# Causes and Consequences of Central Neighborhood Change, 1970-2010*

Nathaniel Baum-Snow, University of Toronto
Daniel Hartley, Federal Reserve Bank of Chicago

May 16, 2016

**Abstract**

Neighborhoods within 2 km of most central business business districts of U.S. metropolitan areas experienced population declines from 1970 to 2000 but have rebounded markedly since 2000 at greater magnitudes than would be expected from simple mean reversion. Statistical decompositions reveal that 1980-2000 departures of all groups without a college degree generated most of the declines while the returns of college educated whites and the stability of neighborhood choices by lower socioeconomic status whites promoted most of the post-2000 rebounds. The rise of childless households and increases in college fraction in the population played important roles in promoting 1980-2010 central area population increases and in these areas' post-2000 gentrification. Estimates from a neighborhood choice model indicate that rising amenity values for educated whites and less rapidly deteriorating labor market opportunities near central business districts led to the 2000-2010 inflows of this group into central neighborhoods. Stabilization of central area employment declines after 2000 also slowed the outflows of most other groups, but were not enough to completely halt them in the face of their continued declines in valuations of local amenities in central neighborhoods.

*We thank Jason Bram, Ingrid Gould Ellen, Randall Walsh, Gilles Duranton and seminar and conference participants for their helpful comments. The views expressed are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Chicago, the Board of Governors of the Federal Reserve System, or its staff.

# 1 Introduction

In the decades following WWII, the central regions of most U.S. metropolitan areas were in decline. Between 1960 and 2000, the aggregate central city population share in the 100 largest metropolitan areas fell from 0.49 to 0.24 while the employment share declined from 0.61 to 0.34 (Baum-Snow, 2015). After 1980, however, the populations of many large cities began to stabilize and the socioeconomic characteristics of central areas of large cities began to rebound. Following sharp declines during the 1970s, neighborhoods within 2 km of central business districts (CBDs) in most medium and large U.S. cities experienced slow 1980-2000 declines and post-2000 growth in population, income, fraction white and fraction college graduate. Indeed, downtown neighborhoods have been the most rapidly gentrifying regions of metropolitan areas during the 2000-2010 period. This paper investigates the causes of central neighborhood 1980-2000 decline and 2000-2010 gentrification.

Our evaluation of the causes of neighborhood change proceeds in two stages. First, using a procedure akin to that proposed by Dinardo, Fortin & Lemieux (1996) for decomposing wage distributions, we systematically decompose the sources of changes in demand for central neighborhoods since 1980 into those due to demographic shifts (holding neighborhood choices constant) versus those due to changes in the neighborhood choices of demographic subgroups. These subgroups are determined based on census cross-tabulations of race by education, age, family type, or income. To better understand the component attributed to changes in choices, we use a conditional choice probability (CCP) procedure as in Hotz & Miller (1994) to recover valuations of each neighborhood in each decade 1980-2010 by narrowly defined demographic groups in the context of a standard neighborhood choice model. The model shows how to combine information about observed neighborhood choices and housing costs to recover neighborhood valuations that reflect a combination of sub-metropolitan area labor market opportunities and local amenities. We evaluate the extent to which shifts in local labor market and microgeographic labor demand conditions explain the increasing propensity of high socioeconomic status (SES) individuals and households to choose central neighborhoods and the declining propoensity of low SES individuals to do so.

Differential shifts in neighborhood choices by high versus low SES individuals have driven the majority of central neighborhood change. Declines in central neighborhood choice probabilities by low SES nonwhites over the full 1980-2010 period began to be offset by increases in such probabilities by high SES whites after 2000. 1980-2000 departures of low SES whites from central neighborhoods contributed to losses during this period, with their neighborhood choices stabilizing after 2000. Changing choices of high SES minorities had only small impacts. The 1980-2000 departures of low SES households from central neighborhoods promoted income growth and some racial change before 2000, despite declining population, which then accelerated after 2000.

Shifts in the racial composition of the population have consistently pushed in favor of downtown population growth, since central areas have a disproportionately high fraction of minorities.

However, racial shifts have pushed against other dimensions of gentrification that began after 1980. Shifts in the education distribution conditional on race have pushed in favor of gentrification but slightly against population growth. Shifts in the distribution of family types conditional on race have pushed in favor of population growth. Shifts in the income distribution and the age structure of the population conditional on race have had small effects. The broad conclusion is that while some of the increase in the educational composition of downtown residents is mechanical, most of the mechanical impact comes from declining white fraction, which if anything has pushed against gentrification. Indeed, since 2000 areas less than 2 km of CBDs is the only CBD distance range with increases in fraction white in the average CBSA.

Combining neighborhood choice probabilities and observations about home prices in the context of the neighborhood choice model reveals that white college graduates/high income white households are the only types with 2000-2010 increases in valuations of the combination of amenity values and labor market opportunities in central neighborhoods. Other whites' valuations stabilized after 2000 while other groups had smaller declines in central neighborhood valuations than in prior decades. We attribute these relative changes in central area neighborhood valuations to post-2000 stabilization of central area employment declines. 2000-2010 increases in central neighborhood valuations by educated whites indicates these neighborhoods' increasing amenity value for this group given no change in central area employment potential during this period.

Rising home prices, neighborhood choices probabilities and valuations amongst high SES whites of downtown neighborhoods can partly but not entirely be explained with shifts in the spatial structure of labor demand. College educated whites in CBSAs with the largest CBD-oriented labor demand shifts experienced the largest increases in valuations of downtown neighborhoods in every decade 1980-2010. As such, stabilization of job losses in downtown areas has driven a sizable fraction of the post-2000 shift in demand amongst high SES whites. However, exogenous positive metro area level labor demand shocks have pulled other demographic groups out of downtown areas. This could be because of they reflect improved labor market opportunities outside of downtown cores or because they reflect rising incomes throughout the CBSA, allowing for residents to seek out neighborhoods with higher amenity values.

Evidence from the choice model indicates that central area gentrification has imparted little harm on poor white residents on average, as despite rising home prices, their central neighborhood valuations have been dropping relative to other areas. However, we do find some weak evidence that less educated black residents have been pushed out by rising home prices, thereby reducing their welfare.

The remarkable post-2000 demographic change in central neighborhoods comes in the context of convergence in racial composition and income across neighborhoods in most CBSAs since 1980. However, consistent with evidence in Chetty et al. (2014) using individual level data, there exists considerable variation across CBSAs in the prevalence of such convergence. In particular, we provide evidence that more rapidly growing metropolitan areas experienced more rapid neighborhood convergence in the 1980s and 2000s and that these cities experienced greater flight of low SES households from downtown areas. However, cities with downtown employment mixes that were specialized in growing industries have experienced strong downtown residential demand growth by high SES whites in particular.

While our focus is on central neighborhoods, the methodology we develop in this paper can be used more broadly to understand neighborhood demographic dynamics. A better understanding of the drivers of neighborhood change may provide clues about some reasons for the growth in income inequality nationwide since 1980. Gould, Lavy & Paserman (2011) and Damm & Dustmann (forthcoming) provide independent evidence of the effects of neighborhood environments for youth on long-run outcomes. To the extent that neighborhood quality influences long run labor market outcomes, it is important to better isolate the mechanisms that have driven changes in neighborhood inequality. In particular, it will be important to understand the extent to which gentrifying neighborhoods retain incumbent residents (who can benefit from positive spillovers) or price them out. Existing evidence for census tracts with incomes that grew by at least $10,000 during the 1990s indicates that most incumbents are able to remain (McKinnish, Walsh & White, 2010). Our evidence is that this phenomenon is almost entirely driven by poor whites; blacks with less than college education experienced declining valuations of gentrifying central neighborhoods.

This paper proceeds as follows. Section 2 describes how we process the data and presents descriptive evidence on the changing fortunes of downtown areas and trends in neighborhood inequality. Section 3 evaluates potential explanations for changes in neighborhood inequality in a descriptive way and examines the extent to which changes in central neighborhoods' fortunes have been driven by these CBSA-level trends. Section 4 lays out a methodology for constructing counterfactual neighborhood compositions and presents decompositions of sources of neighborhood change using these counterfactuals. Section 5 develops a neighborhood choice model that is used to quantify the changes in neighborhood valuations by various demographic groups and evaluate the roles of local labor demand shocks. Finally, Section 6 concludes.

# 2 Characterizing Neighborhood Change

## 2.1 Data Construction

We primarily use 1970-2010 decennial census data and the 2008-2012 American Community Survey (ACS) data tabulated to 2000 definition census tract boundaries for this analysis. Central to our investigation is the need for joint distributions of population by race, education, household income, age and family structure across census tracts in each CBSA. To recover as many of these joint distributions in the most disaggregated form possible, we make use of both summary tape file

(STF) 3 and 4 census tabulations. We also use information about family structure and age by race from STF1 data from the 2010 census. Because the 2010 census did not collect information about income or education, we must rely in the 5 year ACS data for these tract distributions. We also make use of some census micro data to estimate parameters governing shapes of household income distributions above topcodes and to generate weights used to assign some in the tract aggregate data to different types of families. All census tracts are normalized to year 2000 geographies using census bureau reported allocation factors.

We construct three different joint distributions for people and one for households in 1980, 1990, 2000 and 2010. For each one, the race categories are white, black and other. In the other dimension of population joint distributions, we have 4 education groups (less than high school, high school only, some college, college +), 18 age groups (0-4, 5-9, ..., 80-84, 85+) or 6 family type groups (in married couple families with no kids, in married couple families with kids, in single female headed families with kids, in single male headed families with kids, not in a family, in group quarters). In the other dimension of the household joint distribution, we construct the number of households in each decile of the household income distribution of those in our sample area in that year. We do this in order to facilitate comparisons across CBSAs and years in a sensible way while taking into account the secular increase in nationwide income inequality during our sample period.

For the purpose of succinct descriptive analysis, we construct a summary measure of neighborhood change that incorporates fraction white, fraction college educated and median household income. This summary measure for tract $i$ is the average number of standard deviations tract $i$ is away from its mean in each year for each of these components. We call this equally weighted tract z-score the socioeconomic status (SES) index. For tract $i$ in CBSA $j$ in year $t$ and variables indexed by $k$ the SES index is calculated as

$$SES_{ijt} = \frac{1}{3} \sum_{k} \frac{y^k_{ijt} - \overline{y}^k_{jt}}{\sigma^k_{jt}},$$

where $\overline{y}^k_{jt}$ and $\sigma^k_{jt}$ are calculated with tract population or household weights. While we also experimented with using the first principal component of these same three underlying variables, we prefer the equally weighted z-score approach as it does not mechanically assign more weight to a variable only because it has more variation. We think that all three measures indexed by $k$ are roughly equally important indicators of neighborhood status.[1]

The Census Transportation Planning Package (CTPP) reports aggregated census or ACS micro data to microgeographic units for place of work in 1990, 2000 and 2005-2009. We use these data broken out by industry to construct localized labor demand shocks. Where available, we take CBD definitions from the 1982 Economic Census. Otherwise, we use the CBD location as assigned by

[1] There is evidence that conditional on income and education, black households have lower wealth than white households, meaning that fraction white is a proxy for unobserved elements of socioeconomic status.

ESRI. Each CBSA is assigned only one CBD.

Our sample includes regions of all year 2008 definition metropolitan areas (CBSAs) that were tracted in 1970 and had a population of at least 250,000 except Honolulu.[2] The result is a sample of 120 CBSAs. In order for our analysis to apply for the average metropolitan area rather than the average resident, much of the analysis weights tracts such that each CBSA is weighted equally.

The Data Appendix provides more details about data construction.

Figure 1a shows a map of the 120 CBSAs in our sample shaded by the fraction of census tracts within 4 km of the central business district that are in the top half of the tract distribution of our SES index in 1980 (top) and 2010 (bottom) in each CBSA. Those CBSAs above 0.5 have central areas that are less distressed than would be expected given random assignment of SES status to census tracts. Particularly striking is the number of CBSAs whose central areas experience gentrification between 1980 and 2010 (moving up the distribution of blue-green-yellow-red shades). Santa Barbara and New York are the only CBSAs with downtown areas that were more affluent than average in 1980. By 2010, 9 additional CBSAs had relatively affluent downtown areas. While central areas of other CBSAs remained less affluent than average, most became more affluent between 1980 and 2010. Of the 120 CBSAs in our sample, the fraction of the population within 4 km of a CBD living in a tract in the top half of the SES index distribution increased by more than 0.25 in 15 CBSAs, by 0.10 to 0.25 in 35 CBSAs and by 0.00 to 0.10 in 23 CBSAs between 1980 and 2010. Central areas of the remaining 47 CBSA experienced only small declines in their SES indexes on average. These patterns of changes are seen in Figure 1b, with red shaded CBSAs experiencing central area SES growth and blue shaded CBSAs with central area declines.

## 2.2 Facts About Neighborhood Change

Figure 2 reports statistics describing various aspects neighborhood change as functions of CBD distance since 1970. All plots show medians across CBSAs in our sample. We choose medians in order to emphasize that changes are not driven by just a few large notable cities. Analogous results using means across CBSAs or aggregates are similar. The broad message from Figure 2 is that downtown gentrification since 2000 is evident in many dimensions and is very localized. Neighborhoods within 2 km of CBDs exprienced the fastest 2000-2010 growth in terms of population, white fraction, college fraction and income of all CBD distance bands. The seeds of this gentrification started to form after 1980 with even more localized upticks in these indicators.

Panel A shows that the 1970s population declines in central neighborhoods reversed in the 1980s and 1990s, but only within 0.5 km of CBDs. After 2000, population growth rates within 1.5 km of CBDs were the fastest of any CBD distance band. Panel B shows a similar pattern for fraction white. Tracts within 3 km of CBDs experienced faster than average declines in fraction white during

[2] 100% of the 2000 definition tract must have been tracted in 1970 to be in our sample.

the 1970s, typical changes in fraction white during the 1980s, less rapid than average declines during the 1990s and rapid growth 2000-2010. Indeed, this is the only CBD distance band that experienced increases in fraction white after 2000, counteracting the 2000-2010 decline in fraction white in the population of about 5 percentage points. Evident in Panel B is an important racial component to downtown gentrification.

Panel C shows changes in the fraction of the population over 25 with a college degree. Consistent with Couture & Handbury's (2015) evidence, this graph shows modest relative declines in the 1970s, 1980s and 1990s and rapid growth in the 2000-2010 period within 4 km of CBDs. Once again, central neighborhoods were the most rapidly gentrifying in this dimension of any CBD distance ring. Couture & Handbury (2015) document that larger cities experienced more rapid growth in central area college fraction relative to their suburbs than did smaller cities but that even amongst smaller cities this 2000-2010 growth was greater than in the 1990s.[3] Figure 2 Panel D shows that mean income of residents in downtown neighborhoods rose faster than average starting in the 1990s out to about 6 km from CBDs, with less rapid additional growth in the 2000-2010 period, except immediately adjacent to CBDs.

Evidence in Figure 2 Panels A-D show that while some of the gentrification in central neighborhoods has to do with population growth, most of it has to do with shifts in the composition of a declining population. The formal decompositions performed in Section 4 below demonstrate that much of the 1980-2010 gentrification within 4 km of CBDs is explained by departures of lower SES individuals from central areas rather than arrivals of higher SES individuals.

While central neighborhoods have been gentrifying since 2000, their 2010 demographic composition remains of lower socioeconomic status than the suburbs. Of the three indicators in Figure 2 Panels B-D, the only one for which the central area looks like the suburbs is college fraction. White share and household incomes in central areas of cities remain well below those in the suburbs. This observation brings up the possibility that some of the patterns in Figure 2 can be attributed to mean reversion. Below we demonstrate that while neighborhoods do experience mean reversion, magnitudes of demographic change shown here are well beyond the typical amount experienced by central neighborhoods before 2000 and amongst other relatively low SES neighborhoods 2000-2010.

Figure 2 Panel E shows decadal changes in mean reported home value as functions of CBD distance. There are two reasons to look at home values. First, assuming housing supply is not perfectly elastic, changes in home values are indicators of changes in demand for neighborhoods. Outward neighborhood demand shifts associated with income growth can drive reduced population and higher housing prices as smaller homes are combined for households with greater housing demand. Second, home values are an input into welfare calculations, something which we develop further in Section 5. One of the mechanisms through which gentrification may make some worse

[3] Couture & Handbury's (2015) use CBD distance rings within which 5% of the CBSA populations live as their measure of downtown. Using 1970 data, this amounts to a median of 1.75 km and a range of 0.75 to 5 km. We found Figure 1 to be noisier when using such population percentiles instead as the x-axis variable.

off is through higher housing costs. It is important to additionally recognize that home values also capitalize changes in expected housing investment returns, which makes them more volatile than could be justified with fundamentals. The steep rise in home values during the 1980s between 0.5 and 2 km from CBDs may reflect rational expectations about future gentrification that eventually came to pass in the 2000-2010 period. The steep 2000-2010 rise in home values 2-10 km away from CBDs may reflect expectations about future gentrification in these neighborhoods.

The main mechanism that we explore as a potential driver of gentrification is shifts in the spatial structure of labor demand. To get a sense of how important this mechanism could be, Figure 2 Panel F shows employment growth as a function of CBD distance. It shows much more rapid employment growth in suburban areas during the 1990s but 2000-2010 employment growth that is essentially flat as a function of CBD distance. A very similar picture emerges for total payroll rather than employment counts. This look at the data indicates that employment growth may play a role, but is likely not the only driver of downtown gentrification. Indeed, Couture & Handbury's (2015) evidence of rising 2000-2010 rates of reverse commuting is evidence that some other forces like local amenities must also be important. Our systematic empirical investigation below confirms this claim. We do not have employment location information prior to 1990.

Table 1 reports transitions of individual census tracts through the distributions of the same three indicators in Panels B-D of Figure 2 plus the composite SES index. We present this evidence about the nature of demographic change in central neighborhoods to provide a sense of the heterogeneity around the summary statistics presented in Figure 2 and in order to show that a few neighborhoods moving quickly up the distribution are not driving central area gentrification. Table 1 shows the fraction of the population within 4 km of a typical CBSA's CBD living in tracts moving more than 20 percentile points or 0.5 standard deviations up or down the CBSA tract distribution. These numbers are calculated weighting by tract share of CBSA population in the base year, meaning all CBSAs are weighted equally.

Commensurate with evidence in Figure 2, three of our four measures indicate that central area tracts were on balance in decline during the 1970s. Results for the overall SES index in Panel D show that central neighborhoods' declines slowly reversed sometime in the 1980s or 1990s, when 2.8 percent of the central area population moved up at least 20 percentile points of the SES index distribution, relative to 1.9 percent in rapidly declining central tracts. Similarly, 4.6 percent of this population lived in tracts moving up at least 1/2 a standard deviation relative to 3.1 percent living in tracts moving down this much. This increase in the SES index of central tracts during the 1990s was mostly driven by income gains which had begun already in the 1980s. As in Figure 2, evidence in Table 1 shows that the resurgance of central areas really took off between 2000 and 2010. During this period, 7.9 percent of central area population lived in tracts moving up 20 percentile points in SES index distributions relative to only 1.1 percent living in tracts moving down in the typical CBSA.

# 3 Neighborhood Inequality

Downtown neighborhoods were the poorest and had among the lowest education levels and fraction white of any CBD distance ring in 1980. One potential explanation for their gentrification is thus simple mean reversion. In this section, we provide evidence that while mean reversion in neighborhood income and racial composition does exist, it is not the only force behind downtown revitalization. More broadly, we put the fortunes of downtown neighborhoods in the context of trends in overall neighborhood inequality.

We use our three demographic measures and the SES index to generate summary measures of changes in neighborhood inequality for each CBSA since 1980. The process for doing so resembles that in Chetty et al. (2014) but as applied to census tracts over time instead of parent-child pairs. In particular, we calculate correlations between CBSA demeaned outcomes between year $t$ and $t-10$ or 1980, applying tract population weights in the base year. Correlations of 1 indicate no change in neighborhood inequality on average while correlations of less than 1 indicate neighborhood convergence. Chetty et al. (2014) and Lee & Lin (2014) use percentile ranks in each year rather than demeaned outcomes as a basis for describing intergenerational mobility and neighborhood popoulation change respectively. However, our analysis benefits from distinguishing neighborhoods experiencing small changes from those experiencing large changes in their outcomes relative to CBSA means, even if they had the same changes in rank.

