Federal Reserve Bank of Chicago

Accounting for Central Neighborhood Change, 1980-2010

Nathaniel Baum-Snow and Daniel Hartley

REVISED
November 29, 2019

WP 2016-09

Accounting for Central Neighborhood Change, 1980-2010

Nathaniel Baum-Snow, Daniel Hartley*

November 29, 2019

Abstract

After 1980-2000 population decline and economic stagnation, downtown neighborhoods in most large US cities experienced 2000-2010 population growth and gentrification. Stark racial differences in valuations of downtown amenities and suburban labor market opportunities among those with less than college that only emerged after 2000 are the primary drivers of these downtown neighborhood dynamics. As college-educated whites moved in, increases in valuations of downtown amenities encouraged other whites to remain in downtowns, a reversal from prior decades. Continued departures of less than college educated minorities were driven both by relative improvements in suburban employment opportunities for this group and their continued declines in valuations of downtown amenities. Our evidence highlights the importance of racial differences in valuations of potentially endogenous local amenities for understanding recent downtown gentrification.

*Baum-Snow: University of Toronto, 105 St. George St. Toronto, ON M5S 3E6, Nate.Baum-Snow@rotman.utoronto.ca; Hartley: Federal Reserve Bank of Chicago, 230 S. LaSalle St., Chicago, IL 60604, Daniel.A.Hartley@chi.frb.org. We thank participants in various seminars and conferences 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 their staff.

1 Introduction

In the decades following WWII, the downtown neighborhoods of most US metropolitan areas were in decline. A host of mechanisms responsible for this decline have been considered in the literature. These include highway construction (Baum-Snow, 2007), decentralization of low-skilled work (Kain, 1992), white flight from rising minority populations in cities (Boustan, 2010), rising incomes (Margo, 1992), mortgage lending "redlining" (Aaronson, Hartley & Mazumder, 2017), Federal Housing Authority mortgage insurance provision favoring the suburbs (Jackson, 1985) and vintage housing in cities filtering down to lower-income occupants (Brueckner and Rosenthal, 2009). While downtown neighborhood declines continued through the 1980-2000 period when measured in terms of population, they began to reverse when measured using demographic indicators including average household income and the share of the population with a college degree. In the 2000-2010 period, the average city's downtown experienced both population growth and gentrification.

This paper emphasizes the importance of stark racial differences in amenity valuations of downtown neighborhoods and improvements in suburban labor market opportunities among those with less than a college degree for explaining the demographic stabilization and gentrification that occurred in central neighborhoods after 1980. Because 92 percent of downtown residents in the average metro area had no college degree in 1980, shifts in the spatial distribution of neighborhood amenity values and labor market opportunities for this group account for the majority of 1980-2010 demographic change and population dynamics in downtown neighborhoods. Our results indicate that rapidly declining 1980-2000 downtown populations were driven mostly by outflows of less educated whites and minorities alike, mechanically generating some gentrification through population composition effects during this period. Departures were driven primarily by improving suburban employment opportunities (relative to those in city centers), with an additional quantitatively important push given by declining valuations of downtown amenities. For 2000-2010, departures of less educated whites nearly ceased while those of less educated minorities continued unabated. Racial differences in both the suburban employment pull factor and valuation of downtown amenities drive the divergence for this later period. College educated whites and minorities both experienced newly increasing amenity valuations of downtown neighborhoods in the 2000-2010 period. But since only 14 percent of downtown residents were college educated in 2000 in an average metro area, this amenity force accounts for a small part of 2000-2010 downtown demographic change. Stabilization of 1980-2000 downtown employment declines after 2000 also played a small role in generating the reversal of downtown population declines.

Our evidence that the large reductions in rates of downtown population declines after year 2000 are mostly driven by changes in race-specific influences among those with less than a college degree is of particular note. As a baseline, our analysis shows that neighborhoods within 4 km of CBDs (central business districts) lost 33 percent of population on average across the largest 120 CBSAs in the 1980-2000 period because of changes in neighborhood choices, holding the aggregate race-education joint population distributions constant at 1980 levels in each CBSA (core-based statistical area). Holding these joint distributions constant at 2000 levels, central neighborhoods lost 10 percent of population on average in the 2000-2010 period. For 1980-2000, changes in neighborhood choices among those with less than college accounted for 31 of the 33 percent decline, split equally between whites and minorities. For 2000-2010, however, 8 of the 10 percent decline is accounted for by changes in neighborhood choices by less than college minorities, with the remaining 2 by less than college whites. These same patterns appear with greater magnitudes in neighborhoods within 2 km of CBDs.

The post-2000 racial divergence in both valuations of downtown amenities and labor market opportunities in other areas among those with less than a college degree account for these facts. For 1980-2000, all racial groups experienced similar rates of declining amenity valuations of downtown neighborhoods and improving labor market oppurtunities elsewhere, drawing them out of central neighborhoods. For 2000-2010, unobserved amenity valuations among whites with less than a college degree drove a 2 point growth in downtown population, offset by a 2 point decline among minorities. Changing suburban labor market opportunities drove a 1 point decline among whites and a 5 point decline among minorities.

Methodologically, our analysis proceeds in two stages. First, using a procedure akin to that proposed by DiNardo, Fortin & Lemieux (1996) for decomposing wage distributions, we decompose the sources of changes in demand for central neighborhoods into those due to secular demographic shifts, holding neighborhood choices constant, and those due to changes in neighborhood choices of various demographic groups, holding demographic shares constant. While our focus is on central neighborhoods, this method is broadly applicable for decomposing drivers of demographic change for any type of neighborhood, as is also discussed in Carrillo & Rothbaum (2016). In this portion of the analysis, we highlight the importance of the growth in minority population shares to counteract negative group-specific demand shifts for downtown neighborhoods, mitigating downtown population declines driven by changing neighborhood choices among all demographic groups. For 2000-2010, this force was sufficiently large to fully counteract the smaller within-group demand reductions to generate stable or growing downtown populations. Second, we estimate a standard neighborhood choice model to determine demographic group-specific changes in valuations of neighobrhood amenities and labor market opportunities. These estimates facilitate a unified decomposition of the extent to which changing labor demand conditions versus amenities in CBD areas and competing neighborhoods by demographic group explain the components of 1980-2010 central neighborhood demographic change that we attribute to shifts in group-specific neighborhood choices.

A number of other studies also examine the causes of downtown gentrification since 1980 in US cities. These studies primarily focus on forces driving the rise in college educated populations in these areas. Like this study, Couture & Handbury (2019) finds that rising valuations of downtown amenities is the primary driver of 2000-2010 increases in downtown populations of young and college educated residents. Couture & Handbury (2019) provides compelling evidence that changing tastes for restaurants and bars rather than food stores among the young college educated accounts for the largest component of these amenity effects. In contrast, Su (2019) examines the increasing importance of long work hours in high paying downtown jobs, and how this channel may interact with endogenous amenity production. Through estimation of a neighborhood choice model, Su (2019) demonstrates that the rise in long hours among high income occupations has drawn in population to downtown neighborhoods, where commutes to such jobs are shorter. However, he finds an even bigger multiplier that runs through the general equilibrium channel of endogenous amenity production. Edlund, Machado, & Sviatchi (2019) provide further evidence that longer hours worked among skilled workers has promoted growth in downtown housing values and rents through a labor supply channel. Couture et al. (2019) propose a model that generates these endogenous amenity improvements and facilitates analysis of the welfare consequences of downtown gentrification.

Our emphasis on the importance of racial differences in amenity valuations and labor market opportunities, especially among those with less than a college degree, is driven by two features of the data. First, while indeed there were rapid 2000-2010 growth rates in young college educated residents near CBDs of many large cities, they came off of a small base. Therefore, changing neighobrhood choices of less educated individuals continued to be more important quantitatively for driving downtown population and demographic change. Second, these racial differences show up very starkly in our analysis. In the 1980-2000 period, less than college educated downtown residents of all races acted quite similarly. In the 2000-2010 period, the less than college educated whites living in downtowns began to act more like college educated whites. This change represents the largest force pushing the change in downtown population dynamics. Our results are consistent with observations in Gabriel & Rosenthal (1989) that whites and blacks of similar socioeconomic backgrounds have different revealed preferences for downtown and suburban neighborhoods. They are also consistent with Su's (2019) evidence of a quantitatively important endogenous amenity channel driving increased demand for downtown living, but one that is primarily relevant for all groups except minorities with less than a college degree.

A broad literature across the social sciences documents the "back to the city movement" (Sturtevant & Jung, 2011) and considers the consequences of urban gentrification (e.g. Hyra, 2008; Birch, 2009; McKinnish et al., 2010). A necessary prerequisite to learning about the consequences of neighborhood change is to understand its causes. Florida (2002) and more recently Behrens et al. (2018) provide evidence that pioneer gentrifiers of many neighborhoods are the "Creative Class" and small businesses in the arts (such as artist's studios and galleries) that may bestow positive spillovers on incumbents either by improving neighborhood amenities or by alleviating poor nearby employment opportunities (Wilson, 1996). Indeed, the Kerner Report (1968), commissioned to provide policy recommendations in response to the urban riots of the 1960s, recommended establishing policies encouraging residential integration as a way to reduce the despair in "urban ghettoes". With the difficulties many low income minorities face in accessing more affluent neighborhoods, a natural vehicle for integration is thus through just the sort of urban gentrification we study in this paper.

While, by definition, it improves neighborhood quality by some measures, gentrification may also be onerous for incumbent residents. Rising rents may more than offset any potential increases in local amenities as valued by low income incumbents; such incumbents may thus be strictly worse off in the presence of moving costs.[1] The potential for urban gentrification to increase inequality has thus appropriately received considerable recent attention in the literature (Florida, 2017; Sampson, 2019). Our evidence is that incumbent minorities with less than a college degree had similar rates of decline in amenity valuations of downtown neighborhoods in both the 1980-2000 and 2000-2010 periods. This suggests that the 2000-2010 arrival of additional white college graduates into downtowns had little impact on amenity valuations for minorities, though it may account for the 2000-2010 rising amenity values by whites with less than a college degree. Moreover, we find scant evidence that 2000-2010 house price growth is responsible for pushing out many minority downtown residents. In summary, we find no evidence that the 2000-2010 flight of minorities with less than college from downtown neighborhoods was related to the arrival of college educated whites. Instead, if anything, our evidence is consistent with the reverse causal narrative. That is, prior departures of less educated minorities may have precipitated the arrival of college educated whites.

[1] There is little evidence that gentrification has much impact on incumbent adult outcomes. However, there is evidence that gentrification is beneficial for incumbent children (Baum-Snow et al., 2019; Brummet & Reed, 2019).

This paper proceeds as follows. Section 2 describes our data and some key patterns therein. Section 3 lays out decompositions of sources of neighborhood change and presents these results. Section 4 investigates mechanisms driving shifts in central neighborhood choice probabilities as guided by a neighborhood choice model. Finally, Section 5 concludes.

2 Characterizing Neighborhood Change

2.1 Data

We primarily use 1970-2010 decennial Census data and the 2008-2012 American Community Survey (ACS) data tabulated to 2000-definition census tracts for the analysis. Central to our investigation is the need for joint distributions of population by race and: education, household income, age or family structure across census tracts in each CBSA. To recover as many of these joint distributions in the most disaggregated form possible, we use summary tape file (STF) 3 and 4 census tabulations. We also use information about family structure and age by race from STF1 tabulations in the 2010 Census. Because the 2010 Census did not collect information about income or education, we must rely on 5-year ACS tabulations for 2008-2012 for these tract distributions. We also make use of some census micro data to estimate parameters governing shapes of household income distributions above topcodes, to generate weights used to assign some of the population counts in the tract aggregate data to different types of families, and to estimate housing expenditure shares by demographic group. See Appendix A for details.

We construct three different joint distribution data sets for individuals 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 each joint distribution, we have either 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 children, in married couple families with children, in single female headed families with children, in single male headed families with children, not in a family, in group quarters). For income, we construct the number of households in each decile of the household income distribution of those residing in our national sample area in each 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.

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. From this information, we construct the count of workers by industry in each census tract in each of these time periods. We use these data for descriptive evidence and to construct labor demand shocks oriented toward downtown areas. Where available, we take CBD definitions from the 1982 Economic Census. Otherwise, we use the CBD location as assigned by ESRI. Each CBSA is assigned only one CBD.

Our sample includes the regions of all 120 year 2008 definition metropolitan areas (CBSAs) that were tracted in 1970 and had a population of at least 250,000 in the continental United States.[2] In order for our analysis to apply for the average metropolitan area rather than the average resident, each element of our analysis applies tract-level weights such that each CBSA is weighted equally. Appendix A provides more details about our data construction and how we build weights.

To achieve a succinct descriptive analysis, we construct a summary measure of neighborhood demographics that incorporates the share of residents that are white, the share that are college educated and the median household income of the tract. This summary z-score measure for tract i is the average number of standard deviations tract i is away from its CBSA mean in each year across these three measures. We call this equally weighted tract z-score the socioeconomic status (SES) index.[3]

2.2 Facts About Neighborhood Change

Figure 1 shows the prevalence of 1980-2010 central neighborhood gentrification in our sample. For each CBSA, we show how the share of the residential population within 4 km of the CBD that lives in a census tract in the top half of the CBSA's tract-level SES index distribution evolved over time. The left panel shows 1980-2000 trends and the right panel shows 2000-2010 trends. Cities with positive 2000-2010 employment growth within 4 km of CBDs are indicated as circles and those with declining 2000-2010 downtown employment are indicated as crosses. We make this distinction to provide a descriptive indication of whether improving central employment prospects could be behind 2000-2010 downtown gentrification. Symbol size is proportional to 1980 CBSA population.

New York and Santa Barbara are the only CBSAs whose downtown areas were more affluent than average in 1980 (to the right of the vertical dashed line at 0.5). By 2000 Chicago, Orlando and Portland, OR joined this pair and by 2010, eight additional CBSAs were in this category. For 1980-2000, the downtown areas of 72 CBSAs experienced gentrification

[2] 100% of the 2000 definition tract must have been tracted in 1970 to be in our sample. The smallest CBSA in our sample (Binghamton) has 65 tracts and the largest (New York) has 3,423.

[3] While race is not a measure of socioeconomic status, there is evidence that conditional on income and education, black households have lower wealth than white households (Altonji, Doraszelski, and Segal, 2000). We include the share of residents that are white in our SES index as a proxy for unobserved elements of socioeconomic status such as wealth.

