access to high-paying jobs was a major driver of the racial wage gap. Nevertheless, it aligns with longstanding evidence that Black workers are, if anything, more likely to be employed in unionized workplaces than white workers (e.g., Ashenfelter, 1972; Rosenfeld and Kleykemp, 2012).

Next, we examine the geographic distribution of workers and jobs. We show that on average Black workers' homes are, if anything, closer to potential workplaces, and to workplaces with higher earnings premiums, than are white workers' homes. Moreover, jobs near Black neighborhoods tend to offer *higher* average pay premiums than do jobs near white neighborhoods. Again, these findings are the opposite of what one might expect if greater distances to jobs (and particularly to good jobs) creates barriers for Black workers.

Ellwood (1986) pointed out that in an equilibrium model with spatial mismatch, Black workers will have longer commute distances than white workers.[2] A similar prediction arises if Black and white workers are equidistant to jobs but some employers discriminate in hiring.[3] Our third key finding is that Black workers' average commutes are *shorter* than those of white workers in the older industrial cities that have been the focus of most previous studies of spatial mismatch, though slightly longer than whites in a group of Sunbelt cities. Among both white and Black workers longer commutes are associated with higher-paying jobs (i.e., jobs at establishments with higher AKM pay premiums), as one might expect if workers trade off pay and commuting distance. But the association is slightly weaker for Black than white workers, contrary to the pattern that would be expected if SM limited Black workers' access to high-paying jobs.

Overall, we interpret the collage of evidence as suggesting that spatial mismatch is not a major explanation for Black-white earnings differences in the 2010s. A limitation of our analysis is that we focus exclusively on full time earners -- people who earned at least $3800 in a given calendar quarter. We have not attempted to measure the effects of job accessibility on racial

[2] This prediction is more nuanced if the probability of being observed with a job varies by race.

[3] There is an extensive literature (e.g., Bertrand and Mullainathan, 2006) on the use of "audit study designs" to see if employers are equally likely to follow up with job applicants who signal Black race through their name (or some other feature). A recent large-scale study by Kline, Rose and Walters (2022) finds that job applicants with distinctively Black names have about a 2.1 percentage point lower probability of a follow-up contact than whites – a roughly 10% lower rate.

differences in employment or unemployment, which were the primary focus of the early SM literature. In addition, our data cover only the past decade. The fraction of the Black urban population living in suburbs (versus central city areas) has risen substantially in the past 50 years − from 18% in 1970 to 40% in 2010 − arguably leading to a reduction in spatial mismatch.[4] Whether an analysis parallel to the one we present here but using data from the 1960s or 1970s would show more evidence of SM is an important question for future research.

I. Previous Research on Spatial Mismatch

Kain's (1968) original specification related the share of jobs held by Black workers in a tract (technically, a "workplace area") in Chicago or Detroit to the Black share of residents in the tract and the distance to the nearest ghetto. He then performed a counterfactual analysis assuming that Black residential shares were equalized across tracts. The results suggested that the elimination of residential segregation would significantly increase Black employment, though Offner and Saks (1971) showed that that these conclusions were quite sensitive to Kain's assumption of a linear relationship between the Black share of jobs and the Black share of residents in a tract.

Subsequent work tried to test for spatial mismatch more directly by relating the outcomes of Black residents in different neighborhoods to the potential travel distance to jobs (e.g., Hutchinson, 1974). Ellwood (1986) related tract-level employment rates of youth in different census tracts in Chicago to measures of job accessibility, including the proportion of all jobs in the city within a 30 minute commute. He found uniformly small effects, leading him to conclude that job proximity was not a major determinant of youth joblessness. Raphael (1998) extended this analysis by considering tract-level measures of local demand and local supply of low skill workers. Using employment <u>growth</u> in nearby tracts as a measure of local demand he was able to account for up to 1/3 of the negative correlation between the share of Black

[4] See Massey and Tannen (2018, Figure 1). During the same period the fraction of the white urban population living in suburbs rose from 50% to 63% - thus the relative share of Black versus whites in suburbs rose 0.37 to 0.63.

residents in a tract and the employment rate of teenagers, though the effects of local supply and demand were actually concentrated on whites and were insignificant for Blacks.[5]

A few more recent studies have used information on where people live and work to try to measure the effects of SM. For example, Hellerstein et al. (2008) used restricted micro data from the 2000 Census to construct estimates of the number of *jobs* in a zip code area held by workers with certain characteristics, and the number of *residents* in the same zip code area with those characteristics. They then fit models relating the employment rate of Black men in a zip code area to the local number of jobs per resident. They find that Black male employment rates are higher in zip codes where there are more jobs filled by Black men per Black male resident, but that there is no effect of local jobs employing white men – a pattern they interpret as evidence of racial segregation in jobs.[6]

Andersson et al. (2018) is the only previous study we are aware of that has used LEHD data to examine spatial mismatch. These authors model the effects of local job availability on the elapsed time to a new job for workers affected by mass layoff events. As a measure of job availability they use the proportional gap between supply and demand for jobs faced by residents of a given Census tract, based on weighted averages of employment and population counts in nearby tracts (similar to Raphael, 1998). They show the estimated effects of job availability on time to a new job are somewhat larger for Blacks than whites. When they measure race-group-specific supply and demand, however, the effects are more similar, suggestive of a racial component in job matching (as in Hellerstein et al., 2008).

II. A Simple Model

i) Basic Setup

We begin by sketching a simple model of wage outcomes for workers in a spatially differentiated labor market. The model explicitly builds in an AKM-style model of wage setting

---

[5] Raphael (1998, Figure 4a) showed that the share of all jobs within a given commute time from a neighborhood was higher for the neighborhoods of Black teens than those of white teens – a finding we reproduce in our data. But he argued that employment growth is a better proxy for local job opportunities than the level of employment.
[6] A concern with this specification is that omitted factors could shift both Black jobs per Black resident in a zip code and the employment rate of Black men who live in the zip code.

in which each establishment offers a proportional wage premium that raises or lowers wages of any worker who is employed there relative to other workplaces in the market. As in monopsonistic competition models (e.g., Card et al., 2018), we focus on worker's preferences over available job packages, ignoring frictions in the matching process.

Specifically, assume that person i gets utility from employer j:

$$u_{ij} = a_i + \delta_j - \beta_i d_{ij} + \epsilon_{ij},$$

where δ_j is the pay premium offered by employer j, d_{ij} is a measure of the commute distance for i to get to workplace j, and ϵ_{ij} is a match effect. If worker *i* takes a job at employer *j* her observed wage is:

$$\ln w_{ij} = \alpha_i + \delta_j + v_{ij},$$

where (as in a standard AKM model) α_i is a common component of wages for i across all jobs, and the residual term v_{ij} is assumed to be independent of d_{ij} and ϵ_{ij}.

Next, assume that a worker who is searching for a job has an "offer set" O_i representing the potential set of job opportunities that are available. She takes the job with the highest utility in the set:

$$j^*(i) = argmax_{j \in O_i}\left[\delta_j - \beta_i d_{ij} + \epsilon_{ij}\right],$$

and we observe the combination of the wage premium and commute distance $(\delta_{j^*(i)}, d_{ij^*(i)})$ for that worker. Spatial mismatch can be expressed in this framework as a difference in the offer sets available to workers who live in different neighborhoods.

ii) Comparing Job Opportunities of Different Groups

Suppose there are two groups of workers G_1 and G_2. If the joint distributions of $\left(\delta_j, d_{ij}\right)$ in the offer sets are the same for the two groups, and they have the same distribution of $\beta_i's$, then they will have the same probability distributions over $(\delta_{j^*(i)}, d_{ij^*(i)})$. In particular, the conditional expectation of the wage premium, given commute distance,

$$E\left[\delta_{j^*(i)} \middle| d_{ij^*(i)} = d\right]$$

will be the same for the two groups. This provides the basis for a simple outcome test: if two groups have the same access to jobs, and the same preferences for wages versus commuting distance, then we would expect the observed relationship between wage premiums and commute distances to be the same for the two groups.

To facilitate comparisons between workers with <u>different</u> offer sets, suppose that commute distances are discrete, $d_{ij} \in \{d_1, d_2, \dots d_N\}$, and that wage premiums are also discrete, $\delta_j \in \{\delta_1, \delta_2, .. \delta_M\}$. In this case the offer set for a given worker is represented by a 2-dimensional grid showing which particular combinations of (δ_u, d_v) are available (i.e., the support of the joint distribution of wage premiums and commute distances). For example a high-wage premium job at close proximity may not be available in a given worker's choice set.

Suppose that the offer sets for individuals in group G_1 have the property that jobs with wage premiums $\delta_j \in \{\delta_1, \delta_2, .. \delta_M\}$ are available at every commute distance, while the offer sets for individuals in group G_2 have the property that jobs with wage premiums $\delta_j > \bar{\delta}$ are only available with commute distances $d_{ij} > \bar{d}$. In this case we would say that the job opportunities of group G_2 are negatively affected by their residential locations, relative to group G_1. In particular we would expect that the observed wage premiums for workers in G_2 with relatively short commuting distances would be lower than the premiums for workers in G_1 in the same range of commute distances. We would also expect that the slope of the conditional expectation of the wage premium, given the commute distance, will be higher for the disadvantaged group.

III. Data sources

(i) LEHD Sample

We use data from Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) program. LEHD is derived from the quarterly earnings reports provided by employers to state unemployment insurance (UI) agencies. The core data set includes the total wages paid by a given employer to each worker in a quarter. These data are supplemented with information on employers and workers derived from other sources (e.g., the decennial census and ACS files) – see Abowd et al. (2009). The LEHD covers about 95% of private sector employment, as well as state and local government employees, but excludes federal employees, members of the armed services, and self-employed workers. From 2010 forward it includes data from all 50 states.