## 3.1 Chicago as an Example

Figure 3 depicts four measures of neighborhood change in the Chicago CBSA between 1980 and 2010, allowing for visualization of trends in neighborhood inequality. We calculate demeaned share white (Panel A), college graduate share (Panel B), log median household income (Panel C), and the SES index (Panel D) in each tract in 1980 and 2010, weighting by tract population. These demand indicators are graphed against each other in a scatterplot, with 45 degree and regression lines indicated. Both of these lines pass through (0,0) in each panel by construction. Dark black dots represent tracts within 4 km of the CBD. Regression slopes of less than 1, seen for log mean tract household income, tract share white and the composite SES index, indicate that Chicago neighborhoods have experienced convergence in these dimensions. The slopes of these regression lines are our 1980-2010 neighborhood change measures for Chicago. Points above a regression line that are far to the left of a 1980 mean represent gentrifying census tracts.

Figure 3 reveals considerable heterogeneity in 1980-2010 Chicago neighborhood change, with our three SES status measures clearly capturing distinct things. The masses of points at the bottom left and top right of Panel A represent large concentrations of stable minority and white census tracts respectively. The relatively large number of tracts along the right edge of the graph at almost 100 percent white in 1980 and ending up less than 70 percent white may have experienced tipping

(Card, Mas & Rothstein, 2008). But a handful of tracts went in the other direction between 1980 and 2010, seen in the upper left area of the graph. These largely minority tracts in 1980, that gained white share much faster than the typical Chicago tract, are almost exclusively within 4 km of the CBD. Indeed, all but 4 of the tracts within 4 km of the CBD that were less than 80 percent white in 1980 experienced increases in white share between 1980 and 2010, even though share white decreased on average. Such downtown area gentrification is clearly visible for the other measures as well in Figure 3, with central area tracts clustered in the upper left area of each panel.

Figure 4 contains analogous graphs depicting changes in Chicago tract SES indexes over each decade of our study period. It shows that Chicago experienced a small amount of neighborhood convergence in each decade 1970-2010. Dark black dots clustered on both sides of the regression line to the left of 0 in Panels A and B but only above the line in Panels C and D indicate that central area gentrification began during the 1990s in Chicago. We next document statistically that such patterns of neighborhood change near CBDs apply not just to Chicago, but are pervasive across medium and large metropolitan areas, and that poor tracts near CBDs began to turn around after 1990.

## 3.2 Quantifying Trends in Neighborhood Inequality

We now systematically characterize variation in neighborhood change across and within CBSAs and asses the extent to which this variation is explained by local labor market demand conditions. We apply the same logic discussed above for the Chicago example to each tract in our full sample. In particular, we decompose changes in tracts' outcomes of interest into a CBSA-specific component and a tract-specific component. We then investigate how local labor market conditions drove each of these components in turn.

We summarize a CBSA's neighborhood change over a decade with the coefficient from a tract-level regression of a demeaned outcome of interest in year $t$ on that for year $t-10$. Analogous regressions are estimated separately for each CBSA-decade combination for fraction white, fraction college, median household income and the SES index. The associated regression lines are drawn for Chicago in Figures 3 and 4. Most CBSA-decade combinations have slopes of less than 1, indicating neighborhood convergence on average. Mathematically, the neighborhood change index for outcome $y$ is represented by $\mu_j^y$ in the following regression equation:

$$y_{ijt} = \mu_{jt}^y y_{ijt-10} + u_{ijt}^y. \tag{1}$$

$\mu_{jt}^y < 1$ indicates neighborhood convergence on average whereas $\mu_{jt}^y > 1$ indicates divergence. By construction, both the implied regression line and the 45 degree line (indicating stability) pass through CBSA-specific means in years $t$ and $t-10$. In these regressions, each tract is weighted by 1970 population.

Table 2 presents information about distributions across CBSAs of $\mu^y_{jt}$ and how they differ as functions of local labor demand shocks. Each column for each period reports results from a separate CBSA level regression of $\mu^y_{jt}$ constructed using the outcome at top on a constant, various demeaned base year CBSA characteristics listed in the table notes and the demeaned change in CBSA log employment over the decade, expressed in standard deviation units. Coefficients on control variables are not statistically significant except in a few cases. Normalizations of control variables and $\Delta \ln(Employment)$ to be mean 0 allows for interpretation of the coefficient on the constant to be the average index of neighborhood convergence across CBSAs in the indicated decade.

We instrument for $\Delta \ln(Employment)$ with a Bartik (1991) type industry shift-share measure. This instrument is constructed by interacting the 1-digit industrial composition of employment in each CBSA in 1970 with national employment growth rates in each industry to generate a predicted change in employment for each CBSA.[4] The idea is to isolate demand shocks for living in a CBSA that are driven by national trends in industry composition rather than factors that could be correlated with unobservables driving neighborhood change. Decades in which $\Delta \ln(Employment)$ is not included in regressions exhibit insufficient first stage power.[5]

The first column of Table 2 shows that neighborhood racial composition has seen increased rates of convergence over time. The 1970s experienced diverging neighborhood racial composition. The average white share neighborhood change index was 0.99 in the 1980s and declined to 0.92 2000-2010. This evidence is roughly consistent with that in Cutler, Glaeser & Vigdor (1999) and Glaeser & Vigdor (2012) who document that racial segregation peaked in 1970 and has declined in every decade since. Negative coefficients on $\Delta \ln(Employment)$, all of which are statistically significant, indicate that more rapidly growing CBSAs experienced more rapid convergence in neighborhood racial compositions in the 1980s and 2000-2010.

In contrast to racial compositions, most CBSAs experienced neighborhood divergence in fraction college graduate on average after 1980 after the convergence associated with the rapid suburbanization of the 1970s. However, this divergence abated over the course of our sample period. Neighborhood convergence in fraction college was also somewhat stronger in more rapidly growing CBSAs. As with education, neighborhood income convergence was strongest in the 1970s. However, each decade 1980-2010 also experienced neighborhood income convergence. In the 1970s, the index of household income convergence was 0.79 on average across CBSAs, rising to 0.94 in the 2000-2010 period. But as with the other measures, more rapidly growing CBSAs experienced more rapid neighborhood income convergence. Taken together, the SES index results in the final column of Table 2 are roughly an average of the results for components in the other three columns.

[4] That is, we construct the Bartik instrument for CBSA $j$ that applies to the period $t-10$ to $t$ as: $Bartik_{jt} = \sum_k S_{jk1970} \ln(emp^{-j}_{kt}/emp^{-j}_{kt-10})$, where $S_{jk1970}$ is the fraction of employment in CBSA $j$ that is in industry $k$ at in 1970 and $emp^{-j}_{kt}$ is national employment in industry $k$ at time $t$ excluding CBSA $j$.

[5] We also experimented with using an alternative Bartik style instrument constructed using national trends in wages by industry. This instrument yields similar results.

The central conclusion from Table 2 is that greater CBSA employment growth promoted declines in neighborhood inequality and more associated gentrification off of a base of neighborhood convergence. Results in Section 5 below will show that some of this phenomenon is related to the decentralization of low SES groups into higher SES neighborhoods. The following sub-section systematically explores how central area neighborhoods fared in the context of these CBSA level trends in neighborhood inequality.

### 3.3 Spatial Distributions of Neighborhood Inequality

Using results in Table 3, we investigate sources of demand growth for central neighborhoods beyond the decadal CBSA-level trends in neighborhood inequality explored in Table 2. Table 3 presents results from a series of tract-level regressions described by the following equation:

$$S_{ijt} = \rho_{jt} + \mu_{jt} S_{ijt-10} + \sum{}_{d=1}^{4} \alpha_{dt} cbddis_{ij}^{d} + \alpha_{1t}^{b} cbddis_{ij}^{1} \Delta \ln Emp_{jt} + \alpha_{1t}^{s} cbddis_{ij}^{1} \Delta \ln CBDEmp_{jt}$$
$$+ \sum{}_{d=1}^{4} \beta_{dt} topdis_{ij}^{d} + \sum{}_{m} \delta_{mt} \ln(amendis_{ij}^{m}) + \varepsilon_{ijt} \quad (2)$$

These regressions are of the SES index on its lag interacted with CBSA fixed effects, within 4 km CBD distance ring indicators $cbddis_{ij}^{d}$ interacted with CBSA and CBD labor demand shifters described below, log distances to natural amenities $\ln(amendis_{ij}^{m})$ indexed by $m$ and indicators for 4 km distance bands from top quartile SES index tracts in 1970 $topdis_{ij}^{d}$, with each tract weighted by its 1970 CBSA population share. Because we are particularly interested in tracts that start off poor and because impacts of local labor demand shocks may differ by location in the SES distribution, we split the sample by tercile of each CBSA's 1970 SES index distribution across tracts, calculated with population weights. Panel A reports results for the bottom tercile, Panel B for the middle tercile and Panel C for the top tercile. Note that except for potential correlations between the lagged SES index and tract characteristics, this is equivalent to regressions of $u_{ij}$ from (1) on tract characteristics.

Table 3 reports estimates of $\alpha_1$, $\alpha_1^b$ and $\alpha_1^s$. $\alpha_1$ describes how much more or less gentrification occurred in tracts within 4 km of CBDs relative to what was typical among tracts with the same SES index in the indicated tercile on average. $\alpha_1^b$ describes how this gap differed for CBSAs with larger labor demand shocks, where $\Delta \ln Emp_{jt}$ is identical to the endogenous variable used in Table 2 for CBSA employment growth, instrumented with the same Bartik instrument. $\alpha_1^b$ describes how this gap differed for CBSAs with larger CBD oriented labor demand shocks. Both of these local labor demand shifters are standardized into separate z-scores. Because we do not observe the change in employment within 4 km of CBDs before 1990, we cannot use it as a regressor directly. For this reason, and to maintain consistency across the two Bartik demand shifters, we estimate the reduced form for the 1980-1990 and 1980-2010 periods instead of IV regressions. Therefore, for these periods magnitudes of $\alpha_1^b$ and $\alpha_1^s$ do not accurately capture effects of 1 standard deviation changes in CBSA and CBD oriented employment growth respectively. However, sign and significance of these coefficients remain informative.

The spatial Bartik variable $Spatbartik_{jt}$ is built as follows. For CBSA $j$, denote the fraction of employment near the CBD in industry $k$ in 1990 as $f_{jk}^{emp}$. We think of $f_{jk}^{emp}$ as being driven by the interaction of fundamental attributes of the production process like the importance of agglomeration spillovers to TFP. Therefore, we predict the change in the fraction of employment near the CBD to be

$$Spatbartik_{jt} = \sum_k f_{jk}^{emp} \ln(emp_{kt}^{-j}/emp_{kt-10}^{-j}).$$

Here, $emp_{kt}^{-j}$ denotes national employment in industry $k$ and year $t$ excluding CBSA $j$.

Control variables facilitate interpretation of $\alpha_1$, $\alpha_1^b$ and $\alpha_1^s$ as being determined by shifts in residential demand for areas within 4 km of CBDs while taking into account secular trends in neighborhood inequality and the locations of various amenities. Controls for CBSA fixed effects interacted with the lagged SES index removes CBSA trends in neighborhood inequality. The control for distance to natural amenities accounts for the possibility that CBDs are more likely to be located near such anchors of high income neighborhoods (Lee & Lin, 2014). The control for distance to top quartile tracts excludes the possibility that tracts near CBDs gentrified simply because of expansions of nearby high income neighborhoods (Guerrieri, Hartley & Hurst, 2013). Because we want to distinguish reasons for which poor tracts gentrify from reasons for which richer tracts change, we run these regressions separately for each 1970 defined CBSA tercile of the SES index, weighting by the tract's fraction of CBSA population.

Results in Table 3 demonstrate that the reversal of fortunes experienced by many central neighborhoods after 1980 is not simply an artifact of mean reversion. Coefficients in the top row of Panel A show that on average bottom tercile tracts near CBDs were significantly declining during the 1970s relative to bottom tercile tracts in other areas, with this decline abating during the 1980s and 1990s and reversing after 2000. Results in Panel B show a similar pattern of relative decline and reversal for middle tercile tracts. Panel C shows that top tercile tracts were on the decline in the 1970s, stable 1980-2000 and gentrifying in the 2000-2010 period. 1970s declines and 2000-2010 gentrification, while universal, were both most pronounced in bottom tercile locations.

Evidence in Table 3 indicates that bottom tercile neighborhoods near CBDs in CBSAs with above average central employment growth did significantly better in each decade 1970-2010 except the 1980s relative to their counterparts in average CBSAs. Central employment growth also counteracted declines of top tercile neighborhoods in most study periods. Growing overall CBSA employment had the largest positive effects on middle tercile neighborhoods in 1990-2000 and 2000-2010. Bottom tercile neighborhoods in the pooled 1980-2010 period are also estimated to have gentrified more quickly in CBSAs with positive exogenous labor demand shocks, though these shocks accelerated their decay in the 1970s. Table A1 reports summary statistics about these two types of shocks in each decade. Because 1990-2000 central area employment growth was negative, a simple calculation using coefficients in Panel A reveals that bottom tercile neighborhoods would have had approximately stable SES on average had downtown employment been stable during this period rather than declining by 7 percent. The reduction of central area employment declines to only 1 percent on average for the 2000-2010 period means that the average growth of 0.09 standard deviations of SES is consistent with increases in amenity values of these bottom tercile neighborhoods during this period.

To summarize, evidence in Table 3 indicates that while the bulk of 2000-2010 downtown gentrification was likely not driven by labor demand shocks, CBD-oriented labor demand shocks reinforced the downtown gentrification that occurred in many cities because of improvements in amenity values of downtown neighborhoods. CBD-oriented labor demand shocks primarily drives CBD-oriented residential demand growth through changes in the spatial distribution of wages net of commuting costs, though there may be a multiplier if downtown areas of cities are high amenity locations. The model in Section 5 clarifies this intuition.

Table A2 presents regressions analogous to those in Table 3, except that an index of tract housing value growth rates is used as the dependent variable. In particular, the dependent variable is calculated as the residuals from a regression of log mean tract housing value on various characteristics of owner occupied housing and CBSA fixed effects. Because positive demand shifts for neighborhoods will be reflected as some combination of increases in quantities of residents, potential income of residents and housing prices, we view evidence of house value growth and/or increases in the SES index for neighborhoods as signs of outward demand shifts. Indeed, CBSAs with high housing supply elasticities (Saiz, 2010) may have had some neighborhoods with large outward demand shifts that experienced only small relative changes in housing costs. However, because they have the smallest availability of developable land, central areas of cities are likely to have supply elasticities that are amongst the lowest of any neighborhood in any given CBSA.

Evidence in Table A2 largely follows that in Table 3, though with more noise and less dramatic reversals of declines. Coefficients on the Bartik interaction are not significant in any instance. However, coefficients on the spatial Bartik interaction are positive and significant for the bottom and top terciles in most decades of the study period. Overall, evidence in Table A2 is broadly consistent with the evidence in Table 3, that poor central neighorhoods have seen a resurgance, and especially those in CBSAs with CBD oriented employment growth.[6]

[6] Edlund et al. (2015) find that 26 large CBSAs with stronger skilled labor Bartik shocks experienced more rapid decadal central home price growth and demographic change in central areas than other areas of the city. These patterns are replicated in our data as well if census tracts are equally weighted.

# 4 Counterfactual Neighborhood Compositions

Results in the last section showed two important patterns in the data. First, central neighborhoods have been chosen at higher rates by higher SES demographic groups since 2000. Second, this gentrification has been more pronounced in low SES tracts in CBSAs with improving central area employment prospects and middle SES tracts in CBSAs with improving overall employment prospects. Thus far, our examination of location choices one demographic group at a time has limited our ability to determine the demographic characteristics driving downtown gentrification, especially since college education, high incomes and white fraction are all strongly positively correlated. In addition, the analysis to this point has not evaluated the extent to which demographic change toward more education, a more unequal income distribution and smaller families has accounted for gentrification. To separate out the relative importance of changing race-specific neighborhood choices from other observed demographic factors that may be correlated with race, we use tract level joint distributions of race and education or income over time to construct counterfactual neighborhood compositions absent changes in neighborhood choices for particular race-education and race-income combinations. The analysis simultaneously evaluates the extent to which population and SES growth in central neighborhoods are driven by shifts in the demographic compositions of CBSA populations.

To separate out the roles of CBSA-level demographic change from changes in individual groups' neighborhood demands, we carry out decompositions of the sources of neighborhood change along the lines developed by DiNardo, Fortin & Lemieux (1996) for decomposing wage distributions. To quantify the relative importance of changing neighborhood choices and demographic shifts for neighborhood change, we calculate magnitudes of central area population and demographic change under various counterfactual environments. First, we hold the fraction of CBSA population in various demographic groups fixed over time but allow neighborhood choices by specific groups to shift as in equilibrium one by one. This allows us to evaluate the extent to which changes in the choices of higher SES individuals and whites have driven central neighborhood change while holding the demographic composition of CBSA populations constant. We then additionally calculate how shifts in the CBSA level compositions of various demographic groups conditional on race have mechanically influenced neighborhood change, leaving CBSA level racial change as the residual component. This procedure has similarities to that developed in Carillo & Rothbaum (2016).

The results lain out in this section emphasize distinct forces driving central neighborhood change in the 1980-2000 and 2000-2010 periods. In the earlier period, central neighborhoods experienced flight of the poor, less educated and households with children. This was true for both white and minority households and was sizable enough to counteract a rising minority population, which mechanically increased the population of central area incumbent demographic groups. By 2000, the balance of power had shifted. The movement of higher SES whites into central neighborhoods strengthened as the outflow of lower SES whites ceased or reversed. Over the entire study period, the increasing college fraction in the population, especially among whites, has been important for driving composition shifts of downtown neighborhoods toward more white and educated.

## 4.1 Construction of Counterfactual Neighborhoods

### 4.1.1 Overview of Constructing Counterfactual Distributions

We observe the joint population distribution $f_{jt}(i, r, x)$ of race $r$ and other demographic attribute $x$ across census tracts $i$ in CBSA $j$ in year $t$. The attribute $x$ indexes education group, age group, family structure or household income decile in the national distribution. Given the structure of tabulated census data, we are forced to evaluate counterfactual joint distributions of race (white, black, and other) and only one other demographic attribute at a time across census tracts. Denote $N_{jt}$ as the total population of CBSA $j$ at time $t$ and CBSA density functions of demographics as $g_{jt}(r, x) = \sum_i f_{jt}(i, r, x)$. Crucially, we treat CBSA level allocations $g_{jt}(r, x)$ and populations $N_{jt}$ as exogenous to the allocation of people across neighborhoods, which can be justified in a long run open city model such as Ahlfeldt et al. (2015). Therefore, while aggregate population does not influence conclusions drawn from these mechanical counterfactuals, it will matter in principle when incorporating a consideration of housing costs.

Central to our recovery of counterfactuals is the following decomposition:

$$f_{jt}(i, r, x) = f_{jt}(i|r, x)g_{jt}(x|r)h_{jt}(r) \tag{3}$$

This expression shows how to separate out neighborhood choices of particular demographic groups $f_{jt}(i|r, x)$ from the CBSA level distribution of $(r, x)$ across locations. It additionally shows how to separate out shifts in education, age, income, or family type compositions independent of racial composition. Components of demographic change driven by changes in demand by group $(r, x)$ for tract $i$ are captured by shifts in $f_{jt}(i|r, x)$. Components driven by changes in the demographic makeup of whites, blacks or other minorites holding the racial distribution constant are captured by shifts in $g_{jt}(x|r)$. Components driven by changes in the racial composition of the population holding the demographic makeup of each race constant are captured by shifts in $h_{jt}(r)$. McKinnish, Walsh & White (2010) use a similar decomposition to examine the drivers of neighborhood income growth.

Tables 4-7 report results of counterfactual experiments, all with a similar structure. Table 4 uses counterfactual distributions to separate out mechanisms driving total central area population change. Tables 5 and 6 use counterfactual distributions to decompose sources of changes in central area fraction white and fraction college respectively. Table 7 decomposes changes in median income, expressed as percentiles of the household income distribution in sampled tracts. Table 4 examines

2 km CBD radii only and the other tables present results for both 2 and 4 km radii.[7] Panels A and B report results for 1980-2000 and 2000-2010 respectively. In Table 4, each row uses a different data set with joint distributions of race with education, age, family type and income respectively. Table 5 presents results using race-education and race-income joint distributions. Tables 6 and 7 use race-education and race-income distributions only respectively.

Column 1 in Tables 4-7 reports changes in outcomes of interest for central area geographies calculated using the raw data as a basis for comparison to counterfactuals. Because of sampling variability across the education, age and family type data sets and the use of households rather than people in the income data set, numbers in Column 1 of Tables 4 and 5 do not match perfectly across data sets. Column 2 shows the change that would have occurred had choices and shares not shifted from the base year. In Table 4, this is the CBSA population growth rate. Because objects of interest in Tables 5-7 are invariant to scale, Column 2 is all 0s in these tables.

