The Geography of Opportunity over Time*

Catherine Massey
Welch Consulting†

Jonathan Rothbaum
U.S. Census Bureau‡

September 3, 2021

Abstract

Recent research finds that childhood neighborhoods affect adult economic outcomes, especially for children of low-income parents. Because places are shaped by both contemporary and historical factors, it is important to understand regional differences in opportunity both today and in the past. Using 1940 Census data linked to 1040 tax returns, we examine geographic differences in child outcomes experienced by cohorts born roughly 50 years apart – revealing how intergenerational persistence of status has changed over time both at the national level and at smaller geographic levels. We show that for White children of low-income parents, there has been moderate persistence in which places are associated with better and worse outcomes. For Black children, there has been very little persistence. We also show that the associations between many place characteristics (such as poverty and social capital) and better or worse outcomes are not robust to the inclusion of prior outcomes. This suggests that these associations are driven by omitted characteristics of those places.

Keywords: Intergenerational Mobility;
JEL Codes: J62, R00

*This work was supported by the US 2050 project, with funding from the Peter G. Peterson Foundation and the Ford Foundation. This report is released to inform interested parties of ongoing research and to encourage discussion. Any views expressed on statistical, methodological, technical, or operational issues are those of the author and not necessarily those of the US Census Bureau. The data in this paper has been cleared by the Census Bureau's Disclosure Review Board release authorization numbers CBDRB-FY20-037 and CBDRB-FY20-062.

†1716 Briarcrest Dr, Bryan, TX 77802 (email: cmassey@welchcon.com)

‡4600 Silver Hill Road, Washington, DC 20233 (email: jonathan.l.rothbaum@census.gov)

1 Introduction

A recent surge of intergenerational mobility research finds substantial geographic variation in mobility across the US for children entering the labor force today (Aaronson and Mazumder (2008), Chetty et al. (2014), Chetty et al. (2018), Chetty and Hendren (2018a), Chetty and Hendren (2018b)). These analyses show that mobility varies substantially both across and within regions, states, and counties. This work has also found that mobility is associated with place-based characteristics such as the percentage of the population that is Black, the share of single mothers, inequality measures, and income (Chetty et al., 2014). Work by Chetty and Hendren (2018a), Chetty and Hendren (2018b), and Chetty et al. (2018) aim to determine the extent to which these geographic differences are driven by a causal relationship between location and mobility, as well as which characteristics of a place are associated with more or less mobility. This body of research finds significant differences in mobility rates by neighborhood that are associated with several commuting-zone (CZ) level characteristics including the concentration of poverty, crime rates, school quality, and the prevalence of two-parent households.

Because neighborhood characteristics evolve over long periods of time, prior research suggests the need to examine trends in intergenerational mobility in historical settings as well as contemporary ones. We address this gap in the literature by examining geographic variation in mobility rates across the 20th century. Using linked census and administrative data, we examine intergenerational mobility by county and commuting zone (CZ) for individuals born between 1922 and 1940; producing mobility measures that are as consistent as possible with those available from Chetty et al. (2018) for the cohort born roughly 50 years later (1978-83).

Card et al. (2018) and Derenoncourt (2018) use the education of older children (14 to 17) in the 1940 Census to measure historical variation in child outcomes. Tan (2018) uses linked historical census data from 1940 and prior to evaluate spatial variation in mobility using occupation as a proxy for parent and child status. In this paper, we use the reported

income of the parents in the 1940 Census and measure child outcomes from income observed in linked tax returns from 1974 and 1979.

By comparing the outcomes of children born generations apart, we can assess the extent to which the geographic variation in opportunity found by Chetty et al. (2018) has changed over time, both overall and by race. Furthermore, we can further assess the extent to which location characteristics correlate with better and worse outcomes for children. We do this in two ways. First, we evaluate the robustness of current location characteristics in predicting better outcomes for children when conditioning on past outcomes. This helps us understand the extent to which these location characteristics may proxy for historical factors that could be the underlying cause of the variation in outcomes. Second, we evaluate how levels and changes in location characteristics observable in both periods correlate with levels and changes in child outcomes. In doing so, we hope to shed light on which location characteristics are more likely to be causing (or proxies for the causes of) better outcomes for children.

Overall and for White children, we find moderate persistence of outcomes over time. For children of low-income parents, the CZ-level correlation was 0.68 across the cohorts. However, for Black children, there is little persistence in the outcomes of children of low-income parents, as the CZ-level correlation is 0.12. This echoes the results in Derenoncourt (2018), In addition, certain regions such as the Pacific Northwest and Upper Midwest, experienced declines in outcomes for children of low-income families over time.

We also find that the cross-sectional relationships between location characteristics and child outcomes are substantially attenuated when we control for historical outcomes. However, as in Chetty et al. (2014), we find that the share of single mothers and the share of the population that is Black are both robust predictors of worse outcomes for children of low-income parents at both the CZ and county levels. Other location characteristics, such as poverty and measures of social capital are not robust predictors (and can even change sign) of child outcomes.

2 Data

We combine multiple data sources to measure intergenerational mobility in the past. To construct our sample, we link children from the 1940 full count census to their 1040 tax returns filed as adults in 1974 and 1979. We observe parent-child relationships, various measures of parent economic status (earnings, education, and occupation) as well as the location of residence for each household in the 1940 Census.

To estimate measures of intergenerational mobility and child outcomes, we need to observe measures of economic status for these children as adults. To observe the adult outcomes of children observed in the 1940 Census, we link them forward to the 1974 and 1979 1040 tax returns. This linkage relies on Person Identification Keys (PIK), assigned by the Census Bureau. PIKs are assigned by a probabilistic matching algorithm that compares characteristics of records in census and survey data to characteristics of records in a reference file constructed from the Social Security Administration (SSA) Numerical Identification System (or Numident) as well as other federal administrative data. The Numident is a record of every Social Security Number (SSN) ever issued by the SSA. Once assigned, the PIK uniquely identifies a particular person and is consistent for that person over time (PIKs correspond one-to-one with a specific SSN), allowing us to link individuals across other deidentified data sources using only the PIK.

The information used in the assignment of a PIK varies depending on the information available in the data. For administrative data such as the IRS 1040 forms, the assignment of PIKs relies on a simple merge by SSN with additional checks on name and birth date to confirm correct PIK assignment. Once the PIK is assigned to observations in the IRS data, the observation's SSN is removed to deidentify the data. Given the linkage procedure is based on SSNs, we believe these PIKs are accurate.

The assignment of PIKs to records in the 1940 Census relied on altered version of the Person Identification Validation System (PVS). PVS assigns PIKs by comparing characteristics of individuals in a census or survey to the characteristics of records in the reference file.

The reference file largely consists of the Numident and thus contains detailed information on name (and all name changes associated with a SSN), date of birth, parents' names, and city of birth. To assign PIKs, PVS creates the Cartesian product of the census data and the reference file and then compares and scores the similarity of characteristics of a census record and an administrative record, using a Jaro-Winkler string comparator to score the similarity between any two string variables, such as first, middle, and last name. Numerical variables, such as year of birth, are compared within a certain allowance, or band, around the year of birth (typically no more than two years). For numerical values that do not match exactly, but fall within a specified band, a prorated score is given to the characteristics. Once all the characteristics are scored, a total score for the record pair is created by summing over the scores of the characteristics. The highest scoring reference file-census pair for a census record is then assigned the PIK associated with the reference file record. See Wagner and Layne (2014) and Layne et al. (2014) for more information regarding the PIKing process.

Massey et al. (2018) adapted PVS to PIK the 1940 Census to account for the limited personally identifiable information available in the 1940 Census. Instead of full date of birth, only age is available, and, despite the availability of addresses in the 1940 Census, no known administrative data source contains addresses in the 1930s and 1940s that PVS could match the 1940 Census addresses to. Therefore, the PIKing process for the 1940 Census relies on full name, age on April 1, 1940, state or country of birth, and parents' names for individuals co-residing with their parents. Massey et al. (2018) provides a full description of the PIKing process for the 1940 Census.

Despite an overall PIK rate of 41 percent for the 1940 Census, 70 percent of children in the 1940 Census were successfully assigned a PIK. PIK rates were higher for children due to the use of parent's names in the linkage. The Numident contain parent's first and last names, which allowed Massey et al. (2018) to PIK the 1940 Census using parent's names (for children co-residing with parents, the sample we use in this paper) as additional information to score the similarity of records in the 1940 Census and the reference file.

There are multiple reasons why a child may not receive a PIK in the 1940 Census. Their information may not be unique enough on name, age, and state of birth information for PVS to confidently determine a match to a reference-file record. It is also possible that unmatched children never received a Social Security Number (due to death or entering occupations that did not initially require Social Security Numbers) and are not present in the reference file, which is built from the SSA Numident.

3 Methods

3.1 Measuring Status

For our measure of parent status, we use the earnings information provided in the 1940 Census. Individuals 14 and over were asked their wages and salary earnings as well as whether they had $50 or more of "sources other than money wages and salary." Figure 1 shows the 1940 Census enumeration form. Furthermore, this measure was top coded at $5,000 and enumerators were instructed not to "include the earning of businessmen, farmers, or professional persons derived from business profits, sale of corps, or fees" (IPUMS). Therefore, earnings is missing for a subset of parents.[1] There are nearly 41 million children 0-18 years old in the 1940 Census living with a parent. Of those, 65 percent have at least one parent who reported wage and salary earnings.

This presents several challenges to measuring intergenerational mobility. First, even for those parents with wage and salary earnings, we observe survey-reported earnings in only one year. Therefore, many of the known measurement issues in the intergenerational mobility literature will be present in this data, including life-cycle bias, bias from transitory shocks, survey-reporting error, etc.[2] Second, for a non-random subset of the population, such as

[1] Feigenbaum (2014) acknowledges this limitation and imputes income for the 1940 Census using the 1950 Census. In his analysis using income of parents observed in the 1915 Iowa Census and their children in the 1940 Census, Feigenbaum (2018) finds that using imputed income from the 1950 Census resulted in estimates that implied more mobility than those excluding those with missing earnings.

[2] See Black et al. (2011) for a discussion of measurement error and estimates of intergenerational mobility.

farmers and the self-employed, we do not observe some or all of their earnings, but instead only whether those earnings were $50 or more. Without other data on earnings and income for the parent cohort, these issues will likely attenuate measures of the persistence of status across generations.

Furthermore, because we are not able to assign PIKs to all children, we do not observe adult outcomes for all children in the sample. PIK assignment is not completely random, where observable characteristics such as race, education, and parent earnings are correlated with whether a PIK can be assigned to a given child (Massey et al., 2018).

We use Adjusted Gross Income (AGI) as our proxy for household income for adult children using their 1974 and 1979 tax returns. Pre-tax household income includes capital income, disability benefits, labor income, unemployment insurance, and Social Security income. In 1974 and 1979, our sample of children from the 1940 Census range in age from to 34 to 52 and 39 to 57, respectively.

3.2 Accounting for Sample Selection

Our analysis sample includes the children with assigned PIKs of parents with positive reported earnings. Without PIKs, we cannot link individuals to their tax returns to observe their income as adults. Because over 30 percent of children have parents without reported wage and salary earnings, we cannot easily include non-earners in our estimates. We attempt to control for non-random selection into wage and salary earnings and PIK assignment by using inverse probability weights (IPW) conditional on parent, child, and location characteristics. We run two IPW regressions using the predictors discussed below with the dependent variables: 1) child PIK assignment status and 2) conditional on child PIK assignment, the presence of positive parent wage and salary earnings.

For the IPW regressions, the dependent dummy variable was regressed on 1) child characteristics: race, gender, school attendance, and income as an adult (controlling for both an indicator for nonzero income and rank), 2) parent characteristics: age, citizenship, ed-

ucation, log of wage and salary earnings (child PIK dummy only), wage and salary and self-employment earnings dummies, number of children in the family, 3) location characteristics: urban/rural status of the household, and 4) location summary information at the county and CZ level. The CZ-level characteristics include child race shares, share of children with PIK, child age, child school attendance rates, average grade completed for 18-year-olds, share of foreign-born parents, parent wage and salary earnings, parent self-employment earnings dummy, parent education, parent age, share of farm households, and number of children per household.[3] Basic results showing which characteristics predict PIK assignment and parent earnings from the IPW regressions are shown in Table 1. Importantly, in the regression of the parent-earnings indicator on family and location characteristics, child outcomes are included in the model, including whether the child filed taxes and their adjusted gross income. Our IPW, therefore, adjusts the weights in a given location for parents with earnings that were observably similar to parents without reported wage earnings, conditional on both their observable characteristics in 1940 and the outcomes of their children.

Table 2 shows summary statistics for children and their parents in the 1940 Census. Column (1) shows demographic characteristics for all children aged 0-18 still living with their parents, as well as child-weighted parent characteristics for those children. The average child age in the sample is 9.2 years old, 11 percent of the sample is Black, with 89 percent White.[4] 74 percent of children in 1940 do not have a parent that graduated high school. We were able to assign PIKs to 70 percent of all children observed with their parents in the 1940 Census, for a sample of over 28 million linked children.

Column (2) in Table 2 shows the same summary statistics for children that are selected into our analysis sample, without inverse probability weights. From comparing columns (2) and (1), we see that the analysis sample is younger, more likely to be White, and has more

[3]A model selection step was done to restrict the number of variables to a smaller set as a part of the IPW process. The model also included dummies for missing values for each of the individual child and parent characteristics.

[4]Groups, such as Asian or American Indian and Alaskan Native, are difficult to reliably identify or make up less than 0.5 percent of the sample. Hispanics are also difficult to reliably identify in the 1940 census given the questionnaire, shown in Figure 1.

educated and younger parents. Column (3) shows the summary statistics after controlling for selection using inverse probability weights (IPW). Relative to the unweighted analysis sample, the IPW estimates are closer to the full sample ones for child age, race, most parent education levels, and parent age.

3.3 Measuring Mobility

We estimate mobility statistics at the national, CZ, and county levels using location observed in the 1940 Census. For CZs, we use the 1990 CZ definitions used by Chetty et al. (2018) to allow comparisons over time for consistent geographies, although we acknowledge that labor market areas have likely changed over this time period.

Our outcome of interest is the expected child rank in the distribution of income of all children in our sample. Child income is measured as average inflation-adjusted Adjusted Gross Income from 1040 returns in 1974 and 1979.[5] If a child is assigned a PIK, but is not present in the 1040 data in a given year, they are assigned an income of 0. We estimate this expected child rank for various subgroups in the data, including by parent earnings decile in 1940 and by race and native-born status of parents - Black, White, at least one foreign-born parent, and no foreign-born parents.

The expected child outcomes were estimated at the most disaggregated level, an example of which is expected child rank of Black sons, aged 0-6 in 1940, of parents in the 3rd parent earnings decile, in Marion County, Indiana. For disclosure protection, those estimates had $(\epsilon, \delta)-$differentially private Gaussian noise added (Nissim et al., 2007) with $\epsilon = 2$ in each estimate.[6] These baseline, disclosure protected estimates were aggregated up to other possible levels of aggregation, including at the CZ and national levels, pooled genders (aggregating sons and daughters), pooled race/native-born status of parents (aggregated across

[5] We used the CPI Research Series Using Current Methods (CPI-U-RS) to adjust income for inflation.

[6] As approved by the Disclosure Review Board of the U.S. Census Bureau in CBDRB-FY20-037. Because individuals can be in two estimates (by race and native-born status of parents) and in our estimates for 1) all individuals and 2) children born in and observed in the county in 1940 (non-movers), this yields the total approved $\epsilon = 8$. By releasing our estimates at only the most disaggregated level, we can have a higher value of ϵ for each released statistic, and as a result, add less noise to the data.

native-born status of parents to cover all children), and pooled ages.

As noted, our cohort of children is between the ages of 0-18 in 1940, and was born between 1921 and early 1940. We will compare this group to the cohort observed in Chetty et al. (2018), born between 1978 and 1983. We will call these two cohorts the 1940 and 2000 cohorts respectively. The results for the 2000 cohort are generated using parent and child income observed in the administrative tax data, which includes additional information returns, such as W-2s and 1099-Rs. We do not have as much information for the 1940 cohort. Because of this, we would expect measurement error to be greater in the 1940 cohort for both children and parents. For example, 1940 parent earnings are only observed from the survey, and only for wage and salary workers. Furthermore, the 1940 estimates use parent wage and salary earnings vs. parent income for the 2000 cohort, making the comparisons imperfect. As a result, we urge caution in directly comparing point estimates across samples and making statements about how intergenerational income/earnings mobility has changed over time. We focus primarily on relative outcomes across groups and locations and how they have changed over time rather than comparisons over time within the same group.

4 Results

4.1 National

With that caution in mind, Figure 2 shows the relationship between parent rank and child rank in the two cohorts overall and by race and nativity.

Figure 3 plots all of the within-cohort comparisons by race and parent nativity together. There are several relationships that are similar in both cohorts. First, at low income levels children of foreign-born parents have higher expected income ranks than White children or children with no foreign-born parent. Second, Black children have substantially lower expected ranks at every parent rank in both cohorts.

4.2 Child Outcomes by Place

Figure 4 shows the CZ-level expected child rank for children of parents in the third decile of the earnings distribution for the 1940 cohort in Panel A and the expected rank of children of parents at the 25th percentile of the income distribution for the 2000 cohort in Panel B.[7] For both cohorts, the maps look relatively similar, with worse outcomes particularly concentrated in the Southeast. Table 3 shows the unweighted and weighted correlation between the CZ- and county-level outcomes between the two cohorts. At the CZ level, the unweighted correlation is 0.62 and the weighted correlation is 0.68.[8]

Because of the aforementioned differences in the data, we focus our comparisons on the relative changes in outcomes across places. To do that, we standardize the expected child ranks to have mean zero and standard deviation one in each cohort. We then calculate the standardized difference in outcomes in each place, shown in Panel C of Figure 4. In Panel C, darker areas indicate areas where children experience less absolute mobility in 2000 than they did in 1940. In this map, we see that the expected child rank of children of low-income parents has gotten worse in the upper Midwest and along the West coast.

Some of the patterns in Panel C could be driven by sorting. The second wave of the Great Migration (1940 to 1960) resulted in many Blacks relocating from rural regions in the South to urban areas in the North. As Figure 3 shows, in both cohorts Black children have lower absolute mobility than White children, thus changes in the spatial distribution of Blacks can affect the spatial variation in mobility outcomes without affecting the spatial variation in outcomes for any individual race group. Furthermore, both Black and White children who participated in this migration may have experienced higher income in the North than their southern-staying counterparts (Alexander et al., 2017). Both factors could explain the improvement in absolute mobility we observe in the South from 1940 to 2000.

Figure 5 shows the spatial variation in outcomes of Black children of low-income parents

[7] For ease of exposition, we will call each of these groups children of low-income parents.

[8] The two are not statistically significantly different at the CZ or county level.

for the 1940 cohort in Panel A and the 2000 cohort in Panel B. In Panel C we show the standardized change in outcomes. For Blacks, outcomes for children of low-income parents got worse in CZs along the Mississippi River and in the Upper Midwest.[9]

Table 3 shows the correlation between outcomes at the CZ and county level for Blacks. The overall unweighted correlation is 0.62, but for Blacks, the point estimate is 0.05, and it is not statistically different from 0.[10] In other words, whereas there was a high degree of persistence in which places were associated with better and worse outcomes overall, there was no or very little persistence for Black children.

If we further disaggregate by gender, we see roughly the same persistence for sons and daughters overall.[11] However, when we look at Black sons and daughters, we do see differences. For Black sons, whether weighted or unweighted, we do not find a statistically significant correlation in outcomes over time. For Black daughters, the correlation is roughly 0.2, with or without weights.

