The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility

by

Raj Chetty
Harvard University and NBER

John N. Friedman
Brown University and NBER

Nathaniel Hendren
Harvard University and NBER

Maggie R. Jones
U.S. Census Bureau

Sonya R. Porter
U.S. Census Bureau

CES 18-42 September, 2018

The research program of the Center for Economic Studies (CES) produces a wide range of economic analyses to improve the statistical programs of the U.S. Census Bureau. Many of these analyses take the form of CES research papers. The papers have not undergone the review accorded Census Bureau publications and no endorsement should be inferred. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. Republication in whole or part must be cleared with the authors.

To obtain information about the series, see www.census.gov/ces or contact Christopher Goetz, Editor, Discussion Papers, U.S. Census Bureau, Center for Economic Studies 5K028B, 4600 Silver Hill Road, Washington, DC 20233, CES.Working.Papers@census.gov. To subscribe to the series, please click here.

Abstract

We construct a publicly available atlas of children's outcomes in adulthood by Census tract using anonymized longitudinal data covering nearly the entire U.S. population. For each tract, we estimate children's earnings distributions, incarceration rates, and other outcomes in adulthood by parental income, race, and gender. These estimates allow us to trace the roots of outcomes such as poverty and incarceration back to the neighborhoods in which children grew up. We find that children's outcomes vary sharply across nearby tracts: for children of parents at the 25th percentile of the income distribution, the standard deviation of mean household income at age 35 is $4,200 across tracts within counties. We illustrate how these tract-level data can provide insight into how neighborhoods shape the development of human capital and support local economic policy using two applications. First, we show that the estimates permit precise targeting of policies to improve economic opportunity by uncovering specific neighborhoods where certain subgroups of children grow up to have poor outcomes. Neighborhoods matter at a very granular level: conditional on characteristics such as poverty rates in a child's own Census tract, characteristics of tracts that are one mile away have little predictive power for a child's outcomes. Our historical estimates are informative predictors of outcomes even for children growing up today because neighborhood conditions are relatively stable over time. Second, we show that the observational estimates are highly predictive of neighborhoods' causal effects, based on a comparison to data from the Moving to Opportunity experiment and a quasi-experimental research design analyzing movers' outcomes. We then identify high-opportunity neighborhoods that are affordable to low-income families, providing an input into the design of affordable housing policies. Our measures of children's long-term outcomes are only weakly correlated with traditional proxies for local economic success such as rates of job growth, showing that the conditions that create greater upward mobility are not necessarily the same as those that lead to productive labor markets.

Keyword: intergenerational social mobility, data tool

JEL Classification:

---

[*] Any opinions and conclusions expressed herein are those of the authors and do not necessarily reflect the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. The statistical summaries reported in this paper have been cleared by the Census Bureau's Disclosure Review Board release authorization number CBDRB-FY18-319. We thank John Abowd, Peter Bergman, David Deming, Edward Glaeser, David Grusky, Lawrence Katz, Enrico Moretti, Robert Sampson, Salil Vadhan, and numerous seminar participants for helpful comments and discussions. We are indebted to Caroline Dockes, Michael Droste, Benjamin Goldman, Jack Hoyle, Federico Gonzalez Rodriguez, Jamie Gracie, Matthew Jacob, Martin Koenen, Sarah Merchant, Donato Onorato, Kamelia Stavreva, Wilbur Townsend, Joseph Winkelmann, and other Opportunity Insights pre-doctoral fellows for their outstanding contributions to this work. This research was funded by the Bill and Melinda Gates Foundation, Chan-Zuckerberg Initiative, Robert Wood Johnson Foundation, Kellogg Foundation, and Harvard University.

I Introduction

Recent experimental and quasi-experimental studies have established that the neighborhood in which a child grows up has substantial causal effects on his or her prospects of upward income mobility (Chetty et al. 2016; Chetty and Hendren 2018a; Chyn 2018; Deutscher 2018; Laliberté 2018).[1] The emerging consensus that neighborhoods play a key role in shaping children's outcomes (Sharkey 2016) raises a natural question: which neighborhoods in the United States currently offer the best and worst opportunities for children? Answering this question can allow policy makers to target interventions to improve economic opportunity more effectively and allow researchers to better understand the production function of opportunity.

There is currently no comprehensive data on children's outcomes across neighborhoods in the U.S., as prior work has focused either on a small subset of neighborhoods (e.g., the Moving to Opportunity Experiment) or on coarse geographies such as commuting zones (e.g., Chetty et al. 2014). In this paper, we construct a new, publicly available dataset – which we term the *Opportunity Atlas* – that provides estimates of the long-term outcomes of children who grew up in each Census tract, which are small geographic units that have an average population of 4,250 people. Our estimates differ from traditional indicators of local conditions based on cross-sectional data – such as rates of poverty or crime – by tracing the roots of such outcomes back to the neighborhoods in which children grew up.

We construct these estimates using an individual-level panel dataset from the U.S. Census Bureau that covers virtually the entire American population from 1989-2015. Following Chetty et al. (2018), we use de-identified data from the 2000 and 2010 decennial Censuses linked to data from federal income tax returns and the 2005-2015 American Community Surveys to obtain information on children's outcomes in adulthood and their parents' characteristics. We focus primarily on children in the 1978-1983 birth cohorts who were born in the U.S. or are authorized immigrants who came to the U.S. in childhood. Our primary analysis sample consists of 20.5 million children, approximately 96.2% of the total number of children in the birth cohorts we study.

We construct tract-level estimates of children's incomes in adulthood and other outcomes such as incarceration rates and teenage birth rates by race, gender, and parents' household income level. We assign children to locations in proportion to the amount of their childhood they spent growing

[1] These recent studies build on a longstanding literature documenting variation in outcomes across neighborhoods in observational data; see e.g., Jencks and Mayer (1990), Case and Katz (1991), Brooks-Gunn et al. (1993), Cutler and Glaeser (1997), Leventhal and Brooks-Gunn (2000), and Sampson et al. (2002).

up in each Census tract. In each tract-by-gender-by-race cell, we estimate the conditional expectation of children's outcomes given their parents' household income using a univariate regression whose functional form is chosen based on estimates at the national level to capture potential non-linearities. We suppress estimates for cells with fewer than 20 children and infuse noise in the estimates that is inversely proportional to sample size to protect privacy, following methods from the literature on differential privacy (Dwork 2008; Abowd and Schmutte 2015).[2] Despite their granularity, the tract-level estimates are generally quite precise, with most of the spatial variation driven by true differences in mean outcomes rather than sampling variation. For example, pooling racial groups and genders, 91% of the cross-tract variance in the predicted income ranks of children with parents at the 25th percentile is due to signal rather than noise (population-weighted).

We provide the tract-level data for public use both in the form of downloadable datasets and in an interactive mapping tool that facilitates visualization of the spatial patterns. We illustrate how these data can be used to understand how neighborhoods shape children's outcomes and support local policy using two applications: one that focuses on observational variation in children's outcomes and another that studies the causal effects of neighborhoods.

Our first application aims to inform the design of economic policies that target disadvantaged families. Many policies to increase opportunity – ranging from tax credits such as Opportunity Zones to educational interventions such as Head Start centers – use location as a predictor or "tag" (Akerlof 1978; Nichols and Zeckhauser 1982) for having a low level of income. From the perspective of predicting children's future incomes, observed outcomes in a given area are of direct interest. It does not necessarily matter whether these outcomes arise from the causal effect of the neighborhood or from selection.[3] We therefore present a descriptive characterization of how children's outcomes vary across neighborhoods, with the goal of informing those who are interested in targeting low-opportunity areas about where they should focus their attention. Our descriptive analysis yields three sets of results.

First, children's outcomes vary sharply across neighborhoods, even conditional on parental income. Among children with parents at the 25th income percentile, the standard deviation (SD) of mean household income in adulthood across tracts is approximately $6,700 (21% of mean income).

[2]The noise added to protect privacy is typically an order of magnitude smaller than the inherent noise in the estimates due to sampling error and hence does not affect the reliability of the estimates significantly.

[3]More precisely, in neoclassical models of optimal taxation, the optimal policy does not depend upon the reason that location predicts future incomes. In other settings, one may wish to distinguish between areas that have poor outcomes because of selection vs. causal effects; for instance, interventions to improve institutions in a given area may be especially valuable if they have poor causal effects. Our point is simply that for certain questions, the observational variation in outcomes itself is of interest; we turn to distinguish causal effects from selection in our second application.

Some of this variation is due to sharp differences in rates of upward mobility across racial and ethnic groups (Chetty et al. 2018), but there is substantial variation even conditional on race. For instance, among white children, the corresponding SD of mean household income across tracts is $6,100.

More than half of the tract-level variance is within counties. There is substantial variance across tracts even within the same school catchment areas, with schools accounting for less than half of observed variance across tracts within counties. Indeed, children's outcomes often vary dramatically even across tracts that are a few miles apart. For example, 44% of black men who grew up in the lowest-income families in Watts, a neighborhood in central Los Angeles, are incarcerated on a single day (April 1, 2010 – the day of the 2010 Census). By contrast, 6.2% of black men who grow up in families with similar incomes in central Compton, 2.3 miles south of Watts, are incarcerated on a single day.

Second, there is substantial heterogeneity in mean tract-level outcomes across subgroups. Places that have good outcomes for one group do not always have good outcomes for others. The correlations in mean earnings between whites, blacks, and Hispanics across tracts conditional on having parents at the 25th percentile are approximately 0.6. Returning to the example above, Hispanic men who grew up in Watts have an incarceration rate of 4.7% – an order of magnitude smaller than for black men raised in the same tract. We also find substantial heterogeneity across outcomes. For instance, incarceration rates have a race-adjusted correlation of -0.3 with mean household incomes in adulthood for men with parents at the 25th percentile across tracts, implying that the factors that drive variation in crime and incarceration across neighborhoods may differ from those that drive differences in earnings. In short, neighborhoods are not uni-dimensional: they matter differently across subgroups and outcomes.

Third, children's outcomes differ from traditional measures of neighborhood disadvantage. We find no association between children's outcomes and rates of job or wage growth. For example, Atlanta and Charlotte have had exceptionally high rates of job and wage growth over the past two decades, yet have among the lowest rates of upward mobility for children who grow up there. These cities achieve high rates of economic growth despite offering local residents limited prospects for upward income mobility by importing talent – i.e., attracting talented individuals to move in and fill high-paying jobs. Job density is also uncorrelated with children's outcomes across neighborhoods within cities, challenging spatial mismatch theories (Kain 1968). The disconnect between the availability of jobs and upward mobility illustrates that the factors that promote the development

of human capital may differ from those that lead to productive labor markets, which have been the focus of existing work in urban economics (e.g., Moretti 2011, Glaeser et al. 2011, Kline and Moretti 2014).

In contrast with measures of job density, we find a strong positive correlation between the employment rates of adults who live in a tract and rates of upward mobility for children who grow up there (race-adjusted correlation = 0.3). Evidently, what predicts upward mobility is not proximity to jobs, but growing up around people who have jobs. This result echoes Case and Katz's (1991) finding that children's outcomes are correlated with the characteristics of the "company they keep." We find similar correlations between children's outcomes and other socioeconomic characteristics of adults in an area, such as mean incomes and the share of single-parent households, consistent with Sampson et al. (1997). We also find strong correlations with proxies for social capital and test scores. Along all of these dimensions, what matters are characteristics in one's own immediate neighborhood rather than nearby areas. For example, conditional on poverty rates in one's own tract, poverty rates in immediately adjacent tracts are not strongly associated with children's outcomes. Zooming down to the block level, we find that the predictive power of poverty rates for a child's outcomes decays within a 0.6 mile radius of where the child grows up. Hence, the "neighborhood" that matters for children's outcomes is very small.

Together, observable characteristics of neighborhoods explain half of the tract-level variance in children's outcomes, implying that, relative to existing data, our estimates provide considerable new information that can help identify areas where opportunity is most lacking.[4] One practical challenge in using our estimates of social mobility to inform policy design is that they necessarily come with a lag, as one must wait until children grow up to observe their earnings. However, the power of tract-level outcomes in forecasting outcomes for future birth cohorts decays by only 10% over a decade. Moreover, historical outcome data are substantially better predictors of more recent outcomes than contemporaneous observables such as poverty rates. Our estimates are thus informative (albeit imperfect) predictors of economic opportunity even for children today.

Next, we turn to our second application, which is motivated by a simple question: "Where should a family seeking to improve their children's outcomes live?" The answer to this question matters

[4] Existing indices of economic opportunity at the local level, such as the Neighborhood Deprivation Index (Messer et al. 2006), the Kirwan Child Opportunity Index (Acevedo-Garcia et al. 2014), and Opportunity Nation's Opportunity Index (Opportunity-Nation 2017), use combinations of cross-sectional covariates such as poverty rates and job density as proxies for opportunity. The correlations between these existing measures and our statistics range from 0 to 0.6. The key advantage of our approach is that we directly measure opportunity based on children's actual outcomes, mitigating the need to determine which predictors provide the most useful proxies for economic opportunity.

not only to individual families but also for policy design. For example, many affordable housing programs have the explicit goal of helping low-income families access higher-opportunity areas. For this application, it is critical to understand whether the observational variation documented above is driven by causal effects or selection. Does moving to an area with better observed outcomes improve a given child's outcomes?

To estimate the fraction of the variance in observed outcomes across tracts that is due to causal effects of place, we first compare our observational estimates to results from the Moving to Opportunity (MTO) experiment.[5] The MTO experiment offered randomly selected families living in high-poverty housing projects housing vouchers to move to lower-poverty neighborhoods. Chetty et al. (2016) use tax data similar to those we use here to show that moving to a lower-poverty neighborhood led to large increases in earnings in adulthood for children who moved at young ages. We find a correlation of 0.6 between the earnings of children who were randomly assigned vouchers to move to different neighborhoods at young ages in the MTO data and the mean earnings of children who grow up in low-income families in those areas in our observational data. A $1,000 increase in mean earnings (conditional on parental income) in the observational data is associated with a $700 increase in earnings in the experimental data, suggesting that about 70% of the variance in the observational outcomes is due to causal effects of place.

To evaluate the extent to which the observational variation reflects causal effects of place more broadly, beyond the small number of neighborhoods included in the MTO experiment, we use the quasi-experimental research design developed by Chetty and Hendren (2018a). We study the outcomes of children whose families move across tracts, exploiting variation in the timing of moves between areas for identification. We find that children who move to areas with better observed outcomes earlier in their childhood have better outcomes themselves. Under the identifying assumption that unobservable determinants of children's outcomes in adulthood are uncorrelated with the age at which they move to a different area – an assumption that we validate using sibling comparisons and a set of placebo tests, as in Chetty and Hendren (2018a) – this result implies that neighborhoods have causal effects on children's outcomes. Growing up in a better neighborhood is beneficial throughout childhood, but where children live as adolescents (rather than at very early ages) is particularly influential in determining their later outcomes, consistent with Deutscher's

[5] As is now conventional in the social sciences, we use a potential outcomes framework (Neyman 1923; Holland 1986) to define notions of causality. We identify the causal effects of neighborhoods using experimental and quasi-experimental methods. However, we emphasize that the correlations with neighborhood characteristics discussed above are associations that should not be interpreted as causal effects.

(2018) recent findings in Australian data. The quasi-experimental estimates imply that about 62% of the observational variation across tracts in the national data is due to causal effects.

Our estimates imply that moving at birth from a neighborhood at the 25th percentile of the distribution of upward mobility within one's county to a neighborhood at the 75th percentile would increase the lifetime earnings of a child growing up in a low-income family by $206,000. Of course, the feasibility of such a move relies on being able to find affordable housing in high-opportunity neighborhoods. Higher-opportunity tracts are more expensive on average, but the correlation between children's mean earnings given parents at the 25th percentile and median rents is only 0.44 across tracts within commuting zones. Hence, there are many affordable high-opportunity neighborhoods in most cities.[6] One explanation for the existence of such areas in spatial equilibrium is that they have other disamenities that make them less desirable. An alternative possibility is that families lack information about high-opportunity areas or face barriers in moving to such places (e.g., DeLuca et al. 2017, Christensen and Timmins 2018). To distinguish between these explanations, we divide our tract-level estimates into a component predicted by observable factors such as poverty rates and a residual "unobservable" component. Only the observable component is capitalized in rents, suggesting that the opportunity bargains may exist partly because families do not know which neighborhoods have the greatest benefits for their children.[7]

These results suggest that housing choice voucher programs could potentially be designed to achieve even larger gains for children than those in the MTO experiment without additional government expenditure. If the families who received experimental vouchers in MTO had moved to the areas with the best observed outcomes in our data that were both less expensive than the areas to which they moved and closer to their origin neighborhoods, their children's earnings would have increased by nearly twice as much as they did. More broadly, changes in housing policies could influence economic mobility by changing the price of opportunity – the cost of moving to a neighborhood that delivers $1 higher earnings for one's child. We find that the price of opportunity is substantially higher in metro areas that have stricter land use regulations. Relaxing such restrictions or increasing access to high-opportunity areas through other tools could benefit not only the families who move, but all taxpayers.[8] If a child growing up in a low-income family were

[6] We show that moving to tracts with low rents but good observed outcomes at an earlier age in childhood produces substantial long-term gains using the quasi-experimental design described above, addressing the concern that these tracts might just have a positive selection of families.

[7] These findings mirror Chetty and Hendren's (2018b) findings at the county level, but show that "opportunity bargains" exist even at the tract level within counties. Hence, one does not need to move to a different county (and potentially find a different job) to find a higher-opportunity area for one's children.

[8] These positive fiscal externalities could potentially be offset by negative externalities on higher-income families

to move at birth from a neighborhood at the 25th percentile of the distribution of upward mobility within his county to a neighborhood at the 75th percentile, taxpayers would gain $41,000 through increased tax revenues over his lifetime.

Although the analysis in this paper focuses on identifying generalizable lessons about neighborhood effects, the strength of the tract-level data is that they allow users to reach different conclusions in different places. To take one example, in the Midwest, children who grow up in rural areas tend to have higher rates of upward mobility than those who grow up in urban areas. In the South, the pattern is reversed: upward mobility is higher in cities. As this example illustrates, instead of attempting to draw general conclusions about whether cities offer better opportunities for upward mobility, it is more informative to directly study the question in the area of interest. In this sense, the Opportunity Atlas provides a customized platform that we hope will help support researchers and policymakers seeking to improve opportunity in communities across the U.S.

The paper is organized as follows. Section II describes the microdata we use in our analysis. Section III describes the methods we use to construct tract-level estimates. Section IV presents a descriptive characterization of variation in outcomes across tracts and discusses implications for policy targeting. Section V focuses on the causal effects of neighborhoods and implications for moving to opportunity. Section VI concludes. Supplementary results and methodological details are provided in an online appendix.

II Data

The microdata we use in our analysis are identical to those used by Chetty et al. (2018); much of this section is therefore taken directly from Chetty et al. (2018, Section II).

We combine three sources of data housed at the Census Bureau: (1) the Census 2000 and 2010 short forms; (2) federal income tax returns in 1989, 1994, 1995, and 1998-2015; and (3) the Census 2000 long form and the 2005-2015 American Community Surveys (ACS). The Census short forms are designed to cover the entire population; the Census 2000 long form is a stratified random sample covering approximately one-sixth of households; and the American Community Survey is a stratified random sample covering approximately 2.5% of households in each year (U.S. Department of Commerce, Bureau of the Census 2000; U.S. Department of Commerce, Bureau of the Census 2003; U.S. Department of Commerce, Bureau of the Census 2014).

from living in more integrated neighborhoods, e.g. because of having fewer resources per capita. They could also be offset by externalities in the origin neighborhoods that families leave behind. We view the quantification of such spillover effects as an important direction for future work.

These datasets are linked by a unique person identifier called a Protected Identification Key (PIK) that is assigned by Census Bureau staff using information such as Social Security Numbers (SSN), names, addresses, and dates of birth. The record linkage algorithm used to assign individuals PIKs is described in Wagner and Layne (2014). The linkage procedure is very accurate and comprehensive: using datasets that have both SSNs and other identifiers, Layne et al. (2014) show that the error rate in assigning PIKs when one does not have SSNs (as in Census surveys) is typically below 1% for government datasets. In the 2010 Census, 90.3% of individuals are successfully assigned a PIK (Wagner and Layne 2014, Table 2). Bond et al. (2014) show that PIK rates vary slightly across population subgroups in the 2010 ACS, but exceed 85% in virtually all subgroups. All analysis in this paper is conducted using a dataset that contains PIKs but is stripped of personally identifiable information.

In the rest of this section, we describe how we construct our analysis sample, define the variables we use, and present summary statistics. See Online Appendix A of Chetty et al. (2018) for further details.

II.A Sample Definition

Our target sample frame consists of all children in the 1978-83 birth cohorts who were (1) born in the U.S. or are authorized immigrants who came to the U.S. in childhood and (2) whose parents were also U.S. citizens or authorized immigrants.[9] We construct this sample frame in practice by identifying all children who were claimed as a child dependent on a 1040 tax form at some point between 1994-2015 by an adult who appears in the 2016 Numident file and was between the ages of 15-50 at the time of the child's birth.[10] We then restrict the sample to children who were born between 1978-83, based on their record in the 2016 Numident. This sample definition excludes children who are unauthorized immigrants or who are claimed as dependents by unauthorized immigrants because unauthorized immigrants do not have SSNs and therefore do not appear in the Numident file.

We define a child's "parent" as the person who first claims the child as a dependent (between 1994-2015). This person must be supporting the child to claim him or her as a dependent, but may not necessarily be the child's biological parent.[11] If the child is first claimed by a single filer, the

[9] We also present some supplementary results, such as analyses of trends and movers' outcomes, using an extended sample that includes the 1978-1991 birth cohorts. We limit our analysis to individuals who are authorized immigrants because coverage rates of tax data for unauthorized immigrants are difficult to determine.

[10] Dependent claiming information is not available in tax returns from 1989. We impose the 15-50 age restriction to limit links to grandparents or other guardians who might claim a child as a dependent.

[11] An alternative method of identifying parents is to use information on relationships to household members in the

child is defined as having a single parent. For simplicity, we assign each child a parent (or parents) permanently using this algorithm, regardless of any subsequent changes in parents' marital status or dependent claiming. Virtually all children in the 1978-83 birth cohorts are linked to parents through this procedure, because nearly all children get claimed as dependents at some point in their childhood (Chetty et al. 2018, Online Appendix Table II). We limit our analysis to children born during or after 1978 because many children begin to leave the household starting at age 17 (Chetty et al. 2014, Appendix Table I), and the first year in which we have dependent claiming information is 1994.

Finally, we exclude the 3.8% of children for whom we have no address information during childhood (i.e., during or before the year in which they turn 23) because the addresses from which their parents filed their tax returns could not be mapped to a tract. The resulting primary analysis sample consists of 20.5 million children, which covers 96% of our target population. When reporting race-specific estimates, we exclude an additional 5% of children for whom race is missing because they could not be linked to the Census or ACS; however, these children are included in the estimates that pool all racial groups, because those estimates can be constructed purely using information from tax returns.

Chetty et al. (2018, Appendix B and Appendix Tables II-IV) show that this analysis sample provides an accurate representation of our target population by establishing that it has income distributions and demographic characteristics very similar to the ACS.[12] They also show that pre-tax income measures in the tax data are closely aligned with those in survey data. For example, the median income in 2015 of children with non-missing race who appear in both our analysis sample and the 2015 ACS is $33,370 based on the tax data, compared with $34,000 based on the ACS data. Individuals recorded as having zero income in the tax records (because they do not file and have no W-2s) have a median income of only $5,000 in the ACS, showing that tax records do not miss substantial amounts of income for non-filers.

2000 Census short form. We find that the tax- and Census-data based measures of parents are well aligned: for instance, among the children claimed as dependents by parents on a 1040 tax form in 2000, 93% live with the same parents in the 2000 Census. We use the tax data to identify parents because many of the children in the oldest cohorts in our sample have left their parents' houses by the 2000 Census.

[12] Chetty et al. (2018) impose slightly different sample restrictions given their objective of studying racial disparities: they include only individuals with non-missing race information and do not exclude children with missing address information during childhood.

II.B Variable Definitions

In this subsection, we define the variables we use in our primary analysis. We measure all monetary variables in 2015 dollars, adjusting for inflation using the consumer price index (CPI-U).

Parental Characteristics and Family Background.

Parental Income. Our primary measure of parent income is total pre-tax income at the household level, which we label parent family or household income.[13] In years where a parent files a tax return, we define household income as Adjusted Gross Income (as reported on the 1040 tax return) plus tax-exempt interest income and the non-taxable portion of Social Security and Disability benefits. In years where a parent does not file a tax return, household income is coded as zero.[14] Following Chetty et al. (2014), we define our baseline parental income measure as the mean of parents' household income over five years: 1994, 1995, and 1998-2000, as tax records are unavailable in 1996 and 1997.[15] We exclude children whose mean parent income is zero or negative (1.0% of children) because parents who file tax returns (as is required to link them to a child) reporting negative or zero income typically have large capital losses, which are a proxy for having significant wealth.

Parental Marital Status. We identify parents' marital status based on their tax filing status in the year the child is first claimed as a dependent by parents. We say that a child has a "father present" if one of the tax filers who claims the child as a dependent in that year is male and define "mother presence" analogously. Children claimed by two people in the year they are first claimed are defined as having two parents.

Locations in Childhood. In each year, parents are assigned an address based on the address from which they filed their 1040 tax return. For non-filers, we use address information from information returns such as W-2s, which are available beginning in 2003.[16] Addresses are coded as missing

[13] We use the term "household" income for simplicity, but we do not include incomes from cohabitating partners or other household members aside from the primary tax filer's spouse.

[14] Prior work (e.g., Chetty et al. 2014) has used information from W-2 forms to measure income for non-filers. We cannot follow that approach here since income data from W-2 forms are unavailable at the Census Bureau before 2005. However, this has little impact on results; Chetty et al. (2014) show that in 2000, the median W-2 income among parents who were non-filers was $29, and only 2.9% of parents do not file in a given year. Information from W-2s is more important when measuring the incomes of children in early adulthood, for whom we do have W-2 data at the Census Bureau.

[15] Formally, we define mean household income as the mother's individual income plus the father's individual income in each year of 1994-95 and 1998-2000 divided by 10 (or divided by 5 if we only identify a single parent). For parents who do not change marital status, this is simply mean household income over the 5 year period. For parents who are married initially and then divorce, this measure tracks the mean household incomes of the two divorced parents over time. For parents who are single initially and then get married, this measure tracks individual income prior to marriage and total household income (including the new spouse's income) after marriage.

[16] Address information from W-2s starts in 2003, but income amounts are not available until 2005.

in years when a parent does not file or does not have an information return. For children whose parents were married when they were first claimed as dependents, we prioritize the mother's location if marital status changes. Addresses are mapped to geographies such as Census tract or Census block using a geocoding algorithm developed by the Census Bureau (see Online Appendix A of Chetty et al. (2018) for details).

Race. We assign race and ethnicity to children using the information they report on the 2010 Census short form. If the child's PIK does not appear in the 2010 Census microdata, we use the 2000 Census short form; if the child does not appear in the 2000 Census, we then use the ACS. We aggregate race and ethnicity categories into a Hispanic ethnicity category and a set of non-Hispanic races: White, Black, Asian, American Indian or Alaskan Native, and Other, following the Office of Management and Budget (1997). Individuals who report two or more races, Native Hawaiian or Pacific Islander, or Some Other Race are classified in the "Other" category.

Gender and Age. Gender and age are obtained from the Numident file.

Children's Outcomes Observed in Tax Records or Census Short Form.

We report statistics at the Census tract, county, and commuting zone (CZ) levels for the following outcomes, which we observe in the full sample using data from either tax records or the Census short form.

Income. We define children's annual household income in the same way as parents' income except in our treatment of non-filers. Since W-2 data are available for the years in which we measure children's incomes, we define income for a child who does not file a tax return as wage earnings reported on form W-2. We define children's individual incomes as their own W-2 wage earnings plus self-employment and other non-wage income, which we define as Adjusted Gross Income minus total wages reported on form 1040 divided by the number of tax filers (thereby splitting non-wage income equally for joint filers). In years in which children have no tax return and no information returns, both individual and household income are coded as zero. We measure children's individual and household incomes as their mean annual incomes in 2014 and 2015, when children are between the ages of 31 and 37. In addition to these baseline definitions, we also report mean incomes by age at ages 24, 26, and 29.

Upper-Tail Income. We define indicators for children reaching the top 20% and top 1% of the income distribution using the baseline definitions of individual and household income above.

Employment. We use two measures of employment, one based on the tax data and one based on the ACS (defined below). In the tax data, children are defined as working if they have positive

W-2 income in 2015. In addition to this baseline definition, we also report employment rates in the tax data by age for ages 24, 26, 29 and 32.

Marriage. A child's marital status is measured based on whether he or she files a tax return jointly in 2015. In addition to this baseline definition, we report marriage rates by age for ages 24, 26, 29, 32.

Incarceration. Using data from the 2010 Census short form, we define an individual as incarcerated on the day of the Census (April 1, 2010) based on whether he or she lives in any of the following types of group quarters: federal detention center, federal prison, state prison, local jail, residential correctional facility, military jail, or juvenile correctional facility. This variable is defined only for children who appear in the 2010 Census.

Teenage Birth. We define a woman as having a teenage birth if she ever claims a dependent who was born while she was between the ages of 13 and 19. This measure is an imperfect proxy for having a teenage birth because it only covers children who are claimed as dependents by their mothers. Nevertheless, the aggregate level and spatial pattern of teenage births in our data are closely aligned with estimates based on the American Community Survey, with an (unweighted) correlation across states of 0.79.

Spouse's Income. For children who were married in 2015, we define spouse income analogously to the child's own individual income using an average of 2014 and 2015 individual income. To capture spouses who are not within our primary analysis sample, we include all spouses born between 1973 and 1989.[17]

Living in a Low-Poverty Neighborhood in Adulthood. We measure children's locations in adulthood based on the address from which they file tax returns in 2015. For non-filers, we obtain address information from W-2 forms and other information returns. If no address information is available in 2015, we use the most recent year in which an address is available. Among children with a non-missing address in adulthood, we identify those living in a "low-poverty" neighborhood as those living in a tract with less than 10% of people living in poverty, as defined by the publicly available Census 2010 estimates.

Staying in Childhood CZ or Tract. Children are defined as staying in their childhood location (tract or CZ) if their most recent address matches any tract or CZ that they lived in during childhood (before age 23). This variable is defined among the set of children with non-missing

[17]Since we cannot link children to parents prior to the 1978 birth cohort, we define spouse income percentile ranks relative to all individuals in the relevant birth cohort, not just those individuals linked to parents (see Section III).

addresses in adulthood.

Staying with Parents. Children are defined as staying with their parents if their 2015 address matches their parents' 2015 address. This variable is defined among the set of children with non-missing addresses in 2015 and whose parents have non-missing addresses in 2015.

Income for those who Stay in Childhood CZ. We measure income among the children who stay in their childhood CZ using the baseline definitions of household and individual income described above. These variables are defined among the set of children who stayed in their childhood CZ as adults.

Fraction of Childhood Years Spent in Tract. We calculate the fraction of childhood years that a child spends in a tract as the total number of years we observe the child in a given tract (based on their parents' tax returns) divided by the total number of years for which address data is available for the child's birth cohort before age 23. For example, for the 1983 cohort, address information before the child is 23 is available in 12 tax years (1989, 1994-1995, and 1998-2006). We construct the fraction of years in tract variable for the 1983 cohort by dividing the total number of years the child is observed in a given tract by 12.

Children's Outcomes Observed in ACS or Census Long Form.

We report statistics at the county and CZ (but not tract) level for the following outcomes, which we observe for the subsample of individuals who appear in the Census 2000 long form or the 2005-2015 ACS.

Employment. In the ACS, children are defined as working if they report positive weeks worked in the past year. This and all other employment-related ACS measures described below are defined only among children who receive the ACS at age 30 or later.