Our sample construction follows Card, Rothstein and Yi (2024a). We begin with person-employer-quarter (PEQ) observations from 2010Q1 to 2018Q2 where the worker is between 22 and 62 years of age. To help screen out part-time jobs and/or partial-quarter job spells we exclude quarters where an individual had multiple jobs or earnings below $3,800 (roughly the earnings from a full-time job at the federal minimum wage), as well as all *transitional* quarters (the first or last quarter of any person-employer spell). We also drop PEQs with an unknown industry and/or establishment location. We exclude individuals with fewer than 8 quarters of earnings that satisfy the previous restrictions over our 8½ year sample window. We also drop people with Hispanic ethnicity and all those whose primary race is not Black or white. For simplicity, in the remainder of the paper we refer to "white" and "Black" workers without repeating the qualifier that these designations exclude Hispanics of any race. We assign PEQ's to 1990 Commuting Zones (CZs) (Tolber and Sizer, 1996) based on the county of the employing establishment in that quarter.

Table 1 reports summary statistics for employed quarters of white and Black workers in three groups of CZs: (1) a set of 10 older industrial cities in the Northeast and Midwest (Philadelphia, Detroit, Pittsburgh, Cleveland, Newark, Buffalo, Baltimore, Chicago, Minneapolis and St. Louis); (2) a set of 7 newer Sunbelt cities in the South and Southwest (Los Angeles, Houston, Atlanta, Miami, Dallas, San Diego, and Phoenix); (3) the remaining 183 of the largest 200 CZs in the country. Not surprisingly, the first two city groups – which are drawn exclusively from the largest 30 CZs – have somewhat higher earnings for both white and Black workers than does the latter group. They also have somewhat longer average distances between home and workplace: e.g., for white workers, the average commute is 12.2 miles in older industrial CZs, 14.4 miles in the Sunbelt CZs, and 10.4 miles in the remainder of the top 200 CZs.

The third panel in the table shows the white-Black differences in mean log quarterly earnings and mean commute distances. The wage gaps are relatively large, reflecting the fact that white workers tend to have higher hourly wages and more hours of work per unit of calendar time than Black workers. As has been found in other settings (though typically with more selective samples), white workers tend to have longer travel-to-work distances than Black workers – on the order of 14% longer in the older industrial CZs, 1% longer in the Sunbelt CZs, and 3% longer in the other CZs.

(ii) Imputation of Establishment Locations in LEHD

The UI data in the LEHD contain an identifier for the employing firm and the state, but not for the specific establishment of firms with multiple workplaces in the same state.[7] The Census Bureau uses the relative distance from a worker's home address to the locations of the establishments owned by the firm to (probabilistically) impute his or her specific establishment (Vilhuber 2018). We use the first of the multiple imputations available in the LEHD to assign PEQs to establishments. To evaluate the implications of imputation errors we divide establishments into categories based on whether the firm has 1 or more establishments in the same state, and on the size of the firm. Specifically, we use the largest number of PEQs at a firm in the quarters of our sample to define three firm size groups: ten or fewer workers; 11-276 workers; and greater than 276 workers.[8]

(iii) Coding of geographic locations in LEHD

The Census Bureau assigns detailed geographic locations at an annual frequency to workers' residences and to establishment locations. We use the latitude and longitude of a worker's home and workplace to compute the as-the-crow-flies commute distance for each worker, in miles. To analyze the number of jobs within a radius r of each worker, we coarsen the locations of firms and workers to a set of grid points spaced 0.5 miles apart in both the North-South and East-West locations. Commute distances computed using this grid are very highly correlated with distances that use the original (uncoarsened) locations, so we do not believe much precision is lost with this coarsening, which dramatically reduces computational burden. Because CZs differ widely in their physical size, and we want to construct averages across CZs in a group, we standardize travel distances by a multiplicative factor so that the 75th

[7] We define a "firm" as a unique State Employer Identification Number.
[8] Our calculations of firm size, unlike our other LEHD analyses, include Hispanic workers.

percentile of commute distances in each CZ is 16 miles. Thus, if the 75th percentile in a particular CZ is 12 miles, we multiply all distances by 4/3, whereas if the 75th percentile is 24 miles, we multiply all distances by ¾.

(iv) Comparisons to the American Community Survey

To help contextualize our LEHD data, we constructed a parallel sample of workers in the 2010-2018 American Community Survey (ACS). We select people ages 22-62 from the ACS with at least 1 year of potential experience (i.e., age-education-6 > 0). For our analyses of earnings outcomes we further limit attention to "full year earners" with annual wage and salary earnings of $15,200 or higher -- a threshold 4x higher than the quarterly threshold for full time work we impose on the LEHD.

We assign 1990 CZs to Public Use Micro Areas (PUMA's) identified in the ACS using PUMA-county population files for the 2000 and 2010 Census created by David Dorn.[9] For the relatively small share of PUMAs that contain observations from multiple CZs we probabilistically assign one CZ based on the relative share of the PUMA population in that CZ.[10] Finally, we limit attention to individuals in the 30 largest CZs (based on counts of person quarter observations in our LEHD samples) and divide the "top 30" CZs into 4 groups: (1) Older industrial cities (as defined above); (2) Newer Sunbelt cities (also as defined above); (3) Northeast Corridor (New York, Washington DC, Boston, and Hartford CT); (4) Other large CZs (San Francisco, Seattle, Denver, Sacramento, San Jose, Portland, Tampa, Orlando and Fort Worth). The resulting sample contains 6.49 million observations, representing a weighted population of roughly 80 million 22-62-year-olds per year.

Table 2 reports summary statistics for the working age populations in the top 30 CZs, as well as the characteristics of the full-time earners with at least $15,200 in earnings. (Note that in contrast to the statistics in Table 1, these results include people of all ethnicities and racial

[9] See https://www.ddorn.net/data.htm. We downloaded two files from this site: [E5] 2000 Census and 2005-2011 ACS Public Use Micro Areas to 1990 Commuting Zones; and [E6] 2010 Census and 2012-ongoing ACS Public Use Micro Areas to 1990 Commuting Zones.

[10] For example, if 70% of the population in a PUMA is from CZ "A" and 30% from CZ "B", then we randomly assign 70% of the workers in our sample from that PUMA to CZ "A", and 30% to CZ "B". We also experimented with a multiple allocation approach, and found it gave nearly identical estimates of the characteristics of larger CZs.

groups.) For reference the first line of the table shows the relative sizes of the 4 groups of CZs in the top 30 (weighted by the ACS sample weights). The samples of person-year observations in the older industrial CZs and newer Sunbelt CZs are fairly similar in size, while the samples for the Northeast corridor CZs and remaining top 30 CZs are smaller.

Within the working-age population, the relative shares of whites and Blacks vary across the 4 groups of CZs, with whites being relatively under-represented in the Sunbelt CZs and Blacks being relatively under-represented in the "remainder" group. Mean years of education and the share of people with a BA or higher also vary somewhat across CZ groups and are relatively low in the Sunbelt (driven in part by the high share of Hispanics in these cities). Average employment rates (based on having positive earnings in the previous year) are fairly similar across city groups, ranging from 79 to 82%; the fraction of full-time earners varies a little more (from 58% to 65%) and is lowest in the Sunbelt cities.

Looking next at the full-year earners, we see that across the CZ groups the share of Black workers is 12-15% except in the remainder CZ group. The share of females is more stable at about 45%. Mean annual earnings range from $61,000 in the Sunbelt cities to $75,000 in the Northeast corridor; mean hourly wages range from $29 to $35 per hour. On average about 86% of full-time earners in the top 30 CZs commute to work in their own car: this rate is higher in the Sunbelt (around 94%) and lower in the Northeast corridor (67%). Mean commute times average about 30 minutes (one way), but are about 5 minutes longer in the Northeast corridor, partly reflecting the fact that commuters by bus and rail have relatively long average commute times and these modes are more common there.

Finally, the bottom three rows of the table show mean log annual earnings of white non-Hispanics, Black non-Hispanics, and the Black-white earnings gap. Importantly, the magnitudes of the Black-white gaps in annual earnings in our ACS sample are similar to the gaps in quarterly earnings in our LEHD sample: the mean gap is about 32 log points and is slightly lower in the older industrial cities than the Sunbelt cities or the Northeast corridor. About one-eighth of the gap in earnings for full-year earners appears to be due to lower hours among Black workers. The Black-white gap in hourly wages for full-year earners in the largest 30 CZs is 28 log points. This gap is not too different from the 26 log point gap in log hourly wages for

2010-2018 reported by Wilson and Darity (2022), based on hourly or weekly wages reported in the monthly Current Population Surveys.

Appendix Table 1 presents a few salient characteristics of each of the CZs in our top 30 ACS sample. There are some similarities and some differences between the CZs in each of the 4 groups. For example, most of the older industrial cities have around a 20% share of Black workers, though Minneapolis and Pittsburgh are exceptions. The Sunbelt cities are more heterogeneous in this dimension, with a Black share of around 35% in Atlanta, 15-20% in Houston, Dallas and Miami, and only 5-6% in Los Angeles, Phoenix, and San Diego. Average annual salaries range from \$54,000 (Cleveland) to \$75,000 (Newark) in the older industrial cities, but are more narrowly clustered between \$53,000 (Miami) and \$64,000 (Houston) in the Sunbelt cities. Average one-way commute times vary somewhat across CZs: highest in New York City (39 minutes) and lowest in the smaller CZs (e.g. about 25 minutes in Cleveland, San Diego, Hartford, and Portland).

IV. AKM Model and the role of establishment pay premiums in the racial wage gap

(i) AKM Model

Using our LEHD sample for each CZ, we fit a standard AKM model to the log of quarterly earnings for person *i* in quarter *t:*

$$y_{it} = \alpha_i + \delta_{f(i,t)} + X_{it}\beta + \epsilon_{it}. \quad (1)$$

Here α_i is a person effect for worker *i*, $f(i,t)$ is an index function giving the workplace for *i* in quarter *t*, δ_f represents the pay premium at establishment *f*, and X_{it} is a vector that includes a full set of calendar quarter indicators and a cubic in worker age. We estimate (1) separately for each CZ, pooling Black and white workers but limiting to the largest connected set in the CZ (which typically includes well over 95% of PEQs in the CZ). We normalize the (person-quarter-weighted) average pay premium of all establishments in the restaurant industry in each CZ to zero. Thus, $\hat{\delta}_f$ can be interpreted as the pay premium at establishment *f* relative to the average pay at restaurants in the same CZ.

