Migration Constraints and Disparate Responses to Changing Job Opportunities

Kalee E. Burns and Julie L. Hotchkiss

Working Paper 2019-1
February 2019

Abstract: Using the Current Population Survey between 1996 and 2018, this paper investigates the role constraints to migration might play in explaining racial/ethnic disparities in the labor market. The Delta Index of dissimilarity is used to illustrate a greater distributional mismatch between race/education specific workers and jobs among minorities relative to white non-Hispanics. Regression analysis then shows that this mismatch is consistent with minorities being less responsive to changes in the distribution of job opportunities. However, minorities are more responsive when the growing job opportunities are located in areas with greater same-racial/ethnic representation, suggesting that social constraints might play a role in the observed distributional mismatch. The analysis focuses on 25–54 year old men.

JEL classification: J61, J15, J18

Key words: racial labor market disparities, migration costs, Delta Index, social costs, place-based, people-based, mismatch

https://doi.org/10.29338/wp2019-01
Migration Constraints and Disparate Responses to Changing Job Opportunities

1 Introduction and Background

Long-standing disparities in labor market outcomes by race are well documented. At the opening of a conference at the Board of Governors in 2017 highlighting these disparities and their sources, Governor Brainard affirmed that labor market disparities might have negative, "implications for the growth capacity of the economy" (Brainard 2017). Many contributors to these disparities have been documented, including discrimination, educational opportunities, and social networks. An additional contributor could be differences in migration patterns. A greater ability to chase economic opportunities should improve one's labor market outcomes (for example, see El Badaoui, Strobl, and Walsh 2017; Niebuhr et al. 2009; Davis and Haltiwanger 2014). In fact, the "Great Black Migration" has been credited with significantly improving the economic conditions of blacks from the U.S. South during the early 20th century (Boustan 2015).

Therefore, racial disparities in the labor market may result, and persist, if a disadvantaged group faces more constraints to migrating.

Constraints to migration can take many forms -- from social/cultural constraints to financial constraints. R. Wilson (2018) demonstrates that access to information can be important for informing migration decisions. Cooke (2011) attributes 20 percent of the overall decline in migration rates between 1999 and 2009 to what he calls "secular rootedness," suggesting a social

---

1 For example, see Antecol and Bedard (2004); Biddle and Hamermesh (2013); Bradbury (2000); Cajner et al. (2017); Chetty et al. (2018); Engemann and Wall (2010); Fallick and Krolikowski 2018; Zavodny and Zha (2000); Hotchkiss and Moore (2018).
2 Not all outcomes from the Great Migration were positive; Black et al. (2015) provide evidence that migration by African Americans from rural southern states to northern urban locations resulted in increased mortality.
3 An additional constraint, theorized by Shimer (2007), could include irrational expectations about future local job prospects.
cost to migration. Spilimbergo and Ubeda (2004) also establish family ties as a factor affecting migration in their study for differences in migration rates between Whites and Blacks in the U.S. They find that the reason that Blacks move less than Whites, despite having many factors commonly associated with high migration, is because Blacks have stronger family ties. Additionally, investigating migration patterns in the 1990s, Frey et al. (2005) confirm that cultural constraints to migration are more prevalent among racial minorities. This constraint would be in addition to any other differences across race that have been long known to impact migration decisions, such as access to resources, information, and education (for example, see Greenwood 1975). There may be other indirect contributors to the relationship between migration and labor market outcome gaps. For example, Blair and Chung (2017) provide evidence that occupational licensing reduces racial and gender wage gaps, yet Johnson and Kleiner (2017) find that occupational licensing increases costs of interstate migration. Even though blacks and Hispanics are less likely to be found in occupations that are licensed (Blair and Chung 2017), such institutional constraints may be contributing to labor market disparities in ways that are not obvious.

This paper first presents evidence consistent with racial minorities facing greater migration constraints than white, non-Hispanics by documenting greater differences in the geographic distributions of jobs requiring a certain education level and workers with that education level. If migration was perfectly costless (free of constraints), jobs requiring a certain level of education and workers with that education level would be equally distributed across states (or some other relevant geography). Of course, the degree to which these distributions differ is only suggestive of migration constraints. However, documenting a difference in the distributions is, in a sense, a necessary condition to make the argument that differences in
migration patterns are contributing to observed labor market disparities. This analysis is related to, but differs from the long-standing literature on spatial mismatch, which in its most recent incarnation focuses on job decentralization as the dominant force in declining labor market outcomes among urban minorities (see Kain 1968; Wilson 1990 and Ihlanfeldt and Sjoquist 1998 for a comprehensive survey and Miller 2018 for more recent evidence).

The paper then investigates the degree to which these distributional differences across race/ethnicity, or mismatch, reflect differential responses to changing job opportunities. We find evidence of weaker response among racial/ethnic minorities to changes in job opportunities across geographic locations. The implication is that worse labor market outcomes among minorities may, at least in part, be the result of greater migration constraints. Additional analysis provides evidence that social costs may play a role in constraining ethnic/minority response to changing labor market opportunities elsewhere.

2 Delta Index of Concentration

Indices of spatial concentration, within a much broader class of dissimilarity indices, have been used extensively to measure the degree of and changes in residential segregation (see Massey and Denton 1988; Iceland, Weinberg, and Steinmetz 2002). The "Delta" index of concentration was first proposed by Hoover (1941) and it's use, often referred to as the "Duncan Index," became popular among labor economists to measure occupational segregation (Duncan and Duncan 1955; Watts 1998; Karmel and Maclachlan 2007; Silber 1992). Pertaining to the question in this paper, the Delta Index can tell us how workers (of a certain education level and race) are distributed across the U.S. relative to the distribution of jobs requiring the same education level held by workers of the same race. If the distribution of jobs typically requiring,
say, a college degree better matches the distribution of whites with a college degree than the
distribution of blacks with a college degree, this suggests that whites, compared to blacks, are
gerographically less concentrated than the distribution of occupations. Hence, geography could be
playing a role in observed labor market disparities. In other words, there is more of a geographic
mismatch between blacks with a college degree and college jobs than between whites with a
college degree and college jobs.

The Delta Index \( D^r_e \) that quantifies the difference between the distribution across some
level of geography, \( g \), of workers of racial group, \( r \), and education level, \( e \), and the distribution of
jobs (or some other measure of labor market opportunity) across locations requiring that
education level and held by workers of that racial group, is calculated as follows:

\[
D^r_e = \frac{1}{2} \sum_{g=1}^{G} \left| \frac{n_{eg}^r}{N_e^r} - \frac{j_{eg}^r}{J_e^r} \right|, \tag{1}
\]

where,

- \( n_{eg}^r \): number of people of racial group, \( r \), in geographic location, \( g \), with education, \( e \)
- \( N_e^r \): total number of people in the U.S. of racial group, \( r \), with education, \( e \)
- \( j_{eg}^r \): number of jobs in location, \( g \), requiring education, \( e \), held by workers of racial group, \( r \)
- \( J_e^r \): total number of jobs in the U.S. requiring education, \( e \), held by workers of racial group, \( r \)
- \( G \): total number of geographic locations across which the distributions are compared

The Index falls between zero and one. If workers of a certain racial group, with a certain
education, are distributed across locations identically to the distribution of jobs requiring that
education level then the Delta Index would be equal to zero -- the smaller the Index, the lower
the mismatch between distributions of jobs and people. The Delta Index tells us what share of the
racial group (or jobs) that would need to be moved in order to produce an equal distribution (see
Watts 1998); of course, in the context of migration, it's more natural to think about changing the
share of workers in a location (through migration), rather than changing the share of jobs in that location, that would be needed to equalize the distributions. We will consider three levels of geography -- states, core-based statistical areas (CBSAs), and commuting zones (CZ).\textsuperscript{4} CBSAs are restricted to more urban locations, while CZs are defined for both rural and urban areas, although identification of the CZ of a person living in a sparsely populated county is limited for confidentiality reasons. Figure 1 illustrates the CBSAs and CZs that are represented in our data.

[Figure 1 about here]

Job opportunities are classified not only by education but also by race (adding \( r \) superscripts to the share of jobs at each education level) because of results from Hellerstein, Neumark, and McInerney (2008) who find that an absence of the availability of jobs, generally, is not enough to explain lower employment rates of blacks, but the absence of jobs available to blacks that matters -- accounting for the distribution of jobs only by education level would ignore this point. Additionally, one may argue that a more dynamic measure of job opportunities is preferred. However, we are not using merely the level of employment, but the relative measure of the distribution of employment across locations. In an appendix, we repeat all of the analyses in this paper using an alternative measure of job opportunities -- the distribution of race/education specific year-to-year unemployment-to-employment transitions. One might also argue that a measure of job vacancies would better reflect job opportunities, but because of the importance of identifying race-specific job opportunities (see Hellerstein, Neumark, and

\footnote{Detailed information on CBSAs can be found at https://www.census.gov/topics/housing/housing-patterns/about/core-based-statistical-areas.html. Since metropolitan identifiers change over time, we created a cross-walk to create synthesized CBSA; details can be found in Appendix D. CZ definitions are based on county-to-county commuting patterns; details can be found at https://usa.ipums.org/usa-action/variables/COMZONE#description_section.}
McInerney 2008), it is not possible to use vacancies for this purpose; it is illegal to specify race when advertising a job opening.

3 Data and Measurement Issues

The monthly Current Population Survey (CPS) between January 1996 and November 2018 is used for the analyses in this paper. The starting year of 1996 is chosen since this is when county information necessary for identifying CZs is available. The CPS is administered each month by the U.S. Bureau of Labor Statistics to roughly 60,000 households. This is the nationally representative cross-sectional survey from which we get reports of the unemployment rate and the labor force participation rate, among other monthly labor market statistics. Using monthly CPS data, the number of workers and jobs are summed within each month for each year, using the CPS person weight. Then, this monthly total is averaged across months to get an annual average total.

The analysis in this paper only includes 25-54 year old men. Across several dimensions, labor market racial disparities are often found to be worse among men than among women (for example, see Cajner et al. 2017). The analysis was also performed for 25-54 year-old women, as well, but the patterns across race for women mirror the patterns found among men. For each year observations are classified into three racial groups (White, non-Hispanic; Black, non-Hispanic; non-Hispanic).

There is a growing body of research using online vacancy data, such as Glassdoor or Vault (for example, see Kureková, Beblavý, and Thum-Thysen 2015). Additionally, the Bureau of Labor Statistics makes available measures of job openings (vacancies) in their Job Openings and Labor Turnover Survey (JOLTS). But these data are available only by industry or broad Census region, not both. In addition, occupation is more reflective of educational requirements than industry, which will employ workers of a much broader range of educational attainment. But more importantly neither online vacancy data nor JOLTS identifies race-specific job opportunities.
and Hispanic) and four education levels (less than high school, high school degree, some college, and bachelor degree and above).\textsuperscript{6}

3.1 Education "Required" for Each Occupation

For each year, among those employed (both men and women and all races), excluding the armed forces, the median education level is determined for each detailed occupation.\textsuperscript{7} Table 1 reports the distribution of occupations across median education. The CPS has a different set of detailed occupation codes for 1992-2002, 2003-2010, and 2011 to the present. Not being able to match codes across years is not a concern since the median education for each occupation is constructed within these year groups.\textsuperscript{8}

[Table 1 about here]

Table 1 shows that across all years, most occupations have a median education level of a high school degree only, followed by some college, then college and above. Less than one percent of all occupation codes have a median education level of less than a high school degree.