Hours Worked Per Week. Annual hours of work are measured in the ACS as the product of hours worked per week and weeks worked per year. Individuals report weeks worked in bins; we use the midpoint of the bin to assign each individual a single value (e.g., those who choose "27 to 39 weeks" are assigned a value of 33). We convert the annual measures to average weekly hours worked by dividing annual hours worked by 51 (the midpoint of the top bin of weeks worked). Those not working in any week are coded as having zero hours of work.

Hourly Wage. Hourly wages are measured in the ACS by dividing reported annual wage and salary income by annual hours worked. The hourly wage is coded as missing for those with zero hours worked.

Educational Attainment. We measure children's educational attainment based on the high-

est level of education they report having completed in the ACS or the 2000 Census long form (prioritizing the ACS, since it is more recent). We say a student dropped out of high school if their educational attainment is "12th grade- no diploma" or less (hence, those with GEDs are not counted as having dropped out). We define some college attendance as having obtained "at least some college credit." We define community college completion as having an Associate's degree and 4-year college completion as having a Bachelor's degree. Graduate degree completion is defined as having a Master's, professional, or doctorate degree. All education variables are defined as having at least that level of education or higher. High school completion is defined among individuals who are at least 19 at the time they are surveyed. When measuring some college attendance, community college completion, and college completion we require that individuals are at least 24. We require that respondents are at least 30 years old at the time of the survey when measuring graduate degree completion.

Public Assistance Receipt. Children are coded as receiving public assistance in adulthood if they report positive public assistance income in the ACS. This variable is defined among individuals who receive the ACS at age 30 or older.

Income for Children with Native or Immigrant Mothers. We measure income among the children whose mothers are U.S. natives or immigrants using the baseline definitions of household and individual income described above. Children are defined as having a "native-born" mother if their mother was surveyed in the 2000 Census long form or the ACS and reported being born in the United States in either survey. Children are defined as having an immigrant mother if their mother received either survey and reported being born outside of the United States.

II.C Summary Statistics

Table I lists the variables included in the Opportunity Atlas and provides summary statistics for those variables in our primary analysis sample.[18] We present statistics for the full population as well as by gender. Online Appendix Table I presents analogous statistics by race and ethnicity.

Pooling all races and ethnicities, the median household income of parents in our primary analysis sample is $56,730. Parental income and marital status vary sharply across racial and ethnic groups, as is well known from prior work. For example, median household income is $71,470 for whites, $29,600 for blacks, and $33,470 for Hispanics. 79.6% of white children are raised in two-parent families, compared with 32.5% of black children and 57.2% of Hispanic children.

[18] Table I presents statistics for the 96.2% of children in our sample who have non-missing information on childhood tract (for at least one year of their childhood), the sample we use to estimate our tract-level statistics.

The median household income among children in 2014-15 (between the ages of 31-37) is $42,360, while median individual earnings is $29,440. 1.5% of children are incarcerated on April, 2010 (between ages 27-32). 19.7% of women have a teenage birth. 69.5% earned some college credits. These outcomes again vary sharply across subgroups; for example, incarceration rates range from 0.2% for white women to 10.3% for black men, while teenage birth rates range from 13.5% for white women to 41.3% for black women.

III Tract-Level Estimates: Methodology

In this section, we describe how we construct our publicly available tract-level estimates. We first define our estimands of interest, then describe our estimator and its properties, and finally discuss the estimates we release publicly.

Let y_i denote an outcome for child i, such as his or her income in adulthood. Throughout our analysis, we measure both children's and parents' incomes using percentile *ranks* rather than dollar levels. Chetty et al. (2014) show that measuring income using ranks yields more robust estimates by reducing the influence of outliers and mitigating lifecycle bias because individuals' income ranks stabilize earlier in their lives than their income levels (Solon 1999, Haider and Solon, 2006, Grawe, 2006). We define child i's percentile rank y_i based on his position in the *national* distribution of incomes (measured between ages 31-37 in our baseline analysis) relative to all others in his birth cohort who are in our primary analysis sample.[19] Similarly, we measure the percentile rank of the parents of child i, $p(i)$, based on their position in the national distribution of parental income for child i's birth cohort. We always hold the definition of these ranks fixed based on positions in the national aggregate income distribution, even when analyzing subgroups or local areas.

Our objective is to estimate children's expected outcomes given their parents' income percentile p, racial and ethnic group r, and gender g, conditional on growing up in Census tract c from birth:

$$\bar{y}_{cprg} = E[y_i | c(i) = c, p(i) = p, r(i) = r, g(i) = g]. \tag{1}$$

We focus on characterizing how the neighborhoods in which children *grow up* affect their outcomes, which may differ from the neighborhoods in which they live as adults. We focus on childhood neighborhoods because of prior evidence that rates of intergenerational mobility depend on where children grow up rather than where they live as adults (Chetty et al. 2016; Chetty and Hendren 2018a).

[19] We include children with zero income, assigning them the mean rank of the individuals in that group. For example, if 10% of a birth cohort has zero income, all children with zero income would receive a percentile rank of 5.

There are two empirical challenges in estimating the conditional mean $\bar{y}_{prgc}$ in practice. First, there are insufficient observations to estimate $\bar{y}_{prgc}$ non-parametrically in each parent income percentile by race by gender by tract cell. Second, most children do not grow up in a single tract from birth, forcing us to account for movement across tracts when estimating mean outcomes. We construct an estimator that addresses each of these two challenges in turn.

Parametric Estimator. To address the first challenge, we estimate the conditional expectation of children's outcomes given their parents' household income using a univariate regression in each tract by gender by race cell. We choose the parametric form of the regression by examining the relationship between outcomes and parental income rank non-parametrically at the national level. To illustrate, Figure Ia plots the mean household income rank of children within each percentile bin of the parent income distribution for whites, blacks, and Hispanics, $E[y_i|p(i) = p, r(i) = r]$. Figure Ib replicates Figure Ia using incarceration rates as the outcome, restricting attention to male children. There are significant non-linearities in the conditional expectation functions, especially for incarceration.

To capture these non-linearities, we regress children's outcomes on a tract-invariant transformation of parental income rank $f_{rg}(p_i)$:

$$y_i = \alpha_{crg} + \beta_{crg} \times f_{rg}(p_i) + \varepsilon_i, \tag{2}$$

where $f_{rg}(p_i)$ is estimated using a lowess regression of $\bar{y}_{prg}$ on p in each race by gender subgroup at the national level. Intuitively, we fit a lowess regression to the non-parametric conditional expectation functions plotted in Figure I to find a transformation of parental income rank $f_{rg}(p_i)$ that renders the relationship between y_i and $f_{rg}(p_i)$ linear at the national level. We then run a linear regression of the outcome on transformed parental income in each tract-race-gender cell as in (2) and use the predicted values of this regression at each percentile p as our estimate of $\bar{y}_{cprg}$.[20] We obtain standard errors for $\hat{y}_{cprg}$ directly from the regression in (2), assuming homoskedastic errors ε_i at the individual level and taking f_{rg} as known with certainty.

This estimation approach, which is analogous to a Box-Cox transformation (Box and Cox 1964), allows us to summarize the conditional expectation function in each tract using just two parameters,

[20] Mechanically, in each tract-race-gender cell, we estimate α_{crg} and β_{crg} in (2) by regressing the outcome y_i on the predicted values from a lowess regression (with bandwidth 0.3) of $\bar{y}_{prg}$ on p estimated by race and gender at the national level. For outcomes obtained from the ACS samples and for all outcomes for small racial subgroups (Asians, American Indians, Other), the lowess estimates at the national level are themselves noisy and are sometimes non-monotonic. In these cases, if the race-by-gender specific lowess fit exhibits any non-monotonicities, we use the functional form estimated from the lowess regression of the outcome on parent income percentile pooling all races and genders.

thereby yielding precise estimates of expected outcomes. The assumption underlying this estimator is that the shape of the conditional expectation of the outcome given parental income at the national level is preserved in each tract up to an affine transformation (within each race-gender subgroup). We evaluate this assumption by adding a quadratic term $(f(p_i))^2$ to (2) and examining whether the estimates change. In general, we find the estimates remain very similar; for example, tract-level estimates of children's mean household income ranks given parents at $p = 25$ based on the linear and quadratic models are correlated 0.99. We also estimate the conditional expectation functions non-parametrically in large CZs and counties and find that non-parametric estimates at these levels of geography are well approximated by an affine transformation of the national relationship. Hence, although the shape-preservation assumption underlying our estimator is strong, it appears to be a reasonable approximation in practice.

Exposure Weighting. If children spent their entire childhood in a single tract c, (2) would yield an unbiased estimate of $\bar{y}_{prgc}$ under the assumption described above. In practice, many children move across tracts in childhood.[21] To address this issue, first note that the CZ in which one grows up has causal effects on earnings and other outcomes in adulthood until approximately age 23 (Chetty and Hendren 2018a), a finding that we replicate at the tract level in Section V below. We therefore assign children to tracts in proportion to the amount of time they spend before age 23 in each tract over the years observed in our sample. For example, if a child spent half of his childhood in the tract 1 and half of his childhood in tract 2, he would effectively receive 50% weight in each of the two tracts $c = \{1, 2\}$ in the regression in (2).[22]

The estimates $\hat{y}_{cprg}$ obtained from this exposure-weighted regression differ from our target $\bar{y}_{prgc}$ – the mean outcomes of children who spend their *entire* childhood in tract c – in two ways. First, they assign equal weight to all years of childhood, effectively assuming that each year of exposure between ages 0 and 23 contributes equally to children's long-term outcomes. In practice, neighborhoods appear to have slightly larger effects during adolescence than in the earliest years of childhood (see Figure XI below). However, the differences in marginal effects by age are small enough that our baseline estimates based on equal weighting are very similar to estimates obtained

[21] Chetty and Hendren (2018) address this issue by restricting the sample to "permanent residents" – children who never move across CZs during their entire childhood. That approach yields imprecise estimates at the tract level because few children stay in a single tract for their entire childhood; for instance, among children born in 1991 (for whom we have a near-complete address history from birth to age 23), only 27% of children spend their entire childhood in a single tract.

[22] More precisely, we estimate the regression in (2) weighting by the total number of years the child is claimed as a dependent in tract c before age 23. Since the tax records begin in 1989, the earliest age at which we observe children's locations in our primary analysis sample is age 6 (for the 1983 birth cohort).

using age-specific weights corresponding to the marginal age-specific treatment effects estimated below (e.g., correlation = 0.997 for mean household income rank given $p = 25$ across tracts).

Second, the estimates $\hat{y}_{cprg}$ are biased because the outcomes y_i observed for children who live in other tracts incorporate the other tracts' causal effects, which may differ from $\bar{y}_{prgc}$. In general, we expect this bias to lead us to understate the true variance in observed outcomes across tracts relative to what one would observe if children grew up in a single tract for their entire childhood. Intuitively, children who spend part of their childhood in a tract with very positive observed effects will tend to spend the rest of their childhood in a worse tract on average because of mean reversion, pulling down y_i (and hence $\hat{y}_{cprg}$) relative to $\bar{y}_{prgc}$. This bias also turns out to be small empirically, for two reasons. First, although children move, most children spend the majority of their childhood in one tract. Children observed in a given tract spend 75% of their childhood in that tract (using the same weights as in our regression specification). Second, even when children move, they tend to move to an area very similar to the one in which they previously lived. Among children who move, the correlation between $\bar{y}_{prgc}$ and $\bar{y}_{prgc'}$ in the two tracts c and c' in which they spend the most time is 0.86 for mean household income rank given $p = 25$.[23] Given these parameters, if the variation across tracts were driven entirely by causal effects of place that are proportional to childhood exposure, the correlation between $\hat{y}_{cprg}$ and $\bar{y}_{cprg}$ would be

$$Corr(\hat{y}_{cprg}, \bar{y}_{prgc}) = Corr(0.75\bar{y}_{prgc} + 0.25\bar{y}_{prgc'}, \bar{y}_{prgc}) = 0.75 + 0.25 \times 0.86 = 0.96. \tag{3}$$

We conclude that the exposure-weighted regression estimates obtained from the parametric model in (2) are likely to provide accurate estimates of the conditional means of interest in (1).[24]

Publicly Available Estimates. After constructing the estimates described above for all tracts and subgroups, we take three final steps to construct the publicly available statistics.

First, we suppress estimates that are based on 20 or fewer children both to protect privacy and because such estimates typically have very large standard errors. Since most subgroups tend to be concentrated in specific tracts, this suppression leads us to omit relatively little data: the publicly available statistics that pool racial groups and genders cover 99.9% of the individuals in the microdata and the race-specific statistics cover 96% of individuals.

[23] To eliminate spurious correlations driven by having the same individual's data appear in both tracts, we estimate this correlation by first constructing two sets of tract-level estimates, randomly splitting our sample into two at the individual level, and then estimating the correlations *across* these two samples.

[24] The correlation of 0.96 is not an exact estimate because it assumes individuals move at most once for simplicity and, more importantly, it assumes that the observational variation across places is entirely due to causal childhood exposure effects. In the presence of selection effects (which may not be proportional to childhood exposure), the correlation might differ. Nevertheless, this calculation illustrates that the exposure weighted estimates are unlikely to differ substantially from the mean outcomes one would observe if children did not move across tracts.

Second, to further reduce the risk of privacy loss, we add noise to the estimates we release. We release estimates for each race-gender-tract cell at two parent income percentiles, typically $p = 25$ and $p = 75$.[25] We add independent, normally distributed noise to each of these estimates. The standard deviation of the noise distribution is chosen based on the sensitivity of the estimates to a single individual's data, as in the literature on statistical disclosure limitation (Dwork 2008; Abowd and Schmutte 2015). We calculate the sensitivity of the estimates by calculating the maximum change in the estimate $\hat{y}_{cprg}$ that can result from adding a single observation to the relevant tract-race-gender cell.[26] We then compute the maximum sensitivity of the estimates across all tracts within a given state. We add noise with SD proportional to this maximum state-level sensitivity divided by the number of observations in the relevant subgroup (n_{crg}) and a parameter ε that controls the degree of disclosure risk, which we set at $\varepsilon = 8$. We add noise to the standard errors of the estimates following an analogous procedure. Further details on this noise-infusion procedure are provided in Chetty et al. (2018).

We report the SD of the noise added to each tract-level estimate in the publicly available data. Importantly, because the SD of the noise we add is proportional to $1/n_{cprg}$, it is typically significantly smaller than the inherent noise in the estimate from sampling error, which is proportional to $1/\sqrt{n_{cprg}}$.[27]

Third, to facilitate interpretation of our estimates of mean ranks, we provide a crosswalk that maps percentile ranks to dollar values for each monetary variable (e.g., individual income, household income). This crosswalk is constructed based on the income distribution of children in our primary analysis sample in 2014-2015 for our baseline income measures. We report estimates of the dollar value (in real 2015 dollars) corresponding to the mean predicted rank of children using this crosswalk.[28]

The final tract-level estimates in the Opportunity Atlas are available through an interactive mapping tool and in downloadable flat files. A complete list of the variables along with a codebook are available in the Online Data Tables. For each of the outcome variables listed in Section II.B, we

[25] In tracts where fewer that 10% of parents are above the median ($p > 50$), we release estimates at $p = 1$ and $p = 50$; conversely, if fewer than 10% of parents are below the median, we release estimates at $p = 50$ and $p = 100$.

[26] Because all of the outcome variables we analyze are bounded, this value is well-defined.

[27] Because the noise added is independent across subgroups, the estimates we report do not aggregate perfectly: for example, the estimates for men and women in a tract may not aggregate to the pooled estimate we report. The addition of noise can also result in values that fall outside the bounds of the original variables – for example negative incomes or employment rates above 100%. We report these values directly in the raw downloadable data so that researchers can compute unbiased aggregate moments, but bottom-code and top-code values at their natural bounds in the data visualization tool.

[28] Of course, this estimate cannot be interpreted as the mean income of children because the rank transformation is not linear; it merely provides a useful way to interpret the mean rank estimates.

release means by parental income group, race, and gender as well as unconditional means (pooling all parental income levels) by race and gender. We report estimates at 5 parental income levels: lowest income ($p = 1$, incomes of approximately \$1,000), low income ($p = 25$, \$27,000), middle income ($p = 50$, \$55,000), high income ($p = 75$, \$93,000), and highest income ($p = 100$, \$1,100,000).

We also release estimates at the county and CZ levels, which are constructed using methods analogous to those described above. Because outcomes in the American Community Survey (such as college attendance) are available for a small sample of individuals, we only report estimates at the county and CZ levels for those outcomes.

IV Observational Variation and Targeting

In this section, we present a descriptive characterization of children's outcomes across Census tracts. This analysis is motivated by the fact that many policies aimed at improving children's outcomes are targeted based on neighborhood characteristics. For example, the Tax Cuts and Jobs Act of 2017 provides preferential tax treatment for investment in Opportunity Zones – low-opportunity neighborhoods designated by states. Children who live in disadvantaged Census tracts are granted preferential status for admission to Chicago's selective public high schools. Head Start centers often locate in high-poverty neighborhoods to serve children with the most limited opportunities. And environmental remediation projects, such as lead removal, target specific Census tracts partly based on proxies such as poverty rates and the age of the housing stock, which are predictive of lead exposure.

In the terminology of models of optimal taxation, all of these policies use neighborhoods characteristics as a predictor or "tag" for being disadvantaged (Akerlof 1978; Nichols and Zeckhauser 1982). By targeting disadvantaged neighborhoods, these policies have the potential to more efficiently provide resources to those who need them most. However, it is not clear that the proxies that are currently used for targeting – such as neighborhood poverty rates – necessarily identify the areas where children have the lowest chances of rising out of poverty. Here, we explore whether our new measures of children's outcomes by Census tract can be useful for targeting policies to neighborhoods where children have the poorest prospects of climbing the income ladder.[29]

From the perspective of identifying predictors for children's future incomes, the observed out-

[29] Of course, the policies described above may be designed with other objectives in mind beyond predicting where children grow up to have the lowest incomes. We do not take a normative stance on the right objective function; our goal here is simply to evaluate whether and how the Opportunity Atlas measures of children's outcomes can be useful *if* one's objective were to identify areas where children have poor prospects for upward mobility.

comes in each Census tract are of direct interest. It does not necessarily matter whether those outcomes arise from the causal effect of the neighborhood itself or from selection in who chooses to live in that neighborhood.[30] Motivated by this reasoning, we present a descriptive characterization of the observational variation in children's outcomes across Census tracts.

We organize our analysis into four parts, each of which addresses a different question about the utility of our tract-level estimates as tags and, more broadly, sheds new light on the nature of neighborhood-level variation in children's outcomes. First, we characterize the amount of variation in upward income mobility across tracts to quantify the extent to which location is a useful tag for future outcomes. Second, we examine whether tract-level outcomes are heterogeneous across subgroups to understand whether there are gains from targeting based on the interaction of location and other individual characteristics. Third, we study the correlations between children's outcomes and the observable characteristics of neighborhoods to determine whether our new measures provide additional information beyond the measures currently used to target place-based policies. Finally, we analyze how our tract-level estimates change over time to assess whether our estimates of upward mobility (based on children born in the 1980s) provide useful information for targeting policies to children growing up today.

IV.A Variation in Outcomes Across Census Tracts

We begin by characterizing the variation in children's outcomes across Census tracts using some illustrative examples. In much of our analysis, we focus on children's mean household income ranks given parents at the 25th percentile of the national income distribution (roughly $27,000), which we term "upward mobility" for convenience.

Illustrative Example: Los Angeles. Figure IIa maps upward mobility in the Los Angeles metro area, by the tract in which children grew up, estimated using the method described in Section III. Children's outcomes vary substantially across tracts, with much lower rates of upward mobility in the center of Los Angeles than in surrounding areas. Children growing up in the bottom decile of neighborhoods shown on this map in terms of upward mobility (shown in the darkest red colors) have predicted mean income ranks in adulthood below 35.0 (roughly $25,700). Those growing up in the highest decile of neighborhoods (shown in the darkest blue colors) have predicted income

[30] More precisely, in neoclassical models of optimal taxation, the optimal policy does not depend upon the reason that location predicts future incomes. In other settings, one may wish to distinguish between areas that have poor outcomes because of selection vs. causal effects; for instance, interventions to improve institutions in a given area may be especially valuable if they have poor causal effects. Our point is simply that for certain questions, the observational variation in outcomes itself is of interest; we turn to distinguish causal effects from selection in our second application.

ranks above 49.7 (roughly $42,000).[31]

Part of the variation across neighborhoods in Figure IIa is due to differences in rates of upward mobility across racial and ethnic groups. In particular, Chetty et al. (2018) document that black children – and black men in particular – have much lower rates of upward mobility even if they grow up on exactly the same block. Since central Los Angeles has a much larger black population than the suburbs, the lower rates of upward mobility in this part of the city partly reflect the challenges that black Americans face in climbing the income ladder rather than something specific about those neighborhoods.

Even within racial groups, however, there is considerable variation across neighborhoods (click for an interactive illustration). Panel B of Figure II plots upward mobility for black male children. Black men's rates of upward mobility vary widely within central Los Angeles.[32] To take one example, low-income black men who grow up in Watts reach the 17th percentile of the household income distribution as adults on average.[33] This corresponds to a household income of roughly $7,300 (s.e. $2,600) in 2015, when these men are in their mid-thirties.[34] These low income levels are perhaps not surprising given the widely documented challenges that Watts has faced in terms of poverty and violence. What may be more surprising is that in nearby Compton – about 2.3 miles South of Watts – the outcomes of black men growing up in families with comparable incomes are quite different. In Central Compton, black men grow up to have average household incomes of $19,100 (s.e. $2,100).

The disparities across areas are particularly stark for incarceration. Since incarceration rates vary non-linearly with parental income and are much higher among children growing up in the very poorest families (Figure Ib), we focus on incarceration rates among children growing up in the lowest-income families (bottom 1%). Figure IId shows that 44.1% (s.e. 9.3%) of black males growing up in the poorest (bottom 1%) families in Watts were incarcerated on April 1, 2010.[35]

[31] An important caveat to keep in mind is that some of these estimates are based on statistical extrapolation, as there are few low-income families living in coastal areas of Los Angeles such as Santa Monica. Nevertheless, under the maintained assumptions of our statistical model described in Section III above, these estimates constitute a consistent prediction of how well low-income children would do if they were able to grow up in such areas.

[32] We have insufficient data to report estimates in the suburbs for black men, reflecting the degree of racial segregation in Los Angeles.

[33] This particular statistic is for a tract in Watts that contains the Nickerson Garden public housing project. For convenience, we refer to tracts by neighborhood names in this paper, but quote statistics from specific tracts within those neighborhoods.

[34] The standard errors we report include both sampling error and the error due to noise infusion to protect privacy (see Section III). For monetary variables, our regression models yield standard error estimates in ranks; we translate these estimates into dollars by averaging the absolute changes in dollar income resulting from moving up or down the income distribution in 2015 by one rank standard error starting from the point estimate.

[35] This is not an out-of-sample extrapolation, as the mean household income rank of black parents in our sample

In contrast, 6.2% (s.e. 5.0%) of black men who grew up in the lowest-income families in central Compton were incarcerated on April 1, 2010.

National Statistics. Table II generalizes this example and shows that there is considerable variation between nearby neighborhoods not just in Los Angeles but throughout the United States. Panel A shows statistics on the distribution of upward mobility across tracts, pooling all races (column 1) and separately for each race (columns 2-6). The first row of the table reports mean upward mobility across all tracts, weighting by the number of children from below-median income households in each tract. On average, children with parents at the 25th percentile reach the 40th percentile of the household income distribution, consistent with Figure Ia.

The second row of Table IIa reports the raw standard deviation (SD) of our estimates across tracts in the U.S. Pooling all racial groups, the SD of our estimates of upward mobility ($\hat{y}_{cprg}$) is 6.50 percentiles. Part of the variation in $\hat{y}_{cprg}$ is due to noise – both from sampling error (as there are only approximately 500 children per tract in our sample on average) and the random noise added to the estimates to protect privacy (see Section III). As a result, $SD(\hat{y}_{cprg})$ overstates the degree of variation in the conditional expectation of children's outcomes ($\bar{y}_{cprg}$) across tracts, $SD(\bar{y}_{cprg})$.

Under the statistical model in Section III, we can estimate the signal variance $Var(\bar{y}_{cprg})$ by subtracting the variance due to sampling error and the variance due to noise infusion from $Var(\hat{y}_{cprg})$, as both of these errors are orthogonal to our point estimates. The variance of the noise added to protect privacy is exogenously specified. We estimate the variance due to the sampling error as the mean standard error squared, using the standard errors from the regression specification in (2). In the pooled sample, the noise SD from these two sources is 1.95 percentiles, as shown in the third row of Table IIa. Hence, $(1.95/6.50)^2 = 9\%$ of the raw variance in our tract-level estimates is due to noise and 91% is due to signal, as shown in the fourth row of Table IIa.

These estimates imply that the signal SD of upward mobility across tracts is 6.20 percentiles (approximately \$6,686), as shown in the fifth row of Table IIa, showing that there is substantial variation across tracts nationwide, as illustrated by the maps of Los Angeles. We also find substantial variation across areas within racial groups, especially for whites, Asians, and American Indians, for whom the signal SD is around 5.7 percentiles. There is less variability for black and Hispanics across areas, with an SD of approximately 3.3 percentiles. Naturally, the reliability of the race-specific estimates is slightly lower (around 0.7-0.8) because samples are smaller for racial

in Watts is 21.

subgroups.

Panel B of Table II replicates the preceding analysis for incarceration. Here, we again find substantial variation across tracts relative to the mean. This is especially the case for low-income black men, for whom the signal SD in incarceration rates is 4.1 percentage points, consistent with the variation observed in Figure Ib. Reliability for incarceration is generally lower than for income because it is a relatively rarely observed outcome. However, as illustrated by the standard errors of the estimates reported in Figure IId, there are large outliers in rates of incarceration (such as Watts) that can be detected with great precision.[36]

Panel C of Table II considers mean household income for children with high-income parents (parents at the 75th percentile). We continue to find substantial variation across neighborhoods in the outcomes of children from relatively high-income families, although the degree of variation is slightly smaller than at low income levels, especially as a percentage of income.

Geographical Decomposition. Next, we decompose the variation across tracts geographically, by estimating the fraction of the signal variance across tracts that is within counties, across counties within commuting zones, and between commuting zones.[37] We estimate these variance components as the adjusted R-squared from regressions of the tract-level estimates on CZ and county fixed effects, removing the variance due to noise. Figure III shows that 32% of the variation in upward income mobility across tracts, pooling all racial groups, is explained by variation at the CZ level, while 13.5% is across counties within CZs. The remaining 54.5% of the variation in tract-level upward mobility is across tracts within counties. This implies that prior studies of the geography of intergenerational mobility (e.g., Chetty et al. 2014) that focus on variation across counties or CZs miss a substantial share of the variation in outcomes across places that one can observe when one "zooms in" to finer levels of geography. The standard deviation of the within-county variation in upward mobility is 4.65 percentiles ($5,000), as reported in the last row of Table IIa, consistent with the degree of variation evident within Los Angeles county in Figure IIa.

One natural hypothesis for the variation across tracts within counties is that children in different parts of a county attend different schools. As a simple method of assessing the potential explanatory power of schools, we examine the fraction of variance that is across tracts within high school

[36] For subgroups that have small sample sizes and very low incarceration rates, such as Asian men, there is essentially no signal in the estimates – reliability is 0 – and we therefore omit these data from the interactive tool and analyses that follow.

[37] Commuting zones are aggregations of counties analogous to metropolitan statistical areas, but provide a complete partition of the entire United States.

catchment areas vs. between high school catchment areas.[38] Figure III shows that 28% of the total variance in outcomes – and about half of the local tract-within-county variation – can be explained by school catchment area fixed effects.[39] Conditional on race, school boundaries explain less than half of the variance across tracts within counties. School boundaries explain more of the variation when pooling racial groups than within race because racial segregation tends to occur along school boundaries. In sum, although a significant share of the tract-level variation in outcomes could potentially be due to school effects, there is clearly substantial variation in outcomes even among children who attend the same school.

Overall, the analysis above shows that the neighborhood in which a child grows up is a significant predictor of his or her later life outcomes, even at a very local level. As a rough benchmark, note that a 1 SD increase in parental income rank is associated with approximately a 10 percentile increase in children's household income ranks. Since the signal SD of upward mobility across tracts is 6.2 percentiles, growing up in a tract with 1 SD higher upward mobility is associated with an income gain that is roughly equivalent to the income gain from a 0.62 SD increase in parental income. Given that parental income is used as a screen for many policies, this comparison suggests that the outcome-based measures of neighborhood quality constructed here may also be valuable tags.

IV.B Multi-Dimensional Neighborhoods

Neighborhoods are often conceptualized in a one-dimensional manner – as either "good" or "bad" for everyone, for all outcomes. In this subsection, we investigate whether that simplification is a good approximation. We again begin with the example of Los Angeles and then present national statistics.

[38] We assign Census tracts to high school catchment areas in 2017 using data generously provided to us by Peter Bergman on the intersection of Census tracts with high school catchment boundaries in 2017, obtained from Maponics (2017); see Online Appendix A for details. We match 71,720 tracts to school catchment zones, covering roughly 97% of the population. Since school catchment areas do not perfectly nest Census tracts, we assign tracts to the school catchment zone that contains the largest share of their land area. Using information on exact school catchment boundaries in Mecklenburg County, NC we estimate that only 9.6% of the population gets misclassified into the wrong school catchment area using this approach because high school catchment boundaries follow tract boundaries fairly closely (see Online Appendix Figure I).

[39] Insofar as there is spatial autocorrelation in outcomes across tracts for reasons unrelated to schools, this estimate likely provides an upper bound on the portion of the variance in outcomes that can be attributed to schools, since any randomly drawn set of contiguous tracts would share a common variance component in the presence of spatial autocorrelation. However, in the other direction, our use of 2017 high school catchment boundaries may lead us to understate the role of schools because they do not reflect the boundaries faced by children in our sample, who attended school in the 1990s and early 2000s. In practice, tract boundaries appear to be reasonably stable over time: 87% of tract pairs that fell on different sides of school catchment boundaries in 2002 in Charlotte did so in 2017 as well. Moreover, when examining variation in outcomes for more recent birth cohorts up to the 1989 birth cohort, we find no evidence that schools explain a larger share of the variance for more recent cohorts.

Heterogeneity Across Demographic Groups. Figure IIc presents a map of upward mobility for black women in central Los Angeles. Black women have considerably higher rates of upward mobility than black men growing up in the same neighborhoods. For example, black women who grow up in low-income (25th percentile) families in Watts have a mean household income rank of 29.0 in adulthood, corresponding to an income of $19,489 (s.e. $1,985) – nearly three times higher than income of black men who grow up in exactly the same area. But it is not always the case that black women earn much more than black men: in Compton, black men earn $19,141, while black women earn $21,509. Watts and Compton are roughly equivalent in terms of earnings outcomes for women, but differ sharply for men, illustrating that a "one-dimensional" comparison of these neighborhoods that does not account for heterogeneity by gender can be quite misleading. We find analogous heterogeneity by race and ethnicity. For example, 4.7% of Hispanic males in the lowest income (p1) families who grew up in Watts were incarcerated on April 1, 2010 – far less than the 44.1% rate for black males.

Panel A of Table III generalizes these examples by presenting correlations of upward mobility across racial and ethnic groups.[40] We estimate these correlations across tracts within CZs by demeaning all variables by CZ and weighting by the number of low-income children in each tract. The correlations are all positive – places where one racial group does better tend to have better outcomes for other racial groups as well. However, they are far from 1, showing that neighborhoods are not well described by a single factor model. We find similar results when examining heterogeneity in outcomes by parental income level. This is shown in the last column of Table III, which lists the correlation between mean household income ranks for children with parents at the 25th and 75th percentiles for each racial group. In general, children from high-income families tend to have higher incomes in areas where children from low-income families do better, but the correlations are again well below 1.[41]

Heterogeneity Across Outcomes. We find heterogeneity not just across groups but also across *outcomes* for a given group. Panel B of Table III shows correlations of mean outcomes across tracts for children with parents at the 25th percentile.[42] We control for race when estimating these

[40] These correlations are signal correlations; we adjust for attenuation in the raw correlations due to sampling error and noise infusion by dividing the raw correlation the product of the square root of the reliability estimates for the two subgroups.