(an increase in the SES index) for a median rise of 0.01. For 2000-2010, an additional 5 CBSA downtowns experienced gentrification for a median rise of 0.03, consistent with the fact that many CBSAs appear to the left of the 45 degree line in the right panel of Figure 1. As seen in both panels of Figure 1, while downtown gentrification was more pronounced in larger cities, it shows up throughout the city size distribution and is not limited to cities with growing 2000-2010 downtown employment. However, we note that most small CBSAs with declining 2000-2010 downtown employment did not experience 2000-2010 gentrification. We emphasize that downtown gentrification is very localized near CBDs. Looking within 2 km of CBDs instead, magnitudes of gentrification are on average greater. Further evidence to this effect is seen in Figure 2.

Figure 2 reports statistics describing various aspects of neighborhood change since 1970 as functions of the distance from the CBD. All plots show medians across the CBSAs in our sample in order to emphasize that changes are not driven by just a few large notable cities. Analogous results using means across CBSAs, or using 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 experienced the fastest 2000-2010 growth in terms of population, the share of residents that are white, and the share of residents that are college-educated of all CBD distance bands. The seeds of this gentrification started to form after 1980, as evidenced by more localized upticks in 1980-1990 and 1990-2000 population growth within 1 km of the CBDs.

The evidence in Figure 2 shows that while central area population growth relative to that in the suburbs is a useful indicator of downtown gentrification, two additional features in the data also indicate a turnaround in overall demand for downtown neighborhoods. First, consider the growth in population growth (the second derivative) at each distance relative to that at a CBD distance of 20 km. This object, which is not explicitly plotted but can be discerned in Panel A by comparing the 4 lines, is positive well beyond 2 km from the CBD. At each distance out to 6 or 7 km, the population growth rate relative to that at 20 km from the CBD increased in each decade after the 1970s. That is, accounting for overall CBSA growth, central area population growth rates accelerated over time in a way that was particularly oriented toward CBDs. Second, the results in Panels B and C show that even areas within about 5 km of CBDs that experienced declining 2000-2010 populations experienced faster than average growth in the share of residents that are white and the share that are college educated. Such shifts in demographic composition counteract population declines and reflect increasing demand for living in central neighborhoods by higher SES demographic groups.[4]

[4]Analogous figures which use growth in population density for Panel A or which index space instead by the cumulative distribution function of 1970 CBSA population moving outward from the CBD generate

Similar figures (not reported) show that mean household income started to tick up in central areas relative to outlying areas in the 1990s while mean home values did so in the 1980s, with both relative increases persistently greater near CBDs than in other areas in each subsequent decade up to 2010.

Along with 2000-2010 downtown residential gentrification came a shift in downtown employment growth. Evidence in Panel D of Figure 2 indicates that median near-CBD employment density grew by about 150 jobs per sq km in the 2000-2010 period after steep 1990-2000 declines. Despite this absolute growth, within 2 km remains the CBD distance band with the least rapid 2000-2010 median employment growth in percentage terms. More systematic quasi-experimental evidence discussed in Section 4 confirms the suggestive evidence from Figure 1 and Panel D of Figure 2 that improved post-2000 employment opportunities have at best a minor role in explaining post-2000 downtown gentrification.

The results in Table 1 provide a sense of the heterogeneity around the summary statistics presented in Figure 2. Table 1 reports the fraction of the population within 4 km of a typical CBSA's CBD living in tracts that moved more than 20 percentile points or 0.5 standard deviations up or down CBSA tract distributions of fraction white, fraction college educated and median income. We weight by the tract's share of CBSA population in the base year, meaning all CBSAs are weighted equally. Our results show that central area gentrification was not driven by just a few neighborhoods moving quickly up these distributions. Commensurate with the evidence in Figure 2, two of the three measures indicate that central area tracts were, on balance, in decline during the 1970s, with these declines slowly reversing sometime in the 1980s or 1990s and then advancing quickly between 2000 and 2010.[5]

To help visualize typical trends in neighborhood inequality within CBSAs, Figure 3 depicts the same three demographic indicators and the composite SES Index for each census tract in the Chicago CBSA in 1980 and 2010. Each 2010 indicator is graphed against itself in 1980 in a scatterplot with regression lines (sloped dashed lines), and 1980 and 2010 means (vertical and horizontal dashed lines, repectively). Dark black dots represent tracts within 4 km of the CBD. Regression line slopes of less than 1, for tract income percentile and tract share white, indicate that Chicago neighborhoods have experienced convergence in these dimensions. Points above the regression line and to the left of the 1980 mean represent gentrifying neighborhoods.

Figure 3 reveals considerable heterogeneity in Chicago neighborhood change over the

similar results.

[5]Downtown neighborhoods were the poorest and had among the lowest education levels and shares of white residents of any CBD distance ring in 1980. One potential explanation for downtown gentrification is, thus, simple mean reversion. In Section 4.2 we provide evidence that while mean reversion in neighborhood income and racial composition does exist, it is not the main force behind downtown revitalization.

1980-2010 period, with our three neighborhood change 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 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 number of tracts went in the other direction between 1980 and 2010, seen along the left side of the graph. These 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 12 of the 54 tracts within 4 km of the CBD that were less than 75 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 clear from the measures in the remaining panels of Figure 3 as well. For example, 49 of the 61 tracts within 4 km of the CBD with an SES Index of less than 1 experienced 1980-2010 increases in SES Index.

In summary, the descriptive evidence in Figures 1-3 and Table 1 shows that areas within 4 km of CBDs experienced demographic reversals after 1980, with particularly strong gentrification in the 2000-2010 period and closer to CBDs. While the magnitude of this gentrification was greater in larger cities, it appears throughout the top 100 metropolitan areas, as measured by 1980 population, though not in the 101-120 ranked metros.

3 Decomposing Central Neighborhood Demographic Change

In this section, we decompose 1980-2010 changes in central neighborhood population and demographic composition into two sets of forces. First, we isolate contributions from changes in the propensity of each demographic group to live in central neighborhoods while holding such propensities of other groups and demographic composition constant. Second, we break out the remaining contributions to central neighborhood change that come from changes in overall CBSA demographic compositions. To separate out the relative importance of changing race-specific neighborhood choices from those due to other observed demographic factors that may be correlated with race, we use tract-level joint distributions of race and: education, age, family structure or income over time to construct counterfactual neighborhood compositions. While our focus is on isolating the extent to which increases in the downtown neighborhood choice probabilities and/or population shares of each demographic group have driven downtown population and demographic dynamics, the methodology we use can be applied to recover the main drivers of any dimension of neighborhood demographic change.

Our decomposition procedure has similarities to that developed in Carillo and Rothbaum (2016), following the logic formalized in DiNardo, Fortin & Lemieux (1996) for decomposing wage distributions.

We caution that this decomposition exercise is fundamentally descriptive, thereby generating counterfactuals that do not incorporate endogenous congestion or agglomeration forces, including the house price adjustments that clear residential markets in each neighborhood. In Section 4 we show that a more completely specified economic model of neighborhood choice delivers almost identical implications about group-specific changes in neighborhood choices because house price changes impact each group's neighborhood choice probability similarly on the margin.

3.1 Construction of Counterfactual Neighborhoods

We observe the joint population distribution $f_{jt}(i, r, x)$ of race r and the 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. That is, $f_{jt}(i, r, x)$ denotes the fraction of CBSA j population at time t that is in demographic group (r, x) and lives in tract i. Given the structure of the tabulated census data, we are forced to evaluate counterfactual joint distributions of race (white, black, or other) and only one other demographic attribute at a time across census tracts. We denote CBSA density functions of demographics as $g_{jt}(r, x) = \sum_i f_{jt}(i, r, x)$. We treat $g_{jt}(r, x)$ as exogenous to the allocation of people across neighborhoods, which can be justified in a long-run open-city model in which households first choose a CBSA and then a neighborhood within a CBSA.

Central to our recovery of counterfactuals is the following decomposition:

$$f_{jt}(i, r, x) = f_{jt}(i|r, x)g_{jt}(x|r)h_{jt}(r) \tag{1}$$

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 minorities 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 $h_{jt}(r)$.

Tables 2 and 3 report results of the counterfactual experiments. Table 2 focuses on mechanisms driving average central area population change, equally weighting across CBSAs in our sample. Table 3 similarly decomposes sources of changes in central areas' share white, share college-graduate and median income. In each table, Panels A and B report results for 1980-2000 and 2000-2010, respectively. The remainder of this subsection describes the construction of the counterfactual distributions that are used to generate the output for each column in Tables 2 and 3.

Column 1 in Tables 2 and 3 reports changes in outcomes of interest for central areas calculated using the raw data as a basis for comparison with 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, the numbers in Column 1 of Table 2 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 2, this is the CBSA population growth rate. Because the objects of interest in Table 3 are invariant to scale, Column 2 is all 0s in this table.

The remaining columns of Tables 2 and 3 are built using counterfactual distributions. Our notation indicates column number superscripts on these probability density functions. Column 3 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. The results in Column 4 show the effects of holding choices constant but allowing demographic shares to shift as in equilibrium. These statistics 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, baselines in Column 1 are closer to the results in Column 3 than they are to the than the results in Column 4. This means that changes in neighborhood choices have been more important than changes in demographic composition for generating the observed trends in central area demographics.

Columns 5-10 in Tables 2 and 3 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. Adding the effects of changing demographic shares (Columns 9-10) yields the full difference between the actual data in Column 1 and the "no changes" baseline in Column 2.

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 children, or households in the top three deciles of the income distribution of the full sample area, depending on the data set used. In particular, the set of results for counterfactual column 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 the column headers and $f_{jt}^{c}(i|r,x) = f_{jb}(i|r,x)$ for the remaining elements of (r,x). In Tables 2 and 3, results are expressed as the effect of imposing $f_{jt}^{c}(i,r,x)$ relative to the counterfactual distribution that sets choices of all groups to the base year b, $(f_{jb}(i,r,x))$. As such, Column 5 shows the impacts of target whites' changes in choices only, Column 6 shows the impacts of target nonwhites' changes in choices only, etc.[6]

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 but also on the fraction of that group in the population in the base year. That is, neighborhoods can 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 is driving the results, Table 2 reports the average fraction of the near-CBD populations in each of the four demographic categories in parentheses.

After quantifying the roles of changes in neighborhood choices holding demographic composition constant, the remaining neighborhood change must be due to shifts in demographic composition. To measure 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 couples without children, and the aging of the population

[6] The order of demographic groups for which we impose year t choices does not affect the results because the change in the fraction of the population in tract i as a result of imposing any of these counterfactuals is linear. Thus the full impact of changing neighborhood choices in counterfactual 3 relative to 2 can be achieved by imposing counterfactuals 5, 6, 7 and 8 cumulatively in any order.

have had on downtown neighborhood change while holding CBSA white, black and other race population shares constant. Doing so avoids including the mechanical effects that rising minority shares have on the education, age, family type and income distributions. These results are reported in Column 9 of Tables 2-4, and are built using the expression

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

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 and downtown neighborhoods have above average base year minority shares. Table A1 mathematically specifies the construction of each counterfactual distribution and Table A2 reports the average shares of target groups across CBSAs overall and within 2 km and 4 km CBD distance rings.

We use the counterfactual distributions $f^c_{jt}(i,r,x)$ and base year distributions $f_{jb}(i,r,x)$ to calculate counterfactual central neighborhood change as follows. Population growth for counterfactual c between years b and t reported in Table 2 is constructed using the following expression:

$$g^c_{bt} = \frac{1}{J}\sum_j \left(\ln \frac{N_{jt}}{N_{jb}} + \ln \frac{\sum_x \sum_r \sum_{i \subseteq CBD_j} f^c_{jt}(i,r,x)}{\sum_x \sum_r \sum_{i \subseteq CBD_j} f_{jb}(i,r,x)} \right) \tag{2}$$

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

We construct the counterfactual white share, college graduate share and median income of neighborhoods within 4 km of CBDs, appearing in Table 3, analogously. The exact expressions used to construct these counterfactuals are presented in Appendix B.

Since choices and shares matter multiplicatively for the overall population distribution across tracts, the ordering matters for quantifying the influence of each channel. Table A3 shows results analogous to those in Table 2 but imposes the counterfactuals in the reverse order: shares adjustments first and sub-group-specific choice adjustments second. In practice, it shows that the ordering does not materially affect conclusions from this decomposition exercise.

3.2 Counterfactual Results

The results of the exercise primarily point to a shifting balance between departures of low SES minorities and inflows of high SES whites driving 1980-2010 neighborhood change. In the 1980-2000 period, central neighborhoods experienced the flight of the poor and less educated. This was true for both white and minority households. Their flight was sizable enough to counterbalance a growing minority population at the national-level, which mechanically pushed to increase central area populations. By 2000, there was a clear shift in the racial and SES makeup of near CBD neighborhoods. The 2000-2010 movement of high-SES whites into central neighborhoods strengthened while the outflow of low-SES whites ceased or reversed. The net result was 2000-2010 population growth in the most central neighborhoods of about the same rate as CBSAs as a whole.

Table 2 shows what population growth in 1980-2010 would have been under the various counterfactual scenarios laid out in the prior sub-section. The evidence in Column 1 reiterates the Figure 2 result that populations near CBDs declined until 2000, after which the population within 2 km of CBDs grew at about the same rate as the overall urban population growth reported in Column 2, while that within 4 km was almost unchanged, on average. The results in the remaining columns are quite similar across demographic attributes except "Family Type". As such, we focus our discussion on the race-education joint distribution. Because reversals of 1980-2000 population declines are starkest within 2 km of CBDs, we mainly discuss the results for this region, keeping in mind that similar but more muted patterns hold for regions within 4 km of CBDs as well.

The results holding shares constant in Column 3 show somewhat lower population growth 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. Had the race-education distribution not changed from 1980 through 2000, the central area population would have declined by 12 percentage points (Column 3) rather than the actual decline of 7 percentage points (Column 1) in the average CBSA. In the 2000-2010 period, the central area populations within 2 km of CBDs would have grown by 4 percentage points (Column 3) rather than the 6 percentage points (Column 1) they actually grew, on average. That is, even in the 2000-2010 period, central neighborhood choice probabilities declined in the overall population, with growth in minority shares large enough to counteract these declines and generate small rates of central area population growth. The effects of secular demographic change are roughly the same within 4 km of CBDs as within 2 km of CBDs.