Figure 6 shows the outcomes for White children of low-income parents in the 1940 and 2000 cohorts. Panel C shows the standardized change in outcomes. For White children, outcomes for children of low-income parents got worse in various locations, but with concentrations in the Pacific Northwest and the Upper Midwest.[12]

We show the same maps at the county level for the overall population (Figure 7), for Blacks (Figures 8), and for Whites (Figures 9). These maps reveal similar patterns as the CZ-level maps.

[9]Figure A.2 shows the same maps as in Panels A and B of Figure 5, but with the same cutoffs as Figure 4, rather than using cutoffs based on the variation in outcomes for Black children.

[10]If we weight by the number of children in the 1940 sample, the correlation is 0.12, as compared to 0.68 overall, and it is statistically different from 0.

[11]For sons and daughters respectively, 0.60 vs. 0.57 unweighted and 0.64 vs. 0.69 weighted, with none of the differences statistically significant.

[12]Figure A.3 shows the same maps as in Panels A and B of Figure 6, but with the same cutoffs as Figure 4, rather than using cutoffs based on the variation in outcomes for White children.

4.3 Relationship between Location Characteristics and Child Outcomes

As in Chetty et al. (2014), we evaluate the relationship between location characteristics and the outcomes of children of low-income parents to better understand the spatial variation in child outcomes. We take three approaches here:

1. Evaluate the correlation between location characteristics and child outcomes in both cohorts;
2. Use the 1940 cohort as a proxy for historical circumstances to help us understand which correlations between child outcomes and location characteristics in the 2000 cohort are more likely to be due to omitted historical circumstances, which may indicate a lower likelihood of being causally related to child outcomes;
3. Evaluate how changes in location characteristics are associated with changes in child outcomes over time.

4.3.1 Correlation between Location Characteristics and Child Outcomes in Both Cohorts

To allow comparisons with Chetty et al. (2014), we calculate, to the extent possible, many of the same statistics they use to evaluate correlations between location characteristics and child outcomes. These include the poverty rate, mean income, inequality (the Gini coefficient for the bottom 99 percent), the top 1 percent share, the share of people within the 25th and 75th percentiles, the share of single mothers, and the fraction of the population that is Black.[13]

[13]For poverty, we adjust the current poverty threshold for a two-parent, two-child family back to 1939 dollars. We then use a square-root equivalence scale to adjust the poverty line for different family sizes. We use our IPW-adjusted analysis sample to calculate poverty by location to control for how earnings were reported in 1940. For mean earnings/income, we calculate mean parent earnings for families where all parents were aged 18 to 64 in 1940, again from our analysis sample, whereas Chetty et al. (2014) use the mean household income, which is not available in 1940. For the Gini coefficient, top 1 percent share, and the share between the 25th and 75th percentile (which Chetty et al. (2014) call the share in the middle class), we also use the IPW-adjusted analysis sample using earnings. For the share of single mothers and the fraction

In Table 4, we show the results of regressions of child outcomes (for all, Black, and White children, respectively) on the location characteristics for the 1940 and 2000 cohorts separately weighted by location population. In each regression, the variables are standardized to have mean of zero and standard deviation of one, again, weighted by the population in each period. The results are also shown in Figure 10.

For all children, poverty, inequality, the share of single mothers, and the fraction Black are negatively correlated with child outcomes in both periods, whereas mean income is positively correlated with child outcomes. The direction of the relationships for the top 1 percent share is negative in the 1940 cohort and positive in the 2000 cohort. Two robust predictors of child outcomes in the 2000 cohort (as reported by Chetty et al. (2014)) are the share of single mothers and the fraction Black, which have coefficients that are not statistically different across the cohorts, whereas all of the others are statistically significantly different.[14]

When we run the same regressions on Black and White children, we see greater differences in the relationships between location characteristics and child outcomes over time. For both Black and White children, CZ-level poverty and inequality are negatively correlated with child outcomes in the 1940 cohort, but not in the 2000 cohort.[15] The relationship between Black and White child outcomes and the share between the 25th and 75th percentile changes sign from positive in the 1940 cohort to negative in the 2000 cohort. The negative correlation between Black and White child outcomes and 1) the share of single mothers and 2) the fraction Black is also lower in for the 2000 cohort than the 1940 cohort.

Black, we use the full unweighted 1940 sample.

[14] Although the coefficient for the share between the 25th and 75th percentiles are only statistically different at the 10 percent level.

[15] At the county level, shown in Table A.1, both are negatively correlated with Black child outcomes in both cohorts, but for poverty, the correlation is 50 percent smaller for the 2000 cohort, and for inequality, the coefficient is 67 percent smaller for the 2000 cohort. For White children at the county level, poverty is negatively correlated with child outcomes in both periods, but the coefficient declines by 79 percent, whereas for inequality, the sign changes.

4.3.2 Location Characteristics using 1940 as a Proxy for Historical Circumstances

Many location characteristics are correlated with better or worse child outcomes. To the extent possible, we would like to try to separate out those which are more likely to be related to contemporaneous causal factors from those that are more likely to be proxies for other factors that affect child outcomes, such as historical circumstances. To do that, we re-evaluate the regression relationships between outcomes in the 2000 cohort and location characteristics with and without the 1940 cohort outcomes, which serve as a proxy for those historical circumstances. To the extent that a given relationship holds after proxying for past outcomes, it is more likely to be associated with a factor directly related to child outcomes than to be related to current outcomes through an omitted historical factor.

In Table 5 and Figure 11, we evaluate both the characteristics from Table 4 as well as several additional characteristics from Chetty et al. (2014), including spatial mismatch (proxied by the share with a commute under 15 minutes), the high school drop out rate (controlling for income), and an index of social capital (Putnam, 1995).

We find that when child outcomes are regressed against each CZ-level characteristic individually, many of the relationships are attenuated substantially or change sign. Poverty and the top 1 percent share change sign when 1940 cohort outcomes are included. The coefficients for mean income and the social capital index are no longer statistically significant. The coefficients for share with a commute under 15 minutes, inequality, and the share between the 25th and 75th percentiles each decline by about 80 percent. Although the share of single parents and share of Blacks are less affected, they too decline, by 40 to 45 percent when including historical outcomes.

Using county-level characteristics, the results differ slightly, shown in Table A.2 and Figure A.7. For the two characteristics whose sign changes at the CZ level, the coefficient for poverty declines 40 percent, but does not change sign, and the coefficient for the top 1 percent share is no longer significant when controlling for 1940 cohort outcomes. For those

characteristics at the CZ level that are no longer significant when controlling for historical outcomes, the coefficient for mean income declines 47 percent, and the social capital index is no longer a significant predictor of child outcomes at the county level either. At the county level, the coefficients for share with a commute under 15 minutes, inequality, and the share between the 25th and 75th percentile decline by 87, 75, and 94 percent respectively. As before, among the most robust predictors of child outcomes, the share of single mothers and the fraction Black decline 43 and 29 percent, respectively, with the inclusion of 1940 cohort outcomes.

We run the same regressions on outcomes for Black and White children as well. For Black children, as noted above, we do not see strong associations between many CZ-level characteristics and Black child outcomes, with only share of single mothers and fraction Black significant at the 5 percent level. Both are attenuated when we control for historical outcomes. With the historical control, mean household income and the top 1 percent share become positively associated with Black child outcomes. Also, the share with incomes between the 25th and 75th percentiles becomes negatively associated with Black child outcomes. At the county level, the coefficient on share with a commute under 15 minutes becomes more negative and the social capital index becomes negatively related to child outcomes.

In the same regressions using White child outcomes, we again find that most relationships are not robust to the inclusion of the historical outcomes. For example, the sign of the coefficient changes for the Gini coefficient, the top 1 percent share, the share between the 25th and 75th percentiles, the fraction with a commute under 15 minutes, and the social capital index. For poverty, the relationship changes from not statistically significant to negative when controlling for past outcomes. The share of single mothers and the fraction Black are more robust to the inclusion of historical outcomes, but the coefficients for each still decline by about 70 percent. At the county level, the signs for the Gini coefficient, the top 1 percent share, the fraction with a commute under 15 minutes, and the social capital index also change sign. Poverty does not change sign, but is not statistically significantly

related to White child outcomes after controlling for historical outcomes, which is also the case for the high school dropout rate. The coefficients for the share of single mothers and the fraction Black decline by 59 and 56 percent respectively.

4.3.3 Changes in Location Characteristics and Changes in Child Outcomes

In our final approach to evaluating how changes in places over time relate to changes in child outcomes, we regress the change in child outcomes against changes in the location characteristics. For this, we use the characteristics we can calculate from the 1940 Census that are also available for the 2000 cohort, including the poverty rate, mean earnings/income, the Gini coefficient for the bottom 99 percent, the top 1 percent share, the share between the 25th and 75th percentile, the share of single mothers, and the fraction Black. The results are shown in Table 6 and Figure 12 at the CZ level.[16] Changes in each characteristic, with the exception of the share between the 25th and 75 percentiles, are associated with changes in child outcomes. Increases in poverty, inequality, the share of the top 1 percent, the share of single mothers and the fraction Black are all associated with declines in child outcomes. Only for mean income/earnings is a the relationship positive.

For Black children changes in mean income/earnings, inequality, and the top 1 percent share are positively associated with changes in child outcomes. However, increases in the share of single mothers and the fraction Black are associated declines in outcomes for Black children. At the county level, mean income/earnings is still positively associated with child outcomes. However, the coefficient for inequality is negative, and for the top 1 percent share, it is not statistically significant. At the county level, changes to poverty, the share between the 25th and 75th percentile, the share of single mothers and the fraction Black predict worse outcomes.

For White children at the CZ level, an increase in poverty is associated with a slight improvement in child outcomes. Increases in the share between the 25th and 75th percentile,

[16]For the same regressions at the county level, the results are in Table A.3 and Figure A.8.

the share with single mothers and the fraction Black are associated with declines in child outcomes. At the county level, an increase in poverty is associated with a slight decline in outcomes, as are increases in inequality, the top 1 percent share (at the 10 percent level), the share between the 25th and 75th percentiles, the share of single mothers and the fraction Black. County-level increases in mean income/earnings are associated with improvements in child outcomes for White children.

4.3.4 Summarizing the Location Results

With these various approaches, we hope to shed light on which location characteristics are more strongly and robustly predictive of child outcomes over time. We provide a basic summary of the results in Table 7, which brings together the results from Tables 4, 5, and 6 at the CZ level and Tables A.1, A.2, and A.3 at the county level.

At the CZ level, only the share of single mothers and fraction Black are significant predictors of worse child outcomes across all the analyses. This results holds for all subsamples, including for all, Black, and White children. This gives us more confidence that they are proxying for or capturing something more robust about how location characteristics affect child outcomes at the commuting zone level. Mean earnings/income also generally predicts better child outcomes, but not across all regressions and child subgroups.

At the county level, the same three characteristics are robust to the various analyses, with a higher share of single mothers or fraction Black associated with worse child outcomes and higher mean income/earnings associated with better child outcomes across all specifications and subgroups. At the county level, in nearly all of our specifications and subgroups, increased poverty is also associated with worse child outcomes.

These relationships do not imply causality. However, understanding correlations between variables, and how robust and persistent they are, is often a useful step in exploring potential causes, in this case of better and worse child outcomes.

5 Conclusion

This paper creates intergenerational mobility estimates at the county, CZ, and national level using linked census and administrative data for the population born between 1922 and 1940. These estimates were created to be consistent with recent work by Chetty et al. (2018), which produced mobility estimates for the 2000 cohort born between 1978 and 1983. Using our new 1940 cohort estimates combined with those from Chetty et al. (2018), we are able to determine how mobility varied by geography in the the 1940s as well as compare spatial patterns of mobility in the past to those observed today. We also examine the extent to which characteristics of a geographic unit predict mobility today and how it has changed over time.

We found that absolute mobility of Black children was lower at every parent decile for the 1940 cohort. This is consistent with results found for contemporary cohorts in Chetty et al. (2018) and Chetty et al. (2020). Similarly, mobility of children with foreign-born parents was higher than for children of native-born parents in both cohorts.

Consistent with Chetty et al. (2018), we observe significant spatial variation in child outcomes at the CZ and county level in 1940, with the South exhibiting the lowest levels of absolute mobility. We find that the South experienced significant improvements in absolute mobility for Black children in 1940 relative to 2000. White children in the Pacific Northwest, on the other hand, experienced a substantial reduction in mobility from 1940 to 2000, as well as both Black and White children in CZs near the Great Lakes.

We find that the share of single mothers and the fraction Black are the best predictors of lower absolute mobility rates for children in 1940 for all subsamples at both the CZ and county level. The fraction of single mothers and fraction Black also remain significant predictors of mobility today after controlling for child outcomes in 1940. In terms of changes over time, we find that increases in poverty, the percentage Black, the percentage of single mothers, the share in the top 1 percent, and the share between the 25th and 75th percentiles resulted in decreases in mobility from 1940 to 2000 at the CZ level.

References

Aaronson, D. and Mazumder, B. (2008). Intergenerational economic mobility in the united states, 1940 to 2000. *Journal of Human Resources*, 43(1):139–172.

Alexander, T., Leibbrand, C., Massey, C., and Tolnay, S. (2017). Second-generation outcomes of the great migration. *Demography*, 54(6):2249–2271.

Black, S. E., Devereux, P. J., et al. (2011). Recent developments in intergenerational mobility. *Handbook of Labor Economics*, 4:1487–1541.

Card, D., Domnisoru, C., and Taylor, L. (2018). The intergenerational transmission of human capital: Evidence from the golden age of upward mobility. *NBER Working Paper No. 25000, National Bureau of Economic Research*.

Chetty, R., Friedman, J. N., Hendren, N., Jones, M. R., and Porter, S. R. (2018). The opportunity atlas: Mapping the childhood roots of social mobility. *NBER Working Paper 25147*.

Chetty, R. and Hendren, N. (2018a). The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. *The Quarterly Journal of Economics*, 133(3):1107–1162.

Chetty, R. and Hendren, N. (2018b). The impacts of neighborhoods on intergenerational mobility ii: County-level estimates. *The Quarterly Journal of Economics*, 133(3):1163–1228.

Chetty, R., Hendren, N., Jones, M. R., and Porter, S. R. (Forthcoming, 2020). Race and economic opportunity in the united states: An intergenerational perspective. *Quarterly Journal of Economics*.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics*, 129(4):1553–1623.

Derenoncourt, E. (2018). Can you move to opportunity? evidence from the great migration. Technical report, Harvard University, mimeo.

Feigenbaum, J. (2014). A new old measure of intergenerational mobility: Iowa 1915 to 1940. *Unpublished Manuscript*.

Feigenbaum, J. (2018). Multiple measures of historical intergenerational mobility: Iowa 1915 to 1940. *Economic Journal*, 128(612):F446–F481.

IPUMS. 1940 census: Instructions to enumerators.

Layne, M., Wagner, D., and Rothhaas, C. (2014). Estimating record linkage false match rate for the person identification validation system. *Center for Administrative Records Research and Applications Working Paper*, 2.

Massey, C. G., Genedak, K., Alexander, J. T., Gardner, T., , and O'Hara, A. (2018). Linking the 1940 U.S. Census with Modern Data. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 51(4):246–257.

Nissim, K., Raskhodnikova, S., and Smith, A. (2007). *Smooth Sensitivity and Sampling in Private Data Analysis*. Proceedings of the 39th Annual ACM Symposium on Theory of Computing. STOC'07.

Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy*, 6(1):65–78.

Tan, H. R. (2018). A different land of opportunity: The geography of intergenerational mobility in the early 20th century us. Technical report, Working paper.

Wagner, D. and Layne, M. (2014). The person identification validation system: Applying the center for administrative records and research and applications' record linkage software. *Center for Administrative Records Research and Applications Working Paper*.

Figure 1: 1940 Census Enumeration Form

STATE | COUNTY | TOWNSHIP OR OTHER DIVISION OF COUNTY | INCORPORATED PLACE | WARD OF CITY | BLOCK NO. | UNINCORPORATED PLACE

1940 Federal Census

ENUMERATION DISTRICT NO. | SUPERVISOR'S DISTRICT NO. | SHEET NO.

ENUMERATED BY ME ON , 1940

, ENUMERATOR

INSTITUTION

Line No.	LOCATION: Street, Avenue, road, etc.	LOCATION: House Number	HOUSEHOLD DATA: No. of Household in order of visitation	HOUSEHOLD DATA: Home owned (O) or rented (R)	HOUSEHOLD DATA: Value of home or Monthly rental if rented.	HOUSEHOLD DATA: Farm? (Yes or No)	NAME: Name of each person whose *usual* place of residence on April 1, 1940, was in this household. BE SURE TO INCLUDE: 1. Persons temporarily absent from household. Write "Ab" after names of such persons. 2. Children under 1 year of age. Write "Infant" if child has not been given a first name. Enter (X) after name of person furnishing information.	RELATION: Relationship of this person to the head of the household, as wife, daughter, father, mother-in-law, grandson, lodger, lodger's wife, servant, hired hand, etc.	CODE (Leave Blank)	PERSONAL DESCRIPTION: Sex	PERSONAL DESCRIPTION: Color or Race	PERSONAL DESCRIPTION: Age at Last Birthday	PERSONAL DESCRIPTION: Marital Status	EDUCATION: Attended school or college at any time since March 1, 1940?	EDUCATION: Highest grade of school completed	CODE (Leave Blank)	PLACE OF BIRTH: If born in U.S. give state, territory or possession. If foreign born, give country in which birthplace was situated on Jan. 1, 1937. Distinguish: Canada-French from Canada-English and Irish Free State from Northern Ireland.	CODE (Leave Blank)	CITIZENSHIP: Citizenship of the foreign born	RESIDENCE, APRIL 1, 1935: **In what place did this person live on April 1, 1935?** For a person who lived in a different place, enter city or town, county, and State. City, town, or village having 2,600 or more inhabitants If less, enter "R."	RESIDENCE, APRIL 1, 1935: County	RESIDENCE, APRIL 1, 1935: State (or Territory or foreign country)	RESIDENCE, APRIL 1, 1935: On a Farm? (Y or N)	RESIDENCE, APRIL 1, 1935: CODE (Leave Blank)
	1	2	3	4	5	6	7	8	A	9	10	11	12	13	14	B	15	C	16	17	18	19	20	D
1.																								
2.																								
3.																								
4.																								
5.																								
6.																								
7.																								

PERSONS 14 YEARS OLD AND OVER – EMPLOYMENT STATUS

Line No.	Was this person AT WORK for pay or profit in private or nonemergency Govt. work during week of March 24-30? (Y or N)	If not, was he at work on, or assigned to, public EMERGENCY WORK (WPA, NYA, CCC, etc.) during week of March 24-30? (Y or N)	If neither at work nor assigned to public emergency work. ("No" in cols. 21 & 22): Was this person SEEKING WORK? (Y or N)	If neither at work nor assigned to public emergency work. ("No" in cols. 21 & 22): If not seeking work, did he HAVE A JOB, business, etc.? (Y or N)	For persons answering "No" to questions 21-24. Indicate whether engaged in home housework (H), in school (S), unable to work (U), or other (Ot).	CODE	If at private or nonemergency Govt. work. "Yes" in col. 21: Number of hours worked during week of March 24-30, 1940.	If seeking work or assigned to public emergency work. "Yes" in col. 22 or 23: Duration of unemployment up to March 30, 1940 – in weeks.	OCCUPATION, INDUSTRY, AND CLASS OF WORKER. For a person at work, assigned to public emergency work, or with a job ("Yes" in col. 21, 22, or 24), enter present occupation, industry, and class of worker. For a person seeking work ("Yes" in col. 23): (a) if he has previous work experience, enter last occupation, industry, and class of worker; or (b) if he does not have previous work experience, enter "New worker" in Col. 28, and leave Cols. 29-30 blank. OCCUPATION: Trade, profession, or particular kind of work, as – *Frame spinner, Salesman, Laborer, Rivet heater, Music teacher*	OCCUPATION, INDUSTRY, AND CLASS OF WORKER. INDUSTRY: Industry or business, as— *Cotton mill, Retail grocery, Farm, Shipyard, Public school*	Class of Worker	CODE (leave blank)	Number of weeks worked in 1939 (Equivalent full-time weeks)	INCOME IN 1939 (12 months ending Dec. 31, 1939.): Amount of money, wages or salary received (including commissions)	INCOME IN 1939 (12 months ending Dec. 31, 1939.): Did this person receive income of $50 or more from sources other than money wages or salary (Y or N)	Number of Farm Schedule
	21	22	23	24	25	E	26	27	28	29	30	F.	31	32	33	34
1.																
2.																
3.																
4.																
5.																
6.																
7.																