[41] This correlation could potentially be upward biased because the 25th and 75th percentile outcomes in each tract are estimated using a single regression run on the same underlying sample. In practice, using a split-sample approach – estimating the 25th percentile outcomes only using data for below-median income families and the 75th percentile outcomes only using data for above-median income families – yields very similar results.

[42] These correlations are adjusted for attenuation due to noise infusion (which is independent across outcomes) but not sampling error, since the outcomes are estimated on the same underlying sample. As above, using a split-sample

correlations by estimating separate correlations for each of the five racial and ethnic groups listed in Table II and then taking a mean of these five correlations, weighting by each group's national population share in the 2000 Decennial Census. Once again, the correlations are well below 1. For example, the correlation between incarceration rates (for men) and household income in adulthood is only -0.276. The correlation between teen birth rates (for women) and household income is -0.480.

Figure IV explores the relationship between teenage birth rates and upward mobility in greater depth by plotting upward mobility for white women vs. teen birth rates for white women raised in low-income families by tract (restricting to tracts with at least 100 observations). In neighborhoods where teen birth rates are high, upward mobility is almost always low: the mean income rank of white female children with parents at the 25th percentile in tracts in the top decile of teen birth rates is 35.1 and the SD across tracts is only 3.55. But, when teen birth rates are low, there is a wide spectrum of rates of upward mobility, with a mean rank of 49.0 and and an SD of 7.64 in the bottom decile. Hence, from a predictive perspective, low teen birth rates are a necessary but not sufficient condition for having high rates of upward mobility.[43] Because teenage birth outcomes can be measured at earlier ages than incomes, they can be useful for identifying neighborhoods where upward income mobility is likely to be quite low.

In sum, we should not think of neighborhood quality – or the policies that might improve it – as "one size fits all." It may be more effective to design policies that target specific subgroups in ways that directly address the particular challenges they face using the data on outcomes constructed here.

IV.C Correlations with Neighborhood Characteristics

In this subsection, we ask whether the differences in children's outcomes across tracts documented above can be predicted using data on observable characteristics of neighborhoods. Would one be able to do as well by targeting policies based on neighborhood characteristics such as poverty rates, as is common practice at present? More generally, what are the characteristics of places that have high upward mobility? Are they places with lower poverty rates? Greater proximity to jobs? Fewer single-parent households?

We study these questions in Figure V by reporting correlations between upward mobility and

estimator and adjusting for noise due to sampling error yields similar results.

[43] We find qualitatively similar patterns for other outcomes, such as incarceration rates, but this result is most stark for teenage birth.

various neighborhood characteristics that have been studied in prior work.[44] We report race-controlled correlations (computing correlations separately for each race and taking a population-weighted average). To isolate variation across neighborhoods as opposed to the broad geographies studied in earlier work (e.g., Chetty et al. 2014), we include CZ fixed effects and study the correlations across tracts within CZs. The figure plots the magnitude of univariate correlations with various characteristics; green circles represent positive correlations, while red triangles show negative correlations. Details on the construction and definitions of the neighborhood characteristics used in Figure V are given in Online Appendix A.

Jobs. We begin by analyzing the association between upward mobility and local access to jobs. Using data from the publicly available LEHD Origin-Destination Employment Statistics (LODES) dataset, we count the number of jobs within 5 miles of the centroid of a tract. The first row of Figure V shows that this measure is slightly *negatively* associated with upward mobility, with a correlation of -0.175 (s.e. = 0.004). The number of "high-paying" (annual pre-tax wages above $40,000) jobs exhibits a similar pattern.We also find small correlations with the rate of job growth between 2004-2013, the period when children in our sample were entering the labor market. In short, there is little evidence of a positive association between local availability of jobs and upward mobility, challenging spatial mismatch theories of economic opportunity (Kain 1968).

One potential explanation of these results is that job availability does not matter at short distances, but matters at a labor-market level. To evaluate this hypothesis, Figure VI presents a scatter plot of the relationship between upward mobility and job growth for the fifty largest commuting zones, which are geographic areas constructed to capture local labor markets. Even at the CZ level, there is no association between low-income children's earnings in adulthood and job growth rates. For example, Atlanta and Charlotte both experienced very high rates of job and wage growth over the past two decades, yet they have among the lowest rates of upward mobility for children who grow up there. These cities achieve high rates of economic growth because they are magnets for talent – i.e., they attract high-skilled people to move in and fill high-paying jobs. By contrast, Minneapolis has experienced much slower job growth (18%) but has higher rates of upward mobility for children who grow up there. We find similar results when controlling for race by focusing on upward mobility for whites, when looking at metropolitan areas instead of CZs, and

[44] Correlations between the other outcomes analyzed in Table III and observable characteristics are qualitatively similar to those documented below. Correlations with mean income ranks for children with parents at the 75th percentile are also similar (Online Appendix Figure II).

when using measures of wage growth instead of job growth (Online Appendix Figures III-IV).[45]

The upshot of these findings is that a booming labor market does not automatically translate into greater upward mobility for local residents. Hence, policies targeted based on job growth rates would reach quite different areas from the places where upward mobility is lowest. More broadly, the factors that lead to highly productive labor markets with high rates of job, wage, and productivity growth apparently differ from the factors that promote human capital development and result in high levels of upward income mobility across generations.

In stark contrast to the lack of correlation with measures of job availability, we find a strong positive correlation of 0.349 (s.e. 0.004) between the employment rates of the local residents in a neighborhood and the outcomes of children who grow up there. Evidently, what predicts upward mobility is not proximity to jobs, but growing up around people who have jobs. While we of course cannot conclude that this correlation is driven by a causal effect of peers or neighborhood residents, this result echoes Case and Katz's (1991) finding that children's outcomes are correlated with the characteristics of the "company they keep."

Traditional Proxies for Neighborhood Disadvantage. We find similar correlations between children's outcomes and other socioeconomic characteristics that are commonly used as proxies for neighborhood disadvantage, such as the share of residents above the poverty line (correlation = 0.540). Areas with higher mean household incomes, a larger share of college graduates, and higher test scores also all tend to have higher levels of upward mobility (controlling for race) on average. These results show that traditional proxies for neighborhood disadvantage do in fact predict upward mobility, although they do not capture all of the variation.

Family Structure and Social Capital. Consistent with prior work on family structure and children's outcomes (e.g., Sampson 1987), we find a strong negative correlation of -0.589 (s.e. 0.003) between the fraction of single-parent households in a tract and upward mobility. Importantly, this correlation remains similar even conditional on the marital status of a child's own parents. In particular, children of single parents have higher rates of upward mobility if they grow up in a neighborhood with fewer single parent households (correlation = -0.52). This implies that the correlation is driven not by differences in outcomes between children raised by married vs. single parents, but rather by ecological (neighborhood-level) factors.

There is also a large literature documenting the role of social capital in shaping children's

[45] We also find no association between job growth and (1) the mean incomes of children with high-income (75th percentile) parents across CZs or (2) upward mobility for children who stay in the same CZ as adults (who presumably would be most likely to benefit from local growth).

outcomes in adulthood. Measures of social capital are difficult to obtain at the tract level. However, one proxy that is available at the tract level is the fraction of local residents who return their Census forms by mail, an element of the widely used social capital index constructed by Rupasingha and Goetz (2008). We construct this measure for each Census tract using publicly available data from the 2016 Census Planning Database. Across tracts within CZs, we find a strong positive correlation (0.377, s.e. 0.003) with upward mobility of low-income residents, supporting prior work suggesting that social capital is correlated with better outcomes.

Racial and Ethnic Composition. Several studies have shown that racial and ethnic minorities tend to reside in neighborhoods with lower average incomes and higher poverty rates (Massey and Denton 1987; Denton and Massey 1988; Reardon and Townsend 2015). Figure V shows that a similar pattern holds for upward mobility. Black and Hispanic children tend to grow up in neighborhoods with significantly lower rates of race-adjusted upward mobility. Unpacking the data on upward mobility by race, we find that black children tend to live in neighborhoods with *lower* levels of upward income mobility for blacks. The correlation between the share of black residents in a tract and upward mobility for black children in that tract is -0.385. In contrast, white children tend to live in neighborhoods with *higher* levels of upward income mobility for whites: the correlation between the share of white residents and upward mobility for white children is +0.405. These findings imply that the persistence of black-white disparities across generations is amplified by the differences in the neighborhoods where black and white children live.

Granularity of Neighborhoods. Throughout the preceding analysis, we used Census tracts as our definition of "neighborhoods." Is this the right scale of geography to focus on when we analyze neighborhood characteristics? To answer this question, we regress upward mobility from white children in a given tract c on poverty rates both in tract c and its ten nearest neighbors (based on the distance between tract centroids).[46] Figure VIIa plots the coefficients obtained from running this multivariable regression. Both upward mobility and poverty rates are standardized to have mean 0 and SD 1, so the coefficients that are plotted can be interpreted as partial correlations.

The explanatory power of poverty rates decays very rapidly with distance. The coefficient on poverty rates in the child's own tract is -0.31 (s.e. 0.007); the coefficient on the next closest tract (which is on average 1.9 miles away) is -0.02 (s.e. 0.007). Summing the coefficients on the ten nearest neighbors, we can infer that a 1 SD increase in poverty rates in all ten of the nearest

[46] We focus on white children here as a simple method of controlling for race; results are similar when we pool all racial groups. The decay rates documented below are also similar when we examine the other neighborhood characteristics in Figure V instead of poverty rates.

neighboring tracts (roughly a radius of about 5 miles) is associated with a 0.13 SD reduction in upward mobility for white children in a given tract, controlling for poverty rates in that tract. Hence, poverty rates in one's own tract matter 2.4 times more than those in surrounding tracts for upward mobility.

In Figure VIIb, we replicate the analysis in Figure VIIa at the Census *block* level rather than tract level to obtain a more precise picture of how the correlations decay with distance. To construct this figure, we regress the household income rank of children whose parents are between the 20th and 30th percentiles of the income distribution on block-level poverty rates for their own block and the 200 nearest blocks, binned into groups of 5. Since block-level poverty rates are not publicly available, we construct them using information from tax records as the share of families whose total income (defined as wages, social security income, dividends, interest income, and schedule E gains or losses) falls below the poverty line in 2010.[47] We find more gradual and smooth decay across blocks than tracts, which is intuitive insofar as most plausible mechanisms that might underlie these correlations – peer effects, differences in resources, or selection – would be unlikely to operate purely at the own-block level. But the rate of decay is quite rapid with distance: the coefficients remain statistically significant only until about the 40th nearest block, which is only about 0.6 miles away.

In sum, neighborhood characteristics matter at a hyperlocal level. A child's *immediate* surroundings – within about half a mile – are responsible for almost all of the association between children's outcomes and neighborhood characteristics documented above.

Value of Outcome-Based Targeting Relative to Observables. We conclude this section by returning to the question posed at the start: to what extent do our measures of children's outcomes provide new information beyond what can be captured with existing cross-sectional statistics on neighborhoods? Regressing upward mobility on the full set of characteristics in Figure V yields a signal R-squared of 0.50. That is, about half the variation in outcomes is captured in existing measures, while half is not. Looking directly at outcomes is thus quite valuable if one seeks to identify areas with the lowest level of upward mobility.

Perhaps more important, using outcomes directly eliminates the need to determine which predictors provide the most useful proxies for economic opportunity. The R-squared of 0.50 relies on using the optimal (linear) combination of existing proxies to forecast outcomes. Alternative com-

[47] When constructed at the tract level, these poverty rates are very highly correlated with publicly available poverty rates from the Census. In addition, using an estimation approach at the tract level analogous to what we use at the block level yields results very similar to those in Figure VIIa.

binations that form the basis for existing indices for economic opportunity (e.g., Messer et al. 2006 Acevedo-Garcia et al. 2014, Opportunity-Nation 2017) have correlations with our outcome-based measures that range from 0 to 0.6, depending upon the exact set of variables and weights that are used.

The lack of reliance on a specific model relating neighborhood characteristics to outcome is especially valuable because the relationship between neighborhood characteristics and children's outcomes can itself be highly heterogeneous across areas. To illustrate this point, consider the following question: Do cities have higher rates of upward mobility? On average, we find a small negative correlation between upward mobility and population density across tracts within CZs (correlation = -0.233, s.e. 0.004), as shown in the last row of Figure V.[48] However, this aggregate relationship masks considerable heterogeneity across the country. Figure VIIIa maps upward mobility by tract for white children in North Carolina. This map is predominantly red, with blips of blue around urban centers, showing that cities have higher rates of upward mobility than rural areas in North Carolina (correlation with population density = 0.22, s.e. 0.027). In contrast, Figure VIIIb shows that the pattern is exactly reversed in Iowa: rural areas have higher rates of upward mobility for whites than urban areas (correlation = -0.59, s.e. 0.037). Figure VIIIc plots the correlation between upward mobility for whites and population density for each of the fifty states. In the southeast, rural areas tend to have lower rates of upward mobility than urban areas. In contrast, in the midwest and mountain west, rural areas tend to have significantly higher rates of upward mobility than urban areas.

The sharp heterogeneity in the association between population density and upward mobility illustrates the power of using data on outcomes.[49] Rather than attempting to draw general conclusions about whether cities offer better opportunities for upward mobility based on observable characteristics, one can directly study outcomes in the area of interest and target policies accordingly.

IV.D Changes Over Time

Although the preceding analysis suggests that the data on outcomes add information relative to observable neighborhood characteristics, a drawback of these measures is that they come with a lag because one must wait until children grow up to see their earnings. Are these estimates useful

[48] We find a correlation of -0.236 (s.e. 0.005) for upward mobility for whites.

[49] We find spatial heterogeneity in the association between upward mobility and the other characteristics analyzed in Figure V as well, but the differences are most stark for population density.

for targeting policies to improve outcomes for children growing up today?

Since rates of upward mobility for children growing up today are fundamentally unobservable, it is useful to conceptualize this problem as a prediction exercise. How well can we predict upward mobility for birth cohort t ($\bar{y}_{c,t}$) in tract c using historical estimates of upward mobility for birth cohort $t-j$ ($\hat{y}_{c,t-j}$)? Focusing on linear predictors,

$$\bar{y}_{c,t} = \alpha + \beta_j \bar{y}_{c,t-j} + \varepsilon_{cjt}, \tag{4}$$

the optimal weight placed on the historical estimate for tract c is given by the standard formula for the regression coefficient:

$$\beta_j = \frac{Cov\left(\bar{y}_{c,t}, \hat{y}_{c,t-j}\right)}{Var\left(\hat{y}_{c,t-j}\right)}. \tag{5}$$

If upward mobility in each tract did not change over time ($\bar{y}_{c,t} = \bar{y}_{c,t-j}$), then the optimal weight β_j would be determined purely by the reliability of the estimate $\hat{y}_{c,t-j}$. As noted in Table II, the reliability of our baseline tract-level estimates of upward mobility is 0.91, implying that the optimal weight on historical estimates in forecasting future cohorts' outcomes would be 91% under the assumption of perfect stationarity in upward mobility across cohorts. Here, we are interested in understanding how this weight of 91% decays over time once we allow for the fact that upward mobility changes across cohorts.

We analyze how the coefficients β_j decay over time by estimating upward mobility $\hat{y}_{c,t-j}$ separately by single birth cohort (pooling racial groups) and running the regression in (4). To identify as many lags as possible, we extend our primary analysis sample to children born in the 1978-89 birth cohorts and measure children's income ranks at age 26, the earliest at age at which we can obtain reliable estimates of permanent income ranks (Chetty et al. 2018b). To maximize precision, we use all available cohorts, both leads and lags, to estimate each coefficient; for example, the coefficient at a lag of $j = 1$ is estimated using an average of 11 x 2 = 22 estimates of β_1.

Figure IXa plots β_j/β_1 – the optimal weight with estimates that have a j year lag relative to the optimal weight placed on an estimate that is one year old. There is very gradual decay across cohorts. The optimal weight placed on an outcome observed 10 years apart is only 11% smaller than the weight placed on an outcome observed in the previous year. Hence, over the span of years we are able to study in our data, it appears that upward mobility is sufficiently stable that historical predictions are likely to remain a useful (albeit imperfect) prediction for more recent cohorts.

Ideally, we would extend the series in Figure IXa beyond $j = 11$ to determine whether our baseline estimates for the 1978-83 birth cohorts are relevant for children growing up today. As a

feasible alternative with currently available data, we study how tract-level neighborhood characteristics that are predictive of upward mobility evolve over time. Figure IXb replicates Figure IXa using poverty rates by tract instead of upward mobility. We measure poverty rates using publicly available data from the 1990 and 2000 Decennial Census and ACS data collected between 2006 and 2010 and between 2011 and 2015, which we pool to obtain estimates for 2008 and 2013, respectively. We obtain estimates at $j \in \{5, 8, 10, 13, 18, 23\}$ by combining these datasets. Similar to Figure IXa, we observe relatively little decay: poverty rates 23 years ago are 91% as predictive of current poverty rates in a neighborhood as poverty rates five years ago. Since half of the variance in upward mobility is predicted by such neighborhood characteristics, this finding suggests that upward mobility is itself likely to be fairly stable over longer horizons – presumably because the underlying structural factors that lead to differences in upward mobility across areas themselves change slowly over time.

The preceding results show that historical estimates of upward mobility predict the outcomes of more recent cohorts, but they do not establish that these measures are the best predictors. In particular, current neighborhood characteristics such as poverty rates could potentially be better predictors of upward mobility for current cohorts than historical data on upward mobility. To investigate the relative explanatory power of historical measures of upward mobility vs. neighborhood characteristics, we regress our estimates of upward mobility for the 1989 birth cohort on upward mobility for the 1979 cohort and various observable characteristics. We find that the explanatory power of the historical upward mobility estimates is significantly greater than characteristics such as poverty rates, and that the incremental R-squared from adding more recent measures of poverty rates is small (Online Appendix Table II). In short, given a choice between observable characteristics and ten-year-lagged measures of upward mobility, one would obtain more reliable predictions of outcomes for a given cohort using the upward mobility estimates.[50]

In summary, upward mobility is sufficiently stable over time that estimates of upward mobility for children born in the 1980s are still predictive of potential outcomes for children growing up today. This is not to say that upward mobility does not change over time; it certainly does. However, the historical measures of upward mobility constructed here appear to be the best available predictors of upward mobility for children today.

[50] Of course, this applies only to the set of observable characteristics we study here; it would be very useful to identify measures that are more predictive of changes in upward mobility in future work.

IV.E Summary and Illustrative Applications

In this section, we have shown that our measures of upward mobility (1) capture wide variation in children's outcomes conditional on parental income and demographic characteristics, (2) provide new information beyond existing measures of neighborhood characteristics such as poverty rates, and (3) are informative predictors even for children growing up today. These results suggest that the new estimates in the Opportunity Atlas can potentially be useful as tags for policies that are targeted towards children with the lowest rates of upward mobility.

We conclude this section by revisiting the motivating policy examples discussed at the beginning of this section and showing how tagging based on the new data can lead to meaningful changes in the neighborhoods that the policies target. Consider the Opportunity Zones program, which provides preferential tax treatment for investment in selected neighborhoods. Online Appendix Figure Va outlines the tracts that were designated as Opportunity Zones in Los Angeles county. Online Appendix Figure Vb shows the neighborhoods that would hypothetically be selected if one were to choose the same number of tracts, selecting those with the lowest levels of upward mobility. The neighborhoods change quite substantially, with more neighborhoods in the center of the city selected by targeting the lowest-mobility tracts. Children who grew up in low-income (25th percentile) families in neighborhoods that are currently designated as Opportunity Zones in Los Angeles county earn about $31,000 on average in adulthood; under the hypothetical designation, this figure would be $26,000. This suggests that more disadvantaged neighborhoods could be targeted using the upward mobility estimates.

As a second example, consider admissions to Chicago's selective public high schools. Online Appendix Figure VIa outlines the 2017 "Tier 1" tracts that were designated to give preferential admission status in the Chicago Public School tier-based admission system to children living in the most underserved areas. Appendix Figure VIb shows the neighborhoods that would hypothetically be selected based on upward-mobility targeting. Again, we see a significant shift, in particular with more tracts on the far South Side of the city – areas that do not have particularly high poverty rates (and hence are not included under the current designation) yet have low observed rates of upward mobility.

Of course, these results should not be interpreted as a normative prescription that one *should* target different areas in these programs. Policies may be targeted with many potential objectives in mind, of which upward mobility might be one consideration. Our point here is simply that if

one wanted to target areas with low levels of upward mobility, the new data constructed here could meaningfully change one's choices.

V Causal Effects and Neighborhood Choice

The neighborhood-level variation in outcomes documented above could be driven by two very different sources. One possibility is that neighborhoods have causal effects on children's outcomes: that is, moving a given child to a different neighborhood would change his or her outcomes. Another possibility is that the variation is due to differences in the types of people living in each neighborhood. Distinguishing between these two explanations is important to understand whether the observational estimates in the Opportunity Atlas are useful for families choosing where to live and for the design of policies such as housing vouchers that seek to help families move to opportunity.

In this section, we analyze the extent to which our observational estimates of upward mobility reflect causal effects of place vs. selection. To define the estimand of interest, let $\bar{y}_{cp}$ denote the expected income rank in adulthood for children raised from birth in tract c at parental income rank p. Consider a hypothetical experiment in which a new group of children are randomly assigned to grow up in different neighborhoods at birth. Letting y_i denote child i's income rank in adulthood, our goal is to identify the coefficient λ in a regression of outcomes in the experimental sample on the observational predictions:

$$y_i = \alpha + \lambda \bar{y}_{c(i),p(i)} + \eta_i. \tag{6}$$

Since children's potential outcomes are orthogonal to $\bar{y}_{c(i),p(i)}$ in the experimental sample, the parameter λ represents the average causal effect of growing up in a neighborhood where observed ranks are 1 percentile higher. Furthermore, under the assumption that the causal and selection components of $\bar{y}_{cp}$ are additive and uncorrelated, λ can be interpreted as the fraction of the variance in $\bar{y}_{cp}$ that is due to the causal effects of place.[51]

We estimate λ using two research designs. First, we compare our observational predictions with the experimental treatment effects for children who moved to different neighborhoods in the

[51] To see this formally, suppose that child i's outcome $y_i = \mu_{c(i),p(i)} + \theta_i$, where μ_{cp} denotes the causal effect of growing up in tract c given parents at percentile p and θ_i denotes a selection term that reflects family inputs or other factors unrelated to where a child grows up. In the observational sample, the mean selection effect $\bar{\theta}_{cp} = E[\theta_i|c,p]$ can vary across tracts; in the experimental sample, $\bar{\theta}_{cp}$ does not vary across tracts by construction. If $Cov(\bar{\theta}_{cp}, \mu_{cp}) = 0$ in the observational sample – i.e., if the selection and causal components are uncorrelated across tracts – then $\lambda = Cov(y_i, \bar{y}_{cp})/Var(\bar{y}_{cp}) = Var(\mu_{cp})/Var(\bar{y}_{cp})$, the fraction of the variation in $\bar{y}_{cp}$ that is due to causal effects. If $\bar{\theta}_{cp}$ is correlated with μ_{cp}, λ includes this additional covariance and cannot be directly interpreted as the variance share of causal effects. However, even in this case we can still test the null hypothesis that there are no causal effects by estimating λ, since $\lambda > 0 \Leftrightarrow Var(\mu_{cp}) > 0$.

Moving to Opportunity Experiment. Second, we use the quasi-experimental design of Chetty and Hendren (2018a) to estimate causal effects by comparing the outcomes of children who move across tracts at different ages. Both of these approaches reveal that most of the observational variation in upward mobility is due to causal effects. In light of these results, in the final subsection, we use our estimates to study how opportunity for children is priced in the housing market and discuss implications for the design of affordable housing policies.

V.A Comparison to Estimates from the Moving to Opportunity Experiment

Our first method of estimating λ is to compare our observational predictions to estimates from the Moving to Opportunity (MTO) experiment, which closely approximates the ideal experiment described above.

We begin by briefly summarizing the design of the MTO experiment; see Sanbonmatsu et al. (2011) and Chetty et al. (2016) (hereafter CHK) for a more comprehensive description. The MTO experiment enrolled 4,604 low-income families living in high-poverty public housing projects in five U.S. cities – Baltimore, Boston, Chicago, Los Angeles, and New York – from 1994 to 1998. These families were randomized into three groups: 1) the experimental group, which received housing vouchers that subsidized private-market rents and could only be used in census tracts with 1990 poverty rates below 10%; 2) the Section 8 group, which received regular housing vouchers without any constraints; and 3) a control group, which received no assistance through MTO but retained the option to stay in public housing. Families in all three groups were required to contribute 30% of their annual household income toward rent and utilities. Families remained eligible for vouchers (or public housing) indefinitely as long their income was below 50 percent of the median income in their metro area.

CHK estimate the treatment effects of the MTO experiment by linking the MTO program data to federal income tax records analogous to those we use here. They find that children who were assigned at young ages (below age 13) to the Experimental and Section 8 groups earned significantly more in adulthood than their peers in the control group.[52] Here, we use Chetty et al.'s experimental estimates for children who were below age 13 at the point of random assignment to estimate λ, by regressing the MTO experimental estimates on the observational predictions:

$$\hat{y}_{ws}^{MTO} = \alpha_s + \lambda_{MTO}\hat{y}_{ws} + \varepsilon_{ws}. \tag{7}$$

[52] Treatment effects for children who moved at older ages (above age 13) and adults were not significant, suggesting that the duration of exposure to a better environment during childhood is a key determinant of an individual's long-term outcomes. Here, we focus exclusively on the estimates for young children.

In this specification, $\hat{y}_{ws}^{MTO}$ denotes mean individual earnings for children below age 13 at random assignment in site s and treatment arm w in the MTO experiment, while $\hat{y}_{ws}$ denotes the mean observed level of individual earnings from the Opportunity Atlas in the tracts where children in site s and and treatment arm w lived. We include site fixed effects α_s in (7) because children were randomly assigned to different groups only within sites. We identify λ_{MTO} only from the average differences in neighborhood quality between children in different treatment groups in (7) rather than exploiting variation across neighborhoods within treatment arms because the specific neighborhoods that families chose within these arms might be correlated with family-specific unobservables. Under the assumption that the variation in $\hat{y}_{ws}^{MTO}$ is driven purely by differences in the neighborhoods in which children grew up and not other differences in services that children in the treatment arms may have received, λ_{MTO} directly provides an unbiased estimate of λ – the causal effect of moving to a neighborhood with higher observed rates of upward mobility.

To implement (7), we construct the estimates of $\hat{y}_{ws}^{MTO}$ from the intent-to-treat (ITT) estimates on individual earnings reported for each site by CHK (Appendix Table 7, Panel B). To adjust for the fact that not all families who were offered vouchers took them up and moved, we follow CHK and construct treatment-on-the-treated (TOT) estimates for the Section 8 and Experimental groups as the mean observed earnings for the control group in the relevant site plus the site-specific ITT estimate for each treatment arm divided by the voucher takeup rate in that arm.[53]

We construct the corresponding observational predictions $\hat{y}_{ws}$ in three steps. First, we construct observational predictions of mean individual income ranks for children with parents at the 10th percentile, roughly the median income level of MTO participants (Sanbonmatsu et al. 2011), following the methodology in Section III.[54] Second, we translate the percentile rank predictions to dollar values at age 26 – the median age at which CHK measure young children's incomes in adulthood in the MTO analysis – using a crosswalk from ranks to dollars at age 26. Finally, we take the neighborhoods reported as the most common locations for children in MTO from Appendix Table 1c of CHK and map these neighborhoods to Census tracts. We then calculate the average predicted individual income across these tracts, weighting by the number of children from below-median income families in each tract in the 2000 Census, to arrive at $\hat{y}_{ws}$.

Figure X presents a scatter plot of $\hat{y}_{ws}^{MTO}$ vs. $\hat{y}_{ws}$. There are 15 points, representing each of

[53] These TOT estimates rely on the assumption that being offered an MTO voucher had no effect on those who did not take it up, a standard assumption in experimental studies with non-compliance.

[54] We do not use race- and gender-specific predictions because CHK do not report site-specific treatment effects by race and gender, due to small sample sizes.

the three treatment arms in the five sites. Solid circles represent the control group, while hollow triangles and solid diamonds represent the Section 8 voucher and Experimental voucher groups, respectively. To eliminate variation across sites, we demean both $\hat{y}_{ws}^{MTO}$ and $\hat{y}_{ws}$ within site and add back the values of $\hat{y}_{ws}^{MTO}$ and $\hat{y}_{ws}$ for the control group in Chicago to facilitate interpretation of the scale.

There is a clear positive relationship between the actual outcomes of children in the MTO experiment and our observational predictions. The correlation coefficient is 0.60. The slope of the regression line is $\hat{\lambda}_{MTO} = 0.71$ (s.e. 0.26): moving to an area where children in low-income (10 percentile) families earn \$1,000 more in the observational data increases children's earnings by \$710.

The comparison to the MTO estimates suggests that around 70% of the variation in observational estimates of upward mobility across rates is due to causal effects rather than sorting.[55] However, this conclusion rests on estimates for a small set of neighborhoods in five specific cities. To estimate the causal component of our estimates more broadly, we now turn to a second approach that uses data from all tracts in the U.S.

V.B Quasi-Experimental Estimates of Causal Effects

Our second identification strategy builds on the quasi-experimental research design of Chetty and Hendren (2018a), who study how the outcomes of children who move across areas vary with the age at which they move. Chetty and Hendren (2018a) use this timing-of-move design to establish that neighborhoods have causal effects on children's outcomes at the commuting zone level; here, we use the same design to identify the causal effects of tracts. We extend our primary sample to the 1978-91 birth cohorts for this analysis in order to maximize the set of ages at which we see children moving. Since our approach closely mirrors that of Chetty and Hendren (2018a), we briefly summarize the estimation approach in what follows; further details regarding the sample specification, variable definitions, and estimation equations are in Online Appendix B.

Empirical Specification. We study the outcomes of children who move across tracts exactly

[55] This estimate should be interpreted as a rough approximation of λ for several reasons. First, we use data on the most common initial neighborhoods to which MTO participants moved (from Appendix Table Ic of CHK) to estimate $\hat{y}_{ws}$ rather than the exact locations where MTO children grew up throughout their childhood. Second, children moved on average at age 8 in the MTO sample of young children analyzed by CHK; children who move at birth might experience larger treatment effects. Third, because our estimates of $\hat{y}_{ws}$ themselves contain estimation error, $\hat{\lambda}_{MTO}$ slightly understates the fraction of the variance in the conditional expectation $\bar{y}_{cp}$ due to the causal effects of place. Given the reliability estimate of about 0.9 in Table II, the third source of bias leads us to understate λ by about 10%. Quantifying the magnitude and sign of potential bias from the first two issues is more challenging, but we expect that these sources of error are likely to be smaller than the uncertainty in $\hat{\lambda}_{MTO}$ due to sampling error.

once before age 30 during our sample window. Let i index children and p_i denote their parental income ranks. In the sample of one-time movers, let m_i denote the age at which child i moves from origin tract o to destination tract d. Let $\hat{y}_{pc}$ denote the exposure-weighted mean of y_i for children who grew up in location c with parental household income rank p, estimated using the approach in Section III except using only children who never move or who move more than one time (i.e., the complement to one-time movers).[56] Let $\Delta_{odp} = \hat{y}_{pd} - \hat{y}_{po}$ denote the predicted difference in income ranks in the destination versus origin tract for children.