Column 4 of Table 2 shows what would have happened to central area populations had neighborhood choices not changed from base years but demographic shares had. For 1980-

2000, it shows 31 percent growth and for 2000-2010, it shows 9 percent growth within 2 km of CBDs. These results reflect the positive effects associated with a rising minority population reinforced by the imposed lack of shifts in neighborhood choices away from central neighborhoods. Comparing the magnitudes of the results in Columns 3 and 4 of Table 2 indicates that changing neighborhood choices have been the key generator of 1980-2000 central area population decline, even as shifting demographics have pushed for central area population growth. In the 2000-2010 period, shifts in neighborhood choices continued to hold the most central (within 2 km of CBDs) neighborhood growth rates slightly below those for entire CBSAs, with demographic change almost making up for this deficit.

The results in Columns 5-8 of Table 2 show the amount of central area population change driven by changes in neighborhood choices by each of the indicated demographic groups. The entries in Columns 5-8 sum to the difference between the entries in Columns 3 and 2 (-0.34 for 1980-2000 and -0.03 for 2000-2010 within 2 km of CBDs), or the total impact of changing neighborhood choices holding CBSA demographic composition fixed. These results show that 1980-2000 central area population losses are mostly explained by the flight of less than college educated whites and nonwhites alike, whose effects are similar at -0.14 and -0.18 respectively within 2 km of CBDs. The fraction of the population within 2 km or 4 km of the CBD made up by each demographic group is shown in parentheses. With less than college whites representing the largest shares of CBSA and central area populations, the changing choices of less than college nonwhites must have been of greater magnitudes to generate this group's larger impact on downtown population growth. While college educated whites and nonwhites were also choosing to move away from central neighborhoods during 1980-2000, these outflows were much less pronounced and thus contributed little to 1980-2000 central area population declines.

In the 2000-2010 period, minority flight continued and white flight reversed. While less than college educated nonwhites departed central neighborhoods at similar rates as in 1980-2000, whites of all education levels started to return to central neighborhoods. In particular, changing choices of college-educated whites accounted for population growth within 2 km of CBDs of 4 percentage points. Less educated whites were also returning to central areas, but at lower rates than their college-educated counterparts, accounting for 2 percentage points of growth holding shares constant. However, as the departures of less educated minorities contributed negative 8 percentage points to central area population growth, they outweighed the inflows of whites such that changes in neighborhood choices, when aggregated over all demographic groups, continued to contribute to central area population declines.

The results in Column 9 of Table 2 show how shifts in the composition of the education distribution influenced central area population shares holding racial composition constant.

Estimates of -0.04 before 2000 and -0.01 after 2000 reflect declining shares of less educated groups in the population, groups who disproportionately lived in central area neighborhoods in each base year. The results in Column 10 show that the declining white share of the population promoted increases in downtown populations by 10 percentage points in 1980-2000 and 3 percentage points in 2000-2010.

The results for family type-race joint distributions are a bit different from the others, likely because the target group of childless households is not highly correlated with income nor education. Unlike college graduates and high income households, people in childless households represent a larger share of the downtown area population, at 37 percent within 4 km of CBDs in 1980 on average. The growing fraction of people in childless households in the population have thus contributed 0.06 to 1980-2000 and 0.02 to 2000-2010 population growth within 4 km of CBDs holding neighborhood choices and racial composition fixed, as reported in Column 9. Over 70 percent of this phenomenon is driven by whites in both time periods (unreported). However, childless whites also departed central neighborhoods at much higher rates than young and high income whites during the 1980-2000 period (Column 5 of Panel A of Table 2). After 2000, similar to young, educated and high income whites, childless whites departed central neighborhoods at a lower rate (Column 5 of Panel B of Table 2). The fact that the mechanical effect of the growing share of white childless households in the population on central area growth was positive before 2000 and decelerated after 2000 indicates that this demographic shift cannot be an important driver of the post-2000 shift to central area gentrification.

We also examined analogs of Table 2 for four CBSA sub-samples of roughly equal size: 1980 population above and below 600,000 interacted with 2000-2010 downtown employment growth or decline. Results for changing choices of non-target whites and non-whites are broadly similar across these sub-samples, though in the small-negative downtown employment growth sample, the rising departure rate of non-target non-whites after 2000 almost perfectly offset the declining departure rate of whites to produce little change in the downtown population growth rate. In addition, the return rate of target whites to central neighborhoods was considerably more anemic in small CBSAs than in large CBSAs, regardless of 2000-2010 downtown employment growth or decline. These results are available upon request.

Table 3 reports decompositions of changes in fraction white, fraction college educated and median income of residents within 4 km of CBDs into choice- and share-based components. While the central mechanisms driving changes in these demographic indicators can mostly be inferred from the population results in Table 2, a few observations are of note for the 1980-2000 period. The results in Table 3 show why education and income growth before 2000 were leading indicators of racial change in downtown neighborhoods after 2000. Secular growth in college fraction accounted for an increase of 6 percentage points in the fraction of downtown residents with a college education (Panel A, Row 3, Column 9). Moreover, departures of lower income households from central areas of cities promoted a sizable average increase of 1.8 percentiles in median income of these neighborhoods during this period (Panel A, Row 4, Columns 7 and 8). For the 2000-2010 period, central area median income growth accelerated to 1.7 percentiles, with changes in central neighborhood choices by white high income households contributing 0.5 percentiles to this increase - in addition to persistence in mechanisms that existed before 2000.

In summary, the decomposition exercise reveals that 1980-2000 central neighborhood population declines were mostly driven by departures of all racial groups with less than a college education, with downtown neighborhood choices of white college graduates relatively stable. After 2000, less educated whites stopped fleeing central neighborhoods and college educated whites started returning. We find little role for shifting demographic quantities in driving post-2000 downtown gentrification.

4 Understanding Changes in Neighborhood Choices

The prior section presented decompositions quantifying the extent to which central neighborhood demographic change can be understood through shifts in neighborhood choices by various demographic sub-groups. In this section, we investigate the extent to which shifting labor demand conditions versus observed and unobserved amenities generated these shifts in group-specific neighborhood choices. To do so, we lay out a standard neighborhood choice model that delivers regression equations that are used to separate out these forces. Based on the model, we estimate the relative importance of various factors and corroborate our results with more reduced form evidence using the SES index for central areas. We combine the model estimates to generate a unified accounting of the mechanisms driving the choice component of central neighborhood change in the 1980-2000 and 2000-2010 periods. We show that while residential demand in central neighborhoods is responsive to shifts in labor demand conditions, the low employment growth rates in central neighborhoods seen in Panel D of Figure 2 mean that labor market factors were not an important force driving the return of college educated whites to downtowns, though suburban employment growth drove departures of low SES minorities from downtowns throughout our study period. Instead, we find that 2000-2010 rising amenity valuations of downtown neighborhoods amongst whites drove the resurgence of central neighborhoods.

4.1 Explaining Changes in Neighborhood Valuations by Demographic Type

Each household r of type h chooses to reside in the neighborhood that maximizes its indirect utility in each period. The indirect utility associated with living in neighborhood i at time t is:

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

In this expression, p_i^t is the price of one unit of housing services in tract i in year t, w_{hi}^t is wage net of commuting cost, q_i^t summarizes local amenities, and ε_{rhi}^t is an independent and identically distributed (i.i.d.) random utility shock drawn from a Type I extreme value distribution. q_i^t may be a function of exogenous and endogenous neighborhood attributes including the demographic composition of the tract. 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.

In Appendix C we show how this setup delivers an expression that allows us to compute type-specific valuations of neighborhoods using neighborhood-level population shares (π_{hi}^t), housing prices (p_i^t) and the housing expenditure share of demographic type h (β_h). To achieve this, we assume that households have homothetic utility over a composite good x and housing H of the form

$$U_{ht}(x, H; q) = \delta_{ht} q^{\sigma_{ht}} u_h(x, H).$$

Linearizing around the average neighborhood at a point in time, neighborhood valuations (λ_{hi}^t) can be calculated as $\lambda_{hi}^t \equiv \ln \pi_{hi}^t + \beta_h d \ln p_i^t$ to a first-order approximation for small differences across tracts in house prices and incomes.[7] This expression captures the intuition that higher population shares of type h in neighborhood i reflect greater revealed preference for this neighborhood, conditional on housing prices. Furthermore, a higher housing price in tract i signals higher valuations of tract i by type h, particularly if type h spends a large share of income on housing.

The reason to define λ_{hi}^t is that it can be calculated with our data and captures the combination of relative labor market opportunities and amenities across neighborhoods for group h, as follows:

$$\lambda_{hi}^t = \frac{1}{I} \sum_{i'} \ln(\pi_{hi'}^t) + d \ln w_{hi}^t + \sigma_{ht} d \ln q_i^t. \tag{3}$$

[7] If preferences are Cobb-Douglas as in $U_h(x, H; q) = \delta_{ht} q^{\sigma_{ht}} x^{1-\beta_h} H^{\beta_h}$, the expression for λ_{hi}^t holds exactly, even across tracts with large differences in house prices and incomes.

In Equation (3), $\frac{1}{I}\sum_{i'}\ln(\pi^t_{hi'})$ represents a type-specific scale factor which varies over time but not across neighborhoods (discussed further in Appendix C). In the empirical work, we observe an exogenous component of the decadal change in tract i's labor market opportunities for each demographic group ($\Delta_t d\ln w_{hi}$), allowing us to quantify the importance of improved labor demand conditions for driving growth in downtown area neighborhood choice probabilities.

Holding house prices constant, $\pi^b_{hi}(\lambda^t_{hi} - \lambda^b_{hi})$ is the same as $f_{jt}(i|r,x) - f_{jb}(i|r,x)$ from the decomposition exercise (where b is the base year and t is the terminal year), for small changes in neighborhood choice probabilities. House price changes thus reflect a wedge between counterfactuals delivered by our decompositions in Section 3 and counterfactuals given group-specific neighborhood demand shifts in the model. The decomposition exercise only perfectly captures relative magnitudes of neighborhood demand shifts if housing prices do not change in response to demand shocks, meaning that housing supply is perfectly elastic at the neighborhood level. However, given that β_h is similar for all groups h (see Appendix section A.6), the bias in our use of neighborhood choice probabilities to measure the magnitude of demand shifts is similar across groups. As a result, as we report below, using Equation (3) as a basis for inferring the contributions of house price changes to neighborhood change reveals a minimal role for general equilibrium effects in causing us to mis-measure demand shifts from observed group-specific shifts in neighborhood choice probabilities.

Figure 4 gives a sense of how tract valuations, λ^t_{hi} from Equation (3) have changed since 1980 as functions of CBD distance for four demographic groups (one in each panel). The figure shows the average change (across CBSAs) in λ^t_{hi} for 0.5 km CBD distance rings. Figure 4 shows that college whites and blacks and high school dropout whites and blacks all experienced rising valuations of neighborhoods within 2 km of CBDs after 2000, though the estimates for the blacks are much noisier given their small population shares. However, comparing the results in Panel A to those in other panels reveals that college whites have valuations that increase the most over the broadest distance range, out to about 3 km from CBDs. Next, we evaluate the drivers of these changes and their implications for central area population and demographic changes.

To evaluate the mechanisms driving changes in group-specific tract valuations given by Equation (3) over each study decade, Table 4 reports IV regression results of our measures of changes in group-specific tract valuations on various tract-specific measures of labor demand conditions and amenities for four of our twelve education-race demographic groups. The regression specification is as follows, in which j indexes CBSAs:

$$\Delta\widehat{\lambda}^t_{hij} = \rho_{hjt} + \sum\nolimits_{d=1}^{4} a_{hdt} cbddis^d_{ij} + a^b_{h1t} cbddis^1_{ij} \Delta \ln Emp_{jt} + a^s_{h1t} cbddis^1_{ij} \Delta \ln CBDEmp_{jt}$$
$$+ \sum\nolimits_{d=1}^{4} b_{hdt} topdis^d_{ij} + \sum\nolimits_{m} c_{hmt} \ln(amendis^m_{ij}) + e_{hijt}. \quad (4)$$

At the end of this section, we will use these regression estimates to determine the relative importance of various mechanisms for driving changes in group-specific neighborhood choice probabilities and central neighborhood change. We specify the estimation equation to focus on determining reasons for differential changes in tract valuations within 4 km of CBDs relative to in other neighborhoods. To achieve this, we interact variables summarizing changes in CBSA and CBD area labor demand conditions $\Delta \ln Emp_{jt}$ and $\Delta \ln CBDEmp_{jt}$ with an indicator for the tract being within 4 km of a CBD ($cbddis^1_{ij}$). Since our measures of observed amenities, distance bands to top quartile SES Index neighborhoods in 1970 ($topdis^d_{ij}$) and log distances to nearest coastline, lake and river ($\ln(amendis^m_{ij})$), vary at the tract level, we simply include them as regressors. To capture changes in valuations of unobserved amenities in central neighborhoods relative to suburban areas, we include dummy variables for 4 km CBD distance bands ($cbddis^d_{ij}$) out to 16 km. The coefficient a_{h1t} captures a combination of changes in the prevalence of services, restaurants, retail and crime and other local unobserved amenities valued by group h and changes in group h's valuations of these local unobserved amenities in areas within 4 km of CBDs relative to the suburbs. If there are endogenous neighborhood effects, in which for example college graduates' neighborhood demand depends on the number of college graduates in a neighborhood, their relative size for central areas would also load onto a_{h1t}.

Each parameter can be interpreted in the context of Equation (3). ρ_{hjt} accounts for the intercept $\frac{1}{I}\sum_{i'} \ln(\pi^t_{hi'})$. We think of both CBD area and CBSA labor demand shocks as jointly influencing $d \ln w^t_{hi}$. Conditional on CBSA employment growth, CBD area employment growth increases $d \ln w^t_{hi}$ for central area tracts. Conditional on CBD area employment growth, CBSA employment growth reduces $d \ln w^t_{hi}$ in these neighborhoods. Inclusion of 4 km CBD distance ring indicators $cbddis^d_{ij}$ out to 16 km means that a_{h1t} represents average demand shifts for central neighborhoods that occur for unobserved reasons only, which we attribute to amenities given our inclusion of labor demand shifters in the regression.