Remaining columns of Tables 4-7 are built using counterfactual distributions. Our notation indicates column number superscripts on these probability distribution functions. Column 3 of Tables 4-7 reports counterfactual central neighborhood change given CBSA demographic shares that are unchanged from the base year. In particular, they are constructed using the counterfactual distributions

$$f^3_{jt}(i, r, x) = f_{jt}(i|r, x)g_{jb}(x|r)h_{jb}(r).$$

Here, demographic shares $g_{jb}(x|r)h_{jb}(r)$ are for the base year but neighborhood choices for each group $f_{jt}(i|r, x)$ change as they did in equilibrium. Results in Tables 4-7 Column 4 show the effects of holding choices constant but allowing demographic shares to shift as in equilibrium. The same statistics (5) and (6) are constructed using the counterfactual distribution

$$f^4_{jt}(i, r, x) = f_{jb}(i|r, x)g_{jt}(x|r)h_{jt}(r).$$

In most cases, results in Column 3 are closer to baselines in Column 1 than those in Column 4. This means that changes in neighborhood choices have been more important than changes in neighborhood shares for generating observed patterns in the data.

#### 4.1.2 Counterfactual Choices and Shares for Specific Demographic Groups

The remaining columns in Tables 4-7 decompose the difference between the actual changes in Column 1 and the counterfactuals given no changes in choices or shares in Column 2 into components that are related to changes in neighborhood choices (Columns 5-8) and demographic shares (Columns 9-10). The four effects in Columns 5-8 sum to the total effect of changing choices holding demographic shares constant reported in Column 3 relative to no changes reported in Column 2.

[7] Because 2000-2010 population growth was positive within 2 km of CBDs but negative within 4 km of CBDs, we focus on 2 km only for this outcome.

Adding the effects of changing demographic shares results in the full difference between the actual data in Column 1 and the "no changes" baseline in Column 2. That is, moving from left to right starting at Column 5 can be thought of as piling on additional components of demographic change from a baseline of no changes in Column 2 to full changes in Column 1.

Columns 5-8 report components of changes in equilibrium tract composition due to changing neighborhood choices of target whites, non-target whites, target non-whites and non-target non-whites respectively holding demographic shares at their base year levels. "Target" refers to college graduates, 20-34 year olds, single people and married couples without kids, or households in the top three deciles of the income distribution of the full sample area, depending on the data set used.

The set of results for counterfactual $c$ (5 to 8) is constructed using distributions built as

$$f_{jt}^{c}(i,r,x) = f_{jt}^{c}(i|r,x)g_{jb}(x|r)h_{jb}(r),$$

where $f_{jt}^{c}(i|r,x) = f_{jt}(i|r,x)$ for the elements of $(r,x)$ listed in column headers and $f_{jt}^{c}(i|r,x) = f_{jb}(i|r,x)$ for remaining elements of $(r,x)$. We note that the order of demographic groups for which we cumulatively impose year $t$ choices does not affect results. This is because the change in the fraction of the population in tract $i$ as a result of imposing any of these counterfactuals is linear. Each counterfactual amounts to imposing year $t$ rather than year $b$ choices for a few additional elements of $(x,r)$ at a time. Mathematically, the difference in the fraction of the population living in tract $i$ associated with counterfactual $c$ relative to c-1 is

$$\sum_{x}\sum_{r}[f_{jt}^{c}(i|r,x) - f_{jt}^{c-1}(i|r,x)]g_{jb}(x|r)h_{jb}(r). \tag{4}$$

Because of linearity within the square brackets, (4) indicates that the full choice adjustment counterfactual 3 can be achieved by imposing counterfactuals 5, 6, 7 and 8 cumulatively in any order. (4) also indicates that counterfactual $c$'s influence on tract composition depends not only on the magnitudes of differences in choices made by the group $(x,r)$ in question between $t$ and the base year $[f_{jt}^{c}(i|r,x) - f_{jb}(i|r,x)]$, but also by the fraction of that group in the CBSA population in the base year, $g_{jb}(x|r)h_{jb}(r)$. That is, neighborhoods change the same amount if a large group makes small changes in neighborhood choices or a small group makes large changes in neighborhood choices. To provide information about which one is driving results, Table 4 reports the average fraction of CBSA populations in parentheses for each of the four sets of demographic groups for which we examine the effects of changes in choices.

Having determined the roles of changes in neighborhood choices holding demographic composition constant, the remaining changes must be due to shifts in population composition. To look at this, we first maintain the base year racial distribution and examine how shifts in other demographic attributes conditional on race have influenced neighborhood choices. This allows us to see the influences that rising education levels, changes in income inequality, more single people, and the aging of the population have had on downtown neighborhood change while holding CBSA white, black and other population shares constant. Doing so avoids including the mechanical effects rising minority shares have on the education, age, family type and income distributions in these results. These results are reported in Column 9 of Tables 4-7, and are built using the expression

$$f_{jt}^{9}(i,r,x) = f_{jt}(i|r,x)g_{jt}(x|r)h_{jb}(r).$$

The residual effect (Column 10) is due to changes in racial composition, which typically works against gentrification since the white share of the population has declined over time.

Table A3 mathematically specifies construction of each counterfactual distribution and Table A4 reports average shares of target groups across CBSAs overall and within 2 km and 4 km CBD distance rings.

#### 4.1.3 Calculating Counterfactual Demographic Change

We use the distributions $f_{jt}^{c}(i,r,x)$ for each counterfactual $c$ and base year distributions $f_{jb}(i,r,x)$ to calculate counterfactuals of each measure of central neighborhood change discussed above.

We construct counterfactual population growth within 2 km of the CBD for Table 4 using the following expression:

$$\frac{1}{J}\sum_{j}\left(\ln\frac{N_{jt}}{N_{jb}} + \ln\frac{\sum_x\sum_r\sum_{i\subseteq CBD_j} f_{jt}^{c}(i,r,x)}{\sum_x\sum_r\sum_{i\subseteq CBD_j} f_{jb}(i,r,x)}\right) \tag{5}$$

That is, the central area population growth rate in a CBSA can be expressed as the sum of CBSA growth rate and the growth rate of the fraction of the population in the central area. The objects reported in Table 4 are averages across the 120 CBSAs in our sample, as is captured by the outer summation. The reference "no change" results in Column 2 of Table 4 are simply average CBSA population growth rates, calculated as $\frac{1}{J}\sum_{j}\ln(N_{jt}/N_{jb})$.

For Tables 5 and 6, we calculate changes in central area fraction white and fraction college respectively using the following expressions

$$\frac{1}{J}\sum_{j}\left(\frac{\sum_x\sum_{i\subseteq CBD_j} f_{jt}^{c}(i,r=w,x)}{\sum_x\sum_r\sum_{i\subseteq CBD_j} f_{jt}^{c}(i,r,x)} - \frac{\sum_x\sum_{i\subseteq CBD_j} f_{jb}(i,r=w,x)}{\sum_x\sum_r\sum_{i\subseteq CBD_j} f_{jb}(i,r,x)}\right) \tag{6}$$

$$\frac{1}{J}\sum_{j}\left(\frac{\sum_r\sum_{i\subseteq CBD_j} f_{jt}^{c}(i,r,x=\text{col})}{\sum_x\sum_r\sum_{i\subseteq CBD_j} f_{jt}^{c}(i,r,x)} - \frac{\sum_r\sum_{i\subseteq CBD_j} f_{jb}(i,r,x=\text{col})}{\sum_x\sum_r\sum_{i\subseteq CBD_j} f_{jb}(i,r,x)}\right). \tag{7}$$

In these tables, the reference change is 0 since there is no scale component like CBSA population growth provides for the counterfactuals in Table 4. In Table 5, $x$ indexes education composition or income decile as indicated in the row header. For Table 6, $x$ only indexes education composition.

For Table 7, we calculate counterfactual changes in central area median household income. We use median rather than mean income in order to be more robust to misallocating households to incorrect income deciles.[8] To see how this is built, begin with following expression for the cumulative distribution function of CBSA $j$'s central area households across income deciles $x \subseteq \{1, 2, ..., 10\}$ defined for the full national study area under counterfactual $c$ at time $t$:

$$G^c_{jt}(X) = \frac{\sum_{x \leq X} \left[ \sum_r \sum_{i \subseteq CBD_j} f^c_{jt}(i, r, x) \right]}{\sum_x \sum_r \sum_{i \subseteq CBD_j} f^c_{jt}(i, r, x)}.$$

Using these distributions over deciles $x$, we identify the deciles $D^c_{jt}$ that contain 0.5. We assign the median percentile assuming a uniform distribution of household income within $D^c_{jt}$. For example, if $G^c_{jt}(2) = 0.45$ and $G^c_{jt}(3) = 0.55$, $D^c_{jt} = 3$. In this case, we would assign the median household income $M^c_{jt}$ in CBSA $j$ at time $t$ under counterfactual $c$ to be 25, representing the $25^{th}$ percentile of the full study area's household income distribution. Then, the statistics reported in Table 7 are

$$\frac{1}{J} \sum_j \left( M^c_{jt} - M_{jb} \right). \tag{8}$$

As a result, positive numbers in Table 7 mean that the counterfactual in question pushed central area median incomes up by the indicated number of percentile points out of the national urban household income distribution.

Because choices and shares matter multiplicatively for the overall population distribution across tracts, the ordering of imposing year $t$ distributions matters for the influence of each channel. Tables A5 and A6 show results analogous to those in Tables 4-7 but impose the counterfactuals in the reverse order: shares adjustments first and sub-group specific choice adjustments second. This ordering does not materially affect the results.

## 4.2 Counterfactual Results

Before discussing the results of each counterfactual exercise, it is instructive to take a step back and summarize the broad picture provided by them. They all reflect a pattern of declining 1980-2000 central area population of all demographic groups except stability for some types of high SES whites. This trend continued after 2000 among minorities, though target whites had strong central area population growth and target nonwhites had essentially stable central area populations.

[8] Because cutoffs associated with each decile do not match the dollar cutoffs in the tract data, we assume uniform distributions within census data dollar bands for allocation purposes. The Data Appendix details our procedure for allocating households to income deciles.

### 4.2.1 Population

Table 4 shows what population growth 1980-2010 would have been within 2 km of CBDs under the various counterfactual scenarios laid out in the prior sub-section. Each row uses a different census tabulation that includes joint distributions of population by race and the $x$ characteristic indicated under "Data Set" across census tracts. Evidence in Column 1 reiterates the Figure 1 result that near CBD populations declined until 2000, after which they grew at about the same rate as overall urban population growth reported in Column 2. We do not report analogous results for within 4 km of CBDs because they are similar except for baseline population declines in both study sub-periods.

Results holding shares constant in Column 3 are slightly less than the actual changes in Column 1, meaning that shifting demographics pushed toward central area population growth since growing demographic groups were disproportionately located in downtown neighborhoods. We see below that in practice differences between actuals in Column 1 and results holding shares constant in Column 3 are mostly driven by increases in minority population shares. Had the race-education distribution not changed from 1980 to 2000, central area population would have declined by 12 percent rather than the actual decline of 7 percent in the average CBSA. In the 2000-2010 period, central area population would have grown by 4 percent rather than the 6 percent it actually grew. When using joint distributions of age, family type or income with race instead, changes in demographics are estimated to have bolstered central areas even more in both periods. As we discuss in more detail below, this is fully explained by variation in demographic changes in these non-racial dimensions.

Column 4 shows what would have happened to central area populations had neighborhood choices not changed from base years but demographic shares did. For 1980-2000, it shows over 30 percent growth for all data sets and for 2000-2010 it shows over 9 percent growth for all data sets. This reflects the positive effects associated with rising minority population reinforced by the imposed lack of shifts in neighborhood choices away from central neighborhoods.

Comparison of magnitudes of results in Columns 3 and 4 indicate that changing neighborhood choices have been central generators of 1980-2000 central area population decline, even as shifting demographics have pushed for population growth in central areas of cities. In the 2000-2010 period, shifts in neighborhood choices continued to hold central neighborhoods slightly below CBSA growth rates, with demographic changes almost making up for this deficit. Central areas' relatively high minority population shares and increasing minority populations overall have if anything pushed for more rapid population growth in central areas. Larger effects in Columns 3 and 4 for the family type data set reflect an increasing fraction of the population that is in childless households and the greater propensity for childless households to live near CBDs. Smaller effects for the education data set reflect the lower propensity of highly educated people to live near CBDs, especially in 1980.

Results in Columns 5-8 show the amount of population change due to changes in choices by each of the indicated demographic groups. "Target" groups are identified in the table notes, and are typically of higher socioeconomic status. In parentheses is the fraction of each demographic group in the CBSA population. These results show that 1980-2000 central area population losses are mostly explained by the flight of low SES whites and nonwhites alike, whose effects are similar at -0.14 and -0.18, respectively, for education and -0.24 and -0.21, respectively, for income. With non-target whites representing much larger shares of CBSA and central area populations, the logic discussed in the context of (4) indicates that changing choices of non-target nonwhites must have been of greater magnitudes. While all target groups of whites and nonwhites were also choosing to move away from central neighborhoods during 1980-2000 except young whites, the outflow was least pronounced amongst target whites.

In the 2000-2010 period, minority flight continued while white flight reversed. Non-target and target nonwhites departed central neighbohorhoods at similar rates as in 1980-2000, but all 4 groups of target whites examined started to return to central neighborhoods. For example, changing choices of college educated whites and high income white households accounted for 4 percent and 3 percent population and household growth respectively. Less educated and older whites were again also choosing central areas, but at lower rates than young or college educated whites. Young or college educated minorities were not returning to central neighborhoods like their white counterparts. This evidence of the return of the young college educated to downtown areas is in line with Couture and Handbury's (2015) similar evidence using different census tabulations.

Results in Table 4 Column 9 show how shifts in the composition of the demographic described by each data set influenced central area population share, holding racial composition constant. Positive percentages indicate a growing share of population subgroups that disproportionately chose to live in central area neighborhoods in the base year. The biggest standout in this regard is the fact that childless households were always most prevalent in downtown areas. Their growth as a fraction of the population contributed to a 10 percent increase in downtown populations during the 1980s and a 3 percent increase in the 2000-2010 period. In the other direction, the lower propensity of the educated to live near CBDs hurt these areas' populations. The 0 effect for income in Column 9 is mechanically due to our measurement of income as a percentile in the distribution of incomes in our sample in each year. Results in Table 4 Column 10 consistently show that the declining white fraction of the population promoted increases in downtown populations by 10 percent in 1980-2000 and 3 percent 2000-2010.

#### 4.2.2 Fraction White

Table 5 shows changes in counterfactual fraction white of central areas. We focus on education and income data sets and examine both 2 km and 4 km CBD distance radii. The baseline data in Column 1 shows that central neighborhood tracts within 2 km experienced about an 8 percentage point decline in fraction white between 1980 and 2000 and a 3 percentage point growth between 2000 and 2010. Up to 4 km of CBDs, there was a 9 percentage point decline and 1 percent growth in the two periods respectively. Because of secularly declining white population shares, the patterns in Column 1 are consistent with the 1980-2000 absolute declines in white fraction near CBDs seen in Figure 2.

For the 1980-2000 period, changes in demographic shares have driven secular declines in fraction white. This is seen from the fact that holding choices constant in Column 4 yields numbers similar to the data in Column 1 whereas holding shares constant in Column 3 actually yields a small amount of growth in fraction white. As we saw in Table 4, changes in neigjhborhood choices of nontarget and target whites are both large, but their opposite effects on racial composition approximately offset. 1980-2000 flight of all groups yields entries in Columns 5 and 7, for target and nontarget whites respectively, that are all negative and entries in Columns 6 and 8, for target and nontarget nonwhites respectively, that are all positive. The large changes in choices of low SES nonwhites is enough to change overwhelm the smaller shifts by low SES whites to yield the small net positive impact on fraction white of holding shares constant seen in Column 3. Changing education and income shares conditional on race had small effects. However, shifts in the racial composition caused fraction white in central neighborhoods to decline by about 10 percentage points, holding choices constant, similar to the actual declines in Column 1.

In the 2000-2010 period, increases in central area fraction white was mostly driven by changes in neighborhood choices of target whites. The cessation of departures of nontarget whites from central neighborhoods also contributed slightly positively to the racial turnaround of these areas. However, the continued departures of nontarget nonwhites from central areas at high rates represents the largest contribution. Results in Column 9 indicate that the growing educated population worked against 1980-2000 declines and toward 2000-2010 growth in central area white populations. Results in Column 10 indicate that reductions in the overall white share of the population over the entire sample period consistently pushed the central areas' fraction white downwards.

#### 4.2.3 Fraction College and Household Income

Table 6 examines reasons for changes in the propensities of college graduates to locate in downtown areas. Strong growth in fraction college in Column 1 of about 5-6 percentage points for both study periods reflects the rapid secular shift in the education distribution of the population. Normalizing growth in college fraction to be per decade makes the 2000-2010 growth about twice as fast relative to 1980-2000, reflecting the reversal of this demographic trend relative to other neighborhoods that is evident in Figure 2. The general pattern of impacts of changing shares and choices is similar to that for fraction white discussed above. Secular changes in college fraction primarily drove 1980-2000 changes while changing choices of target whites in particular were an important force influencing 2000-2010 growth in central area college fraction.

With non-college graduates moving out of central areas at slightly higher rates than others during 1980-2000, the net effect of shifts in neighborhood choices is very slightly positive, as is seen in Column 3. The demographic shifts toward a more educated population contributed to an increase of 6.4 percentage points in central area college fraction, with declines in the white population pushing in the other direction by 1.2 percentage points. Over the 2000-2010 period, the return of educated whites to central areas coupled with the continued departures of educated nonwhites became the additional important drivers of growth in central college fraction. Of the 2000-2010 6 percentage point growth in fraction college within 2 km of CBDs, about half is from secular demographic change and about half is from changes in choices. Of the changes in choices, about two-thirds (2.6 percentage points) is from changes in educated whites' neighborhood choices and about one-third (1.1 percentage points) is from such changes by lesser educated non-whites.

Finally, Table 7 examines reasons for changes in central area median household incomes expressed in percentile points of this distribution across all tracts in the study area. Results in Column 1 show that areas within 2 km of CBDs moved up the income distribution by about 1 percentile point 1980-2000 and by an additional 4 percentile points in the subsequent decade. Areas within 4 km of CBDs experienced small 1980-2000 income declines and 2 percentage point 2000-2010 gains. Comparison of results in Columns 3 and 4 reveals that changing choices were more important than changing shares in both periods, with changing choices pushing for greater income growth and changing shares pushing for declining incomes. As with education and race, the 1980-2000 increase in incomes is primarily driven by the departures of lower income whites and nonwhites alike. While these departures continued after 2000, the movement of high income whites into central neighborhoods bolstered central area income growth, especially within 2 km of CBDs. Given the increases in income inequality that occurred over the full study period, especially in larger cities (Baum-Snow & Pavan, 2013), this means that average incomes in city centers increased dramatically during the 2000-2010 period, as the rich were moving in and the poor were moving out. Shifts in racial composition was the main force bringing down central area incomes, at about half a percentage point for each decade 1980-2010.

# 5 Understanding Changes in Neighborhood Choices

The prior section performed an accounting of how much of demographic change in central neighborhoods has been driven by shifts in neighborhood choices by various demographic subgroups. In this section, we interpret this descriptive evidence in the context of a standard unified framework which delivers estimates of changes in neighborhood demand by location. This framework allows us to assess whether rising home prices or inward demand shifts are responsible for the flight of lower SES households from central neighborhoods. Moreover, it allows for recovery of the roles of CBSA and CBD oriented local labor demand shocks for driving these changes in demand for various demographic groups.

We lay out a standard neighborhood choice model that facilitates use of neighborhood choice shares in various demographic categories along with housing prices to recover information about changes in demand for neighborhoods over time. The procedure makes use of conditional choice probabilities, first formalized in Hotz & Miller (1994), in a way similar to Bayer et al's (2015) dynamic analysis of demand for neighborhood attributes. For clarity of exposition, we begin by thinking about the choice of neighborhood within one CBSA only. Couture & Handbury (2015) show that this is equivalent to considering a nested choice of first CBSA and then neighborhood within the chosen CBSA. Discrete household types are indexed by $h$ and there is a continuum of households of each type.

