

School Desegregation, School Choice and Changes in Residential Location Patterns by Race

By Nathaniel Baum-Snow and Byron F. Lutz

Web Appendix

Data Appendix

Our sample is comprised of the 92 metropolitan areas (MSAs) with central school districts identified by Welch and Light (1987) as having a major court-ordered desegregation plan implemented between 1960 and 1990. We define central districts as those school districts that included the central business district of the largest census defined central city as of 1960 in each MSA nationwide. The sample includes all 56 central districts of over 50,000 students with minority enrollment between 20 and 80 percent in 1968 other than New York City, which did not have a major desegregation order. The remaining 36 central districts in our sample, which had enrollment over 15,000 and were between 10 and 90 percent minority in 1968, were randomly sampled with sampling weights proportional to enrollment and stratified by census region. Welch and Light investigated desegregation histories of 33 additional districts that we do not use because they do not contain the central business district of a MSA. We merge this information on major plan implementation year with district level enrollment data from the Common Core of Data and the data set used by Welch and Light from the Office of Civil Rights. The enrollment data is used to calculate dissimilarity and exposure indices.

Welch and Light (1987) report the year in which school desegregation was implemented for each school district. We observe only the year, not the month, of desegregation and must therefore make an assumption as to when in the year desegregation begins. Typically desegregation would have begun in the fall of the implementation year, meaning a desegregation plan implemented in 1970 would have taken force at the start of the 1970-1971 school year, though in some cases implementation may have begun earlier. In order to be conservative, we assume that desegregation begins at the start of the year. The census is mostly completed in late March with questions about school enrollment asking whether the individual has attended school at any time since February 1st. Therefore, implementations occurring in the same year as a census year would have had up to three months to have an effect on studied outcomes. In addition, outcomes may have been influenced by the announcement of impending desegregation. We choose this timing so as to capture the full potential response to desegregation. However, results are very similar if implementation is counted as taking hold beginning in the fall of the implementation year.

The data we use on dissimilarity and exposure indices start in 1967 for most districts with missing years scattered idiosyncratically across districts throughout the 1970s and 1980s. Reber (2005) demonstrates that these objects are persistent over time given the desegregation regime. To fill in some of the missing data we impute missing observations using the following procedure. We first assign each district/year to a desegregation regime based on implementation year. The school year starting in the implementation year and beyond is assigned to one category, the previous year to a second category and earlier years to a third category. Within desegregation category, we assign missing values to adjacent observations up to 3 years away.

Missing observations equidistant from two non-missing observations are imputed as the mean of the non-missing observations. This still leaves almost no data on the indices before the mid-1960s. However, based on evidence in Cascio et al. (2008), we assign school districts in the Confederacy a dissimilarity index of 1 in 1960. Similarly, we assign these districts an exposure index of 0 in 1960. These adjustments to the data significantly increase the sample size for the IV specifications in Table 6.

To construct demographic information on 1970-definition school districts, we compile census data from the tract, place, school district and county levels of aggregation for 1960, 1970, 1980 and 1990. We construct digital (GIS) maps of 1970 geography school districts using the 1969-1970 School District Geographic Reference File from the Census. This file indicates the fraction by population of each census tract that fell in each school district in the country. Those tracts split across school districts we allocate to the school district comprising the largest fraction of the tract's population. In 50 of our sampled central districts, there is no such allocation necessary. Using the resulting 1970 central school district digital maps, we allocate tracts in 1960, 1980 and 1990 to central school districts or suburbs based on the locations of their centroids. The 1970 definition central districts located in regions not tracted in 1970 all coincide with county geography which we use instead.

Accurate allocation of tracts in 1960, 1980 and 1990 to 1970 district geographies, built as amalgamations of 1970 definition census tracts, that did not coincide with counties required several steps. The reason is that tract geographies in periods other than 1970 sometimes include water that was not in the 1970 tract geography. Therefore, some tract centroids from other years are in regions that were not in a 1970 tract only because they are on the water. To handle this issue, we clip 1960, 1980 and 1990 tract geographies to the polygon formed by aggregating 1970 tract geography and recalculate centroids constrained to be within tracts before assigning tracts to 1970 definition central school districts. Tract data for suburban regions utilize this same clipped geography unless the clipping process reduces tract area by more than 90% in which case we use the original unclipped geography to calculate centroids.