In most periods, we instrument for standardized CBSA employment growth $\Delta \ln Emp^d_{jt}$ using a Bartik (1991) type industry shift-share variable, isolating labor demand shocks to CBSAs driven by national trends in industry growth weighted by 1970 CBSA industry shares. We instrument for standardized employment growth within 4 km of CBDs $\Delta \ln CBDEmp^d_{jt}$ with a CBD-oriented industry shift-share variable analogous to the CBSA-level employment growth instrument but calcuated only using data from tracts within 4 km of CBDs and using 1990 as the base year. Because we do not observe CBD area employment in 1980, we replace $\Delta \ln Emp^d_{j1990}$ and $\Delta \ln CBDEmp^d_{j1990}$ with their instruments in the 1980-1990 estimation equation, using OLS rather than IV for this period. We detail the construction of the Bartik instruments in Section A.5 of the Appendix. Table A4 reports summary statistics about these two types of shocks in each decade, allowing for translation into direct impacts of employment changes. We use employment growth rather than wage or income growth to build predictor variables primarily for identification reasons. The first stage regression of Bartik employment shocks on CBSA employment growth is stronger than that for wages during the 1990s. To maintain consistency across our two shocks, we thus found it preferable to consistently use employment quantities rather than prices.

While our measures of observed amenities do not change over time, we include them for three reasons. First, we want to account for the possibility that preferences over these amenities have changed over time. Second, controls for natural amenities may anchor affluent neighborhoods (Lee & Lin, 2018), meaning nearby neighborhoods may be less likely to experience demographic change. Third, controls for distance to top quartile tracts account for the possibility that tracts near CBDs gentrified simply because of expansions of nearby high-income neighborhoods (Guerrieri, Hartley, & Hurst, 2013).

We give equal weight to each CBSA-region within 4 km of a CBD and a separate equal weight to each residual CBD region (more than 4 km from the CBD). To achieve this, we weight each tract by $\frac{1}{\#\text{ of tracts in CBSA-region}}$, where a CBSA-region is either the area within 4 km of the CBD or the area more than 4 km from the CBD, as explained further in Appendix A.4. As a result, each CBSA contributes equally to identification of key parameters of interest a_{h1t}, α^b_{h1t} and α^s_{h1t}.[8]

Coefficients on the dummy for a tract being in a neighborhood within 4 km of the CBD shown in the top row of each panel of Table 4 are greater in the 2000-2010 period than for the earlier periods. However, these coefficients only turn significantly positive for college educated whites. These coefficients capture the additional changes in central area residential demands due either to changes in tastes for existing unobserved amenities or increases in such amenities, holding incomes constant.

All groups' demands for central area residency are estimated to respond positively to central area employment shocks in the 2000-2010 period (3rd row of each panel).[9] The asso-

[8] Giving equal weight to all tracts within each CBSA instead yields quantitatively similar results that are slightly muted by the fact that smaller CBSAs receive greater weight in the identification of α_{h1t} since the share of tracts within 4 km of the CBD is higher in smaller CBSAs.

[9] To connect these estimates directly to the model, which expresses labor demand shocks in terms of prices rather than quantities, one must rescale these estimates by the (positive) elasticity of labor supply.

ciated coefficients reflect some combination of improved job opportunities and income effects on existing local amenities. Conditional on nearby labor market opportunities, rising 2000-2010 CBSA employment (2nd row of each panel) significantly drove decreased valuations for near-CBD neighborhoods relative to the suburbs for all but college educated whites. While the responses to CBSA employment shocks are mixed across groups in the 1990s, in the 1980s each group's central neighborhood demand is also estimated to respond negatively to CBSA employment shocks.

The results for whites and blacks who completed high school but not college (not reported in Table 4) are in between the college graduate and high school dropout results for each race. Conditional on educational attainment, the results for the "other" demographic group are between those for whites and blacks, though somewhat more similar to those for whites. We also performed the same exercise as in Table 4 using income deciles instead of education groups and found similar results. The background changes in central neighborhood valuations improved more for the high income deciles than for the low income deciles, but only turned significantly positive after 2000 for high-income whites, not blacks.

4.2 Evidence of Relative Demand Shifts for Central Area Neighborhoods

To corroborate the demographic group-specific evidence in Table 4, we generalize the logic discussed previously for the Chicago example presented in Figure 3 Panel D to each tract in our full sample. In particular, we investigate patterns of changes in central area tracts' demographic composition while accounting for CBSA-specific trends in neighborhood inequality and observable natural amenities whose valuations may have changed over time. Similar to (4), the following regression equation measures such average differences in central area neighborhood change relative to other neighborhoods:

$$\begin{aligned}\Delta S_{ijt} = \rho_{jt} &+ \sum\nolimits_{d=1}^{4} \alpha_{dt} cbddis_{ij}^{d} + \alpha_{1t}^{b} cbddis_{ij}^{1} \overline{\Delta \ln Emp}_{jt}^{d} + \alpha_{1t}^{s} cbddis_{ij}^{1} \overline{\Delta \ln CBDEmp}_{jt}^{d} \\ &+ \sum\nolimits_{d=1}^{4} \beta_{dt} topdis_{ij}^{d} + \sum\nolimits_{m} \delta_{mt} \ln(amendis_{ij}^{m}) + \varepsilon_{ijt} \qquad (5)\end{aligned}$$

where ΔS_{ijt} is the change in tract i's SES index in CBSA j at time t, reflecting a composite of shifts in groups' neighborhood choices. We estimate coefficients in Equation (5) over each decade 1970-2010 and for the full 1980-2010 period.

Panel A of Table 5 reports estimates of α_{1t}, α_{1t}^{b} and α_{1t}^{s}. The first parameter, α_{1t} describes how much more or less gentrification occurred in tracts within 4 km of CBDs relative to what was typical among tracts beyond 16 km from the CBD, which is the excluded distance category. So that α_{1t} measures the average demographic change in central area tracts, we demean $\Delta \ln Emp^d_{jt}$ and $\Delta \ln CBDEmp^d_{jt}$ (as indicated by over-bars). As with the analysis in Table 4, we estimate reduced forms for the 1970-1980, 1980-1990 and 1980-2010 periods instead of instrumental variable (IV) regressions since we do not observe the change in employment within 4 km of CBDs before 1990. Therefore, for these periods magnitudes of α^b_{1t} and α^s_{1t} do not accurately capture causal effects of 1 standard deviation changes in CBSA- and CBD-oriented employment growth, respectively. However, the sign and significance of these coefficients remain informative.

The results in Table 5 Panel A parsimoniously quantify the rebounds experienced by central neighborhoods as shown in Figures 1 and 2. Our estimate of α_1 in the first row is significantly negative for the 1970s, becomes near zero for the 1980s and 1990s, and strengthens further in the 2000-2010 period, showing that, on average, central areas experienced a 0.15 greater increase in the SES index than the typical suburban neighborhood. Over the longer 1980-2010 period, central areas experienced a 0.21 higher increase in the SES index relative to suburban neighborhoods. Due to the interaction terms being normalized to be mean zero and our tract weighting scheme, the interpretation of this first row of coefficients is as an average across CBSAs.

The second and third rows present estimates of α^b_1 and α^s_1, respectively. One consistent finding is that central neighborhoods of CBSAs with more robust central area employment growth experienced relatively more gentrification (seen in the positive α^s_1 coefficients), even in the 1970s. However, this phenomenon was strongest in the 2000-2010 period, when 1 standard deviation greater downtown employment growth generated a 0.13 relative increase in the central area SES index. The effects of CBSA employment growth on downtown neighborhood change depend a lot more on the time period and seem to follow average trends. In the 1970s, central areas of CBSAs with more robust exogenous employment growth deteriorated more than was typical, whereas from 1980-2010 the reverse was true. That is, broader forces buffeting central area neighborhoods appear to be reinforced by aggregate CBSA labor demand shocks. In unreported robustness specifications, we find similar patterns within each tercile of the 1970 SES index distribution. That is, these results are not only being driven by low-SES central neighborhoods.

The evidence from Chicago, shown in Figure 3, revealed that neighborhoods experienced mean reversion in their SES index. This mean reversion is pervasive across CBSAs, and it may be relevant to our setting because central area tracts disproportionately appear in the bottom half of the SES index distribution. We account for such potential mean reversion by including an additional control for S_{ijt-10} in Equation (5) and consolidate S_{ijt-10} onto the right-hand side of the regression equation. This yields an AR(1) specification with CBSA fixed effects fully interacted with the lagged SES index. This specification generates regression lines for each CBSA*decade combination analogous to those in Figure 3 for Chicago.

$$
\begin{aligned}
S_{ijt} = {} & \rho'_{jt} + \mu'_{jt} S_{ijt-10} + \sum{}_{d=1}^{4} \alpha'_{dt} cbddis^{d}_{ij} \\
& + \alpha^{b'}_{1t} cbddis^{1}_{ij} \overline{\Delta \ln Emp}^{d}_{jt} + \alpha^{s'}_{1t} cbddis^{1}_{ij} \overline{\Delta \ln CBDEmp}^{d}_{jt} \\
& + \sum{}_{d=1}^{4} \beta'_{dt} topdis^{d}_{ij} + \sum{}_{m} \delta'_{mt} \ln(amendis^{m}_{ij}) + \varepsilon'_{ijt}
\end{aligned} \tag{6}
$$

These regressions feature the same remaining set of regressors as in Equation (5). Table 5, Panel B reports estimates of coefficients in Equation (6).

The results in Panel B of Table 5 are quite similar to those in Panel A. Whichever assumption we impose about the underlying data-generating process, three main facts persist. First, there is a clear statistically meaningful demographic rebound of central neighborhoods in the 2000-2010 period. Second, central area employment growth/decline bolstered central neighborhood demographic change, especially in the 1970-1980 and 2000-2010 periods. Third, CBSA employment growth had impacts on central neighborhoods which changed over time, lowering near-CBD SES in the 1970-1980 period, and raising it in the 1980-2010 period.

While the empirical approach used in Panel B addresses mean reversion, it is well known that in short panels OLS estimates of μ_{jt} may be biased downward. Such "Hurwicz bias" will influence coefficients of interest α_1, α_1^b and α_1^s if the lagged SES index is correlated with CBD distance, which is likely as near-CBD areas are more likely to be poor. To deal with this bias, we experimented with implementing an Arellano-Bond (1991) type correction. Beginning with Equation (6), we impose that $\mu_{jt} = \mu_{jt-1}$ and, without loss of generality, add a census tract fixed effect and first-difference. As in the Arellano-Bond (1991) estimator, we instrument for ΔS_{ijt-1} with S_{ijt-2}. Unfortunately, this procedure did not generate sufficiently precise estimates to merit reporting them. However, the point estimates of the coefficients are similar to those reported in Panel B of Table 5.

Overall, the evidence in Table 5 indicates that despite causal evidence that positive near-CBD labor demand shocks increase central neighborhood SES, the bulk of 2000-2010 downtown gentrification could not have been driven by shifts in the spatial structure of labor demand. Since 2000-2010 CBD-area employment growth averaged -1 percent across CBSAs, downtown neighborhood growth must have come about for other reasons in most CBSAs.[10]

[10] Regression results analogous to those in Table 5 using an index of tract housing value growth rates as the dependent variable give similar results. These results appear in Table A5.

4.3 Factors Driving Shifts in Downtown Neighborhood Choices

We employ Equation (3) to generate unified decompositions of the relative importance of various mechanisms driving shifts in downtown neighborhood choices. As inputs, we use our tract house price index and education-race specific estimates from Equation (4) for each of the twelve education-race combinations, which we aggregate to the same four groups used in Table 2. This exercise decomposes the contributions of shifts in neighborhood choices to the central neighborhood population change, reported in Columns 5-8 of Table 2, into 6 components: home price changes, CBD-oriented employment growth, employment growth in other areas, observed amenities, central area fixed effects, and a residual. To carry out the decompositions, we construct a series of counterfactual census tract choice shares for each of the twelve education-race groups in 2000 and 2010, taking group-specific 1980 and 2000 neighborhood choice shares as given. To build counterfactual group-specific year 2000 neighborhood choice shares, we apply regression results like those in Table 4 for each of these demographic groups from the 1980s and the 1990s sequentially. Since we do not observe central area employment growth between 1980 and 1990, yet reduced form coefficients on the CBD-oriented Bartik variable are close to zero and statistically insignificant for all groups, we set α^{s}_{h11990} in Equation (4) to zero and estimate the resulting equation by IV for the 1980s (instrumenting for $\Delta \ln Emp_{jt}$ only). For the following decades, we estimate Equation (4) by IV, as specified, for each group. We calculate standard errors using a parametric boostrap assuming joint normality of coefficient estimates with 100 replications.

There are two reasons why this decomposition procedure does not completely describe decadal changes in group-specific neighborhood choice shares. First, we calibrate the impacts of home price changes, imposing our model assumptions about neighborhood own-price demand elasticities for each group. We have nothing to say about which fundamental forces have driven home price changes, but we find that this component is always small. Second, we plug actual CBD-area and CBSA employment growth, a component of which may be endogenous, into the decompositions. The result gives us the "Unexplained" component of our decomposition, which is the wedge between actual and predicted neighborhood demographic change driven by model mis-specification and general equilibrium effects. We view our decompositions as an imperfect substitute to the ideal but infeasible exercise of decomposing the impacts of the fundamental exogenous shocks that have driven downtown neighborhood change, whatever they may be.

Table 6 presents the components of population growth within 4 km of CBDs driven by changes in neighborhood choices of each indicated demographic group, holding demographic shares constant. Each entry can be interpreted as the impact of the indicated force listed at left on the shift in central neighborhood choices for the group listed in the column heading on the average growth rate of central area population. The entries in the "Total" row do not exactly match the numbers in Columns 5-8 of Table 2 because the sample used to estimate the components in Table 6 is slightly more restrictive than the full set of tracts used to construct the numbers in Table 2. For Table 6, we exclude 3 CBSAs for which we have no information on observed amenities. When generating the inputs to Table 6 for demographic group, h, we also exclude any tract with zero population of that group in any year from 1980 through 2010.