The indirect utility of household $r$ of type $h$ residing in census tract $i$ at time $t$ is

$$\widetilde{v}_{rhi}^t = v_h(p_i^t, w_{hi}^t, q_{hi}^t) + \varepsilon_{rhi}^t \equiv v_{hi}^t + \varepsilon_{rhi}^t.$$

In this expression, $p_i^t$ is the price of one unit of housing services in tract $i$, $w_{hi}^t$ is wage net of commuting cost, $q_{hi}^t$ summarizes local amenities as valued by type $h$ and $\varepsilon_{rhi}^t$ is an i.i.d. random utility shock distributed extreme value Type I. $q_{hi}^t$ may be a function of endogenous neighborhood attributes like the population composition itself. [9] $w_{hi}^t$ can depend on human capital characteristics and access to employment locations from tract $i$. We think of a long-run equilibrium in which moving costs are negligible. This setup delivers the following population shares of household type $h$ in each census tract $i$, which are observed in the data.

$$\pi_{hi}^t = \frac{\exp(v_{hi}^t)}{\sum_{i'} \exp(v_{hi'}^t)},$$

suggesting the relationship

$$\ln \pi_{hi}^t = v_{hi}^t - \ln\left(\sum_{i'} \exp(v_{hi'}^t)\right). \tag{9}$$

This equation shows that we can use conditional choice probabilities to recover the mean, median or modal utility associated with each tract up to a scale.[10]

We now consider the derivation of estimates of components of indirect utility that capture neighborhood attributes for a reference household type $\overline{h}$ and use it as a basis for recovering such components for other types. The broad goal here is to show how to control for differences in living costs across locations. Impose as a normalization that average modal utility across neighborhoods $\frac{1}{I}\sum_{i'} v_{\overline{h}i'}^t = 1$. This allows for inversion of (9) to an expression relating neighborhood choice

[9] The more standard way to model amenities would be to have $q_{hi} = q_i$, meaning that all household types care about the same bundle of amenities, but this unnecessary restriction on preferences makes it difficult for the model to match various patterns in the data.

[10] Given the extreme value assumption for the errors, the mean tract utility is $v_{hi}^t + 0.58$ given normalization of the scale parameter to 1, the median is $v_{hi}^t - \ln(\ln(2))$ and the mode is $v_{hi}^t$.

probabilities with indirect utility.

$$\ln \pi^t_{\overline{h}i} - \frac{1}{I}\sum_{i'}(\ln \pi^t_{\overline{h}i'}) + 1 = v_{\overline{h}}(p^t_i, w^t_{\overline{h}i}, q^t_{\overline{h}i})$$

Fully differentiating yields an expression that tells us that $\ln v_{\overline{h}i}$ equals a weighted average of wages net of commuting costs, home prices and neighborhood attributes relative to those in the average location. This expression assumes utility $U(x, H, q)$ takes the form $qu(x, H)$, where $u$ is homothetic.

$$\ln \pi^t_{\overline{h}i} - \frac{1}{I}\sum_{i'}\ln(\pi^t_{\overline{h}i'}) = d\ln w^t_{\overline{h}} - \beta_{\overline{h}} d\ln p^t_i + \sigma_{\overline{h}} dq^t_{\overline{h}i}$$

Here we are expressing utility as relative to the reference location, which has a utility normalized to 1. As in Rosen (1979) and Roback (1982), we see that differences in neighborhood choice probabilities reflect differences in incomes, housing costs and amenity values of locations. We can recover the combination of differences in wages net of commuting costs and local amenities across tracts for the average household type $\overline{h}$ by imposing $d\ln p_i = \ln p_i - \frac{1}{I}\sum_{i'}\ln p_{i'}$.

To recover analogous expressions for household types other than $\overline{h}$, differentiate indirect utility, holding location constant, to reveal $d\ln v = d\ln w$. Therefore, the reference utility level for households of type $h$ is $1 + \ln w_h - \ln w_{\overline{h}}$, where $w_h$ is the wage net of commuting cost for type $h$ in the reference (average) location. Therefore, for generic type $h$ we have

$$\ln \pi^t_{hi} - \frac{1}{I}\sum_{i'}\ln(\pi^t_{hi'}) - (\ln w^t_h - \ln w^t_{\overline{h}}) + \beta_h d\ln p^t_i = d\ln w^t_{hi} + \sigma_h dq^t_{hi} \equiv \lambda^t_{hi}. \tag{10}$$

This formulation takes into account the fact that richer households' marginal utilities of income are lower. The result is a greater discount on share differences across locations to reflect the fact that it is less onerous for higher income people to live in high cost relative to low cost areas, when compared against low income people.

(10) summarizes how to recover the component of differences in neighborhood demands that are driven by differences in wages net of commuting costs and neighborhood amenities. We directly observe $\pi^t_{hi}$ in the data as $f_{jt}(i|x, r)$ in the context of the counterfactual calculations of the prior section. 0 shares do not match the model well, so we assign tracts with 0 share to the smallest observed positive share for that demographic group for the purpose of calculating shares only. We set valuations of tracts with 0 shares to missing. To recover estimates of $d\ln p^t_i$, we take residuals from tract level regressions of log reported median home price on average home characteristics and CBSA fixed effects in each year. The Data Appendix provide further details about this calculation.

Based on evidence form the Consumer Expenditure Survey, we calibrate $\beta_h = 0.17$.[11][12] Remaining terms in (10) will get subsumed into CBSAXtime fixed effects in the empirical work described below.

Assuming the home price component of relative utilities $\beta d \ln p_i^t$ is the same across demographic groups, the model tells us that changes in neighborhood choice probabilities for a particular group must reflect some combination of changes in employment potential and amenity value of the neighborhood. Reintroducing the index $j$ for CBSAs, we decompose changes in neighborhood choice probabilities as follows from (10):

$$\pi_{hij}^t - \pi_{hij}^{t-10} \approx \pi_{hij}^{t-10}(\rho_{hj}^t + \Delta\lambda_{hij}^t - \beta\Delta d \ln p_{ij}^t). \tag{11}$$

In this expression, $\rho_{hj}$ is a type specific CBSA fixed effect. This expression shows that because all residents of the same neighborhood face the same home prices, variation in $\Delta\lambda$ across demographic groups is what generates differential changes in neighborhood choice probabilities relative to some CBSA baseline and a tract baseline driven by home price changes. (11) implicitly takes into account the fact that demand shifts by higher SES groups push up home prices, thereby dissuading lower SES groups from choosing these neighborhoods, even if their valuations have been rising too.

The following sub-section empirically examines variation in $\Delta_{hij}$ amongst demographic subgroups to recover an accounting for their shifts in neighborhood choices.

## 5.1 Using the Model

Figures 5 and 6 show levels of and changes in neighborhood valuations for white college graduates, black college graduates, white high school dropouts and black high school dropouts over the study period. Figure 5 shows that during the 1980-2000 period, central neighborhoods were most attractive for less educated blacks, educated blacks, less educated whites and educated whites respectively. This ordering is entirely driven by differences in relative neighborhood choice probabilities, since housing prices paid by each group are imposed to be identical. Figure 6 shows that central neighborhoods experienced declining attractiveness by all four of these groups in both the 1980s and the 1990s. Figure 5 Panel D shows that in 2010 white college graduates' valuation of neighborhoods adjacent to CBDs jumps dramatically relative to 2000, giving them valuations similar to college educated blacks.

[11] This number excludes utilities, whose costs should not differ across tracts within a CBSA. Limited demographic information in the Consumer Expenditure Survey indicates little variation in this expenditure share across demographic groups.

[12] A second approach is to instrument for price with spatially lagged price changes, as in Bayer, Ferreira & MicMillan (2007), or natural amenities, as in Couture & Handbury (2015). However, given the explicit linkages across local housing sub-markets through upward sloping housing supply and market clearing, the first approach may be problematic. Because natural amenities enter as part of the error term in $\lambda$, the second approach does not fit this context well.

We investigate the extent to which CBSA level and localized labor demand shocks have driven changes in $\lambda$ using regression equations similar to (2) group by group. We think of CBD-oriented labor demand shocks as influencing $d\ln w_{hi}^{t}$ and CBSA level labor demand shocks as potentially changing groups' demands for local amenities through an income effect. We report IV regression results from estimating the following equation for 1990-2000 and 2000-2010, since we observe the change in employment within 4 km of CBDs starting in 1990. For other time periods, we report the reduced form.

$$\Delta\widehat{\lambda}_{hij}^{t} = \rho_{hjt} + \sum{}_{d=1}^{4}\alpha_{hdt}cbddis_{ij}^{d} + \alpha_{1ht}^{b}cbddis_{ij}^{1}\Delta\ln Emp_{jt} + \alpha_{1ht}^{s}cbddis_{ij}^{1}\Delta\ln CBDEmp_{jt}$$
$$+\sum{}_{d=1}^{4}\beta_{hdt}topdis_{ij}^{d} + \sum{}_{m}\delta_{hmt}\ln(amendis_{ij}^{m}) + \varepsilon_{hijt}. \tag{12}$$

This estimation equation is the empirical analog to a differenced version of (10). $\rho_{hjt}$ accounts for the intercept $-\frac{1}{I}\sum_{i'}\ln(\pi_{hi'}^{t}) - (\ln w_{h}^{t} - \ln w_{\underline{h}}^{t})$ and the remaining terms allow us to measure variation tract labor market opportunities and local amenities relative to the average location. So that $\alpha_{h1t}$ can be interpreted as the average change in $\lambda$ for central area tracts for group $h$, we standardize $\Delta\ln Emp_{jt}$ and $\Delta\ln CBDEmp_{jt}$ to have means of 0 and standard deviations of 1. Tracts are weighted by their 1970 CBSA population share, so that each CBSA is weighted equally. Table A1 reports descriptive statistics about CBD area and CBSA employment changes and their instruments. Equation (11) indicates that comparisons of $\Delta\lambda_{hij}^{t}$ across demographic groups is what matters for understanding relative percent changes in neighborhood choices. This observation leads us to use the specification in (12) rather than a specification that controls for mean reversion. Note that measurement error will lead to more noise in neighborhood choice shares among smaller demographic groups, thereby inflating standard errors for these groups.

There are two potential concerns with using (12) to infer changes in neighborhood valuations. First is the issue of whether we have accurately measured housing costs. To get around this, instead of (12) one could estimate a unified equation for all household types simultaneously with typeXCBSA and tract fixed effects. Because the housing cost is common across types, the tract fixed effect would control for these costs assuming the housing expenditure share is the same for all types. The cost of this approach is that the absolute change in tract valuation is lost to a normalization, meaning that one can only recover relative changes in tract valuations across demographic groups. Our experimantation with such unified regression specifications yield very similar conclusions about relative changes in central area tract valuations across demographic groups to the results reported below.

A second concern is sample selection. Many tracts are dropped from the sample for small demographic groups because they have 0 choice shares for that group. The result is potential overestimation of demand for the types of neighborhoods these tracts are in. To address this concern, we built a version of the data in which we combine all tracts within 2 km CBD distance radii into single observations. Results using this aggregate data set are very similar to the results presented below.

Table 8 reports the coefficient estimates for select demographic groups defined by race and education. The dependent variable can be interpreted as the change in the percent difference in wages net of commuting costs plus amenity values associated with living in a tract relative to the average CBSA location. Coefficients in the first row of each panel describe average changes in valuations of central neighborhoods across CBSAs, with coeffiicients in remaining rows measuring the variation around these averages that are related to labor demand shocks. Significant negative coefficients are shaded blue and significant positive coefficients are shaded red.

Results in Panel A show that white college graduates had declining valuations of central neighborhoods on average until 2000, after which their valuations significantly rebounded. We evaluate the extent to which these averages are driven by shifts in localized labor market opportunities versus amenity values by considering what they would be given downtown employment growth of 0. Table A1 reports average central area employment declines of 7 percent and 1 percent in the 1990-2000 and 2000-2010 periods respectively, with 0 at 0.58 and 0.08 standard deviations above these means for the two periods. The significant coefficient of 0.26 on the downtown area employment interaction for the 1990s thus implies that a CBSA with no downtown employment change during this decade would have had almost no change in central area valuation. That is, the average reduction of valuation within 4 km of CBDs by college whites of 13 percent is entirely driven by reductions in nearby labor market opportunities rather than reductions in amenities. During the 2000-2010 period, the significantly positive coefficient on the <4 km CBD interaction of 0.09 gets boosted by an additional 0.01 if downtown employment growth is restricted to 0. This is evidence of improving amenity values of downtown neighborhoods for college educated whites. We also find some evidence that CBSA employment growth hurt educated whites' valuations of downtown neighborhoods in the 1980s and 1990s but not in the 2000-2010 period. This result is consistent with income growth driving residents out of central neighborhoods into higher amenity outlying neighborhoods (Margo, 1992).

Results for college educated blacks are reported in Panel B. This group's much greater declines in central neighborhood valuations than those for whites indicates their declining relative amenity values of central neighborhoods. Unstable coefficients on employment interactions in the 1990s reflect weak first stages, as seen in the low F-statistics listed at bottom. For the 2000-2010 period, the negative coefficient on CBSA employment growth and the positive coefficient on CBD area employment growth is consistent with declining amenities of downtown neighborhoods for college educated blacks outweighing the improved employment opportunities that arose in some CBSAs. Results in Panel D for high school dropout blacks are similar, though this group has significantly greater declines in amenity valuations of central neighborhoods.

Results for high school dropout whites are somewhere in between. This group had reduced declines in valuations of central neighborhoods in the 2000-2010 period relative to prior decades. The main 2000-2010 coefficient of -0.050 is significant but much smaller than that for high school dropout blacks of -0.209. High school dropout whites did not significantly benefit from improved CBD area labor market conditions, though the point estimate on this interaction coefficient is positive. As with blacks, better CBSA labor market conditions promoted declining valuations of central neighborhoods and outflows of this group to more suburban areas. Results for middle education whites and blacks not reported in Table 8 are in between the college graduate and high school dropout results for each race. Conditional on education, results for the "other" demographic group are between those for whites and blacks, though somewhat more similar to those for whites.

In Table 9, we repeat the same exercise using income deciles instead of education groups. So as to have a manageable table, we choose the 3rd, 6th and 8th as representative deciles. Patterns in Table 9 reiterate those in Table 8. CBD-oriented labor demand shocks disproportionately affected the 8th income decile. The background changes in central neighborhood valuations improved more for higher deciles than for lower deciles, but only turned positive for high income whites, not blacks. This is evidence that central neighborhoods have become magnets for high SES whites in particular because of high amenity values, with continued declining amenity values for blacks of all incomes. With a few exceptions, results for other deciles can be extrapolated from the results reported in Table 9.

# 6 Conclusions

Neighborhoods near central business districts of of U.S. metropolitan areas have experienced remarkable rebounds in population and especially socioeconomic status of their residents after 2000. Decompositions reveal that this turnaround in population has primarily been driven by the return of college educated and high income whites to these neighborhoods coupled with a halt in the outflows of other white demographic categories. At the same time, the departures of other socioeconomic groups continued unabated.

These changes in neighborhood choices by high socioeconomic status groups boosted the fortunes of central neighborhoods in the bottom tercile of the 1970 distribution of socioeconomic status in particular. During the 2000-2010 period, these neighborhoods moved up the SES distribution by a significant 0.12 standard deviations on average. Conditional on changes in CBSA level labor demand conditions, bottom tercile tracts in CBSAs with 12 percent more rapid CBD area employment growth experienced 0.2 greater standard deviations increases in socioeconomic status.

Statistical decompositions of the components of central area demographic change for the 1980-2000 and 2000-2010 periods show that shifts in neighborhood choices, or group-specific demands have been more important than demographic shifts for generating changes in the populations of these areas and compositions thereof.

Viewed in the context of a model of neighborhood choice, we find evidence that while only educated whites preceived 2000-2010 increases in the amenity values of downtown neighborhoods, all groups valued positive CBD-oriented labor demand shocks during this period. However, because 2000-2010 downtown employment growth was flat on average, the average CBSA only experienced boosts in downtown neighborhood demand by educated whites, with other groups continuing to exhibit declining demand. When compared with the downtown employment losses of prior decades, stabilization of downtown employment thus halted some of the prior declines of residents from these areas.

# A Data Appendix

Here we describe the construction of our sample and provide information about the sources of that we use to construct the sample. A large portion of the data used in our analysis come from tract-level tabulations from the decennial Census of Population from 1970, 1980, 1990, and 2000, and from the American Community Survey from 2008-2012. We use census tract boundaries from the 2000 Census of Population. We begin with the normalized data provided in Geolytics' 1970-2000 Neighborhood Change Database (NCDB) which provides a subset of the tract-level tabulation variables available from the 1970, 1980, 1990, and 2000 Censuses of Population normalized to year 2000 tract boundaries. We augment this data with other tract-level tabulations from these censuses that are not available in the NCDB and tract-level estimates from the 2008-2012 American Community Survey. In these cases, we perform normalizations to 2000 tract boundaries using the appropriate census tract relationship files provided by the Census Bureau.[13]

## A.1 Tract-level Sample

Our sample includes all of the 2008 definition Core Base Statistical Areas (CBSAs) that had a population of at least 250,000 in the area that was tracted in 1970 except Honolulu.[14] Our sample consists of 120 CBSAs and 39,087 year 2000 census tracts.[15] The CBSAs in the sample can be seen in Figure 1.

[13] See https://www.census.gov/geo/maps-data/data/relationship.html?? .

[14] Since we are using year 2000 tract boundaries, we limit our sample slightly further by using only tract for which 100\% of the 2000 definition tract was tracted in 1970.

[15] For CBSAs that are split into Metropolitan Divisions we treat each Division as a separate entity except in the following 4 cases in which we combine Metropolitan Divisions. These are: 1) Bethesda-Rockville-Frederick, MD is combined with Washington-Arlington-Alexandria, DC-VA-MD-WV. 2) Cambridge-Newton-Framingham, MA and Peabody, MA Metropolitan Divisions are combined with Boston-Quincy, MA. 3) Nassau-Suffolk, NY is combine with New York-White Plains-Wayne, NY-NJ. 4) Warren-Troy-Farmington Hills, MI is combined with Detroit-Livonia-Dearborn, MI.

### A.1.1 1970, 1990, 2000 Tract Data

These we take directly from the Neighborhood Change Database (NCDB) STF3A tabulations.

### A.1.2 1980 Tract Data

We read in these data from the summary tape file 4 files. This allows us to incorporate household income distributions by race and age by race into the data set. It also facilitates imposing various appropriate adjustments for suppression that are not handled well in the NCDB.

Suppression results in undercounting of whites and blacks in various tables. To handle this, we use tract level full population or household counts of whites, blacks and others to form inflation factors. We calculate inflation factors which scale up the total number of people in each age, education, family type or income bin in the STF4A data to equal the total reported in the NCDB data.

In particular, in the case of age, when the 1980 STF4A tract tabulations by race and age do not sum to the total population we implement the following algorithm:

1. inflate the total in each age bin so that the total of the age bins sums to the total population in the NCDB data.

2. calculate other race in each age bin by taking the total population in each age bin and subtract the white and black population of that age bin from the STF4A.

3. calculate the number of whites and blacks that are missing in the STF4A data by summing across the age bins for white and for black and subtracting the totals from the NCDB totals

4. calculate the number of people missing from each age bin by subtracting the STF4A total (that uses the recalculated other category) from the NCDB total

5. inflate the number of others in each age bin by the ratio of the NCDB other total to the STF4A other total

6. calculate the residual number of blacks and whites missing from each age bin by subtracting the inflated other from the inflated total for the age bin

7. reassign the residual number of blacks and whites missing from each age bin to either the white or black count in proportion to the share of the total missing that that white and black make up as calculated in 3.

We do the same process for education, and family type for 1980.

### A.1.3 2010 Census and ACS

We use the 2010 census summary tape file 1 for information about age and household structure by race. Because of the lack of a census long form in 2010, we are forced to use the American Community Survey to measure joint distributions of race by education and race by income.