Central district aggregate demographic data is built by aggregating tract data in each year except in two circumstances. If tract data did not exist or incompletely covered a district in 1960 or 1970 and it was a county district, we use census county aggregate data instead. The Lawton, OK and Amarillo, TX districts are the only two that were not fully tracted in 1960 and did not conform to county boundaries. 1960 demographic information for these districts are hand-entered from the printed 1960 census volume place data. Information on 1999 definition MSAs was built using county aggregates and New England County Metropolitan Areas for New England.

Census data from 1960 was the most challenging to compile and process. We obtain some census tract and county information from the National Historical Geographical Information Systems (NHGIS, nhgis.org), which compiles data from various electronic sources and has high quality geospatial information. The 1960 Census breaks out most variables of interest for whites and nonwhites but not blacks. Because blacks represented 92 percent of nonwhites in 1960 nationwide, we found it to be a reasonable approximation to measure black counts simply by rescaling nonwhite counts by the fraction of total nonwhite population in the tract or county that was black. The 1960 tract data, used for the spatial analysis, requires additional adjustments because this data set does not include school enrollment broken out by race. Instead it reports total enrollment and total public enrollment for elementary and high school separately. As an example, we impute tract public elementary enrollment counts for blacks as

$$(\text{total public elementary enrollment}) \left[\frac{\text{nonwhite population 5 - 14}}{\text{total population 5 - 14}} \right] \left[\frac{\text{total black population}}{\text{total nonwhite population}} \right]$$

To build white public enrollment counts we replace nonwhite and black populations with white population in the above expression. High school enrollment counts are calculated analogously using the 15-19 year old age group. (Census aggregate data from 1960 only includes age by race for 5 year intervals.) Because of the high levels of residential segregation in 1960, the assumption that each race in a census tract has the same propensity to send children to public and private school is not strong. Indeed, corroboration with county based enrollment counts by race reveals county estimates of public enrollment counts that are on average 5 percent greater than tract based estimates for blacks, likely because of lower private enrollment rates for blacks than whites. Central district median family income for 1960 is derived by assuming a uniform distribution within \$1,000 intervals for whites and nonwhites separately. We assign blacks the median nonwhite income. Districts not tracted in 1960 are assigned median family income as reported in printed census volumes.

County aggregates from 1960 on age by white/nonwhite we take from the NHGIS. County aggregate information on school enrollment by white/nonwhite and public/private we collect from published census tables. As with the 1960 tract data, we rescale all nonwhite counts by the ratio of total black population to nonwhite population to generate estimated black counts.

Census aggregates from other decades are taken from the Summary Tape File 4 tabulations. In each year after 1960, tract information is from STF4a. County information is from STF4c in 1970 and 1990 and STF4b in 1980. STF4 breaks out all counts by race such that no imputation adjustments are necessary as they were in 1960.

We use the census school district tabulation file to calculate the number of districts in each MSA in 1970.

Central business districts are taken from the 1982 Economic Census. The 1982 economic census reports the set of census tracts that local business people conceive of as being the central business district in each MSA. We take the centroid of the spatial aggregate of these tracts, which checks reveal corresponds to what is typically considered to be the downtown location in most cities.

Online Appendix Table A1 has a detailed description of each central district in our sample.