The results in Panel A of Table 6 indicate that improving employment opportunities outside of CBD areas was the largest force driving 1980-2000 central area population declines. Point estimates indicate that this force was the most important driver of central area departures for all groups except college educated minorities, accounting for 14 and 11 percentage point declines in central area population through impacts on less educated minorities' and less educated whites' neighborhood choices, respectively. Reductions in the valuation or quality of unobserved amenities represents the second most important driver of 1980-2000 central area population decline. This force accounts for declines of 3 percentage points because of impacts on less educated whites' and 4 percentage points through impacts on less educated minorities' neighborhood choices.

The results in Panel B reveal interesting shifts in the relative importance of mechanisms driving 2000-2010 central area population growth as compared to the prior period. Changes in the valuation of- and/or levels of- unobserved amenities becomes the most important force, quantitatively, for college and less than college whites, and this effect turns positive, accounting for 2 and 3 percentage points of 2000-2010 central area population growth through the changing neighborhood choices of these two groups, respectively. Only less than college educated minorities continued to value these unobserved amenity changes negatively, and less so than in the prior period. Also notable is the almost zero effect of central area employment growth and that employment growth in other areas only continues to appreciably impact departures of less than college minorities after 2000, at 5 percentage points out of their 8 percentage point impact on 2000-2010 central area population declines.

5 Conclusions

Neighborhoods near central business districts of US metropolitan areas have experienced remarkable reversals in population decline and demographic change since 1980. In this paper, we present evidence that population turnarounds were primarily driven by the return of college-graduate and high-income whites to these neighborhoods, coupled with a halt in the outflow of other less-educated whites. However, departures of minorities without college degrees continued unabated. We note that the return of college educated whites to central neighborhoods has not been sufficient to reverse their population declines. Without the mechanical demand growth associated with an increasing minority share in the population, downtown areas would have experienced continued population declines after 2000.

We find scant evidence that post-2000 improvements in downtown labor market opportunities drew in many of the new residents nearby, though this force did contribute to halting some of the outflows of less than college educated whites. Instead, we find that all demographic groups except less than college educated minorities experienced growth in their valuations of unobserved central area amenities in the 2000-2010 period after 1980-2000 declines. The fact that suburban employment growth continued to contribute to outflows of less than college educated minorities after 2000, while ceasing to do so for other groups, is consistent with the idea that downtown amenities became normal goods for everyone except less educated minorities. Otherwise, suburban employment growth should have drawn these other groups away from downtowns as well, as they would benefit from the associated shorter commute times. Couture & Handbury's (2019) compelling evidence that the rising demand for local non-tradeable services among young college graduates explains the majority of downtown revival supports our own more indirect evidence to this effect.

The gentrification of cities' central neighborhoods inverts the decentralization of high-income whites that had been occurring for decades prior to 1980. This represents a fundamental change in the demographic structure of cities, for which this paper provides only a starting point from which to build a deeper understanding. This phenomenon may be the beginning of an urban rebirth with many broader consequences for the economy. It may also exacerbate the rise in real income inequality that has occurred over recent decades, as it is a mechanism through which the cost of living may be rising for the poor. A general equilibrium framework which incorporates both a realistic neighborhood demand system for each demographic group and variation in neighborhood level housing supply elasticity is required to measure the associated welfare consequences. We hope that this study prompts other researchers to pursue this line of inquiry.

A Data Appendix

A large portion of the data used in our analysis come from tract-level tabulations from the Decennial Census of Population and Housing for the years 1970, 1980, 1990, and 2000, and from the American Community Survey (ACS) for the years 2008-2012. We use census tract boundaries from the 2000 census. 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 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 ACS. In these cases, we perform normalizations to 2000 tract boundaries using the appropriate census tract relationship files provided by the U.S. Census Bureau.

A.1 Tract-level Sample

Our sample includes all of the 2008 definition Core Based Statistical Areas (CBSAs) that had a population of at least 250,000 in the area that was tracted in 1970 except Honolulu. Since we use year 2000 tract boundaries, we limit our sample slightly further by using only tracts for which 100% of the 2000 definition tract was tracted in 1970. Our sample consists of 120 CBSAs and 39,087 year 2000 census tracts. 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. The 4 cases are as follows: 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; and 4) Warren-Troy-Farmington Hills, MI, is combined with Detroit-Livonia-Dearborn, MI.

A.1.1 1970, 1990, and 2000 Tract Data

We take these 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 that 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 white and black make up as calculated in 3.

We perform the same process for education and family type in 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 ACS to measure joint distributions of race by education and race by income.

A.2 Procedure for Allocating Income To Percentile Bins

The counterfactual analysis uses 10 household income deciles, with dollar cutoffs calculated using census micro data for the CBSAs in our sample. In each year, the census tract data reports the number of households by race in each of up to 20 income bins bounded by fixed dollar cutoffs. To re-allocate into percentile bins, we assume a uniform distribution within each dollar value bin except the top one. For the top one, we use a Pareto distribution with parameters estimated separately for each year using census micro data.

A.3 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 the year 2000 population. We make two exceptions to this rule based on 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 as designated by the mapping software maker ESRI. 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.

A.4 Construction of Weights

The regressions in Tables 4 and 5 give equal weight to each CBSA region within 4 km of a CBD and each region beyond 4 km, provided that valid data exists in the region in question.

For Table 5, the tract-level weight is:

$$weight^5_{ijr} = \frac{1}{J}\left(\sum_j \frac{N_{jr}}{N_j}\right)\frac{1}{N_{jr}}$$

where i indexes tract in ring r ($<$ 4 km from CBD or $>$ 4 km from CBD) and CBSA j. N_{jr} is the number of tracts in ring r of CBSA j and N_j is the total number of tracts in CBSA j.

For Table 4, the tract level weight is analogous except only tracts with at least some people from the demographic group in each year 1980-2010 are included in the sample. All other tracts get zero weight. Denote N_{jr} as the number of tracts in ring r of CBSA j with at least 1 resident of type h in each year 1980-2010.

$$weight^6_{ijhr} = \frac{1}{J}\left(\sum_j \frac{N^h_{jr}}{N^h_j}\right)\frac{1}{N^h_{jr}}$$

For some smaller groups in a few smaller CBSAs, $N_{jr}^{h} = 0$. In this case, all tracts in area jr are assigned 0 weight. For Table 6, we recalculate the weights analogously but only include in the sample tracts that had at least one resident from demographic group h in 1980, 1990, 2000 and 2010.

A.5 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) with 10 broad industry categories that can be consistently constructed from 1970 through 2010 using the county-level U.S. Census and ACS 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; and 10) Military. We refer to these as 1-digit industry categories. This measure uses the exact geographical boundaries included in each of our CBSA definitions over the entire time period. The Bartik instrument for CBSA j that applies to the period $t-10$ to t is constructed as

$$Bartik_{jt} = \sum_{k} S_{jk1970} \ln(emp_{kt}^{-j}/emp_{kt-10}^{-j}),$$

where S_{jk1970} is the fraction of employment in CBSA j that is in industry k in 1970 and emp_{kt}^{-j} is national employment in industry k at time t excluding CBSA j.

The aim of the CBD area Employment Bartik is to predict which CBSAs might be particularly affected 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 1990 Census Transportation Planning Package. We take these shares and interact them with the national growth rate of that industry to form a CBD-focused 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. 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 total factor productivity (TFP). Therefore, we predict the change in the fraction of employment near the CBD to be

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

A.6 Construction of Housing Expenditure Shares β_h

To construct estimates of β_h (type-specific housing expenditure shares) we use the 1980 Census 5% public use microdata sample. We begin with a sample of renters and owner-occupier households with a mortgage that moved in the 5 years leading up to 1980 and are not living in group quarters. This group experiences housing costs that are closest to 1980 market conditions. We include all mortgage payments, rent, utilities and insurance in housing costs. We trim the 1st and 99th percentiles of housing cost and the 1st and 99th percentiles of household income and take their ratio to calculate the housing expenditure share for each household. We use the median expenditure share for each race - educational attainment cell. These shares range from 0.21 for households of college educated whites to 0.31 for households of black youth. Shares are typically lower for the higher education groups conditional on race and higher for whites than for blacks and those of other races conditional on education group.

B Construction of Counterfactuals in Table 3

We calculate changes in central areas' white and college-graduate shares using the following expressions, respectively. The associated results appear in rows 1-3 of each panel of Table 3.

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

$$\frac{1}{J}\sum_j \left(\frac{\sum_r \sum_{i \subseteq CBD_j} f_{jt}^c(i, r, x = \text{col})}{\sum_x \sum_r \sum_{i \subseteq CBD_j} f_{jt}^c(i, r, x)} - \frac{\sum_r \sum_{i \subseteq CBD_j} f_{jb}(i, r, x = \text{col})}{\sum_x \sum_r \sum_{i \subseteq CBD_j} f_{jb}(i, r, x)} \right). \tag{8}$$

In these expressions, x indexes education group or income decile as indicated in the row label of Table 3. $i \subseteq CBD_j$ indicates a summation only over tracts within 4 km of CBSA j's CBD. The reference change for both outcomes is zero (Column 2 of Table 3), since there is no scale component.

The remaining rows in Table 3 report counterfactual changes in central area median household income. We use median rather than mean income in order to be more robust in avoiding the misallocation households into incorrect income deciles. Since the 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. To see how these medians are constructed, begin with the following expression for the cumulative distribution function of CBSA j's central area households across income deciles $x \subseteq \{1, 2, ..., 10\}$.

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

The income deciles are defined for the full national study area, but here we only focus on the cumulative distribution function (cdf) for central neighborhoods under counterfactual c. 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 4 are

$$\frac{1}{J} \sum_j \left(M^c_{jt} - M_{jb} \right). \tag{9}$$

As a result, positive numbers in Table 4 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.

C Choice Model Derivations

Taking the indirect utility of household r of type h residing in census tract i at time t as $\widetilde{v}^t_{rhi} = v_h(p^t_i, w^t_{hi}, q^t_i) + \varepsilon^t_{rhi} \equiv v^t_{hi} + \varepsilon^t_{rhi}$, the extreme value Type I assumption on the distributions of ε^t_{rhi} conditional on h and t delivers the following population shares of household type h in each census tract i :

$$\pi^t_{hi} = \frac{\exp(v^t_{hi})}{\sum_{i'} \exp(v^t_{hi'})}$$

This implies the relationship

$$\ln \pi^t_{hi} = v^t_{hi} - \ln \left(\sum_{i'} \exp(v^t_{hi'}) \right)$$

Given the extreme value assumption for the errors, the mean, median, and mode are simple functions of v_{hi}^t. With normalization of the scale parameter to 1, the mean tract utility is $v_{hi}^t + 0.58$ (Euler's constant), the median is $v_{hi}^t - \ln(\ln(2))$ and the mode is v_{hi}^t.

We impose as a normalization that the average modal utility across neighborhoods $\frac{1}{I}\sum_{i'} v_{hi'}^t = v_{ht}^{ref}$. This allows for inversion of (**??**) to an expression relating neighborhood choice probabilities to indirect utility, as in Berry (1994).

$$
\begin{aligned}
v_{hi}^t &= \ln \pi_{hi}^t + \ln\left(\sum_{i'} \exp(v_{hi'}^t)\right) + v_{ht}^{ref} - \frac{1}{I}\sum_{i'} v_{hi'}^t \\
&= \ln \pi_{hi}^t - \frac{1}{I}\sum_{i'} (v_{hi'}^t - \ln\left(\sum_{i'} \exp(v_{hi'}^t)\right)) + v_{ht}^{ref}
\end{aligned}
$$

Rearranging these terms, we derive the relationship between neighborhood choice shares and modal neighborhood indirect utility.

$$
\ln \pi_{hi}^t - \frac{1}{I}\sum_{i'} (\ln \pi_{hi'}^t) + v_{ht}^{ref} = v_{ht}(p_i^t, w_{hi}^t, q_i^t) \tag{10}
$$

Full differentiation across tracts, using deviations from the average location, yields an expression that tells us that $\ln v_{hi}^t$ equals a weighted average of wages net of commuting costs, home prices and neighborhood attributes relative to those in the average location. To carry this out, we assume that within each demographic group h, preferences are homothetic over goods and housing and take the form

$$
U_{ht}(x, H; q) = \delta_{ht} q^{\sigma_{ht}} u_h(x, H).
$$

Log differencing the left hand side of (10) relative to the average location yields

$$
\begin{aligned}
d \ln v_{hi}^t &= \ln\left[\ln \pi_{hi}^t - \frac{1}{I}\sum_{i'} (\ln \pi_{hi'}^t) + v_{ht}^{ref}\right] - \ln[v_{ht}^{ref}] \\
&\approx \ln \pi_{hi}^t - \frac{1}{I}\sum_{i'} (\ln \pi_{hi'}^t)
\end{aligned} \tag{11}
$$

if $\ln \pi_{hi}^t - \frac{1}{I}\sum_{i'}(\ln \pi_{hi'}^t) + v_{ht}^{ref} - 1$ and $v_{ht}^{ref} - 1$ are small. In the data, $\ln \pi_{hi} - \frac{1}{I}\sum_{i'}(\ln \pi_{hi'})$ is indeed less than 0.1 in absolute value for all but the most outlier tracts. We can set $v_{ht}^{ref} = 1$ without loss of generality by adjusting δ_{ht} appropriately, maintaining ordinality of utility across neighborhoods.

Differencing the right hand side of (10) using the Envelope Theorem, we have the following for any type h and time period t, as is standard from Roback (1982).

$$
\begin{aligned}
dv_{hi}^t &= \mu_{hi}^t dw_{hi}^t - \mu_{hi}^t H_{hi}^t dp_i^t + \frac{\partial v_{hi}^t}{\partial q} dq_i^t \\
\frac{dv_{hi}^t}{\mu_{hi}^t w_{hi}^t} &= d\ln v_{hi}^t = d\ln w_{hi}^t - \beta_h d\ln p_i^t + \frac{\partial v_{hi}^t / \partial \ln q}{\mu_{hi}^t w_{hi}^t} d\ln q_i^t,
\end{aligned} \tag{12}
$$

where $\mu_{hi}^t = \frac{dv_{hi}^t}{dw}$ is the LaGrange multiplier from utility maximization for type h living in tract i at time t. Homothetic preferences mean that $\frac{d\ln v}{d\ln w} = 1$, so $\mu_{hi}^t w_{hi}^t = \frac{dv_{hi}^t}{d\ln w} = v_{hi}^t$ and $\frac{dv_{hi}^t}{\mu_{hi}^t w_{hi}^t} = \frac{dv_{hi}^t}{v_{hi}^t} = d\ln v_{hi}^t$. Given our assumed form of the utility function, $\frac{\partial v_{hi}^t/\partial \ln q}{\mu_{hi}^t w_{hi}^t} = \frac{\partial \ln v_{hi}^t}{\partial \ln q}$ is thus the constant σ_{ht}.