After computing these variables, we regress the income rank of children who move (y_i) on the measures of origin and destination quality and parental income rank, fully interacted with age-at-move fixed effects:

$$y_i = \sum_{m=2}^{28} I(m_i = m) \left[\alpha_m + \phi_m \hat{y}_{po} + \zeta_m p_i + b_m \Delta_{odp}\right] + \varepsilon_i \tag{8}$$

where α_m denotes an age-at-move-specific intercept and the parameters $\{\phi_m, \zeta_m\}$ are age-specific coefficients on the average exposure-weighted outcome in the origin and parental income rank, respectively. To adjust for attenuation bias due to measurement error in our estimates of $\hat{y}_{pc}$ at the tract level, we estimate this specification using a split-sample instrumental variables approach. We construct independent estimates $\hat{y}_{pc,1}$ and $\hat{y}_{pc,2}$ by randomly splitting families into two samples and then instrument for $\hat{y}_{po,1}$ with $\hat{y}_{po,2}$ and $\Delta_{odp,1}$ with $\Delta_{odp,2}$ when estimating (8).[57] Throughout our analysis, we restrict to moves between tracts that are at least 25 miles apart; we show in Online Appendix B that shorter-distance moves often reflect measurement error in location or returns to areas where a child previously lived, especially in single-parent families, creating attenuation bias in the estimates.

The key parameters of interest in (8) are the b_m coefficients, which capture how children's outcomes vary with the age at which they move to an area with higher or lower predicted earnings in the observational data. We can interpret differences in the coefficients b_m, e.g. $b_m - b_{m+1}$, as the causal effect of exposure to a better area (i.e., an area with higher observed incomes) under the assumption that the potential outcomes of children who move to better vs. worse areas do not vary with the age at which they move. We first present a set of baseline estimates and then present

[56] We do not include one-time movers when constructing these exposure-weighted outcomes to ensure that a child's own outcome does not enter our definition of neighborhood quality.

[57] The degree of noise relative to signal is amplified in (8) because we identify purely from residual variation in Δ_{odp}, controlling for origin neighborhood quality. This is why adjusting for noise is particularly important in this analysis.

a series of tests to validate this orthogonality condition.

Baseline Results. Figure XI plots the coefficients $\{b_m\}$ for the specification in equation 8 using household income ranks at age 24 as the outcome.[58] Consistent with the results in Chetty and Hendren (2018a) at the CZ level, the coefficients b_m decline steadily until approximately age 23, after which they are flat at an average level of $\delta = 0.35$. Since moves after age 24 cannot affect income measured at age 24, this δ coefficient reflects selection: children whose parents move to areas with better observed outcomes tend be positively selected in terms of their potential earnings.[59]

Under our identification assumption that the selection effect δ does not vary with the age at which children move (illustrated by the dashed horizontal line in Figure XI), we can interpret the difference between b_m and δ as the causal effect of moving to an area with one percentile higher predicted income ranks at age m. The declining pattern of the coefficients implies that neighborhoods have causal effects on children's outcomes in proportion to childhood exposure prior to age 23. The slope of this relationship is somewhat steeper between ages 13 and 23 than it is at earlier ages, suggesting that where a child lives as an adolescent may be particularly influential in determining his earnings outcomes, consistent with the findings of Deutscher (2018) in Australian data.

Fitting a linear regression to the estimates below age 23, we estimate an annual childhood exposure effect – the effect of growing up for an additional year in a tract with 1 percentile higher observed earnings – of $\gamma = 0.025$ (s.e. 0.002). Extrapolating linearly back to age 0, we obtain a predicted value of $b_0 = 0.92$, implying that the causal effect of moving at birth to an area with 1 percentile higher observed earnings is $\lambda = 0.92 - 0.35 = 0.57$. Hence, under the identification assumption underlying this research design, we can infer that 57% of the variation in the observational estimates is due to the causal effect of neighborhoods.

In the rest of this subsection, we evaluate the robustness of this estimate and the identification assumptions underlying its interpretation as a causal effect using variants of the specification above. For simplicity, we parameterize the age-specific effects plotted in Figure XI using a two-piece linear spline, permitting different slopes above and below age 23, by estimating regressions of the following

[58] We measure income at age 24 here, as in Chetty and Hendren (2018a), in order to maximize the range of ages at move that we are able to analyze. We verify that measuring income at later ages yields similar results, as shown in Chetty and Hendren (2018a).

[59] This selection term is higher than the corresponding estimate of $\delta = 0.22$ the CZ level (Chetty and Hendren 2018, Figure IV), suggesting that there is more unobservable selection across tracts than across CZs, as one might expect intuitively.

form:

$$y_i = \sum_{m=2}^{28} I(m_i = m)\left[\alpha_m + \phi_m \bar{y}_{po} + \zeta_m p_i\right] + I(m_i \leq 23)(\gamma' + \gamma m_i)\Delta_{odp} + I(m_i > 23)(\rho' + \rho m_i)\Delta_{odp} + \varepsilon_i \quad (9)$$

Here, the coefficient of interest is γ, the annual childhood exposure effect, which is the average effect of moving to a tract with 1 percentile higher observed income ranks one year earlier, at or before age 23. The coefficient ρ measures the corresponding slope for moves after age 23.

Column 1 of Table IV presents estimates of γ and ρ using (9). Consistent with the non-parametric estimate in Figure XI, we obtain an estimate of $\gamma = 0.027$, implying that $23 \times 0.027 = 62\%$ of the variation in observational estimates can be attributed to the causal effect of neighborhoods. Column 2 presents estimates from a more parsimonious specification that eliminates the interactions between age-of-move dummies and origin place quality $\bar{y}_{po}$, effectively constraining the coefficient on $\bar{y}_{po}$ to be constant across ages. The coefficients are again very similar. Columns 3 and 4 replicate Column 1, replacing household income with an indicator for being married at age 30 (Column 3) and for being incarcerated on April 1, 2010 (Column 4).[60] We obtain similar estimates of childhood exposure effects – with a convergence rate of approximately 2.5% per year of exposure to the mean outcomes observed in the destination – for these outcomes as well.

Validation of Identification Assumption. The preceding analysis rests on the assumption that the potential outcomes of children who move to better or worse neighborhoods do not systematically vary with age of move. We evaluate the validity of this assumption using two approaches, following Chetty and Hendren (2018a). First, in Column 5 of Table IV, we add family fixed effects to the specification in Column 1. This approach identifies exposure effects from comparisons between siblings, by asking whether the *difference* in earnings outcomes between two siblings who move to a new area is proportional to their age difference interacted with observed outcomes in the destination. The coefficients in Column 5 are very similar to those in Column 1.

These sibling comparisons address confounds due to factors that are fixed within families, but they do not account for *time-varying* factors, such as a change in family environment at the time of the move (e.g., a new job) that directly affects children in proportion to exposure time independent of neighborhoods. To evaluate whether such unobservables might bias our findings, we turn to a second test of our identification assumption: outcome-based placebo tests. These tests are motivated

[60] Note that we replace not only the left-hand side variable but also the neighborhood-specific predictions with these alternative outcome measures.

by the finding in Section IV.B that neighborhoods are multi-dimensional: incomes, incarceration rates, teenage birth rates, and other outcomes are not perfectly correlated across tracts. Under the causal exposure effect model, the incomes of children who move between neighborhoods should be predicted by the difference in observed incomes rather than by differences in other outcomes, such as incarceration rates. In contrast, it is less plausible that time-varying unobservables (such as a new job) would happen to perfectly replicate the entire distribution of outcomes in each area in proportion to exposure time.

We implement these placebo tests in Table V, where we start from the parsimonious specification in Column 2 of Table IV and include not only $\hat{y}_{op}$ and Δ_{odp} as regressors, but also analogous tract-level predictions of marriage rates at age 30 and incarceration rates in 2010.[61] In Column 1, we put children's income ranks at age 24 on the left-hand side of the regression, as in Table IV. We continue to find a childhood exposure effect of $\gamma = 0.026$ for the prediction based on the change in children's mean income ranks. However, moving to an area with higher incarceration rates or marriage rates has no impact on children's incomes, controlling for observed incomes in the destination. Columns 2 and 3 repeat this exercise with marriage and incarceration (respectively) on the left-hand side; as in Column 1, each outcome is only predicted by the neighborhood quality measure based on that outcome itself, while the other predictors are not significantly related. This evidence strongly supports the view that the variation in children's outcomes across neighborhoods for movers is driven by causal effects. Intuitively, it is unlikely that a correlated shock - such as an increase in wages when the family moves - would covary precisely with differences in neighborhood quality between these different outcomes.[62]

Heterogeneity Analysis. Having established the validity of the research design and the robustness of our estimates in the full sample, we now study whether the fraction of the observational variation that is due to causal effects (λ) varies across subsamples of the data. As a reference, we begin by replicating our baseline specification from equation (9) in Column 1 of Table VI.

One concern with these estimates is that they are identified from the set of families who choose to move to a given area, and hence may capture causal effects that apply only to the particular families who chose to move to a neighborhood that is good for their children, rather than a broader population. For example, our estimates of λ might not have external validity to other families who

[61] We use the more parsimonious specification to eliminate the multiple sets of interactions, which reduce power; results are qualitatively similar though noisier with using specification in Column 1 of Table IV.

[62] Formally, this test relies on the assumption that if unobservables θ_i are correlated with exposure to neighborhood quality as measured by a specific outcome variable y, they must also be correlated with neighborhood quality as measured by another outcome variable y' (conditional on control variables).

may be induced to move by other factors, such as changes in housing policies. In the presence of such selection on heterogeneous treatment effects, we would expect moves to better areas ($\Delta_{odp} > 0$) to produce larger exposure effects in absolute value than moves to worse areas ($\Delta_{odp} < 0$). Intuitively, in this scenario, children moving to better neighborhoods would gain even more than we predict based on average earnings, while those moving to worse neighborhoods would suffer less relative to the observational predictions. We test whether this is the case in Column 2 of Table VI by replicating the specification in Column 1, allowing for separate exposure effects for moves to better places (where $\Delta_{odp} > 0$) and moves to worse places. We find completely symmetric effects of moves to better vs. worse places, suggesting that there is relatively little heterogeneity in treatment effects, at least based on observed patterns of selection.

Column 3 replicates Column 1 for the subset of families who make large moves, defined as moves either from the top decile of neighborhoods in terms of $\hat{y}_{c,p=25}$ to the bottom decile or vice versa. Once again, the estimates are very similar, suggesting that roughly 60% of the differences in observed outcomes reflect causal effects even in the tails of the distribution. Columns 4 and 5 of Table VI test whether the component of children's incomes that is predictable by median rents and the observable characteristics analyzed in Figure V above have the same causal content as the "unobservable" component of the variation in observed outcomes. In Column 4, we use the predicted values of $\hat{y}_{op}$ and $\hat{y}_{dp}$ based on observables to construct the key right-hand-side regressors in (9); in Column 5, we conversely use the residuals from the regressions on observables. We find that the estimated exposure effects are nearly identical between these two components, allaying the concern that the "unobserved" portion of the variation in children's outcomes across areas predominantly reflects selection.

Granularity of Causal Effects. Finally, we use the movers design to revisit the analysis on the geographic scale of neighborhoods in Section IV.C, examining whether neighborhoods' *causal* effects also operate at a fine geographic scale. Figure XII plots estimates from a regression that replaces $\hat{y}_{op}$ and $\hat{y}_{dp}$ in the parsimonious specification used in Column 2 of Table IV with poverty rates in the origin and destination tracts along with symmetric interactions between age at move and poverty rates in the ten tracts that are closest to the actual origin and destination tracts, respectively. We plot the eleven coefficients on the interactions between the destination-origin difference in poverty rates and age at move (for moves below age 23). These coefficients can be interpreted as the causal childhood exposure effect of moving to a tract (or near a tract) that has 1 SD higher poverty rates. Moving to a higher poverty tract earlier in childhood significantly reduces

a child's earnings. However, moving to an area where *surrounding* tracts have higher poverty rates (controlling for poverty rates in one's own tract) has essentially no impact on children's outcomes. This figure replicates the correlational finding in Figure V, showing that neighborhood characteristics are predictive not just of sorting patterns at a hyperlocal level but of causal effects.[63]

We conclude based on this quasi-experimental analysis that $\lambda = 62\%$ of the observational variation in outcomes across Census tracts reflects the causal effects of neighborhoods rather than selection, and that this fraction is quite stable across subsamples. Combining this estimate of λ with the signal SD of individual income across tracts within counties for children with parents at the 25th percentile ($2,737), we estimate that moving at birth from a neighborhood at the 25th percentile of the distribution of upward mobility within one's county to a neighborhood at the 75th percentile would increase the lifetime undiscounted earnings of a child growing up in a low-income family by $206,261 (or $53,136 in present value at birth).[64] This earnings gain would benefit not just the children who move but taxpayers as well, through increased tax revenue and lower incarceration rates. These gains could be substantial; for example, assuming a tax rate of 20%, tax revenue would increase by $41,252 over the course of a child's lifetime.

V.C Price of Opportunity

The finding that a large fraction of the variation in the observational estimates of upward mobility is driven by causal effects of place raises the possibility that these estimates could be used by families to move to higher opportunity areas. The feasibility of such moves, however, relies on being able to find affordable housing in such areas. In this subsection, we study the relationship between our observational estimates of upward mobility and rents to understand how opportunity is priced in spatial equilibrium and whether low-income families are likely to be able to "move to opportunity."

Figure XIIIa plots upward mobility vs. the median rent for two-bedroom apartments in 1990 (measured in 2015 $) for tracts in the city of Chicago. On the one hand, upward mobility is clearly correlated with rent, showing that this neighborhood amenity is priced in equilibrium on

[63] We find a similar pattern when using children's mean observed outcomes or other covariates as regressors instead of poverty rates: in all cases, what matters are observed outcomes and characteristics in one's own tract, not nearby tracts.

[64] We estimate this impact by first tabulating mean individual earnings by age in the publicly available 2015 ACS across all ages. We then apply a 1% wage growth rate to obtain an undiscounted sum of lifetime earnings for the average American of $2.81 million. Assuming that the gain in earnings from growing up in a better neighborhood remains constant in percentage terms over the lifecycle, the estimated impact on undiscounted lifetime earnings is $0.62 \times 0.12 \times \$2.81\text{m} = \$206{,}000$. Assuming a 0.5% wage growth rate would give an undiscounted lifetime earnings impact of $162,253, while a 1.5% rate would give $263,755. To obtain the present values at birth, we apply a 3% annual discount rate to this profile to obtain a PDV at age 0 of $723,340 and repeat the preceding calculations.

average. On the other hand, there is considerable residual variation in upward mobility conditional on median rent. For example, both Hyde Park and Alsip had median rents of about $1,000 per month in 1990. Yet children who grow up in low-income families in Hyde Park reach only the 24th percentile of the income distribution on average, while those in Alsip instead reach the 47th percentile, roughly $24,000 higher in the income distribution. More broadly, across the United States, the within-CZ signal correlation between rent and upward mobility in 0.44. The residual SD of upward mobility controlling for median rent is 4.63 percentiles ($4,977). In general, there appears to be considerable scope to find higher-mobility areas without paying a higher level of rent.

Heterogeneity in the Price of Opportunity. In some CZs, rent covaries much more strongly with upward mobility – i.e., opportunity is priced more steeply. To demonstrate this, we quantify the average price of upward mobility in each CZ by first regressing median annual rents on our estimates of upward mobility across tracts within each CZ, weighting by number of children from below-median income families in the tract. We then inflate this regression coefficient by the reliability of our upward mobility estimates (0.91) to adjust for noise. Finally, we map the estimates in ranks to dollars as above to obtain an estimate that can be interpreted as the average annual rental cost of a $1 increase in future annual income for children with parents at the 25th percentile.

Across all CZs, the average population-weighted "price of opportunity" is 0.185 – that is, a $1,000 increase in future annual income for children costs an additional $180 in annual rent for each year they are growing up. However, the standard deviation of these estimates across CZs is 0.114, showing that there is considerable dispersion in the price of opportunity. What kinds of CZs have a steeper price of opportunity? One key predictor is the extent of land-use regulations, which have been widely studied in prior work on the cost of housing. Figure XIIIb plots the price of predicted outcomes against the Wharton Residential Land-Use Regulatory Index (Gyourko et al. 2008). The price of opportunity is substantially higher in cities with tighter land use regulations variables, with a correlation of 0.54. In lightly regulated cities, such as Wichita, KS, the price of predicted outcomes is just 0.045 dollars in annual rent for a $1 increase in future annual income; in highly regulated cities, such as Boston, MA or Baltimore, MD, the price is five times higher, at approximately 0.25.

Why is Upward Mobility Not Fully Priced? It is clear that in many CZs – especially those without substantial land use regulations – the price of opportunity is not very steep on average, and moreover there is considerable residual variation in upward mobility at a given level of rent. Why is the variation in upward mobility not fully priced in the housing market in spatial equilibrium?

There are at least three potential explanations for this finding, as discussed in Chetty and Hendren (2018b).

One possibility is that high-mobility, low-rent tracts have other disamenities, such as longer commute times. Controlling for average commute times (in addition to median rent) reduces the residual signal variation in upward mobility very little, from 4.63 to 4.60 percentiles (Figure XIIIc). Of course, it is still possible that other unobservable neighborhood characteristics explain the remaining variation.

An alternative possibility is that frictions in the housing market may prevent households – especially low-income households – from moving to high-upward-mobility neighborhoods. For instance, households may face discrimination or may move under duress in a way that limits their available options (DeLuca et al. 2017; Christensen and Timmins 2018).

A third possibility is that households may simply lack information on which neighborhoods produce better outcomes for children. As a simple way to evaluate this final hypothesis, we split the variation in upward mobility into the component that is predicted by the observable neighborhood characteristics analyzed in Figure V (such as poverty rates and test scores) and a residual ("unobservable") component. We then estimate correlations between rents and these two components of upward mobility. Figure XIIId shows that the observable component has a correlation of 0.47 with rent, while the unobservable component has a signal correlation of 0.03 with rent. This result provides some support for the hypothesis that at least part of the reason that upward mobility is not fully priced in equilibrium is that families may be unaware that certain areas are "opportunity bargains."

V.D Summary and Illustrative Application

In this section, we have shown that our measures of upward mobility (1) largely reflect variation in the causal effects of neighborhoods rather than selection and (2) vary substantially even conditional on the cost of housing. These results suggest that these new estimates can potentially be useful for families choosing where to live and for the design of affordable housing policies.

To illustrate how the data could be applied, we return to the MTO experiment and ask the following question: if the experimental vouchers had targeted "opportunity bargain" areas – affordable neighborhoods with the highest levels of upward mobility – rather than those with the lowest poverty rates, how would children's incomes have changed?

To answer this question, we begin by identifying a set of "opportunity bargain" neighborhoods

in each city that could have been targeted in the MTO experiment. To do so, we first count the number of tracts that would have been available to experimental voucher holders in the MTO experiment based on the official requirement that the poverty rate should be less than 10% in the 1990 census, which we denote by N_c. Then, we identify the N_c highest ranking tracts in terms of upward mobility that have both lower median rents than the areas to which MTO voucher holders moved and shorter commute times using public transport to the tracts where control group residents lived (see the notes to Figure XIV for details).[65] Online Appendix Figure VII illustrates some of the areas identified as opportunity bargains in Chicago, which include Uptown (North of the Loop) and Alsip/Marrionette and Evergreen (Southwest of downtown).

Once we have identified the set of opportunity bargain tracts in each city, we compute average individual earnings in adulthood across the relevant Census tracts using our observational estimates, conditional on having parents at the 10th percentile of the income distribution. Figure XIV uses these estimates to predict the earnings children in the MTO experiment would have had if they had moved to these opportunity bargain tracts. It replicates Figure X, but adds five points (in open circles) that show the observational estimates of income on the x axis and predicted values in the MTO sample corresponding to those estimates. These predictions are linear extrapolations using the regression line estimated using the actual MTO experimental estimates, which is shown by the solid line on the figure.

Figure XIV shows that young children in the MTO experiment could potentially have had significantly higher earnings had their families moved to opportunity bargain neighborhoods. On average, across the five MTO sites, we predict that children's individual earnings would have been \$3,800 higher had they moved to opportunity bargain areas instead of Experimental areas. For comparison, the mean earnings gain (relative to the control group) that was actually realized by children in the Experimental group who moved to low-poverty neighborhoods was \$3,500 – implying that one could have achieved double the earnings gains by moving to opportunity bargain areas. These findings suggest that one could improve children's outcomes substantially without increasing expenditure on housing vouchers by helping families use their vouchers move to opportunity bargain areas if they wish to do so. More broadly, this example highlights the potential value of using data on children's observed outcomes rather than traditional proxies for neighborhood quality in policy design.

[65] We obtain similar results if we restrict attention to opportunity bargain areas with a high rate of racial diversity (Online Appendix Figure VIII).

VI Conclusion

Cross-sectional statistics on neighborhood characteristics such as poverty rates and job growth have provided a foundation for economic policy and research on labor markets for several decades. In this paper, we have constructed longitudinal statistics that measure children's outcomes in adulthood based on the Census tract in which they grew up, which can provide an analogous foundation for policies to improve social mobility and research on human capital development.

Using these new statistics, we show that neighborhoods have substantial causal effects on children's long-term outcomes at a highly granular level. Moving to a neighborhood that is just a mile or two away can change children's average earnings by several thousand dollars a year and have significant effects on a spectrum of other outcomes ranging from incarceration to teenage birth rates. Much of this variation in children's outcomes is unrelated to traditional neighborhood-level proxies for economic success – such as rates of job and wage growth – showing that the conditions that create greater upward mobility are not necessarily the same as those that lead to productive labor markets.

Going forward, there are many potential applications of the Opportunity Atlas for both policy and research. Policy makers can use these data to better target programs that aim to improve economic opportunities for disadvantaged children by pinpointing the areas within cities that currently have the weakest outcomes. For example, the placement of pre-school programs or eligibility for local programs or tax credits could potentially be informed by these data. The Atlas can also be used to help low-income families find affordable neighborhoods that offer good opportunities for upward mobility. For example, housing authorities in the Seattle metropolitan area are currently conducting a pilot study that provides information and assistance to housing voucher recipients interested in moving to higher-opportunity areas, motivated by the finding that most voucher recipients currently live in neighborhoods with poor prospects for upward mobility.

For researchers, the Opportunity Atlas offers a new tool to study the determinants of economic opportunity. Comparing areas that have similar observable characteristics but produce different rates of upward mobility using qualitative or quantitative methods can yield new insight into the mechanisms that lead to differences in children's outcomes. The data could also be used to test and develop better short-term proxies for children's long-term outcomes beyond conventional indicators such as test scores. And, as tract-level statistics on children's outcomes are constructed for a longer span of years, they can be used to study the effects of changes in local policies and interventions on

economic mobility. Such research has the potential to yield scalable solutions to increasing mobility out of poverty.

References

Abowd, J. M. and I. M. Schmutte (2015). Economic analysis and statistical disclosure limitation. *Brookings Papers on Economic Activity*, 221–267.

Acevedo-Garcia, D., N. McArdle, E. F. Hardy, U. I. Crisan, B. Romano, D. Norris, M. Baek, and J. Reece (2014). The child opportunity index: Improving collaboration between community development and public health. *Health Affairs 33*(11), 1948–1957. PMID: 25367989.

Akerlof, G. A. (1978). The economics of "tagging" as applied to the optimal income tax, welfare programs, and manpower planning. *The American Economic Review 68*(1), 8–19.

Bond, B., J. D. Brown, A. Luque, and A. O. Hara (2014). The Nature of the Bias When Studying Only Linkable Person Records: Evidence from the American Community Survey. *Working Paper Number: CARRA-WP-2014-08*.

Box, G. E. P. and D. R. Cox (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological) 26*(2), 211–252.

Brooks-Gunn, Jeanne, G. J. Duncan, P. K. Klebanov, and N. Sealand (1993). Do neighborhoods influence child and adolescent development? *American Journal of Sociology 99*, 353–95.

Case, A. C. and L. F. Katz (1991, May). The company you keep: The effects of family and neighborhood on disadvantaged youths. Working Paper 3705, National Bureau of Economic Research.

Chetty, R., J. Friedman, and N. Hendren (2018). Releasing tract-level estimates from us census: A disclosure avoidance strategy. *Working Paper*.

Chetty, R. and N. Hendren (2018a). The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. *The Quarterly Journal of Economics*, qjy007.

Chetty, R. and N. Hendren (2018b). The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates. *The Quarterly Journal of Economics*, qjy006.

Chetty, R., N. Hendren, M. R. Jones, and S. R. Porter (2018, March). Race and economic opportunity in the united states: An intergenerational perspective. Working Paper 24441, National Bureau of Economic Research.

Chetty, R., N. Hendren, and L. F. Katz (2016). The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic Review 106*(4), 855–902.

Chetty, R., N. Hendren, P. Kline, and E. Saez (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.

Christensen, P. and C. Timmins (2018, July). Sorting or steering: Experimental evidence on the economic effects of housing discrimination. Working Paper 24826, National Bureau of Economic Research.

Chyn, E. (2018). forthcoming."moved to opportunity: The long-run effect of public housing demolition on labor market outcomes of children.". *American Economic Review*.

Cutler, D. M. and E. L. Glaeser (1997). Are Ghettos Good or Bad? *The Quarterly Journal of Economics 112*(3), 827–72.

DeLuca, S., H. Wood, and P. Rosenblatt (2017). Why poor families move (and where they go): Reactive mobility and residential decisions. *City and Community*.

Denton, N. and D. Massey (1988). Residential segregation of blacks, hispanics, and asians by socioeconomic status and generation. *Population Research Center, University of Chicago* (69).

Deutscher, N. (2018). *Place, Jobs, Peers and the Importance of the Teenage Years: Exposure Effects and Intergenerational Mobility*. Nathan Deutscher.

Dwork, C. (2008). Differential privacy: A survey of results. In M. Agrawal, D. Du, Z. Duan, and A. Li (Eds.), *Theory and Applications of Models of Computation*, Berlin, Heidelberg, pp. 1–19. Springer Berlin Heidelberg.

Glaeser, E. L., G. A. Ponzetto, and K. Tobio (2011). Cities, skills, and regional change.

Grawe, N. D. (2006). Lifecycle bias in estimates of intergenerational earnings persistence. *Labour Economics 13*(5), 551–570.

Gyourko, J., A. Saiz, and A. Summers (2008). A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index. *Urban Studies 45*(3), 693–729.

Haider, S. and G. Solon (2006). Life-cycle variation in the association between current and lifetime earnings. *American Economic Review 96*(4), 1308–1320.

Holland, P. W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association 81*(396), 945–960.

Jencks, C. and S. E. Mayer (1990). The social consequences of growing up in a poor neighborhood. Technical report, National Research Council.

Kain, J. F. (1968). Housing segregation, negro employment, and metropolitan decentralization. *The Quarterly Journal of Economics 82*(2), 175–197.

Kline, P. and E. Moretti (2014). People, places and public policy: Some simple welfare economics of local economic development programs. *Annual Review of Economics 6*(1), forthcoming.

Laliberté, J.-W. P. (2018). Long-term contextual effects in education: Schools and neighborhoods.

Layne, M., D. Wagner, and C. Rothhaas (2014). Estimating Record Linkage False Match Rate for the Person Identification Validation System. *Center for Administrative Records Research and Applications Working Paper 2*.

Leventhal, T. and J. Brooks-Gunn (2000). The neighborhoods they live in: the effects of neighborhood residence on child and adolescent outcomes. *Psychological bulletin 126*(2), 309.

Maponics (2017). Maponics school boundaries. Pitney Bowes. https://www.pitneybowes.com/us/data/boundary-data/school-district-boundaries.html.

Massey, D. and N. Denton (1987). Trends in the Residential Segregation of Blacks, Hispanics, and Asians: 1970-1980. *American Sociological Review 52*(6), 802–825.

Messer, L. C., B. A. Laraia, J. S. Kaufman, J. Eyster, C. Holzman, J. Culhane, I. Elo, J. G. Burke, and P. O'Campo (2006, September). The development of a standardized neighborhood deprivation index. *Journal of Urban Health : Bulletin of the New York Academy of Medicine 83*(PMC3261293), 1041–1062.

Moretti, E. (2011). *Local Labor Markets*, Volume 4 of *Handbook of Labor Economics*, Chapter 14, pp. 1237–1313. Elsevier.

Neyman, J. (1923). On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. *Statistical Science 5*(4), 465–472. Translation by Dorota M. Dabrowska and Terence P. Speed.

Nichols, A. L. and R. J. Zeckhauser (1982). Targeting transfers through restrictions on recipients. *The American Economic Review 72*(2), 372–377.

Office of Management and Budget (1997). Race and Ethnic Standards for Federal Statistics and Administrative Reporting. *Statistical Policy Directive 15*.

Opportunity-Nation (2017). The 2017 opportunity index. Technical report, Opportunity Nation. http://opportunityindex.org/wp-content/uploads/2017/12/2017-Opportunity-Index-Full-Analysis-Report.pdf.

Reardon, S. Fox, L. and J. Townsend (2015). Neighborhood income composition by household race and income, 1990–2009. *ANNALS of the American Academy of Political and Social Science 660*(1), 78–97.

Rupasingha, A. and S. J. Goetz (2008). US county-level social capital data, 1990-2005. *The Northeast Regional Center for Rural Development, Penn State University, University Park, PA*.

Sampson, R. J. (1987). Urban Black Violence: The Effect of Male Joblessness and Family Disruption. *American Journal of Sociology 93*(2), 348–382.

Sampson, R. J., J. D. Morenoff, and T. Gannon-Rowley (2002). Assessing "neighborhood effects": Social processes and new directions in research. *Annual Review of Sociology 28*(1), 443–478.

Sampson, R. J., S. W. Raudenbush, and F. Earls (1997). Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy. *Science 277*(5328), 918–924.

Sanbonmatsu, L., J. Ludwig, L. F. Katz, L. A. Gennetian, G. J. Duncan, R. C. Kessler, E. Adam, T. McDade, and S. T. Lindau (2011). *Moving to Opportunity for Fair Housing Demonstration Program: Final Impacts Evaluation*. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

Sharkey, P. (2016). Neighborhoods, cities, and economic mobility. *RSF 2*(2), 159–177.

Solon, G. (1999). Intergenerational mobility in the labor market. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3, pp. 1761–1800. Elsevier.

U.S. Department of Commerce, Bureau of the Census (2000). United States Census 2000: Informational Copy. https://www.census.gov/dmd/www/pdf/d-61b.pdf. Form D-61B.

U.S. Department of Commerce, Bureau of the Census (2003). Chapter 5, Sample Design and Estimation; 2000 Census of Population and Housing: Public-use Microdata Samples Technical Documentation. Technical report, U.S. Department of Commerce, Bureau of the Census.

U.S. Department of Commerce, Bureau of the Census (2014). American Community Survey Design and Methodology (January 2014); Chapter 4: Sample Design and Selection. Technical report.

Wagner, D. and M. Layne (2014). The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications'(CARRA) Record Linkage Software. *Center for Administrative Records Research and Applications Working Paper 1.*

ONLINE APPENDICES

A. Construction of Neighborhood Characteristics

This appendix provides definitions and sources for covariates used throughout the paper or shown in the Opportunity Atlas as neighborhood characteristics. Our source data are primarily at the tract level. We use 2010 Census tract definitions throughout. For covariates defined using 2000 tract boundaries, we use the 2010 Census Tract Relationship Files from the US Census Bureau to crosswalk 2000 tracts to 2010 tracts, weighting the 2000 tract-level covariates by the fraction of the 2000tract population that lives within the 2010 tract boundaries.

Tract-Level Characteristics:

Jobs Within 5 Miles (2015). The number of jobs within 5 miles of a tract is constructed using block-level information on the total number of jobs from the Workplace Area Characteristics (WAC) data files in the LEHD Origin-Destination Employment Statistics (LODES) provided by the Census Bureau. For each tract we compute the number of jobs within 5 miles as the total number of jobs in own and neighboring tracts whose centroids fall within a radius of 5 miles from the centroid of the tract.

Number of High Paying Jobs Within 5 Miles (2015). The number of high paying jobs within 5 miles of a tract is constructed using block-level information on the number of jobs with earnings greater than $3,333 per month from the Workplace Area Characteristics (WAC) data files in the LEHD Origin-Destination Employment Statistics (LODES) provided by the Census Bureau. For each tract we compute the number of high paying jobs within 5 miles as the number of high paying jobs in own and neighboring tracts whose centroids fall within a radius of 5 miles from the centroid of the tract.

