NBER WORKING PAPER SERIES

THE IMPACTS OF NEIGHBORHOODS ON INTERGENERATIONAL MOBILITY I: CHILDHOOD EXPOSURE EFFECTS

Raj Chetty
Nathaniel Hendren

Working Paper 23001
http://www.nber.org/papers/w23001

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2016

An earlier version of this paper was circulated as Part I of "The Impacts of Neighborhoods on Intergenerational Mobility: Childhood Exposure Effects and County Level Estimates" (Chetty et al. (2015)). The opinions expressed in this paper are those of the authors alone and do not necessarily reflect the views of the Internal Revenue Service or the U.S. Treasury Department. This work is a component of a larger project examining the effects of tax expenditures on the budget deficit and economic activity. All results based on tax data in this paper are constructed using statistics originally reported in the SOI Working Paper "The Economic Impacts of Tax Expenditures: Evidence from Spatial Variation across the U.S.," approved under IRS contract TIRNO-12-P-00374 and presented at the Office of Tax Analysis on November 3, 2014. We thank David Autor, Gary Chamberlain, Gordon Dahl, Max Kasy, Lawrence Katz, and numerous seminar participants for helpful comments and discussions. Sarah Abraham, Alex Bell, Augustin Bergeron, Michael Droste, Jamie Fogel, Nikolaus Hildebrand, Alex Olssen, Jordan Richmond, and Benjamin Scuderi provided outstanding research assistance. This research was funded by the National Science Foundation, the Lab for Economic Applications and Policy at Harvard, and Laura and John Arnold Foundation. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects
Raj Chetty and Nathaniel Hendren
NBER Working Paper No. 23001
December 2016
JEL No. H0,J0,R0

ABSTRACT

We show that the neighborhoods in which children grow up play a significant role in determining their earnings, college attendance rates, and fertility and marriage rates by studying more than 7 million families who move across commuting zones in the U.S. By exploiting variation in the age of children when families move, we find that neighborhoods have significant *childhood exposure effects*: the outcomes of children whose families move to a better neighborhood – as measured by the outcomes of children already living there – improve linearly in proportion to the time they spend growing up in that area, at a rate of approximately 4% per year of exposure. We distinguish the causal effects of neighborhoods from confounding factors by comparing the outcomes of siblings within families, studying moves triggered by displacement shocks, and exploiting sharp variation in predicted place effects across birth cohorts, genders, and quantiles to implement overidentification tests. The findings show that place affects intergenerational mobility primarily through childhood exposure, helping reconcile conflicting results in the prior literature.

Raj Chetty
Department of Economics
Stanford University
579 Serra Mall
Stanford, CA 94305
and NBER
chetty@stanford.edu

Nathaniel Hendren
Harvard University
Department of Economics
Littauer Center Room 235
Cambridge, MA 02138
and NBER
nhendren@gmail.com

I Introduction

To what extent are children's opportunities for economic mobility shaped by the neighborhoods in which they grow up? Despite extensive research, the answer to this question remains debated. Observational studies by sociologists have documented significant variation across neighborhoods in economic outcomes (e.g., Wilson 1987, Jencks and Mayer 1990, Massey 1993, Sampson et al. 2002, Sharkey and Faber 2014). However, experimental studies of families that move have traditionally found little evidence that neighborhoods affect economic outcomes (e.g., Katz et al. 2001, Oreopoulos 2003, Ludwig et al. 2013).

Using de-identified tax records covering the U.S. population, we present new quasi-experimental evidence on the effects of neighborhoods on intergenerational mobility that reconcile the conflicting findings of prior work and shed light on the mechanisms through which neighborhoods affect children's outcomes. Our analysis consists of two papers. In this paper, we measure the degree to which the differences in intergenerational mobility across areas in observational data are driven by causal effects of place. In the second paper (Chetty and Hendren 2016), we build on the research design developed here to construct estimates of the causal effect of growing up in each county in the United States on children's long-term outcomes and characterize the features of areas that produce good outcomes.

Our analysis is motivated by our previous work showing that children's expected earnings conditional on their parents' incomes vary substantially with the area (commuting zone or county) in which they grow up (Chetty, Hendren, Kline, and Saez 2014).[1] This geographic variation in intergenerational mobility could be driven by two very different sources. One possibility is that neighborhoods have causal effects on economic mobility: that is, moving a given child to a different neighborhood would change his or her life outcomes. Another possibility is that the observed geographic variation is due to systematic differences in the types of people living in each area, such as differences in demographic makeup or wealth.

We assess the relative importance of these two explanations by asking whether children who move to areas with higher rates of upward income mobility among "permanent residents" have better outcomes themselves.[2] Since moving is an endogenous choice, simple comparisons of the

[1] We characterize neighborhood (or "place") effects at two geographies: counties and commuting zones (CZs), which are aggregations of counties that are similar to metro areas but cover the entire U.S., including rural areas. Naturally, the variance of place effects across these broad geographies is a lower bound for the total variance of neighborhood effects, which would include additional local variation.

[2] We define "permanent residents" as the parents who stay in the same commuting zone (or, in the county-level analysis, the same county) throughout the period we observe (1996-2012).

outcomes of children whose families move to different areas confound causal effects of place with selection effects (differences in unobservables). We address this identification problem by exploiting variation in the *timing* of children's moves across areas. We compare the outcomes of children who moved to a better (or worse) area at different ages to identify the rate at which the outcomes of children who move converge to those of the permanent residents.[3] The identification assumption underlying our research design is that the selection effects (children's unobservables) associated with moving to a better vs. worse area do not vary with the age of the child when the family moves. This is a strong assumption, one that could plausibly be violated for several reasons. For instance, families who move to better areas when their children are young may be more educated or invest more in their children in other ways. We evaluate the validity of this identification assumption in detail and show that it holds in practice after presenting a set of baseline results.

In our baseline analysis, we focus on families with children born between 1980 and 1988 who moved once across commuting zones between 1997 and 2010, a sample that consists of 1.5 million movers . We find that on average, spending an extra year in a CZ or county where the mean income rank of children of permanent residents (for families at the same income level) is 1 percentile higher increases a child's expected income rank by approximately 0.04 percentiles. That is, the outcomes of children who move converge to the outcomes of permanent residents of the destination area at a rate of approximately 4% per year of exposure. Symmetrically, moving to an area where permanent residents have worse outcomes reduces a child's expected income by 4% per year. Children who move more than once – entering and leaving a given area within our sample – pick up gains that are proportional to the number of years in which they lived in that area.

These convergence patterns imply that neighborhoods have substantial *childhood exposure effects*: every additional year of childhood spent in a better environment improves a child's long-term outcomes. Convergence is linear with respect to age: moving to a better area at age 8 instead of 9 is associated with the same improvement in earnings as moving to that area at age 15 instead of 16. The exposure effects persist until children are in their early twenties. Extrapolating over the duration of childhood, from age 0 to 20, the 4% annual convergence rate implies that children who move at birth to area with one unit better outcomes among permanent residents would pick up about 80% of that effect themselves. We find childhood exposure effects of a similar magnitude for several other outcomes, including rates of college attendance, teenage employment, teenage birth,

[3] Throughout the paper, we refer to areas where children have better outcomes in adulthood as "better" neighborhoods. We use this terminology without any normative connotation, as there are of course many other amenities of neighborhoods that may be relevant from a normative perspective.

and marriage. We also find similar exposure effects when families moves across counties.

As noted above, the identification assumption underlying the interpretation of the 4% convergence rate as a causal exposure effect is that the potential outcomes of children who move to better vs. worse areas do not vary with the age at which they move. We use three approaches to evaluate this assumption: controlling for observable factors, isolating plausibly exogenous moves triggered by aggregate displacement shocks, and implementing a set of outcome-based placebo tests. The first two approaches are familiar techniques in the treatment effects literature, while the third exploits the multi-dimensional nature of the treatments we study to implement overidentification tests.