Combining (11) and (12) and using the definition of λ_{hi}^t from the text, we have:

$$
\lambda_{hi}^t \equiv \ln \pi_{hi}^t + \beta_h d\ln p_i^t = \frac{1}{I}\sum_{i'} \ln(\pi_{hi'}^t) + d\ln w_{hi}^t + \sigma_{ht} d\ln q_i^t. \tag{13}
$$

That is, we can recover the combination of differences in wages net of commuting costs and local amenities across tracts for household type h at time t by looking at log neighborhood choice shares adjusted by housing costs. This formulation incorporates type-specific intercepts $\frac{1}{I}\sum_{i'} \ln(\pi_{hi'}^t)$ that we account for empirically using type-CBSA specific fixed effects in a time-differenced estimating equation.

References

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

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

Aaronson, Daniel, Daniel Hartley, Bhashkar Mazumder (2017) "The Effects of the 1930s HOLC 'Redlining' Maps," Federal Reserve Bank of Chicago Working Paper No. 2017-12.

Arellano, Manuel, and Stephen Bond (1991) "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." The Review of Economic Studies 58.2, 277-297.

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, Nathaniel (2019) "Urban Transport Expansions and Changes in the Spatial Structure of US Cities: Implications for Productivity and Welfare," Review of Economics and Statistics, forthcoming.

Baum-Snow, Nathaniel, Daniel Hartley & Kwan Ok Lee (2019) “The Long-Run Effects of Neighborhood Change on Incumbent Families,” manuscript.

Behrens, Kristian, Brahim Boualam, Julien Martin, Florian Mayneris (2018) “Gentrification and pioneer businesses,” manuscript.

Berry, Steven (1994) “Estimating Discrete-Choice Models of Product Differentiation,” The RAND Journal of Economics, 25 (2), 242–262.

Birch, E.L. (2009) “Downtown in the ‘New American City’,” The ANNALS of the American Academy of Political and Social Science 626: 134–153.

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

Brueckner, J. K., & Rosenthal, S. S. (2009) “Gentrification and neighborhood housing cycles: will America’s future downtowns be rich?” Review of Economics and Statistics, 91(4), 725-743.

Brummet, Quentin & Davin Reed (2019) “The Effects of Gentrification on the Well-Being and Opportunity of Original Resident Adults and Children,” manuscript.

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 SpaHitial Distributions,” Journal of Regional Science, 56(5): 868-894.

Couture, Victor and Jessie Handbury (2019) “Urban Revival in America, 2000 to 2010,” NBER Working Paper #24084.

Couture, Victor, Cecile Gaubert, Jessie Handbury & Erik Hurst (2019) “Income Growth and the Distributional Effects of Urban Spatial Sorting,” NBER Working Paper No. 26142.

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.

Edlund, Lena, Cecilia Machado & Michaela Sviatchi (2019) “Gentrification and the Rising Returns to Skill,” NBER Working Paper #21729.

Ellickson, Bryan (1981) “An Alternative Test of the Hedonic Theory of Housing Markets,” Journal Urban of Economics, 9 (1): 56-79.

Florida, Richard (2002) The rise of the creative class. New York: Basic books.

Florida, Richard (2017) The New Urban Crisis: How Our Cities Are Increasing Inequality, Deepening Segregation, and Failing the Middle Class-and What We Can Do About It. New York: Basic books.

Gabriel, Stuart and Stuart Rosenthal. (1989) “Household Location and Race: Estimates of a Multinomial Logit Model,” The Review of Economics and Statistics, vol. 71, no. 2, 240-249.

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

Hotz, V. Joseph and Robert A. Miller (1993) “Conditional Choice Probabilities and the Estimation of Dynamic Models,” Review of Economic Studies, 60(3), 497-529.

Hyra, Derek (2008) The New Urban Renewal: The Economic Transformation of Harlem and Bronzeville. Chicago, IL: The University of Chicago Press.

Hyra, Derek (2015) “The back-to-the-city movement: Neighbourhood redevelopment and processes of political and cultural displacement,” Urban Studies, 52(10), 1753-1773.

Jackson, Kenneth (1985) Crabgrass Frontier: The Suburbanization of the United States, Oxford University Press, New York.

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

Kerner, Otto, John V. Lindsay, Fred R. Harris, I.W. Abel, Edward W. Brooke, Charles B. Thornton, James C. Corman, Roy Wilkins, William M. McCulloch, Katherine Graham Peden, Herbert Jenkins (1968) Report of the National Advisory Commission on Civil Disorders, US Department of Justice, Washington DC.

Lee, Sanghoon & Jeffrey Lin (2018) “Natural Amenities, Neighbourhood Income Dynamics, and Persistence in the Spatial Income Distribution,” Review of Economic Studies 85(1): 663-694.

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.

Roback, Jennifer. (1982) “Wages, Rents, and the Quality of Life.” Journal of Political Economy, 90: 1257-1278.

Rosen, Sherwin. (1979) “Wages-based Indexes of Urban Quality of Life.” In Current Issues in Urban Economics, edited by P. Mieszkowski and M. Straszheim. Baltimore: Johns Hopkins Univ. Press.

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

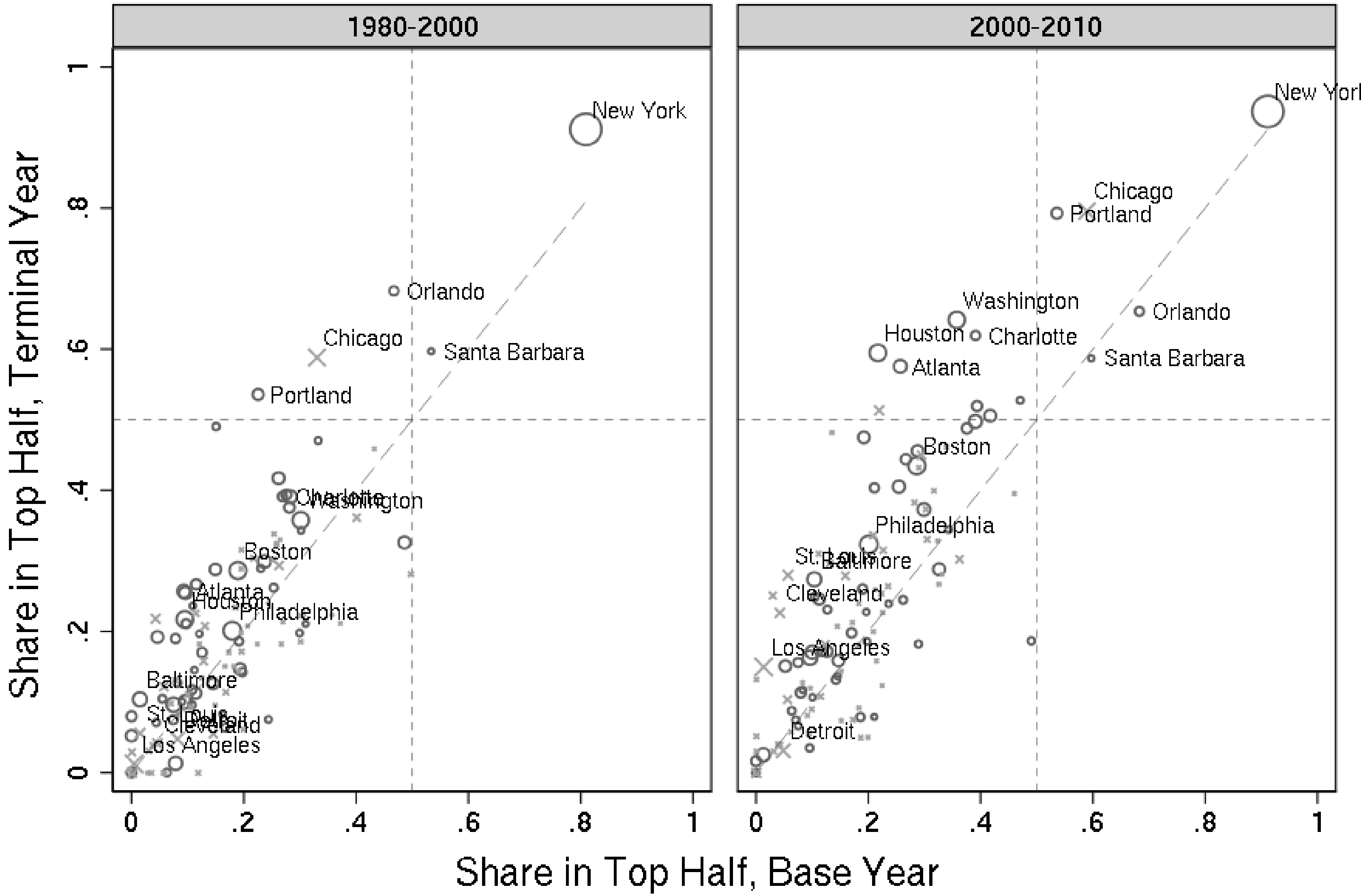
Sampson, Robert (2019) “Neighbourhood effects and beyond: Explaining the paradoxes of inequality in the changing American metropolis,” Urban Studies, 56(1): 3-32.

Sturtevant L.A. and Jung Y.J. (2011) “Are we moving back to the city? Examining residential mobility in the Washington, DC metropolitan area,” Growth and Change 42(1): 48–71.

Su, Yichen (2019) "The Rising Value of Time and the Origin of Urban Gentrification," manuscript.

Wilson W.J. (1996) When Work Disappears: The World of the New Urban Poor. New York: Knopf

Figure 1: SES Index Dynamics in Tracts Within 4 km of CBDs



Graphs show the share of tracts within 4 km of CBDs that are in the top half of the CBSA's SES index distribution, weighted by tract 1980 population, in indicated years. Circles indicate CBSAs with 20000-2010 employment growth in tracts within 4km of the CBD. Crosses indicate CBSAs with 2000-2010 employment declines in this region. Symbols are scaled by 1980 CBSA population. 45-degree lines are also shown.

Figure 2: Measures of Neighborhood Change as a Function of CBD Distance

Medians Across 120 CBSAs, 0.5 km CBD Distance Bands

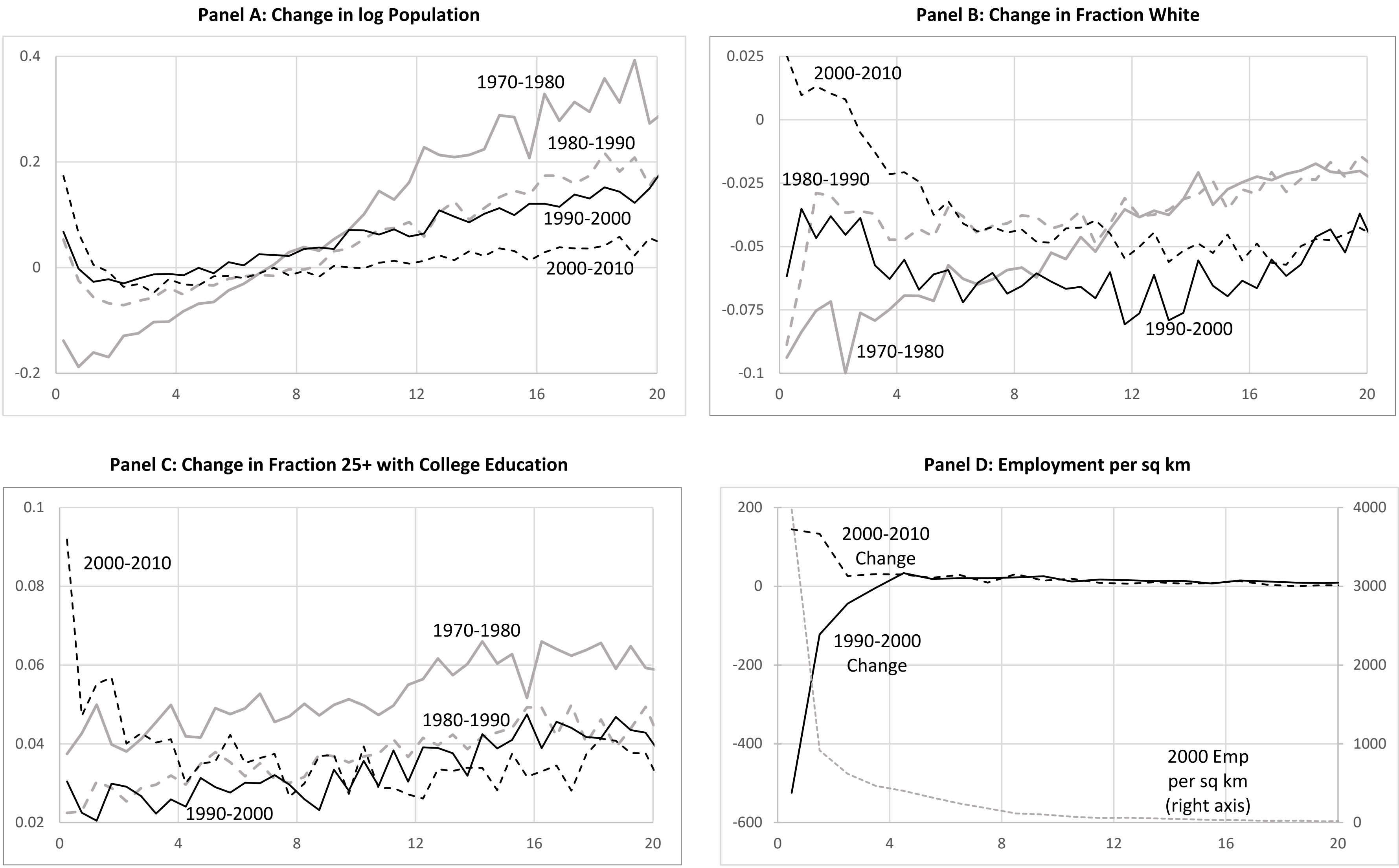
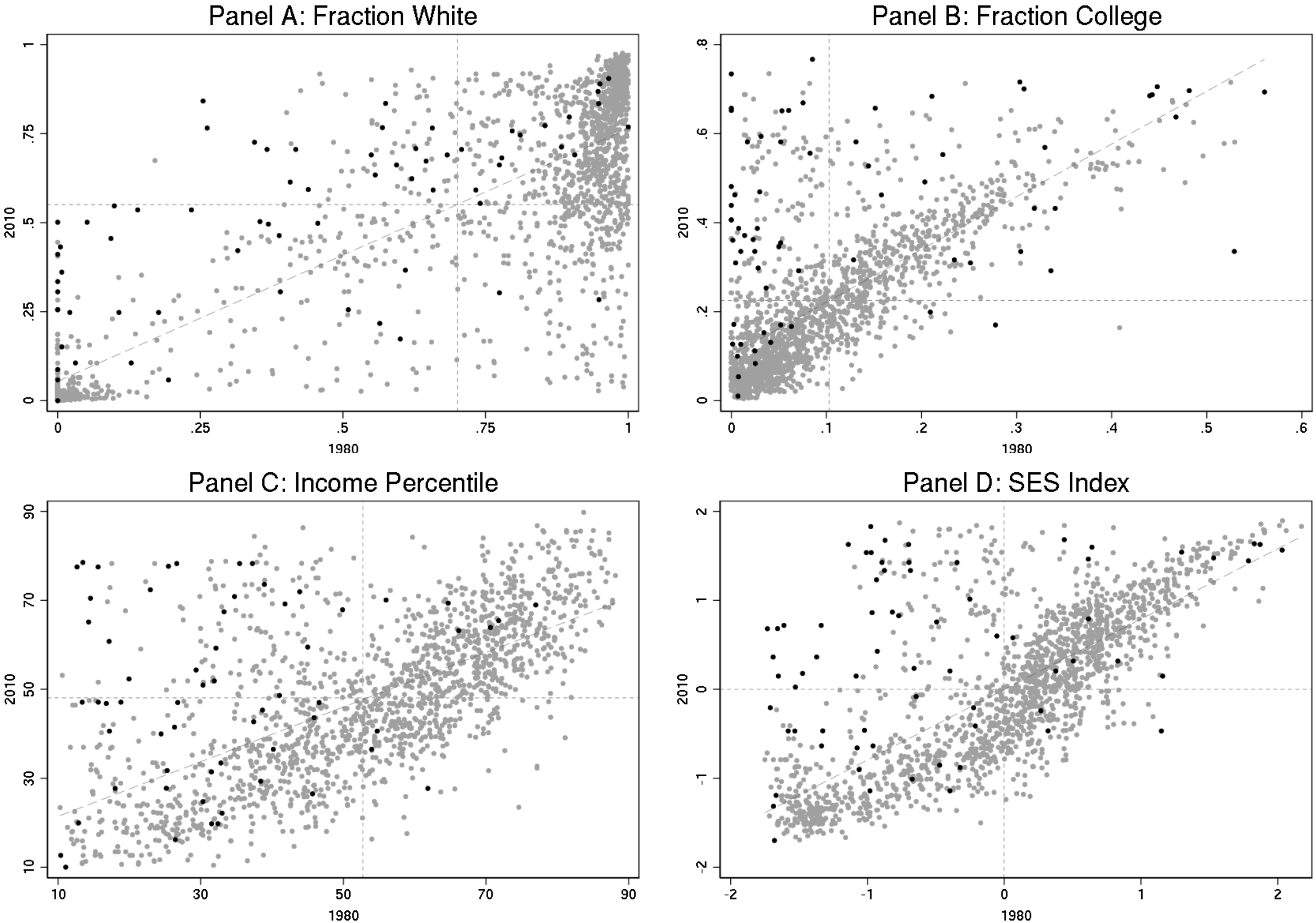