Job Growth (2004-2013). The measure of job growth at the tract level shown in the Opportunity Atlas is constructed using block-level information on the total number of jobs from 2004 to 2013 from the Workplace Area Characteristics (WAC) data files in the LEHD Origin-Destination Employment Statistics (LODES) provided by the Census Bureau. We compute job growth in each tract as the average annualized growth rate from 2004 to 2013.

Job Density (2013). The measure of job density at the tract level shown in the Opportunity Atlas is constructed combining block-level information on total number of jobs in 2013 from the Workplace Area Characteristics (WAC) data files in the LEHD Origin-Destination Employment Statistics (LODES) provided by the Census Bureau and tract-level information on land area in square miles from the 2010 Census Gazetteer Files. We compute job density as the number of jobs per square mile in each tract.

Employment Rate (2000). The rate of employment is constructed using tract-level data on labor market measures from tables NP043E and NP043C of the Census long form SF3a dataset obtained from the National Historical Geographic Information System (NHGIS) database. We construct the rate of employment in 2000 for each tract as the total employed population (the sum of employed females and employed males) divided by the total population 16 years and over.

Poverty Rate (1990, 2000, 2006-2010, 2011-2015). The poverty share variable is constructed as the share of individuals below the federal poverty line in each tract. For the measure in 1990 we use table NP117 from the 1990 Census form SFT3, for the measure in 2000 we use table NP087B from the 2000 Census long form SF3a, and for the measures for 2006-2010 and 2011-2015 we use table C17002 from the American Community Survey in relevant years, all obtained from the NHGIS database.

Single Parent Share (1990, 2000, 2006-2010). We define the share of single parents in each tract as the number of households with female head (and no husband present) or male head (and no wife

present) with own children under 18 years old present divided by the total number of households with own children present. We use table NP19 of the 1990 Census form SFT3 for the measure in 1990, tables NP018E and NP018G of the 2000 Census SF1a form for the measure in 2000, and table B11003 of the 2006-2010 American Community Survey for the measure in 2010. All obtained from the NHGIS database.

Racial Shares (2000, 2010). Racial shares are calculated from the Census long form SF1a, tables NP008A and NP004E, taken from NHGIS database. All races (except Hispanic) exclude Hispanics and Latinos.

Share Foreign Born (2010). The share foreign born variable that is shown in the Opportunity Atlas is constructed as the number of foreign born residents in the 2010 Census divided by the sum of native and foreign born residents (long form SF3a, table NP021A) obtained from the NHGIS database.

Share with Short Commute to Work and Mean Commute Time (2000, 2006-2010). The share of workers with a short commute to work and mean commute time are constructed using tract-level data from table NP031B of the 2000 Decennial Census or tract-level data from table B08303 of the the 2006-2010 American Community Survey, both obtained from the NHGIS database. Fraction with a short to commute to work is computed by taking the share of people who commute less than 15 minutes to work over all workers 16 years and over who did not work at home. Mean commute time is constructed using the share of workers commuting to work in specific bins ($<$ 5 minutes, 5-9 minutes, 10-14 minutes, etc.), imputing the mean time commuted in a given bin (i.e. for 5-9 minutes, imputing mean commute time of 7 minutes), and then calculating a sum of imputed mean commute times within each bin weighted by the share commuting.

Kid Counts (2000). The counts of kids by race and gender used throughout the paper and shown in the Opportunity Atlas are constructed for kids under 18 using tract-level data from tables NP012F and NPCT012H of the 2000 Decennial Census using the NHGIS database.

Census Return Rate (2010). The Census return rate variable used in Figure V and shown in the Opportunity Atlas is obtained from tract-level data from the Census 2016 Planning Database. It is calculated as the number of 2010 Census mail forms completed and returned over the number of valid occupied housing units where a Census form was expected to be delivered for mail return to Census.

Mean Household Income (2000). The measure of mean household income used in Figure V is constructed using tract-level data from table NP052A of the 2000 Decennial Census found in the NHGIS database.

Median Household Income (1990, 2012-2016). The measure of median household income shown in the Opportunity Atlas is constructed using tract-level data from table NP80A of the 1990 Decennial Census and table B19013 of the American Community Survey (2012-2016) found in the NHGIS database.

High School Graduate Wage Growth (2005-2014). The measure of high school graduate wage growth is constructed using data from the 2005-2009 and 2010-2014 American Community Survey provided by NHGIS database. High school graduate wages at the tract level are computed by dividing the average high school graduate annual earnings by the product of overall average weekly hours worked and 52. High school graduate wage growth is then computed as the difference in logarithms between high school graduate wages in 2010-2014 and school graduate wages in 2005-2009.

Share College Graduate (2000, 2006-2010). The share college graduate variable shown in the Opportunity Atlas is constructed using tract-level data from table NP037C of the 2000 Census long form SF3a or tract-level data from table B15002 of the 2006-2010 American Community Survey (both obtained from the NHGIS database), and is calculated as the number of people aged 25 or

older who have a bachelor's degree, master's degree, professional school degree, or doctorate degree, divided by the total number of people aged 25 or older in a tract.

Population Density (2000, 2010). The population density variable used in Figure VIIIc and shown in the Opportunity Atlas is calculated as the total tract-level population in the Census obtained from NHGIS database (long form SF1a, table H7V) divided by tract land area in square miles from the 2010 Census Gazetteer Files.

Median Two-Bedroom Rent (2011-2015). The median two-bedroom rent variable that is used in Figure XIII and shown in the Opportunity Atlas is constructed from tract-level ACS data (2011-2015) and is defined as the median gross rent for renter-occupied housing units with two bedrooms that pay cash rent (table AD79).

Characteristics at Other Levels of Geography:

Job Growth (1990-2010, 2004-2013). The measure of job growth at the CZ or MSA level that we use in Figure VI and Online Appendix Figure III is constructed as the percentage change in employment between 1990 and 2010 in each CZ/MSA using county-level data from the Local Area Unemployment Statistics (LAUS) released by the Bureau of Labor Statistics. The measure of job growth at the county and CZ level that we use in the Opportunity Atlas is constructed as the average annualized growth rate in employment between 2004 and 2013 in each CZ using county level data from the Local Area Unemployment Statistics (LAUS) released by the Bureau of Labor Statistics.

Opportunity Zones. The list of tracts in Qualified Opportunity Zones shown in Online Appendix Figure V was downloaded from the Community Development Financial Institutions Fund.

Wharton Land Use Regulation Index (2008). The Wharton Land Use Regulation Index is constructed using city-level data from Gyourko et al. (2008). The cities in the original dataset are crosswalked to 247 commuting zones (representing 87% of the US population).

3rd Grade Math Score. Data for 3rd grade test scores are downloaded from the Stanford Education Data Archive and measured at the district level. We create a crosswalk from districts to tracts by weighting by the proportion of land area that a given school district covers in a tract.

High School Catchment Areas. We match tracts to high school catchment areas across the U.S. using data on the intersection of census tracts with high school catchment areas in 2017 provided by Peter Bergman. These data come from Maponics (2017). Tracts are not perfectly nested within catchment areas; we create an approximate crosswalk by assigning tracts to the school catchment area that contains the majority of their land area. In a few cases where school catchment areas overlap (e.g. a whole tract belongs to two different school catchment areas) we assign the tract to the largest of the catchment areas that contain it. The results of our variance decomposition analysis are very similar if we alternatively assign these tracts to the smallest catchment area or simply don't use these tracts in the analysis. Shape-files of Mecklenburg County high school catchment areas in 2002 and 2017 come from the Charlotte-Mecklenburg Schools (CMS) education agency. The shape-files for the 2002 Mecklenburg County boundaries were generously provided by David Deming.

Other Measures of Opportunity. We compare our Opportunity Atlas measures with existing indices of economic opportunity. We obtain data for the Kirwan Child Opportunity Index at the metropolitan area level constructed by the Kirwan Institute and the Institute for Child, Youth and Family Policy (ICYFP) from diversitydatakids.org, and we obtain data for the Area Deprivation Index at the block level constructed by the University of Wisconsin School of Medicine and Public Health. Area Deprivation Index.

B. Estimating Causal Effects of Neighborhoods: Methodology

In this appendix, we document the sample, variable construction, and empirical specifications used for the quasi-experimental analysis in Section V.B.

Sample and Variable Construction. Our core sample and data construction is the same as that described in Section II, but expands in two directions that increase our ability to observe moves at younger ages. First, we extend our analysis to include the 1978-1991 cohorts. Second, we focus on income ranks measured at age 24 (as in Chetty and Hendren (2018a)), in addition to marriage at age 30 and incarceration (measured as in Section III), in order to make the measures comparable across cohorts.[66]

Using the location of each child's parents in each year in our sample, we form a sample of one-time movers. These are defined as children whose parents move across tracts exactly once when the children are age 28 or below.[67] We define the year of the move as the tax year in which the parents report living in a different tract relative to the previous year. In cases where we do not observe sequential years of location information (e.g. we do not observe 1990-93 and 1996-97), we assign the year of move as the midpoint between the two nearest years in which different addresses are reported (e.g. if we see a new location in 1994 relative to 1989, we assign the year of move to be 1992.5). In cases where this leads to a non-integer year of move, we randomly select the nearest year for the move. We then define the child's age at the time of the move as the year of the move minus the child's cohort.

Following Chetty and Hendren (2018a), we make three additional sample restrictions. First, we restrict to moves between origins and destinations that have at least 20 observations used to calculate $\bar{y}_{po}$ and $\bar{y}_{pd}$. As shown in Online Appendix A of Chetty and Hendren (2018a) imposing such sample restrictions limits the impact of attenuation bias from sampling error in the $\bar{y}_{pc}$ estimates. Second, we require that we are able to observe the parents for at least two years after the move in order to enter the sample (e.g. for a child born in 1991 we only consider moves through 2013, since s/he is observed until 2015). Third, we require families to move at least 25 miles; moves less than 25 miles suffer from more severe measurement error, as we discuss in more detail below. Appendix Table IV presents summary statistics for the one-time movers sample and the complementary exposure-weighted sample.

For each subgroup of the analysis, g (e.g. g could represent black males, white females, etc), we use the complementary sample of children (who never move or who move more than one time) to provide an estimate of the average outcomes of children in group g who grew up in each tract. Using this sample, we restrict to those in group g and construct exposure-weighted outcomes, $\bar{y}_{pgl}$, for each parental income p, demographic group g, and location l. To do so, we regress children's outcomes on a linear term in transformed parental income rank (based on the national non-parametric relationship), weighting by the number of years below age 23 in which the parents are observed in the tract. We let $\bar{y}_{pgl}$ denote the predicted value from this regression for a child at parental income rank p.

Empirical Specification. Using the sample of one-time movers, we consider the outcomes of child i with parental income rank p_i who moved at age m_i from origin tract, o, to destination tract, d. Weuse a specification analogous to the approach in Chetty and Hendren (2018a). Let $\bar{y}_{pgl}$ denote

[66] Some variables are only defined for a subset of these cohorts. Marriage at age 30 cannot be observed past the 1985 cohort. Because individual income is only well defined starting in 2005, our age 24 individual income measure is missing for cohorts 1978-1980. Finally, we require incarceration to be measured after age 23, and therefore omit cohorts 1987 and later for that outcome.

[67] When constructing the sample, we observe location up to age 30. But, as discussed below, we follow Chetty and Hendren (2018a) and require that we observe the parents in the destination for at least two years. Therefore, the oldest age of move for the children is 28.

the exposure-weighted outcome of y_i for children who grew up in tract l with parental income rank $p = p_i$, where we will substitute $l \in \{o, d\}$ to denote origin tract and destination tract, respectively. Let $\Delta_{odpg} = \bar{y}_{pgd} - \bar{y}_{pgo}$ denote the difference in the income rank of exposure-weighted residents in the destination versus origin for children with parental income rank p. We then run three primary specifications. Our "extended specification" is:

$$y_i = \sum_{m=2}^{28} I(m_i = m) \left[\alpha_m + \phi_m \bar{y}_{pgo} + \zeta_m p_i + b_m \Delta_{odpg}\right] + \varepsilon_i$$

where the parameters $\{\alpha_m, \phi_m, \zeta_m\}$ are age-at-move-specific intercepts, controls for the average exposure-weighted outcome in the origin, and parental income rank, respectively. The key parameters of interest are the b_m coefficients, which capture how children's outcomes vary with the age at which they move to an area with higher or lower predicted earnings. Our "parametric specification" is:

$$\begin{aligned} y_i &= \sum_{m=2}^{28} I(m_i = m) \left[\alpha_m + \phi_m \bar{y}_{pgo} + \zeta_m p_i\right] + I(m_i \leq 23)(\gamma' + \gamma m_i)\Delta_{odpg} \\ &\quad + I(m_i > 23)(\rho' + \rho m_i)\Delta_{odpg} + \varepsilon_i \end{aligned}$$

which differs from the extended specification in that it does not interact age-at-move with the difference in tract predicted outcome measures Δ_{odpg}. One can interpret the coefficients ρ and γ as the slope of the b_m coefficients above and below age 23, respectively. Our "parsimonious specification" is:

$$\begin{aligned} y_i &= \sum_{m=2}^{28} I(m_i = m) \left[\alpha_m + \zeta_m p_i\right] + \phi'_m \bar{y}_{pgo} + I(m_i \leq 23)(\gamma' + \gamma m_i)\Delta_{odpg} \\ &\quad + I(m_i > 23)(\rho' + \rho m_i)\Delta_{odpg} + \varepsilon_i \end{aligned}$$

which further removes the interaction between age-at-move and origin tract predicted outcomes. In all specifications, we address classical measurement error in $\bar{y}_{pgo}$ and Δ_{odpg} using a split-sample instrumental variables approach. We randomly split those in the complementary sample into two groups, requiring that siblings are included in the same group to ensure independence of the samples, and instrument for $\bar{y}_{pgo}$ and Δ_{odpg} measured in one group using the same variables as measured in the second group.

Distance Restriction. In all movers regressions, we include only moves with 25 miles or more between the origin and destination tracts (as measured using the latitude and longitude of the tract centroid). We do so in order to avoid relatively large mismeasurement of $\bar{y}_{pgo}$ and Δ_{odpg} due to mismeasurement of the origin location (since it is in fact the destination from a prior, unobserved move). While this issue arises whenever we do not observe an individual from birth (which is for most cohorts), this issue is most severe for short-distance moves, where many of the single moves that we observe are in fact a return to an original tract of residence (or a very similar tract to the final destination).

In order to assess origin mismeasurement quantitatively, we focus on the later cohorts of children, for whom we can recreate the missing data problem in earlier cohorts by artificially restricting to later ages. Specifically, we focus on cohorts 1986-1991 and use data on location from age 11 onwards (this is the first age at which we observe location for the 1978 cohort). Among those children classified as one-time movers in this truncated sample, we then examine location before

age 11. Denote by o and d the origin and destination tracts for the one-time move after age 11. Denote by o' the origin of the pre-age-11 move. We then classify individuals based on the distance of the post-age-11 move.

This analysis yields several findings about the difference between children making post-age-11 moves that are greater than vs. less than 25 miles. (We refer to these two groups as "short-" and "long-distance" movers, respectively, for this section.) While short- and long-distance movers have similar probabilities of moving before age 11, short-distance movers are disproportionately likely to move to tracts d which are very similar to o'. The signal correlation between $\bar{y}_{pgo'}$ and $\bar{y}_{pgd}$ is higher for short-distance movers than for long-distance movers, and tract o' is more likely to lie within 5 miles of tract d for the former group.

The relationship between $\bar{y}_{pgo'}$ and $\bar{y}_{pgd}$ is particularly pronounced for children with single parents, for whom short-distance moves may in fact simply reflect children being claimed in different years by two separated parents living in different neighborhoods in the same city. (This is far less of a worry for long-distance movers.) This suggests that children of married parents would not suffer from the same measurement issues, even for short-distance moves. To verify this, in Appendix Table III we rerun the specification in Column 1 of Table IV, splitting children by married vs. unmarried parents. For children of married parents, in Columns 3 and 6, the key coefficient on the interaction between move quality and age-at-move (below age 23) is -0.022 (0.002) for short-distance movers, which is similar to the coefficient of -0.030 (0.002) for long-distance movers. In contrast, short- and long-distance movers look very different for children of unmarried parents in Columns 1 and 4, and 2 and 6, respectively. Based on this evidence, we restrict our analysis in the main text to those who move more than 25 miles.

Table I

Summary Statistics for Primary Analysis Sample (1978-1983 Birth Cohorts)

	Pooled	Male	Female
	(1)	(2)	(3)
A. Parental Characteristics			
Median Parent Household Income ($)	56,730	56,890	56,560
Mean Parent Household Income Percentile Rank	50.5	50.6	50.5
SD of Parent Household Income Percentile Rank	28.8	28.8	28.8
Father Present in Household?	78.9%	79.7%	78.2%
Mother Present in Household?	89.7%	89.2%	90.3%
Both Parents Present in Household?	68.7%	68.9%	68.5%
B. Income and Employment Outcomes in Adulthood			
Median Household Income ($)	42,360	41,250	43,590
Mean Household Income Percentile Rank	50.2	48.9	51.6
SD Household Income Percentile Rank	28.9	29.2	28.4
Median Individual Income ($)	29,440	35,120	24,390
Mean Individual Income Percentile Rank	50.2	53.9	46.4
SD of Individual Income Percentile Rank	28.9	29.7	27.4
Household Income in the Top Quintile?	20.2%	19.0%	21.6%
Individual Income in the Top Quintile?	20.2%	25.3%	14.9%
Household Income in the Top 1%?	1.0%	0.9%	1.2%
Individual Income in the Top 1%?	1.0%	1.4%	0.7%
Employed (Individual Income > 0)?	76.5%	77.8%	75.1%
C. Other Outcomes on Full Population			
Married?	45.1%	42.6%	47.8%
Incarcerated on April 1, 2010?	1.5%	2.7%	0.3%
Had a Child as a Teenager?			19.7%
Mean Spouse Individual Income Percentile Rank	62.4	53.5	71.2
SD of Spouse Individual Income Percentile Rank	26.5	27.4	22.3
Lives in Low Poverty Tract?	47.9%	47.4%	48.5%
Stays in Childhood CZ in Adulthood?	66.0%	66.4%	65.5%
Stays in Childhood Tract in Adulthood?	20.5%	22.7%	18.3%
Lives with Parents in Adulthood?	15.0%	16.7%	13.3%
Mean Household Income Rank | Stays in Childhood CZ	47.9	46.6	49.2
SD Household Income Rank | Stays in Childhood CZ	28.0	28.3	27.5
Mean Individual Income Rank | Stays in Childhood CZ	48.8	52.0	45.5
SD Individual Income Rank | Stays in Childhood CZ	27.9	29.1	26.1
D. ACS Outcomes			
Employed?	84.8%	88.6%	81.0%
Hours Worked Per Week	31.9	35.7	28.1
SD of Hours Worked Per Week	18.8	18.4	18.4
Median Hourly Wages ($)	18.2	19.3	17.2
Mean Hourly Wage Percentile Rank	50.2	52.3	48.0
SD of Mean Wage Percentile Rank	28.9	29.0	28.6
Graduated from High School?	86.2%	83.7%	88.6%
Earned Some College Credits?	69.5%	63.8%	75.1%
Graduated with 2-Year College Degree?	46.3%	40.5%	51.9%
Graduated with 4-Year College Degree?	36.4%	31.6%	41.1%
Has Post-Graduate Degree?	13.3%	10.6%	16.0%
Receives Public Assistance?	2.3%	1.4%	3.2%
Mean Household Income Rank | Child of U.S. Native Parents	53.2	52.0	54.4
SD of Household Income Rank | Child of U.S. Native Parents	28.4	28.7	28.0
Mean Individual Income Rank | Child of U.S. Native Parents	52.0	56.4	47.4
SD of Individual Income Rank | Child of U.S. Native Parents	28.4	29.0	27.1
Mean Household Income Rank | Child of Immigrant Parents	53.0	51.1	55.1
SD of Household Income Rank | Child of Immigrant Parents	29.0	29.4	28.5
Mean Individual Income Rank | Child of Immigrant Parents	54.1	56.4	51.7
SD of Individual Income Rank | Child of Immigrant Parents	29.6	30.2	28.7
Fraction Not Suppressed in Tract Collapse	0.9998	0.9996	0.9995
Number of Obs in Full Population	20,500,000	10,400,000	10,000,000
Fraction of Total	100.0%	51.0%	49.0%
Number of Obs in ACS Work Sample	1,409,000	696,000	713,000
Number of Obs in ACS Full Sample	3,979,000	1,979,000	2,000,000

Notes: This table presents summary statistics for children in our primary analysis sample. All values in this and all subsequent tables and figures have been rounded to four significant digits as part of the disclosure avoidance protocol. Counts are rounded in the following manner: numbers between 10,000 and 99,999 are rounded to the nearest 500; between 100,000 and 9,999,999 to the nearest 1,000 and above 10,000,000 to the nearest 10,000. Sources for this and all subsequent tables and figures: authors calculations based on Census 2000 and 2010, tax returns, and American Community Surveys 2005-2015. See Section II for definitions of variables. See Online Appendix Table I for analogous summary statistics by race and ethnicity.

Table II

Variance Decomposition for Tract-Level Estimates of Upward Mobility

	All Races (1)	White (2)	Black (3)	Hispanic (4)	Asian (5)	American Indian and Alaska Native (6)
A. Household Income, for Children of Parents at the 25th Percentile						
Mean	40.46	45.06	31.99	42.85	57.38	30.76
Total SD	6.50	6.09	3.82	4.16	6.81	6.34
Noise SD	1.95	2.80	1.90	2.29	3.73	2.70
Reliability	0.91	0.79	0.75	0.70	0.70	0.82
Signal SD	6.20	5.41	3.31	3.47	5.70	5.73
	($6686)	($6126)	($3435)	($3813)	($8243)	($5872)
Within County Signal SD	4.65	3.52	2.35	2.28	3.71	2.29
	($5002)	($3967)	($2439)	($2500)	($5348)	($2369)
B. Share Incarcerated, for Sons of Parents at the 25th Percentile						
Mean	4.81	2.99	11.27	3.35	0.52	5.45
Total SD	4.11	3.19	5.45	2.90	2.51	4.34
Noise SD	2.52	2.60	3.60	2.05	2.30	2.88
Reliability	0.62	0.34	0.56	0.50	0.16	0.56
Signal SD	3.24	1.85	4.09	2.05	1.00	3.25
Within County Signal SD	2.46	1.47	2.38	1.30	0.81	1.31
C. Household Income, for Children of Parents at the 75th Percentile						
Mean	58.31	60.57	43.40	53.76	65.22	44.79
Total SD	5.64	4.62	5.94	5.96	6.93	10.29
Noise SD	2.05	2.15	4.18	4.51	4.31	5.84
Reliability	0.87	0.78	0.51	0.43	0.61	0.68
Signal SD	5.25	4.09	4.22	3.90	5.43	8.47
	($7768)	($6373)	($4672)	($5173)	($9405)	($9644)
Within County Signal SD	4.20	2.80	3.26	3.09	3.98	2.32
	($6205)	($4368)	($3599)	($4096)	($6895)	($2609)

Notes: This table reports estimates of variance components of predicted children's adult outcomes conditional on parent income at the 25th and 75th percentiles (see Figure 2 for more detail). Panel A and C analyzes upward mobility for children with parent incomes at the 25th and 75th percentiles, respectively. The total SD is simply the national tract-level standard deviation of upward mobility estimates, weighted by the number of qualifying children in each tract. The noise SD is the square-root of the average squared standard error; the signal SD is the square root of the difference between the total variance and noise variance. Reliability is the ratio of signal variance to total variance. In parentheses we report the dollar values corresponding to the standard deviations listed in the row above. Column 1 reports these statistics for upward mobility estimates including all children; Columns 2 through 6 report the same statistics for upward mobility estimates including only children from a specific race (where all groups other than Hispanic include only non-Hispanics). Panel B replicates Panel A using upward mobility estimates for male incarceration, measured as living in group quarters on April 1, 2010, in the 2010 Decennial Census.

Table III

Correlations Between Tract-Level Estimates of Children's Outcomes

A. Mean Household Income Ranks: Correlation Across Racial Groups and Parental Income Levels

	Parents at 25th Percentile					
	White	Black	Hispanic	Asian	American Indian & Alaska Natives	Parents at 75th Pctile, Same Race
	(1)	(2)	(3)	(4)	(5)	(6)
White	1	0.573	0.580	0.523	0.636	0.604
Black		1	0.546	0.357	0.436	0.454
Hispanic			1	0.374	0.602	0.353
Asian				1	0.267	0.465
American Indian & Alaska Natives					1	0.357

B. Correlations Across Outcomes

	Parents at 25th Percentile						
	Household Income Rank	Individual Income Rank	Employment Rate	Incarceration Rate	Fraction w/ Teen Birth	Fraction Married	Household Income Rank (Parent p = 75)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Household Income Rank (Parent p = 25)	1	0.858	0.346	-0.276	-0.480	0.608	0.547
Individual Income Rank		1	0.512	-0.253	-0.418	0.311	0.478
Employment Rate			1	-0.114	-0.071	0.040	0.141
Incarceration Rate				1	0.182	-0.198	-0.154
Fraction with Teen Birth					1	-0.273	-0.350
Fraction Married						1	0.253
Household Income Rank (Parent p = 75)							1

Notes: This table presents correlations between tract-level estimates of various child outcomes conditional on parent income at the 25th percentile (Columns 1- 5) or the 75th percentile (Column 6). Panel A presents correlations between household income rank conditional on parent income at the 25th percentile across different racial groups. These correlations are “signal” correlations (see notes to Figure V for details) and are estimated from variation across tracts within CZs. Panel B presents correlations between seven different tract-level outcomes including all subgroups of children. Six of the outcomes are conditional on parent income at the 25th percentile: household income rank, individual income rank, an indicator for positive earnings (employment), an indicator for incarceration on April 1, 2010 (see Figure I for more details), an indicator for claiming a child born when the child is between 13 and 19 years old (teen birth, defined for women only), and an indicator for filing taxes as a married couple in 2015. The last outcome is household income rank conditional on parent income at the 75th percentile. These correlations control for race (see notes of Figure V for details) and CZ fixed effects, and are adjusted for attenuation due to noise infused in the estimates.

Table IV
Quasi-Experimental Estimates of Tract-Level Exposure Effects

	Baseline (1)	No Age Interactions (2)	Married at 30 (3)	Incarcerated (4)	Family FEs (5)
Age <= 23	-0.027	-0.026	-0.027	-0.025	-0.021
	(0.001)	(0.001)	(0.001)	(0.005)	(0.002)
Age > 23	-0.008	-0.004	0.003	0.010	-0.004
	(0.009)	(0.008)	(0.009)	(0.033)	(0.009)
Num. of Obs.	2,814,000	2,814,000	1,614,000	1,484,000	2,814,000

Notes: This table reports estimates of annual childhood tract level exposure effects on children's household income ranks at age 24 (Columns 1, 2 and 5), marriage (Column 3), and incarceration (Column 4). Standard errors are shown in parentheses. Each column reports estimates from a split-sample IV regression of a child's household income rank at age 24 on the difference between parent income-specific predicted income ranks in the destination vs. the origin, interacted with the age of the child at the time of the move (m). Column 1 reports estimates from equation (9) using all children in the primary analysis sample of one-time movers. We permit separate linear interactions for age m≤23 and m>23. The estimates can be interpreted as the impact of delaying by one year moving to a tract which has a 1 percentile point higher predicted income rank, essentially a linear fit to the coefficients in Figure XI above and below age 23. Column 2 estimates exposure effects omitting the interaction terms between age and predicted ranks in the origin tract that were included in column 1. Columns 3 and 4 replicate column 1 using marriage rates at 30 and incarceration rates respectively (rather than household income ranks) to measure both the child's outcome (dependent variable) and the predictions (independent variables). Column 5 adds family fixed effects to the specification in column 1; here we identify exposure effects from families of one-time movers with two or more children of different ages at the time of move.

Table V

Quasi-Experimental Estimates of Tract-Level Exposure Effects: Placebo Tests

	Income Rank at 24	Married at 30	Incarceration
	(1)	(2)	(3)
Mean Income Rank at 24	-0.026	0.003	-0.001
	(0.002)	(0.004)	(0.001)
Frac. Married at 30	-0.002	-0.025	0.000
	(0.001)	(0.002)	(0.000)
Incarceration Rate	-0.017	-0.019	-0.036
	(0.009)	(0.023)	(0.005)
Num. of Obs.	2,222,000	1,614,000	1,481,000

Notes: This table reports estimates of annual childhood exposure effects when using predictions for multiple outcomes simultaneously. Standard errors are shown in parentheses. Each column reports estimates of annual childhood exposure effects based on children's predicted income ranks at age 24, the marriage rates at age 30, and incarceration rates on April 1, 2010. The estimates in each column can be interpreted as the impact on a given individual outcome of delaying by one year moving to a tract which has a 1 percentile or 1 pp higher predicted value of each of the regressors. Column 1 uses child income rank at age 24 as the outcome, while columns 2 and 3 use marriage at 30 and incarceration as the outcome variables, respectively.

Table VI

Quasi-Experimental Estimates of Tract-Level Causal Exposure Effects: Heterogeneity Analysis

	Baseline (1)	Good and Bad Moves (2)	Large Moves (3)	Observed Components of Opportunity (4)	Unobserved Components of Opportunity (5)
Age <= 23	-0.027		-0.046	-0.020	-0.025
	(0.001)		(0.017)	(0.001)	(0.003)
Age <= 23, Good Moves		-0.031			
		(0.002)			
Age <= 23, Bad Moves		-0.027			
		(0.002)			
Observations	2,814,000	2,814,000	22,500	2,692,000	2,692,000

Notes: This table reports estimates of annual childhood exposure effects on children's family income ranks at age 24 for different subgroups of one-time movers. See notes to Table IV for more details on these specifications. Standard errors are shown in parentheses. Column 1 replicates Column 1 from Table IV as a reference. Column 2 reports exposure effects separately for one-time movers who move to tracts with higher vs. lower predicted income ranks using a specification that allows the effects to vary for these two groups. Column 3 restricts to the subgroup of one-time movers who move either from the top to bottom or bottom to top decile of the within-CZ rankings of upward mobility estimates. In Column 4, we replace mean observed income ranks on the right hand side of the regression with ranks predicted based on the following neighborhood characteristics: the total number of jobs within 5 miles, the total number of high paying jobs within 5 miles, local unemployment rates, local poverty rates, local high school test scores, the fraction attending college locally, the fraction completing high school locally, the median two-bedroom rent in the tract, the share of area residents who are owner-occupiers, the local share of single-parent families, and area population density. In Column 5, we instead use the residuals from this regression (the "unobservable" component of incomes) as the regressor.