3.2 Demand for Educational Skills--Labor Market Opportunities by Education and Race

Only 12 occupations prior to 2003, three occupations between 2003 and 2010, and two occupations since 2011 have a median education level of less than high school -- these include occupations such as farm workers, nursery workers, textile cutting machine operators, launderers and ironers, and graders and sorters.

\textsuperscript{6} "Other, non-Hispanic" is excluded from the analysis due to the small number of observations.

\textsuperscript{7} Using the mode education level proved problematic since several occupations had multiple "modes," or, rather, multiple education levels that tied for mode status. CPS person weights are used when obtaining the median.

\textsuperscript{8} Occupation codes for each year can be found at https://cps.ipums.org/cps-action/variables/OCC#codes_section. This is also why the potential of changing skill requirements within occupation over time (see Goyette 2008) is also not a concern.
The number of jobs (held by workers of any gender, within race groups) in each geographic location requiring a certain education level \((j_{egr})\) is simply the sum of people of that race group employed in that location in occupations requiring that education level, using the median education for occupations. Summing across locations yields the total number of jobs in the U.S. held by workers in that race group requiring that education level \((J_{egr})\). Again, all analyses are repeated using an alternate measure of job opportunities -- the distribution of race/education specific year-to-year unemployment-to-employment (U-to-E) transitions.

### 3.2 Supply of Workers by Race and Education

The supply of potential workers in each geographic location for each race and education group is calculated simply as the sum of workers in the location of that race with that education level \((n_{egr})\). The total number of workers (25-54 year old men) in the U.S. of that race with that education level, then, is just the sum across locations \((N_{egr})\). Table 2 reports the distribution of 25-54 year old men across race/ethnicity for each educational group. This is the for the full sample 1996-2018. White, non-Hispanics make up the largest share in all education groups, except those with less than a high school degree. The shares of black, non-Hispanics and Hispanics declines in educational attainment, whereas the share of white, non-Hispanics increases in education.

[Table 2 about here]

### 4 Results - The Delta Index

#### 4.1 Distributions in the Data

Figure 2 illustrates how the distributions of jobs across U.S. states in 2018 for which the median education is a high school degree and the distribution of workers with that level of
education differ from one another. Panel (a) makes this comparison for black, non-Hispanics and Panel (b) makes the comparison for white, non-Hispanics.

[Figure 2 about here]

The distributions will reflect the largest states (California, Texas, and New York, for example) having among the greatest shares of jobs and people of each race at each education level. However, the Delta Index is able to quantify the subtleties in relative distributional differences. For example, the share of jobs requiring a high school degree held by blacks in Ohio is less than the share of blacks living in Ohio (highlighted with red circles in panel a). Similarly, the share of jobs requiring a high school degree held by whites in Oregon is greater than the share of whites living in Oregon (highlighted with red circles in panel b).

4.2 Delta Index and Migration - Exploring some Stylized Facts

We claim that the Delta Index will potentially reflect differential migration constraints. Specifically, in the presence of migration constraints, all else equal, we should see a greater mismatch between workers and job opportunities. A greater mismatch in the distributions of workers and jobs will produce a higher Delta Index. Therefore, we interpret a higher Delta Index as evidence consistent with lower migration. Of course, all else is not equal, and there are potential sources of worker/job mismatch other than migration constraints. Mismatch across locations might arise if there are differences in substitution elasticities between education groups (Ciccone and Peri 2005), if there are significant differences in amenities (Chen and Rosenthal 2008), or differences in location-specific incentives designed to retain residents with higher levels of education (Clotfelter 1976; Strathman 1994). The regression analysis below will

---

9 Similar distributional comparisons can be made for CBSAs and CZs, but differences are much easier to see visually across states.
address this "all-else-equal" concern, but in this section, we investigate whether the relationship between the Delta Index and migration is at least consistent with some stylized migration facts from the literature.

**4.2.a The Delta Index and the Relationship Between Migration and Education**

A positive relationship between education and migration is well established in the literature (for example, see Molloy, Smith, and Wozniak 2011; Greenwood 1975). The theory behind this observation is that education reduces the informational cost of migrating and moving yields a greater return on general human capital afforded to those with higher education levels. Figure 3 shows that this stylized fact holds for white, non-Hispanics but not for the other racial groups.

For all groups, those will less than a high school degree are most dissimilarly distributed across the locations relative to the jobs requiring that education level (not shown, but available upon request). But, unexpectedly, there is greater mismatch between blacks and Hispanics with a college degree and jobs requiring a college degree, compared with those with high school or some college. This suggests that differences in migration constraints by education level vary by race, as well. Specifically, black and Hispanic workers at higher education levels may face more constraints than those workers at lower education levels. This reflects the importance of considering racial specificity of job opportunities (see Hellerstein, Neumark, and McInerney 2008).

Comparing across geographic locations Figure 3 also illustrates, for all racial/ethnic groups, how the broad geography of states masks mismatch that exists at the more narrow geography of CBSAs and CZs. The share of race/education specific jobs in a state is much too
large a geography to capture the labor market opportunities for any one person or group. Notably, across all races, the Delta Index is higher, and, thus, reflects a greater degree of mismatch when the distribution across a more narrow geography is considered. This makes sense, for example, since there may be exactly the same share of black college graduates in the state of California as the share of jobs held by black college graduates. However, the jobs may be concentrated in Los Angeles, whereas the population may be concentrated in San Francisco. The state level Delta Index does not pick up this mismatch, but the CBSA and CZ do. Additionally, level of geography appears to matter more among racial/ethnic minorities.

4.2.b The Delta Index and Migration Patterns over Time

Since at least the 1980s, overall declines in inter-state migration are well-documented, and many varied explanations have been offered to explain it (for example, see Costa and Kahn 2000; Cooke 2013; Molloy, Smith, and Wozniak 2011; Kaplan and Schulhofer-Wohl 2017). The declining trends in migration have been documented among all racial and education groups. If there is a link between lower migration and greater dissimilarity between the distribution of jobs and workers, we would expect the downward trend in migration rates to manifest itself in rising Delta Indices.

We find the Delta Index to be unambiguously rising over the time period across states only for white high school graduates. This suggests that even if the Delta Index is found to reflect a greater mismatch between people and job opportunities among racial/ethnic minorities than among whites, linking this result to lower migration rates (i.e., migration constraints) may be more tenuous than we thought.\textsuperscript{10} However, if falling migration rates are more related to the

\textsuperscript{10} Also see Hall and Schulhofer-Wohl (2018) who document a reduction in job matching efficiency between 2001 and 2013.
aging population or declines in geographic specificity of occupations (Kaplan and Schulhofer-Wohl 2017), then rising migration may not result in greater mismatch between job opportunities and the working age population.

4.3 Delta Index and Evidence of Greater Mismatch among Racial/Ethnic Minorities

In order to emphasize differences in geographic mismatch across race/ethnicity, Figure 4 re-arranges the Delta Indices presented in Figure 3. Comparing distributions across states, it appears that only white, NHs with a college degree have a distributional advantage in job opportunities. However, at the CBSA or CZ level, that distributional advantage shows up at all education levels, and is most dramatic at the college or above level of education. If differences in the Delta Index across race/ethnicity reflect differences in migration constraints, then Figure 3 suggests that racial/ethnic minorities, at all education levels, indeed face greater migration constraints than white, NHs.

[Figure 4 about here]

Figure A1 in Appendix A reproduces Figure 4 using a different measure of job opportunities reflected in the distribution of year-to-year U-to-E transitions. The conclusion that the Delta Index provides evidence of greater mismatch between job opportunity and population among racial/ethnic minorities than among whites is even stronger when this measure of job opportunity is used.

5 Multivariate Regression

So far, this analysis has interpreted a higher Delta Index (i.e., greater mismatch between people and job opportunities) for a particular racial group as evidence for the presence of greater constraints on migration. This has not been a causal analysis, but merely the presentation of
evidence consistent with that conclusion. All else equal, fewer migration constraints implies a
greater migration response to growing job opportunities. In an effort to get at a more causal
conclusion to the question of whether white, non-Hispanics are more responsive to job
opportunities than racial and ethnic minorities, we estimate a regression model where the change
in local labor market opportunities (total share of race/education specific employment) enters as
a determinant for the change in the share of people (of the race/education group) in that location.

5.1 Model Specification

Again, restricting the analysis to men, ages 25-54, we make use of the same CPS data
between 1996 and 2018 to estimate the relationship between changes in labor market
opportunities and responsiveness. The primary analysis defines job opportunities as the share of
total race/education specific jobs in a certain geographic location. All analyses are repeated using
the year-to-year U-to-E race/education transition shares; this specification and results are
reported in Appendix A.

The baseline model is specified as follows:

\[
\Delta \left( \frac{n^r_e}{N^r_e} \right)_{g,t} = \alpha + \rho \Delta \left( \frac{j^r_e}{j^r_i} \right)_{g,t} + \sum_{j=1}^{2} \left\{ \beta^1_j RACE^j_{g,t} + \beta^2_j RACE^j_{g,t} \ast \Delta \left( \frac{j^r_e}{j^r_i} \right)_{g,t} \right\}
+ \sum_{k=1}^{2} \left\{ \phi^1_k EDUC^k_{g,t} \ast \Delta \left( \frac{j^r_e}{j^r_i} \right)_{g,t} \right\}
+ \sum_{j=1}^{2} \sum_{k=1}^{2} \left\{ \lambda^1_{jk} RACE^j_{g,t} EDUC^k_{g,t} + \lambda^2_{jk} RACE^j_{g,t} EDUC^k_{g,t} \ast \Delta \left( \frac{j^r_e}{j^r_i} \right)_{g,t} \right\}
+ \theta' X_{g,t-1} + \tau_t + \sigma_g + \epsilon_{g,e,r,t},
\]

where,

\[
\Delta \left( \frac{n^r_e}{N^r_e} \right)_{g,t} = \text{change in the share of all people of racial group, } r, \text{ with education, } e, \text{ in geography,}
\]

g, from t-1 to t;
\[ \Delta \left( \frac{r_e}{r_e} \right)_{g,t} = \text{change in the share of total employment of racial group, } r, \text{ with education level, } e, \text{ in geography, } g, \text{ from } t-1 \text{ to } t; \]

\( RACE_{g,t} = \text{set of 0,1 regressors indicating black, non-Hispanic or Hispanic race/ethnicity}; \)

\( EDUC_{g,t} = \text{set of 0,1 regressors indicating some college or college plus education groups}; \)

\( X_{g,t-1} = \text{geography specific additional regressors at time } t-1, \text{ including the unemployment rate and industry shares, which are expected to capture both baseline job opportunities and unemployment risk in the geographic location (e.g., see Devaraj et al. 2017)}; \)

\( \tau_t \text{ and } \sigma_g \text{ are time and geography fixed effects, respectively; and } \varepsilon_{g,e,r,t} \text{ are robust standard errors, clustered at the geography level (each CBSA and CZ are observed multiple times across years).}^{11} \)

Both random and fixed effects (treating the geography/race/education as the unit of observation) versions of the model are estimated and presented. A Hausman test indicates there is no statistical difference between the random and fixed effects model parameter estimates. Since the random effects estimator is more efficient (Clarke et al. 2010), those are the results we will focus on in the discussion.