It is well known that OLS estimation of an AKM style model will lead to biased estimates of the firm and person effects unless the unobserved determinants of pay, included in the error term ϵ_{it}, are uncorrelated with the sequence of jobs held by individual *i*. In Card, Rothstein and Yi (2024a) we present a number of checks for this "exogenous mobility" condition. We find evidence of small but systematic departures from exogenous mobility for samples similar to the ones used here. We conclude, however, that these departures lead to relatively small biases in the estimated establishment effects.

After estimating (1), we average the left-hand and right-hand sides by CZ and race, then take the difference between whites and Blacks in each CZ. This yields:

(2) $$\bar{y}_{cw} - \bar{y}_{cb} = (\bar{\alpha}_{cw} - \bar{\alpha}_{cb}) + (\bar{\delta}_{cw} - \bar{\delta}_{cb}) + (\bar{X}_{cw} - \bar{X}_{cb})\hat{\beta},$$

where $\bar{y}_{cw}$ and $\bar{y}_{cb}$ represent the means of log earnings for white and Black workers in CZ *c,* respectively, $\bar{\alpha}_{cw}$ and $\bar{\alpha}_{cb}$ represent the means of the estimated person effects for white and Black workers in that CZ, $\bar{\delta}_{cw}$ and $\bar{\delta}_{cb}$ represent the means of the estimated establishment effects for the two groups, and $\bar{X}_{cw}$ and $\bar{X}_{cb}$ represent the mean vectors of covariates.

Let s_{fcw} and s_{fcb} represent the shares of all PEQs of white and Black workers in CZ *c* that worked at establishment *f*. Then

(A-3) $$\bar{\delta}_{cw} - \bar{\delta}_{cb} = \sum_{f \in C} (s_{fcw} - s_{fbw})\, \hat{\delta}_f.$$

Thus, the second term in equation (2) can be interpreted as measure of the differential sorting of whites relative to Blacks to workplaces with a higher estimated pay premium. If Black workers are less likely than whites to be employed at such workplaces, this term will be negative.

Table 3 presents a summary of the terms in equation (2) for the three groups of CZs in our LEHD sample. The first row presents the mean white-Black earnings gap in our estimation sample (based on the largest connected set, comprising well over 95% of PEQs, in each CZ). This is very close to the gap across all full-time earners in Table 1. In all three groups of CZs the mean white-Black gap in person effects is slightly *larger* than the gap in earnings, while the mean gap in establishment effects is small and in two of the three CZ groups actually negative – implying that, for example, the mean establishment pay premium for white workers in the third group of CZs is about 1.3 percentage points lower than the mean premium for Black workers. In our two focal groups of CZs the mean gap in establishment effects is not statistically significantly different from zero. Thus, any differential sorting of Black and white workers to high-paying establishments in these two sets of CZs is negligible.

In Card, Rothstein and Yi (2024b) we present a simple decomposition of establishment pay premiums into the mean by industry and the deviation of the establishment premium from the average for its industry, which we call an "industry hierarchy effect". The bottom rows of Table 3 use this approach to decompose the Black-white difference in mean establishment pay premiums into the difference in mean industry wage effects and the difference in mean hierarchy effects. Interestingly, for the two focal groups of CZs these components have opposite signs: Black workers work in slightly lower-paying industries than whites, but within a given industry they are employed at slightly <u>higher-paying</u> establishments.[11]

(ii) Interpretation

The fact that estimated average pay premiums for white and Black workers are nearly the same in our LEHD sample is potentially surprising, and different than the pattern in Brazil reported by Gerrard et al. (2022), where nonwhite workers are employed at lower-premium workplaces. As in other settings, in our sample people with higher values of α_i tend to work at establishments with higher pay premiums – a pattern of "positive assortative matching." In particular, our estimates imply that a 10% increase in α_i is associated with about a 1% increase in $\delta_{f(i,t)}$ within a CZ. Given the 30-35 log point gap in the mean of α_i between Black and white workers, one might have expected a roughly 3 log point gap in average pay premiums between Blacks and whites just because of assortative matching, rather than the 0 that we observe. We do find that Black workers are slightly less likely to work in higher-paying <u>industries</u>, but this is offset by the tendency to be employed at higher-premium workplaces within a given industry.

The pattern of sorting within industries is potentially consistent with a longstanding fact about the U.S. labor market, which is that Black workers are more likely to be covered by unions than whites (e.g., Ashenfelter, 1972). Data from the unionstats.com website shows that the ratio of the Black to white union coverage rate for both male and female workers averaged

[11] We also find a tendency for whites to work in higher-premium industries in the ACS. Using estimated pay premiums for 295 4-digit industries from a standard Mincer-style model we find that the average industry premium is 3 log points higher for whites than Blacks. Card, Rothstein, and Yi (2024b) show that the industry premiums in this specification include a component reflecting differences in worker type (i.e., in α) that are not captured by the available controls in the ACS and that this component is positively correlated with the industry premium. This accounts for the difference in magnitude from Table 3.

about 120% in years 2010-2018. Other things equal, the higher union coverage rate of Black workers presumably led to modest narrowing of the white-Black wage gap

V. Job access and commuting patterns by race

i) Job access

Figure 1 begins our analysis of the relative accessibility of jobs for Blacks relative to whites. For each worker and for varying radii r, we compute the share of all jobs in the CZ that are located within radius r of the worker's residence. We average this across Black and White workers separately in each CZ, then average across CZs in our CZ groups (normalizing distances within each CZ as discussed above). Panel a shows that for every r, the cumulative share of jobs within r miles of the typical Black worker's home is higher than the share for a typical white worker in the older industrial CZs (compare the solid orange line to the solid blue line). Panel b shows the cumulative fraction of jobs within radius *r* for Black workers relative to the cumulative fraction of all jobs in the CZ within the same radius for white workers (a ratio we call the "relative fraction" of jobs within that radius). We can see that at all distances, Black workers in the older industrial CZs are closer to jobs. Moreover, the jobs near Black workers are of somewhat higher quality (as measured by the establishment premium) than those near White workers. The dashed lines in the figures show the share of jobs at establishments with estimated pay premiums in the top tercile for the CZ within radius r; these tell a similar story as the all-jobs series.

Looking at data for the Sunbelt CZs in panels c and d, the racial differences in job accessibility are smaller, possibly reflecting the multi-centric structure of many of these cities. Nevertheless, for radii of 1 mile or more, Black workers in these CZs are also closer to jobs. There is no indication that there is a systematic shortage of jobs, or of good jobs, within a reasonable commuting distance of Black workers.

A potential caveat to this conclusion is that not all workers are qualified for all jobs, and since Black workers have lower education than whites, an analysis of all jobs as a whole may give a misleading impression of access to jobs available for a representative Black worker. Appendix Figure 1 follows the same format as Figure 1, but separates jobs into those held by workers with a high school education or less (non-college jobs) and those held by workers with at least some college education (college jobs). We find that access to non-college jobs is uniformly higher for non-college Blacks than for non-college whites in both the older industrial CZs and the Sunbelt CZs, but access to college jobs at very near proximity (<1 mile) is slightly lower for college-educated Blacks in the Sunbelt CZs.

Another way to measure access to good jobs is via the correlation between the fraction of residents at a location who are Black and the average premium of all establishments within a short commuting range of that location. We assign workers to locations defined by a 0.5-mile-by-0.5 mile grid, and measure the average pay premiums of all establishments within 2.5 miles of each location. The correlation of this measure of nearby job quality with the fraction of Black workers at the location is 0.26 for the older industrial CZs and 0.10 for the newer Sunbelt CZs. This approach confirms that if anything, jobs near Black neighborhoods tend to offer higher pay premiums than those near white neighborhoods.

ii) Commute distances and job quality

While proximity to jobs has been widely used in the literature to assess the SM hypothesis, many previous authors have expressed the concern that Black job-seekers may not have equal access to jobs (e.g., Hellerstein et al. 2008; Andersson et al., 2018). One way to address this concern is to look at realized commute distances. As discussed in section II, if Black workers have access to fewer high-paying jobs within a short commute distance, they will have to travel further to obtain a job with a given pay premium. To what extent is that true?

Figure 2 shows the densities and cumulative distribution functions of commute distances for Black and white workers in our two groups of CZs.[12] (For display purposes we show commute distance on the x-axis using a $\log_2$ scale). In older industrial CZs the CDF for black workers is shifted left, implying that the quantiles of Black commuter distance are uniformly lower. In the Sunbelt cities the CDFs cross once at about the 60th percentile: thus the quantiles of commute distance for Blacks are above the corresponding quantiles for whites up to the 0.6 quantile, and below the quantiles for whites at higher quantiles.

[12] We use a kernel density procedure to estimate the densities of log commute distance by race, then integrate the estimated densities to construct the CDFs.

Table 4 reports selected quantiles of commute distance for white and Black workers in different groups of CZs. For comparison, we also show the quantiles of 1-way commute times from the ACS. (The ACS does not report commute distances.) In the older industrial CZs the distance quantiles for Black workers are lower than the quantiles for white workers, whereas the travel time quantiles are the same.[13] In the newer Sunbelt CZs, consistent with Figure 2, the lower quantiles of distance for Blacks are above those for whites whereas the 75th and 90th percentiles for Blacks are lower. In terms of commute times, however, all the quantiles except the median are the same for Blacks and whites. Looking at commute times in the Northeast corridor we see a larger and more systematic gap between Blacks and whites, with 5 minutes longer commute times at the 10th and 25th percentiles and 10 minutes longer times at the 75th and 90th percentiles. Some share of this extra time may be due to the fact that in the Northeast CZs the white-Black gap in use of a car to get to work is 14 percentage points (75% of whites use a car versus 61% of Blacks) whereas in other CZs the gap is only 5 percentage points (93% or whites versus 88% of Blacks).[14] Finally, in the other top 30 CZs commute times are very similar for whites and Blacks.