## A.2 Central Business District Definitions

For each of our 120 CBSAs, we define the Central Business District (CBD) of the CBSA as that of the most populous Census place within the CBSA based on year 2000 population. We make two exceptions to this rule based our knowledge of the cities. For the Santa Barbara-Santa Maria-Goleta, CA Metropolitan Statistical Area we use the Santa Barbara CBD rather than the Santa Maria CBD even though Santa Maria was more populous in 2000 than Santa Barbara. For the Virginia Beach-Norfolk-Newport News, VA-NC Metropolitan Statistical Area we use the Norfolk CBD rather than the Virginia Beach CBD. For 113 of the our 120 CBSAs we were able to determine the CBD of the most populous city from the 1982 Census of Retail Trade. We use the latitude and longitude of the centroid of the tract or tracts specified as CBD tracts. For the remaining 7 CBSAs, we used the latitude and longitude where designated by ESRI.[16]

## A.3 Bartik Instrument Construction

We construct two Bartik instruments from several data sources. We label these instruments "Employment Bartik" and "Spatial Employment Bartik". The "Employment Bartik" attempts to predict CBSA-level employment growth for each of the 4 decades using initial year employment shares and decadal employment growth (implemented as changes in log employment levels) using 10 broad industry categories that can be consistently constructed from 1970 through 2010 using the county-level Census of Population and American Community Survey tabulations. The 10 industry categories are: 1) Agriculture, forestry, fisheries, and mining. 2) Construction. 3) Manufacturing. 4) Wholesale trade. 5) Retail trade. 6) Transportation, communication, other public utilities, and information. 7) Finance, insurance, and real estate. 8) Services. 9) Public administration. 10) Military. We refer to these as 1-digit industry categories.[17] This measure uses the exact geographical boundaries included in each of our CBSA definitions over the entire time period.

The aim of the "Spatial Employment Bartik" is to predict which CBSAs might be particularly impacted near the CBD by national industry growth. To construct this index, we calculate the share of employment located within 4 km of the CBD made up by each industry for each CBSA using the year 2000 Census Transportation Planning Package. We take these shares and interact them with the national industry growth rate of that industry to form a spatial or CBD-focussed Bartik instrument. Ideally, we would calculate the shares in each initial year, 1970, 1980, 1990, and 2000. However, the data are only available starting in 1990. Therefore, we use the 1990 1-digit industry distribution as the base.

[16] These 7 cities are Duluth, MN, Edison, NJ, Indianapolis, IN, Jacksonville, FL, Nashville, TN, and York, PA. Manual inspection of these 7 cities revealed CBD placement where we would expect it. Also, for the 113 cities where we have both Census of Retail Trade and ESRI CBD definitions the points line up closely.

[17] In practice, we do this once for each CBSA excluding that CBSA to calculate a national-level change that is not influenced by that particular CBSA.

## References

Ahlfeldt, Gabriel, Stephen Redding, Daniel Sturm, Wolf (2015) "The Economics of Density: Evidence from the Berlin Wall," Econometrica 83(6): 2127-2189.

Albouy, David and Bert Lue (2015) "Driving to Opportunity: Local Rents, Wages, Commuting and Sub-Metropolitan Quality of Life," Journal of Urban Economics 89:74-92.

Altonji, Joseph, Ulrich Doraszelski, and Lewis Segal (2000) "Black/White Differences in Wealth," Economic Perspectives, Vol. 24(1): 38–50.

Bajari, Patrick, and Matthew E. Kahn (2005) "Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities," Journal of Business and Economic Statistics 23(1):20–33.

Bartik, Timothy (1991) Who Benefits from State and Local Economic Development Policies? Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

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

Baum-Snow (2015) "Urban Transport Expansions, Employment Decentralization, and the Spatial Scope of Agglomeration Economies," manuscript.

Baum-Snow, Nathaniel and Byron Lutz (2011) "School Desegregation, School Choice and Changes in Residential Location Patterns by Race," American Economic Review, 101(7), 3019-3046.

Baum-Snow & Marion (2009) "The Effects of Low Income Housing Tax Credit Developments on Neighborhoods," Journal of Public Economics 93(5-6): 654-666.

Baum-Snow & Pavan (2013) "Inequality and city Size," Review of Economics and Statistics 95(5): 1535-1548.

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

Bayer, Patrick, Robert McMillan, Alvin Murphy, & Christopher Timmins (2015) "A Dynamic Model of Demand for Houses and Neighborhoods." Econometrica, forthcoming.

Berry, Steven, James Levinsohn and Ariel Pakes (1995) "Automobile Prices in Market Equilibrium," Econometrica, 63(4): 841-890.

Boustan (2010) "Was Postwar Suburbanization "White Flight"? Evidence from the Black Migration." Quarterly Journal of Economics 125.1: 417–43.

Busso, Matias, Jesse Gregory & Patrick Kline (2013) "Assessing the Incidence and Efficiency of a Prominent Place-Based Policy" American Economic Review 103(2): 897-947.

Card, David, Alex Mas & Jesse Rothstein (2008) "Tipping and the Dynamics of Segregation," Quarterly Journal of Economics, 123 (1): 177-218.

Carrillo, Paul & Jonathan Rothbaum (2016) "Counterfactual Spatial Distributions," Journal of Regional Science, forthcoming.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez (2014) "Where is the Land of Opportunity: The Geography of Intergenerational Mobility in the United States," Quarterly Journal of Economics 129: 1553-1623.

Couture, Victor and Jessie Handbury (2015) "Urban Revival in America, 2000 to 2010," manuscript.

Cutler, David, Edward Glaeser & Jacob Vigdor (1999) "The Rise and Decline of the American Ghetto," Journal of Political Economy 129: 1553-1623.

Damm, Anna Piil and Christian Dustmann (forthcoming) "Does Growing Up in a High Crime Neighborhood Affect Youth Criminal Behavior?" American Economic Review, 104(6): 1806-32.

Diamond, Rebecca (2015) "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," American Economic Review, forthcoming.

DiNardo, John, Nicole Fortin and Thomas Lemieux. 1996. "Labor Market Institutions, and the Distribution of Wages 1973-1992: A Semiparametric Approach," Econometrica, 64, 1001-1044.

Ding, Lei, Jackelyn Hwang and Eileen Divringi. 2015. "Gentrification and Residential Mobility in Philadelphia," manuscript.

Edlund, Lena, Cecilia Machado & Michaela Sviatchi (2015) "Bright Minds, Big Rent: Gentrification and the Rising Returns to Skill," NBER Working Paper #21729.

Galiani, Sebastian, Alvin Murphy and Juan Pantano (2012) "Estimating Neighborhood Choice Models: Lessons from a Housing Assistance Experiment." American Economic Review 105:11.

Glaeser, Edward and Jacob Vigdor (2012) "The End of the Segregated Century: Racial Separation in America's Neighborhoods, 1890-2010." Center for State and Local Leadership Civic Report #66, Manhattan Institute.

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

Gould, Eric D., Victor Lavy, and M. Daniele Paserman (2011) "Sixty Years after the Magic Carpet Ride: The Long-Run Effect of the Early Childhood Environment on Social and Economic Outcomes," Review of Economic Studies, 78, 938–973.

Hotz, Joseph & Miller (1994) "Conditional Choice Probabilities and the Estimation of Dynamic Models," The Review of Economic Studies 60(3), 497-529.

Kain, John F. (1992) "The Spatial Mismatch Hypothesis: Three Decades Later," Housing Policy Debate, 3:2, 371-462.

Lee, Sanghoon & Jeffrey Lin (2014) "Natural Amenities, Neighborhood Income Dynamics, and Persistence in the Spatial Income Distribution," manuscript.

Mankiw, N. Gregory & David Weil (1989) "The Baby Boom, the Baby Bust, and the Housing Market" Regional Science and Urban Economics 19.

Margo, Robert (1992) "Explaining the postwar suburbanization of population in the United States: The role of income," Journal of Urban Economics 31(3): 301-310.

McKinnish, Terra, Randall Walsh & T. Kirk White (2010) "Who Gentrifies Low-Income Neighborhoods?" Journal of Urban Economics 67: 180-193.

Neal, Derek & William Johnson (1996) "The Role of Premarket Factors in Black-White Wage Differences," Journal of Political Economy, 104(5): 869-895.

Saiz, Albert (2010) “The Geographic Determinants of Housing Supply” Quarterly Journal of Economics, 125(3):1253-1296.

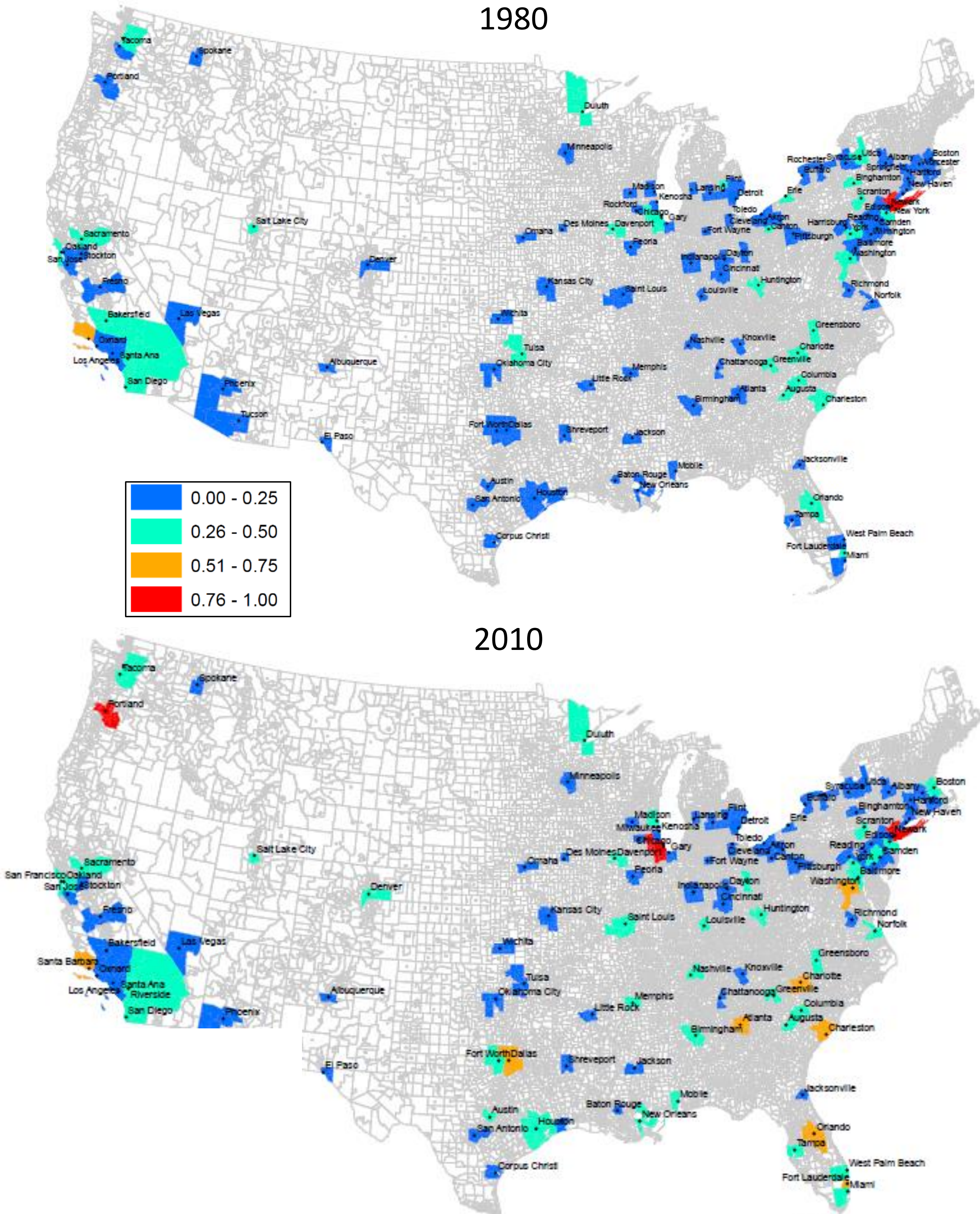



Figure 1a: Share of Residents Within 4 km of the CBD Living in a Top Half SES Distribution Census Tract

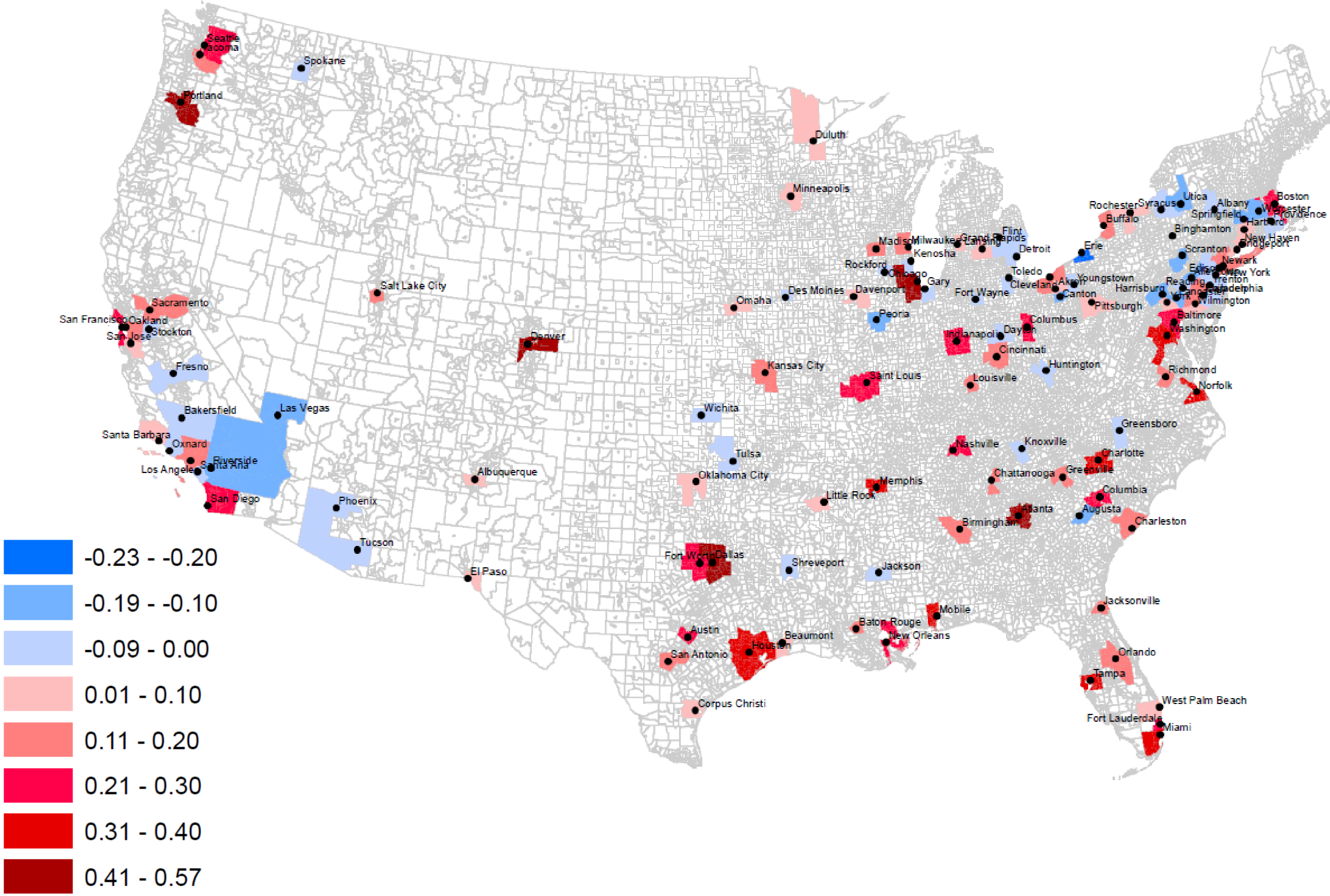
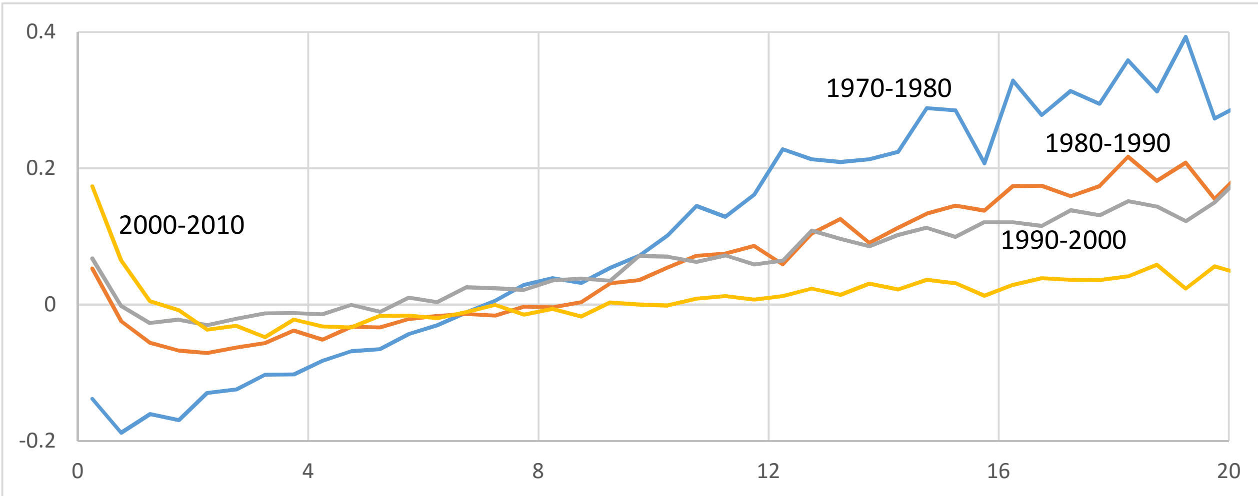



Figure 1b: 1980-2010 Change in Share of Residents Within 4 km of the CBD Living in a Top Half SES Distribution Census Tract

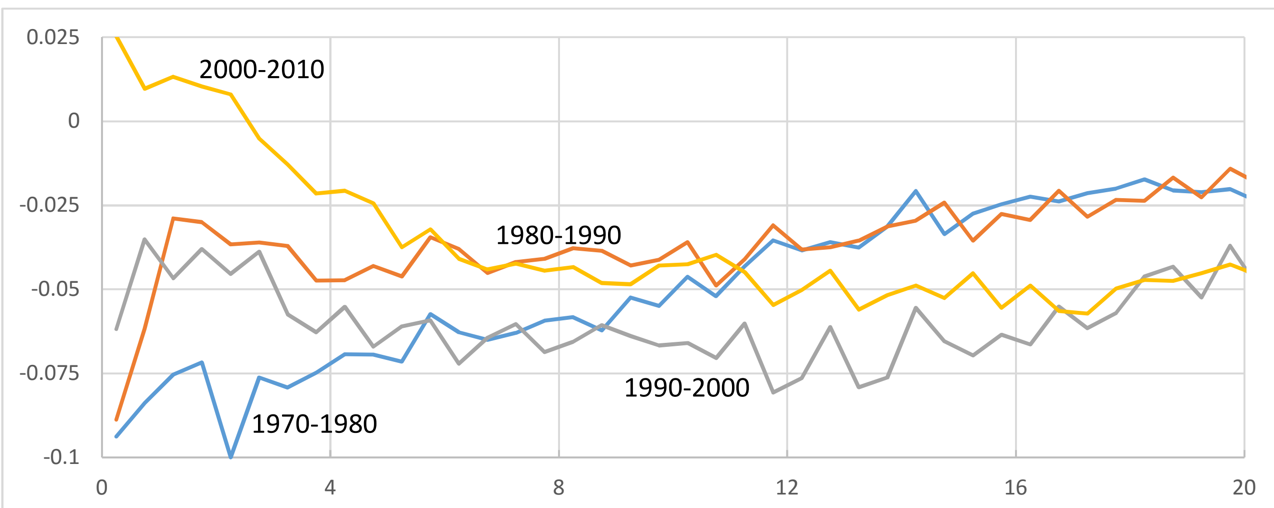
**Figure 2: Measures of Gentrification as a Function of CBD Distance (km)**