This analysis is restricted to location/race/education observations that have non-zero values for current and lagged values of population share and job opportunity share. The reason for this restriction is that we do not know whether a zero share of a race/education specific occupation in a specific location is a true zero, or whether that location was simply not sampled

---

11 Clustering is not done at the state level since CZs and CBSAs cross state boundaries. And, using Census regions or divisions would provide too few clusters (see Cameron, Gelbach, and Miller 2008).
that year. Over 80 percent of CBSA and CZ race/education specific population and occupation shares in the data are non-zero.

The geography fixed effect is expected to control for time-invariant location effects, such as differences in amenities; and the time fixed effect will control for periods of common economic experiences, such as recessions. The analysis is performed for both CBSA and CZ geography levels, and the unit of observation is race/education/geography/year. The analysis excludes less than high school. Note that the definition of commuting zones depends on knowing a person's county, which is often suppressed in public data (due to small county size); so there will be more observations in the CBSA analysis than in the CZ analysis. Since changes in job opportunities may be endogenous to population changes, we will investigate instrumenting \( \Delta \left( \frac{J_t^g}{J^g} \right) \) with a Bartik shift-share (Bartik 1991) in future analyses.\(^{12}\)

This analysis is not unlike that undertaken by Amior and Manning (2018), who find evidence of significant migratory response to labor market opportunity, but that push-migration (from declining economic opportunity) is much weaker than pull-migration. This means that populations never fully adjust to changing employment opportunities and labor market disequilibrium persists across locations. Their analysis, however, does not separate migration responses by education or race.

Table 3 reports the marginal effects of the change in location job shares on change in population shares.\(^{13}\) A positive marginal effect indicates that in an area with a higher change in

\(^{12}\) As a first step to address potential endogeneity through timing, we re-estimate equation (2) with lagged change in job opportunities. Most of the patterns of results are consistent with those presented here, but the marginal effects are estimated less precisely.

\(^{13}\) Full estimation results are contained in Appendix B (Table B1); none of the industry fixed effects or lagged unemployment rate contribute additional explanatory power beyond the locational change in job opportunities.
education/race job share over the previous year, the share of people in that education/race group also increased in that area -- suggestive of a positive net migration response to improved job opportunities in the area. The question is whether racial/ethnic minorities exhibit any different level of response than white, non-Hispanics.

[Table 3 about here]

Table 3 indicates that there is a significant positive correlation between job share changes and population changes at both the CBSA and CZ levels of geography. For the random effects specification, the marginal effect for whites across CZs (0.7339), for example, suggests that a standard deviation increase in the change in a location's job share (about 0.13 percentage point) results in a 0.095 percentage point change in the population share of whites (0.13x0.7339).14

The main result from Table 3 is that the marginal effect of education/race job share changes on education/race specific population shares is smaller for blacks and Hispanics than for whites, generally, and for all education levels, except Hispanics with a high school education level. The marginal estimates for blacks and Hispanics that are statistically different from those for whites are in bold.

The results suggest that, for the most part, racial/ethnic minorities are less responsive to changes in job opportunities. Additionally, responsiveness of white and black non-Hispanics increases in education, whereas less educated Hispanics are more responsive than more educated Hispanics. The marginal effects for responsiveness to changes in U-to-E transition shares are found in Appendix A (Table A1). While much less precise, the general conclusions hold for either the CBSA or CZ analyses.

14 Comparing to the impact of changes in the CZ share of 12 month U-to-E transitions (see Table A1 in Appendix A), a one standard deviation in U-to-E transitions among whites (0.50 percentage points) results in a 0.016 percentage point change in the population share of whites.
6 Importance of Social Costs

The appropriate policy aimed at correcting the mismatch between jobs and workers depends on the reason why racial/ethnic minorities are less responsive to changes in labor market opportunities. If social costs are keeping racial and ethnic minorities from migrating to better opportunities, then a policy aimed at moving people to jobs is likely to be less effective than a policy of moving jobs to people. A graphical analysis of Facebook connections (Bailey et al. 2018, also see Badger and Bui 2018) illustrates how powerful connections from historical events, like the Great Migration in the early 20th century, can dictate geographic connectedness today.\(^\text{15}\) Also, Ananat, Shihe, and Ross (2018) find that as the share of a worker's race in a local area increases, the employment density wage premium for that worker increases, providing yet another reason why we might expect minorities to respond more to employment opportunities in areas with higher own-racial shares. This section explores the role that social costs might be playing in weaker responsiveness of minority workers to changes in job opportunities.

The analysis modifies equation (2) by including the share of the population in the location (either CBSA or CZ) that is black or Hispanic by itself, interacted with education, and also interacted with changing job opportunities in the location. The presence of a greater share of one's own race/ethnic demographic might be expected to reduce the social cost of moving to that location. Marginal effects by ethnic/racial percentiles are presented in Table 4.\(^\text{16}\) The marginal

\(^{15}\) A future analysis will make use of the same data to explore responsive to job opportunities in location in which individuals have greater social media connections. An analysis that restricts locations to Great Migration states does not find results statistically different from the full sample (results available upon request).

\(^{16}\) The estimating equation and full set of parameter estimates are found in Appendix C. Again, we focus on the random effects model results here because a Hausman test fails to reject the null hypothesis of equality between fixed and random effects parameter estimates.
effects at the 75th percentile that are statically significantly different from the marginal effects at the 25th percentile are bolded.

[Table 4 about here]

For Hispanics, the marginal effects at the 75th percentile (for high school and college and above) are higher at the 25th rather than the 75th percentile -- opposite of what we might expect if social costs were a barrier to migration. Hispanics with some college education appear to be more motivated to respond to job opportunities in locations with higher shares of Hispanics as the marginal effect at the 75th percentile is larger than (and statistically different from) the marginal effect at the 25th percentile.

For black, non-Hispanics, all the point estimates progress in the way that we would expect if job market opportunities in locations with higher shares of ethnic/racial minorities was more influential in motivating blacks to migrate to take advantage of those opportunities. The significant difference between the 75th and 25th percentile marginal effects across CZs for blacks overall, however, is being driven by the difference among the college educated. This suggests reducing social costs would increase migration responses of minorities to job market opportunities elsewhere, especially if information and resources were available (as they might expected to be for the more educated) to take advantage of them. The stronger migration response among college-educated blacks across CZs (including more rural areas and suburbs) rather than across CBSAs (reflecting more urban locations) is consistent with the finding by Couture and Handbury (2017, 32) that, "Young college-educated Blacks...are over-represented in 2000, but have grown faster [through 2010] in the suburbs, unlike Whites, Asians, Hispanics, and others who grew faster [in this time period] in urban areas" (also see Sisson 2018).
Table A2 in Appendix A contains analogous results using 12 month U-to-E transitions as job opportunities. The pattern of the point estimates is the same for both blacks and Hispanics, but the point estimates are less precise.

7 Conclusions and Policy Considerations

The analysis in this paper finds that black and Hispanic workers, at each education level, are more geographically concentrated than whites, relative to race/education specific job opportunities. This result holds for different measures of job opportunities and across different levels of geography, including states, Core-based Statistical Areas (CBSA), and Commuting Zones (CZ). The differences in concentration are most dramatic when job opportunity is measured by the distribution of transitions from unemployment to employment (U-to-E) across different geographies.

A regression analysis supports the interpretation of these results as differences in migration responses by education and race to changing job opportunities. At both the CBSA and CZ geographic level, the relationship between the change in education/race specific job opportunity in a location and the change in education/race specific population is significantly larger and more statistically significant among white, non-Hispanics than it is for ethnic/racial minorities. Additional analysis provides evidence that social costs may play a role in constraining ethnic/minority response to changing labor market opportunities elsewhere.

Finding evidence of greater job opportunity/population mismatch and weaker response to changing job opportunities among racial/ethnic minorities is not sufficient to conclude that blacks and Hispanics would be better off if they were spread more thinly across the U.S. to better match the distribution of jobs matching their education. Some have found that racial and ethnic
minorities experience significant gains from social and cultural networks that are accessible when living in close proximity with one another (e.g., Montgomery 1991; Edin, Fredriksson, and Åslund 2003; Elliott 2005). This would suggest that efforts directed toward decreasing disparate labor market outcomes should focus on adjusting the human capital of minorities (e.g., by improving educational opportunities) to better match the occupational demands of the area, or by improving economic opportunities that better match the educational attainment of the population, rather than necessarily promoting migration.

On the other hand, Xie and Gough (2011) don't find any evidence of benefits to immigrants working in "ethnic enclaves" relative to immigrants working outside of the enclave. In addition, Dickerson (2007) finds that employment outcomes are worse for blacks in segregated cities, suggesting that geographic concentration may indeed be harmful for economic outcomes of minorities, and that easing migration might prove useful for improving labor market disparities.

Picard and Zenou (2018) provide a theoretical model showing how minority workers, faced with a mismatch of location and jobs, could benefit from a variety of policy approaches. Place-based policies, such as neighborhood regeneration (which provides incentives for majority workers to move there providing improved networking contacts) and establishment of enterprise zones (attracting firms providing additional employment opportunities) are ways in which specific geographic locales can attract both residents and firms. Contrastingly, people-based policies, such as the Moving to Opportunity programs, provide housing subsidies in order to improve outcomes by moving people closer to jobs. Both of these first two types of policies would improve the measured locational mismatch between minorities and jobs. However, incentivizing people to move is a tall order (for example see Harrison and Raice 2018). Indirect
policies, such as improving public transportation or access to information (see Waldrip et al. 2015; R. Wilson 2018) will also improve employment outcomes among minorities, but may not change the locational mismatch between minorities and jobs. This potential conflict in policies focused on *either* people *or* place is long-standing in the urban literature, described in a phrase coined by Winnick (1966)-- 'Place Prosperity vs. People Prosperity' (also see Bolton 1992; Partridge and Rickman 2007).

Hellerstein, Neumark, and McInerney (2008) find that an absence of the availability of jobs, generally, is not enough to explain lower employment rates of blacks, but it's the absence of jobs *available to blacks* that matters. This suggests that while Marinescu and Rathelot (2018) find that aggregate geographical mismatch between jobs and people may not be very important in the overall unemployment rate, education/race specific mismatch may play a greater role in determining labor market outcome. In other words, combating discrimination and negative neighborhood effects (Cain and Finnie 1990) may be even more important than solving the distribution problem.
References


Devaraj, Srikant, Michael J. Hicks, Emily J. Wornell, and Dagney Faulk. 2017. “How Vulnerable Are American Communities to Automation, Trade, & Urbanization?” Muncie, IN: Rural Policy Research Institute, Center for State Policy, Ball State University.


Figure 1 CBSAs and CZs represented in the data.
Figure 2 Maps of the distribution of high-school-only jobs by race across states.

Panel (a): Distribution of high school jobs held by black, non-Hispanics vs. the distribution of black, non-Hispanics with a high school degree.

Notes: The "required" education for a job is determined by the median education of people employed in that occupation. Data reflects the distribution of jobs in 2018. Maps created using the Stata program _maptile_.

Panel (b): Distribution of high school jobs held by white, non-Hispanics vs. the distribution of white, non-Hispanics with a high school degree.
Figure 3 Delta Index by race across education and time, total number of jobs across U.S. states, CBSAs, and CZs.

(a) Across States  
(b) Across CBSAs  
(c) Across CZs

Source: Authors calculations using the CPS monthly data January 1996 through November 2018. Education level less-than-high-school excluded for illustrative purposes (available upon request).
Figure 4 Delta Index by education across race and time, total number of jobs across U.S. states, CBSAs, and CZs.