Figure 3 shows how commute distance is related to job quality, as measured by the average pay premium earned by workers who travel a given distance. In both groups of cities and for both races, longer commutes are associated with higher pay premiums, as would be expected if workers trade off wages against commute time in job search. In the older industrial CZs Black workers with commute distances up to 10 miles work at establishments with higher average pay premiums than whites who travel the same distance; at further distances the relationship to pay premiums flattens out and even turns slightly negative for Blacks. In the Sunbelt CZs, average pay premiums conditional on commute distance are higher for Blacks up to about a 2 mile commute, then very similar for commute of 2-8 miles, then a little higher for whites. Again, the relationship flattens out after about 20 miles, and goes negative for both race groups in the upper tail of commuting distances (though we caution that the share of jobs at commutes beyond 30 miles is small, and is not evenly balanced across CZs).

---

[13] Travel times reported in the ACS are typically rounded to 5 minute intervals, which accounts for the fact that all the quantiles end with 5 or 0.

[14] Mean commute times are about 20 minutes longer for commuters who use other transit modes relative to car.

The inverse-U shape of the lines in panel b of Figure 3 is unexpected. A possible explanation for this shape is errors in the imputation of establishment for people who work at firms with multiple establishments. The establishment imputation model used by LEHD does not take account of establishment earnings premiums, so if workers are more willing to accept long commutes to work at high-premium establishments, the average premium of imputed establishments with long commutes may overstate the average premium obtained by workers who actually commute such long distances. To assess this hypothesis, we estimated the commute distance-pay premium relationship separately for workers in five groups of firms – small firms, with no more than 10 workers in any quarter; larger firms that are still below median in size (11-276 employees in a quarter), separately by whether they have one or multiple establishments; and above-median-sized firms (>276 employees in any quarter), again separately by whether they have one or multiple establishments. Imputation error should only impact the patterns at multi-establishment firms.

Appendix Figure 2 shows that the estimated relationships between commute distance and mean pay premium for the single establishment firms are fairly well-behaved for both white and Black workers. For multi-establishment firms, however, we see a pattern of declining average premiums after commutes of around 25 miles. These patterns suggest that part of the decline in premiums after about 25 miles observed in Figure 3 is due to imputation errors, though there is also some decline in premiums for Black workers at single establishment firms with 11-276 employees.

iii) Commute distances by worker skills

The comparisons of commute distances in Table 4 make no allowance for the fact that workers with different levels of skill may commute different distances. Most of the existing literature finds that commute distances are increasing in worker income. This pattern is often attributed to the demand for larger housing units by high-income families, but some of the correlation between income and commute distance may be due instead to the tradeoff between distance and pay premiums highlighted in our simple model.

Figure 4 shows quantiles of commuting distance by decile of worker fixed effect, separately by race and CZ group. In both groups of CZs we see that commute distances are longer for higher-α workers, though the gradient in the Sunbelt CZs is relatively modest (only a 17% rise in median commute distance between whites in the bottom and top deciles) compared to the gradient in the older industrial CZs (a 43% rise in median commute distance between whites in the bottom and top deciles). Comparing Blacks and whites in the same skill decile, we see patterns that are quite similar to the unconditional comparisons in Table 4. In older industrial CZs the white commute distance distribution stochastically dominates the Black distribution in each ability decile, except possibly among the lowest-skill workers. In the Sunbelt CZs, the 25th percentile of Black commute distances is higher than the 25th percentile of white commute distances for all skill deciles, but the Black-white gap is relatively small at the median (particularly for the lowest skill deciles), and is negligible at the 75th percentile.

iv) The elasticity of earnings with respect to commute distance

As a final exercise, we examine the partial correlation between commute distance and earnings, and decompose this overall effect into the correlations of distance with the person and wage premium components of earnings. As discussed in Section II, a relative shortage of nearby good jobs for Black workers would be expected to result in a steeper gradient of earnings with respect to commute distance for Black workers, driven by the establishment pay premium component.

Table 5 presents estimates of simple models in which we regress log earnings on the log of individual commute distance and CZ dummies. The first row shows estimates of the elasticity of earnings with respect to distance for whites: the elasticity is 0.059 in the older industrial cities, 0.024 in the Sunbelt CZs, and 0.031 in the other large CZs. (All these estimated elasticities are precisely estimated with robust standard errors of less than 0.002). The corresponding elasticities for Black workers are 0.056, 0.046, and 0.056. When we decompose earnings into the three terms in equation (1), we see that the differences across CZ groups are largely driven by the relationship between commute distance and person effects. The relationship between distance and the workplace pay premium component of earnings is more similar across the three groups of CZs. And importantly, within each group of CZs, the elasticity of pay premiums with respect to distance is lower for Blacks than whites: 0.016 for Blacks versus 0.030 for whites in the older industrial CZs; 0.24 versus 0.019 in the Sunbelt CZs, and 0.023 versus 0.020 in the other large CZs. This pattern suggest that access to nearby better-paying jobs, conditional on worker skills, is if anything better for Black workers.

Our estimates of the relationship between commute distance and workplace pay premiums have to be interpreted cautiously since they are based on purely observational data. One might hope that whatever biases are present in these estimates are similar for Blacks and whites so that a comparison between the elasticities for the two groups is still informative.[15] Interestingly, a recent study by Agarwal et al (2024) using German data obtains quantitatively similar estimates of the elasticity of pay premiums with respect to commuting *time* using individual tax subsidies as instruments for commuting time. Specifically, they report IV estimates of the effect of commuting time on the estimated pay premium for an individual's workplace in the range of 0.02 to 0.04.

Table 6 presents a parallel set of estimates of the effect of commuting time on earnings derived from our ACS sample. Since the ACS only reports the average (one-way) time taken for commuting, we include a set of dummies for the mode of transit, as well as CZ dummies. The elasticities of earnings with respect to commute time range from 0.07 to 0.10 for whites and 0.06 to 0.07 for Blacks. In three of the four city groups we estimate that the elasticity is *lower* for Black workers than whites, though in the Sunbelt cities the elasticity is slightly higher for Blacks (0.077 versus 0.070).

We cannot decompose earnings in the ACS into person and workplace effects. As an alternative, we estimated a relatively rich cross-sectional wage model (separately by race) that included 295 4-digit industry effects. This allows us to decompose an individual earnings observation into a part attributable to the industry of employment, a part due to other observed covariates, and an unexplained part. We then regressed the industry component on commute time and obtained the set of elasticities shown in the second row of each panel in

[15] This would be true, for instance, if Black and white workers have similar preferences and face similar opportunity sets for jobs with different pay premiums at different distances from home.

Table 6. For both Blacks and whites we estimate that longer commute times are associated with employment in higher-paying industries: the elasticities are in the range of 0.026-0.030 for whites and 0.021 to 0.028 for Blacks – not too different from the elasticities of workplace pay premiums with respect to travel distance we obtained in the LEHD. Again, the elasticity of industry pay premiums tends to be slightly lower for Blacks, suggesting that if anything Black workers have slightly *better access* to higher-paying industries, except in the Sunbelt cities where whites may have slightly better access.

VI. Conclusions

We have used information on where people live and work in the LEHD to shed new light on the spatial mismatch hypothesis. While many studies of SM focus on employment probabilities – particularly for very young or less skilled workers -- we focus on earnings of full-time employed workers. In part this is driven by our data source, which lacks direct information on people who are not working. In part it is also driven by our desire to build on the past two decades of research on pay determination, which has shown the importance of employer-specific wage components for understanding wage inequality and differences in pay between groups. The now-standard AKM framework can potentially inform the analysis of spatial mismatch by focusing on accessibility to employers that pay above-average wage premiums.

Our findings suggest that geographic proximity to jobs – or to "good" jobs as measured by their AKM pay premiums – is not a major source of racial wage gaps in large cities in the U.S. today. At the most basic level, we find no evidence that Black workers are systematically under-represented at workplaces with above-average pay premiums. Paralleling results that have been obtained in many earlier studies, we also find that Black workers live, if anything, <u>closer</u> to places where existing jobs are located. Importantly, this relative proximity for Black workers extends to "good jobs."

Building on insights from earlier studies we also examine the distributions of commute distances for whites and Blacks, and the gradient of earnings with respect to distance for the two race groups. We find that in most cities Black workers commute about the same distance or slightly less than whites, though among Sunbelt cities the comparison is reversed. We also find that earnings are higher for people who commute longer. An advantage of our AKM framework is that it allows us to decompose the overall distance gradient into a part attributable to the worker themselves and a part attributable to the pay premium at their current workplace. Here again we find that in most cities the gradient of pay premiums with respect to commute distance is lower for Black workers.

We close by emphasizing four limitations of our work. First, we have not attempted to address the potential importance of mismatch in earlier decades, when housing discrimination was more widespread and a much higher fraction of the Black urban population lived in central cities. Second, we have not tried to look at the effects of mismatch on employment or part time work. Third, we have focused exclusively on Black and white workers, ignoring Hispanic workers who are a major presence in many large cities. And finally, while our analysis includes both men and women, we have not tried to separately estimate the effects of mismatch for female versus male workers.

References

Abowd, John, Francis Kramarz, and David N. Margolis (1999). "High wage workers and high wage firms." *Econometrica* 67(2): 251-333.

Agrawal, David R., Elke J. Jahn and Eckhard Janeba (2024). "Do commuting subsidies drive workers to better firms?" CESifo Working Paper No. 10981, February 2024.

Andersson, Fredrik, John C. Haltiwanger, Mark J. Kutzbach, Henry O. Pollakowski, and Daniel H. Weinberg (2018). "Job displacement and the duration of joblessness: The role of spatial mismatch." *Review of Economics and Statistics* 100 (2): 202-218.

Ashenfelter, Orley (1972). "Racial discrimination and trade unionism." *Journal of Political Economy* 80(3, part 1): 488-508.