National Archives and Records Administration

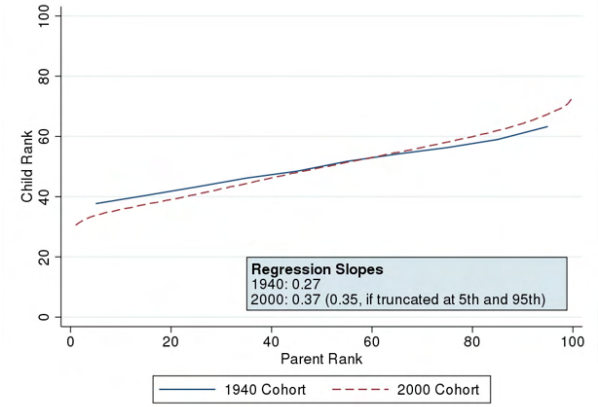
NARA's web site is http://www.archives.gov

NA 14129 (6-09)

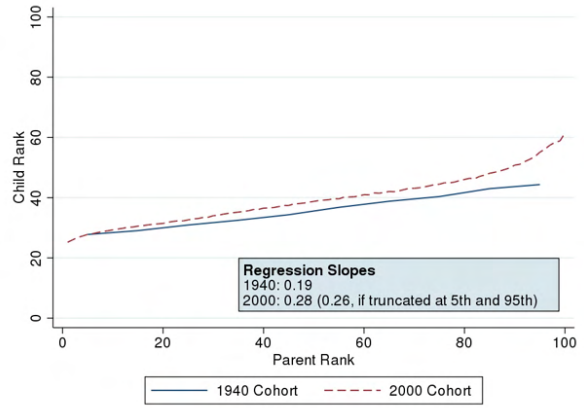
Source: National Archives https://www.archives.gov/files/research/census/1940/1940.pdf, accessed 5/15/19.

Figure 2: Comparison of Mobility over Time by Subgroup

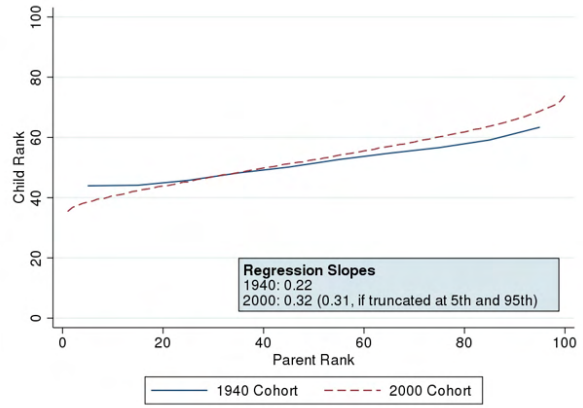
A. All Children



B. Black Children



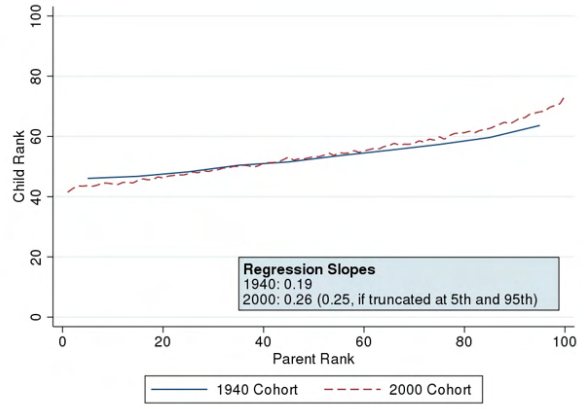
C. White Children



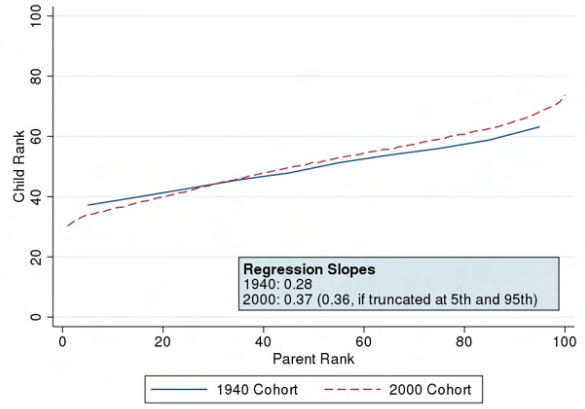
Notes: Each panel compares the expected child rank for children by parent earnings decile in the 1940 cohort and parent income percentile in the 2000 cohort. The 1940 cohort includes children born between 1921 and 1940 and observed in the 1940 Census. The 2000 cohort includes children born from 1978 and 1983 from Chetty et al. (2018). Panel A compares the outcomes of all children, Panel B compares the outcomes for Black children, and Panel C compares the outcomes for White children. Panel D compares children with a foreign born parent (1940 cohort) to children with a foreign-born mother (2000 cohort). Panel E compares children with no foreign-born parent (1940) to children with a native-born mother (2000 cohort). The slopes reported are from regressions of the decile parent-child ranks shown for the 1940 cohort and the percentile parent-child ranks shown for the 2000 cohorts.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

Figure 2: Comparison of Mobility over Time by Subgroup, continued

D. Children of Foreign-Born Parents

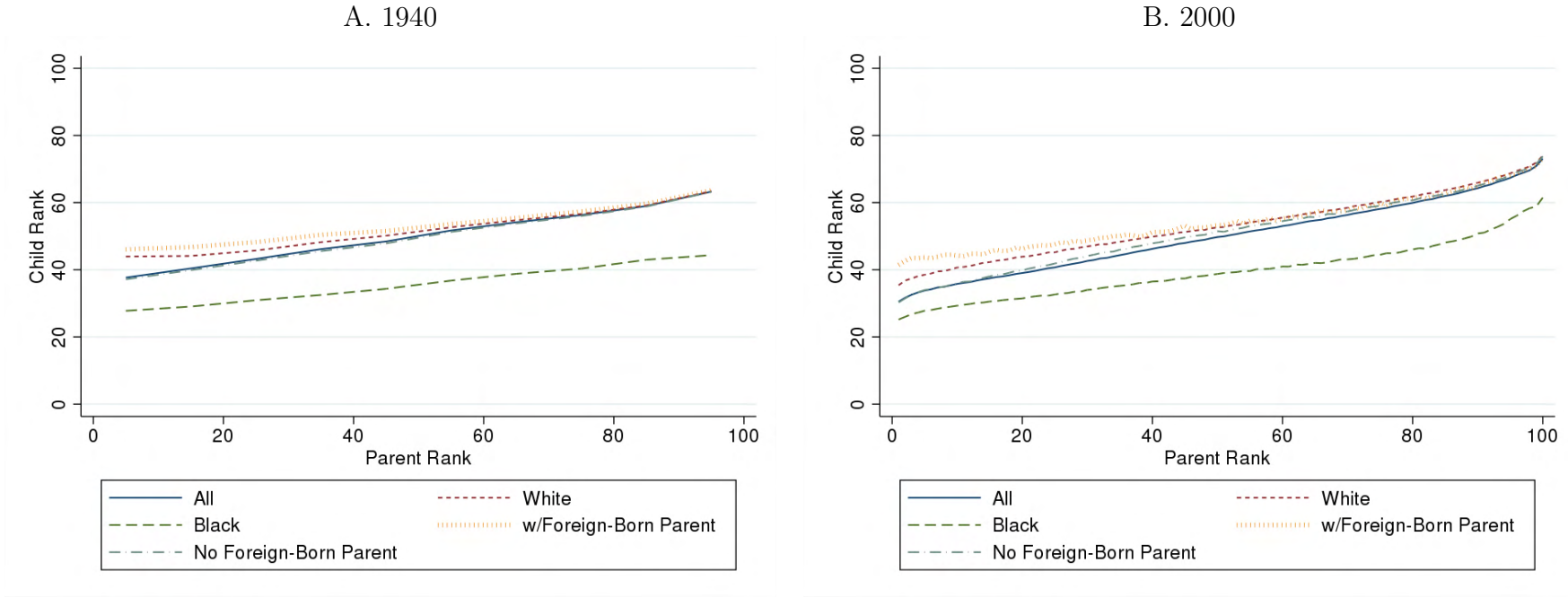


E. Children of Native-Born Parents



Notes: Each panel compares the expected child rank for children by parent earnings decile in the 1940 cohort and parent income percentile in the 2000 cohort. The 1940 cohort includes children born between 1921 and 1940 and observed in the 1940 Census. The 2000 cohort includes children born from 1978 and 1983 from Chetty et al. (2018). Panel A compares the outcomes of all children, Panel B compares the outcomes for Black children, and Panel C compares the outcomes for White children. Panel D compares children with a foreign born parent (1940 cohort) to children with a foreign-born mother (2000 cohort). Panel E compares children with no foreign-born parent (1940) to children with a native-born mother (2000 cohort). The slopes reported are from regressions of the decile parent-child ranks shown for the 1940 cohort and the percentile parent-child ranks shown for the 2000 cohorts.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

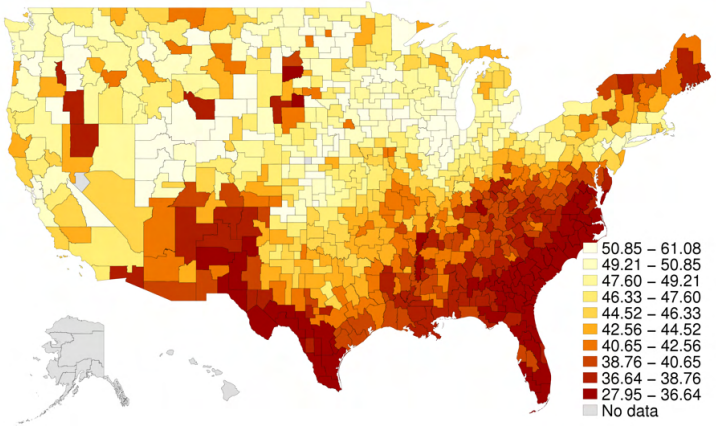
Figure 3: Mobility Comparisons by Race and Nativity of Parents



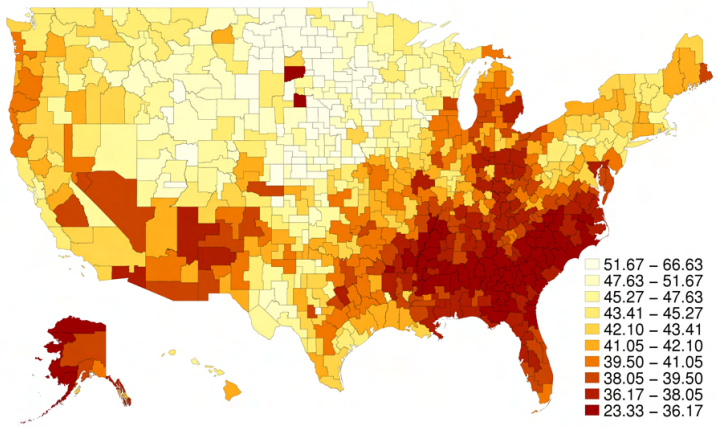
Notes: Each panel compares the expected child rank for children by parent earnings decile in the 1940 cohort and parent income percentile in the 2000 cohort. The 1940 cohort includes children born between 1921 and 1940 and observed in the 1940 Census. The 2000 cohort includes children born from 1978 and 1983 from Chetty et al. (2018). Panel A compares the outcomes by race and parent nativity in the 1940 cohort and Panel B does the same for the 2000 cohort.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

Figure 4: Spatial Variation in Child Outcomes over Time

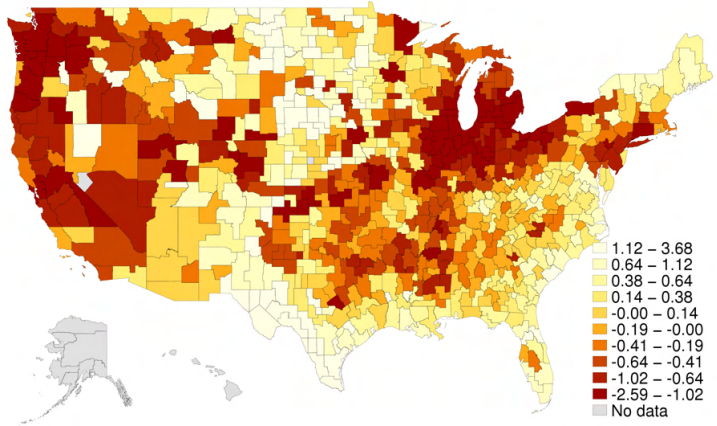
A. 1940 Cohort



B. 2000 Cohort



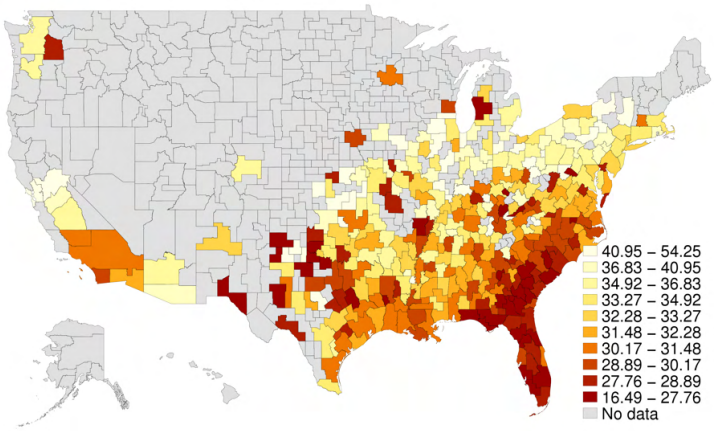
C. Standardized Change in Outcomes (2000-1940)



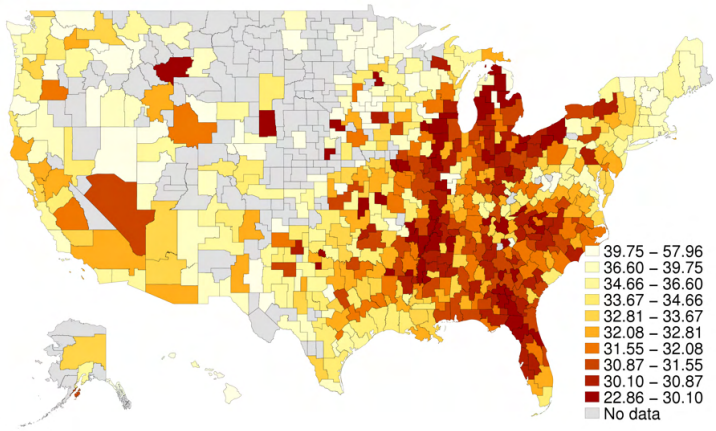
Notes: Panel A shows the expected outcome of children born between 1921 and 1940 and observed in the 1940 Census with parents in the 3rd decile of the parent earnings distribution at the CZ level. Panel B shows estimates of the expected outcome for children of parents with income at the 25th percentile of the distribution born from 1978 to 1983 from Chetty et al. (2018) at the CZ level. For Panel C, the values in A and B are standardized to have mean 0 and standard deviation 1 (in each period). Panel C maps the difference between the standardized outcomes, showing in which areas child outcomes have improved or declined relative to the average over time.

Figure 5: Spatial Variation in Child Outcomes over Time
Black Children

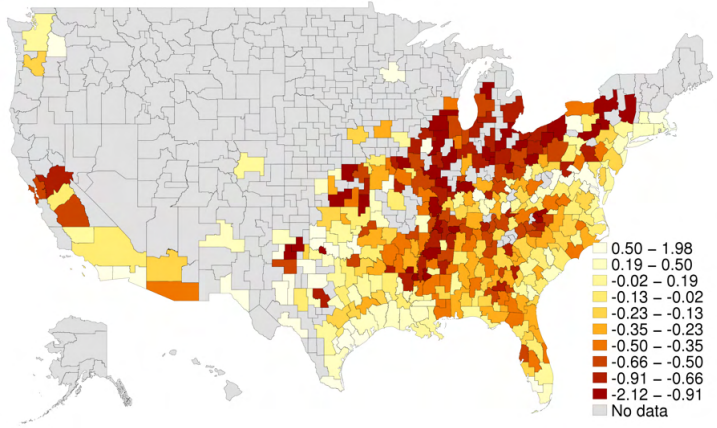
A. 1940 Cohort



B. 2000 Cohort



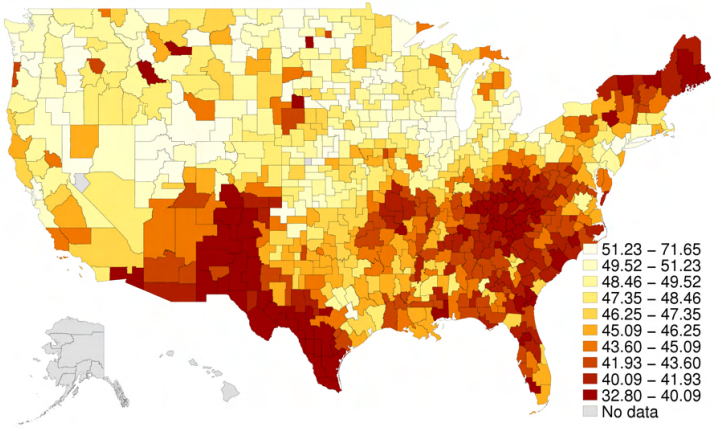
C. Standardized Change in Outcomes (2000-1940)



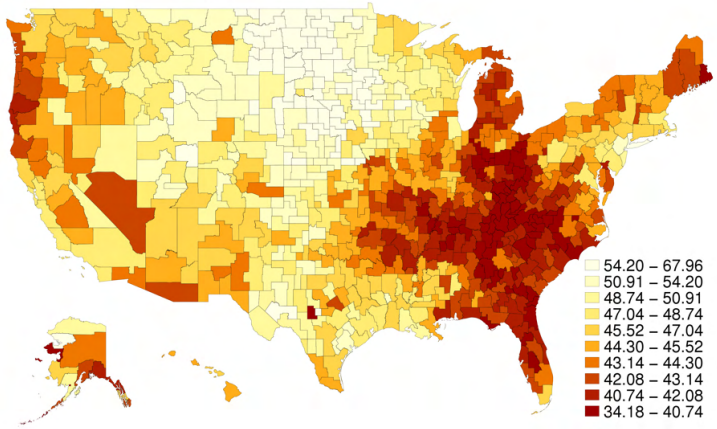
Notes: Panel A shows the expected outcome of Black children born between 1921 and 1940 and observed in the 1940 Census with parents in the 3rd decile of the parent earnings distribution at the CZ level. Panel B shows estimates of the expected outcome for Black children of parents with income at the 25th percentile of the distribution born from 1978 to 1983 from Chetty et al. (2018) at the CZ level. For Panel C, the values in A and B are standardized to have mean 0 and standard deviation 1 (in each period). Panel C maps the difference between the standardized outcomes, showing in which areas Black child outcomes have improved or declined relative to the average Black child outcome over time.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