To implement the first approach, we begin by controlling for factors that are fixed within the family (e.g., parent education) by including family fixed effects, as in Plotnick and Hoffman (1996) and Aaronson (1998). This approach identifies exposure effects from comparisons between siblings, effectively asking whether the *difference* in outcomes between two siblings in a family that moves is proportional to the size of the age gap between them. We estimate an annual exposure effect of approximately 4% per year with family fixed effects, very similar to our baseline estimates. The 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 that directly affects children in proportion to exposure time independent of neighborhoods. We cannot observe all such time-varying factors, but we do observe two particularly important characteristics of the family environment in each year: income and marital status. Controlling flexibly for changes in income and marital status interacted with the age of the child at the time of the move has no impact on the exposure effect estimates.

While the preceding results rule out confounds due to observable factors such as income, they do not address potential confounds due to unobservable factors. In particular, whatever event endogenously induced a family to move (e.g., a wealth shock) could also have had direct effects on their children's outcomes. Our second approach addresses the problem of bias associated with endogenous choice by directly focusing on a subset of moves that are more likely to be driven by exogenous aggregate shocks. In particular, we identify moves that occur as part of large outflows from ZIP codes, often caused by natural disasters or local plant closures. We replicate our baseline design within this subsample of displaced movers, comparing the outcomes of children who move to different destinations at different ages. We obtain similar exposure effect estimates for displaced households, mitigating concerns that our baseline estimates are biased by omitted variables correlated with a household's choice of when to move.[4]

Although the evidence from the first two approaches strongly supports the validity of the identification assumption, each of these approaches itself rests on assumptions – selection on observables and exogeneity of the displacement shocks – that could themselves potentially be violated. We therefore turn to a third approach – a set of placebo (overidentification) tests that exploit heterogeneity in place effects across subgroups – that in our view provides the most compelling method of assessing the validity of the research design. We begin by analyzing heterogeneity in place effects across birth cohorts. Although outcomes within CZs are highly persistent over time, some places improve and others decline. Exploiting this variation, we find using multivariable regressions that the outcomes of children who move to a new area converge to the outcomes of permanent residents of the destination in their *own* birth cohort but are unrelated to those of surrounding birth cohorts (conditional on their own birth cohort's predictions). Such cohort-specific convergence is precisely what one would expect in the causal exposure effect model, but it would be unlikely to emerge from sorting or other omitted variables because the cohort-specific effects are only realized with a long time lag, after children grow up.

We implement analogous placebo tests by exploiting variation in the *distribution* of outcomes across areas. For instance, low-income children who spend their entire childhood in Boston and San Francisco have similar outcomes on average, but children in San Francisco are more likely to end up in the upper (top 10%) or lower tail (bottom 10%) of the income distribution. The causal exposure effects model predicts convergence not just at the mean but across the entire distribution; in contrast, it would be unlikely that omitted variables (such as changes in parent wealth) would happen to perfectly replicate the entire distribution of outcomes in each area in proportion to exposure time. In practice, we find quantile-specific distributional convergence: controlling for mean outcomes, children's outcomes converge to predicted outcomes in the destination across the distribution in proportion to exposure time at a rate of approximately 4% per year.

Finally, we implement placebo tests exploiting heterogeneity in place effects across genders. Though place effects are highly correlated across genders, there are some places where boys do worse than girls (e.g., areas with highly concentrated poverty) and vice versa. When a family with both a daughter and a son moves to an area that is especially good for boys, their son's outcomes converge to those in the destination much more than their daughter's outcomes. Once again, if our

[4] We eliminate variation due to individuals' endogenous choices of *where* to move in these specifications by instrumenting for each household's change in neighborhood quality using the average change in neighborhood quality of those who move out of the ZIP code during the years in our sample.

findings of neighborhood exposure effects were driven by sorting or omitted variables, one would not expect to find gender-specific convergence unless families are fully aware of the exact gender differences in outcomes across areas and sort to neighborhoods on these gender differences.

Putting together these results, we conclude that the baseline timing-of-move design yields consistent estimates of neighborhood exposure effects, of about 4% per year. An important caveat in interpreting this estimate is that it is a local average treatment effect estimated based on households who choose to move to certain areas. The mean exposure effect of moving a randomly selected household to a new area may differ, since households that choose to move to a given area may be more likely to benefit from that move than the average household in the population. The fact that exposure effects are similar within the subset of displaced households and are symmetric for moves to better and worse areas suggest that the endogeneity of choice does not have a substantial effect on the magnitude of exposure effects, but further work is needed to understand how exposure effects vary with households' willingness to move.

Our findings yield three broad lessons. First, place matters for intergenerational mobility: the differences we see in outcomes across neighborhoods are largely due to the causal effect of places rather than differences in the characteristics of their residents. Second, place matters for intergenerational mobility largely through differences in childhood environment, rather than the differences in labor market conditions that have received attention in previous studies of place. Moving to a better area just before entering the labor market has little impact on individual's outcomes, suggesting that place-conscious policies to promote upward mobility should focus primarily on improving the local childhood environment rather than conditions in adulthood. Third, we find that each year of childhood exposure matters roughly equally; there is no "critical age" after which the marginal returns to being in a better neighborhood fall sharply. This result is germane to recent policy discussions regarding early childhood interventions, as it suggests that improvements in neighborhood environments can be beneficial even in adolescence.

Our results help explain why previous experimental studies – most notably, the Moving to Opportunity (MTO) Experiment – failed to detect significant effects of moving to a better neighborhood on economic outcomes. Prior analyses of the MTO experiment focused primarily on the effects of neighborhoods on adults and older youth (e.g. Kling et al. (2007)), because data on the long-term outcomes of younger children were unavailable. In a followup paper (Chetty, Hendren, and Katz 2016), we link the MTO data to tax records and show that the MTO data exhibit childhood exposure effects similar to those identified here. In particular, Chetty, Hendren, and Katz

(2016) find substantial improvements in earnings and other outcomes for children whose families received experimental vouchers to move to low-poverty neighborhoods at young ages. In contrast, children who moved at older ages experienced no gains or slight losses.[5] The findings in the present paper complement the re-analysis of MTO by (a) delivering very precise estimates of the magnitude and linear age pattern of childhood exposure effects (making use of 7 million observations instead of 4,500 observations) and (b) developing a scalable research design that can be used to estimate neighborhood effects in all areas even in the absence of a randomized experiment.

More generally, our findings imply that much of the neighborhood-level variation in economic outcomes documented in previous observational studies does in fact reflect causal effects of place, but that such effects arise through accumulated childhood exposure rather than immediate impacts on adults. The idea that exposure time to better neighborhoods may matter has been noted since at least Wilson (1987) and Jencks and Mayer (1990), and has received growing attention in observational studies in sociology (Crowder and South (2011), Wodtke et al. (2011, 2012); Wodtke (2013), and Sampson 2012; Sharkey and Faber 2014). We contribute to this literature by presenting quasi-experimental estimates of exposure effects, which address the concerns about selection and omitted variable bias that arise in observational studies (e.g., Clampet-Lundquist and Massey (1993); Ludwig et al. (2008)). Although we find evidence of childhood exposure effects that are qualitatively consistent with the observational studies, we find no evidence of exposure effects in adulthood either in this study or our followup MTO study, contrary to the patterns observed in observational data (e.g., Clampet-Lundquist and Massey (1993)).

The paper is organized as follows. Section II describes the data. Section III presents our empirical framework, starting with a description of differences in intergenerational mobility across areas for permanent residents and then specifying our estimating equations. Section IV presents baseline estimates of neighborhood exposure effects on earnings and other life outcomes. Section V presents the tests evaluating our identification assumption. Section VI discusses mechanisms, and Section VII concludes.

[5] One important distinction beteween the two studies is that the analysis sample in the present quasi-experimental study consists entirely of families who moved across commuting zones, whereas the MTO experiment compares families who moved with families who did not move at all or stayed in an area similar to where they lived before. As a result, the analysis here identifies the effects of moving to better vs. worse areas *conditional* on moving to a different area, whereas the MTO analysis compares the effects of moving vs. staying in a given area. The exposure effect estimates here thus net out any fixed disruption costs of moving to a different area, whereas such costs are not netted out in the MTO experiment. This distinction may explain why Chetty, Hendren, and Katz (2016) find slightly negative effects for children who move at older ages in the MTO data, whereas we estimate positive exposure effects of moving to a better area (conditional on moving) at all ages here.