Bertrand, Marianne and Sendhil Mullainathan (2004). "Are Emily and Gregg more employable than Lakisha and Jamal? A field experiment on labor market discrimination." *American Economic Review* 94(4): 991-1013.

Card, David, Ana Rute Cardoso, Jörg Heining and Patrick Kline (2018). "Firms and labor market inequality: Evidence and some theory." *Journal of Labor Economics* 36(S1): 13-70.

Card, David, Jesse Rothstein, and Moises Yi (2024a). "Location, Location, Location." Unpublished working paper. Earlier version available as NBER Working Paper No. 31587, August 2023.

Card, David, Jesse Rothstein, and Moises Yi (2024b). "Industry wage differentials: A firm-based approach." *Journal of Labor Economics*, forthcoming. Available as NBER Working Paper No. 31588, August 2023.

Ellwood, David T. (1986). “The spatial mismatch hypothesis: Are there teenage jobs missing in the ghetto?” In Richard Freeman and Harry Holzer, editors, *The black youth employment crisis*. Chicago: University of Chicago Press, pp. 147-190.

Gerard, François, Lorenzo Lagos, Edson Severnini, and David Card (2021). “Assortative matching or exclusionary hiring? The impact of employment and pay policies on racial wage differences in Brazil.” *American Economic Review* 111(10): 3418-57.

Gobillon, Laurent, Harris Selod and Yves Zenou (2007). “The mechanisms of spatial mismatch.” *Urban Studies* 44 (12): 2401-2427.

Hellerstein, Judith K., David Neumark and Melissa McInerney (2008). “Spatial mismatch or racial mismatch?” *Journal of Urban Economics* 64(2): 464-479.

Holzer, Harry J. (1991). “The spatial mismatch hypothesis: what has the evidence shown?” *Urban Studies* 28(1): 105-122.

Hutchinson, Paul (1974). “The effects of accessibility and segregation on the employment of the urban poor.” In George von Furstenburg, Bennett Harrison, and Ann R. Horowitz, editors. *Patterns of Racial Discrimination* Volume 1 . Lexington, MA: Lexington Books.

Ihlanfeldt, Keith R. and David L. Sjoquist (1998). “The spatial mismatch hypothesis: A review of recent studies and their implications for welfare reform.” *Housing Policy Debate* 9(4): 849-892.

Kain, John F. (1964). “The effect of the ghetto on the distribution and level of nonwhite employment in urban areas.” Proceedings of the Annual Meetings of the American Statistical Association.

Kain, John F. (1968). “Housing segregation, negro employment, and metropolitan decentralization.” *Quarterly Journal of Economics* 82 (2): 175-197.

Kline, Patrick, Evan K. Rose and Christopher R. Walters (2022). “Systematic discrimination among large US Employers.” *Quarterly Journal of Economics* 137(4): 1963-2036.

Knowles, John, Nicola Persico and Petra Todd (2001). “Racial bias in motor vehicle searches: Theory and evidence. *Journal of Political Economy* 109(1): 203-229.

Massey, Douglas S. and Jonathan Tannen (2018). “Suburbanization and segregation in the United States: 1970-2010.” *Ethnic and Racial Studies* 41(9): 1594-1611.

Offner, Paul and Daniel H. Saks (1971). “A note on John Kain's `Housing segregation, Negro employment, and metropolitan decentralization.'” *Quarterly Journal of Economics* 85(1): 147-160.

Raphael, Steven (1998). “The spatial mismatch hypothesis and Black youth joblessness: Evidence from the San Francisco Bay area.” *Journal of Urban Economics* 43: 79-111.

Rosenfeld, Jake and Meredith Kleykamp (2012). “Organized labor and racial wage inequality in the United States.” *American Journal of Sociology* 117(5): 1460-1502.

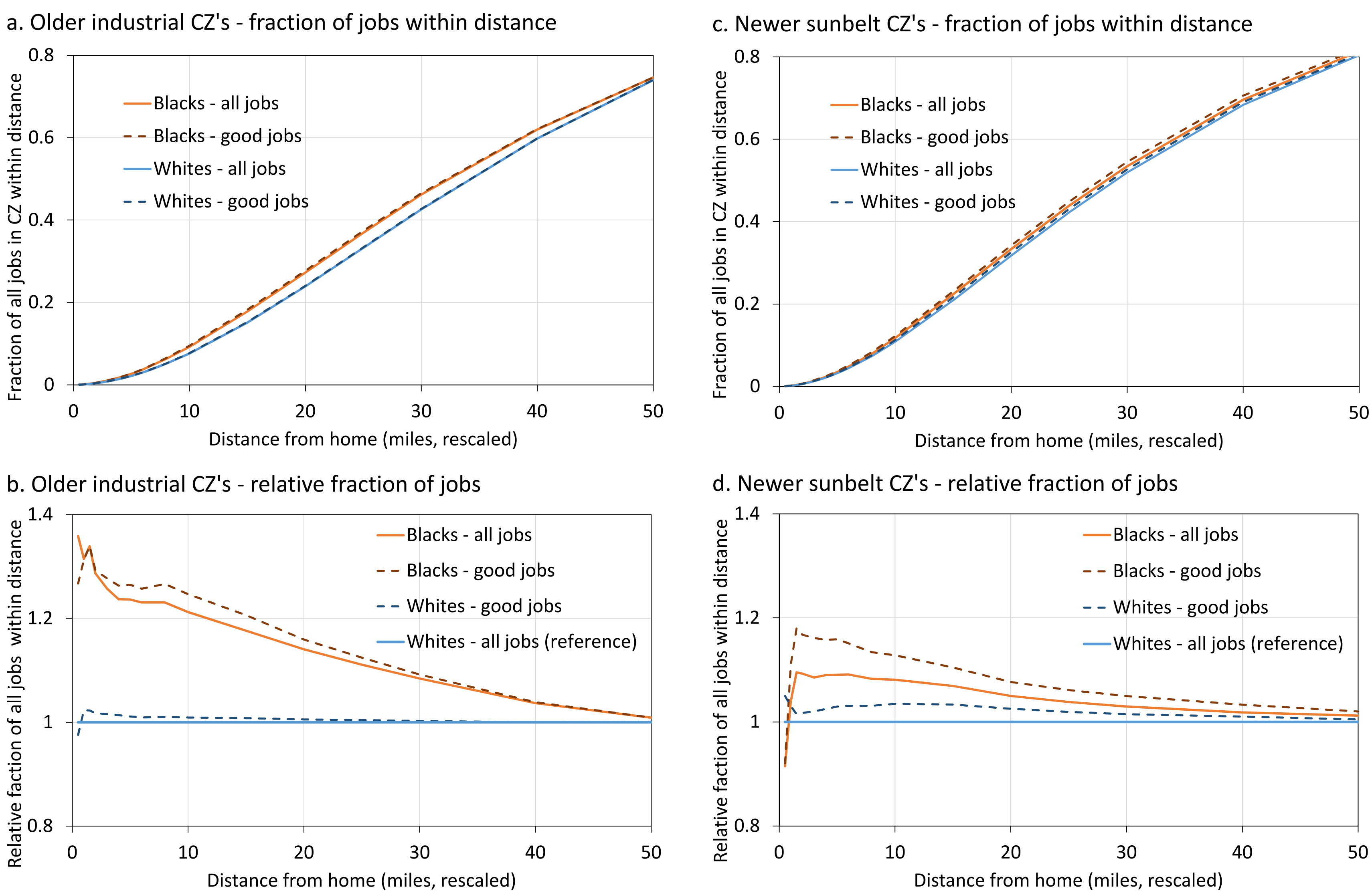
Tolbert, Charles M., and Molly Sizer (1996). “U.S. commuting zones and labor market areas: A 1990 update.” U.S. Department of Agriculture, Economic Research Service, staff paper number AGES-9614.

Vilhuber, Lars (2018). “LEHD Infrastructure S2014 files in the FSRDC.” U.S. Census Bureau, Center for Economic Studies Discussion Paper CES 18-27.

Wang, Liping, Cifang Wu and Songnian Zhao (2022). "A Review of spatial mismatch research: Empirical debate, theoretical evolution and connotation expansion." *Land* 11: 1049.