Figure 6: Spatial Variation in Child Outcomes over Time, White Children

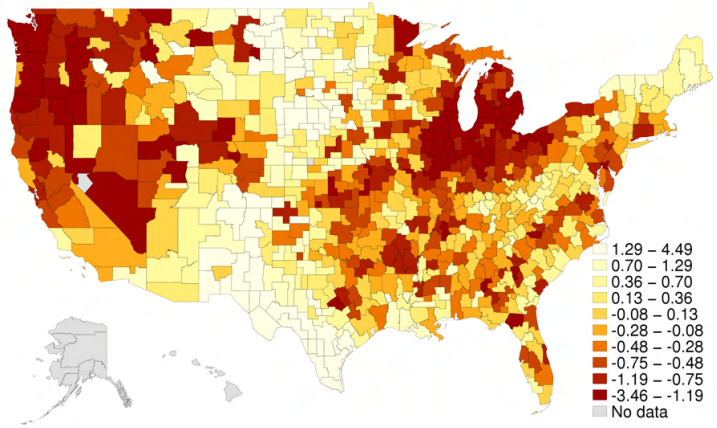
A. 1940 Cohort



B. 2000 Cohort



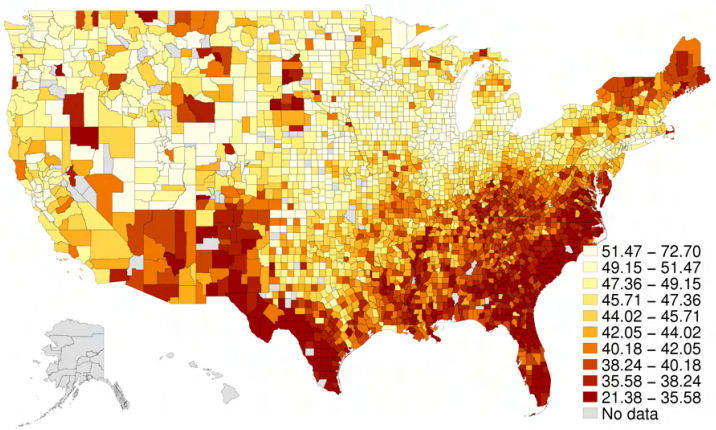
C. Standardized Change in Outcomes (2000-1940)



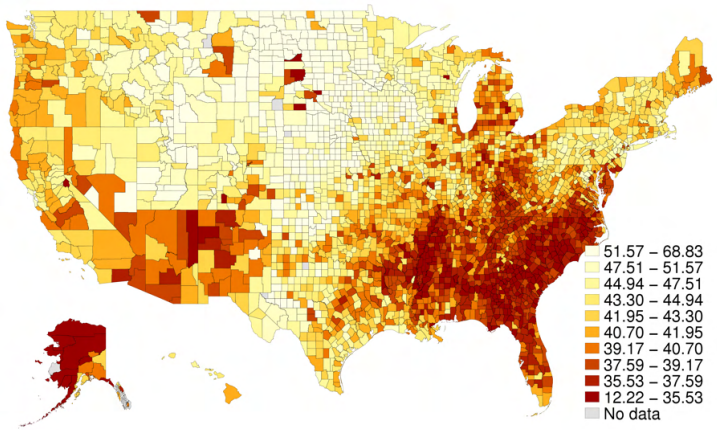
Notes: Panel A shows the expected outcome of White children born between 1921 and 1940 and observed in the 1940 Census with parents in the 3rd decile of the parent earnings distribution at the CZ level. Panel B shows estimates of the expected outcome for White children of parents with income at the 25th percentile of the distribution born from 1978 to 1983 from Chetty et al. (2018) at the CZ level. For Panel C, the values in A and B are standardized to have mean 0 and standard deviation 1 (in each period). Panel C maps the difference between the standardized outcomes, showing in which areas White child outcomes have improved or declined relative to the average White child outcome over time.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

Figure 7: County Level Spatial Variation in Child Outcomes over Time

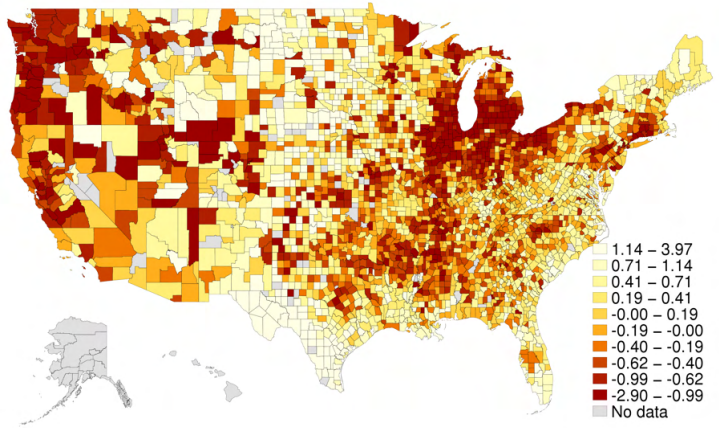
A. 1940 Cohort



B. 2000 Cohort



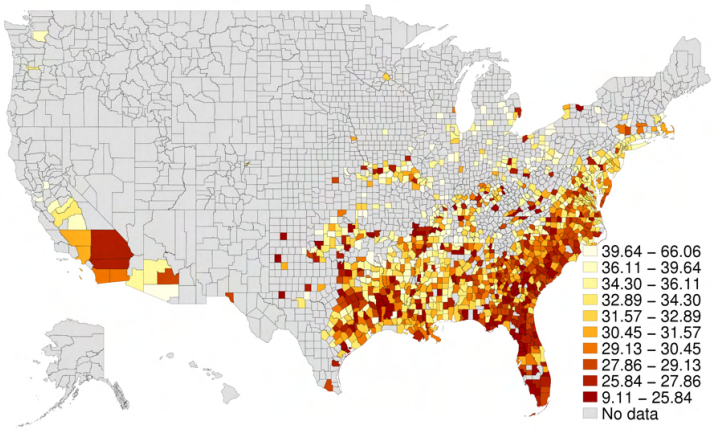
C. Standardized Change in Outcomes (2000-1940)



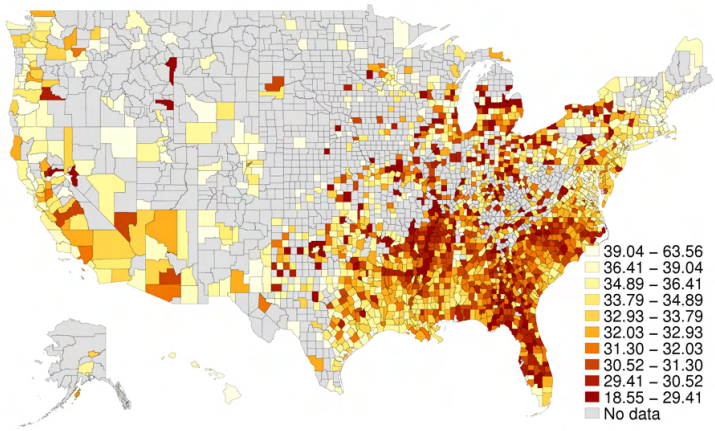
Notes: Panel A shows the expected outcome of children born between 1921 and 1940 and observed in the 1940 Census with parents in the 3rd decile of the parent earnings distribution at the county level. Panel B shows estimates of the expected outcome for children of parents with income at the 25th percentile of the distribution born from 1978 to 1983 from Chetty et al. (2018) at the county level. For Panel C, the values in A and B are standardized to have mean 0 and standard deviation 1 (in each period). Panel C maps the difference between the standardized outcomes, showing in which areas child outcomes have improved relative to the average and in which areas child outcomes have gotten worse over time.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

Figure 8: County Level Spatial Variation in Child Outcomes over Time
Black Children

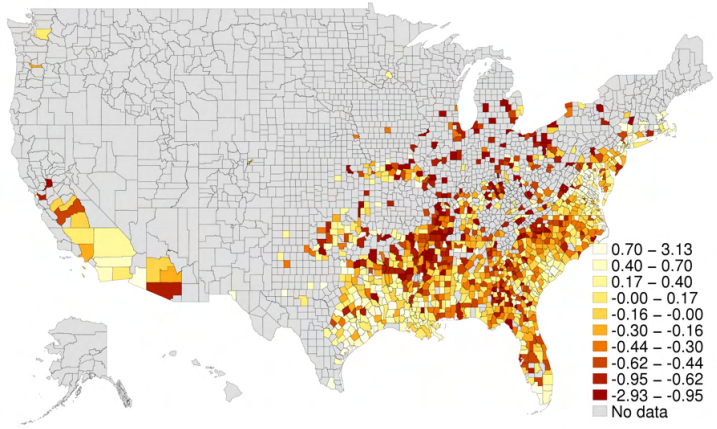
A. 1940 Cohort



B. 2000 Cohort



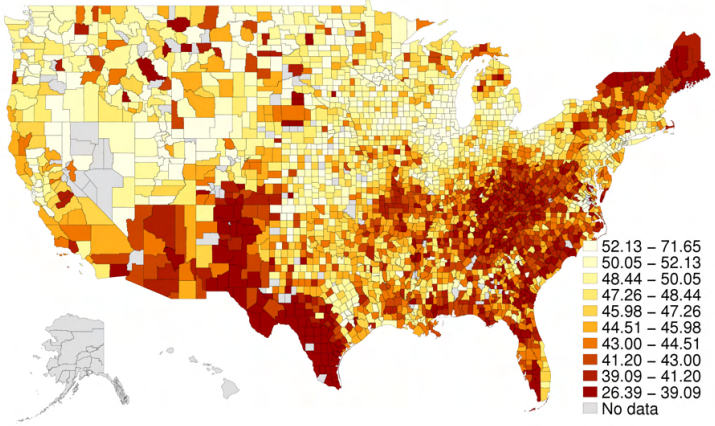
C. Standardized Change in Outcomes (2000-1940)



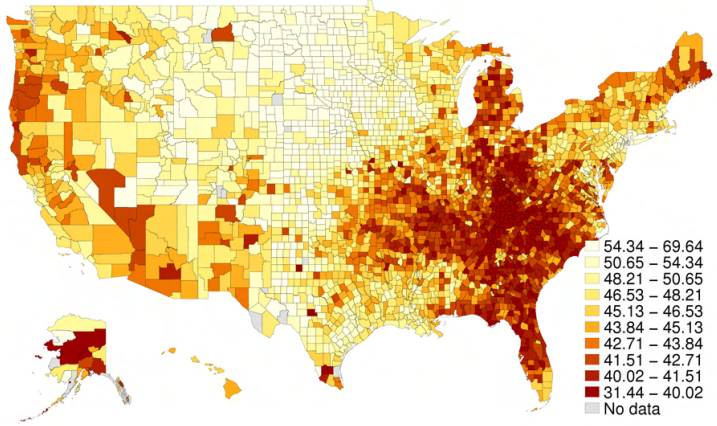
Notes: Panel A shows the expected outcome of Black children born between 1921 and 1940 and observed in the 1940 Census with parents in the 3rd decile of the parent earnings distribution at the county level. Panel B shows estimates of the expected outcome for Black children of parents with income at the 25th percentile of the distribution born from 1978 to 1983 from Chetty et al. (2018) at the county level. For Panel C, the values in A and B are standardized to have mean 0 and standard deviation 1 (in each period). Panel C maps the difference between the standardized outcomes, showing in which areas Black child outcomes have improved or declined relative to the average White child outcome over time.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

Figure 9: County Level Spatial Variation in Child Outcomes over Time
White Children

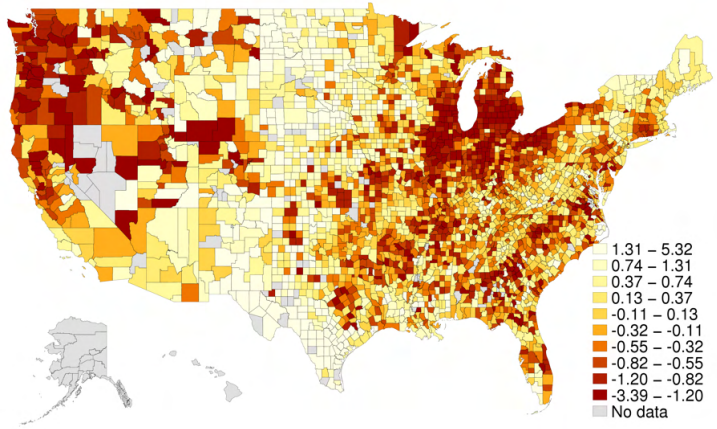
A. 1940 Cohort



B. 2000 Cohort

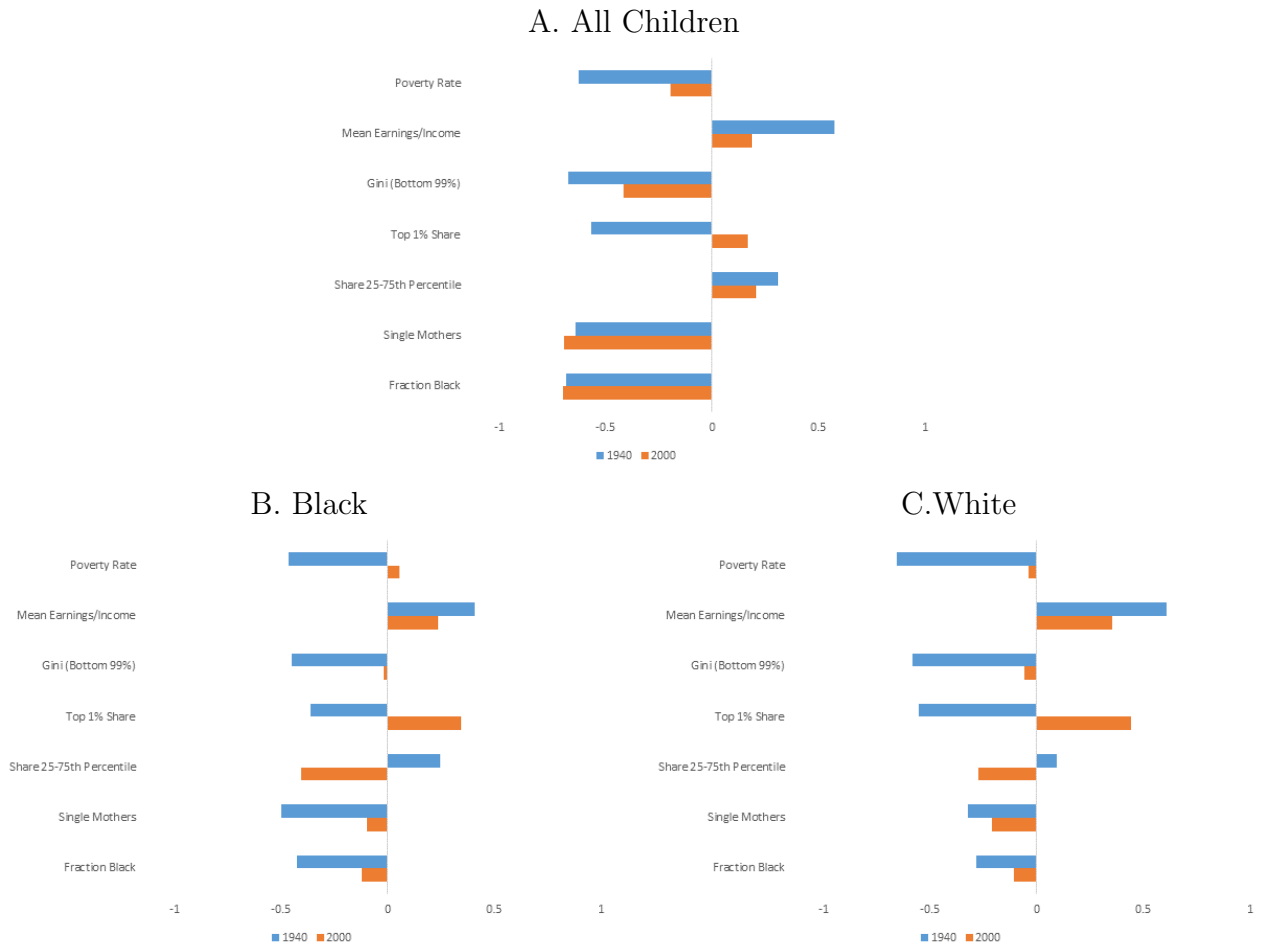


C. Standardized Change in Outcomes (2000-1940)



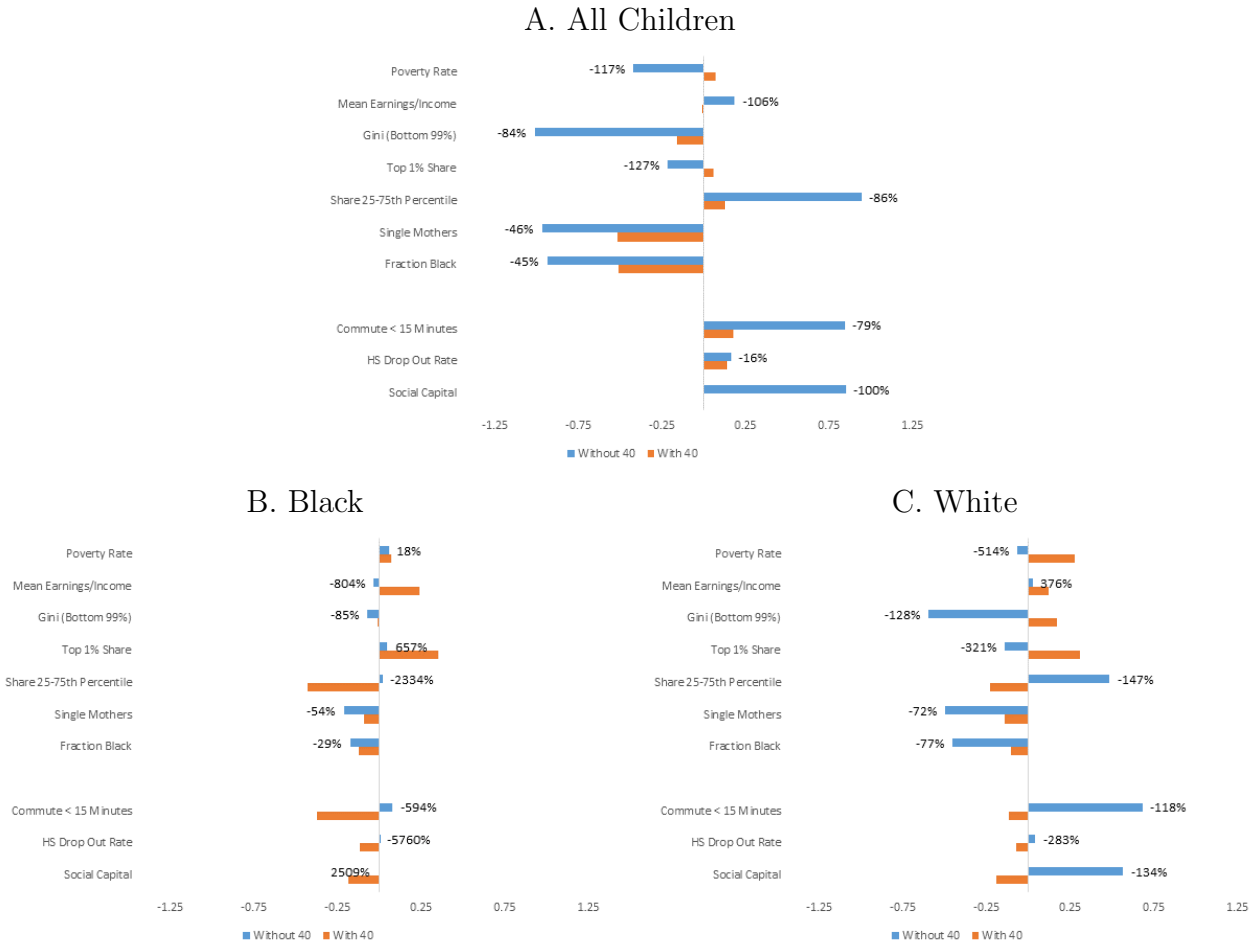
Notes: Panel A shows the expected outcome of White children born between 1921 and 1940 and observed in the 1940 Census with parents in the 3rd decile of the parent earnings distribution at the county level. Panel B shows estimates of the expected outcome for White children of parents with income at the 25th percentile of the distribution born from 1978 to 1983 from Chetty et al. (2018) at the county level. For Panel C, the values in A and B are standardized to have mean 0 and standard deviation 1 (in each period). Panel C maps the difference between the standardized outcomes, showing in which areas White child outcomes have improved or declined relative to the average White child outcome over time.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