(a) Across States  
(b) Across CBSAs 
(c) Across CZs

Source: Authors calculations using the CPS monthly data January 1996 through November 2018. Education level less-than-high-school excluded for illustrative purposes (available upon request).
**Table 1** Distribution of occupations across median education of those employed in the occupation 1996-2018.

<table>
<thead>
<tr>
<th>Median Education in Occupation</th>
<th>Percent of Occupation codes across years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than high school</td>
<td>0.96%</td>
</tr>
<tr>
<td>High school degree only</td>
<td>36.17%</td>
</tr>
<tr>
<td>Some college</td>
<td>35.58%</td>
</tr>
<tr>
<td>College degree and above</td>
<td>27.28%</td>
</tr>
</tbody>
</table>

Notes: Authors calculations using the CPS.

**Table 2** Distribution of 25-54 year old men across race/ethnicity by educational attainment, 1996-2018.

<table>
<thead>
<tr>
<th>Percent of Education Category</th>
<th>White, NH</th>
<th>Black, NH</th>
<th>Hispanic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than HS</td>
<td>38</td>
<td>11</td>
<td>51</td>
</tr>
<tr>
<td>HS degree</td>
<td>68</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>Some College</td>
<td>74</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>College degree or more</td>
<td>84</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

Notes: Authors calculations using the CPS person weight. Row totals may not sum to 100 due to rounding.
Table 3 Marginal effect of a change in location job share $\Delta \left( \frac{j}{J} \right)$ on share of population $\Delta \left( \frac{n}{N} \right)$ in the location, by race/ethnicity and educational attainment.

<table>
<thead>
<tr>
<th>Geographic area:</th>
<th>Fixed Effects</th>
<th>Random Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CBSA</td>
<td>CZone</td>
</tr>
<tr>
<td>White, NH</td>
<td>0.5645***</td>
<td>0.7297***</td>
</tr>
<tr>
<td></td>
<td>[0.0243]</td>
<td>[0.0289]</td>
</tr>
<tr>
<td>High School</td>
<td>0.5638***</td>
<td>0.6229***</td>
</tr>
<tr>
<td></td>
<td>[0.0279]</td>
<td>[0.0361]</td>
</tr>
<tr>
<td>Some College</td>
<td>0.5171***</td>
<td>0.7317***</td>
</tr>
<tr>
<td></td>
<td>[0.0392]</td>
<td>[0.0522]</td>
</tr>
<tr>
<td>College and Above</td>
<td>0.6174***</td>
<td>0.8509***</td>
</tr>
<tr>
<td></td>
<td>[0.0491]</td>
<td>[0.0411]</td>
</tr>
<tr>
<td>Black, NH</td>
<td>0.4905***</td>
<td>0.6034***</td>
</tr>
<tr>
<td></td>
<td>[0.0262]</td>
<td>[0.0449]</td>
</tr>
<tr>
<td>High School</td>
<td>0.5330***</td>
<td>0.5789***</td>
</tr>
<tr>
<td></td>
<td>[0.0247]</td>
<td>[0.0425]</td>
</tr>
<tr>
<td>Some College</td>
<td>0.4010***</td>
<td>0.5963***</td>
</tr>
<tr>
<td></td>
<td>[0.0676]</td>
<td>[0.0564]</td>
</tr>
<tr>
<td>College and Above</td>
<td>0.5385***</td>
<td>0.6394***</td>
</tr>
<tr>
<td></td>
<td>[0.0582]</td>
<td>[0.0670]</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.4805***</td>
<td>0.6503***</td>
</tr>
<tr>
<td></td>
<td>[0.0266]</td>
<td>[0.0459]</td>
</tr>
<tr>
<td>High School</td>
<td>0.5975***</td>
<td>0.7684***</td>
</tr>
<tr>
<td></td>
<td>[0.0416]</td>
<td>[0.0349]</td>
</tr>
<tr>
<td>Some College</td>
<td>0.4334***</td>
<td>0.6373***</td>
</tr>
<tr>
<td></td>
<td>[0.0779]</td>
<td>[0.1243]</td>
</tr>
<tr>
<td>College and Above</td>
<td>0.3944***</td>
<td>0.5279***</td>
</tr>
<tr>
<td></td>
<td>[0.0518]</td>
<td>[0.0731]</td>
</tr>
<tr>
<td>Observations</td>
<td>36,573</td>
<td>20,303</td>
</tr>
<tr>
<td>Number of FE</td>
<td>2,557</td>
<td>1,560</td>
</tr>
<tr>
<td>Clusters</td>
<td>307</td>
<td>188</td>
</tr>
</tbody>
</table>

Note: Robust standard errors are clustered at the level of geography. *, **, *** => statistical significance at the 90, 95, and 99 percent level. Additional controls include the area's baseline unemployment rate and industry shares. Sample includes 25-54 year-old men with at least a high school degree and 1996-2018 years of data. Full estimation results are included in Appendix B. Bolded numbers reflect statistically significant difference between racial/ethnical minority and white, NH marginal effects.
Table 4 Marginal effect of a change in the job share among black, NH and Hispanics on share of population in CBSA, by own race/ethnicity, at different points in the distribution of CBSA race/ethnicity population share; random effects specification.

<table>
<thead>
<tr>
<th>Core Statistical Business Area Analysis</th>
<th>Commuting Zone Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own race population share in CBSA, percentile</td>
<td>Own race population share in CZ, percentile</td>
</tr>
<tr>
<td>25th Percentile</td>
<td>50th Percentile</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td><strong>M.E. for Blacks, NH</strong></td>
<td></td>
</tr>
<tr>
<td>0.4407***</td>
<td>0.4533***</td>
</tr>
<tr>
<td>[0.0289]</td>
<td>[0.0250]</td>
</tr>
<tr>
<td>High School</td>
<td></td>
</tr>
<tr>
<td>0.5210***</td>
<td>0.5233***</td>
</tr>
<tr>
<td>[0.0401]</td>
<td>[0.0338]</td>
</tr>
<tr>
<td>Some College</td>
<td></td>
</tr>
<tr>
<td>0.3313***</td>
<td>0.3468***</td>
</tr>
<tr>
<td>[0.0871]</td>
<td>[0.0748]</td>
</tr>
<tr>
<td>College or Above</td>
<td></td>
</tr>
<tr>
<td>0.4642***</td>
<td>0.4856***</td>
</tr>
<tr>
<td>[0.0700]</td>
<td>[0.0590]</td>
</tr>
<tr>
<td><strong>M.E. for Hispanics</strong></td>
<td></td>
</tr>
<tr>
<td>0.4732***</td>
<td>0.4741***</td>
</tr>
<tr>
<td>[0.0268]</td>
<td>[0.0261]</td>
</tr>
<tr>
<td>High School</td>
<td></td>
</tr>
<tr>
<td>0.6738***</td>
<td>0.6657***</td>
</tr>
<tr>
<td>[0.0557]</td>
<td>[0.0533]</td>
</tr>
<tr>
<td>Some College</td>
<td></td>
</tr>
<tr>
<td>0.2635***</td>
<td>0.2822***</td>
</tr>
<tr>
<td>[0.0760]</td>
<td>[0.0736]</td>
</tr>
<tr>
<td>College or Above</td>
<td></td>
</tr>
<tr>
<td>0.4665***</td>
<td>0.4589***</td>
</tr>
<tr>
<td>[0.0578]</td>
<td>[0.0552]</td>
</tr>
</tbody>
</table>

Note: Robust standard errors are clustered at the geography level. There are 307 CBSAs and 188 CZs. *, **, *** => statistical significance at the 90, 95, and 99 percent level. Regression includes location specific unemployment rate and industry shares, and location/education/race and year fixed effects. Sample includes 25-54 year-old men with at least a high school degree and 1996-2018 years of data. Full estimation results are found in Appendix C. Bolded numbers reflect statistically significant difference between the 75th and 25th percentile marginal effects.
Appendix A: Analyses using the Year-to-year Unemployment to Employment Transitions as Proxy for Job Opportunities

Rather than proxy for job opportunities across geography using the total number of education/race specific total employment or number of jobs, this appendix presents results from the Delta Index and regression analyses using the year-to-year unemployment-to-employment (UR) transitions as the proxy for job opportunities.

A.1 The Delta Index Analysis

It might be argued that the total number of race/education specific jobs in an area does not appropriately account for job opportunities -- that what is needed is a more dynamic measure. While a measure of job vacancies by occupation (for any geography level) is not practical since they would not be available by race, we can create a measure of transitions from unemployment to employment (U-to-E). While total jobs reflects a point-in-time employment opportunity for an area, U-to-E transitions could be argued to reflect greater labor market dynamism, thus changing opportunity.

In this case, the Delta Index is calculated to compare the distribution of people of a particular race and education level with the distribution of year-to-year transitions by workers of the same race and education level (a similar analysis using monthly transitions produces similar results):

\[
D^r_e = \frac{1}{2} \sum_{s=1}^{50} \left| \frac{n^r_{es}}{N^r_e} - \frac{t^r_{es}}{T^r_e} \right|, \tag{A1}
\]

where,

\( n^r_{es} \) and \( N^r_e \) are defined as above in equation (1);

\( t^r_{es} \) = number of U-to-E transitions from one year to the next made by workers of racial group, \( r \), with education level, \( e \), in state, \( s \); and

\( T^r_e \) = total number of U-to-E transitions in the U.S. from one year to the next made by workers of racial group, \( r \), with education level, \( e \).
Figure A1 reproduces Figure 3 from the text, comparing the degree of mismatch between race/education job opportunities (measured by the share of U-to-E transitions in the location) and race/education specific population. If anything, Figure A1 presents even more compelling evidence of greater mismatch, at all education levels, among racial/ethnic minorities than among white, non-Hispanics.

[Figure A1 about here]

A.2 The Multivariate Regression Analysis

Equation (2) in the text is re-specified with changes in the share of year-to-year U-to-E transitions as the measure of local labor market opportunities:

\[
\Delta \left( \frac{n_e^t}{N_e^t} \right)_{g,t} = \alpha + \rho \Delta \left( \frac{t_e^t}{t_e^t} \right)_{g,t} + \sum_{j=1}^{2} \left\{ \beta_j^1 \text{Race}^j_{g,t} + \beta_j^2 \text{RACE}^j_{g,t} \right\} \Delta \left( \frac{t_e^t}{t_e^t} \right)_{g,t} \\
+ \sum_{k=1}^{2} \left\{ \varphi_k^1 \text{EDUC}^k_{g,t} + \varphi_k^2 \text{EDUC}^k_{g,t} \right\} \Delta \left( \frac{t_e^t}{t_e^t} \right)_{g,t} \\
+ \sum_{j=1}^{2} \sum_{k=1}^{2} \left\{ \lambda_{jk}^1 \text{RACE}^j_{g,t} \text{EDUC}^k_{g,t} + \lambda_{jk}^2 \text{RACE}^j_{g,t} \text{EDUC}^k_{g,t} \right\} \Delta \left( \frac{t_e^t}{t_e^t} \right)_{g,t} \\
+ \theta' \chi_{g,t-1} + \gamma_{ger} + \tau_t + \sigma_g + \varepsilon_{g,e,r,t}, \tag{A2}
\]

where,

\[
\Delta \left( \frac{n_e^t}{N_e^t} \right)_{g,t} = \text{change in the share of all people of racial group, } r, \text{ with education, } e, \text{ in geography, } g, \text{ from } t-1 \text{ to } t;
\]

\[
\Delta \left( \frac{t_e^t}{t_e^t} \right)_{g,t} = \text{change in the share of all U-to-E transitions from one year to the next made by workers of racial group, } r, \text{ with education level, } e, \text{ in geography, } g, \text{ from } t-1 \text{ to } t;
\]

\( \text{RACE}_{g,t} \) = set of 0,1 regressors indicating black, non-Hispanic or Hispanic race/ethnicity;

\( \text{EDUC}_{g,t} \) = set of 0,1 regressors indicating some college or college plus education groups;
\(X_{g,t-1} = \text{geography specific additional regressors at time } t-1\), including the unemployment rate and industry shares, which are expected to capture both baseline job opportunities and unemployment risk in the geographic location;

\(\tau_t, \sigma_g, \text{ and } \varepsilon_{g,e,r,t}\) time, geography, and geography/race/education fixed effects respectively.