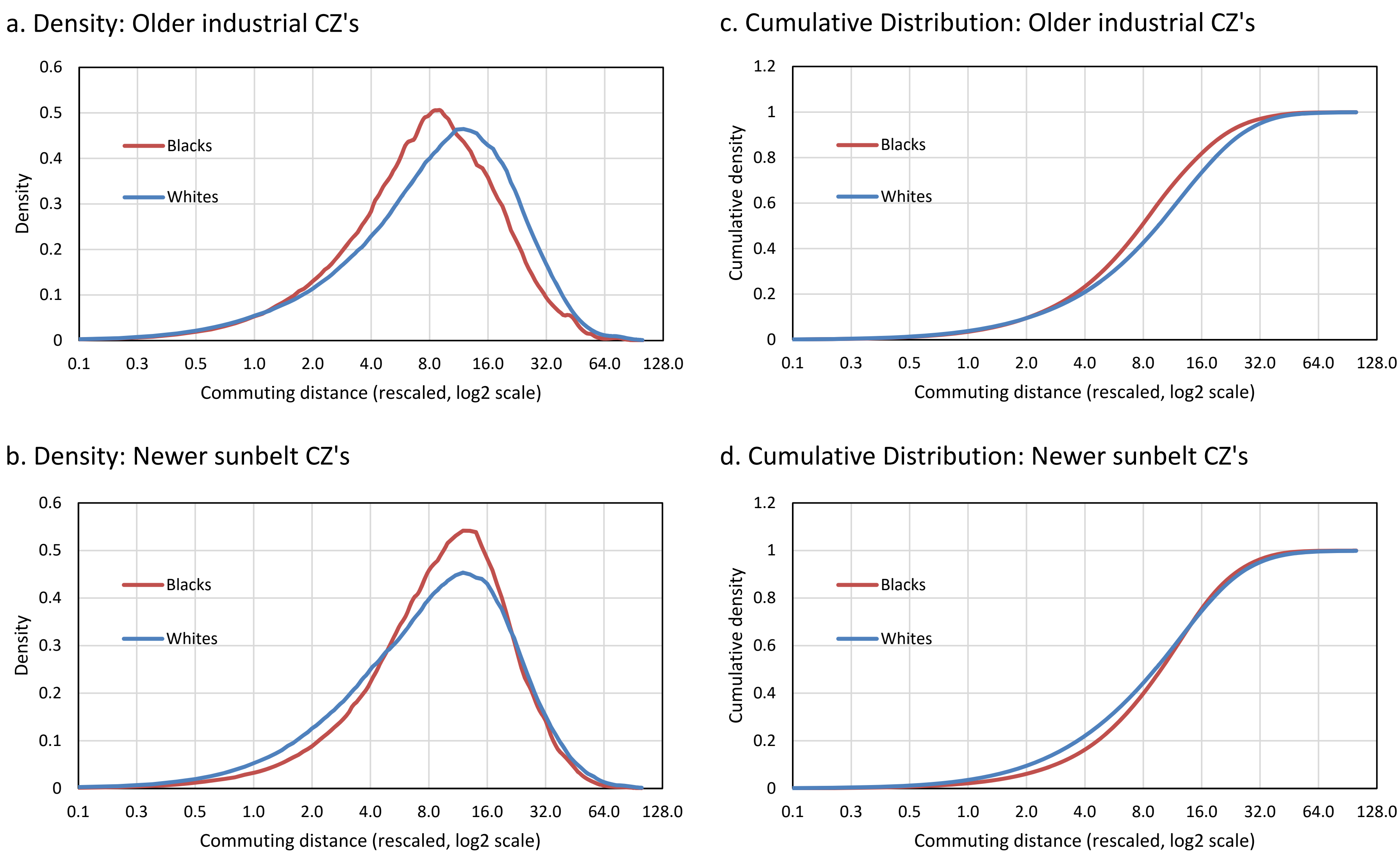
Wilson, Valerie and William Darity Jr. (2022). "Understanding black-white disparities in labor market outcomes requires models that account for persistent discrimination and unequal bargaining power." Unpublished working paper. Washington DC: EPI.

Figure 1: Fraction of all jobs and good jobs within given distance of worker's homes in two groups of CZ's



Note: Based on data on worker's residences and places of work in LEHD, 2010-2018. "Good jobs" are jobs at establishments with pay premiums in top tercile of all premiums in CZ. Distances from home are rescaled in each CZ so that 75th percentile of commute distance in CZ for all workers is 16 miles. See Table 1 for definition of CZ groups.

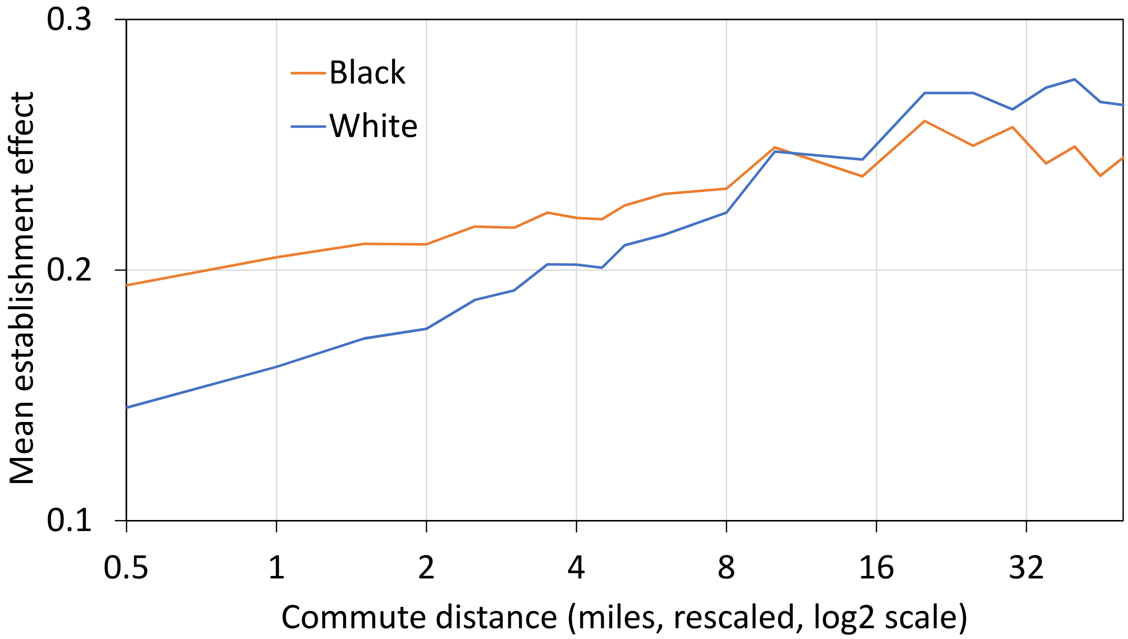
Figure 2: Densities and Cumulative Distribution Functions of Commute Distance



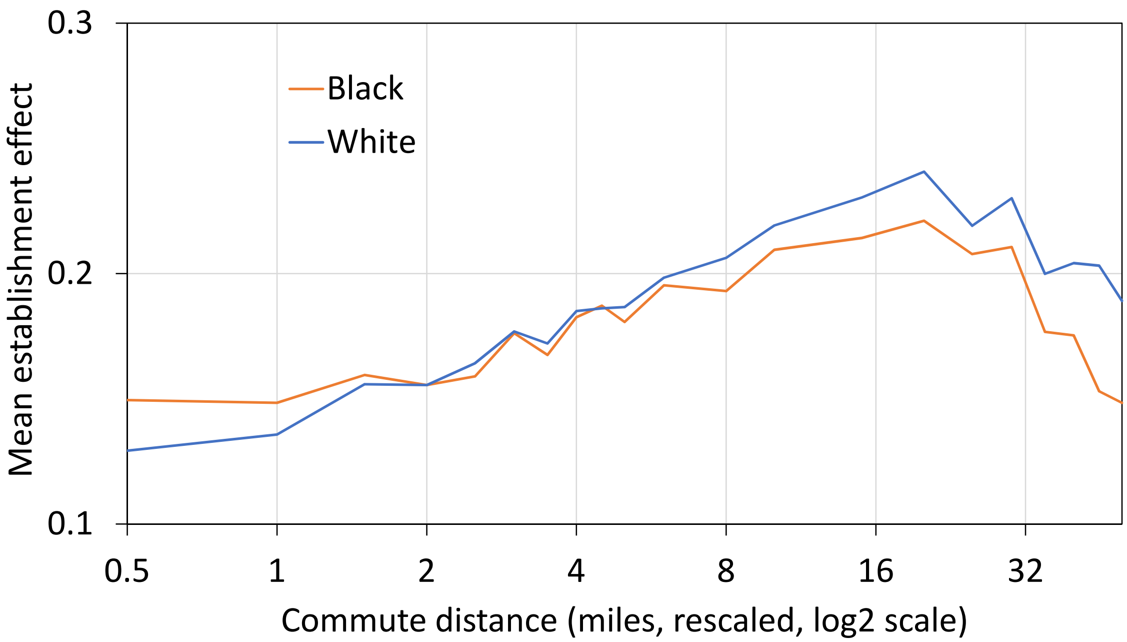
Note: Based on data on worker's residences and places of work in LEHD, 2010-2018. Distances from home are rescaled in each CZ so that 75th percentile of commute distance in CZ for all workers is 16 miles. See Table 1 for definition of CZ groups.

Figure 3: Commute distance and average pay premiums
in two groups of CZ's

a. Older industrial CZ's



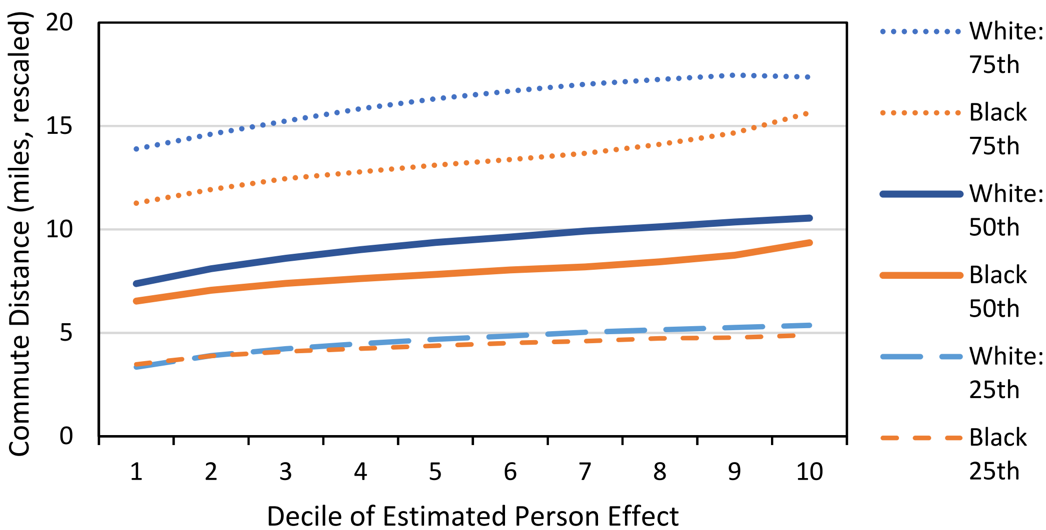
b. Newer sunbelt CZ's



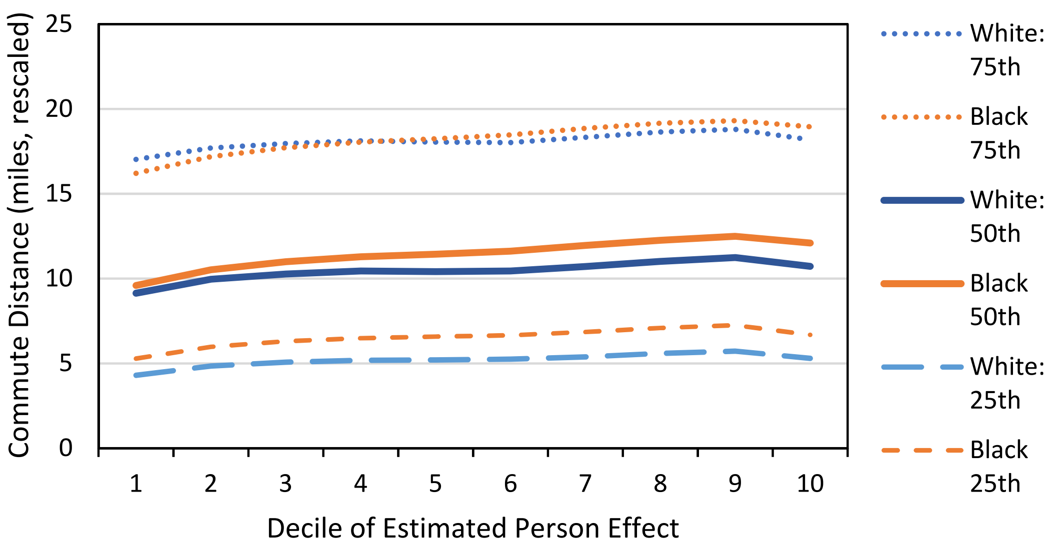
Note: Based on data on worker's residences and places of work in LEHD, 2010-2018. Distances from home are rescaled in each CZ so that 75th percentile of commute distance in CZ for all workers is 16 miles. See Table 1 for definition of CZ groups.