Figure 10: Location Characteristics and Geographic Variation in Child Outcomes



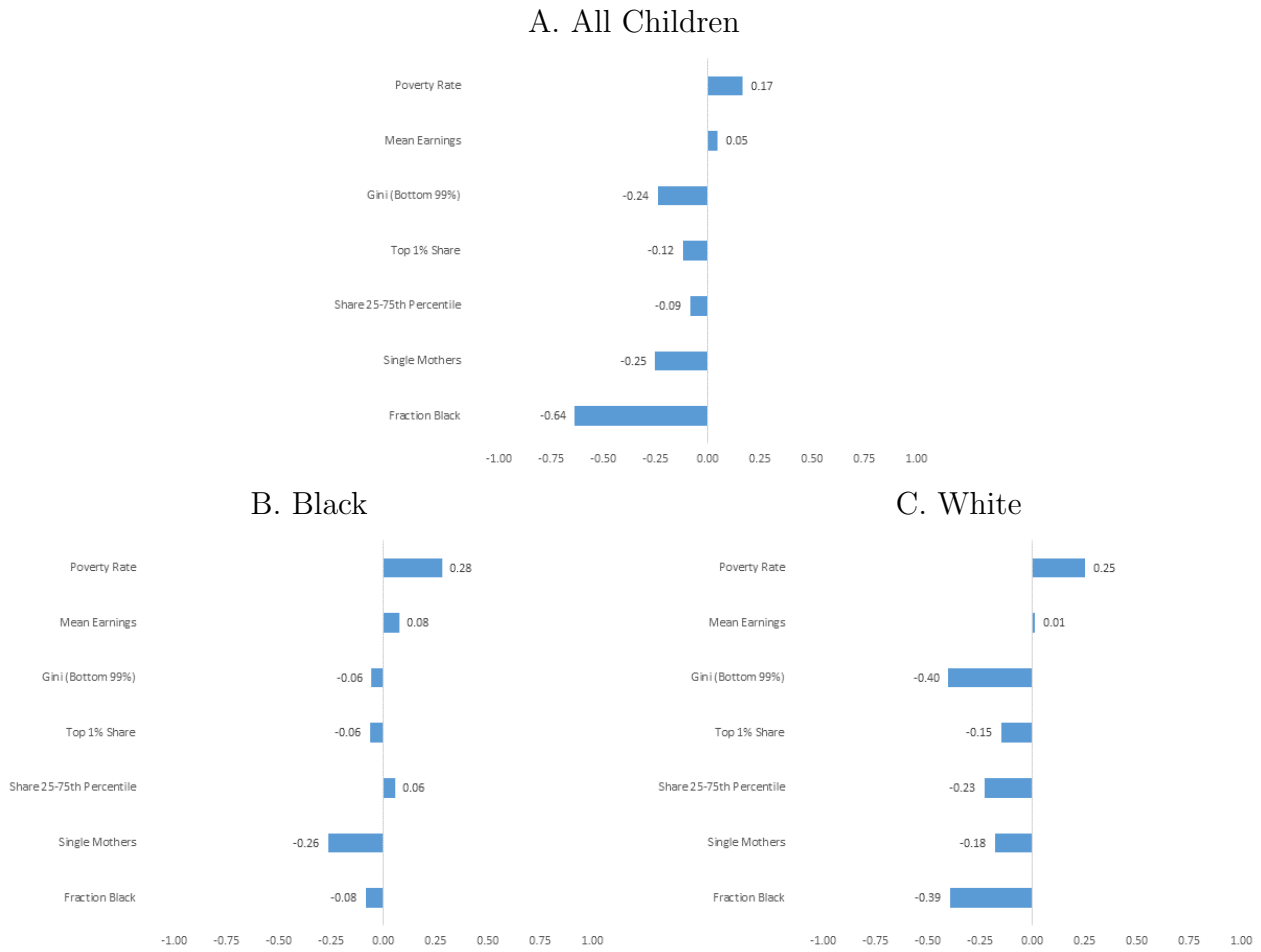
Notes: The results in this figure correspond to Table 4. This Figure shows the correlation (using population weights) between child outcomes and location characteristics at the commuting zone level in the 1940 and 2000 cohorts, for all children in Panel A, Black children in Panel B, and White children in Panel C.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018). 2000 cohort location characteristics from Chetty et al. (2014).

Figure 11: Location Characteristics and Geographic Variation in Child Outcomes for the 2000 Cohort, Controlling for the Past



Notes: The results in this figure correspond to Table 5. This figure shows the correlation between child outcomes in the 2000 cohort and location characteristics at the commuting zone level, controlling for and not controlling for child outcomes for the 1940 cohort. The first six characteristics shown for each group are the "Baseline Location Characteristics" in Panel A of Table 5, which are those that we can calculate from the 1940 Census. The next three characteristics shown are a subset studied in Chetty et al. (2014), including those they note as being particularly correlated with child outcomes. All variables are standardized to mean 0 and standard deviation 1. In the population-weighted regressions the standardization is done using population weights. The numbers indicate the extent to which the inclusion of the 1940 outcomes affect the regression coefficient between each location characteristic and outcomes for the 2000 cohort.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018). 2000 cohort location characteristics from Chetty et al. (2014).

Figure 12: Relationship between Changes in Location Characteristics and Changes in Child Outcomes



Notes: The results in this figure correspond to Table 6. This figure shows the results of regressing the change in expected child outcome from 1940 to 2000 on the change in various location characteristics at the commuting zone level. All variables are standardized to mean 0 and standard deviation 1. In the population-weighted regressions the standardization is done using population weights.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

Table 1: Inverse Probability Weight Predictors

Characteristic	Child PIK	Parent Earnings|Child PIK
Child		
Race		
Black	-	+
Native American	-	+
Asian	-	-
School Attendance	+	+
Female	+	+
1040 Filed		-
Income Rank in 1974		-
Parent		
Mother		
Foreign Born	-	-
Years of Schooling	+	-
Self-employment $\geq$ \$50	+	
Father		
Foreign Born	-	+
Years of Schooling	+	-
Self-employment $\geq$ \$50	+	
Parent Characteristics		
Average Years of Schooling	-	+
Positive wage and salary earnings	+	
Log average wage and salary earnings	+	
Number of Children	+	-
Household		
Urban	-	-
Farm	+	
Tjur R-2	0.06	0.16
N	40,600,000	28,280,000

Notes: This table shows the child and parent characteristics that predict whether the child was successfully assigned a PIK and conditional on PIK assignment, whether parents reported wage and salary earnings. The models include county and CZ summary statistics as well. A "+" indicates a characteristic is positively related to PIK assignment conditional on the other characteristics in the model, whereas a "-" indicates the relationship is negative. Given the predicted probability that the dependent variable $y = 1$, $\hat{P}$, the Tjur R-2 reports $E(\hat{P}|y = 1) - E(\hat{P}|y = 0)$.
Source: 1940 census linked to 1974 and 1979 1040 filings.

Table 2: Summary Statistics for Full and Analysis Sample

		Analysis Sample (Children with PIK and Parents with Earnings)		Ratio of Analysis Sample/Full Sample	
	Full Sample	Unweighted	Inverse Probablity Weighted	Unweighted	Inverse Probablity Weighted
	(1)	(2)	(3)	(4)	(5)
Child					
Age	9.2	8.7	9.0	0.95	0.98
	(0.0011)	(0.0015)	(0.0018)	(0.0002)	(0.0002)
Race					
Black	0.11	0.07	0.12	0.69	1.10
	(0.00010)	(0.00010)	(0.00022)	(0.0014)	(0.0023)
White	0.89	0.92	0.88	1.04	0.99
	(0.00010)	(0.00010)	(0.00023)	(0.0002)	(0.0003)
Native American	0.003	0.002	0.003	0.59	0.92
	(0.00002)	(0.00002)	(0.00003)	(0.0079)	(0.0119)
Asian	0.002	0.001	0.002	0.60	1.16
	(0.00001)	(0.00001)	(0.00004)	(0.0096)	(0.0207)
with PIK	0.70	1.00	1.00	1.43	1.43
	(0.00009)				
Parent					
Education (most educated)					
< High School	0.74	0.71	0.75	0.96	1.01
	(0.00012)	(0.00016)	(0.00018)	(0.0003)	(0.0003)
High School	0.15	0.17	0.15	1.18	1.01
	(0.00009)	(0.00013)	(0.00013)	(0.0011)	(0.0011)
Some College	0.06	0.06	0.06	1.13	0.99
	(0.00006)	(0.00008)	(0.00009)	(0.0018)	(0.0018)
Bachelor's	0.03	0.04	0.03	1.14	0.91
	(0.00004)	(0.00006)	(0.00005)	(0.0023)	(0.0021)
Professional/Graduate	0.01	0.01	0.01	1.07	0.84
	(0.00002)	(0.00003)	(0.00003)	(0.0041)	(0.0038)
Missing/Unknown	0.02	0.01	0.01	0.41	0.74
	(0.00004)	(0.00003)	(0.00008)	(0.0028)	(0.0047)
With Earnings	0.65	1.00	1.00	1.53	1.53
	(0.00014)				
Earnings	812	1,233	1,053	1.52	1.30
	(0.264)	(0.352)	(0.410)	(0.0005)	(0.0006)
Self-Employment > $50 (Father)	0.36	0.18	0.20	0.51	0.56
	(0.00014)	(0.00015)	(0.00019)	(0.0006)	(0.0007)
Self-Employment > $50 (Mother)	0.09	0.06	0.07	0.62	0.76
	(0.00008)	(0.00008)	(0.00016)	(0.0013)	(0.0019)
Age (Father)	36.5	35.6	35.8	0.97	0.98
	(0.0022)	(0.0029)	(0.0038)	(0.0001)	(0.0001)
Age (Mother)	40.9	39.7	40.0	0.97	0.98
	(0.0027)	(0.0032)	(0.0039)	(0.0001)	(0.0001)
Child Observations	40,600,000	19,110,000	19,110,000		

Notes: This table shows parent and child summary statistics for all children observed in the 1940 census as well as the analysis sample used in the paper with and without inverse probability weighting. The standard error of the mean estimate is shown in parenthesis.
Source: 1940 census linked to 1974 and 1979 1040 filings.

Table 3: Correlation in Child Outcomes over Time

A. CZ				B. County			
	Unweighted (1)	Weighted (2)	Number of CZs (3)		Unweighted (4)	Weighted (5)	Number of Counties (6)
All	0.62	0.68	720	Overall	0.56	0.57	3,018
	(0.03)	(0.04)			(0.01)	(0.02)	
Race				Race			
White	0.47	0.60	719	White	0.43	0.52	2,991
	(0.03)	(0.03)			(0.02)	(0.02)	
Black	0.05	0.12	362	Black	0.15	0.10	1,143
	(0.06)	(0.03)			(0.04)	(0.02)	
Foreign-Born Parent	0.06	0.24	327	Foreign-Born Parent	0.08	0.13	599
	(0.07)	(0.04)			(0.06)	(0.03)	
Gender-Race				Gender-Race			
Overall				Overall			
Male	0.60	0.64	715	Male	0.55	0.53	2,943
	(0.03)	(0.04)			(0.02)	(0.02)	
Female	0.57	0.69	713	Female	0.53	0.56	2,936
	(0.03)	(0.04)			(0.01)	(0.02)	
White				White			
Male	0.49	0.61	712	Male	0.44	0.51	2,884
	(0.03)	(0.04)			(0.02)	(0.02)	
Female	0.39	0.57	707	Female	0.39	0.47	2,886
	(0.03)	(0.03)			(0.02)	(0.02)	
Black				Black			
Male	-0.02	0.04	318	Male	0.09	0.06	899
	(0.08)	(0.04)			(0.05)	(0.03)	
Female	0.19	0.23	317	Female	0.29	0.18	898
	(0.06)	(0.04)			(0.05)	(0.02)	

Notes: This table shows the correlation between child outcomes in the 1940 and 2000 cohorts at the CZ and county levels. In the 1940 cohort, the outcome is expected child rank for children of parents in the third decile of the parent earnings distribution. In the 2000 cohort, the outcome is expected rank for children of parents at the 25th percentile of the parent income distribution. The expected ranks are standardized to mean 0 and standard deviation 1. In the population-weighted regressions the standardization is done using sample weights in the 1940 cohort and population weights in the 2000 cohort. The regression weights are the 1940 sample weights. Columns (1) and (4) show the regression coefficient of the unweighted standardized outcomes. Column (2) and (5) show the weighted regression coefficient of the weighted standardized outcomes. To be included in each regression, the location must have at least 20 children in the 1940 cohort for that subgroup. *Source:* For the 1940 cohort, 1940 census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

Table 4: Location Characteristics and Geographic Variation in Child Outcomes

	All		Black		White	
	1940	2000	1940	2000	1940	2000
	(1)	(2)	(3)	(4)	(5)	(6)
Poverty Rate	-0.62***	-0.19***	-0.46***	0.06	-0.65***	-0.04
	(0.03)	(0.04)	(0.04)	(0.05)	(0.03)	(0.04)
Mean Earnings/Income	0.57***	0.19***	0.41***	0.24***	0.61***	0.36***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)
Gini (Bottom 99%)	-0.68***	-0.41***	-0.45***	-0.02	-0.58***	-0.05
	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)
Top 1% Share	-0.57***	0.17***	-0.36***	0.35***	-0.55***	0.44***
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)
Share 25-75th Percentile	0.31***	0.21***	0.25***	-0.40***	0.09**	-0.27***
	(0.04)	(0.04)	(0.05)	(0.05)	(0.04)	(0.04)
Single Mothers	-0.64***	-0.69***	-0.50***	-0.10*	-0.32***	-0.21***
	(0.03)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)
Fraction Black	-0.69***	-0.70***	-0.42***	-0.12**	-0.28***	-0.11**
	(0.03)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)
Observations	692	692	362	362	692	692

Notes: This table shows the correlation between child outcomes and location characteristics at the commuting zone level in the 1940 and 2000 cohorts. In this table, we include the "baseline" location characteristics that we can calculate from the 1940 census and are available for the 2000 cohort. All variables are standardized to mean 0 and standard deviation 1, using population weights in each period. To be included in the 1940 regressions, the location must have at least 20 children in sample in the 1940 cohort. The regressions are run on the same set of locations in both periods. The outcome variable in the 1940 cohort regressions is expected child income rank for children of parents in the third earnings decile. In the 2000 cohort, the outcome is the expected child income rank for children of parents at the 25th percentile of the household income distribution. The outcomes for each subgroups are standardized to have mean 0 and standard deviation 1, using population weights as well. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.
Source: For the 1940 cohort, 1940 census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018). 2000 cohort location characteristics from Chetty et al. (2014).

Table 5: Location Characteristics and Geographic Variation in Child Outcomes for the 2000 Cohort, Controlling for the Past

A. Baseline Location Characteristics

	All			Black			White		
	Without 40	With 40	% Difference	Without 40	With 40	% Difference	Without 40	With 40	% Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Poverty Rate	-0.42***	0.07*	-116.5	0.06	0.07	17.9	-0.07	0.28***	-514.2
	(0.05)	(0.04)		(0.04)	(0.05)		(0.04)	(0.04)	
Mean Earnings/Income	0.18***	-0.01	-105.6	-0.03	0.24***	-804.1	0.02	0.12***	375.8
	(0.07)	(0.03)		(0.04)	(0.04)		(0.06)	(0.03)	
Gini (Bottom 99%)	-1.01***	-0.16***	-84.4	-0.07*	-0.01	-84.8	-0.60***	0.17***	-128.4
	(0.04)	(0.04)		(0.04)	(0.05)		(0.04)	(0.03)	
Top 1% Share	-0.22***	0.06**	-126.7	0.05	0.35***	657.0	-0.14***	0.31***	-321.4
	(0.05)	(0.03)		(0.05)	(0.03)		(0.04)	(0.03)	
Share 25-75th Percentile	0.94***	0.13***	-86.4	0.02	-0.43***	-2334.0	0.49***	-0.23***	-147.0
	(0.04)	(0.03)		(0.04)	(0.05)		(0.04)	(0.03)	
Single Mothers	-0.97***	-0.52***	-46.5	-0.21***	-0.10*	-54.5	-0.50***	-0.14***	-71.5
	(0.03)	(0.03)		(0.03)	(0.05)		(0.04)	(0.04)	
Fraction Black	-0.94***	-0.51***	-45.3	-0.18***	-0.13**	-28.6	-0.45***	-0.10***	-77.2
	(0.05)	(0.04)		(0.03)	(0.05)		(0.05)	(0.04)	
Observations	692	692		362	362		692	692	

B. Location Characteristics from Chetty, Hendren, Klein, and Saez (2014)

	All			Black			White		
	Without 40	With 40	% Difference	Without 40	With 40	% Difference	Without 40	With 40	% Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Commute < 15 Minutes	0.84***	0.18***	-78.9	0.08	-0.37***	-594.0	0.68***	-0.12***	-117.7
	(0.04)	(0.04)		(0.05)	(0.06)		(0.04)	(0.04)	
Gini (Bottom 99%)	-0.92***	-0.13***	-85.7	-0.09**	-0.01	-86.4	-0.51***	0.19***	-137.5
	(0.04)	(0.04)		(0.04)	(0.05)		(0.04)	(0.03)	
HS Drop Out Rate	0.16***	0.14***	-16.0	0.00	-0.12**	-5760.0	0.04	-0.08*	-282.9
	(0.05)	(0.04)		(0.04)	(0.05)		(0.05)	(0.04)	
Social Capital	0.85***	0.00	-100.0	-0.01	-0.19***	2509.0	0.56***	-0.19***	-133.6
	(0.04)	(0.04)		(0.04)	(0.06)		(0.04)	(0.04)	
Single Mothers	-0.87***	-0.49***	-43.4	-0.20***	-0.09	-53.0	-0.39***	-0.12***	-69.6
	(0.04)	(0.04)		(0.03)	(0.06)		(0.04)	(0.04)	
Fraction Black	-0.84***	-0.50***	-41.0	-0.16***	-0.12**	-22.5	-0.37***	-0.09**	-76.1
	(0.04)	(0.04)		(0.03)	(0.05)		(0.05)	(0.04)	
Observations	573	573		335	335		573	573	

Notes: This table shows the relationship between child outcomes in the 2000 cohort and location characteristics at the commuting zone level, controlling for and not controlling for child outcomes for the 1940 cohort. The "Baseline Location Characteristics" in Panel A are those that we can calculate from the 1940 census. The characteristics in Panel B include a subset studied in Chetty et al. (2014), including those they note as being predictive of child outcomes. All variables are standardized to mean 0 and standard deviation 1, using population weights in each period. The expected rank of a child at the 25th percentile in the 2000 cohort (standardized as well using 2000 population weights) is regressed against each location characteristic individually. "Without 40" and "With 40" indicate whether the the regrssion included as a predictor the expected child rank for children from households in the third decile of the parent earnings distribution in the 1940 cohort (standardized using 1940 population weights). To be included in each regression, the location must have at least 20 children in sample in the 1940 cohort. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

Source: For the 1940 cohort, 1940 census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018). 2000 cohort location characteristics from Chetty et al. (2014).