II Data

We use data from federal income tax records spanning 1996-2012. The data include both income tax returns (1040 forms) and third-party information returns (e.g., W-2 forms), which contain information on the earnings of those who do not file tax returns. Because our empirical analysis is designed to determine how much of the geographic variation in intergenerational mobility documented by Chetty et al. (2014) is due to causal effects of place, our analysis sample is essentially identical to the "extended sample" used in Chetty et al. (2014). Online Appendix A of Chetty et al. (2014) gives a detailed description of how we construct the analysis sample starting from the raw population data. Here, we briefly summarize the key variable and sample definitions, following Section III of Chetty et al. (2014).[6]

II.A Sample Definitions

Our base dataset of children consists of all individuals who (1) have a valid Social Security Number or Individual Taxpayer Identification Number, (2) were born between 1980-1988, and (3) are U.S. citizens as of 2013.[7] We impose the citizenship requirement to exclude individuals who are likely to have immigrated to the U.S. as adults, for whom we cannot measure parent income. We cannot directly restrict the sample to individuals born in the U.S. because the database only records current citizenship status.

We identify the parents of a child as the first tax filers (between 1996-2012) who claim the child as a child dependent and were between the ages of 15 and 40 when the child was born. If the child is first claimed by a single filer, 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.

If parents never file a tax return, we do not link them to their child. Although some low-income individuals do not file tax returns in a given year, almost all parents file a tax return at some point between 1996 and 2012 to obtain a tax refund on their withheld taxes and the Earned Income Tax Credit (Cilke 1998). We are therefore able to identify parents for approximately 95% of the children in the 1980-1988 birth cohorts. The fraction of children linked to parents drops sharply prior to

[6]The tax records we use were drawn in the middle of 2013. They include a complete set of information returns (W-2's) for 2012, but exclude a small number of amendments and late filings for 1040s. This slight incompleteness of the data is inconsequential, as using data through 2011 yields very similar results.

[7]For selected outcomes that can be measured at earlier ages, such as teenage labor force participation rates, we extend the sample to include more recent birth cohorts, up to 1996.

the 1980 birth cohort because our data begins in 1996 and many children begin to the leave the household starting at age 17 (Chetty et al. (2014); Online Appendix Table I). This is why we limit our analysis to children born during or after 1980.

Our full analysis sample includes all children in the base dataset who are born in the 1980-88 birth cohorts, for whom we are able to identify parents, and whose mean parent income between 1996-2000 is strictly positive (which excludes 1.2% of children).[8] We divide the full sample into two parts: *permanent residents* (or stayers) and *movers*. We define the permanent residents of each commuting zone (CZ) c as the subset of parents who reside in a single CZ c in all years of our sample, 1996-2012. The movers sample consists of all individuals in the full sample who are not permanent residents.

In our baseline analysis, we focus on the subset of individuals who live in CZs with populations above 250,000 (based on the 2000 Census) to ensure that we have adequate precision to estimate place effects. We also restrict attention to movers who moved at least 100 miles to eliminate moves across CZ borders that do not reflect a true change of location.[9] There are approximately 24.6 million children in the baseline sample, of whom 19.5 million are children of permanent residents and 1.55 million move at least 100 miles. [10]

II.B Variable Definitions and Summary Statistics

In this section, we define the key variables we use in our analysis. We measure all monetary variables in 2012 dollars, adjusting for inflation using the headline consumer price index (CPI-U). We begin by defining the two key variables we measure for parents: income and location.

Parent Income. Our primary measure of parent income is total pre-tax income at the household level, which we label *parent family income*. In years where a parent files a tax return, we define family 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, we define family income as the sum of wage earnings (reported on form W-2), unemployment benefits (reported on form 1099-G), and gross social security and

[8]We limit the sample to parents with positive income because parents who file a tax return (as required to link them to a child) yet have zero income are unlikely to be representative of individuals with zero income and those with negative income typically have large capital losses, which are a proxy for having significant wealth.

[9]We measure the distance of moves as the distance between the centroids of the origin and destination ZIPs. . We show the robustness of our results to using alternative cutoffs for minimum population size and move distances in Online Appendix Table II.

[10]This 1.55M sample selection also imposes the restriction discussed in Section IV that we are able to observe families in the destination for at least 2 years.

disability benefits (reported on form SSA-1099) for both parents.[11] In years where parents have no tax return and no information returns, family income is coded as zero.[12]

Our baseline income measure includes labor earnings and capital income as well as unemployment insurance, social security, and disability benefits. It excludes non-taxable cash transfers such as TANF and SSI, in-kind benefits such as food stamps, all refundable tax credits such as the EITC, non-taxable pension contributions (e.g., to 401(k)'s), and any earned income not reported to the IRS. Income is always measured prior to the deduction of individual income taxes and employee-level payroll taxes.

In our baseline analysis, we average parents' family income over the five years from 1996 to 2000 to obtain a proxy for parent lifetime income that is less affected by transitory fluctuations (Solon 1992). We use the earliest years in our sample to best reflect the economic resources of parents while the children in our sample are growing up.[13] Because we measure parent income in a fixed set of years, the age of the child when parent income is measured varies across birth cohorts. We account for this variation by conditioning on the child's birth cohort throughout our analysis.

Parent Location. In each year, parents are assigned ZIP codes of residence based on the ZIP code from which they filed their tax return. If the parent does not file in a given year, we search information returns (such as W-2) for a ZIP code in that year. Non-filers with no information returns are assigned missing ZIP codes. For children whose parents were married when they were first claimed as dependents, we always track the mother's location if marital status changes. We map parents' ZIP codes to counties using a crosswalk that combines the union of a 1999 Census crosswalk and a 2011 Housing and Urban Development crosswalk (Census (1999); HUD (2011)).[14] We then assign counties to commuting zones using the crosswalk constructed by David Dorn. See Online Appendix A of Chetty et al. (2014) for further details on the mapping of ZIP codes to CZs.

[11] The database does not record W-2's and other information returns prior to 1999, so non-filer's income is coded as 0 prior to 1999. Assigning non-filing parents 0 income has little impact on our estimates because only 2.9% of parents in the full analysis sample do not file in each year prior to 1999 and most non-filers have very low W-2 income (Chetty et al. (2014)). For instance, in 2000, median W-2 income among non-filers was $29.

[12] Importantly, these observations are true zeros rather than missing data. Because the database covers all tax records, we know that these individuals have 0 taxable income.

[13] Formally, we define mean family income as the mother's family income plus the father's family income in each year from 1996 to 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 family income over the 5 year period. For parents who are married initially and then divorce, this measure tracks the mean family 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 family income (including the new spouse's income) after marriage. These household measures of income increase with marriage and naturally do not account for cohabitation; to ensure that these features do not generate bias, we assess the robustness of our results to using individual measures of income.

[14] The 1999 census crosswalk is no longer publicly posted at https://www.huduser.gov/portal/datasets/usps_crosswalk.html, but is available on our project website.

Next, we define the outcomes that we analyze for children.

Income. We define child family income in exactly the same way as parent family income. We measure children's annual incomes at ages ranging from 24-30 and define the child's household based on his or her marital status at the point at which income is measured. For some robustness checks, we analyze individual income, defined as the sum of individual W-2 wage earnings, UI benefits, SSDI payments, and half of household self-employment income (see Online Appendix A of Chetty et al. (2014) for more details).

Employment. We define an indicator for whether the child is employed at a given age based on whether he has a W-2 form filed on his behalf at that age. We measure employment rates starting at age 16 to analyze teenage labor force participation.