\(\varepsilon_{g,e,r,t}\) are robust standard errors, clustered at the geography level (each CBSA and CZ are observed multiple times across years). The geography/race/education fixed effect is expected to control for time-invariant group effects. For example, at the geographic level this would account for differences in amenities; and at the race/education level, the fixed effect would account for underlying race/education specific differences in migration patterns. The unit of observation is race/education/geography/year and the analysis is performed for both CBSA and CZ geography levels. The analysis excludes less than high school. Note that the definition of commuting zones depends on knowing a person's county, which is often missing; so there will be more observations in the CBSA analysis than in the CZ analysis. Table A1 reports the marginal effects from this estimation that are analogous to those reported in Table X in the text.
Figure A1 Delta Index by education across race and time, total number of year-to-year UE transitions across U.S. states, CBSAs, and CZs.

(a) Across States

(b) Across CBSAs

(c) Across CZs

Source: Authors calculations using the CPS monthly data January 1996 through November 2018. Education level less-than-high-school excluded for illustrative purposes (available upon request). Comparable to Figure 3 in text.
Table A1 Marginal effects of a change in location U-to-E 12 month transitions share \( \Delta \left( \frac{t_{12}}{U_{12}} \right) \) on share of population \( \Delta \left( \frac{u}{N} \right) \) in the location, by race/ethnicity and educational attainment; comparable to Table 3 in text.

<table>
<thead>
<tr>
<th>Geographic area:</th>
<th>Fixed Effects</th>
<th>Random Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CBSA (3)</td>
<td>CZone (4)</td>
</tr>
<tr>
<td>White, NH</td>
<td>0.0086</td>
<td>0.0333***</td>
</tr>
<tr>
<td></td>
<td>[0.0061]</td>
<td>[0.0099]</td>
</tr>
<tr>
<td>High School</td>
<td>0.0068</td>
<td>0.0371***</td>
</tr>
<tr>
<td></td>
<td>[0.0092]</td>
<td>[0.0142]</td>
</tr>
<tr>
<td>Some College</td>
<td>0.0082</td>
<td>0.0370**</td>
</tr>
<tr>
<td></td>
<td>[0.0092]</td>
<td>[0.0154]</td>
</tr>
<tr>
<td>College and Above</td>
<td>0.0141**</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>[0.0066]</td>
<td>[0.0177]</td>
</tr>
<tr>
<td>Black, NH</td>
<td>0.0147*</td>
<td>0.0298**</td>
</tr>
<tr>
<td></td>
<td>[0.0079]</td>
<td>[0.0141]</td>
</tr>
<tr>
<td>High School</td>
<td>0.0057</td>
<td>0.0114</td>
</tr>
<tr>
<td></td>
<td>[0.0184]</td>
<td>[0.0228]</td>
</tr>
<tr>
<td>Some College</td>
<td>0.0221</td>
<td>0.0341**</td>
</tr>
<tr>
<td></td>
<td>[0.0141]</td>
<td>[0.0171]</td>
</tr>
<tr>
<td>College and Above</td>
<td>0.0242***</td>
<td>0.0654**</td>
</tr>
<tr>
<td></td>
<td>[0.0087]</td>
<td>[0.0296]</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.0246**</td>
<td>0.0476**</td>
</tr>
<tr>
<td></td>
<td>[0.0116]</td>
<td>[0.0198]</td>
</tr>
<tr>
<td>High School</td>
<td>0.0547*</td>
<td>0.0740**</td>
</tr>
<tr>
<td></td>
<td>[0.0294]</td>
<td>[0.0313]</td>
</tr>
<tr>
<td>Some College</td>
<td>-0.0048</td>
<td>0.0290*</td>
</tr>
<tr>
<td></td>
<td>[0.0113]</td>
<td>[0.0157]</td>
</tr>
<tr>
<td>College and Above</td>
<td>0.0011</td>
<td>0.0192</td>
</tr>
<tr>
<td></td>
<td>[0.0110]</td>
<td>[0.0225]</td>
</tr>
<tr>
<td>Observations</td>
<td>12,768</td>
<td>7,258</td>
</tr>
<tr>
<td>Number of FE</td>
<td>1,430</td>
<td>887</td>
</tr>
<tr>
<td>Clusters</td>
<td>300</td>
<td>180</td>
</tr>
</tbody>
</table>

Note: Robust standard errors are clustered at the level of geography. *, **, *** => statistical significance at the 90, 95, and 99 percent level. Additional controls include the area's baseline unemployment rate and industry shares. Sample includes 25-54 year-old men with at least a high school degree and 1996-2018 years of data. Full estimation results available upon request.
Table A2 Marginal effects of a change in the 12 month U-to-E transition shares share among black, NH and Hispanics on share of population in CBSA, by own race/ethnicity, at different points in the distribution of CBSA race/ethnicity population share; comparable to Table 4 in text; random effects specification.

<table>
<thead>
<tr>
<th></th>
<th>Core Statistical Business Area Analysis</th>
<th></th>
<th>Commuting Zone Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Own race population share in CBSA, percentile</td>
<td>Own race population share in CZ, percentile</td>
<td></td>
</tr>
<tr>
<td></td>
<td>25th Percentile</td>
<td>50th Percentile</td>
<td>75th Percentile</td>
</tr>
<tr>
<td><strong>M.E. for Blacks, NH</strong></td>
<td>0.0181*</td>
<td>0.0174*</td>
<td>0.0164**</td>
</tr>
<tr>
<td></td>
<td>[0.0109]</td>
<td>[0.0091]</td>
<td>[0.0074]</td>
</tr>
<tr>
<td>High School</td>
<td>0.012</td>
<td>0.0102</td>
<td>0.0075</td>
</tr>
<tr>
<td></td>
<td>[0.0227]</td>
<td>[0.0203]</td>
<td>[0.0182]</td>
</tr>
<tr>
<td>Some College</td>
<td>0.0255</td>
<td>0.0251</td>
<td>0.0245*</td>
</tr>
<tr>
<td></td>
<td>[0.0194]</td>
<td>[0.0162]</td>
<td>[0.0130]</td>
</tr>
<tr>
<td>College or Above</td>
<td>0.0196</td>
<td>0.0212*</td>
<td>0.0237***</td>
</tr>
<tr>
<td></td>
<td>[0.0159]</td>
<td>[0.0128]</td>
<td>[0.0089]</td>
</tr>
<tr>
<td><strong>M.E. for Hispanics</strong></td>
<td>0.01</td>
<td>0.0123</td>
<td><strong>0.0186</strong></td>
</tr>
<tr>
<td></td>
<td>[0.0096]</td>
<td>[0.0094]</td>
<td><strong>[0.0098]</strong></td>
</tr>
<tr>
<td>High School</td>
<td>0.0534**</td>
<td>0.0538**</td>
<td>0.0549**</td>
</tr>
<tr>
<td></td>
<td>[0.0260]</td>
<td>[0.0258]</td>
<td>[0.0270]</td>
</tr>
<tr>
<td>Some College</td>
<td>-0.0404**</td>
<td>-0.0353**</td>
<td><strong>-0.0211</strong></td>
</tr>
<tr>
<td></td>
<td>[0.0177]</td>
<td>[0.0164]</td>
<td><strong>[0.0135]</strong></td>
</tr>
<tr>
<td>College or Above</td>
<td>-0.0096</td>
<td>-0.0076</td>
<td>-0.0021</td>
</tr>
<tr>
<td></td>
<td>[0.0120]</td>
<td>[0.0112]</td>
<td>[0.0106]</td>
</tr>
</tbody>
</table>

Note: Robust standard errors are clustered at the CBSA level. There are 300 CBSAs and 180 CZs. *, **, *** => statistical significance at the 90, 95, and 99 percent level. Regression includes location specific unemployment rate and industry shares. Sample includes 25-54 year-old men with at least a high school degree and 1996-2018 years of data. Full estimation results available upon request.
### Appendix B: Parameter Coefficient Estimates