Table 6: Relationship between Changes in Location Characteristics and Changes in Child Outcomes

	All	Black	White
	(1)	(2)	(3)
Poverty Rate	-0.17***	0.05	0.07**
	(0.03)	(0.07)	(0.03)
Mean Earnings	0.17***	0.36***	0.03
	(0.04)	(0.09)	(0.04)
Gini (Bottom 99%)	-0.16***	0.16**	0.02
	(0.03)	(0.07)	(0.03)
Top 1% Share	-0.10***	0.10**	0.02
	(0.02)	(0.04)	(0.02)
Share 25-75th Percentile	0.01	-0.38***	-0.11**
	(0.05)	(0.10)	(0.04)
Single Mothers	-0.44***	-0.46***	-0.26***
	(0.04)	(0.08)	(0.04)
Fraction Black	-1.04***	-0.56***	-0.40***
	(0.06)	(0.15)	(0.07)
Observations	692	362	692

Notes: This table shows the results of regressing the change in expected child outcome from 1940 to 2000 on the change in various location characteristics at the commuting zone level, each individually. All variables are standardized to mean 0 and standard deviation 1, using population weights in each period. To be included in each regression, the location must have at least 20 children in sample in the 1940 cohort. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.
Source: For the 1940 cohort, 1940 census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

Table 7: Summary of Location Characteristic Regressions

A. CZ

	All						Black						White						Summary				
	Location Chars (Table 4)		Controlling for 40 (Table 5)				Location Chars (Table 4)		Controlling for 40 (Table 5)				Location Chars (Table 4)		Controlling for 40 (Table 5)				Share of Estimates			Changes	Always
	1940	2000	Without 40	With 40	Declines by 50%+ or Changes Sign	Changes in Chars (Table 6)	1940	2000	Without 40	With 40	Declines by 50%+ or Changes Sign	Changes in Chars (Table 6)	1940	2000	Without 40	With 40	Declines by 50%+ or Changes Sign	Changes in Chars (Table 6)	+	-	Not Sig	Sign	Significant
Poverty Rate	-	-	-	+	Yes	-	-	No	No	No	No	No	-	No	No	+	Yes	+	0.20	0.40	0.40	1	0
Mean Earnings/Income	+	+	+	No	Yes	+	+	+	No	+	Yes	+	+	+	No	+	No	No	0.73	0.00	0.27	0	0
Gini (Bottom 99%)	-	-	-	-	Yes	-	-	No	-	No	Yes	+	-	No	-	+	Yes	No	0.13	0.60	0.27	1	0
Top 1% Share	-	+	-	+	Yes	-	-	+	No	+	No	+	-	+	-	+	Yes	No	0.47	0.40	0.13	1	0
Share 25-75th Percentile	+	+	+	+	Yes	No	+	-	No	-	Yes	-	+	-	+	-	Yes	-	0.47	0.40	0.13	1	0
Single Mothers	-	-	-	-	No	-	-	-	-	-	Yes	-	-	-	-	-	Yes	-	0.00	1.00	0.00	0	1
Fraction Black	-	-	-	-	No	-	-	-	-	-	No	-	-	-	-	-	Yes	-	0.00	1.00	0.00	0	1
Commute < 15 Minutes			+	+	Yes				No	-	Yes				+	-	Yes		0.50	0.33	0.17	1	0
HS Drop Out Rate			+	+	No				No	-	Yes				No	-	Yes		0.33	0.33	0.33	1	0
Social Capital			+	No	Yes				No	-	No				+	-	Yes		0.33	0.33	0.33	1	0

B. County

	All						Black						White						Summary				
	Location Chars (Table A.1)		Controlling for 40 (Table A.2)				Location Chars (Table A.1)		Controlling for 40 (Table A.2)				Location Chars (Table A.1)		Controlling for 40 (Table A.2)				Share of Estimates			Changes	Always
	1940	2000	Without 40	With 40	Declines by 50%+ or Changes Sign	Changes in Chars (Table A.3)	1940	2000	Without 40	With 40	Declines by 50%+ or Changes Sign	Changes in Chars (Table A.3)	1940	2000	Without 40	With 40	Declines by 50%+ or Changes Sign	Changes in Chars (Table A.3)	+	-	Not Sig	Sign	Significant
Poverty Rate	-	-	-	-	No	-	-	-	-	-	No	-	-	-	-	No	Yes	-	0.00	0.93	0.07	0	0
Mean Earnings/Income	+	+	+	+	No	+	+	+	+	+	No	+	+	+	+	+	No	+	1.00	0.00	0.00	0	1
Gini (Bottom 99%)	-	-	-	-	Yes	-	-	-	-	-	No	-	-	+	-	+	Yes	-	0.13	0.87	0.00	1	1
Top 1% Share	-	+	-	No	Yes	-	-	+	No	+	No	No	-	+	-	+	Yes	-	0.33	0.47	0.20	1	0
Share 25-75th Percentile	+	+	+	+	Yes	-	+	-	+	-	Yes	-	+	-	+	-	Yes	-	0.53	0.47	0.00	1	1
Single Mothers	-	-	-	-	No	-	-	-	-	-	No	-	-	-	-	-	Yes	-	0.00	1.00	0.00	0	1
Fraction Black	-	-	-	-	No	-	-	-	-	-	No	-	-	-	-	-	Yes	-	0.00	1.00	0.00	0	1
Commute < 15 Minutes			+	+	Yes				-	-	No				+	-	Yes		0.50	0.50	0.00	1	1
HS Drop Out Rate			+	No	Yes				No	-	No				+	No	Yes		0.33	0.17	0.50	1	0
Social Capital			+	No	Yes				No	-	Yes				+	-	Yes		0.33	0.33	0.33	1	0

Notes: This table summarizes the results of the various approaches to comparing child outcomes to location characteristics over time, from Tables 4, 5, and 6 at the CZ level (Panel A) and Tables A.1, A.2, and A.3 at the county level (Panel B).

Source: For the 1940 cohort, 1940 census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018). 2000 cohort location characteristics from Chetty et al. (2014).

Appendices

Table A.1: County Level Location Characteristics and Geographic Variation in Child Outcomes

	All		Black		White	
	1940	2000	1940	2000	1940	2000
	(1)	(2)	(3)	(4)	(5)	(6)
Poverty Rate	-0.53***	-0.42***	-0.29***	-0.15***	-0.59***	-0.12***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)
Mean Earnings/Income	0.48***	0.26***	0.26***	0.21***	0.55***	0.31***
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Gini (Bottom 99%)	-0.59***	-0.32***	-0.29***	-0.10***	-0.51***	0.09***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Top 1% Share	-0.47***	0.03*	-0.22***	0.06***	-0.48***	0.28***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Share 25-75th Percentile	0.25***	0.10***	0.16***	-0.35***	0.07***	-0.30***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)
Single Mothers	-0.54***	-0.70***	-0.26***	-0.24***	-0.22***	-0.24***
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Fraction Black	-0.64***	-0.73***	-0.28***	-0.18***	-0.22***	-0.13***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)
Observations	2,941	2,941	1,145	1,145	2,925	2,925

Notes: This table shows the correlation between child outcomes and location characteristics at the county level in the 1940 and 2000 cohorts. In this table, we include the "baseline" location characteristics that we can calculate from the 1940 census and are available for the 2000 cohort. All variables are standardized to mean 0 and standard deviation 1, using population weights in each period. To be included in the 1940 regressions, the location must have at least 20 children in sample in the 1940 cohort. The regressions are run on the same set of locations in both periods. The outcome variable in the 1940 cohort regressions is expected child income rank for children of parents in the third earnings decile. In the 2000 cohort, the outcome is the expected child income rank for children of parents at the 25th percentile of the household income distribution. The outcomes for each subgroups are standardized to have mean 0 and standard deviation 1, using population weights as well. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

Source: For the 1940 cohort, 1940 census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018). 2000 cohort location characteristics from Chetty et al. (2014).

Table A.2: County Level Location Characteristics and Geographic Variation in Child Outcomes for the 2000 Cohort, Controlling for the Past

A. Baseline Location Characteristics

	All			Black			White		
	Without 40	With 40	% Difference	Without 40	With 40	% Difference	Without 40	With 40	% Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Poverty Rate	-0.48***	-0.29***	-39.6	-0.14***	-0.14***	-2.9	-0.20***	0.00	-99.5
	(0.02)	(0.02)		(0.02)	(0.03)		(0.02)	(0.02)	
Mean Earnings/Income	0.33***	0.17***	-46.8	0.17***	0.20***	20.8	0.24***	0.21***	-12.1
	(0.03)	(0.01)		(0.02)	(0.02)		(0.03)	(0.01)	
Gini (Bottom 99%)	-0.87***	-0.22***	-74.9	-0.16***	-0.08***	-46.3	-0.50***	0.10***	-120.3
	(0.03)	(0.02)		(0.03)	(0.02)		(0.03)	(0.02)	
Top 1% Share	-0.22***	-0.01	-95.7	0.00	0.06***	2035.0	-0.09***	0.19***	-320.6
	(0.03)	(0.01)		(0.03)	(0.02)		(0.03)	(0.01)	
Share 25-75th Percentile	0.76***	0.05***	-93.5	0.06**	-0.37***	-731.5	0.41***	-0.25***	-161.4
	(0.02)	(0.02)		(0.02)	(0.03)		(0.02)	(0.02)	
Single Mothers	-1.08***	-0.62***	-42.8	-0.25***	-0.23***	-6.5	-0.62***	-0.26***	-59.1
	(0.02)	(0.01)		(0.02)	(0.02)		(0.03)	(0.02)	
Fraction Black	-0.86***	-0.61***	-29.1	-0.13***	-0.17***	31.0	-0.39***	-0.17***	-56.9
	(0.02)	(0.02)		(0.02)	(0.03)		(0.03)	(0.02)	
Observations	2,941	2,941		1,145	1,145		2,925	2,925	

B. Location Characteristics from Chetty, Hendren, Klein, and Saez (2014)

	All			Black			White		
	Without 40	With 40	% Difference	Without 40	With 40	% Difference	Without 40	With 40	% Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Commute < 15 Minutes	0.41***	0.05*	-87.3	-0.15***	-0.36***	144.3	0.34***	-0.23***	-167.8
	(0.03)	(0.03)		(0.03)	(0.04)		(0.03)	(0.03)	
Gini (Bottom 99%)	-0.69***	-0.18***	-73.5	-0.17***	-0.09***	-42.9	-0.29***	0.13***	-146.2
	(0.03)	(0.02)		(0.04)	(0.03)		(0.04)	(0.02)	
HS Drop Out Rate	0.17***	0.03	-80.5	-0.02	-0.06*	238.0	0.09***	0.01	-90.1
	(0.03)	(0.03)		(0.03)	(0.04)		(0.03)	(0.03)	
Social Capital	0.60***	0.00	-100.6	0.03	-0.12***	-470.0	0.40***	-0.13***	-133.4
	(0.03)	(0.03)		(0.03)	(0.03)		(0.03)	(0.03)	
Single Mothers	-0.88***	-0.60***	-31.9	-0.25***	-0.23***	-7.6	-0.37***	-0.25***	-33.3
	(0.03)	(0.02)		(0.03)	(0.03)		(0.03)	(0.02)	
Fraction Black	-0.75***	-0.61***	-19.7	-0.16***	-0.18***	12.0	-0.27***	-0.16***	-39.2
	(0.03)	(0.02)		(0.03)	(0.04)		(0.03)	(0.03)	
Observations	1,284	1,284		652	652		1,284	1,284	

Notes: This table shows the correlation between child outcomes in the 2000 cohort and location characteristics at the county level, controlling for and not controlling for child outcomes for the 1940 cohort. The "Baseline Location Characteristics" in Panel A are those that we can calculate from the 1940 census. The characteristics in Panel B include a subset studied in Chetty et al. (2014), including those they note as being particularly correlated with child outcomes. All variables are standardized to mean 0 and standard deviation 1, using population weights in each period. The expected rank of a child at the 25th percentile in the 2000 cohort (standardized as well using 2000 population weights) is regressed against each location characteristic individually. "Without 40" and "With 40" indicate whether the regression includes as a dependent variable the expected child rank for children from households in the third decile of the parent earnings distribution in the 1940 cohort (standardized using 1940 population weights). To be included in each regression, the location must have at least 20 children in sample in the 1940 cohort. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

Source: For the 1940 cohort, 1940 census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018). 2000 cohort location characteristics from Chetty et al. (2014).

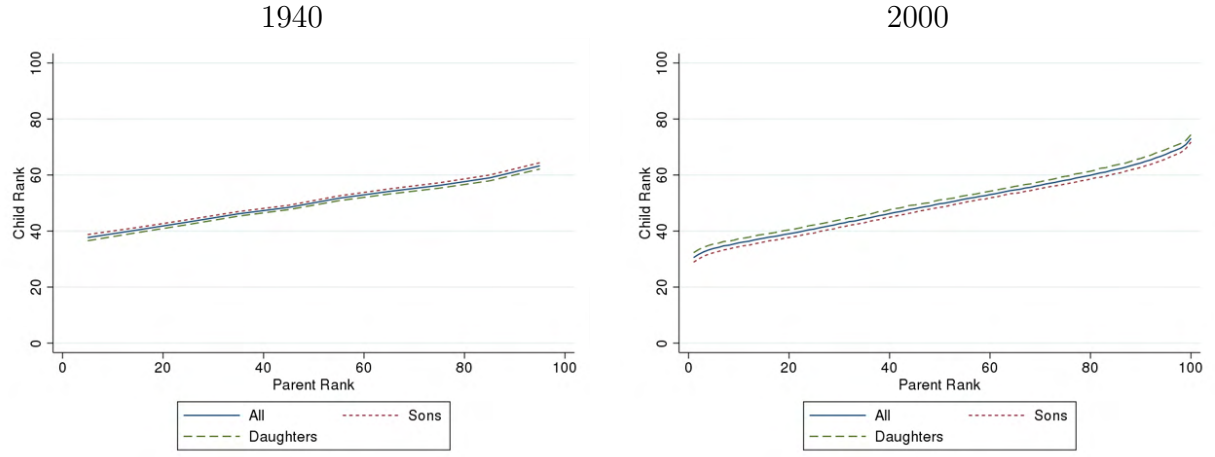
Table A.3: County Level Relationship between Changes in Location Characteristics and Changes in Child Outcomes

	All	Black	White
	(1)	(2)	(3)
Poverty Rate	-0.34***	-0.11***	-0.07***
	(0.01)	(0.03)	(0.02)
Mean Earnings	0.33***	0.20***	0.15***
	(0.02)	(0.03)	(0.02)
Gini (Bottom 99%)	-0.27***	-0.09***	-0.09***
	(0.01)	(0.03)	(0.01)
Top 1% Share	-0.14***	-0.04	-0.02*
	(0.01)	(0.02)	(0.01)
Share 25-75th Percentile	-0.04**	-0.15***	-0.12***
	(0.02)	(0.04)	(0.02)
Single Mothers	-0.58***	-0.27***	-0.35***
	(0.02)	(0.04)	(0.02)
Fraction Black	-0.84***	-0.33***	-0.30***
	(0.02)	(0.05)	(0.03)
Observations	2,941	1,145	2,925

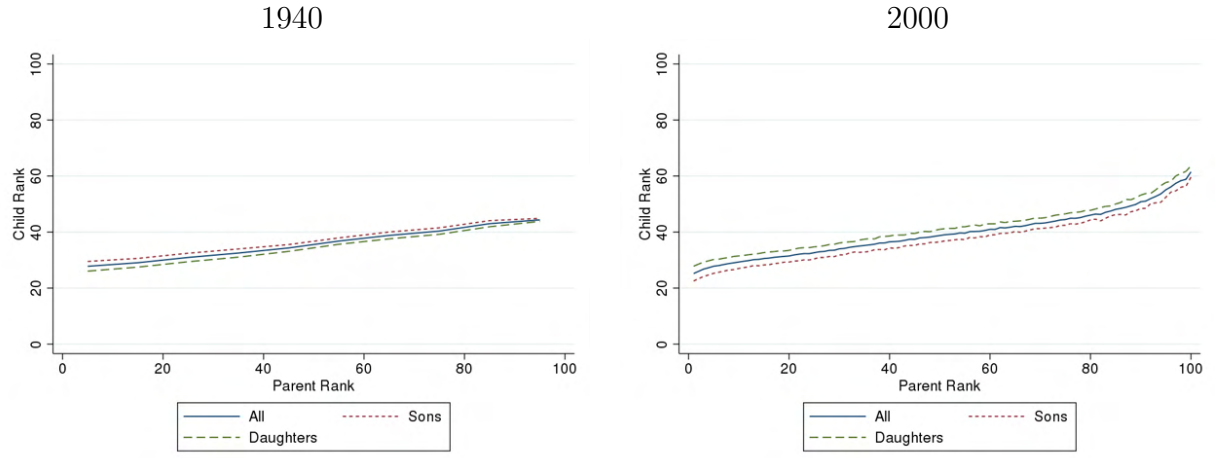
Notes: This table shows the results of regressing the change in expected child outcome from 1940 to 2000 on the change in various location characteristics at the county level, each individually. All variables are standardized to mean 0 and standard deviation 1, using population weights in each period. To be included in each regression, the location must have at least 20 children in sample in the 1940 cohort. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.
Source: For the 1940 cohort, 1940 census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

Figure A.1: Mobility Comparisons by Gender

A. All Children



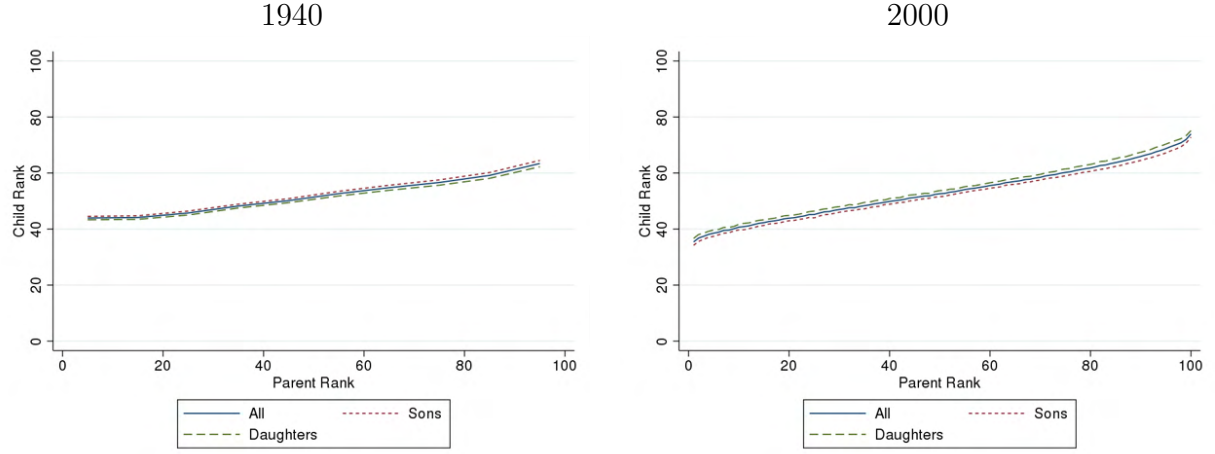
B. Black Children



Notes: Each panel compares the expected child rank for children (all, sons, and daughters) by parent earnings decile in the 1940 cohort and parent income percentile in the 2000 cohort. The 1940 cohort includes children born between 1921 and 1940 and observed in the 1940 Census. The 2000 cohort includes children born from 1978 and 1983 from Chetty et al. (2018). Each compares the outcomes by race and parent nativity in the 1940 cohort and Panel B does the same for the 2000 cohort.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