College Attendance. We define college attendance as an indicator for having one or more 1098-T forms filed on one's behalf when the individual is aged 18-23. Title IV institutions – all colleges and universities as well as vocational schools and other post-secondary institutions eligible for federal student aid – are required to file 1098-T forms that report tuition payments or scholarships received for every student. The 1098-T forms are filed directly by colleges independent of whether an individual files a tax return and are available from 1999-2012. Comparisons to other data sources indicate that 1098-T forms capture more than 95% of college enrollment in the U.S. (see Chetty et al. (2014), Appendix B).[15]

Teenage Birth. For women, we define an indicator for teenage birth if they are listed as a parent on a birth certificate when they are between the ages of 13 and 19, using data from the Social Security Administration's DM-2 database.[16]

Marriage. We define an indicator for whether the child is married at at a given age based on the marital status listed on 1040 forms for tax filers. We code non-filers as single because linked CPS-IRS data show that the vast majority of non-filers below the age of 62 are single (Cilke 1998).

Summary Statistics. Table I reports summary statistics for our analysis sample and various

[15]Colleges are not required to file 1098-T forms for students whose qualified tuition and related expenses are waived or paid entirely with scholarships or grants. However, the forms are frequently available even for such cases because of automated reporting to the IRS by universities. Approximately 6% of 1098-T forms are missing from 2000-2003 because the database contains no 1098-T forms for some small colleges in these years (Chetty et al. (2014)). To verify that this does not affect our results, we confirm that our results are very similar when we exclude data from these years (not reported).

[16]The total count of births in the SSA DM-2 database closely matches vital statistics counts from the Center for Disease Control prior to 2008; however, the DM-2 database contains approximately 10% fewer births between 2008-2012. Using an alternative measure of teenage birth that does not suffer from this missing data problem – in which we define a woman as having a teen birth if she ever claims a dependent who was born while she was between the ages of 13 and 19 – yields very similar results (not reported). We do not use the dependent-claiming definition as our primary measure of teenage birth because it only covers children who are claimed as dependents by their mothers.

subgroups. In general, movers are slightly negatively selected on observables relative to permanent residents. For permanent residents, median parent family income is $59,200, as compared to $58,700 for our sample of one-time movers. Children of permanent residents have a mean family income of $48,377 when they are 30 years old, compared with $47,882 for one-time movers. Roughly 70% of children of permanent residents and one-time movers are enrolled in a college at some point between the ages of 18 and 23 and roughly 11% of daughters of permanent residents and one-time movers have a teenage birth.

III Empirical Framework

In this section, we first present a descriptive characterization of the earnings outcomes of children who grow up in different areas in the U.S. We then formally define our estimands of interest – childhood exposure effects – and describe the research design we use to identify these exposure effects in observational data.

III.A Geographical Variation in Outcomes of Permanent Residents

We conceptualize "neighborhood" effects as the sum of place effects at different geographies, ranging from broad to narrow: commuting zones, counties, ZIP codes, and blocks. In this paper, we focus on variation across commuting zones (CZs). CZs are aggregations of counties based on commuting patterns in the 1990 Census constructed by Tolbert and Sizer (1996). There are 741 CZs in the U.S.; on average, each CZ contains 4 counties and has a population of 380,000. We also replicate the results reported in the main text at the county level in Online Appendix Table IV. We focus on variation across relatively broad geographies to maximize statistical precision, as some of our research designs require large sample sizes to discern fine variation in place effects. Of course, the variation across CZs and counties we document is a lower bound for the total variance of neighborhood effects, which would include additional variation at narrower geographies.

We characterize children's outcomes in each CZ using the same approach as in Chetty et al. (2014), except that we focus here on "permanent residents" – the subset of children whose families never move between 1996 and 2012 – to measure outcomes for children who spent their entire childhoods in a single area.[17] Importantly, our definition of permanent residents conditions on

[17]Because our data start in 1996, we cannot measure parents' location over their children's entire childhood. For the 1980 birth cohort, we measure parents' location between the ages of 16 and 32; for the 1991 birth cohort, we measure parents' location between 5 and 21. This creates measurement error in children's childhood environment that is larger in earlier birth cohorts. Fortunately, we find that our results do not vary significantly across birth cohorts, and in particular remain similar for the most recent birth cohorts. The reason such measurement error turns

parents' locations, not children's locations in adulthood. The CZ where a child grew up may differ from the CZ where he lives when we measure his earnings in adulthood.

Since places can have different effects across parent income levels and over time, we characterize children's mean outcomes conditional on their parents' income separately for each CZ c and birth cohort s. Chetty et al. (2014) show that measuring incomes using percentile *ranks* (rather than dollar levels) has significant statistical advantages. Following their approach, we define child i's percentile rank y_i based on his position in the *national* distribution of incomes relative to all others in his birth cohort. Similarly, we measure the percentile rank of the parents of child i, $p(i)$, based on their positions in the national distribution of parental income for child i's birth cohort.

Let $\bar{y}_{pcs}$ denote the mean rank of children with parents at percentile p of the income distribution in CZ c in birth cohort s. Figure I illustrates how we estimate $\bar{y}_{pcs}$ for children born in 1982 to parents who are permanent residents of the Chicago CZ. This figure plots the mean child rank at age 30 within each percentile bin of the parent income distribution, $E[y_i|p(i) = p]$. The conditional expectation of a child's rank given his parents' rank is almost perfectly linear, a property that is robust across CZs (Chetty et al. (2014), Online Appendix Figure IV). Exploiting linearity, we parsimoniously summarize the relationship between children's mean income ranks and their parents' ranks by regressing children's ranks on their parents' ranks in each CZ c and birth cohort s:

$$y_i = \alpha_{cs} + \psi_{cs} p_i + \varepsilon_i. \tag{1}$$

We then estimate $\bar{y}_{pcs}$ using the fitted values from this regression:

$$\bar{y}_{pcs} = \hat{\alpha}_{cs} + \hat{\psi}_{cs} p. \tag{2}$$

For example, in Chicago, $\bar{y}_{25,c,1985} = 40.8$ for children growing up at the 25th percentile of the national income distribution and $\bar{y}_{75,c,1985} = 56.1$ for children growing up at the 75th percentile.

Figure II maps children's mean income ranks at age 30 by CZ for children with parents at the 25th percentile (Panel A) and 75th percentile (Panel B). We construct these maps by dividing CZs into deciles based on their estimated value of $\bar{y}_{25,c,s}$ and $\bar{y}_{75,c,s}$; lighter colors represent deciles with higher mean outcomes. As documented by Chetty et al. (2014), children's outcomes vary substantially across CZs, especially for children from low-income families. Chetty et al. (2014, Section V.C) summarize the spatial patterns in these maps in detail. Here, we focus on investigating

out to be modest empirically is that most families who stay in a given area for several years tend not to have moved in the past either. For example, among families who stayed in the same CZ c when their children were between ages 16-24, 81.5% of them lived in the same CZ when their children were age 8.

whether the variation in these maps is driven by causal effects of place or heterogeneity in the types of people living in different places.

III.B Definition of Exposure Effects

Our objective is to determine how much a given child's potential outcomes would improve on average if he were to grow up in an area where the permanent residents' outcomes are 1 percentile point higher. We answer this question by studying children who move across areas, focusing on identifying childhood *exposure effects*. We define the exposure effect at age m as the impact of spending year m of one's childhood in an area where permanent residents' outcomes are 1 percentile point higher.

Formally, consider a hypothetical experiment in which we randomly assign children to new neighborhoods d starting at age m for the rest of their childhood. The best linear predictor of children's outcomes y_i in the experimental sample, based on the permanent residents' outcomes in CZ d ($\bar{y}_{pds}$), can be written as

$$y_i = \alpha + \beta_m \bar{y}_{pds} + \theta_i, \tag{3}$$

where the error term θ_i captures family inputs and other determinants of children's outcomes. Since the random assignment guarantees that θ_i is orthogonal to $\bar{y}_{pds}$, estimating (3) using OLS yields a coefficient β_m that represents the mean impact of spending year m of one's childhood onwards in an area where permanents residents have 1 percentile better outcomes. We define the exposure effect at age m as $\gamma_m = \beta_m - \beta_{m+1}$.[18] Note that if the earnings y_i is measured at age T, $\beta_m = 0$ for $m > T$, as moving after the outcome is measured cannot have a causal effect on the outcome.