**Table B1** CBSA and CZ results; the dependent variable is the year-to-year change in share of population $\Delta \left( \frac{a}{N} \right)$, change in job shares; produces marginal effects found in Table 3 in text.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Fixed Effects Model</th>
<th>Random Effects Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CBSA</td>
<td>CBSA</td>
</tr>
<tr>
<td></td>
<td>CZ</td>
<td>CZ</td>
</tr>
<tr>
<td>Change in Job Share</td>
<td>0.5637749***</td>
<td>0.6229410***</td>
</tr>
<tr>
<td></td>
<td>(0.0278852)</td>
<td>(0.0360957)</td>
</tr>
<tr>
<td></td>
<td>0.5701856***</td>
<td>0.6346650***</td>
</tr>
<tr>
<td></td>
<td>(0.0285874)</td>
<td>(0.0353831)</td>
</tr>
<tr>
<td>Black, NH*Change in Job Share</td>
<td>-0.0307846</td>
<td>-0.0440886</td>
</tr>
<tr>
<td></td>
<td>(0.0382499)</td>
<td>(0.0528527)</td>
</tr>
<tr>
<td></td>
<td>-0.0388586</td>
<td>-0.0523334</td>
</tr>
<tr>
<td>Hispanic*Change in Job Share</td>
<td>0.0336981</td>
<td>0.1454763***</td>
</tr>
<tr>
<td></td>
<td>(0.0460081)</td>
<td>(0.0466190)</td>
</tr>
<tr>
<td></td>
<td>0.0382927</td>
<td>0.1319804***</td>
</tr>
<tr>
<td></td>
<td>(0.0439190)</td>
<td>(0.0449022)</td>
</tr>
<tr>
<td>Some College*Change in Job Share</td>
<td>-0.0466737</td>
<td>0.1087971**</td>
</tr>
<tr>
<td></td>
<td>(0.0464151)</td>
<td>(0.0539799)</td>
</tr>
<tr>
<td></td>
<td>-0.0498561</td>
<td>0.0947603*</td>
</tr>
<tr>
<td></td>
<td>(0.0478723)</td>
<td>(0.0524134)</td>
</tr>
<tr>
<td>College and Above*Change in Job Share</td>
<td>0.0535944</td>
<td>0.2279417***</td>
</tr>
<tr>
<td></td>
<td>(0.0626347)</td>
<td>(0.0480742)</td>
</tr>
<tr>
<td></td>
<td>0.0418313</td>
<td>0.2187585***</td>
</tr>
<tr>
<td></td>
<td>(0.0653168)</td>
<td>(0.0445213)</td>
</tr>
<tr>
<td>Black, NH<em>Some College</em>Change in Job Share</td>
<td>-0.0853157</td>
<td>-0.0913255</td>
</tr>
<tr>
<td></td>
<td>(0.1056606)</td>
<td>(0.0704872)</td>
</tr>
<tr>
<td></td>
<td>-0.0774849</td>
<td>-0.0806739</td>
</tr>
<tr>
<td>Black, NH<em>College and Above</em>Change in Job</td>
<td>-0.0480971</td>
<td>-0.1674409**</td>
</tr>
<tr>
<td></td>
<td>(0.0727114)</td>
<td>(0.0684606)</td>
</tr>
<tr>
<td></td>
<td>-0.0346695</td>
<td>-0.1645122**</td>
</tr>
<tr>
<td>Hispanic<em>Some College</em>Change in Job Share</td>
<td>-0.1174401</td>
<td>-0.2398716**</td>
</tr>
<tr>
<td></td>
<td>(0.0915795)</td>
<td>(0.0931641)</td>
</tr>
<tr>
<td></td>
<td>-0.1190162</td>
<td>-0.2201955**</td>
</tr>
<tr>
<td>Hispanic<em>College and Above</em>Change in Job</td>
<td>-0.2566847***</td>
<td>-0.4684578***</td>
</tr>
<tr>
<td></td>
<td>(0.0909986)</td>
<td>(0.0853731)</td>
</tr>
<tr>
<td></td>
<td>-0.25464726***</td>
<td>-0.4514762***</td>
</tr>
<tr>
<td></td>
<td>(0.0880296)</td>
<td>(0.0880005)</td>
</tr>
<tr>
<td>Lag CBSA Unemployment Rate</td>
<td>0.00000007</td>
<td>0.00000005</td>
</tr>
<tr>
<td></td>
<td>(0.0000027)</td>
<td>(0.00000073)</td>
</tr>
<tr>
<td></td>
<td>0.00000006</td>
<td>0.00000008</td>
</tr>
<tr>
<td></td>
<td>(0.0000096)</td>
<td>(0.0000277)</td>
</tr>
<tr>
<td>Black, NH</td>
<td>-0.000067</td>
<td>-0.0000119</td>
</tr>
<tr>
<td></td>
<td>(0.0000093)</td>
<td>(0.0000214)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.0000004</td>
<td>0.00000069</td>
</tr>
<tr>
<td></td>
<td>(0.0000063)</td>
<td>(0.0000141)</td>
</tr>
<tr>
<td>Some College</td>
<td>0.0000024</td>
<td>0.00000049</td>
</tr>
<tr>
<td></td>
<td>(0.0000024)</td>
<td>(0.00000144)</td>
</tr>
<tr>
<td>College and Above</td>
<td>-0.0000155</td>
<td>-0.0000218</td>
</tr>
<tr>
<td></td>
<td>(0.0000132)</td>
<td>(0.0000366)</td>
</tr>
<tr>
<td>Black, NH*Some College</td>
<td>-0.0000101</td>
<td>-0.0000423</td>
</tr>
<tr>
<td></td>
<td>(0.0000168)</td>
<td>(0.0000389)</td>
</tr>
<tr>
<td>Hispanic*Some College</td>
<td>-0.0000055</td>
<td>-0.0000171</td>
</tr>
<tr>
<td></td>
<td>(0.0000152)</td>
<td>(0.0000326)</td>
</tr>
<tr>
<td>Hispanic*College and Above</td>
<td>-0.0000313**</td>
<td>-0.0000545</td>
</tr>
<tr>
<td>VARIABLES</td>
<td>Fixed Effects Model</td>
<td>Random Effects Model</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>---------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td></td>
<td>CBSA</td>
<td>CZ</td>
</tr>
<tr>
<td>Constant</td>
<td>0.0003495 (0.0005250)</td>
<td>0.0006889 (0.0011198)</td>
</tr>
<tr>
<td>Observations</td>
<td>36,573</td>
<td>20,303</td>
</tr>
<tr>
<td>Number of FE (location/education/race)</td>
<td>2,557</td>
<td>1,560</td>
</tr>
<tr>
<td>R-squared Within</td>
<td>0.132</td>
<td>0.243</td>
</tr>
<tr>
<td>R-squared Between</td>
<td>0.271</td>
<td>0.669</td>
</tr>
<tr>
<td>R-squared Overall</td>
<td>0.137</td>
<td>0.254</td>
</tr>
<tr>
<td>Number of Clusters (location)</td>
<td>307</td>
<td>188</td>
</tr>
</tbody>
</table>

Note: Robust standard errors are clustered at the geography level. *, **, *** => statistical significance at the 90, 95, and 99 percent level. Sample includes 25-54 year-old men with at least a high school degree and 1996-2018 years of data. Additional controls include location specific industry shares and location and year fixed effects. A Hausman test rejects equality between the random and fixed effects parameter estimates.
Appendix C: Regression specification including interactions with geographic location share of population that is black/Hispanic.

This appendix describes an estimation specification that modifies equation (2) to take into account the share of the minority population in the location where growing job opportunities are observed. If social costs are important to the migration decision, we should observe that blacks and Hispanics are more willing to respond to growing labor market opportunities, all else equal, in locations with larger population shares of racial minorities. Equation (2) is modified as follows:

\[
\begin{align*}
\Delta \left( \frac{N_e}{N_t} \right)_{g,t} &= \alpha + \rho \Delta \left( \frac{I_e}{I_t} \right)_{g,t} + \sum_{j=1}^{2} \left\{ \beta_j^1 \text{RACE}_{g,t}^j + \beta_j^2 \text{RACE}_{g,t}^j \Delta \left( \frac{I_e}{I_t} \right)_{g,t} \right\} \\
&+ \sum_{k=1}^{2} \left\{ \varphi_k \text{EDUC}_k^g + \varphi_k \text{EDUC}_k^g \Delta \left( \frac{I_e}{I_t} \right)_{g,t} \right\} \\
&+ \sum_{j=1}^{2} \sum_{k=1}^{2} \left\{ \lambda_{jk} \text{RACE}_{g,t}^j \text{EDUC}_k^g + \lambda_{jk} \text{RACE}_{g,t}^j \text{EDUC}_k^g \Delta \left( \frac{I_e}{I_t} \right)_{g,t} \right\} \\
&+ \sum_{l=1}^{2} \left\{ \omega_l \text{SHARErace}_{g,t}^l + \omega_l \text{SHARErace}_{g,t}^l \Delta \left( \frac{I_e}{I_t} \right)_{g,t} \right\} \\
&+ \sum_{k=1}^{2} \sum_{l=1}^{2} \left\{ \eta_{kl} \text{EDUC}_{g,t}^k \text{SHARErace}_{g,t}^l + \eta_{kl} \text{EDUC}_{g,t}^k \text{SHARErace}_{g,t}^l \Delta \left( \frac{I_e}{I_t} \right)_{g,t} \right\} \\
&+ \theta'X_{g,t-1} + \tau_t + \sigma_g + \epsilon_{g,e,r,t}.
\end{align*}
\]