Medians Across 120 CBSAs, 0.5 km CBD Distance Bands

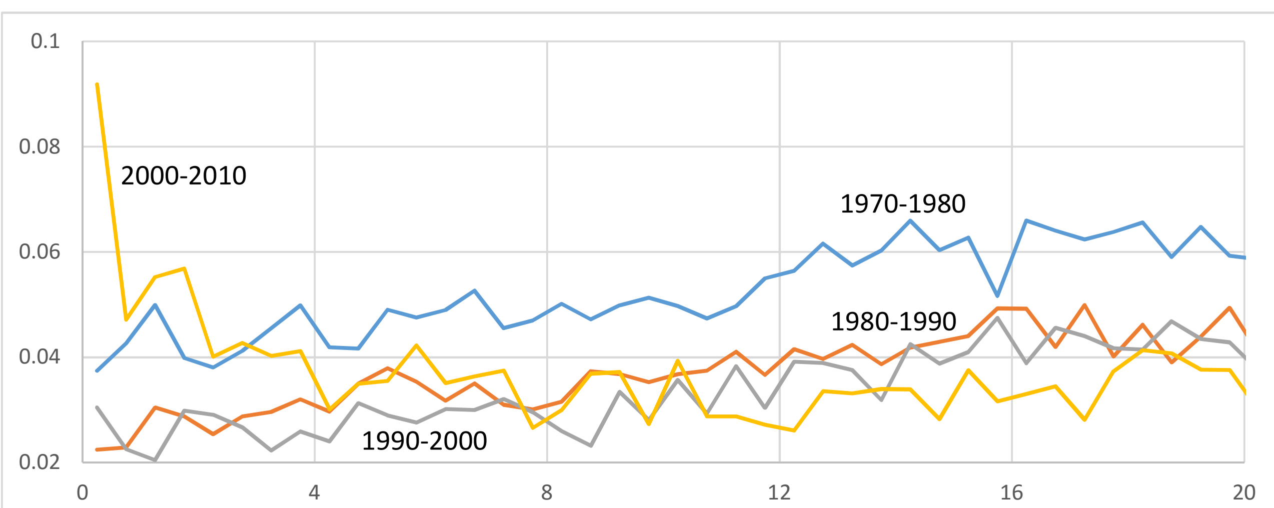
**Panel A: Percent Change in Population**



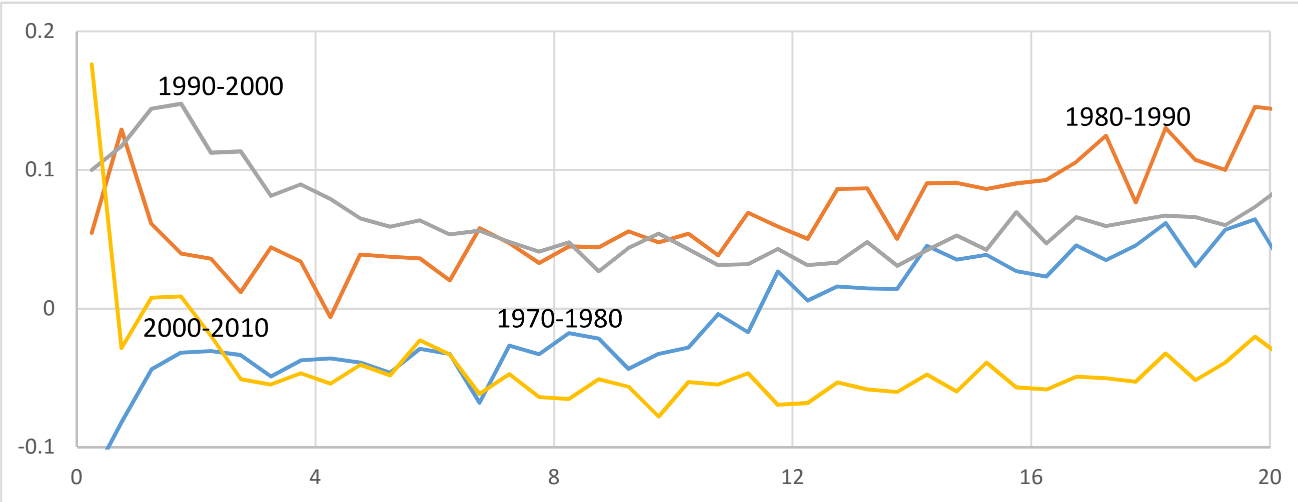

**Panel B: Change in Fraction White**



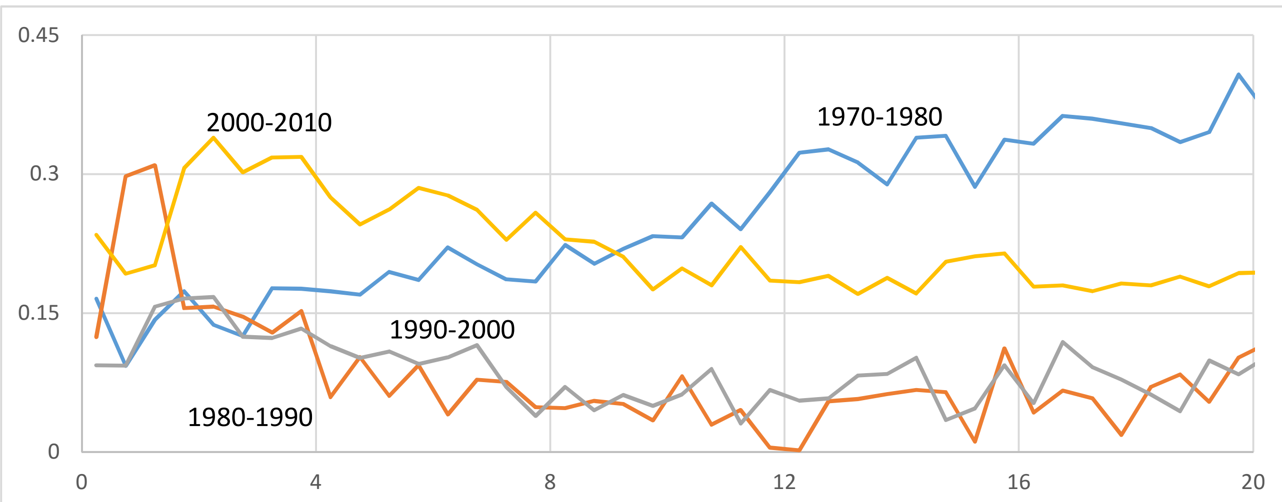

**Panel C: Change in Fraction 25+ with College Education**



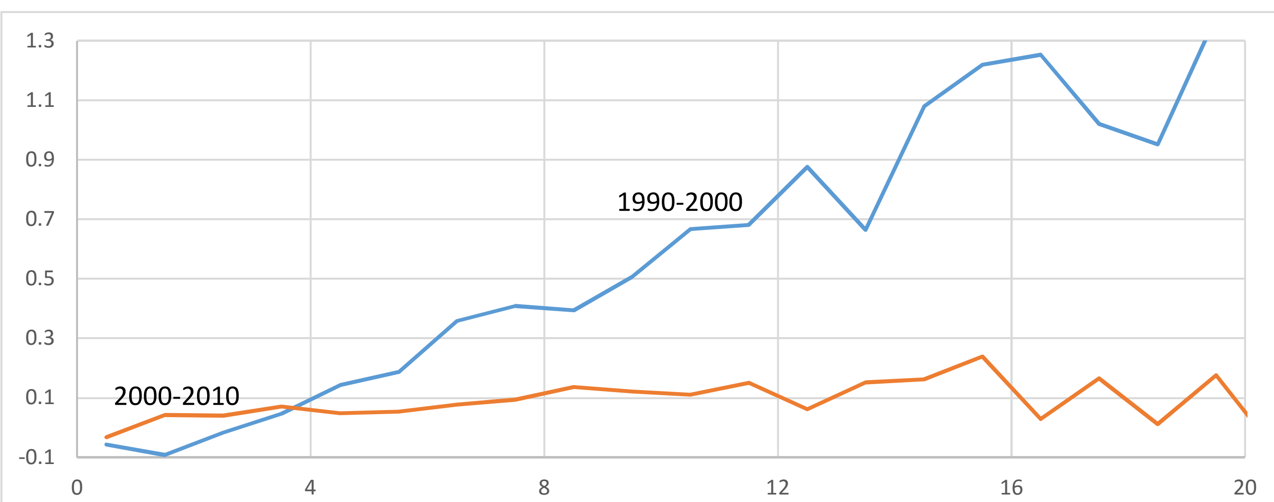
**Panel D: Percent Change in Mean HH Income (2010 $)**



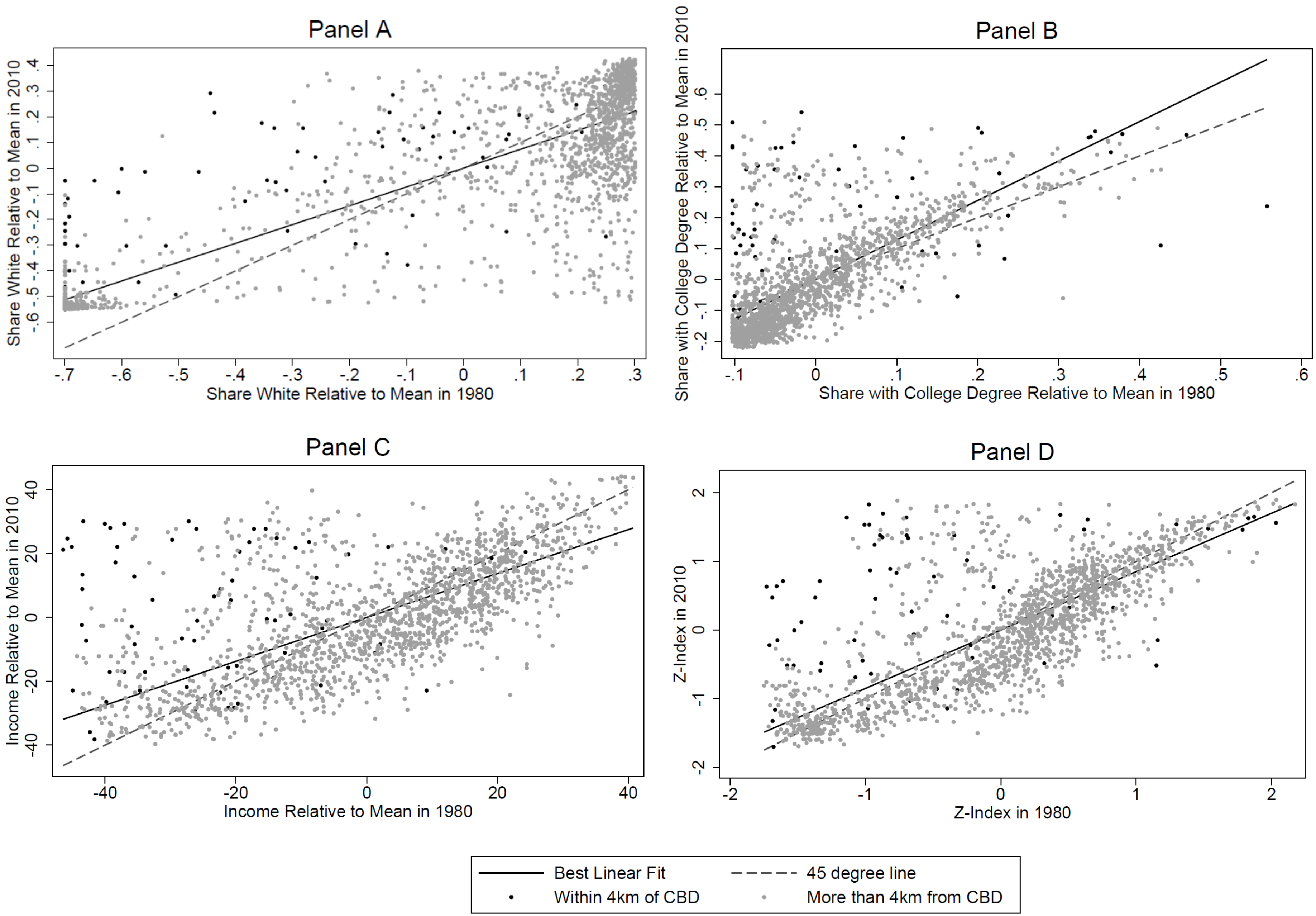

**Panel E: Percent Change in Mean Housing Value (2010 $)**




**Panel F: Median Change in Employment**



## Figure 3: 1980-2010 Neighborhood Change in Chicago



# Figure 4: Decadal Tract Changes in SES Index, Chicago

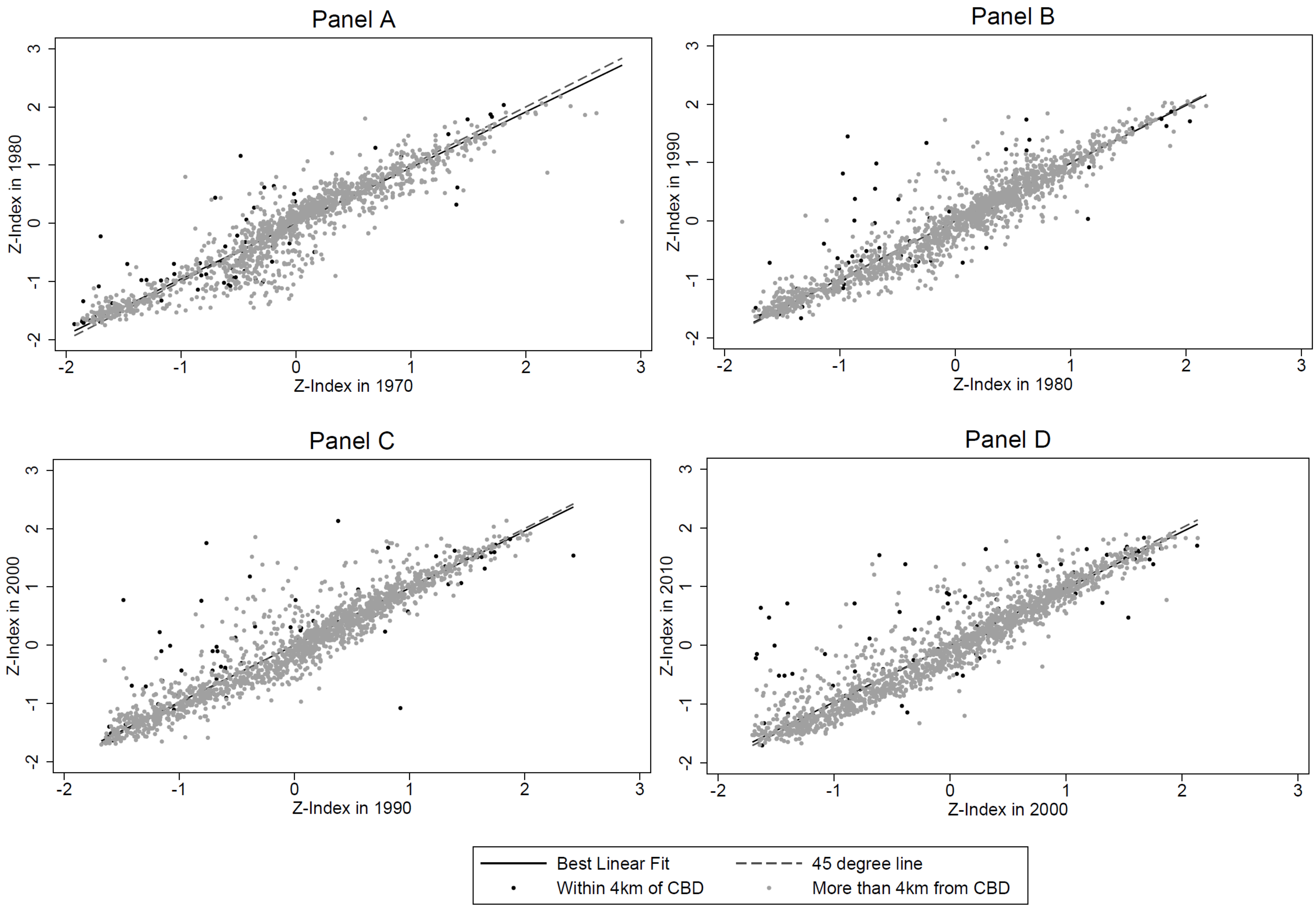
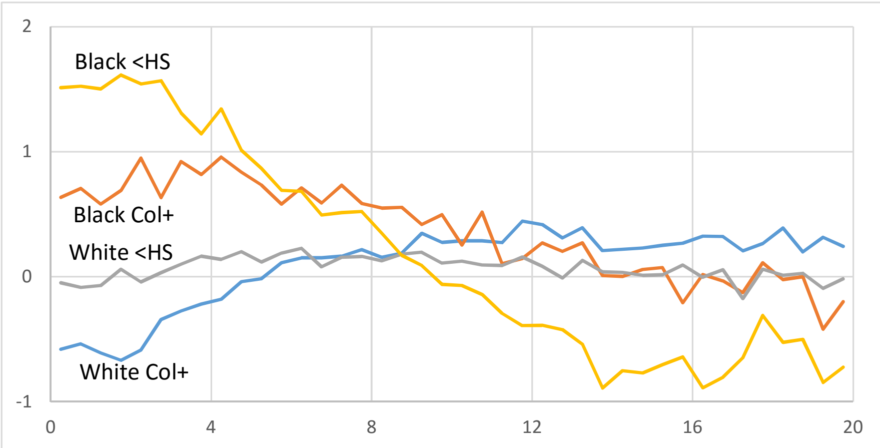
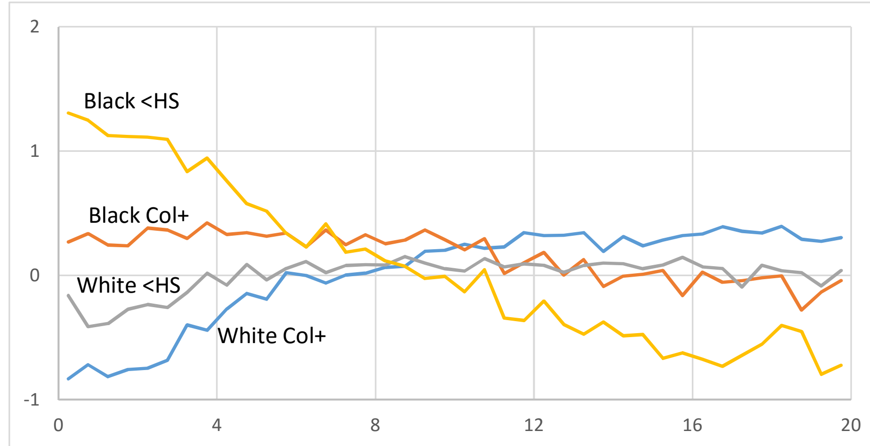


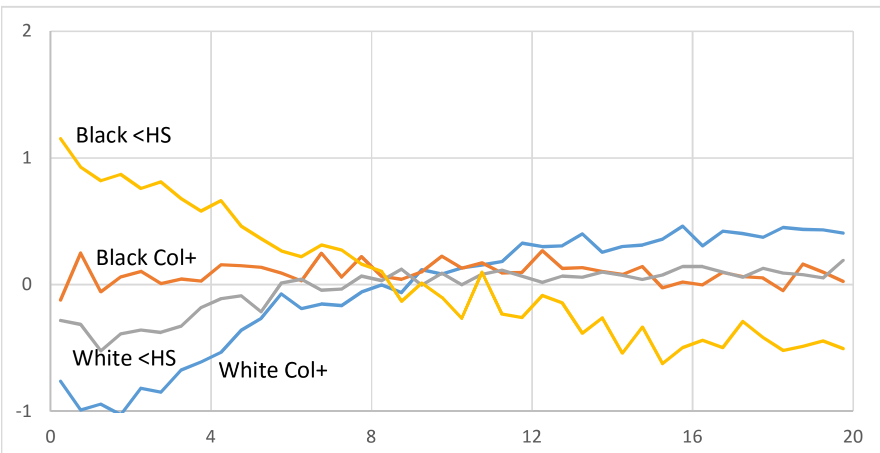
**Figure 5: Lambdas in Each Year by Education and Race**