Figure A.1 Mobility Comparisons by Gender, continued

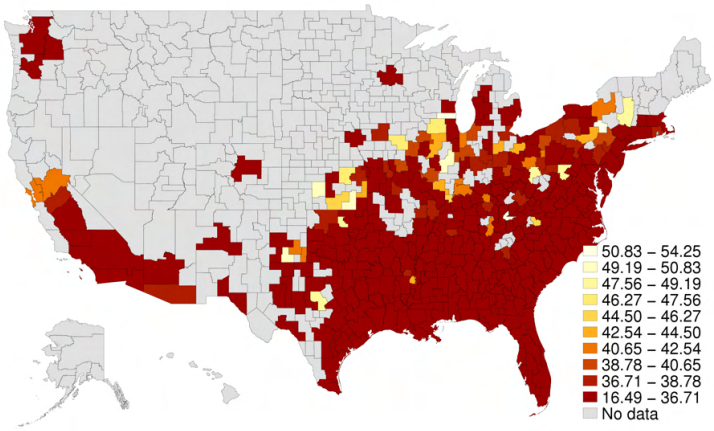
C. White Children



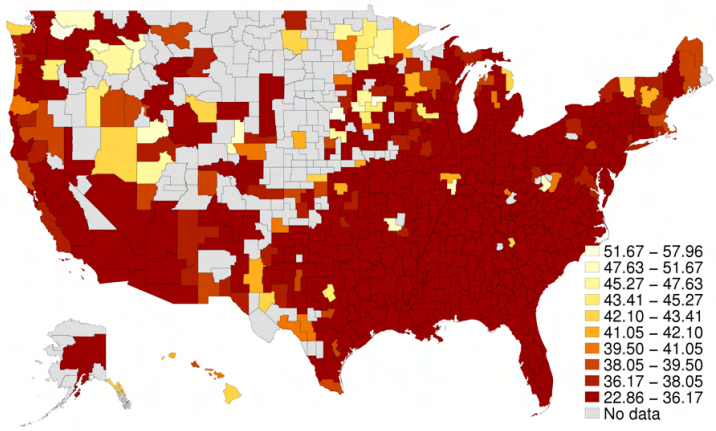
Notes: Each panel compares the expected child rank for children (all, sons, and daughters) by parent earnings decile in the 1940 cohort and parent income percentile in the 2000 cohort. The 1940 cohort includes children born between 1921 and 1940 and observed in the 1940 Census. The 2000 cohort includes children born from 1978 and 1983 from Chetty et al. (2018). Each compares the outcomes by race and parent nativity in the 1940 cohort and Panel B does the same for the 2000 cohort.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

Figure A.2: Spatial Variation in Child Outcomes over Time
Black Children
(Using Cutoffs from Pooled Sample)

A. 1940 Cohort



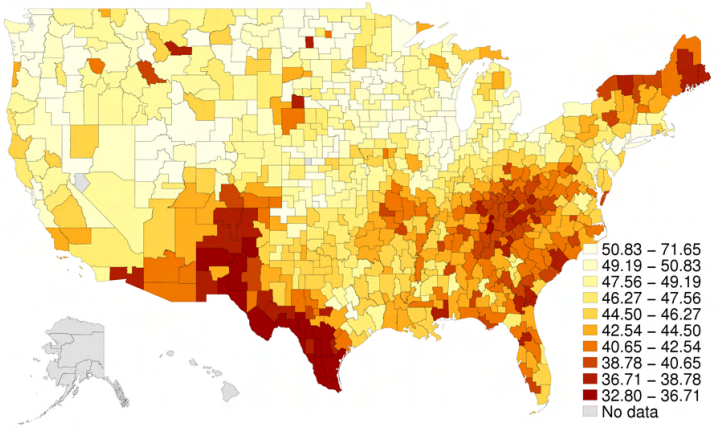
B. 2000 Cohort



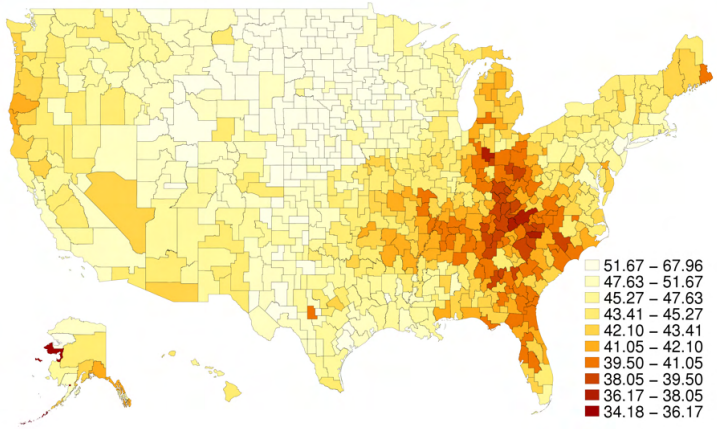
Notes: Panel A shows the expected outcome of Black children born between 1921 and 1940 and observed in the 1940 Census with parents in the 3rd decile of the parent earnings distribution. Panel B shows estimates of the expected outcome for Black children of parents with income at the 25th percentile of the distribution born from 1978 to 1983 from Chetty et al. (2018). The map cutoffs are held fixed at those from Figure 4.

Figure A.3: Spatial Variation in Child Outcomes over Time, White Children
(Using Cutoffs from Pooled Sample)

A. 1940 Cohort



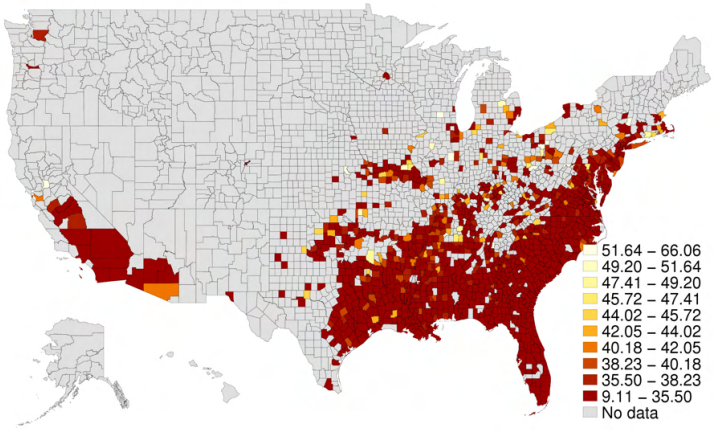
B. 2000 Cohort



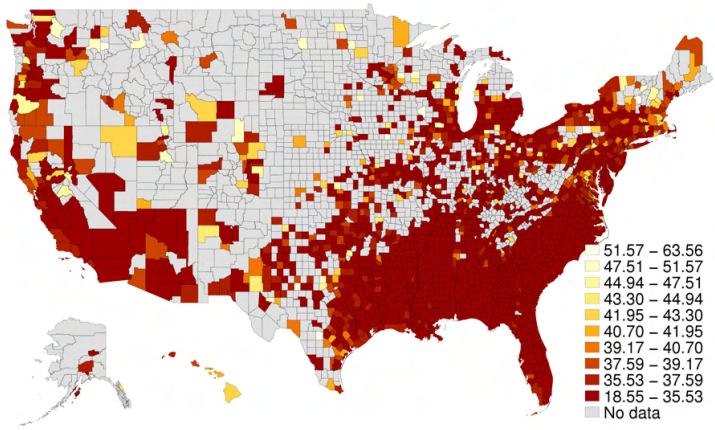
Notes: Panel A shows the expected outcome of White children born between 1921 and 1940 and observed in the 1940 Census with parents in the 3rd decile of the parent earnings distribution. Panel B shows estimates of the expected outcome for White children of parents with income at the 25th percentile of the distribution born from 1978 to 1983 from Chetty et al. (2018). The map cutoffs are held fixed at those from Figure 4.

Figure A.4: County Level Spatial Variation in Child Outcomes over Time
Black Children (Using Cutoffs from Pooled Sample)

A. 1940 Cohort



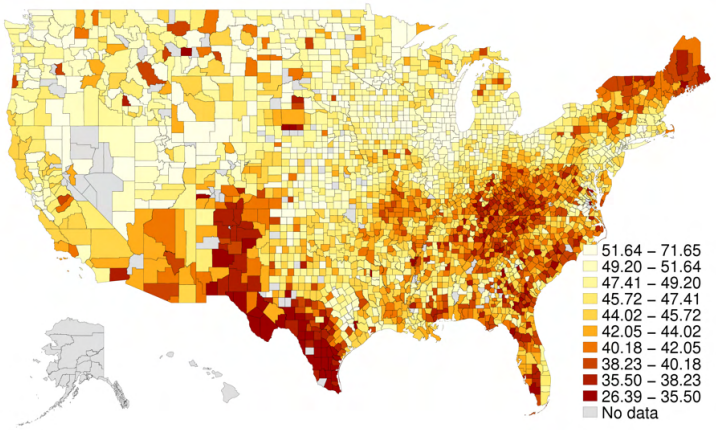
B. 2000 Cohort



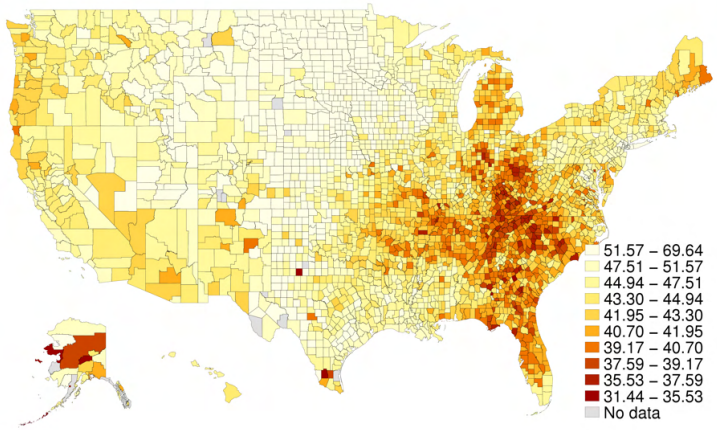
Notes: Panel A shows the expected outcome of Black children born between 1921 and 1940 and observed in the 1940 Census with parents in the 3rd decile of the parent earnings distribution. Panel B shows estimates of the expected outcome for Black children of parents with income at the 25th percentile of the distribution born from 1978 to 1983 from Chetty et al. (2018). The map cutoffs are held fixed at those from Figure 7.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

Figure A.5: County Level Spatial Variation in Child Outcomes over Time, White Children
(Using Cutoffs from Pooled Sample)

A. 1940 Cohort

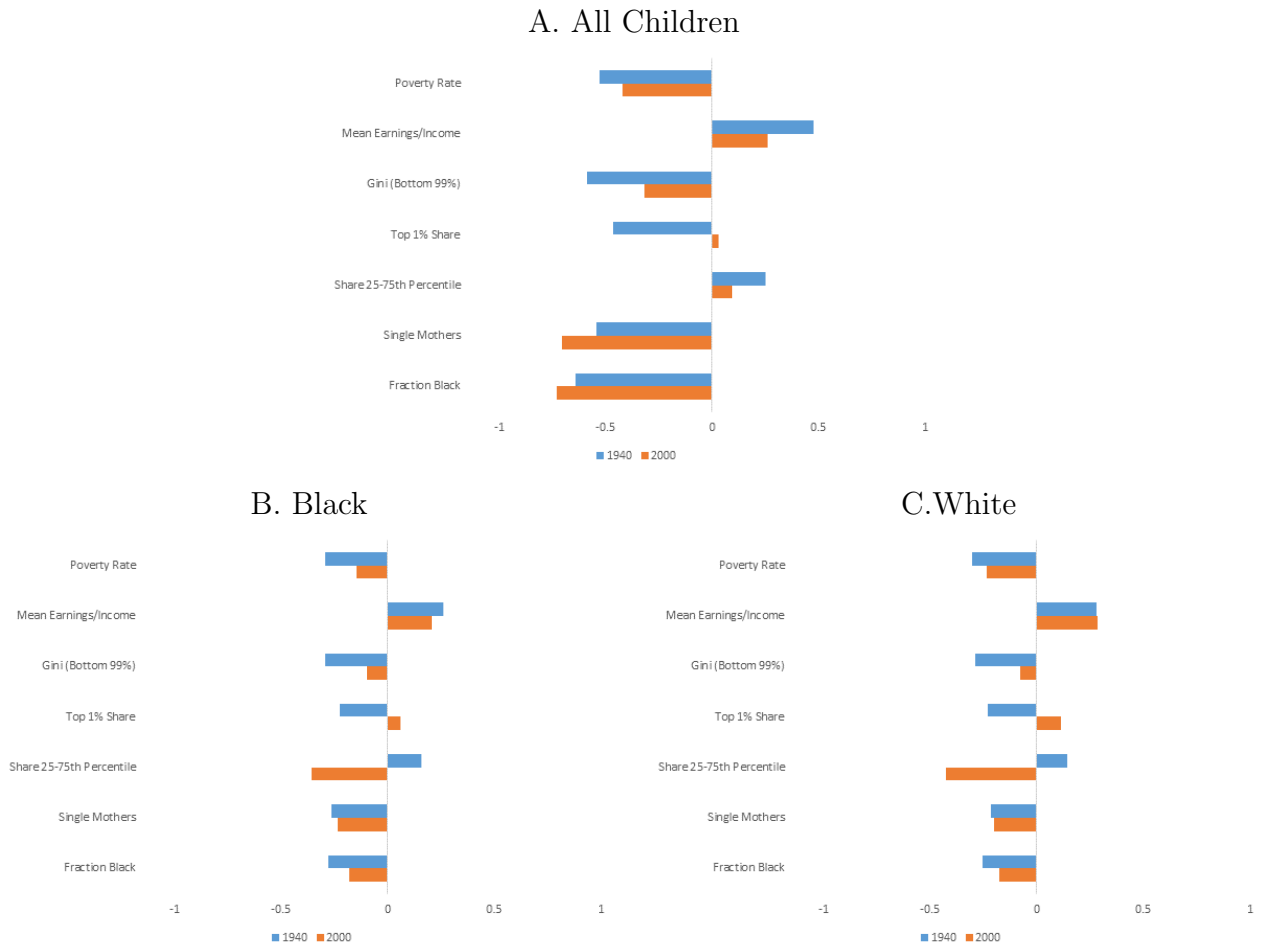


B. 2000 Cohort



Notes: Panel A shows the expected outcome of White children born between 1921 and 1940 and observed in the 1940 Census with parents in the 3rd decile of the parent earnings distribution. Panel B shows estimates of the expected outcome for White children of parents with income at the 25th percentile of the distribution born from 1978 to 1983 from Chetty et al. (2018). The map cutoffs are held fixed at those from Figure 7.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

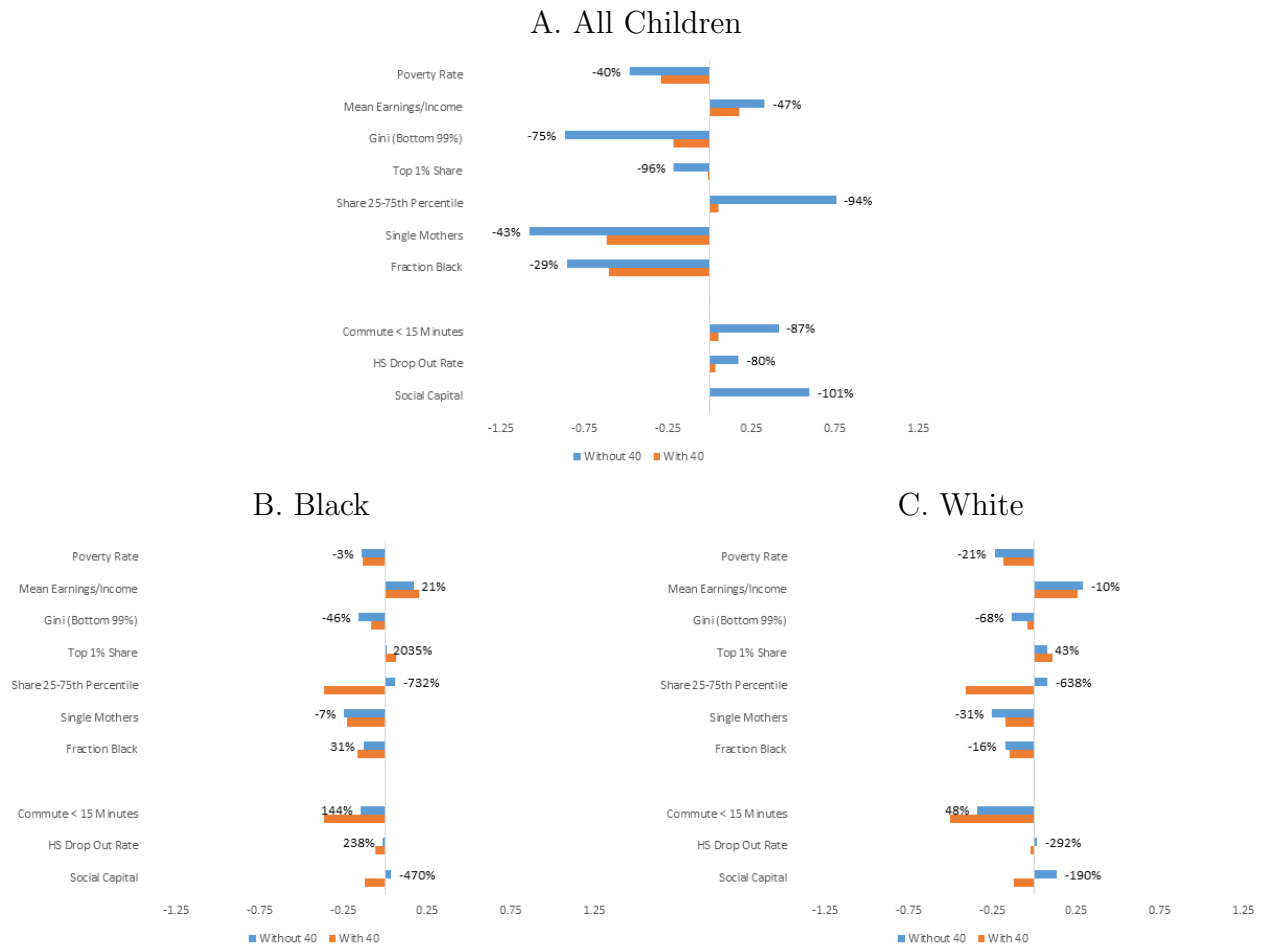
Figure A.6: County Level Location Characteristics and Geographic Variation in Child Outcomes



Notes: The results in this figure correspond to Table A.1. This Figure shows the correlation (using population weights) between child outcomes and location characteristics at the county level in the 1940 and 2000 cohorts, for all children in Panel A, Black children in Panel B, and White children in Panel C.

Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018). 2000 cohort location characteristics from Chetty et al. (2014).