In this specification, the share of the population in location \( g \) that is black, non-Hispanic or Hispanic enters the regression by itself, interacted with education, and also interacted with changing job opportunities in location \( g \). Increasing responsiveness to job market opportunities in CBSAs with higher shares of same ethnic/racial population would suggest that social costs
could be constraining migration of ethnic/racial minorities. Marginal effects of changing transition shares on changes in population shares at different points in the population race/ethnicity share distribution are reported in the text and the full set of estimation results for CBSA and CZ are found in Table C1.
Table C1 Regression results of a change in the job share among black, NH and Hispanics on share of population in CBSA or CZ, by own race/ethnicity, at different points in the distribution of CBSA race/ethnicity population share; produces marginal effects found in Table 4 in text.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Fixed Effects Model</th>
<th>Random Effects Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CBSA</td>
<td>CZ</td>
</tr>
<tr>
<td>Change in Job Share</td>
<td>0.5956279***</td>
<td>0.5789529***</td>
</tr>
<tr>
<td></td>
<td>(0.0412233)</td>
<td>(0.0611491)</td>
</tr>
<tr>
<td>Black, NH*Job Share</td>
<td>-0.0345456</td>
<td>-0.0811573</td>
</tr>
<tr>
<td></td>
<td>(0.0409039)</td>
<td>(0.0558995)</td>
</tr>
<tr>
<td>Hispanic*Job Share</td>
<td>0.0745500</td>
<td>0.1489807***</td>
</tr>
<tr>
<td></td>
<td>(0.0509310)</td>
<td>(0.0530096)</td>
</tr>
<tr>
<td>Some College*Job Share</td>
<td>-0.2043757**</td>
<td>0.0051501</td>
</tr>
<tr>
<td></td>
<td>(0.0966772)</td>
<td>(0.0814370)</td>
</tr>
<tr>
<td>College and Above*Job Share</td>
<td>-0.0123773</td>
<td>0.1116080</td>
</tr>
<tr>
<td></td>
<td>(0.1037141)</td>
<td>(0.0805704)</td>
</tr>
<tr>
<td>Black, NH<em>Some College</em>Job Share</td>
<td>-0.1140743</td>
<td>-0.0896577</td>
</tr>
<tr>
<td></td>
<td>(0.1133889)</td>
<td>(0.0803677)</td>
</tr>
<tr>
<td>Black, NH<em>College and Above</em>Job Share</td>
<td>-0.0709120</td>
<td>-0.2609688***</td>
</tr>
<tr>
<td></td>
<td>(0.0656861)</td>
<td>(0.0735753)</td>
</tr>
<tr>
<td>Hispanic<em>Some College</em>Job Share</td>
<td>-0.2675755***</td>
<td>-0.3874840***</td>
</tr>
<tr>
<td></td>
<td>(0.0895653)</td>
<td>(0.0911077)</td>
</tr>
<tr>
<td>Hispanic<em>College and Above</em>Job Share</td>
<td>-0.2415866***</td>
<td>-0.4033302***</td>
</tr>
<tr>
<td></td>
<td>(0.0899247)</td>
<td>(0.0972957)</td>
</tr>
<tr>
<td>Share Black, NH</td>
<td>0.0004604**</td>
<td>0.0010173</td>
</tr>
<tr>
<td></td>
<td>(0.0002235)</td>
<td>(0.0006837)</td>
</tr>
<tr>
<td>Share Hispanic</td>
<td>0.0002590</td>
<td>0.0003404</td>
</tr>
<tr>
<td></td>
<td>(0.0002195)</td>
<td>(0.0005008)</td>
</tr>
<tr>
<td>Some College*Share Black, NH</td>
<td>0.0004002</td>
<td>0.0001121</td>
</tr>
<tr>
<td></td>
<td>(0.0003941)</td>
<td>(0.0010872)</td>
</tr>
<tr>
<td>College and Above*Share Black, NH</td>
<td>-0.0002754</td>
<td>-0.0010732</td>
</tr>
<tr>
<td></td>
<td>(0.0003185)</td>
<td>(0.0007672)</td>
</tr>
<tr>
<td>Some College*Share Hispanic</td>
<td>0.0005178</td>
<td>0.0011673</td>
</tr>
<tr>
<td></td>
<td>(0.0003526)</td>
<td>(0.0007836)</td>
</tr>
<tr>
<td>College and Above*Share Hispanic</td>
<td>-0.0002740</td>
<td>0.0010405</td>
</tr>
<tr>
<td></td>
<td>(0.0004092)</td>
<td>(0.0009243)</td>
</tr>
<tr>
<td>Share Black, NH*Job Share</td>
<td>0.0334551</td>
<td>0.3958826</td>
</tr>
<tr>
<td></td>
<td>(0.2004900)</td>
<td>(0.3359306)</td>
</tr>
<tr>
<td>VARIABLES</td>
<td>Fixed Effects Model</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>---------------------</td>
<td>-------------</td>
</tr>
<tr>
<td></td>
<td>CBSA</td>
<td>CZ</td>
</tr>
<tr>
<td>Share Hispanic*Job Share</td>
<td>-0.2896950**</td>
<td>-0.0053539</td>
</tr>
<tr>
<td></td>
<td>(0.1235172)</td>
<td>(0.1677579)</td>
</tr>
<tr>
<td>Some College<em>Share Black, NH</em>Job Share</td>
<td>0.3139990</td>
<td>-0.0358843</td>
</tr>
<tr>
<td></td>
<td>(0.6236709)</td>
<td>(0.3655042)</td>
</tr>
<tr>
<td>Coll. and Above<em>Share Black, NH</em>Job Share</td>
<td>0.4962632</td>
<td>1.1553783***</td>
</tr>
<tr>
<td></td>
<td>(0.5156328)</td>
<td>(0.3570404)</td>
</tr>
<tr>
<td>Some College<em>Share Hispanic</em>Job Share</td>
<td>0.9083513***</td>
<td>0.7446835***</td>
</tr>
<tr>
<td></td>
<td>(0.2553295)</td>
<td>(0.2130460)</td>
</tr>
<tr>
<td>College and Above<em>Share Hispanic</em>Job Share</td>
<td>0.0218304</td>
<td>-0.1917244</td>
</tr>
<tr>
<td></td>
<td>(0.2051024)</td>
<td>(0.1849187)</td>
</tr>
<tr>
<td>Lag CBSA Unemployment Rate</td>
<td>-0.0000001</td>
<td>-0.0000018</td>
</tr>
<tr>
<td></td>
<td>(0.0000028)</td>
<td>(0.0000070)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.0001746</td>
<td>0.0005044</td>
</tr>
<tr>
<td></td>
<td>(0.0005199)</td>
<td>(0.0009977)</td>
</tr>
<tr>
<td>Black, NH</td>
<td>0.00000075</td>
<td>-0.000017</td>
</tr>
<tr>
<td></td>
<td>(0.0000096)</td>
<td>(0.0000211)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.00000117</td>
<td>-0.0000433</td>
</tr>
<tr>
<td></td>
<td>(0.0000118)</td>
<td>(0.0000290)</td>
</tr>
<tr>
<td>Some College</td>
<td>0.0000031</td>
<td>-0.0000202</td>
</tr>
<tr>
<td></td>
<td>(0.0000101)</td>
<td>(0.0000271)</td>
</tr>
<tr>
<td>College and Above</td>
<td>-0.0000200**</td>
<td>-0.0000299</td>
</tr>
<tr>
<td></td>
<td>(0.0000119)</td>
<td>(0.0000351)</td>
</tr>
<tr>
<td>Black, NH* Some College</td>
<td>-0.00000136</td>
<td>-0.0000615</td>
</tr>
<tr>
<td></td>
<td>(0.0000170)</td>
<td>(0.0000387)</td>
</tr>
<tr>
<td>Hispanic*Some College</td>
<td>-0.0000074</td>
<td>-0.0000138</td>
</tr>
<tr>
<td></td>
<td>(0.0000145)</td>
<td>(0.0000319)</td>
</tr>
<tr>
<td>Hispanic*College and Above</td>
<td>-0.0000313**</td>
<td>-0.0000485</td>
</tr>
<tr>
<td></td>
<td>(0.0000152)</td>
<td>(0.0000376)</td>
</tr>
<tr>
<td>Observations</td>
<td>36,573</td>
<td>36,573</td>
</tr>
<tr>
<td>Number of FE</td>
<td>2,557</td>
<td>1,560</td>
</tr>
<tr>
<td>R-squared Within</td>
<td>0.135</td>
<td>0.251</td>
</tr>
<tr>
<td>R-squared Between</td>
<td>0.136</td>
<td>0.540</td>
</tr>
<tr>
<td>R-squared Overall</td>
<td>0.135</td>
<td>0.258</td>
</tr>
<tr>
<td>Number of Clusters</td>
<td>307</td>
<td>188</td>
</tr>
</tbody>
</table>
Note: Robust standard errors are clustered at the CBSA or CZ level. *, **, *** => statistical significance at the 90, 95, and 99 percent level. Additional controls include location specific industry shares and location and year fixed effects. Sample includes 25-54 year-old men with at least a high school degree and 1996-2018 years of data. A Hausman test rejects equality between the random and fixed effects parameter estimates.
Appendix D: Synthesizing CBSA Codes Across Time

In 2003, OMB changed classifications of metropolitan areas from “Metropolitan Statistical Areas” (MSAs) to “Core Based Statistical Areas” (CBSAs). MSAs had four digit codes and CBSAs have 5 digit codes. For many MSAs, a CBSA code is directly comparable. However, there are several problems with CBSA code consistency post 2003.

Here we document problems encountered in trying to create a crosswalk between MSA and CBSA codes contained in the Current Population Survey (CPS) across time and the solutions employed. Much of the difficulty arises from the fact that post 2003, CBSAs also include micropolitan areas that were never part of MSAs.

For all but a small number of CBSAs, a one-to-one match was possible based on the names of the CBSA and MSA. The locations for which a direct name match was not possible are detailed in Section D.1. Additionally, there were a handful of occurrences where other geographic codes were incorrectly recorded as CBSA codes. This problem and its solutions are detailed Section D.2. Lastly, there were four CBSAs that simply changed codes over time. The solution to this problem is detailed in Section D.3.

Another source for consistent classifications for metropolitan areas is IPUMS:CPS (https://cps.ipums.org/cps/index.shtml). However, they convert CBSA codes to MSA codes. Since CBSA codes more often than not encompass a larger geographic area than an MSA code, we choose to convert MSA codes prior to 2003 to their more recent CBSA counterpart.

---

2 Sources used for matching on names and counties: https://www.census.gov/population/estimates/metro-city/0312msa.txt
https://www.census.gov/population/estimates/metro-city/99mfips.txt
The final MSA to CBSA crosswalk was verified by summing up the population in the synthesized CBSAs to make sure there is a smooth transition across the introduction of CBSAs.

D.1 Problem: MSAs and CBSAs cannot be matched by name only

**Solution:** Smaller areas previously known as MSAs are combined into larger CBSAs

<table>
<thead>
<tr>
<th>CBSA Name</th>
<th>CBSA Code (MSA Code below)</th>
</tr>
</thead>
<tbody>
<tr>
<td>New York-Northern New Jersey-Long Island, NY-NJ-PA</td>
<td>35620</td>
</tr>
<tr>
<td>Middlesex-Somerset-Hunterdon, NJ</td>
<td>5015</td>
</tr>
<tr>
<td>Monmouth-Ocean, NJ</td>
<td>5190</td>
</tr>
<tr>
<td>Newark, NJ</td>
<td>5640</td>
</tr>
<tr>
<td>Bergen-Passaic, NJ</td>
<td>875</td>
</tr>
<tr>
<td>Jersey City, NY</td>
<td>3640</td>
</tr>
<tr>
<td>Nassau-Suffolk, NY</td>
<td>5380</td>
</tr>
<tr>
<td>New York, NY</td>
<td>5600</td>
</tr>
<tr>
<td>San Francisco-Oakland-Fremont, CA</td>
<td>41860</td>
</tr>
<tr>
<td>San Francisco-Oakland-San Jose, CA</td>
<td>5775</td>
</tr>
<tr>
<td>San Francisco, CA</td>
<td>7360</td>
</tr>
<tr>
<td>Los Angeles-Long Beach-Santa Ana</td>
<td>31100</td>
</tr>
<tr>
<td>Los Angeles-Long Beach, CA</td>
<td>4480</td>
</tr>
<tr>
<td>Orange County, CA</td>
<td>5945</td>
</tr>
<tr>
<td>Boston-Cambridge-Quincy, MA-NH</td>
<td>14460</td>
</tr>
<tr>
<td>Boston, MA-NH</td>
<td>1120</td>
</tr>
<tr>
<td>Lowell, MA-NH</td>
<td>4560</td>
</tr>
<tr>
<td>Lawrence, MA-NH</td>
<td>4160</td>
</tr>
<tr>
<td>Brockton, MA</td>
<td>1200</td>
</tr>
<tr>
<td>Portsmouth-Rochester, NH-MEa</td>
<td>6450</td>
</tr>
<tr>
<td>Worcester, MA</td>
<td>49340</td>
</tr>
<tr>
<td>Worcester, MA-CTb</td>
<td>9240</td>
</tr>
<tr>
<td>Fitchburg-Leominster, MA</td>
<td>2600</td>
</tr>
<tr>
<td>Seattle-Tacoma-Bellevue, WA</td>
<td>42660</td>
</tr>
<tr>
<td>Seattle-Bellevue-Everett, WA</td>
<td>7600</td>
</tr>
<tr>
<td>Tacoma, WA</td>
<td>8200</td>
</tr>
<tr>
<td>Miami-Fort Lauderdale-Miami Beach, FL</td>
<td>33100</td>
</tr>
<tr>
<td>Miami, FL</td>
<td>5000</td>
</tr>
<tr>
<td>Fort Lauderdale, FL</td>
<td>2680</td>
</tr>
<tr>
<td>West Palm Beach-Boca Raton, FL</td>
<td>8960</td>
</tr>
<tr>
<td>Philadelphia-Camden-Wilmington, PA-NJ-DE-MD</td>
<td>37980</td>
</tr>
<tr>
<td>Philadelphia, PA-NJ</td>
<td>6160</td>
</tr>
<tr>
<td>Wilmington-Newark, DE-MD</td>
<td>9160</td>
</tr>
<tr>
<td>Poughkeepsie-Newburgh-Middletown, NY</td>
<td>39100</td>
</tr>
<tr>
<td>Dutchess County, NY</td>
<td>2281</td>
</tr>
<tr>
<td>CBSA Name</td>
<td>CBSA Code</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>(names of MSAs combined are listed below the CBSA name)</td>
<td>(MSA Code below)</td>
</tr>
<tr>
<td>Newburgh, NY-PA&lt;sup&gt;c&lt;/sup&gt;</td>
<td>5660</td>
</tr>
<tr>
<td>Houston-Baytown-Sugar Land, TX</td>
<td>26420</td>
</tr>
<tr>
<td>Houston, TX</td>
<td>3360</td>
</tr>
<tr>
<td>Galveston-Texas City, TX</td>
<td>2920</td>
</tr>
<tr>
<td>Brazoria, TX</td>
<td>1145</td>
</tr>
<tr>
<td>Bridgeport-Stamford-Norwalk, CT</td>
<td>14860</td>
</tr>
<tr>
<td>Danbury, CT</td>
<td>1930</td>
</tr>
<tr>
<td>Bridgeport, CT</td>
<td>1160</td>
</tr>
<tr>
<td>Stamford-Norwalk, CT</td>
<td>8040</td>
</tr>
<tr>
<td>Dallas-Fort Worth-Arlington, TX</td>
<td>19100</td>
</tr>
<tr>
<td>Dallas, TX</td>
<td>1920</td>
</tr>
<tr>
<td>Fort Worth-Arlington, TX</td>
<td>2800</td>
</tr>
<tr>
<td>New Haven-Milford, CT</td>
<td>35300</td>
</tr>
<tr>
<td>New Haven-Meriden, CT</td>
<td>5480</td>
</tr>
<tr>
<td>Waterbury, CT</td>
<td>8880</td>
</tr>
<tr>
<td>Chicago-Naperville-Joliet, IL-IN-WI</td>
<td>16980</td>
</tr>
<tr>
<td>Chicago, IL</td>
<td>1600</td>
</tr>
<tr>
<td>Gary, IN</td>
<td>2960</td>
</tr>
<tr>
<td>Sacramento--Arden-Arcade--Roseville, CA</td>
<td>40900</td>
</tr>
<tr>
<td>Sacramento, CA</td>
<td>6920</td>
</tr>
<tr>
<td>Yolo, CA</td>
<td>9270</td>
</tr>
<tr>
<td>Cincinnati-Middletown, OH-KY-IN</td>
<td>17140</td>
</tr>
<tr>
<td>Cincinnati, OH-KY-IN</td>
<td>1640</td>
</tr>
<tr>
<td>Hamilton-Middletown, OH</td>
<td>3200</td>
</tr>
<tr>
<td>Manchester-Nashua, NH</td>
<td>31700</td>
</tr>
<tr>
<td>Manchester, NH</td>
<td>4760</td>
</tr>
<tr>
<td>Nashua, NH</td>
<td>5350</td>
</tr>
<tr>
<td>Providence-New Bedford-Fall River, RI-MA</td>
<td>39300</td>
</tr>
<tr>
<td>Providence-Fall River-Warwick, RI-MA</td>
<td>6480</td>
</tr>
<tr>
<td>New Bedford, MA</td>
<td>5400</td>
</tr>
</tbody>
</table>