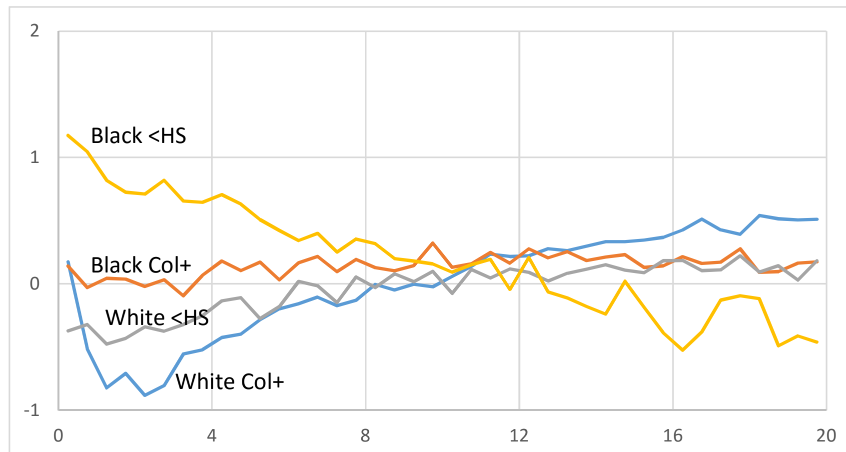

**Panel A: 1980**




**Panel B: 1990**

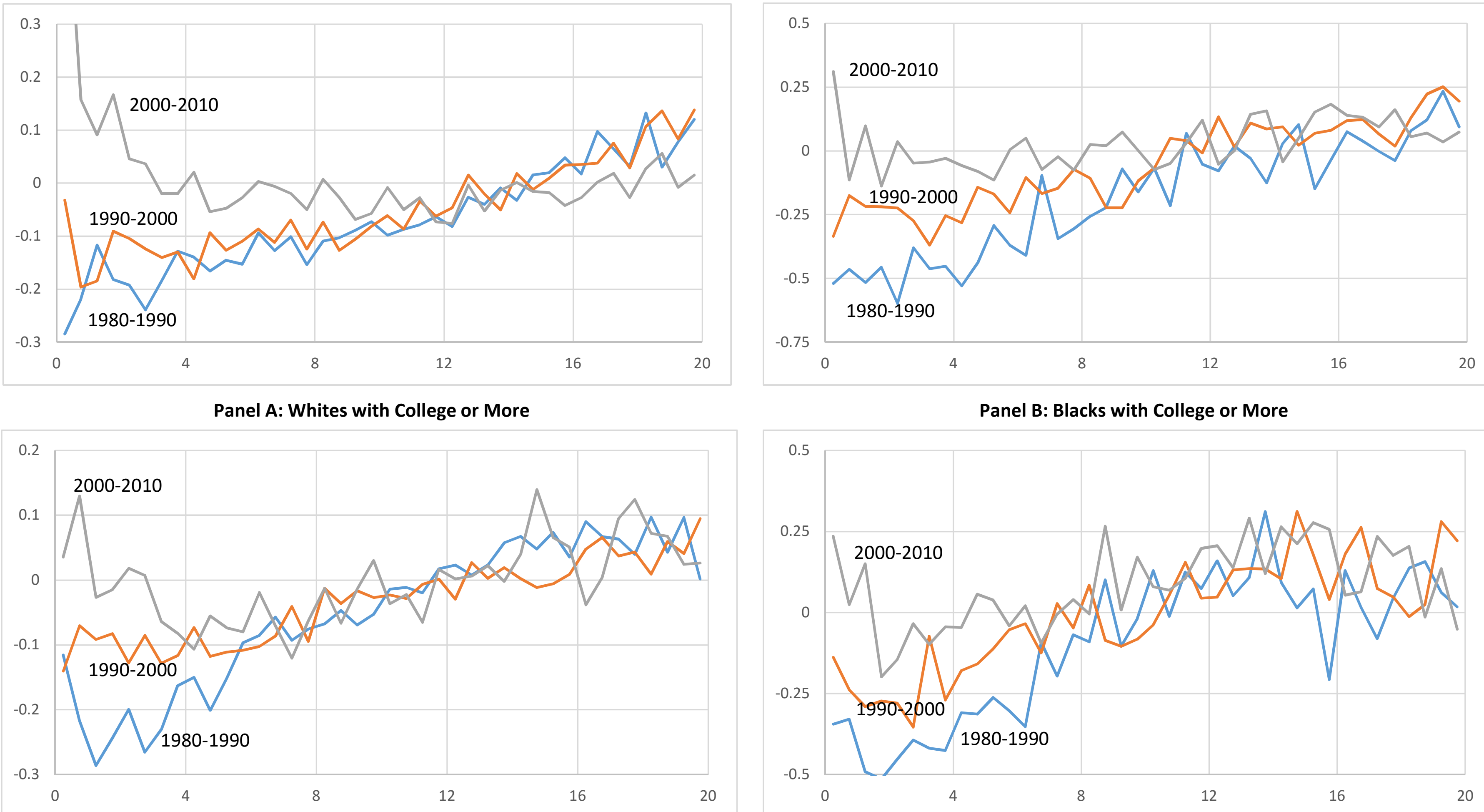



**Panel C: 2000**




**Panel B: 2010**

**Figure 6: Changes in Neighborhood Valuation as a function of CBD Distance by Race and Education**




**Panel A: Whites with College or More**

**Panel B: Blacks with College or More**

**Panel C: White High School Dropouts**

**Panel D: Black High School Dropouts**

**Table 1: Share of Population within 4km of CBD in Tract Changing by at Least**

| | 20 Percentile Points | | 1/2 Standard Deviation | |
|---|---|---|---|---|
| | up | down | up | down |
| **Panel A: Fraction White** | | | | |
| 1970-1980 | 6.5% | 13.3% | 14.5% | 20.8% |
| 1980-1990 | 4.4% | 6.0% | 8.1% | 13.9% |
| 1990-2000 | 4.0% | 3.1% | 12.1% | 11.0% |
| 2000-2010 | 5.2% | 1.3% | 14.2% | 5.5% |
| 1980-2010 | 5.3% | 1.3% | 34.8% | 23.2% |
| **Panel B: Fraction College Educated** | | | | |
| 1970-1980 | 10.3% | 10.0% | 14.7% | 7.6% |
| 1980-1990 | 5.2% | 5.8% | 6.0% | 7.5% |
| 1990-2000 | 3.8% | 6.1% | 5.5% | 7.6% |
| 2000-2010 | 10.3% | 4.0% | 14.4% | 5.3% |
| 1980-2010 | 10.8% | 4.0% | 18.8% | 16.6% |
| **Panel C: Median Income** | | | | |
| 1970-1980 | 0.7% | 11.9% | 3.3% | 21.3% |
| 1980-1990 | 3.5% | 1.1% | 7.8% | 3.3% |
| 1990-2000 | 3.3% | 1.4% | 7.7% | 2.9% |
| 2000-2010 | 8.2% | 1.4% | 14.6% | 4.4% |
| 1980-2010 | 8.1% | 1.3% | 30.7% | 8.9% |
| **Panel D: SES Index** | | | | |
| 1970-1980 | 2.6% | 7.7% | 4.6% | 12.5% |
| 1980-1990 | 2.4% | 1.9% | 3.8% | 3.2% |
| 1990-2000 | 2.8% | 1.9% | 4.6% | 3.1% |
| 2000-2010 | 7.9% | 1.2% | 10.8% | 1.6% |
| 1980-2010 | 7.9% | 1.1% | 24.5% | 13.1% |

Notes: Distributions are within each of the 120 large CBSAs in our sample. Each tract is weighted by its share of CBSA population.

**Table 2: CBSA Demand Shifts and Neighborhood Inequality**

| | | Inequality Criterion | | | |
|---|---|---|---|---|---|
| | | Fraction | Fraction | Median HH | SES |
| Period | | White | College Ed | Income | Index |
| 1970-1980 | Constant | 1.111 | 0.686 | 0.785 | 0.889 |
| | | (0.023) | (0.008) | (0.014) | (0.006) |
| 1980-1990 | ∆Ln(Employment), | -0.135 | -0.068 | -0.101 | -0.050 |
| | standard devs. | (0.070) | (0.024) | (0.036) | (0.020) |
| | Constant | 0.990 | 1.210 | 0.949 | 0.962 |
| | | (0.019) | (0.007) | (0.009) | (0.005) |
| 1990-2000 | Constant | 0.977 | 1.082 | 0.917 | 0.964 |
| | | (0.011) | (0.005) | (0.004) | (0.004) |
| 2000-2010 | ∆Ln(Employment), | -0.076 | -0.009 | -0.023 | -0.025 |
| | standard devs. | (0.022) | (0.010) | (0.012) | (0.006) |
| | Constant | 0.920 | 1.007 | 0.940 | 0.969 |
| | | (0.011) | (0.004) | (0.005) | (0.003) |
| 1980-2010 | ∆Ln(Employment), | -0.216 | -0.108 | -0.126 | -0.095 |
| | standard devs. | (0.075) | (0.045) | (0.033) | (0.020) |
| | Constant | 0.828 | 1.321 | 0.809 | 0.853 |
| | | (0.027) | (0.014) | (0.010) | (0.007) |

Notes: Each column in each block reports coefficient(s) from a separate regression of the CBSA neighborhood convergence index, built using the variable at top, on the indicated variables and share married, share of population that are children, share college and share white in the base year. ∆ln(Employment) is expressed in standard deviation units and is instrumented with a Bartik quantity instrument using industry shares from 1970, as is explained in the text. Only periods with sufficiently strong first stages have reported employment coefficients. Reported coefficients on the constant can be interpreted as the mean index across CBSAs. Each regression has 120 observations and is weighted by initial year tract population share of the CBSA. First stage F statistics are 10.31 (1980-1990), 33.48 (2000-2010) and 12.53 (1980-2010).

**Table 3: Patterns in SES Index Gentrification of Tracts within 4 km of CBDs**

| | 1970-1980 | 1980-1990 | 1990-2000 | 2000-2010 | 1980-2010 |
|---|---|---|---|---|---|
| Estimator | RF | RF | IV | IV | RF |
| **Panel A: Bottom Tercile Neighborhoods** | | | | | |
| 1(< 4km to CBD) | -0.224 | -0.056 | -0.029 | 0.117 | 0.017 |
| | (0.028) | (0.017) | (0.011) | (0.013) | (0.047) |
| Employment Bartik * 1(< 4km to CBD) | -0.060 | 0.015 | -0.073 | 0.002 | 0.103 |
| | (0.027) | (0.018) | (0.122) | (0.032) | (0.046) |
| Spatial Employment Bartik * 1(< 4km to CBD) | 0.049 | 0.017 | 0.167 | 0.215 | 0.093 |
| | (0.027) | (0.017) | (0.073) | (0.065) | (0.040) |
| N | 12,592 | 12,581 | 12,576 | 12,571 | 12,576 |
| R-Squared | 0.771 | 0.879 | 0.890 | 0.858 | 0.632 |
| **Panel B: Middle Tercile** | | | | | |
| 1(< 4km to CBD) | -0.203 | -0.029 | -0.085 | 0.056 | -0.049 |
| | (0.028) | (0.018) | (0.025) | (0.017) | (0.052) |
| Employment Bartik * 1(< 4km to CBD) | -0.049 | -0.001 | 0.400 | 0.106 | 0.165 |
| | (0.036) | (0.018) | (0.156) | (0.048) | (0.057) |
| Spatial Employment Bartik * 1(< 4km to CBD) | 0.027 | 0.003 | -0.280 | -0.037 | -0.013 |
| | (0.026) | (0.022) | (0.146) | (0.137) | (0.048) |
| N | 12,645 | 12,645 | 12,643 | 12,636 | 12,633 |
| R-Squared | 0.292 | 0.770 | 0.811 | 0.887 | 0.547 |
| **Panel C: Top Tercile** | | | | | |
| 1(< 4km to CBD) | -0.078 | 0.021 | 0.003 | 0.080 | 0.081 |
| | (0.044) | (0.022) | (0.021) | (0.025) | (0.053) |
| Employment Bartik * 1(< 4km to CBD) | -0.088 | 0.015 | -0.097 | -0.036 | 0.031 |
| | (0.057) | (0.024) | (0.119) | (0.040) | (0.051) |
| Spatial Employment Bartik * 1(< 4km to CBD) | 0.150 | 0.001 | 0.135 | 0.224 | 0.072 |
| | (0.052) | (0.024) | (0.090) | (0.107) | (0.065) |
| N | 12,674 | 12,667 | 12,662 | 12,660 | 12,661 |
| R-Squared | 0.528 | 0.856 | 0.886 | 0.905 | 0.649 |

Notes: Each column in each panel reports results from a separate regression of the tract SES gentrification index on indicated variables and indicators for 4-8, and 8-12 km from a CBD and 0-4, 4-8 and 8-12 km from the nearest top 1970 quartile SES index tract. Log of distance to the nearest coastline, lake, and river are also included as controls. See Equation (2) in the text. Bartik variables are standardized to be mean 0 and standard deviation 1. Regressions are weighted by share of 1970 tract population in 1970 CBSA population. Tercile samples are defined as of 1970. Coefficients that are significant at the 10% level are shaded red if positive and blue if negative.

**Table 4: Decomposition of Percent Changes in Population within 2 km of CBDs**
**Fraction of Group in Base Year Totals in Parentheses**

| | | | | | Contribution to Difference Between (1) and (2) from | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | ∆choices of | | | | ∆shares of | |
| Choices in year t | All | None | All | None | Target | Target | NonTarget | NonTarget | X\|Race | Race |
| Shares in year t | All | None | None | All | White | NonWhite | White | NonWhite | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Data Set | | | | | | | | | | |
| **Panel A: 1980-2000** | | | | | | | | | | |
| Education | -0.07 | 0.21 | -0.12 | 0.31 | -0.01 | 0.00 | -0.14 | -0.18 | -0.04 | 0.10 |
| | | | | | (0.09) | (0.01) | (0.74) | (0.15) | | |
| Age | -0.07 | 0.21 | -0.14 | 0.34 | 0.01 | -0.04 | -0.15 | -0.17 | -0.03 | 0.10 |
| | | | | | (0.22) | (0.05) | (0.62) | (0.12) | | |
| Family Type | -0.07 | 0.21 | -0.27 | 0.43 | -0.11 | -0.06 | -0.12 | -0.19 | 0.10 | 0.10 |
| | | | | | (0.29) | (0.04) | (0.55) | (0.12) | | |
| Income | -0.11 | 0.27 | -0.19 | 0.37 | 0.00 | -0.01 | -0.24 | -0.21 | 0.00 | 0.09 |
| | | | | | (0.32) | (0.03) | (0.54) | (0.11) | | |
| **Panel B: 2000-2010** | | | | | | | | | | |
| Education | 0.06 | 0.07 | 0.04 | 0.09 | 0.04 | 0.00 | 0.02 | -0.08 | -0.01 | 0.03 |
| | | | | | (0.14) | (0.03) | (0.61) | (0.22) | | |
| Age | 0.06 | 0.07 | 0.03 | 0.12 | 0.04 | -0.01 | 0.01 | -0.08 | 0.00 | 0.03 |
| | | | | | (0.15) | (0.06) | (0.60) | (0.19) | | |
| Family Type | 0.05 | 0.08 | -0.01 | 0.15 | 0.02 | -0.03 | -0.01 | -0.08 | 0.03 | 0.03 |
| | | | | | (0.24) | (0.06) | (0.50) | (0.20) | | |
| Income | 0.05 | 0.08 | 0.03 | 0.11 | 0.03 | 0.00 | 0.00 | -0.08 | 0.00 | 0.02 |
| | | | | | (0.39) | (0.08) | (0.40) | (0.13) | | |

Notes: Each line usess a different baseline data set as is explained in the text. Results in (1) and (2) report actual data and average CBSA population growth rates respectively. Results in remaining columns use counterfactual data. Results in (5)-(10) sum to actuals in (1) minus CBSA growth in (2). X in (9) refers to the demographic characteristic that is jointly distributed with race in each block. Results weight each CBSA equally. Target groups are college graduates, 20-34 year olds, singles not in group quarters or maried couples without children and households in the top 30 percent of the income distribution of tracts in the sample for each data set respectively. See Table A3 for mathematical expressions used to construct each counterfactual tract population. See the text for a full explanation.

**Table 5: Decompositions of Changes in Fraction White**

| | | | | | | Contribution to All in (1) from | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | $\Delta$choices of | | | | $\Delta$shares of | |
| Choices in year t | | All | None | All | None | Target | Target | NonTarget | NonTarget | X\|Race | Race |
| Shares in year t | | All | None | None | All | White | NonWhite | White | NonWhite | | |
| | | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Data Set | CBD Radius | | | | | | | | | | |
| **Panel A: 1980-2000** | | | | | | | | | | | |
| Education | 2 km | -0.08 | 0.00 | 0.02 | -0.11 | -0.00 | 0.00 | -0.05 | 0.08 | 0.01 | -0.11 |
| Education | 4 km | -0.10 | 0.00 | 0.01 | -0.11 | -0.01 | 0.00 | -0.06 | 0.07 | 0.00 | -0.11 |
| Income | 2 km | -0.08 | 0.00 | 0.02 | -0.10 | 0.00 | 0.00 | -0.09 | 0.10 | 0.00 | -0.10 |
| Income | 4 km | -0.09 | 0.00 | 0.00 | -0.09 | -0.01 | 0.01 | -0.08 | 0.08 | 0.00 | -0.10 |
| **Panel B: 2000-2010** | | | | | | | | | | | |
| Education | 2 km | 0.03 | 0.00 | 0.06 | -0.04 | 0.02 | -0.00 | 0.01 | 0.04 | 0.00 | -0.04 |
| Education | 4 km | 0.01 | 0.00 | 0.04 | -0.04 | 0.01 | 0.00 | -0.00 | 0.04 | 0.00 | -0.04 |
| Income | 2 km | 0.03 | 0.00 | 0.06 | -0.02 | 0.01 | 0.00 | 0.00 | 0.04 | 0.00 | -0.03 |
| Income | 4 km | 0.02 | 0.00 | 0.04 | -0.02 | 0.00 | 0.00 | -0.01 | 0.04 | 0.00 | -0.03 |

Notes: Entries are analogous to those in Table 6 except that the CBSA level statistic of interest differs and both 2 km and 4 km CBD distance rings are examined. See the notes to Table 6 for a description of target groups and Table A3 for mathematical expressions used to calculate these counterfactuals.

**Table 6: Decompositions of Changes in Fraction College Educated**

| | | | | | Fraction of All in (1) from Δchoices of | | | | from Δshares of | |
|---|---|---|---|---|---|---|---|---|---|---|
| Choices in year t | All | None | All | None | Target | Target | NonTarget | NonTarget | X\|Race | Race |
| Shares in year t | All | None | None | All | White | NonWhite | White | NonWhite | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| CBD Radius | | | | | | | | | | |
| **Panel A: 1980-2000** | | | | | | | | | | |
| 2 km | 0.060 | 0.000 | 0.007 | 0.046 | -0.011 | -0.004 | 0.009 | 0.014 | 0.064 | -0.012 |
| 4 km | 0.052 | 0.000 | 0.002 | 0.049 | -0.016 | -0.005 | 0.010 | 0.012 | 0.064 | -0.013 |
| **Panel B: 2000-2010** | | | | | | | | | | |
| 2 km | 0.059 | 0.000 | 0.031 | 0.024 | 0.026 | 0.001 | -0.006 | 0.011 | 0.031 | -0.005 |
| 4 km | 0.043 | 0.000 | 0.018 | 0.023 | 0.006 | -0.002 | 0.001 | 0.013 | 0.030 | -0.006 |

Notes: Entries are analogous to those in Table 6 except that the CBSA level statistic of interest differs and both 2 km and 4 km CBD distance rings are examined. See the notes to Table 6 for a description of target groups and Table A3 for mathematical expressions used to calculate these counterfactuals.

**Table 7: Decompositions of Change of Median Income in terms of Percentiles of Sample Income Distribution**

| | | | | | Fraction of All in (1) from Δchoices of | | | | from Δshares of | |
|---|---|---|---|---|---|---|---|---|---|---|
| Choices in year t | All | None | All | None | Target | Target | NonTarget | NonTarget | X\|Race | Race |
| Shares in year t | All | None | None | All | White | NonWhite | White | NonWhite | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| CBD Radius | | | | | | | | | | |
| **Panel A: 1980-2000** | | | | | | | | | | |
| 2 km | 1.18 | 0.00 | 1.65 | -0.23 | 0.08 | -0.22 | 0.77 | 1.01 | 0.46 | -0.93 |
| 4 km | -0.45 | 0.00 | 0.40 | -0.63 | -1.07 | -0.34 | 0.84 | 0.98 | 0.23 | -1.08 |
| **Panel B: 2000-2010** | | | | | | | | | | |
| 2 km | 3.84 | 0.00 | 4.19 | -0.17 | 1.81 | 0.03 | 1.27 | 1.08 | 0.07 | -0.42 |
| 4 km | 1.79 | 0.00 | 2.06 | -0.18 | 0.50 | -0.14 | 0.75 | 0.95 | 0.19 | -0.46 |

Notes: Entries are analogous to those in Table 6 except that the baseline is the change in the fraction of tracts within the indicated CBD radius within the top tercile of the CBSA distribution of median household income. See the notes to Table 6 for a description of target groups and Table A3 for mathematical expressions used to calculate these counterfactuals.