Table A1: Sample Districts and Attributes

| City of Central District | State of Central District | 1970 MSA Enrollment | 1970 Central District Enrollment | 1970 CD Public Enrollment | 1970 CD Public % Black | Desegregation Year | County District |
|--|---------------------------|---------------------|----------------------------------|---------------------------|------------------------|--------------------|------------------|
| Panel A: Large Central School Districts in Sample | | | | | | | |
| Birmingham | AL | 178 | 69 | 65 | 52% | 1970 | No |
| Mobile | AL | 99 | 84 | 72 | 42% | 1971 | Yes |
| Tucson | AZ | 83 | 56 | 53 | 4% | 1978 | No |
| Fresno | CA | 115 | 59 | 57 | 9% | 1978 | No |
| Los Angeles | CA | 1,538 | 698 | 614 | 23% | 1978 | No |
| Oakland | CA | 377 | 69 | 60 | 56% | 1966 | No |
| Sacramento | CA | 189 | 53 | 48 | 12% | 1976 | No |
| San Diego | CA | 299 | 132 | 122 | 12% | 1977 | No |
| San Francisco | CA | 279 | 108 | 84 | 26% | 1971 | Yes |
| Denver | CO | 270 | 104 | 89 | 14% | 1974 | Yes ^c |
| Wilmington* | DE | 110 | 18 | 14 | 73% | 1978 | No |
| Fort Lauderdale | FL | 122 | 122 | 110 | 21% | 1970 | Yes |
| Jacksonville | FL | 151 | 129 | 120 | 28% | 1971 | Yes |
| Lakeland | FL | 53 | 53 | 51 | 23% | 1969 | Yes |
| Miami | FL | 263 | 263 | 233 | 23% | 1970 | Yes |
| Orlando | FL | 126 | 84 | 79 | 18% | 1972 | Yes |
| Tampa | FL | 213 | 112 | 102 | 18% | 1971 | Yes |
| Titusville | FL | 63 | 62 | 58 | 11% | 1969 | Yes |
| West Palm Beach | FL | 72 | 72 | 64 | 27% | 1970 | Yes |
| Atlanta | GA | 411 | 95 | 88 | 64% | 1973 | No |
| Chicago | IL | 1,674 | 724 | 542 | 54% | 1982 | No |
| Indianapolis | IN | 305 | 110 | 98 | 35% | 1973 | No |
| Wichita | KS | 101 | 67 | 61 | 13% | 1971 | No |
| Louisville* | KY | 226 | 60 | 50 | 45% | 1975 | No |
| Baton Rouge | LA | 97 | 72 | 65 | 34% | 1970 | Yes |
| New Orleans | LA | 293 | 141 | 105 | 66% | 1961 | Yes |
| Shreveport | LA | 84 | 57 | 53 | 45% | 1969 | Yes |
| Boston | MA | 943 | 118 | 85 | 30% | 1974 | No |
| Baltimore | MD | 498 | 208 | 178 | 65% | 1974 | Yes |
| Detroit | MI | 1,131 | 322 | 261 | 63% | 1975 | No |
| Minneapolis | MN | 502 | 77 | 63 | 8% | 1974 | No |
| Kansas City | MO | 329 | 75 | 65 | 47% | 1977 | No |
| St. Louis | MO | 604 | 134 | 106 | 64% | 1980 | Yes |
| Charlotte | NC | 197 | 86 | 82 | 29% | 1970 | Yes |