Estimating the exposure effects $\overrightarrow{\gamma}_m$ is of interest for several reasons. First, a positive effect (at any age) allows us to reject the null hypothesis that neighborhoods do not matter, a null of interest given experimental evidence to date. Second, $\overrightarrow{\gamma}_m$ is informative about the ages at which neighborhood environments matter most for children's outcomes. Third, the magnitude of $\beta_0 = \sum_{t=0}^{T} \gamma_m$ – the impact of assigning children to better neighborhood from birth – provides an estimate of the degree to which the differences in children's outcomes across areas are due to place effects vs. selection. If place effects are homogeneous within birth cohorts and parent income groups, $\beta_0 = 0$ would imply that all of the variation across areas is due to selection, while $\beta_0 = 1$ would imply that all of the variation would reflect causal effects of place. More generally, the magnitude of β_0 tells us how much of the differences across areas in Figure II rub off on children

[18]We assume that β_m does not vary across parent income percentiles p for simplicity, but one could estimate (3) separately by p to identify β_{mp} for each percentile p. Empirically, we find that β_{mp} does not vary significantly across percentiles.

who are randomly assigned to live there from birth.

Although identifying exposure effects sheds light on the importance of place effects on average, it does not identify the causal effect of any given area on a child's potential outcomes. The causal effect of growing up in a given CZ c will generally differ from the mean predicted impact $\beta_0 \bar{y}_{pds}$ based on permanent residents' outcomes because the degree of selection and causal effects can vary across areas. We build on the methodology developed in this paper to estimate the causal effect of growing up in each CZ in the second paper in this series (Chetty and Hendren (2016)).

III.C Estimating Exposure Effects in Observational Data

We estimate exposure effects by studying families who move across CZs with children of different ages in observational data. In observational data, the error term θ_i in (3) will generally be correlated with $\bar{y}_{pds}$. For instance, parents who move to a good area may have latent ability or wealth that produces better child outcomes. Estimating (3) in an observational sample of families who move exactly once yields a regression coefficient

$$b_m = \beta_m + \delta_m,$$

where $\delta_m = \frac{cov(\theta_i, \bar{y}_{pds})}{var(\bar{y}_{pds})}$ is a standard selection effect that measures the extent to which parental inputs and other determinants of children's outcomes for movers covary with permanent residents' outcomes. Fortunately, the identification of exposure effects does not require that *where* people move is orthogonal to child's potential outcomes. Instead, it requires that *timing* of moves to better areas is orthogonal to children's potential outcomes, as formalized in the following assumption.

Assumption 1. Selection effects do not vary with the child's age at move: $\delta_m = \delta$ for all m.

Assumption 1 allows for the possibility that the families who move to better areas may differ from those who move to worse areas, but requires that the extent of such selection does not vary with the age of the child when the parent moves. Under this assumption, we immediately obtain consistent estimates of exposure effects $\gamma_m = \beta_m - \beta_{m+1} = b_m - b_{m+1}$ because the selection effect δ cancels out when estimating the exposure effect. We can go further and estimate the selection effect δ itself by studying the outcomes of children whose families move *after* their income is measured, e.g. at age $a \geq 30$ if income is measured at age $T = 30$. Because moves at age $a > T$ cannot have a causal effect on children's outcomes at age 30, $b_m = \delta$ for $m > T$ under Assumption 1. Using the estimated selection effect, we can identify the causal effect of moving to a better area at age m as $\beta_m = b_m - b_{T+1}$ and thereby identify β_0, the total causal effect of growing up in area from birth.

Of course, Assumption 1 is a strong restriction that may not hold in practice. We therefore evaluate its validity in detail after presenting a set of baseline estimates in the next section.

IV Baseline Estimates of Childhood Exposure Effects

This section presents our baseline estimates of exposure effects. We begin with a set of semi-parametric estimates that condition on origin fixed effects and correspond most closely to the hypothetical experiment described in Section III.B. We then present estimates from parametric models that show how movers' outcomes can be parsimoniously modeled as a linear combination of the outcomes of permanent residents in origins and destination. Finally, we present results from variants of the baseline specification to assess the sensitivity of our estimates to specification choices.

In our baseline analysis, we focus on children whose parents moved across CZs exactly once between 1996 and 2012 and are observed in the destination CZ for at least two years. We also restrict attention to families who moved at least 100 miles to exclude moves across CZ borders that do not reflect a true change of neighborhood and limit the sample to CZs with populations above 250,000 to mitigate measurement error in the estimates of permanent residents' outcomes $\bar{y}_{pds}$. We present estimates that include families who move more than once in Section 6 and show that the findings are robust to alternative cutoffs for population size and distance in Online Appendix Table I.

In prior work (Chetty et al. 2014), we found that the intergenerational correlation between parents' and children's incomes stabilizes when children turn 30, as college graduates experience steeper wage growth in their 20s (Haider and Solon 2006). Measuring income at age 30 limits us to estimating exposure effects only after age 15 given the time span of our dataset.[19] Fortunately, measuring income at earlier ages (from 24-30) turns out not to affect the exposure effect estimates. The reason is that our estimates of b_m correlate the incomes of children who move with the incomes of permanent residents in the destination measured at the *same age*. The incomes of permanent residents serve as goalposts that allow us to measure the degree of convergence in incomes at any age, even before we observe children's permanent income. For example, if a given area c sends many children to college and therefore generates relatively low incomes at age 24, we will obtain a *higher* estimate of b_m if a child who moves to area c has a *low* level of income at age 24. We

[19] The most recent birth cohort for which we observe income at age 30 (in 2012) is the 1982 cohort; since our data begin in 1996, we cannot observe moves before age 15.

therefore measure income at age 24 in our baseline specifications to estimate exposure effects for the broadest age range.[20]

IV.A Semi-Parametric Estimates

To begin, consider the set of children whose families moved when they were exactly m years old. We analyze how these children's earnings are related to those of the permanent residents in their destination CZ using the following linear regression:

$$y_i = \alpha_{qos} + b_m \Delta_{odps} + \varepsilon_{1i}, \tag{4}$$

where y_i denotes the child's household income rank at age 24, α_{qos} is a fixed effect for the origin CZ o by parent income decile q by birth cohort s and $\Delta_{odps} = \bar{y}_{pds} - \bar{y}_{pos}$ is the difference in predicted income rank (at age 24) of permanent residents in the destination versus origin for the relevant parent income rank p and birth cohort s. Equation (4) can be interpreted as an observational analog of the specification in (3) that we would ideally estimate in experimental data.[21]

Figure III presents a non-parametric binned scatter plot corresponding to the regression in (4) for children who move at age $m = 13$. To construct the figure, we first demean both y_i and Δ_{odps} within the parent decile (q) by origin (o) by birth cohort (s) cells in the sample of movers at age $m = 13$ to construct residuals: $y_i^r = y_i - E[y_i|q, o, s, m]$ and $\Delta_{odps}^r = \Delta_{odps} - E[\Delta_{odps}|q, o, s, m]$. We then divide the Δ_{odps}^r residuals into twenty equal-size groups (ventiles) and plot the mean value of y_i^r vs. the mean value of Δ_{odps}^r in each bin.

Figure III shows that children who move to areas where children of permanent residents earn more at age 24 themselves earn more when they are 24. The relationship between y_i and Δ_{odps} is linear. The regression coefficient of $b_{13} = 0.618$, estimated in the microdata using (4), implies that a 1 percentile increase in $\bar{y}_{pds}$ is associated with a 0.629 percentile increase in y_i for the children who move at age 13.