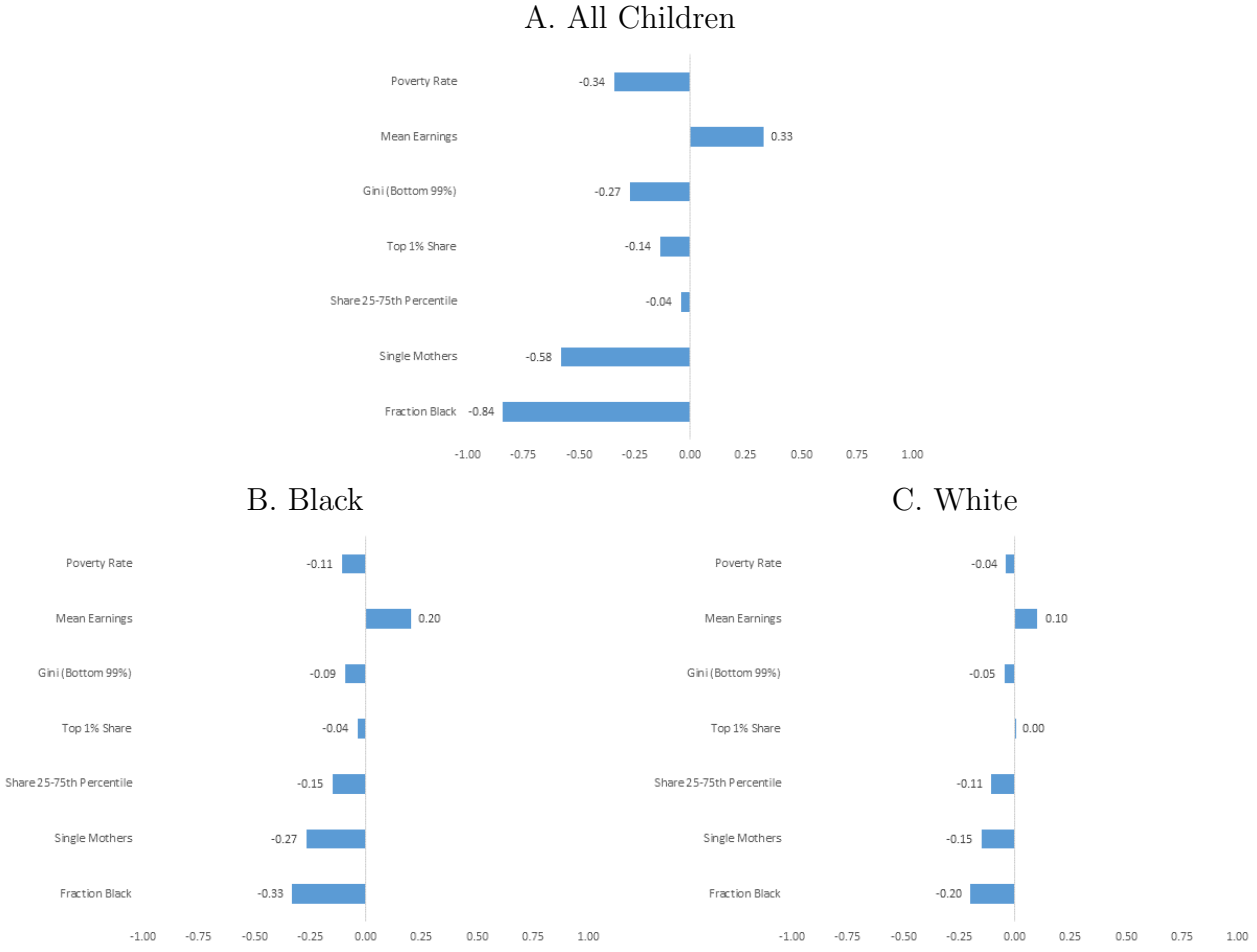
Figure A.7: County Level Location Characteristics and Geographic Variation in Child Outcomes for the 2000 Cohort, Controlling for the Past



Notes: The results in this figure correspond to Table A.2. This figure shows the correlation between child outcomes in the 2000 cohort and location characteristics at the county level, controlling for and not controlling for child outcomes for the 1940 cohort. The first six characteristics shown for each group are the "Baseline Location Characteristics" in Panel A of Table 5, which are those that we can calculate from the 1940 Census. The next three characteristics shown are a subset studied in Chetty et al. (2014), including those they note as being particularly correlated with child outcomes. All variables are standardized to mean 0 and standard deviation 1. In the population-weighted regressions the standardization is done using population weights. The numbers indicate the extent to which the inclusion of the 1940 outcomes affect the regression coefficient between each location characteristic and outcomes for the 2000 cohort.

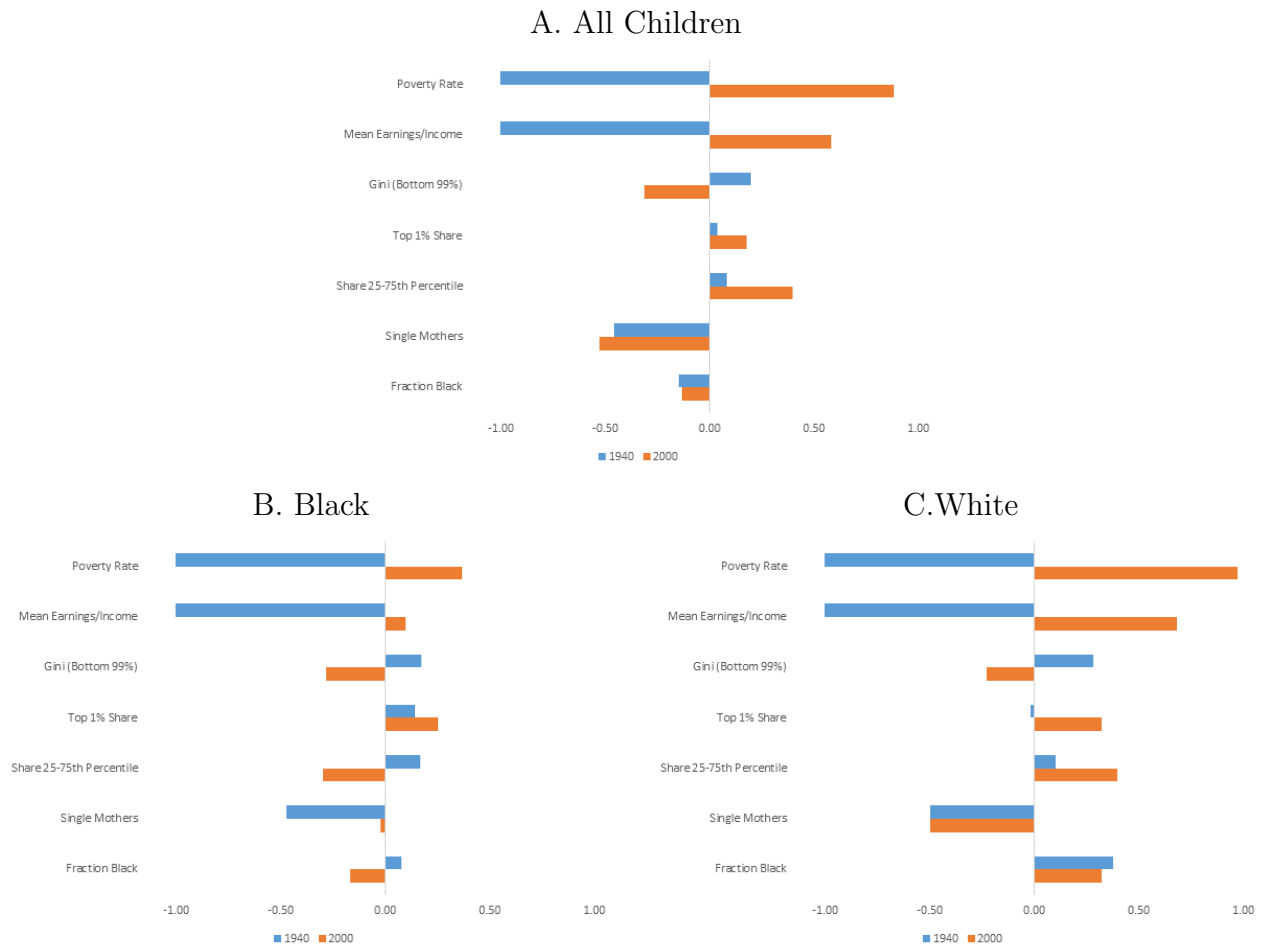
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018). 2000 cohort location characteristics from Chetty et al. (2014).

Figure A.8: County Level Relationship between Changes in Location Characteristics and Changes in Child Outcomes



Notes: The results in this figure correspond to Table A.3. This figure shows the results of regressing the change in expected child outcome from 1940 to 2000 on the change in various location characteristics at the county level. All variables are standardized to mean 0 and standard deviation 1.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

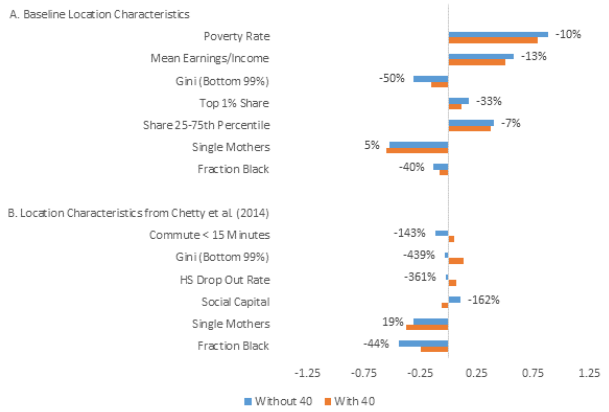
Figure A.9: Location Characteristics and Geographic Variation in Child Outcomes, Regressed Together



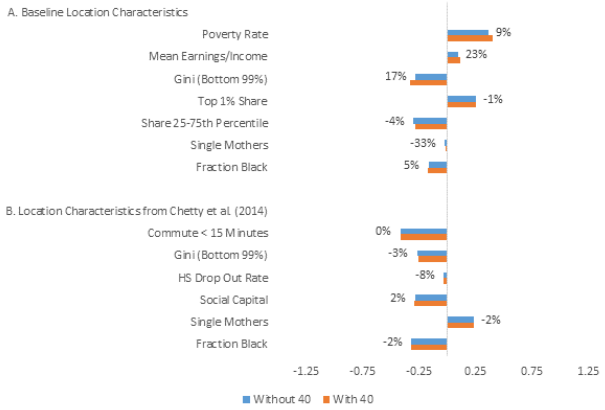
Notes: This Figure shows the coefficients from a regression of child outcomes on location characteristics at the commuting zone level in the 1940 and 2000 cohorts, for all children in Panel A, Black children in Panel B, and White children in Panel C. These regressions includes all location characteristics simultaneously rather than each separately as in Table 4 and Figure 10.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018). 2000 cohort location characteristics from Chetty et al. (2014).

Figure A.10: Location Characteristics and Geographic Variation in Child Outcomes for the 2000 Cohort, Controlling for the Past, Regressed Together

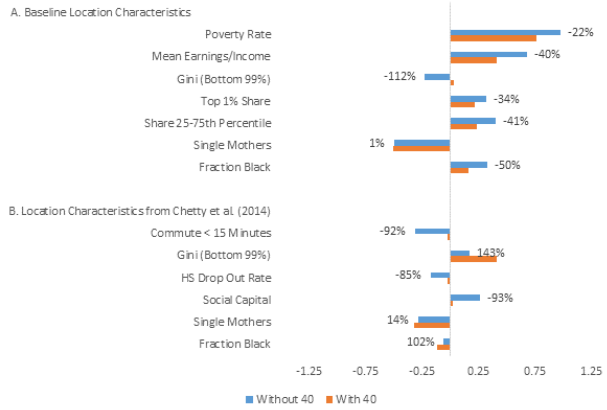
A. All Children



B. Black

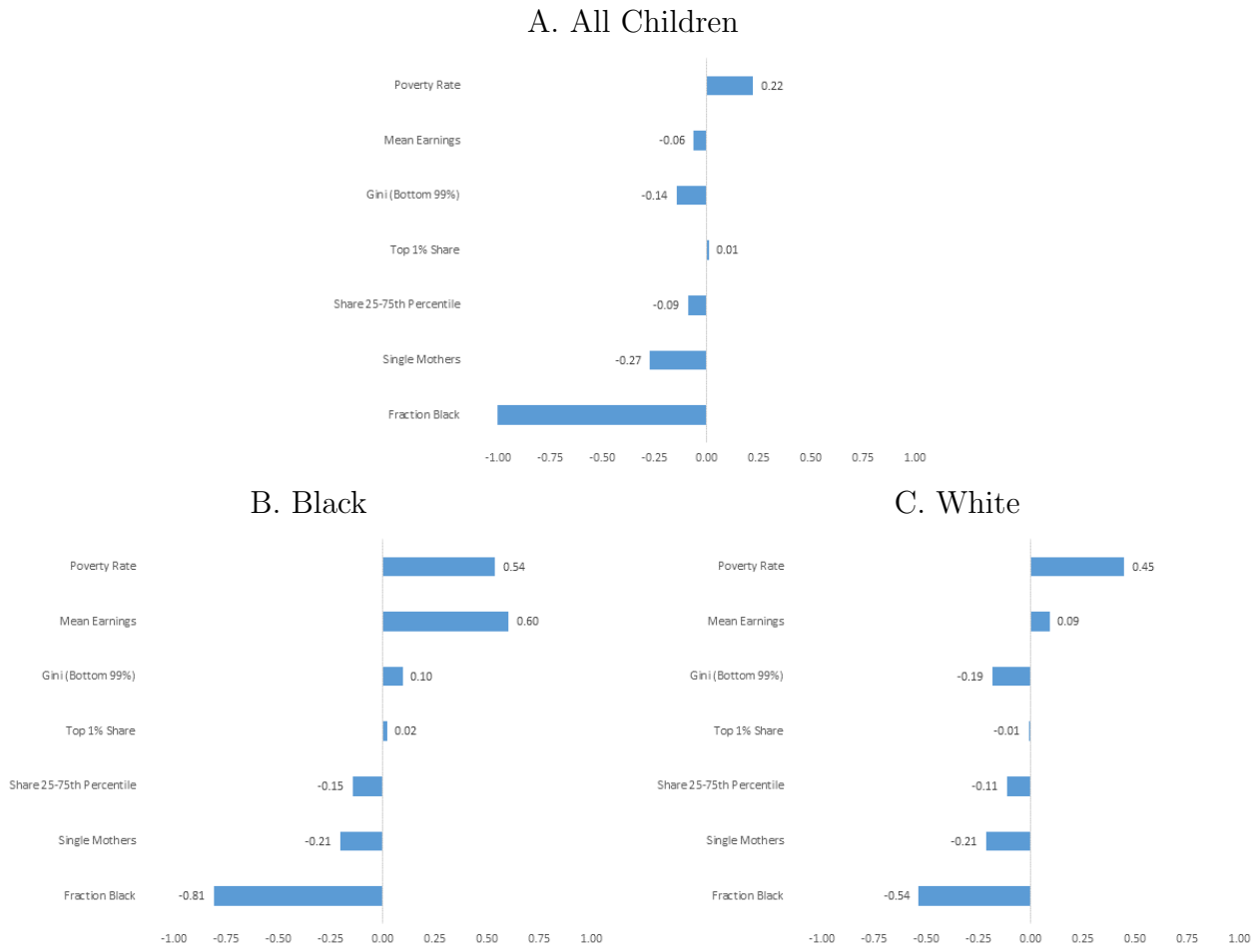


C. White



Notes: This figure shows the coefficients from a regression of child outcomes in the 2000 cohort on location characteristics at the commuting zone level, controlling for and not controlling for child outcomes for the 1940 cohort. The first six characteristics shown for each group are the "Baseline Location Characteristics" in Panel A of Table 5, which are those that we can calculate from the 1940 Census. The next three characteristics shown are a subset studied in Chetty et al. (2014), including those they note as being particularly correlated with child outcomes. All variables are standardized to mean 0 and standard deviation 1. In the population-weighted regressions the standardization is done using population weights. The numbers indicate the extent to which the inclusion of the 1940 outcomes affect the regression coefficient between each location characteristic and outcomes for the 2000 cohort. These regressions includes all location characteristics simultaneously rather than each separately as in Table 5 and Figure 11.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018). 2000 cohort location characteristics from Chetty et al. (2014).

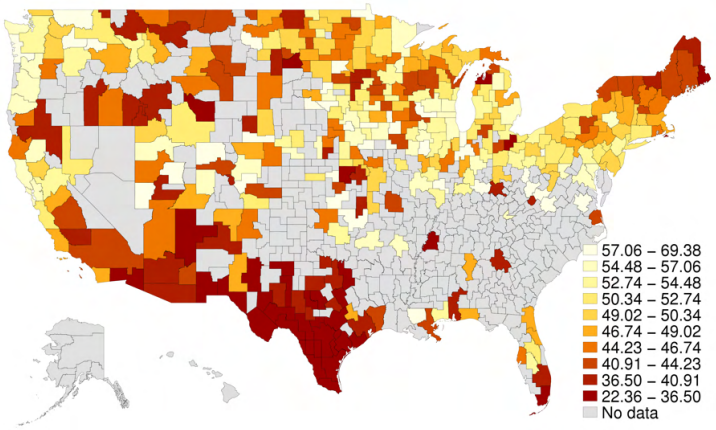
Figure A.11: Relationship between Changes in Location Characteristics and Changes in Child Outcomes, Regressed Together



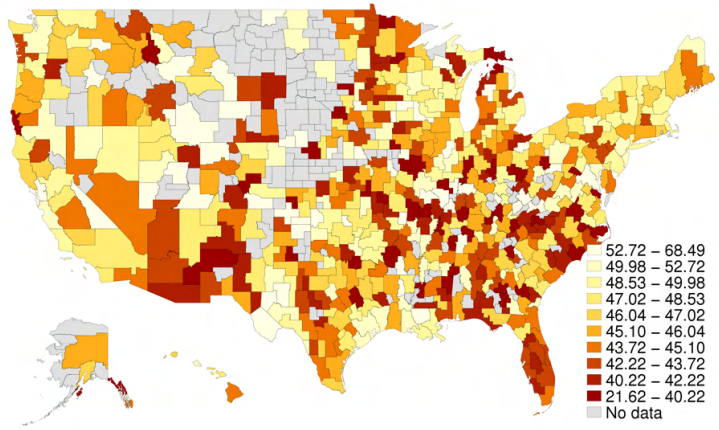
Notes: The results in this figure correspond to Table A.3. This figure shows the results of regressing the change in expected child outcome from 1940 to 2000 on the change in various location characteristics at the commuting zone level. All variables are standardized to mean 0 and standard deviation 1. In the population-weighted regressions the standardization is done using population weights.
Source: For the 1940 cohort, 1940 Census linked to 1974 and 1979 1040 filings. For the 2000 cohort, Chetty et al. (2018).

Figure A.12: Spatial Variation in Child Outcomes over Time
Children of Foreign-Born Parents

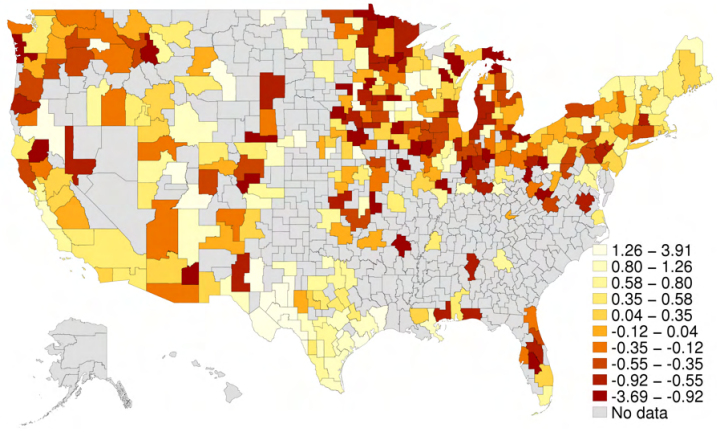
A. 1940 Cohort



B. 2000 Cohort



C. Standardized Change in Outcomes (2000-1940)



Notes: Panel A shows the expected outcome of children with a foreign-born parent born between 1921 and 1940 and observed in the 1940 Census with parents in the 3rd decile of the parent earnings distribution. Panel B shows estimates of the expected outcome for children with a foreign-born mother with parent income at the 25th percentile of the distribution born from 1978 to 1983 from Chetty et al. (2018). For Panel C, the values in A and B are standardized to have mean 0 and standard deviation 1 (in each period). Panel C maps the difference between the standardized outcomes, showing in which areas outcomes of children of foreign-born parents have improved or declined relative to the average outcome for children of foreign-born parents over time.