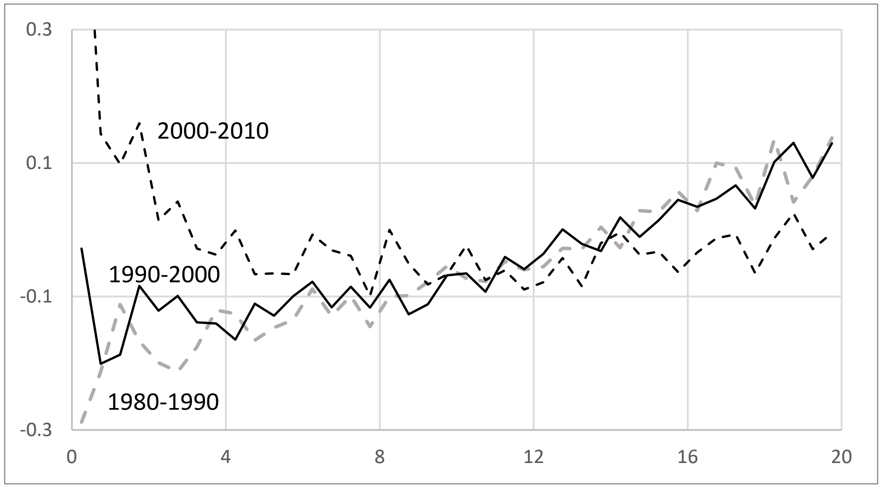
Figure 3: 1980-2010 Neighborhood Change in Chicago



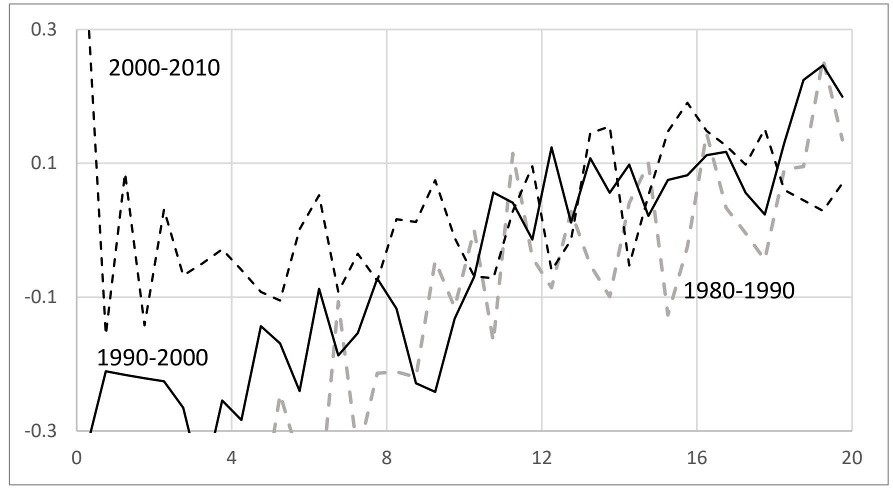
Darker circles indicate tracts within 4 km of the Chicago CBD. Means in 1980 and 2010 and regression lines are indicated in the graphs.

Figure 4: Changes in Neighborhood Valuations as a function of CBD Distance by Race and Education

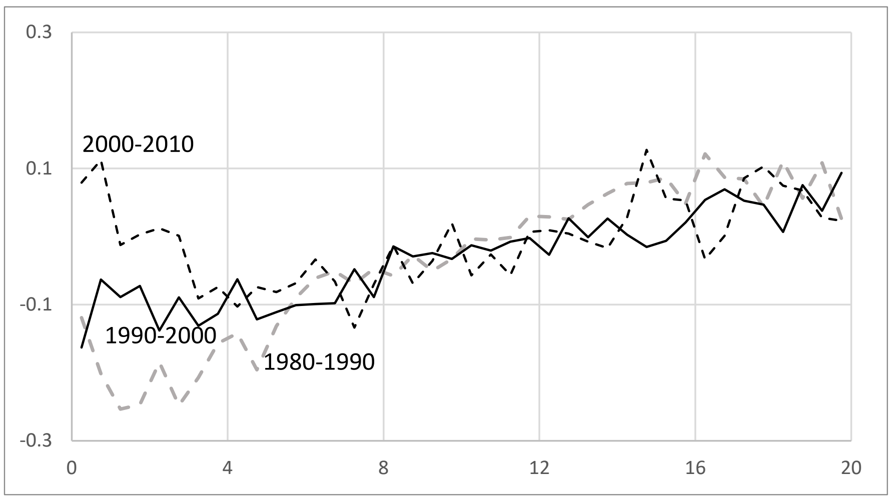
Panel A: Whites with College or More



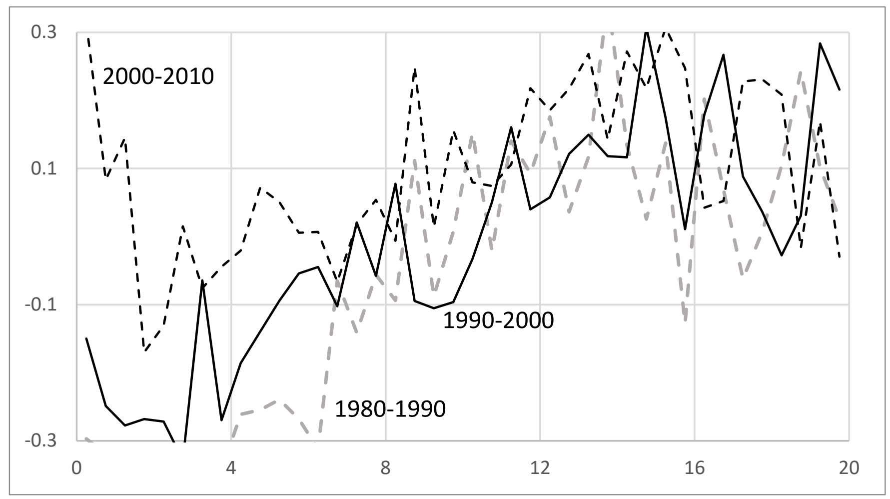
Panel B: Blacks with College or More



Panel C: White High School Dropouts



Panel D: Black High School Dropouts



Each plot indicates the average change across CBSAs in λ for the indicated demographic group over the indicated decade. λ is calculated for each 0.5 km CBD distance band using $\lambda_{hi}^{t} = \ln \pi_{hi}^{t} + \beta_h \, dlnp_i^{t}$.

Table 1: Share of Population within 4 km of CBD in Tracts 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	0.8%	12.2%	3.3%	21.3%
1980-1990	3.5%	1.1%	7.8%	3.3%
1990-2000	3.3%	1.4%	7.6%	2.9%
2000-2010	8.1%	1.5%	14.6%	4.7%
1980-2010	8.0%	1.4%	30.6%	9.1%

Notes: Distributions of fraction white, fraction college educated and median income are calculated separately for each CBSA using tracts weighted by base year population. Each CBSA is weighted equally in reporting.

Table 2: Decomposition of Percent Changes in Central Area Population by Demographic Category

					Contribution to Difference Between (1) and (2) from					
					Δchoices of				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)
Data Set										
Panel A: 1980-2000										
Education, 2 km	-0.07	0.21	-0.12	0.31	-0.01	0.00	-0.14	-0.18	-0.04	0.10
					(0.07)	(0.01)	(0.52)	(0.40)		
Education, 4 km	-0.07	0.21	-0.12	0.28	-0.02	-0.01	-0.16	-0.15	-0.04	0.09
					(0.07)	(0.01)	(0.56)	(0.36)		
Age, 4 km	-0.07	0.21	-0.14	0.31	-0.01	-0.03	-0.16	-0.14	-0.02	0.09
					(0.19)	(0.10)	(0.44)	(0.27)		
Family Type, 4 km	-0.07	0.21	-0.22	0.37	-0.10	-0.05	-0.14	-0.15	0.06	0.09
					(0.27)	(0.10)	(0.36)	(0.27)		
Income, 4 km	-0.07	0.27	-0.15	0.35	-0.03	-0.01	-0.22	-0.16	0.00	0.07
					(0.12)	(0.03)	(0.56)	(0.29)		
Panel B: 2000-2010										
Education, 2 km	0.06	0.07	0.04	0.09	0.04	0.00	0.02	-0.08	-0.01	0.03
					(0.11)	(0.03)	(0.40)	(0.46)		
Education, 4 km	-0.01	0.07	-0.03	0.08	0.01	0.00	-0.02	-0.09	-0.01	0.03
					(0.11)	(0.03)	(0.42)	(0.44)		
Age, 4 km	-0.01	0.07	-0.04	0.11	0.01	-0.02	-0.02	-0.09	0.00	0.03
					(0.15)	(0.12)	(0.38)	(0.35)		
Family Type, 4 km	-0.02	0.08	-0.07	0.14	-0.01	-0.03	-0.03	-0.08	0.02	0.03
					(0.26)	(0.14)	(0.27)	(0.33)		
Income, 4 km	0.00	0.09	-0.02	0.12	0.01	-0.01	-0.04	-0.08	0.00	0.03
					(0.11)	(0.04)	(0.48)	(0.38)		

Notes: All results are averages over the 120 CBSAs in our sample weighting each CBSA equally. Results in (1) and (2) report actual percent changes in population in the indicated CBD distance ring and average CBSA population growth rates respectively. Results in remaining columns use counterfactual data for the indicated demographic group. Table A1 presents the mathematical expression for each counterfactual. Results in (5)-(10) sum to actuals in (1) minus CBSA growth in (2). Entries in parentheses show the average base year fraction of the near-CBD population in the indicated demographic group. The income joint distribution uses households rather than people. Target groups are college graduates, ages 20-34, single or married without children and in the top 3 deciles of the sample area household income distribution.

Table 3: Decompositions of Changes in Demographic Composition within 4 km of CBDs

					Contribution to All in (1) from					
					Δchoices of				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)
Outcome										
Data Set										
Panel A: 1980-2000										
Fraction White										
1 Education	-0.10	0.00	0.01	-0.11	-0.01	0.00	-0.06	0.07	0.00	-0.11
2 Income	-0.09	0.00	0.00	-0.09	-0.01	0.01	-0.08	0.08	0.00	-0.10
Fraction College										
3 Education	0.05	0.00	0.00	0.05	-0.02	0.00	0.01	0.01	0.06	-0.01
Median Income (Percentile Points of Sample Area Distribution)										
4 Income	-0.45	0.00	0.40	-0.63	-1.07	-0.34	0.84	0.98	0.23	-1.08
Panel B: 2000-2010										
Fraction White										
1 Education	0.01	0.00	0.04	-0.04	0.01	0.00	-0.00	0.04	0.00	-0.04
2 Income	-0.00	0.00	0.03	-0.04	0.00	0.00	-0.01	0.04	0.00	-0.04
Fraction College										
3 Education	0.04	0.00	0.02	0.02	0.01	-0.00	0.00	0.01	0.03	-0.01
Median Income (Percentile Points of Sample Area Distribution)										
4 Income	1.74	0.00	2.01	-0.19	0.51	-0.10	0.71	0.89	0.26	-0.54

Notes: Entries are analogous to those in Tables 2 except that the CBSA-level statistic of interest differs. See the notes to Table 2 for a description of target groups and Table A1 and Appendix B for mathematical expressions used to calculate these counterfactuals.