Building on this approach, we estimate analogous regression coefficients b_m for children whose parents move at each age m from 9 to 30. We estimate $\{b_m\}$ using the following regression specification:

[20] We do not study income before age 24 because a large fraction of children are enrolled in college at earlier ages and because we find that exposure effects persist until age 23 when income is measured at any point between 24 and 30. We study college attendance as a separate outcome in Section VI.

[21] We use parent income deciles rather than percentiles to define the fixed effects α_{qos} to simplify computation; using finer bins to measure parent income groups has little effect on the estimates. Conditional on parent percentile, origin, and birth cohort, the variation in Δ_{odps} is entirely driven by variation in the destination outcomes ($\bar{y}_{pds}$). Hence, b_m is identified from variation in $\bar{y}_{pds}$, as in (3), up to the approximation error from using parent deciles instead of exact percentiles.

$$y_i = \alpha_{qosm} + \sum_{m=9}^{30} b_m I(m_i = m)\Delta_{odps} + \sum_{s=1980}^{1987} \kappa_s I(s_i = s)\Delta_{odps} + \varepsilon_{2i}, \tag{5}$$

where α_{qosm} is an origin CZ by parent income decile by birth cohort by age at move fixed effect and $I(x_i = x)$ is an indicator function that is 1 when $x_i = x$ and 0 otherwise. This specification generalizes (4) by fully interacting the age at move m with the independent variables in (4). In addition, we permit the effects of Δ_{odps} to vary across birth cohorts (captured by the κ_s coefficients) because our ability to measure parent's locations during childhood varies across birth cohorts. We observe children's locations starting only at age 16 for the 1980 cohort, but starting at age 8 for the 1988 cohort. This leads to greater measurement error in Δ_{odps} for earlier birth cohorts, which could potentially confound our estimates of b_m since the distribution of ages at move is unbalanced across cohorts. By including cohort interactions, we identify $\{b_m\}$ from within-cohort variation in ages at move.[22]

Figure IVa plots estimates of b_m from (4). The estimates exhibit two key patterns: *selection effects* after age 24 and *exposure effects* before age 24. First, the fact that $b_m > 0$ for $m > 24$ is direct evidence of selection effects ($\delta_m > 0$), as moves after age 24 cannot have a causal effect on earnings at 24. Families who move to better areas have children with better unobservable attributes. The degree of selection δ_m does not vary significantly with m above age 24: regressing b_m on m for $m \geq 24$ yields a statistically insignificant slope of 0.001 (s.e. $= 0.011$). This result is consistent with Assumption 1, which requires that selection does not vary with the child's age at move. The mean value of δ_m for $m \geq 24$ is $\delta = 0.126$, i.e. families who move to an area where permanent residents have 1 percentile better outcomes have 0.126 percentile better outcomes themselves purely due to selection effects. Assumption 1 allows us to extrapolate the selection effect of $\delta = 0.126$ back to earlier ages $m < 24$, as shown by the dashed line in Figure 1, and thereby identify causal exposure effects at earlier ages.

This leads to the second key pattern in Figure IVa, which is that the estimates of b_m decline steadily with the age at move m for $m < 24$. Under Assumption 1, this declining pattern constitutes evidence of an exposure effect, i.e. that moving to a better area earlier in childhood generates larger long-term gains. The linearity of the relationship between b_m and the age at move m in Figure

[22] To avoid collinearity, we omit the most recent birth cohort (1988 for income at age 24) interaction with Δ_{odps}. The inclusion of the cohort interactions has little impact on the estimates obtained from (5), as shown in Table II, Column (5), presumably because the fraction of the variance in Δ_{odps} due to measurement error is small. The cohort interactions play a larger role in specifications that include family fixed effects, as the portion of the residual variance in Δ_{odps} that is due to measurement error is larger in those specifications.

IVa below age 23 implies that the exposure effect $\gamma_m = b_{m+1} - b_m$ is approximately constant with respect to age at move m. Regressing $\hat{b}_m$ on m for $m < 24$, we estimate an average annual exposure effect of$\gamma = 0.044$ (s.e. $= 0.0018$). That is, the outcomes of children who move converge to the outcomes of permanent residents of the destination area at a rate of 4.4% per year of exposure until age 23.[23]

Because some children do not move with their parents, the estimates of b_m in (5) should be interpreted as intent-to-treat (ITT) estimates, in the sense that they capture the causal effect of moving (plus the selection effect) for children whose *parents* moved at age m. We can obtain treatment-on-the-treated (TOT) estimates for the children who move themselves by inflating the ITT estimates by the fraction of children who move at each age m. In Online Appendix Figure IV, we show that the TOT estimate of the exposure effect is $\gamma^{TOT} = 0.040$. This estimate is very similar to our baseline estimate because virtually all children move with their parents below age 18 and roughly 60% of children move with their parents between ages 18-23. Because the treatment effects converge toward zero as the age at move approaches 23, inflating the coefficients by 1/0.6 at later ages has little impact on exposure effect estimates.

IV.B Parametric Estimates

Equation (5) includes more than 200,000 fixed effects (α_{qosm}), making it difficult to estimate in smaller samples and introduce additional controls such as family fixed effects. As a tractable alternative to controlling non-parametrically for parent income, origin, birth cohort, and age at move using fixed effects, we now estimate a more parsimonious model in which we control parametrically for two key factors captured by the α_{qosm} fixed effects: (1) the quality of the origin location, which we model by interacting the predicted outcomes for permanent residents in the origin with birth cohort fixed effects and (2) disruption costs of moving that may vary with the age at move and parent income, which we model using age at move fixed effects linearly interacted with parent

[23]Figure IVa is identified from variation in movers' destinations holding their origin fixed. An alternative approach is to exploit variation in origins, holding destinations fixed. Online Appendix Figure III presents estimates of b_m identified from variation in origins by replacing the origin (α_{qosm}) fixed effects in (5) with destination (α_{qdsm}) fixed effects. The resulting estimates yield a qualitative pattern that is the mirror image of those in Figure IVa: the later the family moves to the destination, the more the child's outcomes match the permanent residents in the origin, up to age 23. The estimated exposure effect of 0.03 is smaller than the estimates above because we measure children's origins with greater error than destinations, as our location data is left-censored. This is why we focus on variation in destinations in most of our specifications.

income percentile p_i. This leads to the following regression specification:

$$\begin{aligned} y_i &= \sum_{s=1980}^{1988} I(s_i = s)(\alpha_s^1 + \alpha_s^2 \bar{y}_{pos}) + \sum_{m=9}^{30} I(m_i = m)(\zeta_m^1 + \zeta_m^2 p_i) \\ &+ \sum_{m=9}^{30} b_m I(m_i = m)\Delta_{odps} + \sum_{s=1980}^{1987} \kappa_s^d I(s_i = s)\Delta_{odps} + \varepsilon_{3i}, \end{aligned} \tag{6}$$

The first two terms of this specification control for origin quality and disruption effects. The third term represents the exposure effects of interest, and as in equation (5), the fourth consists of cohort interactions with Δ_{odps} to control for differential measurement error across cohorts.[24]

Figure IVb plots the coefficients $\{b_m\}$ obtained from estimating (6). The coefficients are very similar to those obtained from the more flexible specification used to construct Figure IVa. Regressing the b_m coefficients on m for $m \leq 23$, we obtain an average annual exposure effect estimate of $\gamma = 0.038$ (s.e. 0.002). The exposure effect estimate is similar to that obtained from the fixed effects specification because controlling for the quality of the origin using the permanent residents' outcomes is adequate to account for differences in origin quality. Put differently, movers' outcomes can be modeled as a weighted average of the outcomes of permanents residents in the origin and destination, with weights reflecting the amount of childhood spent in the two places.

When measuring income at age 24, we cannot determine whether b_m stabilizes after age 24 because moving after age 24 has no causal effect on earnings or because we measure income at that point. In Online Appendix Figure II, we replicate the analysis measuring income at ages 24, 26, 28, and 30. All four series display very similar patterns of exposure effects in the overlapping age ranges, showing that our estimates of b_m are insensitive to the age at which we measure children's incomes in adulthood. In particular, all four series decline linearly at a rate of approximately $\gamma = 0.04$ until age 23 and are flat thereafter. These results imply that neighborhood exposure before age 23 is what matters for earnings in subsequent years.