Figure 4: Commute distance percentiles, by decile of estimated person effect and race

a. Older industrial CZ's



b. Newer sunbelt CZ's



Note: Based on data on worker's residences and places of work in LEHD, 2010-2018. Distances from home are rescaled in each CZ so that 75th percentile of commute distance in CZ for all workers is 16 miles. See Table 1 for definition of CZ groups.

Table 1. LEHD summary statistics

	Older industrial CZ's (1)	Newer sunbelt CZ's (2)	All other CZ's among largest 200 (3)
White workers			
Mean quarterly earnings (x4)	77,960	90,400	71,320
Mean log quarterly earnings	9.59	9.68	9.50
Mean commute distance (miles)	12.22	14.37	10.43
No. of person-quarters (millions)	262.1	156.6	931.7
Black workers			
Mean quarterly earnings (x4)	49,440	52,960	36,150
Mean log quarterly earnings	9.25	9.31	9.22
Mean commute distance (miles)	10.47	14.22	9.47
No. of person-quarters (millions)	49.2	46.5	145.3
White-Black gap			
Mean log earnings	0.34	0.38	0.28
Mean commute distance (miles)	1.75	0.15	0.96

Notes: Source is 2010-2018 LEHD. Sample includes only white and Black non-Hispanic people with one employer in the quarter and quarterly earnings above $3,800, and excludes the first and last quarter of each employment spell. Commute distance is based on location of person's place of residence and place of work. Older industrial group in column 1 includes 10 CZ's: Chicago, Philadelphia, Detroit, Pittsburgh, Cleveland, Newark, Buffalo, Baltimore, Minneapolis, and St. Louis. New sunbelt group in column 2 includes 7 CZ's: Los Angeles, Houston, Atlanta, Miami, Dallas, San Diego, Phoenix. Remaining group in column 3 includes 183 CZ's.

Table 2. Summary Statistics for Four Groups of Larger CZ's from American Community Survey

	Top 30 CZ's (1)	Older Industrial CZ's (2)	Newer Sunbelt CZ's (3)	Northeast Corridor CZ's (4)	Remainder of Top 30 CZ's (5)
Pct. working-age pop. in top 30	100.0	28.6	32.4	19.0	20.0
Demographics of working-age population					
White non-Hispanic (%)	52.0	64.1	38.7	51.9	56.4
Black non-Hispanic (%)	13.7	16.6	13.2	16.7	7.5
Hispanic (%)	22.6	11.3	36.3	19.0	20.0
Asian non-Hispanic (%)	11.8	8.1	11.8	12.5	16.3
Mean years of education	13.6	13.8	13.1	14.0	13.7
BA or higher (%)	36.3	37.1	31.2	43.2	36.8
Employed (%)	80.5	81.0	79.0	81.8	81.0
Full-time earner (%)	61.8	63.3	58.4	64.7	62.2
Mean earnings (with 0's)	41,976	42,197	36,913	49,322	42,878
Characteristics of full-time earners					
Black non-Hispanic (%)	12.3	13.6	12.7	15.7	6.6
Female	45.1	46.0	43.7	46.8	44.4
Mean years of education	14.2	14.4	13.8	14.6	14.2
BA or higher (%)	44.4	45.1	39.2	51.7	44.2
Mean earnings	66,064	64,812	60,986	74,618	67,170
Mean hourly wage	31.33	30.62	29.14	35.09	31.99
Use car to commute (%)	86.0	88.4	94.3	67.2	88.8
Mean commute time (minutes)	30.9	29.9	30.0	35.3	29.4
Earnings of white and Black workers					
Mean log earnings - white NH	10.96	10.90	10.97	11.09	10.93
Mean log earnings - Black NH	10.64	10.61	10.61	10.74	10.61
White-Black gap	0.32	0.30	0.36	0.35	0.32

Source: 2010-2018 ACS public use files. Adult population includes people age 22-62 with age> education+6. Full time earners have annual earnings above $15,200. Older industrial CZs and new sunbelt CZs are the same as in Table 1. Northeast Corridor CZs include New York, Washington DC, Boston, and Hartford. Remaining CZ's are San Francisco, Seattle, Denver, Sacramento, San Jose, Portland, Tampa, Orlando and Fort Worth.

Table 3. AKM decomposition of racial earnings gap

	Older industrial CZ's (1)	Newer sunbelt CZ's (2)	All other CZ's among largest 200 (3)
White-Black gap in mean log earnings	0.350	0.365	0.251
Components of AKM decomposition:			
Person effect	0.367	0.380	0.278
(% of overall gap)	(105.1)	(104.1)	(111.0)
Establishment effect	-0.003	0.004	-0.013
(% of overall gap)	(-0.8)	(1.2)	(-5.1)
Covariates (Xβ)	-0.015	-0.019	-0.015
(% of overall gap)	(-4.3)	(-5.3)	(-5.9)
Decomposition of establishment effect:			
Between-industry	0.016	0.011	0.001
Within-industry ("hierarchy")	-0.019	-0.007	-0.014

Notes: Source is 2010-2018 LEHD. Sample includes only white and Black non-Hispanic people with quarterly earnings above $3,800, and excludes the first and last quarter of each employment spell. Covariates in AKM model are cubic in age and dummies for calendar quarter. All means for CZ groups are unweighted averages of means for CZ's in group.

Table 4. Quantiles of commute distance / commute time by CZ group and race

	Older Industrial CZ's		Newer Sunbelt CZ's		Northeast Corridor CZ's	Remainder of Top 30 CZ's
	Miles (LEHD)	Minutes (ACS)	Miles (LEHD)	Minutes (ACS)	Minutes (ACS)	Minutes (ACS)
	(1)	(2)	(3)	(4)	(5)	(6)
Percentiles for White Non-Hispanics						
10	2.1	10	2.1	10	10	10
25	4.7	15	4.5	15	15	15
50	9.5	25	9.2	25	30	25
75	16.6	40	16.2	40	45	40
90	25.3	60	24.9	60	60	55
Percentiles for Black Non-Hispanics						
10	2.1	10	2.8	10	15	10
25	4.2	15	5.5	15	20	15
50	7.8	25	9.9	30	30	25
75	13.4	40	15.9	40	55	40
90	20.9	60	23.3	60	70	60

Notes: Miles represent distances from home to work, from LEHD. Distances are standardized across CZs to set the 75th percentile commute distance in each CZ to 16 miles. Minutes represent commute times, from ACS, unstandardized.

Table 5: Elasticity of Earnings and Earnings Components w.r.t. Commute Distance

	Older industrial CZ's (1)	Newer sunbelt CZ's (2)	All other CZ's among largest 200 (3)
White non-Hispanic			
Log earnings	0.059	0.024	0.031
Person effects	0.027	0.001	0.007
Establishment effects	0.030	0.024	0.023
Covariates and match effects	0.002	-0.001	0.001
Black non-Hispanic			
Log Quarterly Earnings	0.056	0.048	0.056
Person Effects	0.035	0.026	0.032
Establishment effects	0.016	0.019	0.020
Covariates and match effects	0.005	0.003	0.004