Table A1: Sample Districts and Attributes, continued

| City of Central District | State of Central District | 1970 MSA Enrollment | 1970 Central District Enrollment | 1970 CD Public Enrollment | 1970 CD Public % Black | Desegregation Year | County District |
|---|---------------------------|---------------------|----------------------------------|---------------------------|------------------------|--------------------|-----------------|
| Panel A: Large Central School Districts in Sample, continued | | | | | | | |
| Omaha | NE | 143 | 78 | 58 | 18% | 1976 | No |
| Newark | NJ | 455 | 90 | 78 | 71% | 1961 | No |
| Las Vegas | NV | 71 | 64 | 61 | 12% | 1972 | Yes |
| Buffalo | NY | 326 | 97 | 69 | 37% | 1976 | No |
| Akron | OH | 166 | 63 | 54 | 26% | 1977 | No |
| Cincinnati | OH | 352 | 102 | 77 | 43% | 1973 | No |
| Cleveland | OH | 582 | 169 | 138 | 56% | 1979 | No |
| Columbus | OH | 263 | 110 | 96 | 27% | 1979 | No |
| Dayton | OH | 236 | 63 | 54 | 36% | 1976 | No |
| Toledo | OH | 144 | 75 | 57 | 26% | 1980 | No |
| Oklahoma City | OK | 168 | 74 | 69 | 21% | 1972 | No |
| Tulsa | OK | 125 | 77 | 73 | 13% | 1971 | No |
| Portland | OR | 251 | 81 | 70 | 8% | 1974 | No |
| Philadelphia | PA | 1,134 | 411 | 265 | 60% | 1978 | Yes |
| Pittsburgh | PA | 622 | 107 | 70 | 39% | 1980 | No |
| Charleston | SC | 88 | 63 | 56 | 45% | 1970 | Yes |
| Greenville | SC | 141 | 58 | 55 | 21% | 1970 | Yes |
| El Paso | TX | 101 | 63 | 60 | 3% | 1978 | No |
| Fort Worth | TX | 187 | 87 | 83 | 26% | 1973 | No |
| Houston | TX | 469 | 252 | 235 | 32% | 1971 | No |
| San Antonio | TX | 231 | 83 | 73 | 15% | 1969 | No |
| Norfolk | VA | 262 | 61 | 55 | 41% | 1970 | Yes |
| Seattle | WA | 342 | 95 | 82 | 12% | 1978 | No |
| Milwaukee | WI | 343 | 158 | 120 | 25% | 1976 | No |
| Panel B: Large Central School Districts Not in Sample | | | | | | | |
| Washington | DC | 754 | 153 | 135 | 93% | None | Yes |
| Albuquerque | NM | 103 | 86 | 80 | 3% | None | Yes |
| New York | NY | 1,746 | 1,468 | 1,088 | 35% | None | Yes |
| Panel C: Medium Sized Central School Districts in Sample | | | | | | | |
| Little Rock | AR | 87 | 28 | 26 | 38% | 1971 | No |
| San Bernardino | CA | 277 | 38 | 36 | 14% | 1978 | No |
| San Jose | CA | 268 | 37 | 35 | 2% | 1986 | No |
| Vallejo | CA | 58 | 14 | 13 | 25% | 1975 | No |

Table A1: Sample Districts and Attributes, continued

| City of Central District | State of Central District | 1970 MSA Enrollment | 1970 Central District Enrollment | 1970 CD Public Enrollment | 1970 CD Public % Black | Desegregation Year | County District |
|--|---------------------------|---------------------|----------------------------------|---------------------------|------------------------|--------------------|------------------|
| Panel C: Medium Sized Central School Districts in Sample, continued | | | | | | | |
| Hartford | CT | 242 | 31 | 26 | 46% | 1966 | No |
| Daytona Beach | FL | 33 | 32 | 30 | 22% | 1969 | Yes |
| Fort Myers | FL | 20 | 20 | 19 | 18% | 1969 | Yes |
| Albany | GA | 25 | 23 | 22 | 40% | 1980 | Yes |
| Columbus | GA | 58 | 40 | 39 | 31% | 1971 | Yes |
| Rockford | IL | 77 | 44 | 38 | 11% | 1973 | No |
| Fort Wayne | IN | 104 | 51 | 39 | 14% | 1971 | No |
| South Bend | IN | 58 | 41 | 34 | 16% | 1981 | No |
| Lexington-Fayette | KY | 64 | 36 | 34 | 16% | 1972 | Yes |
| Alexandria | LA | 30 | 30 | 28 | 32% | 1969 | Yes |
| Houma | LA | 40 | 21 | 20 | 17% | 1969 | Yes |
| Lake Charles | LA | 40 | 40 | 38 | 25% | 1969 | Yes |
| Springfield | MA | 133 | 37 | 28 | 20% | 1974 | No |
| Grand Rapids | MI | 205 | 48 | 33 | 21% | 1968 | No |
| Lansing | MI | 88 | 32 | 29 | 12% | 1972 | No |
| Wilmington | NJ | 25 | 19 | 18 | 29% | 1969 | Yes |
| Jersey | NY | 121 | 54 | 35 | 41% | 1976 | No |
| Rochester | NC | 237 | 56 | 41 | 33% | 1970 | No |
| Lawton | OK | 23 | 21 | 20 | 14% | 1973 | No |
| Columbia | SC | 74 | 42 | 40 | 43% | 1970 | No |
| Amarillo | TX | 34 | 30 | 28 | 6% | 1972 | No |
| Lubbock | TX | 42 | 34 | 33 | 11% | 1978 | No |
| Odessa | TX | 43 | 24 | 24 | 6% | 1982 | Yes |
| Waco | TX | 32 | 19 | 18 | 21% | 1973 | No |
| Roanoke | VA | 44 | 19 | 18 | 26% | 1970 | Yes ^C |
| Tacoma | WA | 95 | 39 | 35 | 9% | 1968 | No |