The kink at age 23 motivates the baseline regression specification that we use for much of our analysis. We parameterize both the exposure and selection effects shown in Figure IV linearly, replacing the non-parametric $\sum_{m=9}^{30} b_m I(m_i = m) \Delta_{odps}$ term in (6) with two separate lines above

[24] In addition to having much fewer fixed effects, this specification uses variation in both the quality of the origin ($\bar{y}_{pos}$) and the destination ($\bar{y}_{pds}$) to identify $\{b_m\}$. In contrast, the semi-parametric model in (5) is identified purely from variation in destinations because it includes origin fixed effects. Estimating a parametric model that identifies $\{b_m\}$ from variation in destinations by controlling for outcomes of permanent residents in the origin interacted with the age of the child at the time of the move ($\sum_{m=9}^{30} b_m I(m_i = m) y_{pos}$) yields very similar estimates.

and below ages 23:

$$y_i = \sum_{s=1980}^{1988} I(s_i = s)(\alpha_s^1 + \alpha_s^2 \bar{y}_{pos}) + \sum_{m=9}^{30} I(m_i = m)(\zeta_m^1 + \zeta_m^2 p_i) + \sum_{s=1980}^{1987} \kappa_s^d I(s_i = s)\Delta_{odps} \quad (7)$$
$$+ I(m_i \leq 23)\left(b_0 + (23 - m_i)\gamma\right)\Delta_{odps} + I(m_i > 23)(\delta + (23 - m_i)\delta')\Delta_{odps} + \varepsilon_{3i},$$

Estimating this specification directly in the microdata yields an average annual exposure effect $\gamma = 0.040$ (s.e. 0.002), as shown in Column 1 of Table II.[25]

IV.C Alternative Specifications

The model in (7) is one of many potential parametric specifications one could use to estimate exposure effects. In Table II, we show that several natural variants of (7) all yield very similar estimates of γ.

We begin in Columns 2 and 3 of Table II by showing that estimating γ using data only up to age 18 or 23 – i.e., excluding the data at older ages that identifies the selection effect in (7) – yields similar estimates of γ. Column 4 shows that restricting the sample to children claimed in the destination CZ (to ensure that the children moved with the parents) also yields similar estimates. Column 5 shows that dropping the cohort interactions, $\sum_{s=1980}^{1988} I(s_i = s)\alpha_s^2 \bar{y}_{pos}$ and $\sum_{s=1980}^{1988} \kappa_s^d I(s_i = s)\Delta_{odps}$, in (7) has little effect on the results. Column 6 shows we obtain similar estimates when using a child's individual income rank as the outcome, y_i, as opposed to household income rank.

The exposure effect estimates also remain roughly similar across subgroups (Online Appendix Table II) We find similar but slightly higher estimates for children from above-median versus below-median income families ($\gamma = 0.047$ versus $\gamma = 0.031$). We also find that moves to better and worse areas have symmetric effects. Standard models of learning predict that moving to a better area will improve outcomes but moving to a worse area will not. In practice, we find little evidence of such an asymmetry: if anything, the point estimate of exposure effects for negative moves is larger. These findings suggest that what matters for children's mean long-term outcomes is the total duration of exposure to a better environment rather than a permanent effect obtained from short-term exposure.

To distinguish the role of childhood environment from differences caused by variation in labor market conditions or local costs of living across areas, in Column 7 we add fixed effects for the

[25] This coefficient differs slightly from the coefficient of $\gamma = 0.038$ that we obtain when regressing the coefficients b_m on m in Figure IVb because estimating the regression in the microdata puts different weights on each age (as we have more data at older ages), while estimating the regression using the b_m coefficients puts equal weight on all ages.

CZ in which the child lives at age 24 (when income is measured) to the baseline model. This specification compares the outcomes of children who live in the same labor market in adulthood but grew up in different neighborhoods. We obtain an annual exposure effect of $\gamma = 0.031$ in this specification, indicating that the majority of the exposure effect in our baseline specification is driven by differences in exposure to a better childhood environment, holding fixed labor market conditions.[26] This conclusion is consistent with the fact that moving to an area where permanent residents have higher earnings just before entering the labor market (e.g., in one's early 20s) has little effect on earnings, as shown in Figure IV.

Our baseline model only includes families who move across CZs exactly once during our sample frame (1996-2012). In Online Appendix C, we generalize our approach to include families who move more than once by estimating a variant of (7) that replaces Δ_{odps} with a duration-weighted measure of exposure to different areas over childhood. We obtain an annual exposure effect estimate of $\gamma = 0.039$ from this multiple movers specification. The similarity of this coefficient to our estimate for one-time movers implies neighborhoods affect children's long-term outcomes through an exposure (or dosage) effect rather than "critical age" effects in which children's long-term outcomes are a function of the specific ages at which they live in a neighborhood.

Finally, we replicate the analysis in Table II at the county level in Online Appendix Table IV. We obtain slightly smaller exposure effect estimates of $\gamma \simeq 0.035$ at the county level. This is consistent with the hypothesis that selection effects account for a larger fraction of the variance in permanent resident outcomes at smaller geographies. This is plausible insofar as families are more likely to sort geographically (e.g., to better school districts) within rather than across metro areas.

IV.D Summary

The results in this section yield three lessons. First, place matters: children who move at earlier ages to areas where prior residents have higher earnings earn more themselves as adults. Second, place matters via *childhood* exposure. Every year of exposure to the better area during childhood contributes to higher earnings in adulthood. Third, each year of childhood exposure matters roughly equally. The returns to growing up in a better neighborhood remain substantial well beyond early childhood. All of these results are predicated on our assumption that selection effects do not vary with the child's age at move. We evaluate this critical assumption in the next section.

[26] This specification likely over-adjusts for differences in labor market conditions and underestimates γ because the CZ in which the child resides as an adult is itself an endogenous outcome that is likely related to the quality of a child's environment. For example, one of the effects of growing up in a good area may be an increased probability of getting a high-paying job in another city.

V Validation of Baseline Design

We assess the validity of the key identifying assumption – that the potential outcomes of children who move to better vs. worse areas do not vary with the age at which they move – using a series of tests that focus on different forms of selection and omitted variable bias. To organize this analysis, it is useful to partition the unobserved determinant of children's outcomes, represented by θ_i in equation (3), into two components: a component $\bar{\theta}_i$ that reflects inputs that are *fixed* within families, such as parent genetics and education, and a residual component $\tilde{\theta}_i = \theta_i - \bar{\theta}_i$ that may vary over time within families, such as parents' jobs.

We implement four tests for bias in this section. First, we address bias due to selection on fixed family factors by $\bar{\theta}_i$ by comparing siblings' outcomes. Second, we control for changes in parents' income and marital status, two key time-varying factors of $\tilde{\theta}_i$ that we observe in our data. Our remaining tests focus on *unobservable* time-varying factors, such as changes in wealth, that may have triggered a move to a better area. In our third set of tests, we isolate moves that occur due to local area displacement shocks that induce many families to move. Finally, we conduct a set of outcome-based placebo (overidentification) tests of the exposure effect model, exploiting heterogeneity in place effects across subgroups to generate sharp testable predictions about how children's outcomes should change when they move to different areas. In our view, this last approach, although least conventional, provides the most compelling evidence that the identifying assumption holds and that neighborhoods have causal exposure effects on children's long-term outcomes.