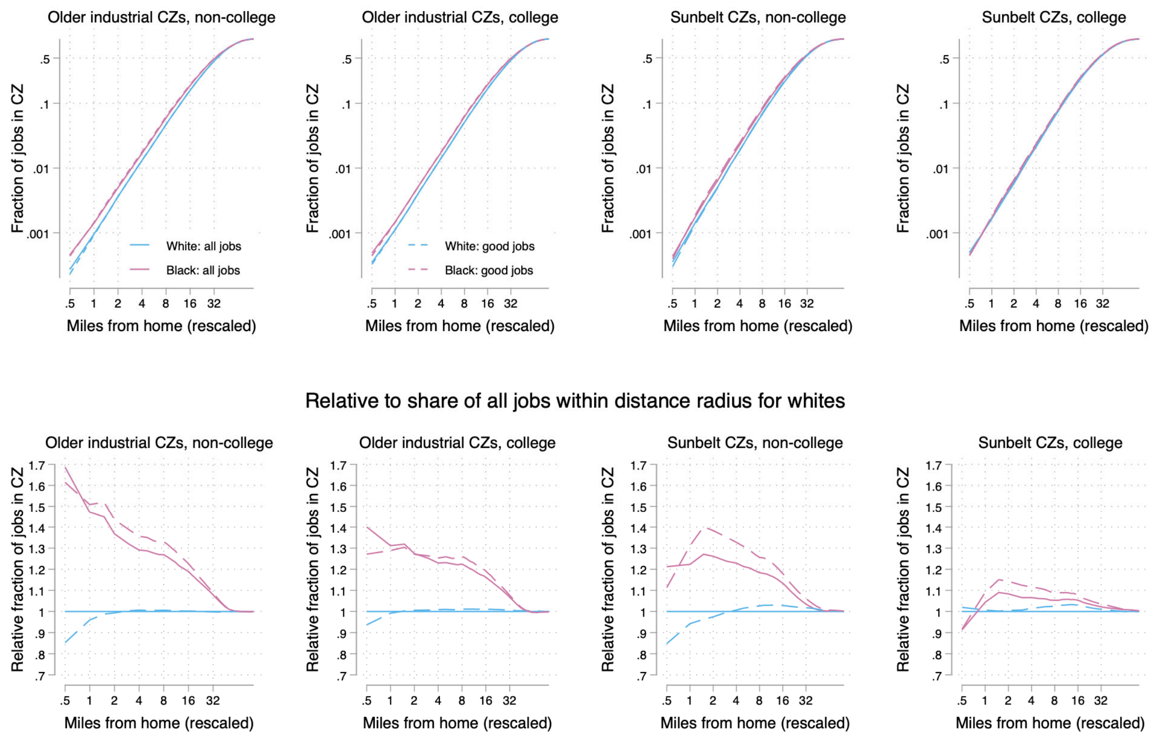
Source: Based on LEHD data -- see notes to Table 1. Coefficient estimates in table are obtained from specifications that regress log of quarterly earnings and AKM components on log of commute distance with dummies for individual CZ's. Standard errors of estimated coefficients are less than 0.002 in all cases

Table 6: Elasticity of Annual Earnings w.r.t. Commute Time

	Top 30 CZ's (1)	Older industrial CZ's (2)	Newer sunbelt CZ's (3)	Northeast corridor CZ's (4)	Remainder of top 30 CZ's (5)
<u>White non-Hispanic</u>					
Log annual earnings	0.090	0.103	0.070	0.099	0.079
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Industry premium	0.028	0.030	0.026	0.028	0.027
(295 industries)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
<u>Black non-Hispanic</u>					
Log annual earnings	0.069	0.063	0.077	0.066	0.069
	(0.002)	(0.003)	(0.003)	(0.003)	(0.005)
Industry premium	0.023	0.021	0.028	0.021	0.023
(295 industries)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)

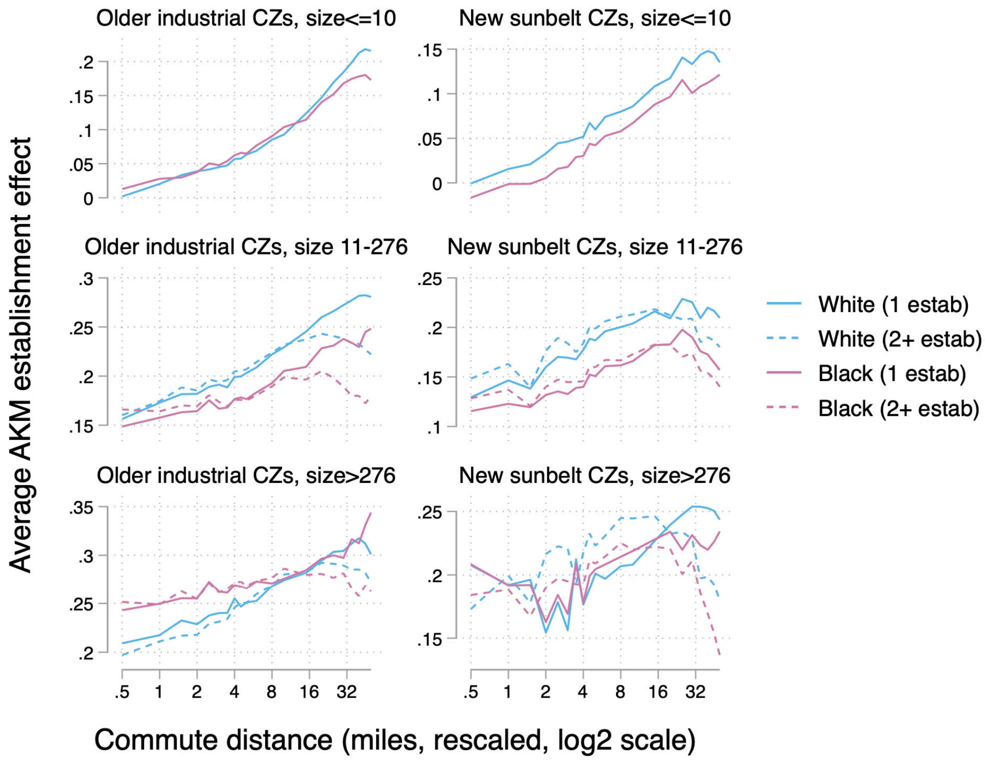
Source: 2010-2018 ACS public use files. Sample includes only people age 22-62 with positive experience (age-education>6) and annual earnings above $15,200. Industry premium represents estimated industry wage effect received by worker, obtained from model fit by gender to all 30 of the largest CZ's, with controls for education, experience, race, immigrant status and CZ effects. Coefficient estimates in table (with robust standard errors) are obtained from specifications that regress log of annual earnings on log of commute time with controls for gender, mode, and CZ.

Appendix Figure 1. Job access in college and non-college labor markets



Notes: Distances for each CZ are rescaled to set the 75th percentile commute distance to 16 miles. “Good jobs” are those at establishments with AKM establishment effects in the top tercile. Bottom panels show fraction of jobs in a CZ relative to fraction of all jobs available within distance radius for whites with the same level of education.

Appendix Figure 2. Estimated pay premiums and commute distance, by firm type, CZ group, and race



Notes: Commute distances are standardized to set the 75th percentile commute distance in each CZ to 16 miles. Rows categorize firms by the maximum number of workers observed at the firm in any quarter; within each panel, series distinguish firms with a single establishment vs. multiple establishments.

Appendix Table 1. Characteristics of population and earners in four groups of Commuting Zones

	Working age population (×1,000)	Working age population (22-62 with positive experience)							Full-time earners only	
		White NH	Black NH	Hispanic	Asian NH	Immigrant	BA or more	Full-time earner(%)	Annual Wage & Sal. Earnings	One-way Comm. Time
Older Industrial Cities:										
Chicago	4,764	53.4	16.4	21.5	8.8	25.6	38.9	63.7	66,155	33.3
Newark	3,429	48.6	13.8	23.4	14.2	36.6	42.3	65.0	75,162	34.2
Philadelphia	3,237	63.2	19.5	9.4	7.9	14.0	36.3	62.4	65,846	30.7
Detroit	2,827	69.2	20.8	3.8	6.2	11.1	30.9	57.2	60,445	27.8
Minneapolis	1,871	77.7	7.7	5.3	9.3	14.2	41.8	70.1	65,041	25.8
Baltimore	1,517	57.8	29.1	5.2	7.9	13.2	39.0	66.7	67,591	31.9
Cleveland	1,377	74.2	18.0	3.6	4.2	6.6	31.5	61.1	56,743	25.2
St Louis	1,336	73.4	19.4	2.6	4.6	6.5	34.9	63.6	59,286	26.2
Pittsburgh	1,334	87.1	7.7	1.5	3.7	4.2	35.4	63.1	57,925	28.0
Buffalo	1,257	79.3	10.9	5.2	4.7	7.4	32.7	62.3	54,583	22.2
Newer Sunbelt Cities:										
Los Angeles	10,272	32.6	6.5	44.6	16.4	40.3	29.1	55.7	61,037	31.0
Houston	3,446	38.0	17.3	35.3	9.4	32.4	30.2	59.8	64,352	30.3
Atlanta	2,822	46.5	34.8	10.5	8.2	20.1	38.3	61.7	62,790	32.3
Dallas	2,562	46.0	16.6	27.7	9.7	28.4	34.8	63.7	63,323	28.8
Miami	2,577	24.2	20.3	51.6	3.9	53.1	29.7	57.1	53,723	30.2
Phoenix	2,421	57.5	5.3	29.0	8.2	20.2	28.4	59.4	57,027	27.0
SanDiego	1,826	47.7	5.0	31.5	15.8	30.6	35.1	59.8	62,988	26.0
Northeast Corridor Cities:										
NYC	6,993	41.6	18.3	25.5	14.6	40.5	39.5	60.8	73,448	39.2
Washington DC	3,296	45.9	25.4	15.1	13.5	30.3	50.4	70.9	78,587	35.4
Boston	2,967	72.2	7.3	10.4	10.1	23.0	46.8	67.0	73,401	32.0
Hartford	1,946	68.3	10.1	14.9	6.7	18.5	39.0	64.9	73,128	27.1
Remaining Cities in Top 30:										
San Francisco	2,981	41.6	7.8	21.6	29.1	36.8	45.5	63.5	81,992	33.1
Seattle	2,616	68.9	5.1	8.7	17.4	20.1	37.8	64.8	67,805	30.2
Denver	1,765	68.4	4.9	20.1	6.7	15.8	42.7	67.0	65,303	27.6
Sacramento	1,711	52.1	6.7	23.6	17.6	25.3	27.3	55.3	60,366	29.4
Tampa	1,571	64.3	11.7	18.3	5.7	17.0	28.8	59.4	54,652	27.8
San Jose	1,488	35.5	2.4	31.1	31.0	46.3	43.4	62.5	83,949	27.6
Ft. Worth	1,311	55.7	13.4	24.2	6.7	19.8	28.0	62.3	57,988	28.6
Orlando	1,326	50.4	15.2	27.8	6.6	21.9	29.8	58.9	51,650	28.5
Portland	1,286	75.5	2.8	10.4	11.3	16.8	37.1	62.3	61,163	26.7

Source: 2010-2018 ACS Public Use Files. Working age population includes people 22-62 with positive experience. Full year earnings have at least $15,200 in annual wage and salary earnings. Size of working age population is based on average weighted count of ACS sample in 2010-2018. Commuting zones are based on 1990 CZ definitions.