Online Appendix Table Ia

Summary Statistics for Primary Analysis Sample, by Race and Ethnicity (1978-1983 Birth Cohorts)

	White			Black			Hispanic		
	Pooled	Male	Female	Pooled	Male	Female	Pooled	Male	Female
	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. Parental Characteristics									
Median Parent Household Income	71,470	71,510	71,430	29,600	29,910	29,300	33,470	33,400	33,540
Mean Parent Household Income Percentile Rank	58.4	58.4	58.3	33.1	33.4	32.8	36.5	36.4	36.6
SD of Mean Parent Household Income Percentile Rank	27.1	27.0	27.1	24.4	24.5	24.2	25.4	25.4	25.5
Father Presence	86.2%	86.8%	85.5%	49.6%	50.7%	48.4%	73.7%	74.8%	72.7%
Mother Presence	93.4%	92.9%	93.9%	83.0%	82.3%	83.6%	83.5%	82.5%	84.4%
Two Parent Family	79.6%	79.7%	79.5%	32.5%	33.0%	32.0%	57.2%	57.3%	57.2%
B. Income and Employment Outcomes in Adulthood									
Median Household Income ($)	53,920	52,120	55,970	20,740	17,780	22,820	35,250	35,310	35,190
Mean Household Income Percentile Rank	55.8	54.7	57.0	34.9	32.7	37.0	45.7	44.7	46.8
SD of Mean Household Income Percentile Rank	28.3	28.5	28.1	23.2	24.3	21.9	26.1	26.5	25.6
Median Individual Income ($)	33,760	40,830	26,730	19,630	18,270	20,510	27,220	32,280	23,060
Mean Individual Income Percentile Rank	53.4	58.6	48.0	42.1	40.9	43.3	48.2	51.7	44.6
SD of Mean Individual Income Percentile Rank	28.7	28.7	27.8	25.9	27.7	23.9	27.1	28.3	25.4
P(Child Household Income in Q5)	25.3%	23.8%	26.9%	5.5%	5.5%	5.6%	12.6%	11.6%	13.5%
P(Child Individual Income in Q5)	23.3%	29.7%	16.5%	10.7%	11.9%	9.5%	15.7%	20.4%	10.9%
P(Child Household Income in p100)	1.3%	1.1%	1.4%	0.1%	0.1%	0.1%	0.3%	0.3%	0.4%
P(Child Individual Income in p100)	1.3%	1.7%	0.8%	0.2%	0.3%	0.2%	0.4%	0.5%	0.3%
Employment Measured in Tax Data	78.6%	81.7%	75.4%	76.2%	71.0%	81.2%	76.9%	77.8%	76.1%
C. Other Outcomes on Full Population									
Marriage Rate	54.7%	51.5%	58.1%	16.3%	16.9%	15.8%	37.3%	35.0%	39.7%
Incarceration Rate	0.9%	1.5%	0.2%	5.1%	10.3%	0.6%	1.5%	2.9%	0.2%
Teen Birth Rate			13.5%			41.3%			29.3%
Spouse Individual Income Percentile Rank	63.2	54.2	72.2	57.4	52.8	62.6	58.2	48.5	67.6
SD of Spouse Individual Income Percentile Rank	26.3	27.3	21.8	25.5	26.2	23.6	26.5	26.8	22.6
Fraction Living in Low Pov. Neighborhood	54.6%	53.9%	55.2%	27.7%	27.4%	27.9%	36.3%	35.6%	37.1%
Fraction Living in Childhood CZ in adulthood	63.2%	63.9%	62.5%	71.7%	71.1%	72.3%	75.0%	75.1%	74.9%
Fraction Living in Childhood Tract in adulthood	19.2%	21.1%	17.1%	22.6%	25.5%	19.8%	24.2%	26.4%	22.0%
Fraction Living with Parents in adulthood	11.6%	13.1%	10.1%	21.0%	23.9%	18.6%	23.0%	24.8%	21.2%
Mean HH Income Rank for Children who stay in CZ	52.8	51.7	54.0	33.0	30.9	34.9	44.9	43.9	45.9
SD of Mean HH Income Rank for Children who stay in CZ	28.1	28.2	27.9	21.6	22.8	20.3	25.4	25.9	24.9
Mean Indiv. Income Rank for Children who stay in CZ	51.3	55.8	46.6	40.8	39.6	41.9	48.0	51.1	44.9
SD of Mean Indiv. Income Rank for Children who stay in CZ	28.0	28.5	26.7	24.9	26.8	22.8	26.5	27.9	24.7
D. ACS Outcomes									
Employment in the ACS	86.6%	91.5%	81.6%	75.1%	70.1%	79.6%	81.5%	85.5%	77.8%
Hours Worked Per Week	32.98	37.67	28.34	26.01	24.88	27.05	29.72	33.02	26.60
SD of Hours Worked Per Week	18.51	17.43	18.38	19.63	20.89	18.34	18.88	18.80	18.43
Median Wages ($/hr)	18.89	19.76	17.71	14.71	14.72	14.57	16.19	16.90	15.69
Mean Wage Percentile Rank	51.5	53.8	49.0	41.2	41.5	40.9	45.6	47.1	44.1
SD of Mean Wage Percentile Rank	28.8	28.8	28.5	27.4	27.8	27.1	27.9	28.3	27.3
Fraction with HS Degree	88.7%	86.7%	90.8%	78.0%	73.0%	82.6%	76.9%	73.4%	80.3%
Fraction with Some 4-Year College Credits	72.7%	67.4%	78.0%	56.7%	47.2%	65.5%	56.4%	50.3%	62.1%
Fraction with Community College Degree	50.4%	44.5%	56.2%	29.0%	21.6%	35.7%	30.2%	25.3%	34.9%
Fraction with 4-Year College Degree	40.1%	35.0%	45.0%	21.0%	15.4%	26.0%	21.3%	17.4%	25.0%
Fraction with Grad Degree	14.6%	11.7%	17.4%	8.0%	4.9%	10.8%	7.0%	5.2%	8.7%
Fraction Receiving Public Assistance	1.9%	1.3%	2.5%	4.6%	2.0%	6.9%	3.1%	1.5%	4.5%
Mean Household Income Rank for Natives	57.0	56.0	58.1	35.5	33.5	37.5	47.0	46.0	47.9
SD of Mean Household Income Rank for Natives	27.8	27.9	27.6	23.2	24.4	21.9	26.8	27.1	26.4
Mean Individual Income Rank for Natives	54.2	59.6	48.6	42.9	41.9	43.8	48.5	52.5	44.8
SD of Mean Individual Income Rank for Natives	28.3	28.1	27.5	25.8	27.7	23.8	27.4	28.4	25.8
Mean Household Income Rank for Immigrants	58.5	56.8	60.4	44.9	42.2	47.5	48.3	47.0	49.6
SD of Mean Household Income Rank for Immigrants	29.4	29.8	28.8	26.5	27.5	25.2	26.0	26.5	25.4
Mean Individual Income Rank for Immigrants	56.8	60.5	52.9	51.1	49.5	52.5	50.8	53.9	47.6
SD of Mean Individual Income Rank for Immigrants	30.2	30.2	29.7	28.2	29.6	26.7	27.1	28.2	25.5
Fraction Not Suppressed in Tract Collapse	0.9987	0.9981	0.9980	0.9585	0.9541	0.9579	0.9420	0.9253	0.9236
Number of Obs in Full Population	13,000,000	6,639,000	6,360,000	2,640,000	1,294,000	1,346,000	2,517,000	1,262,000	1,255,000
Fraction of Total	0.6356	0.3246	0.3110	0.1291	0.0633	0.0658	0.1231	0.0617	0.0614
Number of Obs in ACS Work Sample	1,033,000	514,000	519,000	139,000	66,500	73,000	148,000	72,000	76,000
Number of Obs in ACS Full Sample	2,855,000	1,429,000	1,426,000	433,000	207,000	226,000	443,000	220,000	224,000

Notes: This table replicates Table I, presenting summary statistics by race and gender for children in our primary analysis sample. All racial groups except Hispanics exclude individuals of Hispanic ethnicity.

Online Appendix Table Ib

Summary Statistics for Primary Analysis Sample, by Race and Ethnicity (1978-1983 Birth Cohorts)

	Asian			American Indian and Alaska Native		
	Pooled	Male	Female	Pooled	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
A. Parental Characteristics						
Median Parent Household Income	53,350	52,680	54,040	36,710	36,820	36,610
Mean Parent Household Income Percentile Rank	49.4	49.1	49.8	38.2	38.3	38.1
SD of Mean Parent Household Income Percentile Rank	30.7	30.7	30.7	25.7	25.8	25.7
Father Presence	88.5%	88.6%	88.3%	71.1%	72.2%	70.1%
Mother Presence	92.1%	91.8%	92.5%	88.5%	87.8%	89.1%
Two Parent Family	80.6%	80.4%	80.8%	59.6%	60.0%	59.2%
B. Income and Employment Outcomes in Adulthood						
Median Household Income ($)	63,850	56,660	72,050	23,490	22,320	24,550
Mean Household Income Percentile Rank	60.7	57.6	64.0	37.8	36.7	38.8
SD of Mean Household Income Percentile Rank	29.6	30.0	28.8	26.2	26.3	26.0
Median Individual Income ($)	43,790	45,640	41,860	17,440	20,370	15,270
Mean Individual Income Percentile Rank	60.4	61.6	59.1	40.4	43.0	37.7
SD of Mean Individual Income Percentile Rank	30.4	30.4	30.3	26.6	28.3	24.5
P(Child Household Income in Q5)	34.6%	30.4%	38.9%	8.6%	8.0%	9.3%
P(Child Individual Income in Q5)	35.8%	37.5%	33.9%	10.6%	13.8%	7.2%
P(Child Household Income in p100)	3.1%	2.5%	3.8%	0.3%	0.2%	0.3%
P(Child Individual Income in p100)	3.0%	3.6%	2.4%	0.3%	0.4%	0.2%
Employment Measured in Tax Data	79.6%	80.5%	78.6%	70.3%	70.5%	70.0%
C. Other Outcomes on Full Population						
Marriage Rate	50.0%	45.4%	54.7%	32.3%	30.2%	34.3%
Incarceration Rate	0.3%	0.5%	0.0%	2.9%	5.0%	0.8%
Teen Birth Rate			6.8%			31.4%
Spouse Individual Income Percentile Rank	67.4	58.8	75.4	55.6	47.4	63.8
SD of Spouse Individual Income Percentile Rank	28.4	30.1	24.0	25.5	25.3	22.9
Fraction Living in Low Pov. Neighborhood	60.7%	59.3%	62.1%	29.2%	29.2%	29.2%
Fraction Living in Childhood CZ in adulthood	66.5%	69.2%	63.8%	65.7%	66.3%	65.2%
Fraction Living in Childhood Tract in adulthood	22.4%	26.3%	18.4%	27.1%	29.0%	25.2%
Fraction Living with Parents in adulthood	27.1%	31.8%	22.4%	20.3%	22.7%	18.0%
Mean HH Income Rank for Children who stay in CZ	57.7	54.6	61.1	36.9	36.1	37.7
SD of Mean HH Income Rank for Children who stay in CZ	29.3	29.5	28.7	25.3	25.6	25.1
Mean Individual Income Rank for Children who stay in CZ	58.5	59.0	58.0	40.0	42.5	37.5
SD of Mean Individual Income Rank for Children who stay in CZ	29.8	30.1	29.5	26.0	27.7	23.8
D. ACS Outcomes						
Employment in the ACS	88.2%	90.5%	85.9%	73.6%	77.0%	70.2%
Hours Worked Per Week	34.14	36.31	31.95	24.99	27.15	22.84
SD of Hours Worked Per Week	17.96	17.46	18.19	20.03	20.63	19.16
Median Wages ($/hr)	23.96	23.54	24.43	13.96	14.71	13.26
Mean Wage Percentile Rank	61.2	60.7	61.8	39.3	41.2	37.2
SD of Mean Wage Percentile Rank	29.9	30.1	29.6	27.0	28.0	25.6
Fraction with HS Degree	91.4%	90.1%	92.8%	77.3%	74.5%	80.1%
Fraction with Some 4-Year College Credits	84.6%	81.6%	87.7%	51.2%	44.2%	58.2%
Fraction with Community College Degree	67.3%	62.4%	72.3%	22.8%	18.4%	27.2%
Fraction with 4-Year College Degree	58.9%	53.9%	64.1%	14.6%	12.0%	17.3%
Fraction with Grad Degree	23.4%	19.9%	26.9%	4.2%	3.2%	5.2%
Fraction Receiving Public Assistance	1.1%	1.0%	1.3%	4.8%	2.6%	7.0%
Mean Household Income Rank for Natives	58.2	55.8	60.4	38.5	37.5	39.5
SD of Mean Household Income Rank for Natives	28.8	29.2	28.3	26.2	26.4	25.9
Mean Individual Income Rank for Natives	57.5	60.5	54.8	41.0	43.6	38.3
SD of Mean Individual Income Rank for Natives	29.4	29.4	29.1	26.4	28.0	24.4
Mean Household Income Rank for Immigrants	63.3	60.1	66.7	42.2	40.9	43.6
SD of Mean Household Income Rank for Immigrants	29.3	29.9	28.3	28.0	28.7	27.2
Mean Individual Income Rank for Immigrants	63.1	63.9	62.2	44.6	46.4	42.5
SD of Mean Individual Income Rank for Immigrants	30.2	30.2	30.1	28.4	29.7	26.7
Fraction Not Suppressed in Tract Collapse	0.7781	0.7290	0.7069	0.4897	0.4426	0.4591
Number of Obs in Full Population	673,000	344,000	330,000	134,000	68,000	66,000
Fraction of Total	0.0329	0.0168	0.0161	0.0066	0.0033	0.0032
Number of Obs in ACS Work Sample	46,000	23,000	23,000	12,000	6,100	6,100
Number of Obs in ACS Full Sample	128,000	65,000	63,000	31,000	15,500	15,500

Notes: This table replicates Table I, presenting summary statistics by race and gender for children in our primary analysis sample. All racial groups except Hispanics exclude individuals of Hispanic ethnicity.

Online Appendix Table II

Explanatory Power of Historical Mobility Estimates vs. Tract-Level Covariates: OLS Regression Estimates

	Dep Var: Upward Mobility for 1989 Cohort		
	(1)	(2)	(3)
Share of Variance by Specification			
Upward Mobility for 1979 Cohort	X	X	X
Share Single Parent (1990)		X	X
Share in Poverty (1990)		X	X
Share Single Parent (2010)			X
Share in Poverty (2010)			X
Adjusted R-squared	49.04%	60.19%	63.48%

Notes: This table presents the adjusted R-squared from regressions of upward mobility for the 1989 birth cohort (the mean child household income rank at age 26 given parents at the 25th percentile of income distribution) on upward mobility for the 1979 birth cohort and tract level characteristics measured in various years. We correct for sampling and noise-infusion error by rescaling the adjusted R-squared by the reliability ratio – the ratio of the signal variance to total variance of upward mobility for the 1989 cohort.

Online Appendix Table III

Quasi-Experimental Exposure Effect Estimates by Distance of Move and Marital Status

	Moves <= 25 miles			Moves > 25 miles		
	Single Dad	Single Mom	Two Parent	Single Dad	Single Mom	Two Parent
	(1)	(2)	(3)	(4)	(5)	(6)
Age <= 23	-0.006	-0.007	-0.022	-0.027	-0.022	-0.030
	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)
Observations	931,000	1,705,000	2,329,000	287,000	417,000	898,000

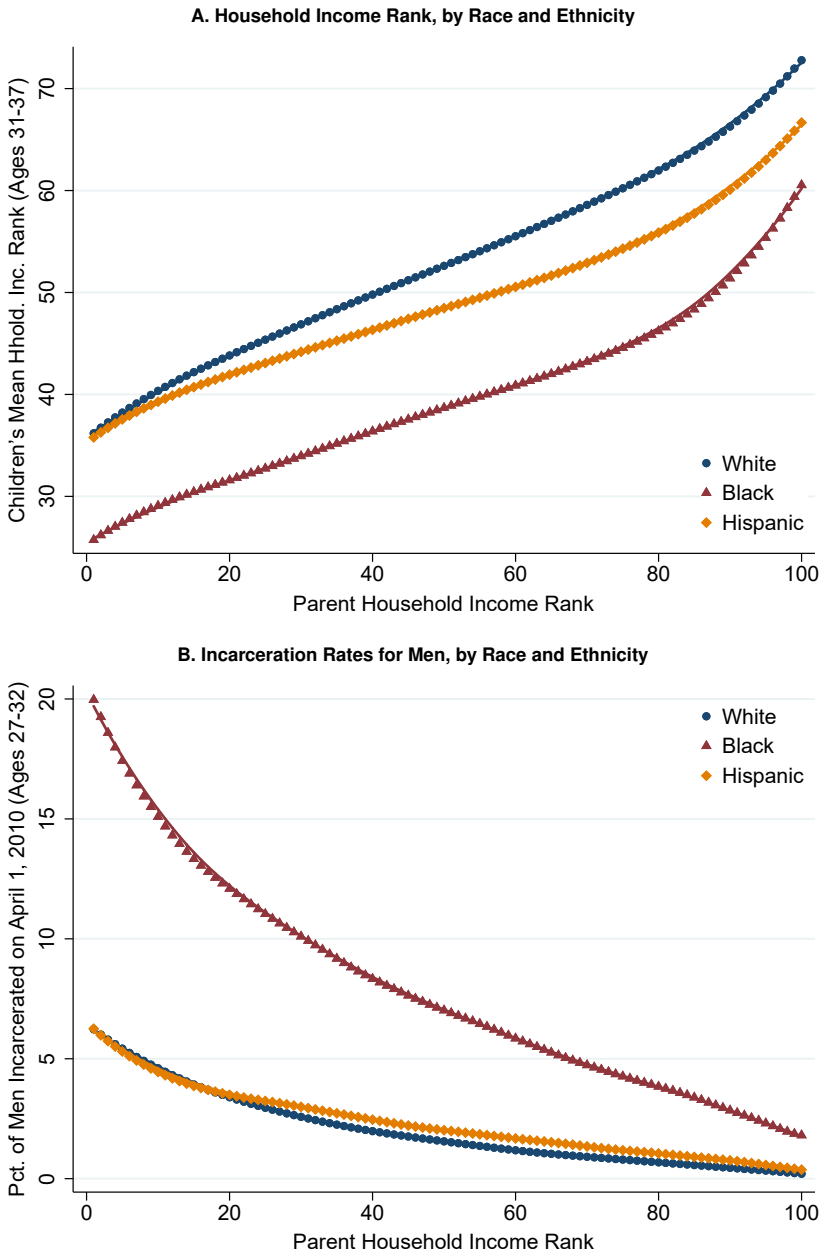
Notes: This table reports estimates of annual childhood tract-level exposure effects on children's family income ranks at age 24 for different subgroups of one-time movers. Columns 1-3 show exposure effects for moves between tracts that are less than 25 miles apart; columns 4-6 show estimates for moves between tracts that are more than 25 miles apart. Columns 1 and 4 restrict to movers with a single invariant claimer (in the tax records) who is male, columns 2 and 5 restrict to movers with a single invariant claimer who is female, and columns 3 and 6 restrict to movers with two invariant claimers who are married in every available tax year. See notes to Table IV for details on these specifications.

Online Appendix Table IV
Summary Statistics for Tract Movers Analysis Samples

		1-time Movers	Non 1-time Movers (0 & 2+ Movers)
Parent Family Income Rank	Mean	56.9	48.6
	Std. Dev.	29.2	28.7
	Num. of Obs.	3,100,000	42,000,000
Child Individual Income Rank at 24	Mean	51.5	49.5
	Std. Dev.	29.1	28.8
	Num. of Obs.	2,400,000	34,000,000
Child Household Income Rank at 24	Mean	51.8	49.6
	Std. Dev.	29.1	28.8
	Num. of Obs.	3,100,000	42,000,000
Child Incarcerated in 2010	Mean	0.9%	1.4%
	Std. Dev.	9.4%	11.8%
	Num. of Obs.	2,500,000	33,000,000
Child Married at 30	Mean	42.6%	38.1%
	Std. Dev.	49.4%	48.6%
	Num. of Obs.	1,800,000	22,000,000

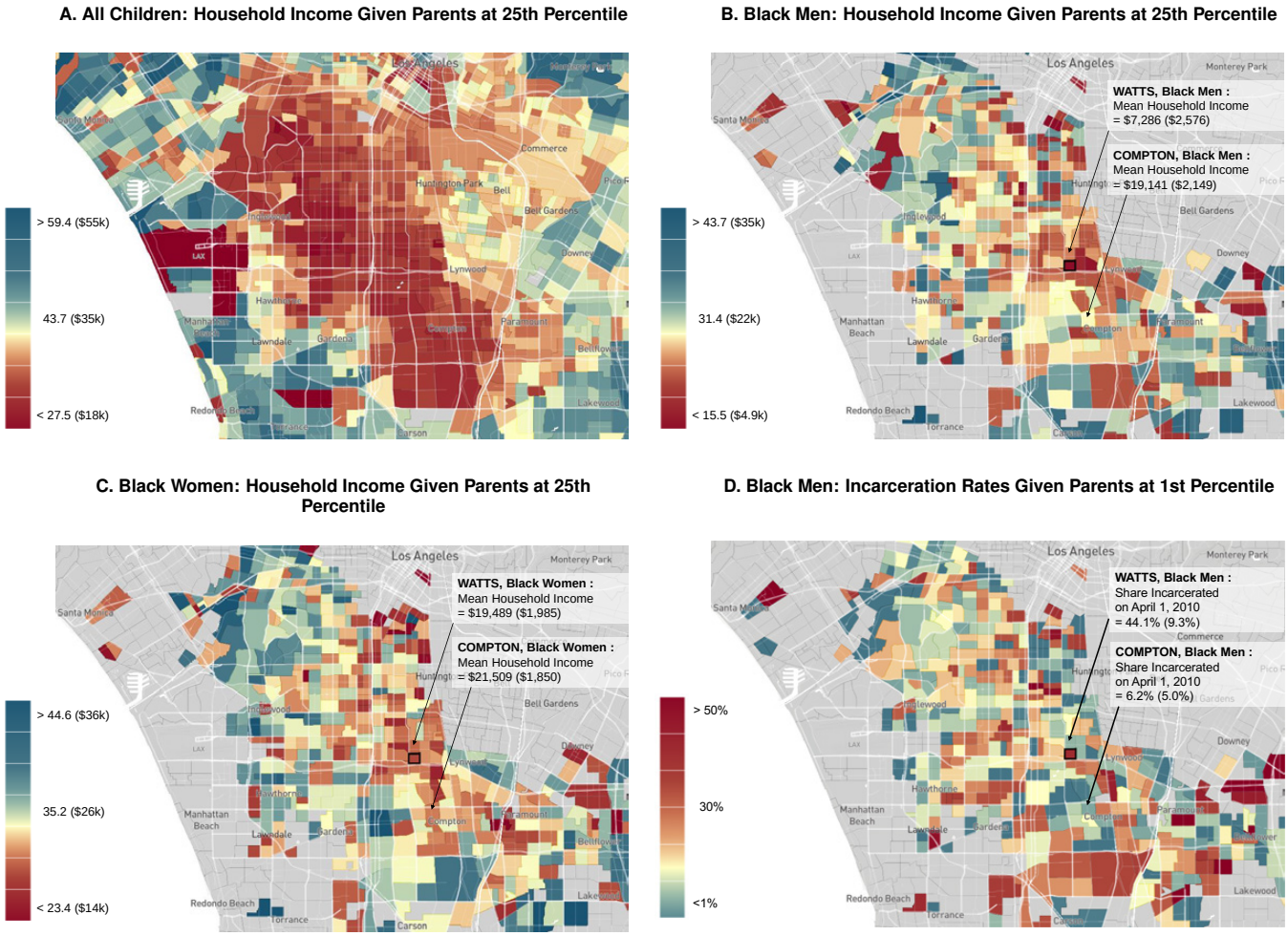
Notes: This table presents summary statistics for the samples used in our analyses of tract-level exposure effects. The full analysis sample extends the core sample described in Section III by including additional cohorts up until 1991 in order to observe moves at younger ages. Column (1) reports summary statistics for children whose parents moved across tracts exactly once throughout our sample window (1989-2015), are observed in their destination for at least two years, and moved at least 25 miles (based on their tract centroids). Column (2) reports summary statistics for children whose parents do not move across tracts throughout our sample window and for children whose parents move more than once across tracts. Parent family income is the average pre-tax household income from 1994-2000 measured as AGI. Child individual income is defined as the sum of individual W-2 wage earnings and half of household self-employment income. Incarceration is based on the individual's group home status in the 2010 US population census. Marital status is defined based on the marital status listed on 1040 forms for tax filers; non-filers are coded as single. See Section III for further details on variable and sample definitions.

FIGURE I: Children's Outcomes vs. Parental Income Rank



Notes: This figure plots the relationship between children's outcomes in adulthood and the income of their parents for non-Hispanic black children, non-Hispanic white children, and Hispanic children in our primary analysis sample (1978-83 birth cohorts). Panel A plots children's mean household income ranks in adulthood vs. their parents' income percentile. In each series, each point represents the mean income rank of children with parents in a single income percentile. Child income is the mean of 2014-2015 household income (when the child is between 31-37 years old), while parent income is mean household income from 1994-1995 and 1998-2000. Children are assigned percentile ranks relative to all other children in their birth cohort, while parents are ranked relative to all parents with children in the same birth cohort. Panel B replicates Panel A, replacing the outcome with an indicator for being incarcerated on April 1, 2010, as recorded on the 2010 Decennial Census Short Form, and focusing solely on male children. Incarceration is defined living in a federal detention center, federal prison, state prison, local jail, residential correctional facility, military jail, or juvenile correctional facility. For each series, we plot curves showing the lowess fit (with a bandwidth of 0.3) that we use as our estimate of the conditional expectation function $f_{rg}(p)$ discussed in Section III.

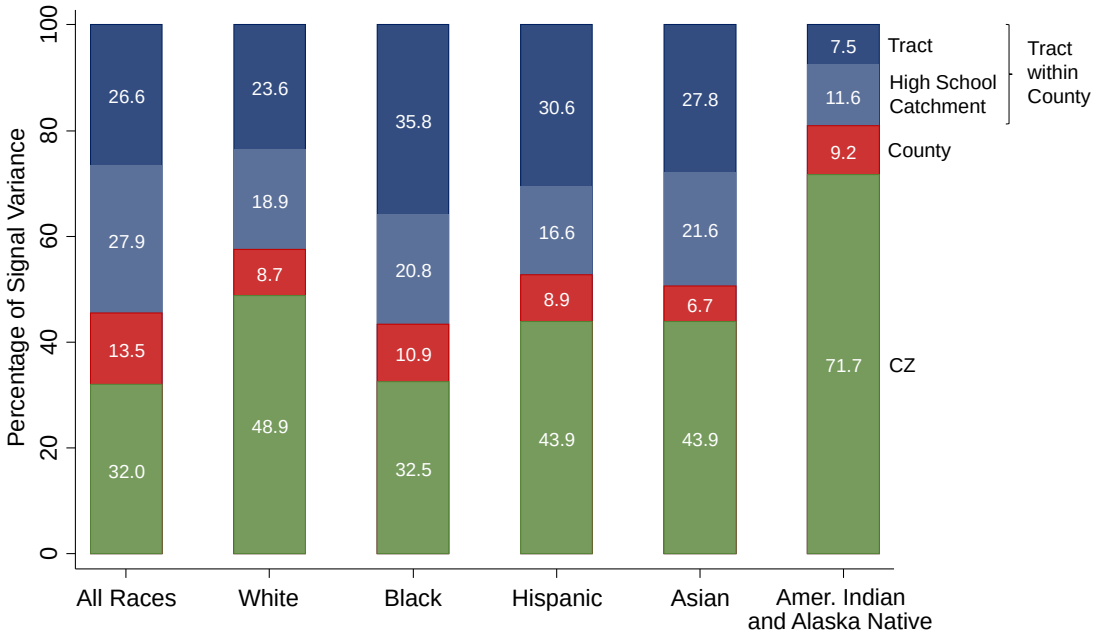
FIGURE II: Children's Outcomes in Adulthood, by Census Tract in Los Angeles



These maps must be printed in color to be interpretable.

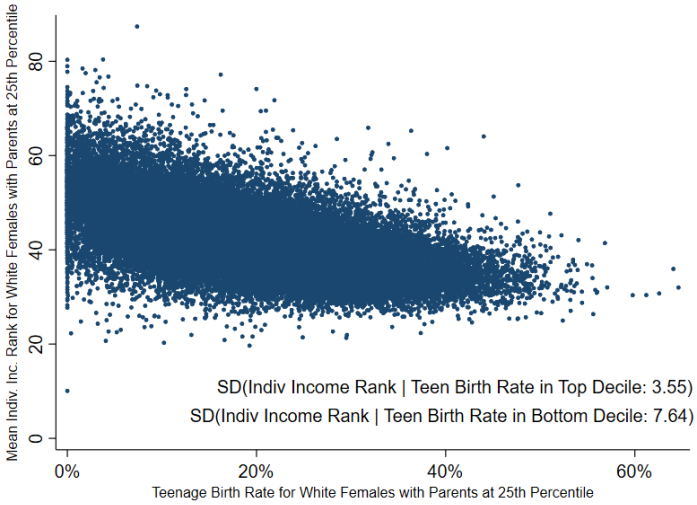
Notes: These maps display mean outcomes in adulthood of children who grew up in the Los Angeles metro area, by the tract in which they grew up. Panel A plots our estimates of mean household income ranks for children whose parents were at the 25th percentile of the national household income distribution (an income of approximately $27,000) using our primary analysis sample (1978-83 birth cohorts), which we hereafter refer to as "upward mobility." Upward mobility is estimated separately in each tract using linear regressions of children's income ranks on a tract-invariant transformation of parent income rank $f_{rg}(p)$ that is estimated at the national race-by-gender level using a lowess fit, as shown in Figure I. We weight each child by the number of years they lived in each tract before the age of 23 when estimating these regressions. Finally, we add independent Gaussian noise to the resulting tract-level estimates to protect privacy; the standard deviation of this noise is typically less than one-tenth of the standard error due to sampling variation. To facilitate interpretation, we report both mean income ranks and, in parentheses, the dollar values corresponding to those ranks based on the income distribution of children in 2015 in the legend. Panels B and C replicate Panel A, limiting the sample to non-Hispanic black male and female children, respectively. Panel D replicates Panel B for black men with parents at the 1st percentile, using an indicator for being incarcerated on April 1, 2010 as the outcome. In each panel, we report point estimates and standard errors (in parentheses) for selected tracts. The standard errors reported include the noise added to protect privacy. Tracts shown in gray are areas with no estimate due to insufficient data (fewer than 20 observations in the race-by-gender cell). See notes to Figure I for definitions of income and incarceration.

FIGURE III: Geographic Decomposition of Variance in Upward Mobility



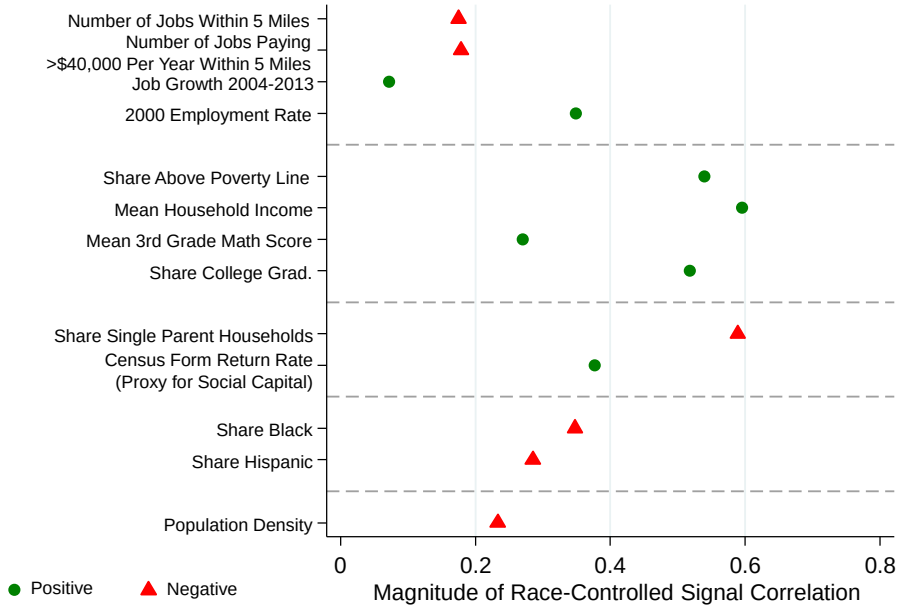
Notes: This figure presents a geographical variance decomposition of the tract-level estimates of upward mobility (children's mean household income ranks given parents at the 25th percentile), which are constructed as described in the notes to Figure II. We estimate the share of variance explained by each level of geography as the adjusted R-squared in a regression of the tract-level estimates on fixed effects for different levels of nested geographies, weighted by the number of children in each tract whose parents earn less than the national median income. We correct for sampling and noise-infusion error by rescaling the adjusted R-squared by the reliability ratio – the ratio of the signal variance to total variance of the tract-level estimates reported in Table II. We plot the share of signal variance explained by CZ fixed effects, county fixed effects, high school catchment area fixed effects, and the residual (attributed to tract-within-school catchment area). Tracts are not perfectly nested within catchment areas; we create an approximate crosswalk by assigning tracts to the school catchment area that contains the majority of their land area, as discussed in Online Appendix A.