Table 4: Changes in Tract Valuations by Race and Education

Estimator	1980-1990 RF	1990-2000 IV	2000-2010 IV
Panel A: White College+			
1(< 4 km to CBD)	0.004	-0.058	0.149
	(0.065)	(0.226)	(0.049)
CBSA Employment Growth* 1(< 4 km to CBD)	-0.200	-0.065	-0.023
	(0.042)	(0.159)	(0.046)
CBD Area Employment Growth* 1(< 4 km to CBD)	0.030	0.106	0.269
	(0.021)	(0.082)	(0.094)
Observations	33,891	35,169	35,077
R-Squared (First Stage F)	0.058	(20.3)	(48.2)
Panel B: Black College+			
1(< 4 km to CBD)	-0.427	-2.111	0.140
	(0.215)	(0.662)	(0.084)
CBSA Employment Growth* 1(< 4 km to CBD)	-0.174	1.146	-0.291
	(0.129)	(0.470)	(0.075)
CBD Area Employment Growth* 1(< 4 km to CBD)	0.010	-0.539	0.299
	(0.050)	(0.210)	(0.121)
Observations	17,423	21,844	23,380
R-Squared (First Stage F)	0.047	(10.2)	(72.2)
Panel C: White <HS			
1(< 4 km to CBD)	-0.206	-0.443	-0.007
	(0.044)	(0.211)	(0.045)
CBSA Employment Growth* 1(< 4 km to CBD)	-0.073	0.179	-0.089
	(0.030)	(0.150)	(0.042)
CBD Area Employment Growth* 1(< 4 km to CBD)	-0.016	-0.107	0.105
	(0.014)	(0.071)	(0.100)
Observations	34,883	36,014	35,252
R-Squared (First Stage F)	0.079	(22.0)	(57.0)
Panel C: Black <HS			
1(< 4 km to CBD)	0.067	-0.643	-0.020
	(0.139)	(0.513)	(0.087)
CBSA Employment Growth* 1(< 4 km to CBD)	-0.298	0.106	-0.269
	(0.085)	(0.340)	(0.073)
CBD Area Employment Growth* 1(< 4 km to CBD)	0.018	-0.229	0.263
	(0.032)	(0.196)	(0.143)
Observations	17,808	19,698	19,642
R-Squared (First Stage F)	0.071	(11.8)	(60.9)

Notes: Reported coefficients are from regressions analogous to those in Table 5 Panel A, except using changes in λ utility components for each group indicated in panel headers rather than the unified SES index. Equation (8) in the text shows the full regression specification used. CBSA and CBD area employment shocks are normalized to have a standard deviation of 1. All regressions have CBSA fixed effects. Tracts with valid data 1980-2010 are equally weighted within 0-4 km and beyond 4 km in each CBSA, such that each distance ring gets equal weight across CBSAs. Because of sample differences, weights differ across groups. Coefficients that are significant at the 10% level are shaded red if positive and blue if negative.

Table 5: SES Index Regressions
Equally weighted Rings

	1970-1980	1980-1990	1990-2000	2000-2010	1980-2010
Estimator	RF	RF	IV	IV	RF
Panel A: Difference Specification					
1(< 4 km to CBD)	-0.203	0.013	0.024	0.153	0.205
	(0.022)	(0.013)	(0.008)	(0.008)	(0.043)
Standardized CBSA Emp Growth	-0.043	0.011	0.120	0.025	0.108
X 1(< 4 km to CBD)	(0.020)	(0.015)	(0.091)	(0.020)	(0.044)
Standardized CBD Area Emp Growth	0.047	0.017	0.052	0.129	0.064
X 1(< 4 km to CBD)	(0.015)	(0.013)	(0.043)	(0.035)	(0.038)
Observations	37,924	38,329	38,275	38,249	38,279
R-Squared (First Stage F)	0.123	0.028	(19.0)	(71.6)	0.114
Panel B: AR(1) Specification					
1(< 4 km to CBD)	-0.308	-0.033	-0.001	0.153	0.082
	(0.025)	(0.016)	(0.008)	(0.009)	(0.044)
Standardized CBSA Emp Growth	-0.040	0.000	0.143	0.018	0.100
X 1(< 4 km to CBD)	(0.022)	(0.018)	(0.091)	(0.023)	(0.046)
Standardized CBD Area Emp Growth	0.043	0.020	0.059	0.136	0.071
X 1(< 4 km to CBD)	(0.021)	(0.016)	(0.040)	(0.037)	(0.037)
Observations	37,924	38,329	38,306	38,281	38,279
R-Squared (First Stage F)	0.780	0.882	(22.0)	(87.3)	0.666

Notes: Each column in each panel reports results from a separate regression of the change in (Panel A) or level of (Panel B) the tract SES index on variables listed at left 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 Equations (4) and (5) in the text for specifications used in Panels A and B respectively. Employment growth variables and their Bartik instruments are standardized to be mean 0 and standard deviation 1. "RF" refers to "reduced form" and "IV" stands for "instrumental variables" in column headers. Tracts with valid data 1980-2010 are equally weighted within 0-4 km and beyond 4 km in each CBSA, such that each distance ring gets equal weight across CBSAs. Coefficients that are significant at the 10% level are shaded red if positive and blue if negative. RF standard errors are clustered by CBSA.

Table 6: Contributions to Changes in Central Area Neighborhood Choices by Demographic Group

Due to …	College White	College NonWhite	< College White	< College NonWhite	All
Panel A: 1980-2000					
Chg. In Home Prices	0.000	0.000	0.002	-0.002	0.00
Central Area Employment	-0.002	0.001	-0.017	0.006	-0.01
	(0.001)	(0.001)	(0.006)	(0.009)	(0.01)
Employment in Other Areas	-0.011	0.002	-0.107	-0.135	-0.25
	(0.010)	(0.004)	(0.025)	(0.032)	(0.04)
Val. Of Observed Amenities	-0.002	0.000	-0.029	-0.005	-0.04
	(0.002)	(0.001)	(0.008)	(0.006)	(0.01)
Unobserved Amenities	-0.004	-0.009	-0.031	-0.035	-0.08
	(0.001)	(0.000)	(0.002)	(0.003)	(0.00)
Unexplained	0.002	0.001	0.022	0.022	0.05
	(0.009)	(0.004)	(0.023)	(0.035)	(0.04)
Total	-0.02	0.00	-0.16	-0.15	-0.33
Panel B: 2000-2010					
Chg. In Home Prices	0.000	0.000	0.000	-0.004	0.00
Central Area Employment	0.005	0.000	0.001	0.010	0.02
	(0.001)	(0.001)	(0.001)	(0.004)	(0.00)
Employment in Other Areas	0.000	-0.004	-0.008	-0.049	-0.06
	(0.003)	(0.001)	(0.003)	(0.005)	(0.01)
Val. Of Observed Amenities	-0.003	-0.001	-0.012	-0.006	-0.02
	(0.001)	(0.000)	(0.001)	(0.001)	(0.00)
Unobserved Amenities	0.016	0.004	0.026	-0.016	0.03
	(0.001)	(0.001)	(0.002)	(0.003)	(0.00)
Unexplained	-0.014	-0.002	-0.028	-0.017	-0.06
	(0.002)	(0.001)	(0.003)	(0.006)	(0.01)
Total	0.00	0.00	-0.02	-0.08	-0.10

Notes: Elements in the "Total" rows match elements in the "Education, 4 km" rows and Columns 5-8 of Table 2, except numbers in this table are built excluding some tracts to maintain consistent samples over time. Regressions used as a basis for calculating results are the same as those in Table 4 (plus those for the other demographic groups) but maintaining consistent samples over time. Tracts with 0 reported population of demographic group h in at least one year during the study period receive 0 weight for that group. Patterns of coefficient sign and statistical significance are not affected. Coefficients that are significant at the 10% level are shaded red if positive and blue if negative. Standard errors are constructed using a parametric bootstrap with 100 replications.

Table A1: Explanation of Counterfactual Experiments
Population Distributions Used to Construct Counterfactuals

Column in Tables 2-4	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	$f_{jt}(i\|r,x)g_{jb}(r,x)$
4	All Base Yr	All t	All	All	$f_{jb}(i\|r,x)g_{jt}(r,x)$
5	Target Whites t	All Base Yr	Whites	Target	$f_{jt}(i\|r,x)g_{jb}(r,x)$
			Blacks, Others	Target	$f_{jb}(i\|r,x)g_{jb}(r,x)$
			Whites	Non-Target	$f_{jb}(i\|r,x)g_{jb}(r,x)$
			Blacks, Others	Non-Target	$f_{jb}(i\|r,x)g_{jb}(r,x)$
6	Target t	All Base Yr	Whites	Target	$f_{jt}(i\|r,x)g_{jb}(r,x)$
			Blacks, Others	Target	$f_{jt}(i\|r,x)g_{jb}(r,x)$
			Whites	Non-Target	$f_{jb}(i\|r,x)g_{jb}(r,x)$
			Blacks, Others	Non-Target	$f_{jb}(i\|r,x)g_{jb}(r,x)$
7	Target+Whites t	All Base Yr	Whites	Target	$f_{jt}(i\|r,x)g_{jb}(r,x)$
			Blacks, Others	Target	$f_{jt}(i\|r,x)g_{jb}(r,x)$
			Whites	Non-Target	$f_{jt}(i\|r,x)g_{jb}(r,x)$
			Blacks, Others	Non-Target	$f_{jb}(i\|r,x)g_{jb}(r,x)$
8	All t	All Base Yr	All	All	$f_{jt}(i\|r,x)g_{jb}(r,x)$
9	All t	X|r in t, r in Base Yr	All	All	$f_{jt}(i\|r,x)g_{jt}(x\|r)h_{jb}(r)$
10	All t	All t	All	All	$f_{jt}(i\|r,x)g_{jt}(x\|r)h_{jt}(r)$

Notes: Entries show the basis for the construction of each counterfactual in Tables 2-4. See Section 3.1 of the text for an explanation of notation. 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. Entries in Columns 1-4 of Tables 2-4 only are built using the indicated counterfactual distributions. Entries in Column 5 are built using the indicated distribution to calculate statistics relative to those calculated using the distribution in Column 2. Entries in remaining columns c>5 use the indicated distribution relative to statistics built using the distributions associated with columns c-1.

Table A2: 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 A3: Decomposition of Percent Changes in Population - Reverse Order

	Contribution to Difference Between (1) and (2) in Table 2					
	from Δshares of		from Δchoices 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 & CBD Distance Ring						
Panel A: 1980-2000						
Education, 2 km	-0.04	0.13	-0.02	-0.01	-0.11	-0.24
Education, 4 km	-0.03	0.10	-0.03	-0.01	-0.12	-0.19
Age, 4 km	0.00	0.10	-0.01	-0.04	-0.15	-0.18
Family Type, 4 km	0.06	0.09	-0.10	-0.06	-0.11	-0.17
Income, 4 km	0.00	0.08	-0.02	-0.01	-0.19	-0.20
Panel B: 2000-2010						
Education, 2 km	-0.02	0.05	0.04	0.00	0.02	-0.09
Education, 4 km	-0.02	0.04	0.01	0.00	-0.01	-0.09
Age, 4 km	0.00	0.04	0.01	-0.02	-0.02	-0.09
Family Type, 4 km	0.02	0.04	-0.01	-0.03	-0.03	-0.09
Income, 4 km	0.00	0.03	0.01	-0.01	-0.04	-0.08

Notes: Results are analogous to those in Table 2. The only difference is the order in which the counterfactuals are imposed.

Table A4: 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
1980-1990	0.17	0.12	1.42		Not Available	
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	-0.09

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 actual employment shocks for the 1990-2000 and 2000-2010 periods in Tables 4, 5 and 6, instrumented with variables whose summary statistics are reported in Panel B. For other periods, those tables report reduced form results. Statistics above are for the 120 CBSAs in the sample.

Table A5: Patterns of Housing Costs in Tracts within 4 km of CBDs

	1970-1980	1980-1990	1990-2000	2000-2010	1980-2010
Estimator	RF	RF	IV	IV	RF
Panel A: Difference Specification					
1(< 4 km to CBD)	-0.065	-0.029	-0.026	0.017	0.005
	(0.016)	(0.012)	(0.007)	(0.008)	(0.022)
Standardized CBSA Emp Growth	-0.046	0.013	0.072	0.015	0.037
X 1(< 4 km to CBD)	(0.013)	(0.013)	(0.084)	(0.019)	(0.028)
Standardized CBD Area Emp Growth	0.032	0.048	0.062	0.074	0.085
X 1(< 4 km to CBD)	(0.015)	(0.016)	(0.044)	(0.040)	(0.028)
Observations	31,011	35,704	37,096	36,715	35,078
R-Squared (First Stage F)	0.400	0.568	(21.5)	(61.8)	0.365
Panel B: AR(1) Specification					
1(< 4 km to CBD)	-0.061	-0.003	0.007	0.039	0.047
	(0.015)	(0.012)	(0.008)	(0.008)	(0.022)
Standardized CBSA Emp Growth	-0.042	0.000	-0.126	0.019	0.010
X 1(< 4 km to CBD)	(0.011)	(0.013)	(0.074)	(0.020)	(0.026)
Standardized CBD Area Emp Growth	0.037	0.024	0.146	0.030	0.072
X 1(< 4 km to CBD)	(0.014)	(0.013)	(0.040)	(0.040)	(0.026)
Observations	31,011	35,704	35,572	36,330	35,078
R-Squared (First Stage F)	0.033	0.009	(25.1)	(73.4)	0.024

Notes: Each column in each panel reports results from a separate regression of the change in tract owner occupied housing price index using the same specification as in Table 5. 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. See the notes to Table 5 for a description of variables and weights.