Note: Panel A lists information on central school districts that had enrollment exceeding 50,000 in 1968 as measured in Welch and Light (1987). Welch and Light use school district reported enrollments as of 1968, in contrast to the 1970 Census based enrollment measure reported here. Panel B lists information on those central districts with enrollment exceeding 50,000 in 1968 which were not included in the sample because they never had a major desegregation order or had very high or low minority enrollment. Panel C lists all remaining central districts in the sample as investigated by Welch and Light. These districts had enrollment between 15,000 and 50,000 in 1968 and were 10 to 90 percent black. All numbers are for 1970 district geographies. *District consolidated into one or more surrounding districts at some point after 1970. ^CThis district is or was comprised fully of one county that changed over time.

Table A2: Impacts of Desegregation on Total Enrollment

| | <u>ln(white total enrollment)</u> | <u>ln(black total enrollment)</u> |
|--------------------------------|---------------------------------------|---------------------------------------|
| | 1 | 2 |
| Panel A: National | | |
| Desegregated | -0.09** (0.04) | |
| Desegregated (5+) | | 0.12*** (0.03) |
| Panel B: Regional | | |
| (Deseg)*(South) | -0.12** (0.06) | |
| (Deseg)*(Non-South) | -0.02 (0.05) | |
| (Deseg 5+)*(South) | | -0.01 (0.04) |
| (Deseg 5+)*(Non-South) | | 0.18*** (0.04) |
| MSA & Year-South FE | X | X |

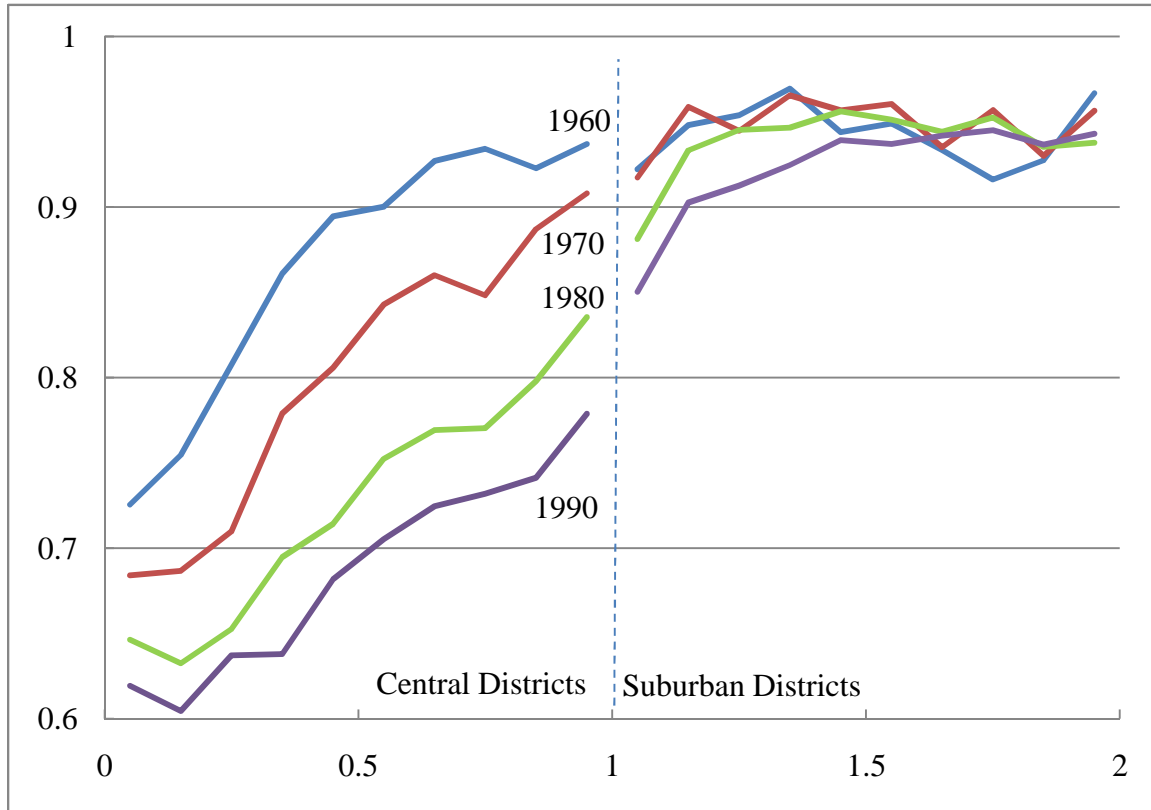
Note: See note to Table 2 for an explanation of the sample and variables.