**Table 8: Patterns of Lambdas for Tracts within 4 km of CBDs**

| Estimator | 1980-1990<br>RF | 1990-2000<br>IV | 2000-2010<br>IV | 1980-2010<br>RF |
|---|---|---|---|---|
| **Panel A: White College+** | | | | |
| 1(< 4km to CBD) | -0.195 | -0.133 | 0.090 | -0.239 |
| | (0.022) | (0.017) | (0.018) | (0.060) |
| Employment Bartik * 1(< 4km to CBD) | -0.040 | -0.250 | 0.046 | 0.061 |
| | (0.021) | (0.151) | (0.051) | (0.071) |
| Spatial Employment Bartik * 1(< 4km to CBD) | 0.033 | 0.265 | 0.230 | 0.146 |
| | (0.024) | (0.103) | (0.126) | (0.078) |
| N | 33,770 | 34,983 | 34,742 | 33,311 |
| R-Squared (First Stage F) | 0.107 | (27.3) | (31.3) | 0.152 |
| **Panel B: Black College+** | | | | |
| 1(< 4km to CBD) | -0.539 | -0.414 | -0.123 | -0.984 |
| | (0.069) | (0.054) | (0.039) | (0.094) |
| Employment Bartik * 1(< 4km to CBD) | -0.030 | 1.128 | -0.313 | -0.010 |
| | (0.060) | (0.547) | (0.088) | (0.083) |
| Spatial Employment Bartik * 1(< 4km to CBD) | 0.028 | -0.725 | 0.376 | -0.036 |
| | (0.057) | (0.352) | (0.155) | (0.078) |
| N | 17,373 | 21,747 | 23,144 | 17,108 |
| R-Squared (First Stage F) | 0.053 | (8.1) | (49.5) | 0.107 |
| **Panel C: White <HS** | | | | |
| 1(< 4km to CBD) | -0.270 | -0.129 | -0.050 | -0.453 |
| | (0.023) | (0.014) | (0.020) | (0.047) |
| Employment Bartik * 1(< 4km to CBD) | -0.014 | 0.074 | -0.081 | -0.008 |
| | (0.021) | (0.132) | (0.045) | (0.047) |
| Spatial Employment Bartik * 1(< 4km to CBD) | -0.001 | -0.056 | 0.109 | -0.044 |
| | (0.022) | (0.085) | (0.121) | (0.060) |
| N | 34,760 | 35,831 | 34,941 | 33,701 |
| R-Squared (First Stage F) | 0.130 | (29.8) | (40.4) | 0.133 |
| **Panel D: Black <HS** | | | | |
| 1(< 4km to CBD) | -0.319 | -0.310 | -0.209 | -0.898 |
| | (0.055) | (0.047) | (0.043) | (0.108) |
| Employment Bartik * 1(< 4km to CBD) | -0.071 | 0.219 | -0.228 | -0.186 |
| | (0.045) | (0.315) | (0.082) | (0.081) |
| Spatial Employment Bartik * 1(< 4km to CBD) | -0.013 | -0.337 | 0.393 | -0.019 |
| | (0.048) | (0.265) | (0.170) | (0.087) |
| N | 17,769 | 19,644 | 19,546 | 16,404 |
| R-Squared (First Stage F) | 0.098 | (11.6) | (44.0) | 0.124 |

Notes: Reported coefficients are from regressions analogous to those in Table 3, except using estimated $\lambda$ utility components for each indicated group rather than the SES index. See Equation (9) in the text. Coefficients that are significant at the 10% level are shaded red if positive and blue if negative. Standard errors are clustered by CBSA.

**Table 9: Patterns in Lambdas of Tracts within 4 km of CBDs**

| | Whites | | | | Blacks | | | |
|---|---|---|---|---|---|---|---|---|
| | 1980-1990 | 1990-2000 | 2000-2010 | 1980-2010 | 1980-1990 | 1990-2000 | 2000-2010 | 1980-2010 |
| Estimator | RF | IV | IV | RF | RF | IV | IV | RF |
| **Panel A: 20th-30th Percentiles** | | | | | | | | |
| 1(< 4km to CBD) | -0.431 | -0.153 | -0.080 | -0.654 | -0.838 | -0.317 | -0.238 | -1.298 |
| | (0.029) | (0.019) | (0.018) | (0.044) | (0.174) | (0.093) | (0.056) | (0.181) |
| CBSA Employment Growth | -0.013 | -0.492 | 0.055 | 0.072 | 0.082 | -0.523 | -0.108 | -0.407 |
| * 1(< 4km to CBD) | (0.030) | (0.173) | (0.049) | (0.050) | (0.164) | (0.468) | (0.122) | (0.153) |
| Near CBD Employment Growth | 0.007 | 0.381 | -0.012 | -0.010 | -0.213 | 0.370 | 0.002 | 0.120 |
| * 1(< 4km to CBD) | (0.032) | (0.111) | (0.129) | (0.062) | (0.110) | (0.303) | (0.244) | (0.139) |
| N | 34,086 | 34,900 | 34,261 | 33,229 | 15,507 | 16,656 | 16,335 | 13,821 |
| R-Squared (First Stage F) | 0.147 | | | 0.199 | 0.098 | | | 0.163 |
| **Panel B: 50th-60th Percentiles** | | | | | | | | |
| 1(< 4km to CBD) | -0.321 | -0.106 | 0.022 | -0.384 | -0.755 | -0.378 | -0.134 | -1.304 |
| | (0.028) | (0.017) | (0.020) | (0.051) | (0.166) | (0.056) | (0.093) | (0.149) |
| CBSA Employment Growth | -0.051 | -0.005 | 0.146 | 0.087 | -0.193 | -0.222 | -0.367 | 0.014 |
| * 1(< 4km to CBD) | (0.027) | (0.166) | (0.056) | (0.058) | (0.207) | (0.360) | (0.191) | (0.158) |
| Near CBD Employment Growth | 0.037 | 0.059 | -0.023 | 0.050 | 0.142 | 0.081 | 0.467 | -0.120 |
| * 1(< 4km to CBD) | (0.042) | (0.110) | (0.139) | (0.073) | (0.105) | (0.273) | (0.350) | (0.152) |
| N | 33,549 | 34,382 | 34,032 | 32,931 | 14,402 | 15,963 | 16,590 | 13,786 |
| R-Squared (First Stage F) | 0.127 | | | 0.157 | 0.187 | | | 0.130 |
| **Panel C: 70th-80th Percentiles** | | | | | | | | |
| 1(< 4km to CBD) | -0.330 | 0.004 | 0.066 | -0.144 | -0.840 | -0.316 | -0.120 | -1.587 |
| | (0.034) | (0.023) | (0.021) | (0.088) | (0.150) | (0.090) | (0.079) | (0.197) |
| CBSA Employment Growth | 0.007 | -0.246 | 0.063 | 0.194 | -0.323 | 0.877 | -0.162 | 0.073 |
| * 1(< 4km to CBD) | (0.034) | (0.206) | (0.059) | (0.090) | (0.156) | (0.558) | (0.147) | (0.117) |
| Near CBD Employment Growth | 0.012 | 0.336 | 0.161 | 0.142 | 0.195 | 0.172 | 0.542 | 0.223 |
| * 1(< 4km to CBD) | (0.037) | (0.129) | (0.151) | (0.078) | (0.099) | (0.332) | (0.302) | (0.116) |
| N | 33,374 | 34,419 | 33,960 | 32,674 | 15,191 | 17,851 | 17,638 | 13,854 |
| R-Squared (First Stage F) | 0.100 | | | 0.107 | 0.087 | | | 0.105 |

Notes: Each column in each panel shows results of a separate regression of the change in log neighborhood choice shares on the indicated variables and various additional CBD distance indicators and distances to exogenous local amenities. See the notes to Table 8 for additional explanation.

**Table A1: Descriptive Statistics for Employment Shocks**

**Panel A: Employment Shocks**

| | Δ ln(CBSA Employment) | | | Δ ln(Employment Within 4 km of CBD) | | |
|---|---|---|---|---|---|---|
| | Mean | SD | Coeff of Var | Mean | SD | Coeff of Var |
| 1990-2000 | 0.10 | 0.09 | 1.11 | -0.07 | 0.12 | -0.58 |
| 2000-2010 | 0.08 | 0.09 | 0.89 | -0.01 | 0.13 | -0.08 |

**Panel B: Instruments**

| | Bartik | | | Spatial Bartik | | |
|---|---|---|---|---|---|---|
| | Mean | SD | Coeff of Var | Mean | SD | Coeff of Var |
| 1970-1980 | 0.11 | 0.02 | 5.15 | 0.14 | 0.02 | 6.29 |
| 1980-1990 | 0.17 | 0.03 | 5.99 | 0.20 | 0.02 | 8.27 |
| 1990-2000 | 0.05 | 0.03 | 1.49 | 0.10 | 0.03 | 3.00 |
| 2000-2010 | 0.07 | 0.03 | 2.44 | 0.08 | 0.02 | 3.54 |
| 1980-2010 | 0.29 | 0.08 | 3.64 | 0.39 | 0.07 | 5.23 |

Notes: We only use employment shocks for the two indicated periods in Table 3 and Table 8. Statistics are across the 120 CBSAs in the sample.

**Table A2: Patterns of Housing Cost in Tracts within 4 km of CBDs**

| | 1970-1980 | 1980-1990 | 1990-2000 | 2000-2010 | 1980-2010 |
|---|---|---|---|---|---|
| **Panel A: Bottom Tercile Neighborhoods** | | | | | |
| 1(< 4km to CBD) | -0.115 | -0.025 | -0.016 | 0.022 | 0.013 |
| | (0.020) | (0.014) | (0.015) | (0.022) | (0.026) |
| Employment Bartik * 1(< 4km to CBD) | -0.041 | -0.004 | 0.007 | 0.015 | 0.008 |
| | (0.024) | (0.013) | (0.015) | (0.025) | (0.030) |
| Spatial Employment Bartik * 1(< 4km to CBD) | 0.028 | 0.046 | 0.027 | 0.043 | 0.118 |
| | (0.016) | (0.015) | (0.014) | (0.026) | (0.030) |
| N | 9,708 | 11,536 | 11,967 | 11,577 | 11,116 |
| R-Squared | 0.419 | 0.607 | 0.632 | 0.481 | 0.357 |
| **Panel B: Middle Tercile** | | | | | |
| 1(< 4km to CBD) | -0.001 | -0.003 | -0.039 | 0.004 | -0.050 |
| | (0.023) | (0.015) | (0.015) | (0.018) | (0.028) |
| Employment Bartik * 1(< 4km to CBD) | -0.040 | 0.025 | 0.020 | 0.029 | 0.065 |
| | (0.030) | (0.018) | (0.016) | (0.023) | (0.041) |
| Spatial Employment Bartik * 1(< 4km to CBD) | 0.035 | 0.019 | 0.017 | 0.015 | 0.049 |
| | (0.027) | (0.015) | (0.016) | (0.024) | (0.039) |
| N | 10,655 | 12,042 | 12,503 | 12,398 | 11,921 |
| R-Squared | 0.284 | 0.576 | 0.740 | 0.604 | 0.373 |
| **Panel C: Top Tercile** | | | | | |
| 1(< 5km to CBD) | 0.073 | -0.019 | 0.001 | 0.007 | -0.019 |
| | (0.022) | (0.022) | (0.019) | (0.017) | (0.028) |
| Employment Bartik * 1(< 4km to CBD) | -0.038 | -0.011 | 0.012 | 0.020 | -0.004 |
| | (0.018) | (0.019) | (0.016) | (0.021) | (0.033) |
| Spatial Employment Bartik * 1(< 4km to CBD) | 0.073 | 0.027 | 0.030 | 0.008 | 0.068 |
| | (0.027) | (0.018) | (0.017) | (0.024) | (0.036) |
| N | 10,648 | 11,990 | 12,413 | 12,378 | 11,908 |
| R-Squared | 0.435 | 0.690 | 0.784 | 0.772 | 0.570 |

Notes: Each column in each panel reports results from a separate regression of the tract owner occupied housing cost index on indicated variables and indicators for 4-8 and 8-12 km from the CBD and 0-4, 4-8 and 8-12 km from the nearest highest quartile SES index tract as of 1970. The log of distance to the nearest coast line, lake, and river are also included as controls. The Bartik variables are standardized to be mean 0 and standard deviation 1. Regressions are weighted by share of 1970 tract population in 1970 CBSA population. The housing cost index is formed from the residuals of a regression of log mean owner occupied home value on housing unit structure characteristics (number of units in building, number of bedrooms in unit, age of building) of the tract and CBSA fixed effects. Coefficients that are significant at the 10% level are shaded red if positive and blue if negative.

**Table A3: Explanation of Counterfactual Experiments**
**Population Distributions Used to Construct Counterfactuals**

| Column in Tables 4-7 | Choices | Shares | Group: Race | Group: X-Dimension | Math Notation |
|---|---|---|---|---|---|
| 1 | All t | All t | All | All | $f_{jt}(i\|r,x)g_{jt}(r,x)$ |
| 2 | All Base Yr | All Base Yr | All | All | $f_{jb}(i\|r,x)g_{jb}(r,x)$ |
| 3 | All t | All Base Yr | All | All | $f_{jt}(i\|r,x)g_{j8}(r,x)$ |
| 4 | All Base Yr | All t | All | All | $f_{j8}(i\|r,x)g_{jt}(r,x)$ |
| 5 | Target Whites t | All Base Yr | Whites | Target | $f_{jt}(i\|r,x)g_{j8}(r,x)$ |
| | | | Blacks, Others | Target | $f_{j8}(i\|r,x)g_{j8}(r,x)$ |
| | | | Whites | Non-Target | $f_{j8}(i\|r,x)g_{j8}(r,x)$ |
| | | | Blacks, Others | Non-Target | $f_{j8}(i\|r,x)g_{j8}(r,x)$ |
| 6 | Target t | All Base Yr | Whites | Target | $f_{jt}(i\|r,x)g_{j8}(r,x)$ |
| | | | Blacks, Others | Target | $f_{jt}(i\|r,x)g_{j8}(r,x)$ |
| | | | Whites | Non-Target | $f_{j8}(i\|r,x)g_{j8}(r,x)$ |
| | | | Blacks, Others | Non-Target | $f_{j8}(i\|r,x)g_{j8}(r,x)$ |
| 7 | Target+Whites t | All Base Yr | Whites | Target | $f_{jt}(i\|r,x)g_{j8}(r,x)$ |
| | | | Blacks, Others | Target | $f_{jt}(i\|r,x)g_{j8}(r,x)$ |
| | | | Whites | Non-Target | $f_{jt}(i\|r,x)g_{j8}(r,x)$ |
| | | | Blacks, Others | Non-Target | $f_{j8}(i\|r,x)g_{j8}(r,x)$ |
| 8 | All t | All Base Yr | All | All | $f_{jt}(i\|r,x)g_{j8}(r,x)$ |
| 9 | All t | X\|r in t, r in Base Yr | All | All | $f_{jt}(i\|r,x)g_{jt}(x\|r)h_{j8}(r)$ |
| 10 | All t | All t | All | All | $f_{jt}(i\|r,x)g_{jt}(x\|r)h_{jt}(r)$ |

Notes: Target groups are college graduates, households in the top three deciles of the income distribution, people aged 20-34 and singles or married couples with no kids.

**Table A4: Aggregate Quantities**

| | Fraction White | Fraction College | Median HH Income | Share in Families without Kids | Share 20-34 |
|---|---|---|---|---|---|
| | | | **Panel A: Entire Sample** | | |
| 1970 | 0.883 | 0.116 | 47881 | | |
| 1980 | 0.836 | 0.102 | 44266 | 0.328 | 0.266 |
| 1990 | 0.809 | 0.138 | 52310 | 0.357 | 0.255 |
| 2000 | 0.753 | 0.167 | 58308 | 0.384 | 0.211 |
| 2010 | 0.717 | 0.196 | 55532 | 0.401 | 0.209 |
| | | | **Panel B: Within 2 km of CBDs** | | |
| 1970 | 0.683 | 0.082 | 32626 | | |
| 1980 | 0.590 | 0.085 | 26281 | 0.404 | 0.300 |
| 1990 | 0.548 | 0.115 | 30991 | 0.376 | 0.317 |
| 2000 | 0.507 | 0.144 | 36770 | 0.420 | 0.298 |
| 2010 | 0.533 | 0.204 | 38423 | 0.454 | 0.324 |
| | | | **Panel C: Within 4 km of CBDs** | | |
| 1970 | 0.722 | 0.089 | 36523 | | |
| 1980 | 0.629 | 0.087 | 31055 | 0.366 | 0.288 |
| 1990 | 0.584 | 0.115 | 35777 | 0.358 | 0.289 |
| 2000 | 0.531 | 0.139 | 40934 | 0.396 | 0.267 |
| 2010 | 0.537 | 0.183 | 39882 | 0.423 | 0.286 |

Notes: Each entry is an average across CBSAs in the sample.

**Table A5: Decomposition of Percent Changes in Population within 2 km of CBDs - Reverse Order**

| | Contribution to Difference Between (1) and (2) in Table 4 | | | | | |
|---|---|---|---|---|---|---|
| | from $\Delta$shares of | | from $\Delta$choices of | | | |
| Choices in year t | X\|Race | Race | Target | Target | NonTarget | NonTarget |
| Shares in year t | | | White | NonWhite | White | NonWhite |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Data Set | | | | | | |
| | | | **Panel A: 1980-2000** | | | |
| Education | -0.04 | 0.13 | -0.02 | -0.01 | -0.11 | -0.24 |
| Age | 0.00 | 0.13 | 0.01 | -0.04 | -0.14 | -0.23 |
| Family Type | 0.10 | 0.12 | -0.11 | -0.09 | -0.09 | -0.21 |
| Income | 0.00 | 0.10 | 0.00 | -0.01 | -0.20 | -0.27 |
| | | | **Panel B: 2000-2010** | | | |
| Education | -0.02 | 0.05 | 0.04 | 0.00 | 0.02 | -0.09 |
| Age | 0.01 | 0.05 | 0.04 | -0.01 | 0.01 | -0.09 |
| Family Type | 0.03 | 0.04 | 0.02 | -0.03 | -0.01 | -0.09 |
| Income | 0.00 | 0.03 | 0.03 | 0.00 | 0.00 | -0.09 |

Notes: Results are analogous to those in Table 4. The only difference is the ordering in which the counterfactuals are imposed.

**Table A6: Decompositions of Changes in Fraction White, Fraction College Educated and Percentile of Median Income - Reverse Order**

| | | Fraction White (See Table 5) | | | | | | Fraction College Educated (T. 6) or Median Income (T. 7) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | from Δshares of | | Δchoices of | | | | from Δshares of | | Δchoices of | | | |
| Choices in year t | | X\|Race | Race | Target | Target | NonTarget | NonTarget | X\|Race | Race | Target | Target | NonTarget | NonTarget |
| Shares in year t | | | | White | NonWhite | White | NonWhite | | | White | NonWhite | White | NonWhite |
| | | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Data Set | CBD Radius | | | | | | | | | | | | |
| **Panel A: 1980-2000** | | | | | | | | | | | | | |
| Education | 2 km | 0.01 | -0.12 | -0.01 | 0.00 | -0.05 | 0.09 | 0.06 | -0.01 | -0.02 | -0.01 | 0.01 | 0.03 |
| Education | 4 km | 0.00 | -0.11 | -0.01 | 0.00 | -0.06 | 0.07 | 0.06 | -0.01 | -0.02 | -0.01 | 0.01 | 0.02 |
| Income | 2 km | 0.00 | -0.10 | 0.00 | 0.00 | -0.09 | 0.11 | 0.28 | -0.51 | 0.11 | -0.27 | 0.46 | 1.11 |
| Income | 4 km | 0.00 | -0.09 | -0.01 | 0.01 | -0.08 | 0.09 | 0.15 | -0.78 | -0.86 | -0.44 | 0.48 | 1.01 |
| **Panel B: 2000-2010** | | | | | | | | | | | | | |
| Education | 2 km | 0.00 | -0.04 | 0.02 | -0.00 | 0.01 | 0.04 | 0.03 | 0.00 | 0.03 | 0.00 | -0.01 | 0.01 |
| Education | 4 km | 0.00 | -0.04 | 0.01 | 0.00 | -0.00 | 0.04 | 0.03 | -0.01 | 0.01 | 0.00 | 0.00 | 0.02 |
| Income | 2 km | 0.00 | -0.03 | 0.01 | 0.00 | 0.00 | 0.04 | 0.14 | -0.31 | 1.72 | 0.06 | 1.21 | 1.02 |
| Income | 4 km | 0.00 | -0.03 | 0.00 | 0.00 | -0.01 | 0.04 | 0.20 | -0.38 | 0.47 | -0.14 | 0.68 | 0.95 |

Notes: Results in Columns 1-6 are analogous to those in Columns 5-10 of Table 5. Results for Education in Columns 7-12 are anologus to those in Columns 5-10 of Table 6. Results for Income in Columns 7-12 are analogous to those in Columns 5-10 of Table 7.