<sup>a</sup> NH residents only; ME residents from this MSA are assigned to a different CBSA.
<sup>b</sup> MA residents only; CT residents from this MSA are assigned to a different CBSA.
<sup>c</sup> NY residents only; PA residents from this MSA are assigned to a different CBSA.

Sources of information:
Crosswalk between PMSA and MSA codes: https://www.census.gov/population/estimates/metro-city/99mfips.txt
Crosswalk between CBSA and CBSA divisions:
https://www.census.gov/population/estimates/metro-city/0312msa.txt
**D.2 Problem: Incorrect CBSA Codes Reported**

_D.2.a New England City and Town Area (NECTA) codes were incorrectly reported as CBSA codes for some observations in New England._

**Solution:** Match NECTA and CBSA names to assign correct CBSA code. All NECTA codes all begin with a "7" and no CBSA codes begin with "7," so this error was easily identified.

**Changes Made:**

<table>
<thead>
<tr>
<th>NECTA</th>
<th>NECTA Name</th>
<th>CBSA</th>
<th>CBSA Name</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>70750</td>
<td>Bangor, ME</td>
<td>12620</td>
<td>Bangor, ME</td>
<td></td>
</tr>
<tr>
<td>70900</td>
<td>Barnstable Town, ME</td>
<td>12700</td>
<td>Barnstable Town, ME</td>
<td></td>
</tr>
<tr>
<td>71650</td>
<td>Boston-Cambridge-Quincy, MA-NH</td>
<td>14460</td>
<td>Boston-Cambridge-Quincy, MA-NH</td>
<td></td>
</tr>
<tr>
<td>71950</td>
<td>Bridgeport-Stamford-Norwalk, CT</td>
<td>14860</td>
<td>Bridgeport-Stamford-Norwalk, CT</td>
<td></td>
</tr>
<tr>
<td>72400</td>
<td>Burlington-South Burlington, CT</td>
<td>15540</td>
<td>Burlington-South Burlington, CT</td>
<td></td>
</tr>
<tr>
<td>72850</td>
<td>Danbury, CT</td>
<td>14860</td>
<td>Bridgeport-Stamford-Norwalk, CT</td>
<td>Danbury, CT is in Fairfield County in CT which is in the Bridgeport CBSA</td>
</tr>
<tr>
<td>73450</td>
<td>Hartford-West Hartford-East Hartford, CT</td>
<td>25540</td>
<td>Hartford-West Hartford-East Hartford, CT</td>
<td></td>
</tr>
<tr>
<td>74500</td>
<td>Leominster-Fitchburg-Gardner, MA</td>
<td>49340</td>
<td>Worcester, MA-CT</td>
<td>All observation are in Worcester County, MA which is in the CBSA</td>
</tr>
<tr>
<td>74950</td>
<td>Manchester, NH</td>
<td>31700</td>
<td>Manchester-Nashua, NH</td>
<td></td>
</tr>
<tr>
<td>75550</td>
<td>New Bedford</td>
<td>39300</td>
<td>Providence-New Bedford-Fall River, RI-MA</td>
<td></td>
</tr>
<tr>
<td>75700</td>
<td>New Haven, CT</td>
<td>35300</td>
<td>New Haven-Milford, CT</td>
<td></td>
</tr>
<tr>
<td>76450</td>
<td>Norwich-New London, CT-RI</td>
<td>35980</td>
<td>Norwich-New London, CT</td>
<td>All observations are our sample are in CT</td>
</tr>
<tr>
<td>76750</td>
<td>Portland-South Portland, ME</td>
<td>38860</td>
<td>Portland-South Portland-Biddeford, ME</td>
<td></td>
</tr>
<tr>
<td>NECTA</td>
<td>NECTA Name</td>
<td>CBSA</td>
<td>CBSA Name</td>
<td>Notes</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------</td>
<td>------</td>
<td>------------------------------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>77200</td>
<td>Providence-Fall River-Warwick, RI-MA</td>
<td>39300</td>
<td>Providence-New Bedford-Fall River, RI-MA</td>
<td></td>
</tr>
<tr>
<td>77350</td>
<td>Rochester-Dover, NH-ME</td>
<td>14460</td>
<td>Boston-Cambridge-Quincy, MA-NH</td>
<td>The observations we have are in NH</td>
</tr>
<tr>
<td>78100</td>
<td>Springfield, MA-CT</td>
<td>44140</td>
<td>Springfield, MA</td>
<td>The observations we have are MA</td>
</tr>
<tr>
<td>78700</td>
<td>Waterbury, CT</td>
<td>35300</td>
<td>Waterbury, CT</td>
<td></td>
</tr>
<tr>
<td>79600</td>
<td>Worcester, MA-CT</td>
<td>49340</td>
<td>Worcester, MA-CT</td>
<td></td>
</tr>
</tbody>
</table>

Sources of information used for solving this problem included:

- Division codes to CBSA: https://www.census.gov/population/estimates/metro-city/0312msa.txt
- List of NECTA codes and names: https://www2.census.gov/programs-surveys/cps/methodology/Geographic%20Coding%20-%20Metro%20Areas%20(since%20August%202005).pdf
- Counties to MSA to CBSA crosswalk: http://www.nber.org/data/cbsa-msa-fips-ssa-county-crosswalk.html
- Larger list of NECTA codes and names: https://www.census.gov/population/estimates/metro-city/03nfips.txt

**D.2.b CBSA Division codes were incorrectly recorded as CBSA codes across multiple states.**

**Solution:** CBSA Division codes are mapped to a unique, larger CBSA, so the CBSA Division Codes are simply changed to the corresponding CBSA codes. This error is easily identified since all CBSA Division codes end with a "4" and all CBSA codes end with a "0."

**Changes Made**

<table>
<thead>
<tr>
<th>CBSA</th>
<th>State</th>
<th>CBSA Division Codes re-assigned</th>
</tr>
</thead>
<tbody>
<tr>
<td>14460</td>
<td>25</td>
<td>14484 15764 21604</td>
</tr>
<tr>
<td>14460</td>
<td>33</td>
<td>40484</td>
</tr>
<tr>
<td>16980</td>
<td>17</td>
<td>16974 29404</td>
</tr>
<tr>
<td>16980</td>
<td>18</td>
<td>23844</td>
</tr>
<tr>
<td>16980</td>
<td>55</td>
<td>29404</td>
</tr>
<tr>
<td>19100</td>
<td>48</td>
<td>19124 23104</td>
</tr>
<tr>
<td>19820</td>
<td>26</td>
<td>19804 47644</td>
</tr>
<tr>
<td>31100</td>
<td>6</td>
<td>31084 42044</td>
</tr>
<tr>
<td>33100</td>
<td>12</td>
<td>22744 33124</td>
</tr>
<tr>
<td>35620</td>
<td>34</td>
<td>20764 35644 35084</td>
</tr>
<tr>
<td>35620</td>
<td>36</td>
<td>35004 35644</td>
</tr>
<tr>
<td>37980</td>
<td>10</td>
<td>48864</td>
</tr>
<tr>
<td>37980</td>
<td>24</td>
<td>48864</td>
</tr>
<tr>
<td>37980</td>
<td>34</td>
<td>15804 48864</td>
</tr>
<tr>
<td>37980</td>
<td>42</td>
<td>37964</td>
</tr>
</tbody>
</table>
Sources of Information: crosswalk between CBSA and CBSA divisions:
https://www.census.gov/population/estimates/metro-city/0312msa.txt

D.2.c MSA codes were incorrectly recorded as CBSA codes in the years 2004 and 2005. These are easily identified because MSA codes are 4 digits and CBSA codes are 5 digits.

Solution: Replace missing MSA FIPS code with the code listed incorrectly in the CBSA code.

**MSA Codes that were incorrectly labeled CBSA codes**

<table>
<thead>
<tr>
<th>MSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0460</td>
</tr>
<tr>
<td>3000</td>
</tr>
<tr>
<td>3160</td>
</tr>
<tr>
<td>3720</td>
</tr>
</tbody>
</table>

D.3 Problem 3: Two CBSA’s had codes that changed over time.

Solution: Identified the correct CBSA change in these sources:

**Changes Made**

<table>
<thead>
<tr>
<th>CBSA</th>
<th>State</th>
<th>CBSA (old) re-assigned</th>
</tr>
</thead>
<tbody>
<tr>
<td>42060</td>
<td>6</td>
<td>42200</td>
</tr>
<tr>
<td>22520</td>
<td>1</td>
<td>22460</td>
</tr>
<tr>
<td>31100</td>
<td>6</td>
<td>31080</td>
</tr>
<tr>
<td>26180</td>
<td>15</td>
<td>46520</td>
</tr>
<tr>
<td>14060</td>
<td>17</td>
<td>14010</td>
</tr>
<tr>
<td>42260</td>
<td>12</td>
<td>35840</td>
</tr>
</tbody>
</table>

Sources of information:
https://www2.census.gov/programs-surveys/cps/methodology/Geographic%20Coding%20Metro%20Areas%20(since%20August%202005).pdf (for CBSA 22520) and http://mcdc.missouri.edu/data/georef/Tools/cbsa_changes.lst (for CBSA 3110).
http://mcdc.missouri.edu/data/georef/Tools/cbsa_changes.lst
https://www.uspto.gov/web/offices/ac/ido/oeip/taf/cls_cbsa/cbsa_countyassoc.htm