FIGURE IV: Upward Mobility vs. Teenage Birth Rates for White Women



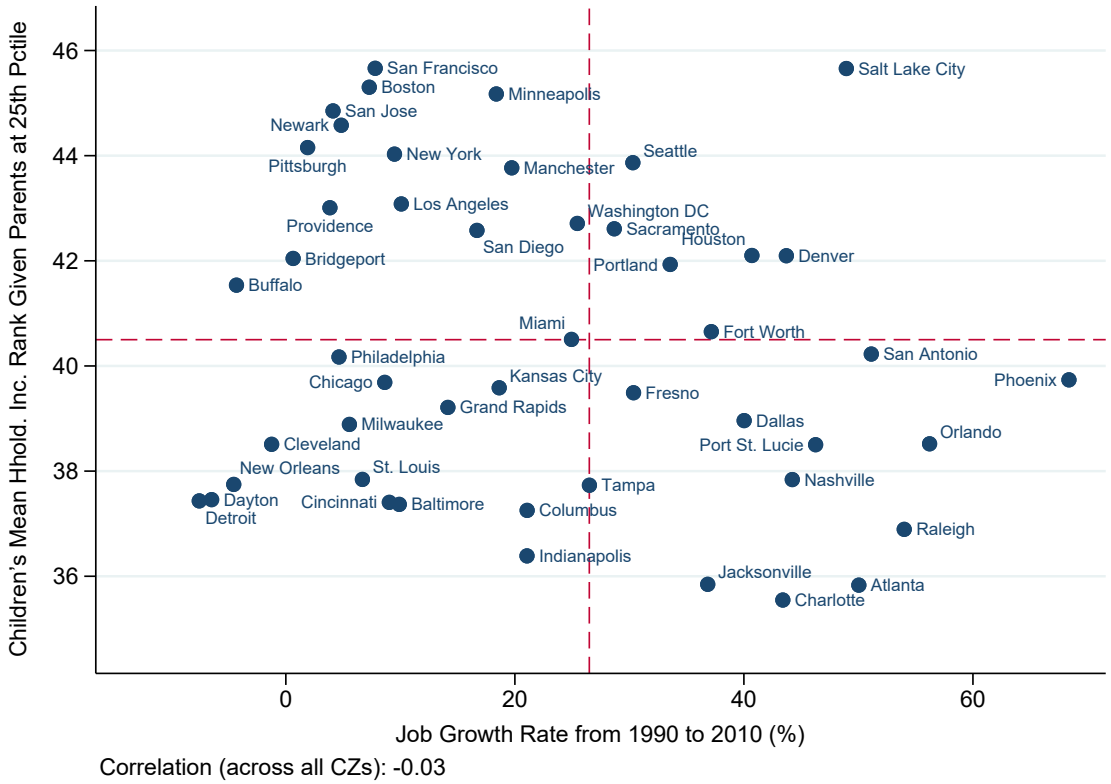
Notes: This figure presents scatter plots of mean individual income ranks vs. the teenage birth rate for white women with parents at the 25th percentile, by Census tract. Mean individual ranks are estimated as described in the notes to Figure II. Teenage birth is an indicator for ever claiming a dependent on a tax return who was born while the claimer was between ages 13 and 19. We limit the sample to tracts in which there are at least 100 observations for white women and bottom-code tracts with negative teenage birth rates to zero (negative values arise due to the addition of noise to the estimates). The standard deviations of mean income ranks reported conditional on having teenage birth rates in the bottom or top decile of the distribution are weighted by the number of white women in each tract whose parents earn less than the national median. We omit one tract in Canton, Michigan for scaling purposes; the x and y coordinates for this tract are (90%, 65.35).

FIGURE V: Tract-Level Correlations Between Neighborhood Characteristics and Upward Mobility



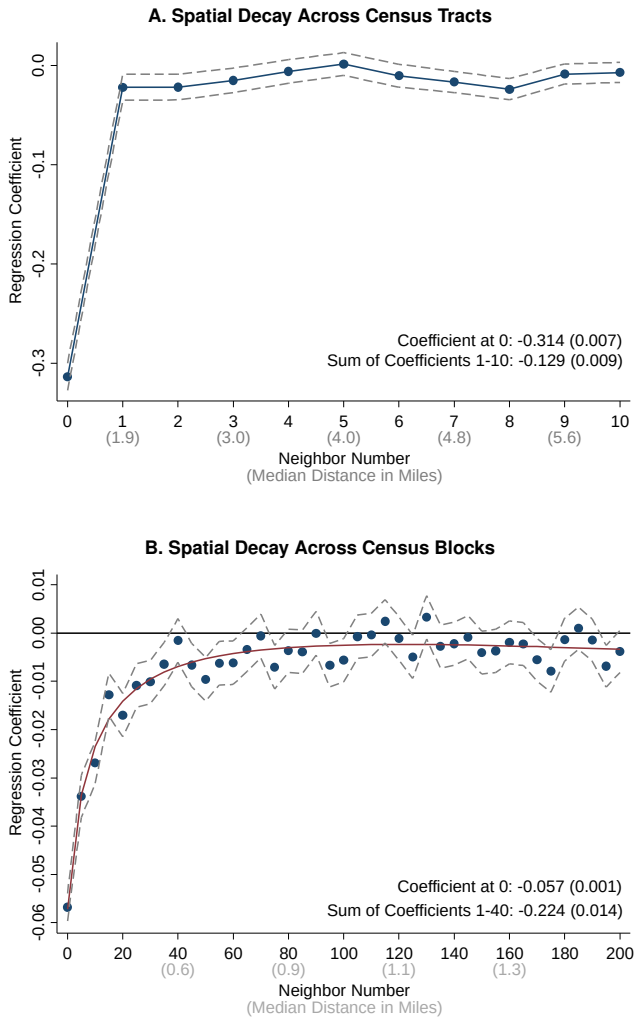
Notes: This figure plots univariate, race-controlled correlations between various tract-level characteristics and our estimates of upward mobility in each tract (children's mean household income ranks given parents at the 25th percentile, constructed as described in the notes to Figure II). The correlations are weighted by the number of children in each tract whose parents earn less than the national median income and are estimated using tract-within-CZ variation by demeaning all variables by CZ prior to estimating the correlations. We control for race when estimating each correlation coefficient by first estimating five separate correlations for each race (non-Hispanic Asians, non-Hispanic Blacks, non-Hispanic whites, American Indians, and Hispanics) and then taking a mean of the five correlations, weighting each of the five groups by its national population share in the 2000 Decennial Census. We estimate "signal" correlations that adjust for attenuation due to sampling error and noise infusion in our upward mobility estimates by rescaling the raw correlations by the square root of the reliability ratio, which is one minus the ratio of the noise variance (estimated as the mean standard error squared) to the total within-CZ variance of the upward mobility estimates. Red triangles denote negative correlations, while green circles denote positive correlations. See Online Appendix A for definitions of each of the characteristics.

FIGURE VI: Upward Mobility vs. Job Growth in the 50 Largest CZs



Notes: This figure presents a scatter plot of upward mobility in each CZ vs. the rate of job growth between 1990 and 2010 in the 50 largest CZs based on their populations in 2000. Upward mobility is constructed as described in the notes to Figure II. Job growth rates are defined as the percentage change in employment in each CZ using data from the Local Area Unemployment Statistics from the Bureau of Labor Statistics. We omit Las Vegas and Austin from the figure for scaling purposes as they have exceptionally high growth rates; the x and y coordinates for these CZs are: Las Vegas (107.7, 38.9) and Austin (87.9, 40.3). We also report the population-weighted correlation across all CZs (not just the top 50 CZs) as a reference.

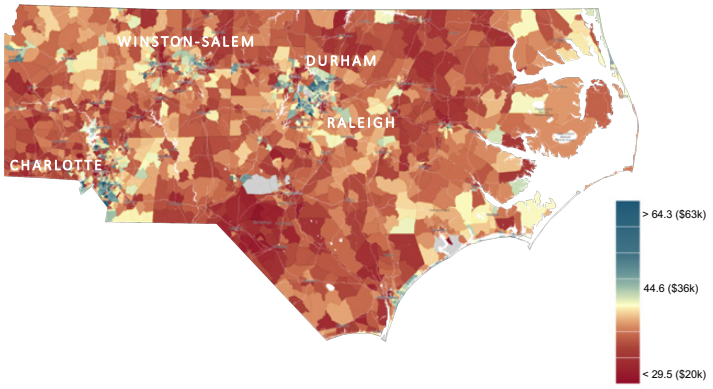
FIGURE VII: Spatial Decay of Correlation Between Upward Mobility for Whites and Poverty Rates



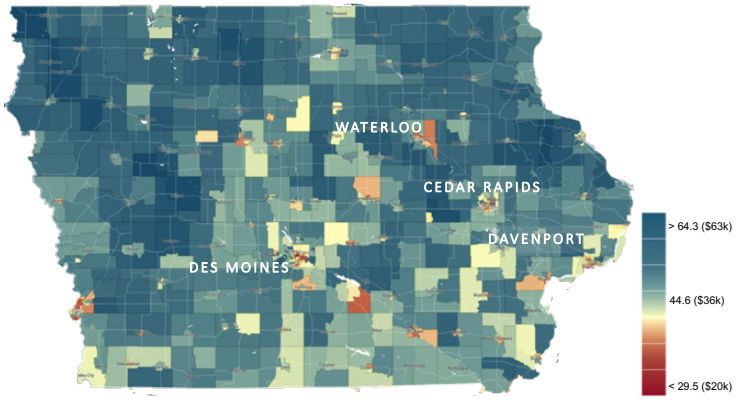
Notes: This figure plots the spatial decay of the relationship between upward mobility for whites and poverty rates at two different levels of geography: Census tracts (Panel A) and Census blocks (Panel B). Upward mobility refers to children's mean household income ranks given parents at the 25th percentile, constructed as described in the notes to Figure II. Tract-level poverty rates are obtained from the publicly available 2000 Decennial Census. Block-level poverty rates are estimated using tax records as the share of families whose total income (wages, social security income, dividends, interest income, and schedule E gains or losses) falls below the poverty line in 2010. To construct Panel A, we first standardize both the upward mobility and poverty rate measures, weighting by the number of children whose parents earn less than the national median. We then regress upward mobility on poverty rates in the same tract and the ten nearest neighbors (defined by the minimum cardinal distance between centroids) and plot the coefficients. To construct Panel B, we regress the household income rank of white children whose parents are between the 20th and 30th percentiles of the income distribution on block-level poverty rates for their own block and the 200 nearest blocks, binned into groups of 5. 95% confidence intervals for the estimates are shown by the dashed lines. In both panels, we also report the median distance between the own-tract (or block) and neighboring tracts (or blocks) in each of the bins as a reference.

FIGURE VIII: Heterogeneity in the Relationship between Upward Mobility and Population Density

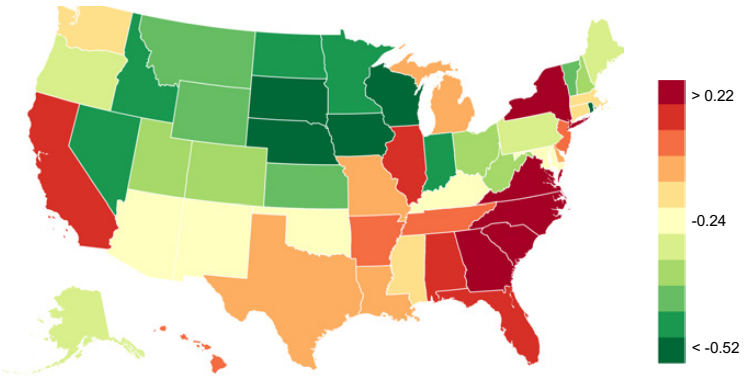
A. Mean Household Income Rank of White Children with Parents at 25th Percentile, North Carolina



B. Mean Household Income Rank of White Children with Parents at 25th Percentile, Iowa



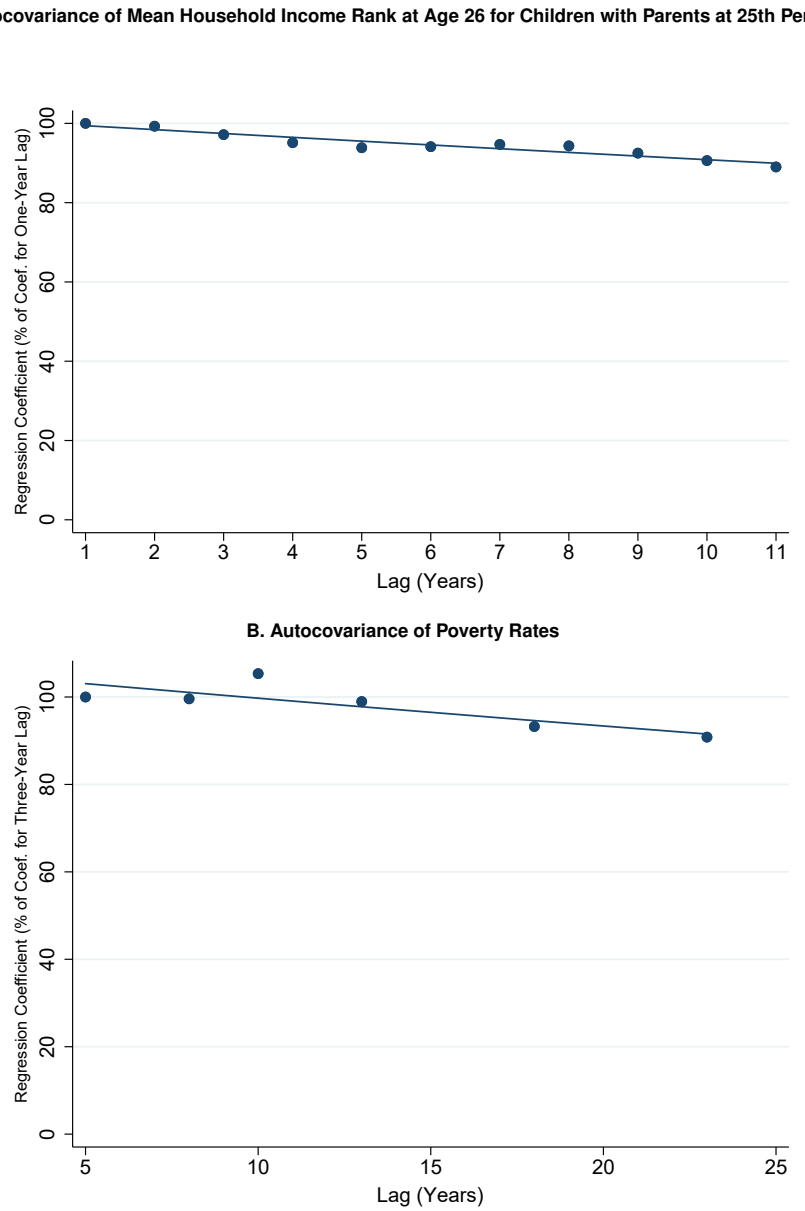
C. Correlations between Population Density and Upward Mobility for White Children, by State



These maps must be printed in color to be interpretable

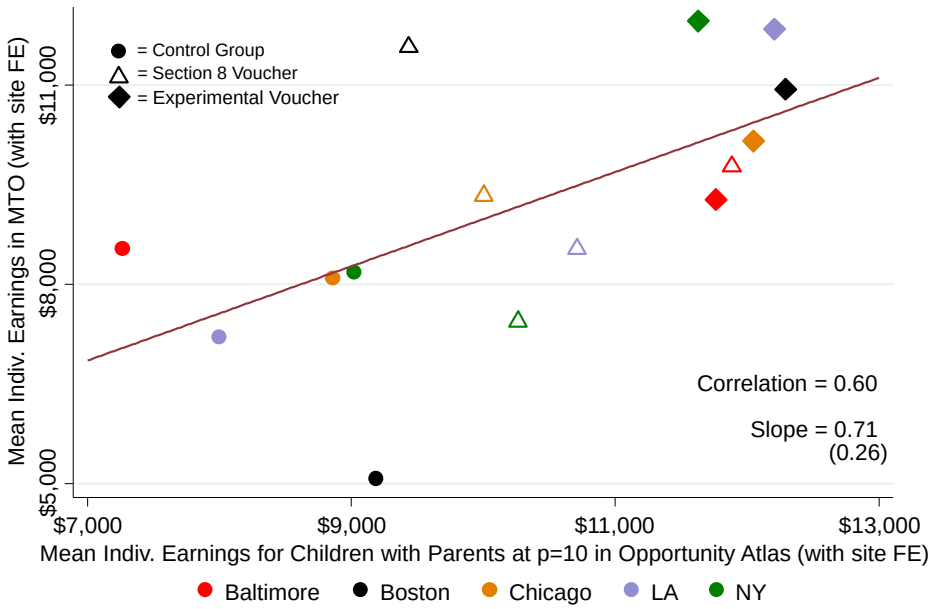
Notes: This figure analyzes the relationship between upward mobility (constructed as described in the notes to Figure II) and population density across the U.S. Panels A and B replicate Figure IIa, for white children in North Carolina (Panel A) and Iowa (Panel B). Panel C shows the signal correlation between upward mobility for white children and population density (measured using the 2000 Decennial Census) within each state, weighted by the number of children in each tract whose parents earn less than the national median. We estimate "signal" correlations that adjust for attenuation due to sampling error and noise infusion in our upward mobility estimates by rescaling the raw correlations by the square root of the reliability ratio, which is one minus the ratio of the noise variance (estimated as the mean standard error squared) to the total variance of the upward mobility estimates.

FIGURE IX: Changes in Tract-Level Outcomes and Characteristics Over Time



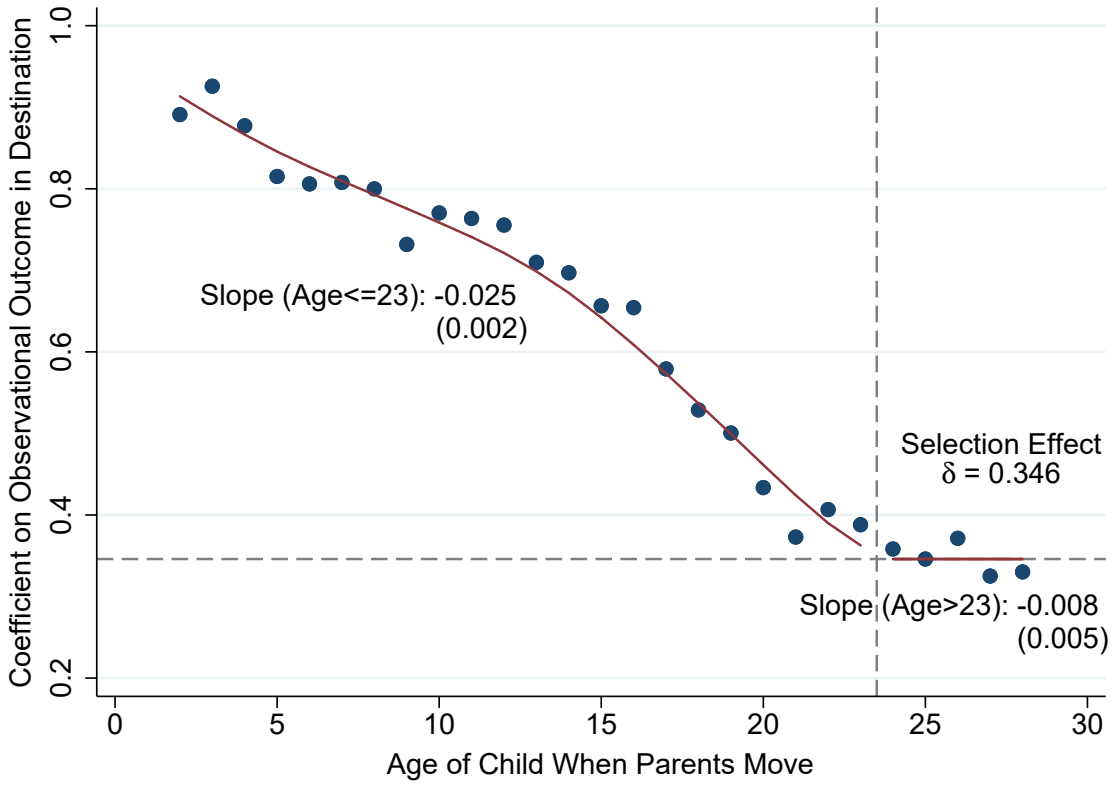
Notes: This figure examines how upward mobility and neighborhood characteristics change over time. Panel A shows the rate of decay in one's ability to forecast future cohorts' outcomes using historical data. It plots the coefficients from regressions of tract-level estimates of upward mobility for a given cohort t (constructed as described in the notes to Figure II) on estimates of upward mobility from a different birth cohort $t \pm x$, varying x from 1 to 11. We normalize the estimates by the coefficient of the regression with the one year lag/lead so that the estimates that are plotted can be interpreted as the percentage decay in the forecast coefficient. We extend our primary analysis sample to children born in the 1978-89 birth cohorts and measure children's incomes at age 26 in this figure in order to estimate as many lags as possible. To maximize precision, we use all available cohorts to estimate each covariance; for instance, the covariance at a lag of 1 is estimated using 11 pairs of cohorts. Panel B plots the autocovariance of tract-level poverty rates using publicly available data from the 1990, and 2000 Decennial Census and ACS data collected between 2006 and 2010 and between 2011 and 2015, which we pool to obtain an estimate for 2008 and 2013, respectively. This figure is constructed in the same way as Panel A, estimating the relationship between poverty rates at lags and leads of 5, 8, 10, 13, 18, and 23 years. See Online Appendix A for definition of poverty rates.

FIGURE X: Experimental Estimates of Earnings from Moving to Opportunity Experiment vs. Observational Estimates from Opportunity Atlas



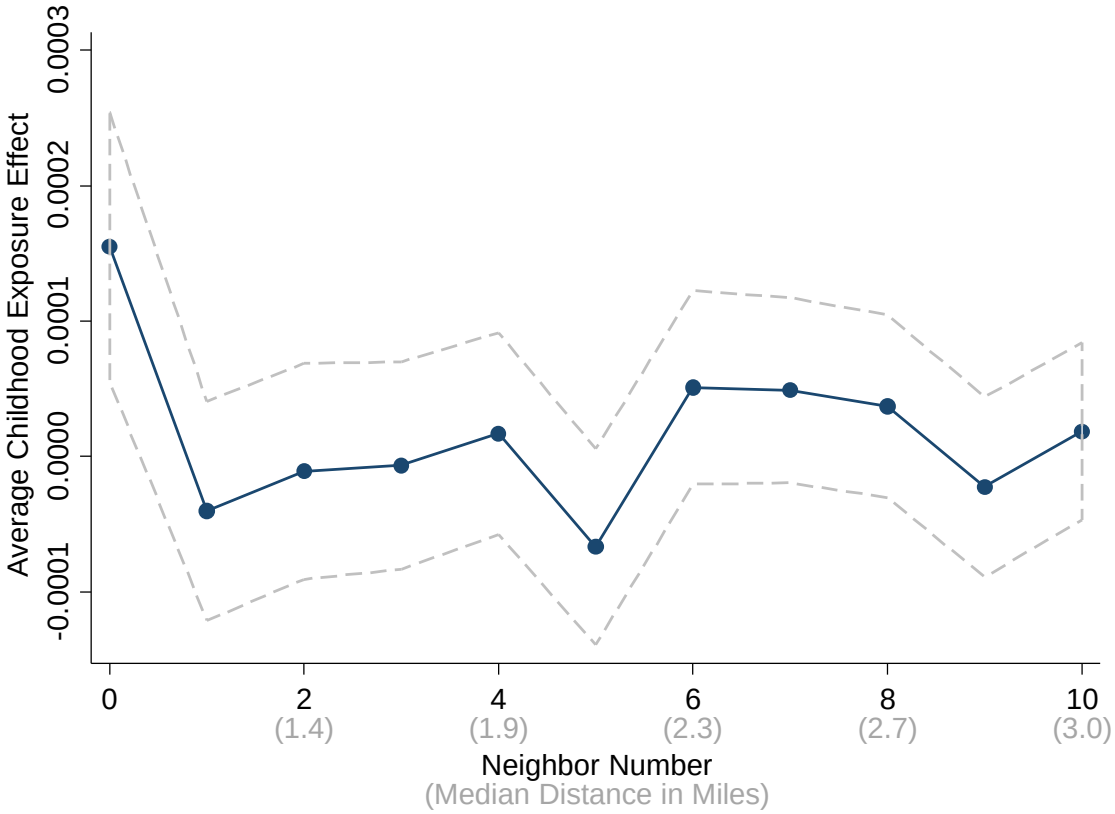
Notes: This figure plots estimates of children's earnings in adulthood from the Moving to Opportunity (MTO) experiment vs. children's mean observed earnings in adulthood in the Opportunity Atlas. The y-axis plots fifteen MTO estimates of earnings outcomes for children who were younger than 13 at the time of the experiment, for each of the five cities (sites) where MTO was conducted and for each of the three treatment arms (Control group, Section 8 Voucher group, and Experimental Voucher group). To construct the values, we start from the ITT estimates reported in Chetty, Hendren, and Katz (2016, Online Appendix Table 7, Panel B). We then construct implied treatment-on-the-treated (TOT) values for the Section 8 and Experimental groups as the mean observed earnings for the control group in the relevant site plus the site-specific ITT estimate for each treatment arm divided by the voucher takeup rate in that arm. To eliminate non-experimental variation across sites, we demean each set of estimates within site, and then add the mean income value observed for those in the MTO control group in Chicago (thereby normalizing estimates to observed earnings levels in Chicago). The x-axis plots observational estimates from the Opportunity Atlas of children's mean earnings in adulthood conditional on having low-income parents for the neighborhoods corresponding to those where children in each of the MTO groups grew up. To construct these estimates, we first identify these neighborhoods by mapping the neighborhood names listed in Online Appendix Table 1c of Chetty, Hendren, and Katz (2016) to Census tracts. We then take a population-weighted mean of children's predicted individual income ranks in adulthood across the relevant Census tracts, conditional on having parents at the 10th percentile of the income distribution (approximately the average income of parents in the MTO sample). Finally, we translate these mean ranks to dollar values at age 26 (the average age at which children's earnings were measured in the MTO sample) using a crosswalk from ranks to dollars in 2015. As with the MTO estimates, we demean children's incomes within site and add back the estimate for the mean over the set of tracts we use for the control group in Chicago. The best-fit line and slope estimates are based on an unweighted regression of the MTO estimates on the Opportunity Atlas estimates. The figure reports both the regression coefficient (with standard error in parentheses) and the corresponding correlation coefficient.

FIGURE XI: Childhood Exposure Effects: Quasi-Experimental Estimates Using Movers



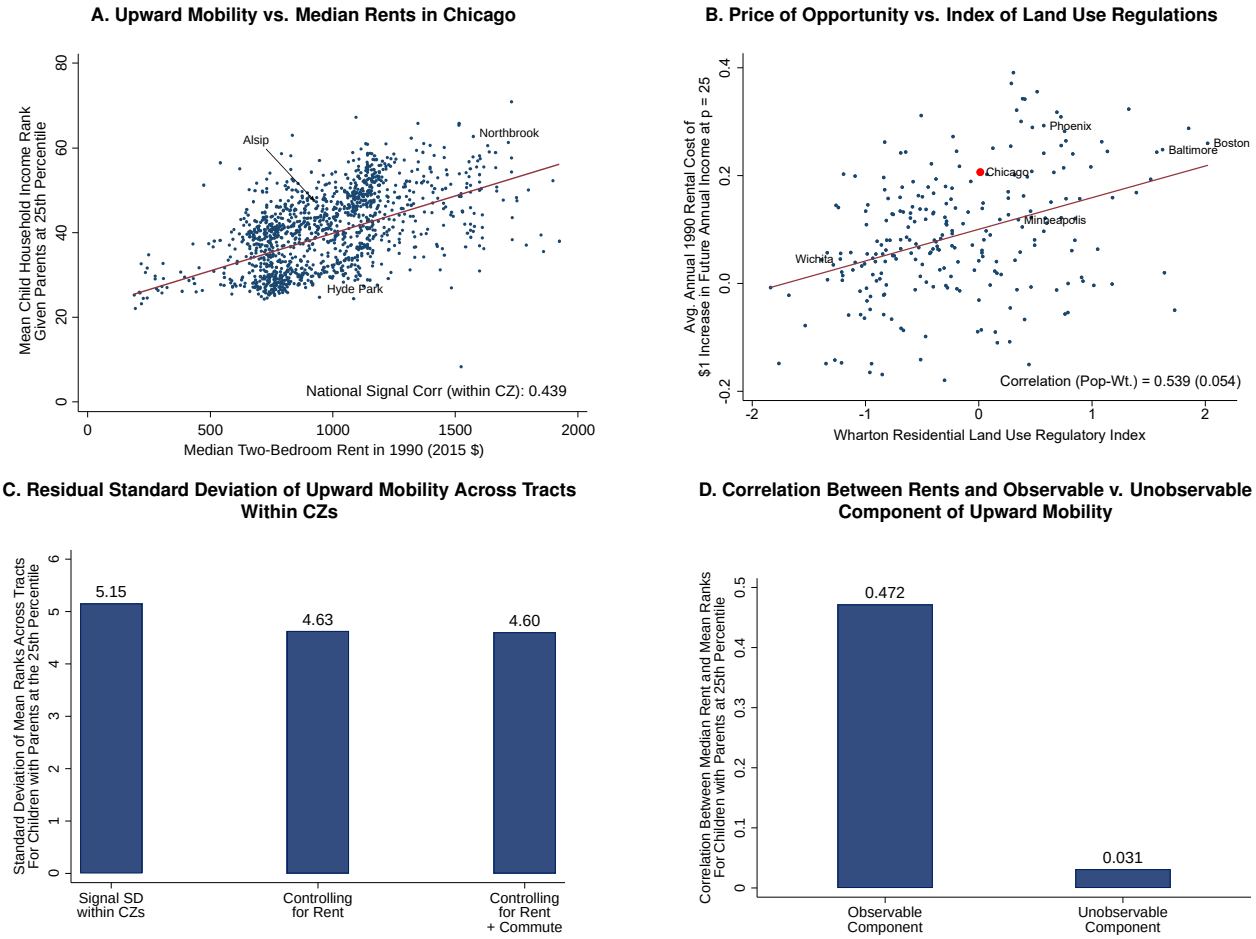
Notes: This figure plots the effect of moving to a tract where children have one percentile point higher household income ranks in adulthood, by the age at which children move. To construct the figure, we first estimate mean observed outcomes in each tract following the methodology described in Figure II, except that we (1) pool data from the 1978-91 birth cohorts and measure income at age 24 and (2) exclude all children who move exactly once between 1989-2015. We extend our primary sample to the 1978-91 birth cohorts for this analysis in order to observe moves at earlier ages and exclude one-time movers to avoid having the same observations on the left- and right-hand side of the regression specifications we use in what follows. We then take the set of children who move exactly once between two tracts that are at least 25 miles apart, and regress their household income ranks at age 24 on the difference in the observational predictions between their destination and origin tracts (at the relevant parental income percentile) interacted with indicators for their ages at move as well as the other controls specified in equation (8). The figure plots the resulting regression coefficients (b_m) vs. children's ages at move (m), along with a lowess fit to these points below age 23. We also report linear slopes and standard errors using unweighted OLS regressions of b_m on m, separately for moves at or below age 23 and above age 23. The parameter δ – defined as the mean value of the age-of-move-specific coefficients for moves older than age 23 – represents a selection effect because moves after age 24 cannot affect income measured at age 24. The dashed horizontal line shows the value of the selection effect δ; the identification assumption underlying the analysis is that the selection effect δ does not vary with the child's age at move m. Under this assumption, the magnitude of the slope for moves below age 23 represents an estimate of the average annual causal childhood exposure effect.