Table A3: Impacts of Desegregation on MSA Level Outcomes

| | <u>ln(public enrollment)</u> | <u>ln(private enrollment)</u> | <u>ln(total population)</u> |
|--------------------------------|----------------------------------|-----------------------------------|---------------------------------|
| | 1 | 2 | 3 |
| Panel A: Whites | | | |
| (Deseg)*(South) | -0.02 (0.04) | 0.09 (0.07) | -0.04 (0.03) |
| (Deseg)*(Non-South) | 0.00 (0.03) | 0.08 (0.08) | 0.02 (0.03) |
| Panel B: Blacks | | | |
| (Deseg)*(South) | 0.00 (0.02) | -0.43*** (0.16) | 0.01 (0.02) |
| (Deseg)*(Non-South) | -0.13 (0.09) | -0.01 (0.18) | -0.10 (0.09) |
| (Deseg 5+)*(South) | -0.01 (0.03) | -0.33 (0.24) | -0.01 (0.04) |
| (Deseg 5+)*(Non-South) | 0.14** (0.03) | -0.06 (0.08) | 0.08** (0.03) |
| MSA & Year-South FE | X | X | X |

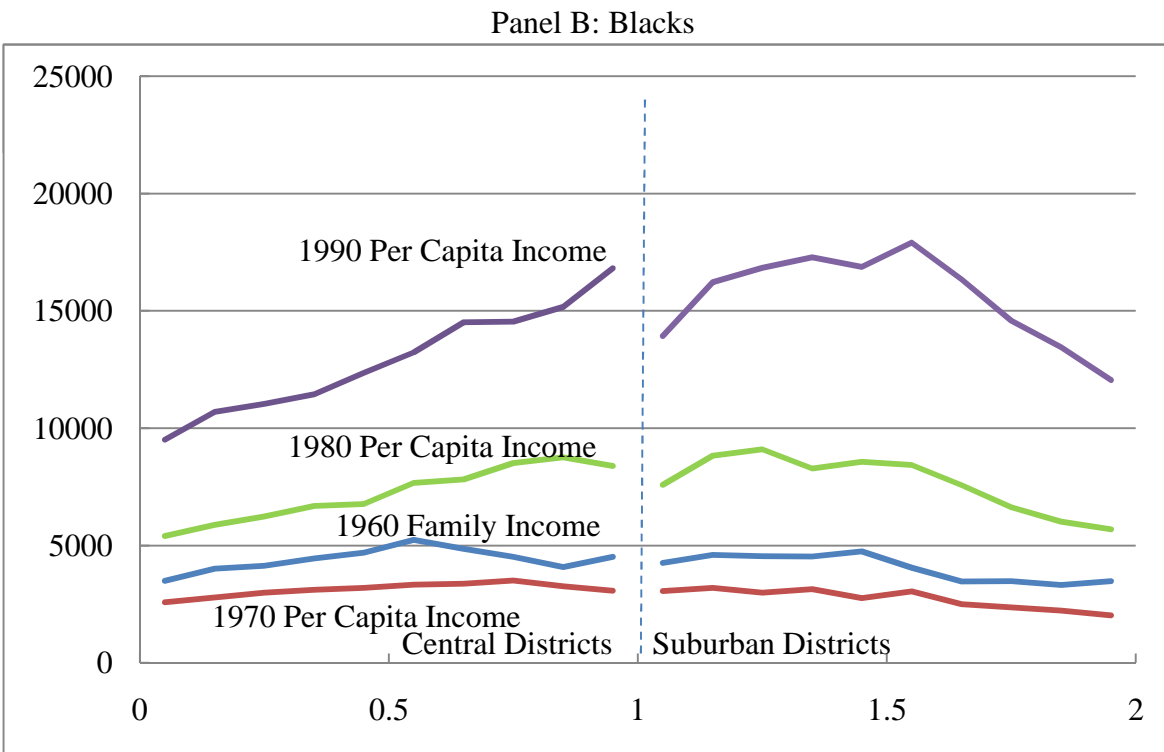
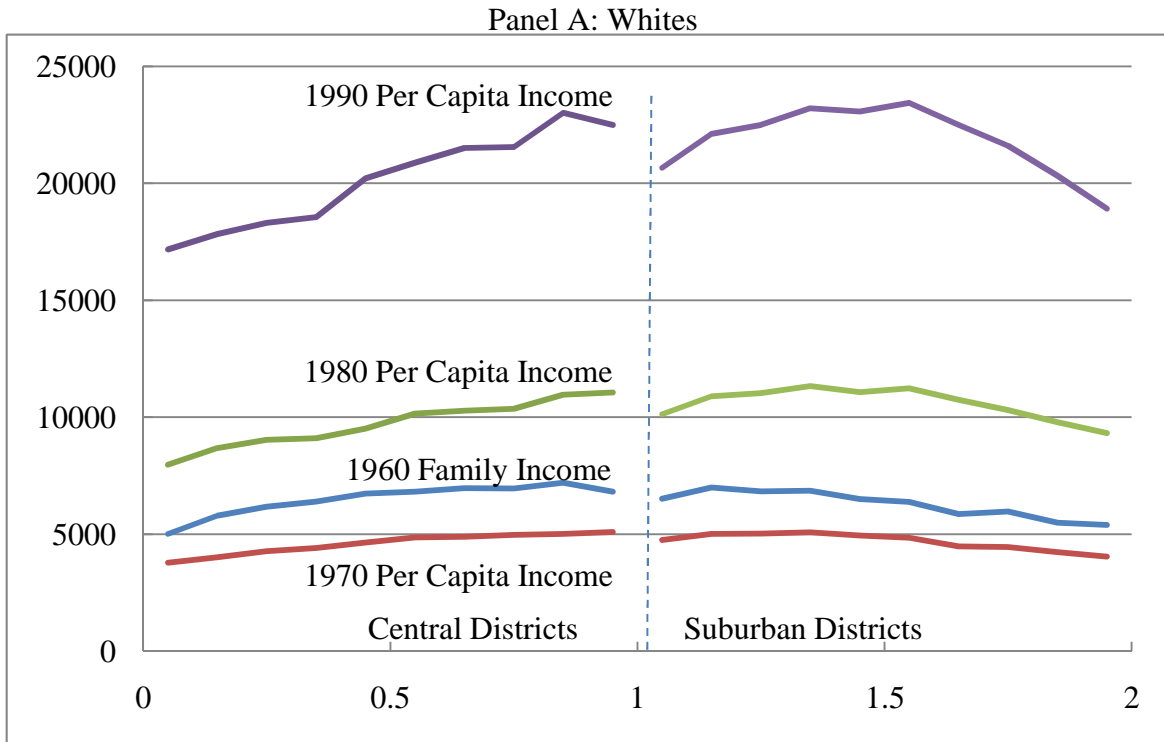
Note: The unit of observation is the MSA; otherwise the sample is the same as used in Table 2. See note to Table 2 for an explanation of the sample and variables.

Figure A1: Fraction White by Residential Location



Note: Graph shows the average ratio of residential white to white plus black population as a function of CBD distance across metropolitan areas in our sample for which census tract data are available. The sample includes the 64 metropolitan areas with central districts that were tracted in 1960 and experienced major desegregation orders. Metropolitan areas with fewer than 6 suburban tracts in any year are excluded. Each metropolitan area is weighted equally at all locations on the graph. The horizontal axis shows locations indexed as the cumulative distribution functions of 1990 population with respect to CBD distance inside and outside of central districts.

Figure A2: Income by Race and Residential Location



Note: Graphs show median family or per capita income by race as a function of CBD distance across metropolitan areas. See the notes to Figure A1 for explanations of the sample, distance metric and weighting.