FIGURE XII: Predictive Power of Poverty Rates in Actual Destination vs. Neighboring Tracts



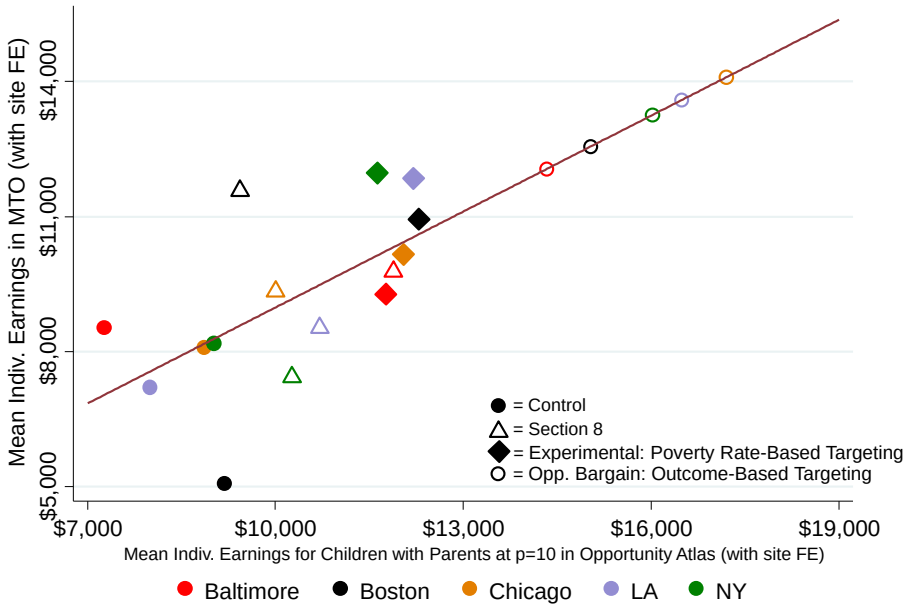
Notes: This figure plots coefficients from a regression that identifies childhood exposure effects using a specification analogous to that in Column 2 of Table IV, which is estimated on the sample of one-time movers who moved at least 25 miles. In the specification in Table IV, we regress children's household income ranks at age 24 on the difference in the observational predictions between their destination and origin tracts linearly interacted with their age at move (below age 23) and other controls specified in equation (9). Here, we replace the observational predictions on the right hand side with the poverty rates in the origin and destination tracts. We also include symmetric interactions between age at move and poverty rates in the ten tracts that are closest to the actual origin and destination tracts, respectively. We plot the eleven coefficients on the interactions between the destination-origin difference in poverty rates and age at move (for moves below age 23). These coefficients can be interpreted as the causal childhood exposure effect of moving to a tract that is x neighbors away from a tract has 1 SD higher poverty rates. Dashed lines show 95% confidence intervals for the point estimates. We also report the median distance between the own-tract and neighboring tracts in each of the bins as a reference.

FIGURE XIII: The Price of Opportunity



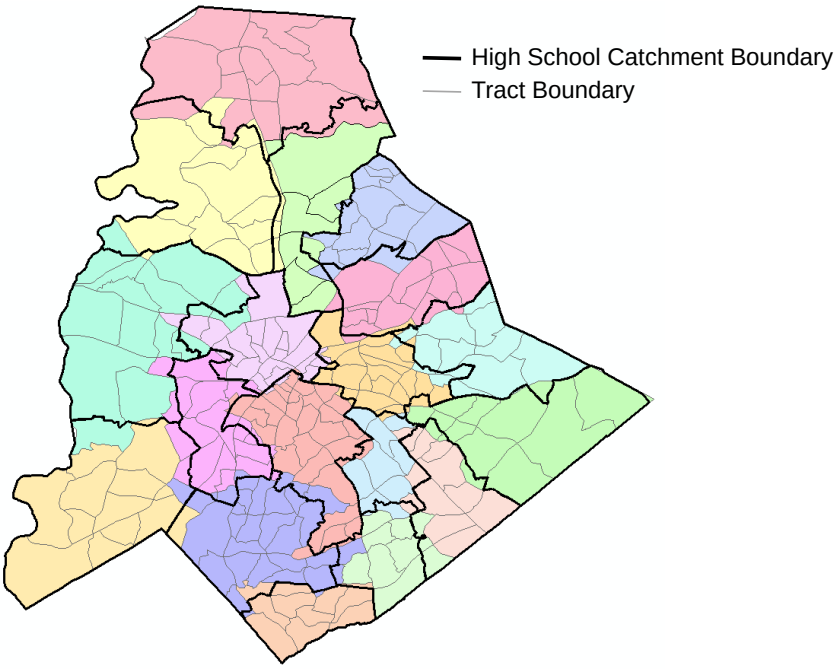
Notes: This figure assesses the relationship between our tract-level estimates of upward mobility (constructed as described in the notes to Figure II) and tract-level median rents (measured in the 1990 American Community Survey as the median rent in a tract for a two-bedroom apartment and inflated to 2015 dollars). Panel A presents a scatter plot of upward mobility vs. median rent, by tract in the Chicago CZ. We also report the signal correlation between upward mobility and rent within CZs nationally. This and all subsequent correlations and standard deviations that follow are weighted by the number of children with below-median income parents. This signal correlation is estimated by first demeaning both variables within CZs and then adjusting for attenuation due to sampling error and noise infusion in our upward mobility estimates by rescaling the raw correlations by the square root of the reliability ratio, which is the ratio of the noise variance to the total within-CZ variance of the upward mobility estimates. Panel B presents a scatter plot of the CZ-specific price of opportunity vs. the Wharton Residential Land Use Regulation index (WRLURI). To calculate the CZ-specific price of opportunity, we first regress median annual rents on upward mobility across tracts within a CZ, weighting as specified above. We then inflate this regression coefficient by the reliability of our upward mobility estimates (0.9) to adjust for noise. Finally, we map the estimates in ranks to dollars as above to obtain an estimate that can be interpreted as the average annual rental cost of a $1 increase in future annual income for children with parents at the 25th percentile. The WRLURI is obtained from Gyourko, Saiz, and Summers (2007, Wharton Land Regulation Data File) and is available for 247 CZs; we limit our sample to these CZs in Panel B. Honolulu is excluded from Panel B because it is an extreme outlier (0.41, 2.30). In Panel C, we report the signal root-mean-squared-error (RMSE) – i.e., the residual standard deviation – from a tract-level regression of our upward mobility estimates on CZ fixed effects alone (first bar); median rents and CZ fixed effects (second bar); and median rents, mean commuting times (measured in the 2000 ACS), and CZ fixed effects (third bar). The signal RMSE is computed by taking the raw RMSE from these regressions and multiplying by the square root of the reliability of the estimates across tracts within CZs. Panel D reports tract-level within-CZ correlations between median rents and the observable and unobservable components of our upward mobility estimates.We define the observable component as the predicted value from a national regression of upward mobility on the set of tract-level characteristics used in Figure V. We define the unobservable component as the residuals from the same regression. We adjust for noise in the unobservable component by reporting a signal correlation.

FIGURE XIV: Predicted Impacts of Moving to "Opportunity Bargain" Areas in MTO Cities



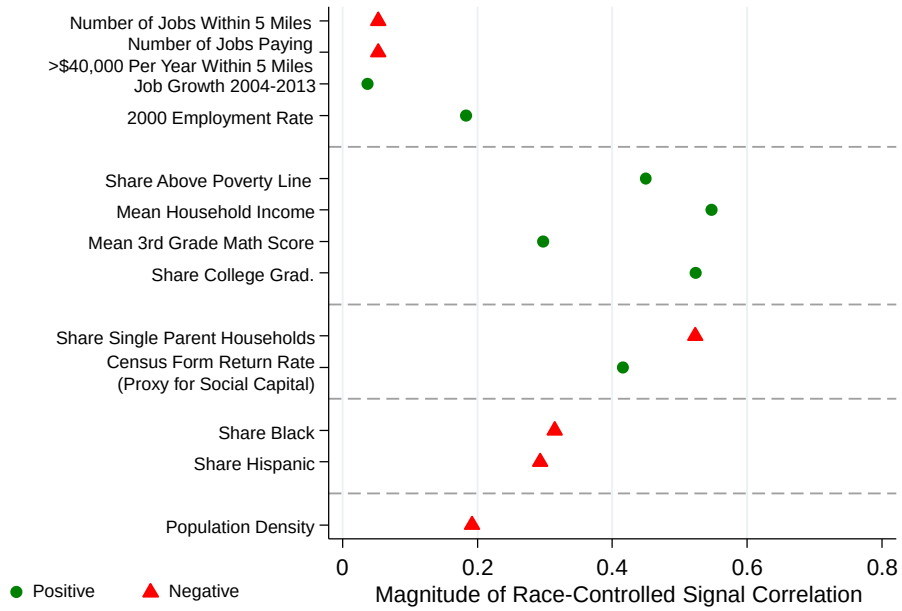
Notes: This figure replicates Figure X, adding five additional points (open circles) that show the predicted outcomes of children who grow up in "opportunity bargain" tracts in each of the five MTO cities. We define opportunity bargain areas in two steps. First, we count the number of tracts that would have been available to experimental voucher holders in the MTO experiment based on the official requirement that the poverty rate should be less than 10% in the 1990 census, which we denote by N_c. In the second step, we rank tracts within each city in descending order based on their mean observed individual income ranks conditional on having parents at the 10th percentile in the Opportunity Atlas data. We then take the N_c highest ranking tracts from the second step that satisfy the following two criteria: (1) median rent in 2000 (based on publicly available 2000 Decennial Census data) is less than or equal to the 90th percentile of the distribution of rents across the tracts where experimental or Section 8 voucher recipients moved in the same city and (2) the commute time using public transportation (as of May 29, 2018 at 8:00 AM, obtained from Google Maps) from the MTO control group tracts is less or equal to the 90th percentile of the distribution of commute times from the control locations to the tracts where experimental or Section 8 voucher recipients moved in the same city. Tracts where control group members and voucher recipients lived are identified by mapping the neighborhood names listed in Online Appendix Table 1c of Chetty, Hendren, and Katz (2016) to Census tracts. Once we have identified the set of opportunity bargain tracts in each city, we compute a population-weighted mean of children's predicted individual income ranks in adulthood across the relevant Census tracts, conditional on having parents at the 10th percentile of the income distribution. We then translate these mean ranks to dollar values at age 26 using a crosswalk from ranks to dollars in 2015 to obtain the five values plotted on the x-axis. The values on the y-axis are predicted values corresponding to these x values using the linear fit estimated in Figure X based on the 15 points from the MTO experimental data.

ONLINE APPENDIX FIGURE I: School Catchment Zone Boundaries in Mecklenburg County, NC



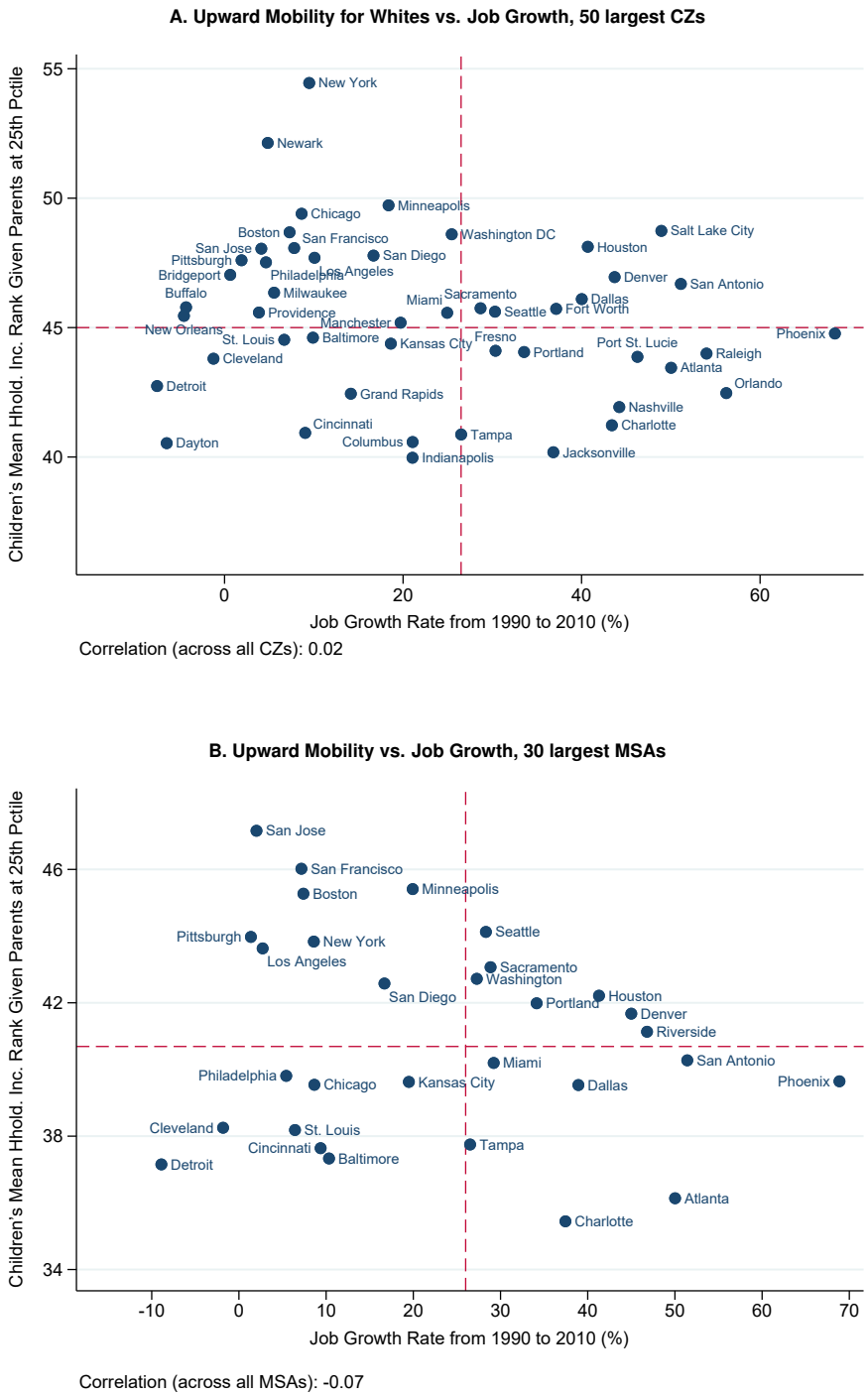
Notes: This figure presents a map of exact high school catchment areas (bold lines) in Mecklenburg County, North Carolina overlayed on tract boundaries (thin lines).

ONLINE APPENDIX FIGURE II: Tract-Level Correlations Between Neighborhood Characteristics and Children's Outcomes Given Parents at the 75th Percentile



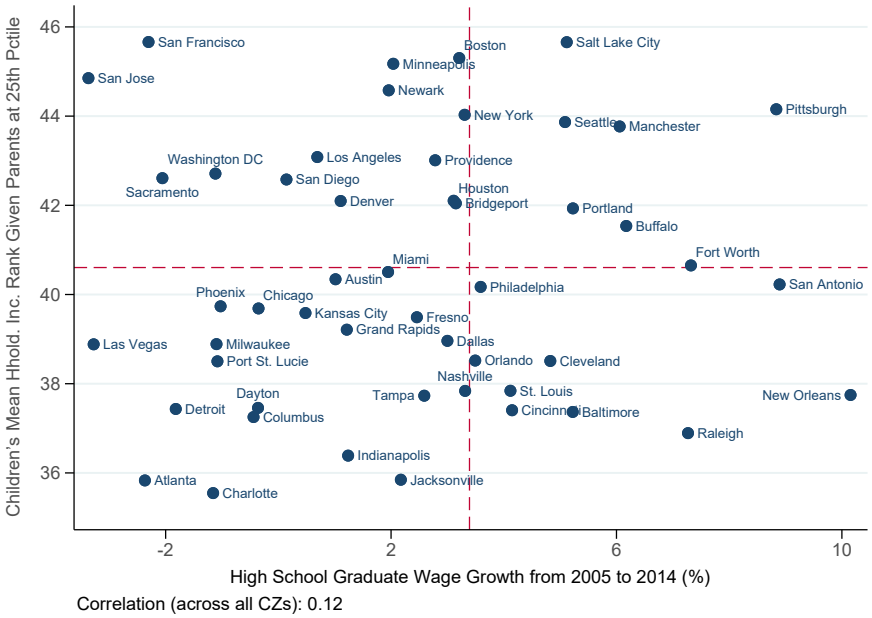
Notes: This figure replicates Figure V using children's mean household income ranks given parents at the 75th percentile, instead of the 25th percentile. See notes to Figure V for details.

ONLINE APPENDIX FIGURE III: Upward Mobility vs. Job Growth



Notes: Panel A replicates Figure VI using upward mobility for whites on the y-axis. Panel B replicates Figure VI for the 30 largest metropolitan statistical areas instead of the 50 largest commuting zones. See notes to Figure VI for details.

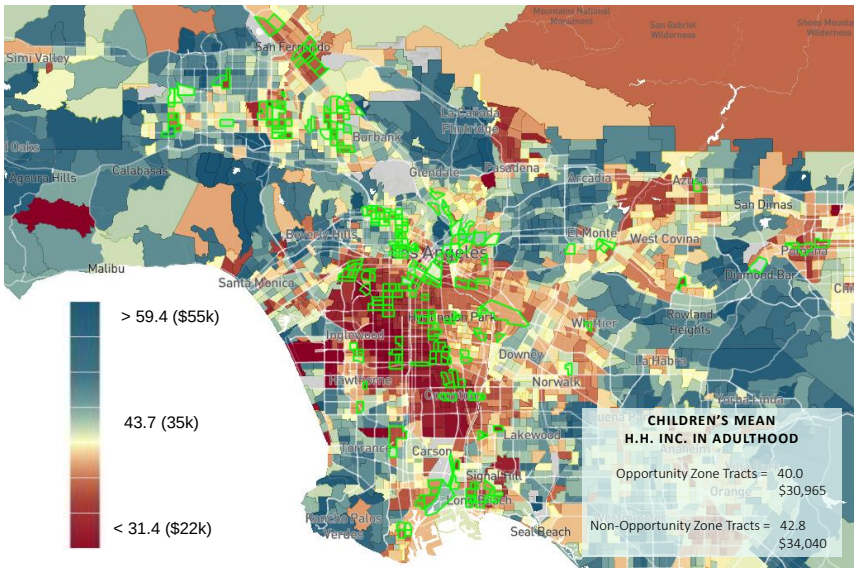
ONLINE APPENDIX FIGURE IV: Upward Mobility vs. Wage Growth for High School Graduates



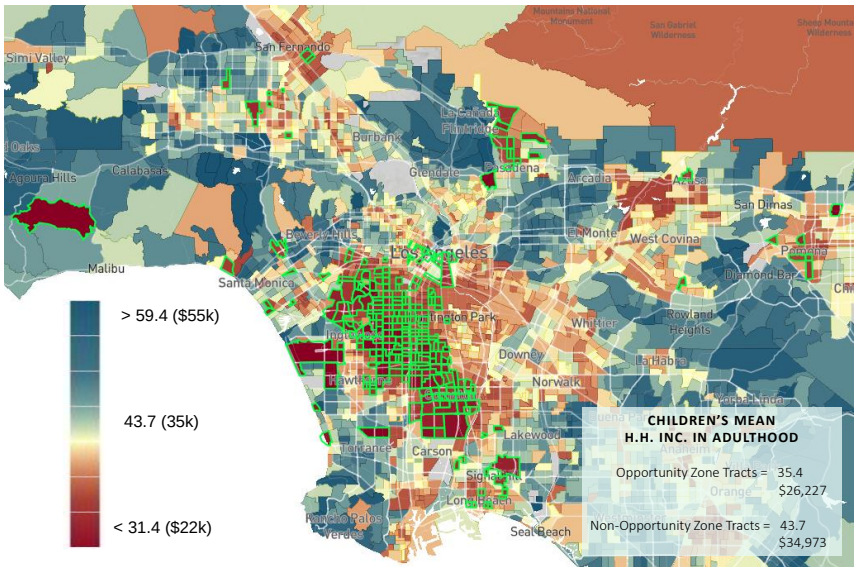
Notes: This figure replicates Figure VI, changing the x variable to log wage growth for high school graduates in each CZ between 2005 and 2014. Wage growth is measured using high school graduate annual earnings and weekly hours worked variables from the American Community Survey. See notes to Figure VI for details.

ONLINE APPENDIX FIGURE V: Targeting Opportunity Zones in Los Angeles

A. Actual Tracts Designated as Opportunity Zones



B. Hypothetical Opportunity Zones using Upward Mobility Estimates

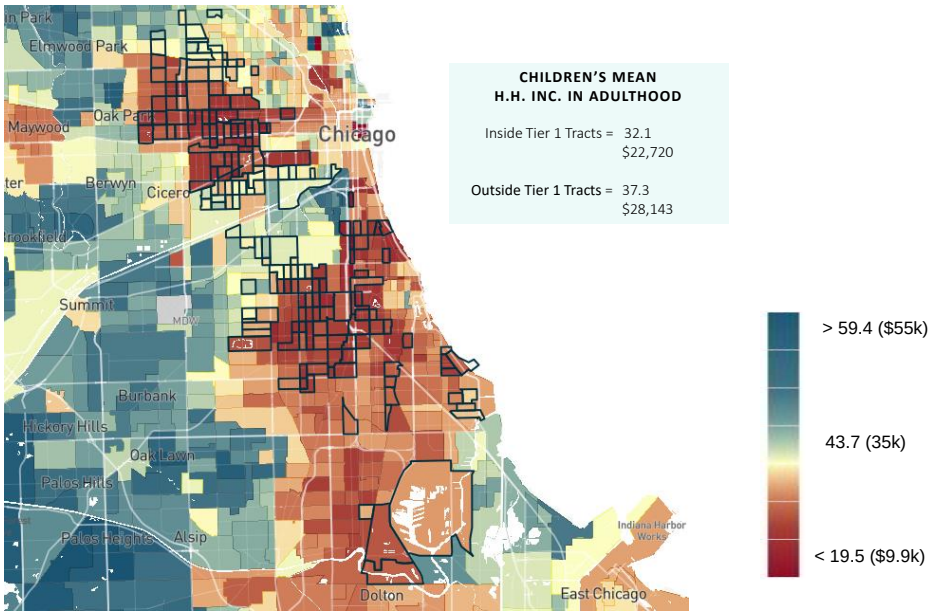


These maps must be printed in color to be interpretable.

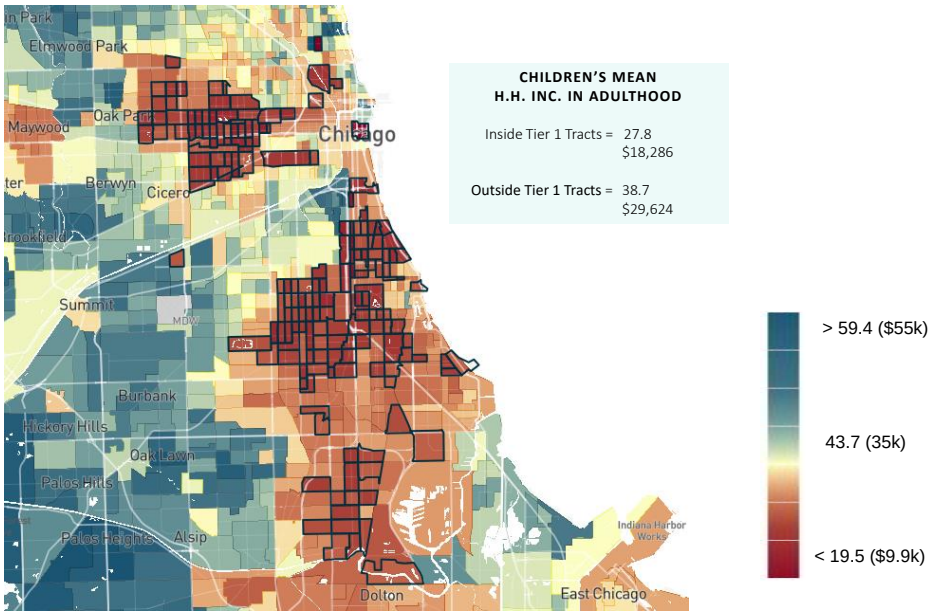
Notes: These maps replicate Figure IIa, plotting children's mean household income ranks given parents at the 25th percentile in Los Angeles. In Panel A, we outline in green borders the tracts that have been designated as Opportunity Zones in Los Angeles. Opportunity Zones are a federal incentive included in the Tax Cuts and Jobs Act to spur investment and improve economic opportunity in low-opportunity neighborhoods. State governments designated qualified areas to receive a host of tax benefits based on poverty and income. In Panel B, we consider a hypothetical alternative targeting strategy, designating the same number of zones in Panel A, but choosing the tracts with the lowest rates of upward mobility in Los Angeles county. In each case, we also report the mean household income rank in adulthood (and the corresponding dollar value) of children with parents at the 25th percentile for areas designated as Opportunity Zones vs. those that are not.

ONLINE APPENDIX FIGURE VI: Targeting Selective High School Admissions in Chicago

A. Actual Tracts Granted Tier 1 Preferential Admission Status



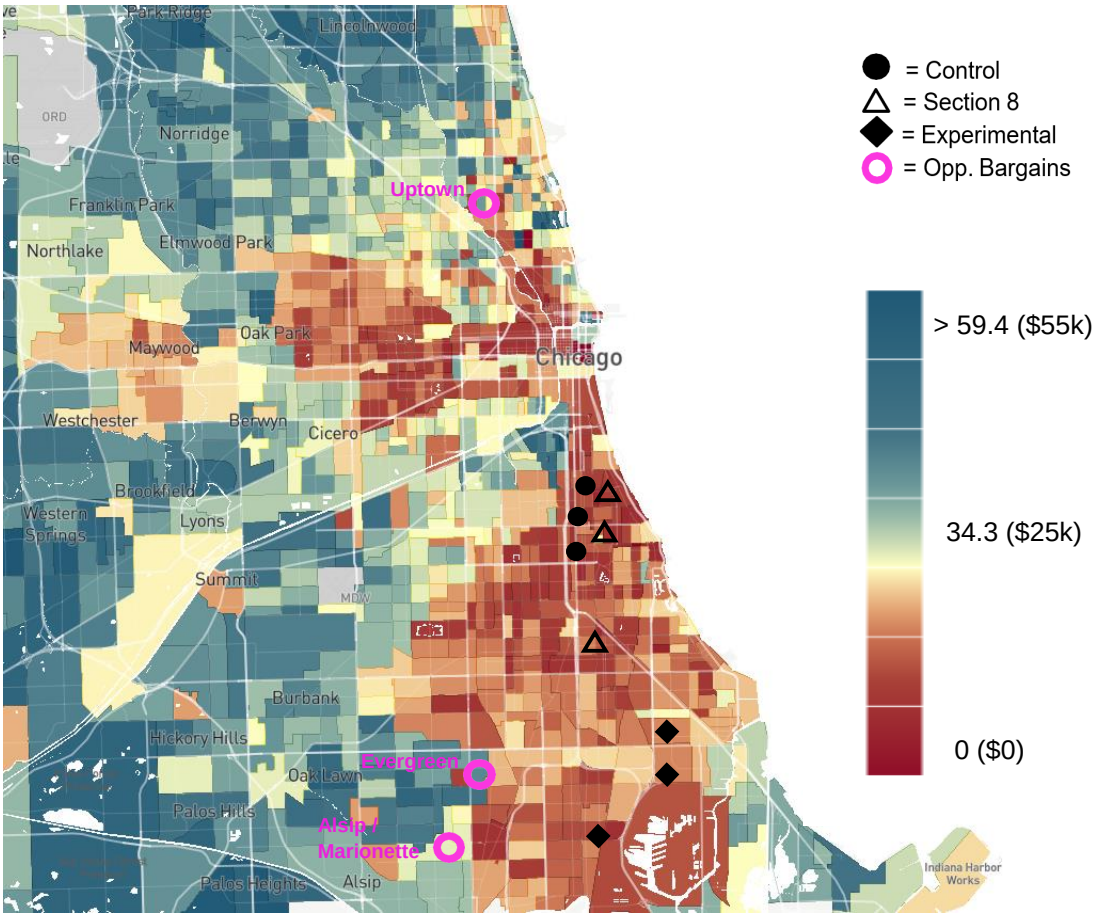
B. Hypothetical Tier 1 Tracts using Upward Mobility Estimates



These maps must be printed in color to be interpretable.

Notes: These maps replicate Figure IIa, plotting children's mean household income ranks given parents at the 25th percentile in Chicago. In Panel A, we outline in black borders the tracts that have been designated as Chicago Exam School Tier 1 tracts. The Chicago Public School tier-based admission system was created to give students from underserved areas greater access to selective schools. Chicago tracts are placed in one of four tracts, where Tier 1 tracts are the most underserved. In Panel B, we consider a hypothetical alternative targeting strategy, designating the same number of zones in Panel A, but choosing the tracts with the lowest rates of upward mobility in Cook County. In each case, we also report the mean household income rank in adulthood (and the corresponding dollar value) of children with parents at the 25th percentile for areas designated as Tier 1 tracts vs. those that are not.

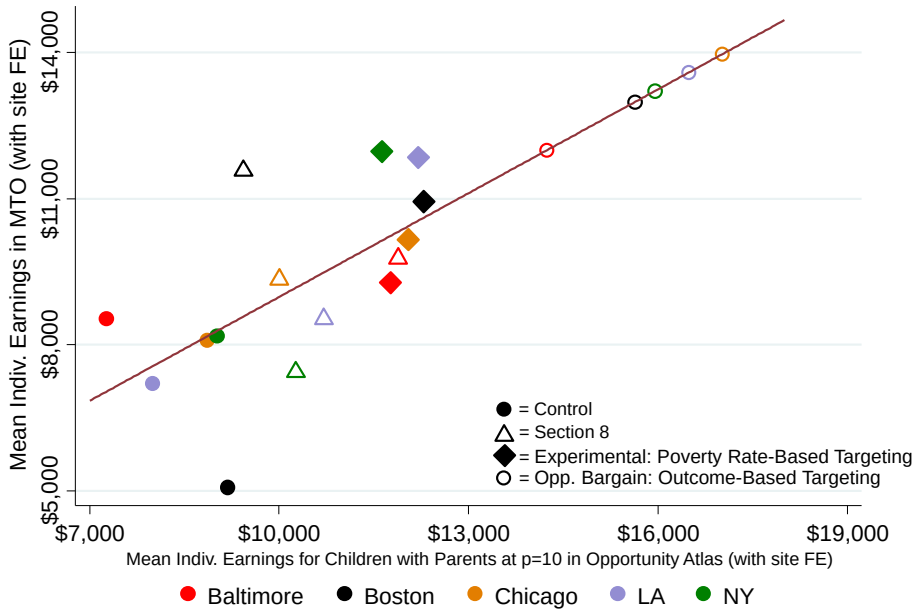
ONLINE APPENDIX FIGURE VII: Most Common Neighborhoods for MTO Participants vs. Opportunity Bargain Tracts in Chicago



This map must be printed in color to be interpretable.

Notes: This map replicates Figure IIa for tracts in Chicago instead of Los Angeles, plotting children's mean household income ranks given parents at the 1st percentile. We mark the most common neighborhoods where families in each of the three treatment arms of MTO lived on the map, using the list in Online Appendix Table 1c of Chetty, Hendren, and Katz (2016). We also mark selected "opportunity bargain" neighborhoods in Chicago, which are identified as described in the notes to Figure XIV.

ONLINE APPENDIX FIGURE VIII: Predicted Impacts of Moving to "Opportunity Bargain" Areas with High Minority Shares in MTO Cities



Notes: This figure replicates Figure XIV, except adding one additional criterion that an "opportunity bargain" must satisfy: the fraction of residents in a tract who do not self-identify as non-Hispanic white alone must be at least 20%, as measured in the 2000 Decennial Census. See notes to Figure XIV for details.