V.A Sibling Comparisons

If families with better unobservables (higher $\bar{\theta}_i$) move to better neighborhoods at earlier ages, Assumption 1 would be violated and our estimated exposure effect $\hat{\gamma}$ would be biased upward. We control for differences in such family-level factors $\bar{\theta}_i$ by including family fixed effects when estimating (6).[27] For example, consider a family that moves to a better area with two children,

[27] The idea of using sibling comparisons to better isolate neighborhood effects dates was discussed in the seminal review by Jencks and Mayer (1990). Plotnick and Hoffman (1996) and Aaronson (1998) implement this idea using data on 742 sibling pairs from the Panel Study of Income Dynamics, but reach conflicting conclusions due to differences in sample and econometric specifications. More recently, Andersson et al. (2013) use a siblings design to estimate the impact of vouchers and public housing provision. Our analysis also relates to papers that seek to identify critical periods by studying immigrants using sibling comparisons (Basu (2010); van den Berg et al. (2014)). Our approach differs from these studies in that we focus on how the difference in siblings' outcomes *covaries* with the outcomes of permanent residents in the destination neighborhood, whereas the immigrant studies estimate the mean difference in siblings' outcomes as a function of the age gap. This allows us to separate the role of neighborhood exposure from changes in the family that also generate exposure-dependent differences across siblings, such as changes in wealth

who are ages m_1 and m_2 at the time of the move. When including family fixed effects, the exposure effect γ is identified by the extent to which the *difference* in sibling's outcomes, $y_1 - y_2$, covaries with their age gap interacted with the quality of the destination CZ, $(m_1 - m_2)\Delta_{odps}$.[28]

Figure Va replicates Figure IVb, adding family fixed effects to equation (6). The linear decline in the estimated values of b_m until age 23 is very similar to that in the baseline specification. Children who move to a better area at younger ages have better outcomes than their older siblings. Regressing the b_m coefficients on m for $m \leq 23$ yields an average annual exposure effect estimate of $\gamma = 0.043$ (s.e. 0.03), very similar to our estimates above. The selection effect (i.e., the level of b_m after age 24) falls from $\delta = 0.23$ in the baseline specification to $\delta = 0.01$ (not significantly different from zero) with family fixed effects.[29] The introduction of family fixed effects thus reduces the *level* of the b_m coefficients by accounting for differential selection in which types of families move to better vs. worse areas, but does not affect the *slope* of the b_m coefficients. This is precisely what we should expect if selection effects in where families choose to move do not vary with children's ages at the point of move, as required by Assumption 1.

Column 8 of Table II shows that adding family fixed effects to the linear specification in equation (7) and estimating the model directly on the micro data yields an estimate of $\gamma = 0.044$. Other variants of this regression specification, analogous to those in Columns 2-7 of Table II, all yield very similar estimates of γ, with one exception: excluding cohort interactions with $\bar{y}_{pos}$ and Δ_{odps}, as in Column 5, yields $\gamma = 0.031$, slightly lower than the other estimates (Column 9). This attenuation occurs because the level of the selection effect δ is smaller for more recent cohorts, as children's origins are measured more accurately in more recent cohorts. Since the specification with family fixed effects is identified purely from comparisons across birth cohorts, this differential measurement error across cohorts biases the estimate of γ downward unless one allows for cohort-specific interactions.[30]

when a family moves to a new country.

[28]Since siblings of different ages must be in different cohorts, β is also partly identified from variation in the outcomes of permanent residents in differing cohorts s. We focus on this variation across cohorts in Section V.D below.

[29]δ is identified even with family fixed effects because Δ_{odps} varies across birth cohorts.

[30]This attenuation due to measurement error is particularly large when we include CZs with smaller population, where Δ_{odps} is measured with greater error. See Online Appendix B for a discussion of the impact of measurement error on the family fixed effect specifications.

V.B Controls for Time-Varying Observables

The research design in Figure Va accounts for bias due to fixed differences in family inputs $\bar{\theta}_i$, but it does not account for time-varying inputs $\tilde{\theta}_i$. For example, moves to better areas may be triggered by events such as job promotions that directly affect children's outcomes in proportion to their time of exposure to the destination. Such shocks could bias our estimate of β upward even with family fixed effects.

Prior research has focused on parents' income and marital status as two of the key determinants of children's outcomes in adulthood. We can directly control for these two time-varying factors in our data, as we observe parents' incomes and marital status in each year from 1996-2012. We control for the effects of changes in income around the move when estimating (6) by including controls for the change in the parent's income rank from the year before to the year after the move interacted with indicators for the child's age at move. The interactions with age at move permit the effects of income changes to vary with the duration of childhood exposure to higher vs. lower levels of family income. Similarly, we control for the impact of changes in marital status by interacting indicators for each of the four possible changes in marital status of the mother in the year before vs. after the move (married to unmarried, unmarried to married, unmarried to unmarried, and married to married) with indicators for the child's age at move.

Figure Vb replicates Figure Va, controlling for all of these variables in addition to family fixed effects. Controlling for changes in parent income and marital status has little effect on the estimates of $\{b_m\}$. The estimates of $\gamma = 0.042$ and $\delta = 0.015$ are virtually identical to those when we do not control for these time-varying factors. Column 10 of Table II confirms that including these controls in a linear regression specification estimated on the micro data yields similar estimates.

These results show that changes in income and family structure are not a significant source of bias in our design. However, other unobserved factors could still be correlated with moving to a better or worse area in a manner that generates omitted variable bias. The fundamental identification problem is that any unobserved shock that induces child i's family to move to a different area could be correlated with parental inputs θ_i. These changes in parental inputs could potentially increase the child's earnings y_i in proportion to the time spent in the new area even in the absence of neighborhood effects. For example, a wealth shock might lead a family to both move to a better neighborhood and increase investments in the child in the years after the shock, which could improve y_i in proportion to exposure time independent of neighborhood effects. In the next two subsections, we address concerns about bias due to such unobserved, time-varying factors.

V.C Displacement Shocks

One approach to accounting for unobservable shocks is to identify moves where we have some information about the shock that precipitated the move. Suppose we identify families who were forced to move from an origin o to a nearby destination d because of an exogenous displacement shock such as a natural disaster. Such displacement shocks can induce differential changes in neighborhood quality as measured by permanent residents' outcomes (Δ_{odps}). For instance, Hurricane Katrina displaced families from New Orleans (an area with relatively poor outcomes compared to surrounding areas), leading to an increase in neighborhood quality for displaced families ($\Delta_{odps} > 0$). In contrast, Hurricane Rita hit Houston, an area with relatively good outcomes, and may have reduced neighborhood quality ($\Delta_{odps} < 0$). If these displacement shocks do not have direct exposure effects on children that are correlated withΔ_{odps} – e.g., the direct effects of the disruption induced by hurricanes does not covary with neighborhood quality changes – then Assumption 1 is satisfied and we obtain unbiased estimates of the exposure effect γ. Conceptually, by isolating a subset of moves caused by known exogenous shocks, we can more credibly ensure that changes in children's outcomes are not driven by unobservable factors.[31]

To operationalize this approach, we first identify displacement shocks based on population outflows at the ZIP code level. Let K_{zt} denote the number of families who leave ZIP code z in year t in our full sample and $\bar{K}_z$ mean outflows between 1996 and 2012. We define the shock to outflows in year t in ZIP z as $k_{zt} = K_{zt}/\bar{K}_z$.[32]

Though many of the families who move in subsamples with large values of k_{zt} do so for exogenous reasons, their destination d is still the result of an endogenous choice that could lead to bias. For example, families who choose to move to better areas (higher $\bar{y}_{pds}$) when induced to move by an exogenous shock might also invest more in their children. To eliminate potential biases arising from endogenous choices of destinations, we isolate variation arising purely from the *average* change in neighborhood quality for individuals who are displaced. Let $E[\Delta_{odps}|q, z]$ denote the change in the mean predicted outcome in the destination CZ relative to the origin CZ for individuals in origin

[31]This research design is closely related to Sacerdote's (2012) analysis of the effects of Hurricanes Katrina and Rita on student test score achievement. Although we use similar variation, we do not focus on the direct effects of the displacement itself, but rather on how children's long-term outcomes changed in relation to the outcomes of permanent residents in the destination to which they were displaced.

[32]Searches of historical newspaper records for cases with the highest outflow rates k_{zt} reveal that they are frequently associated with events such as natural disasters or local plant closures. Unfortunately, there is insufficient power to estimate exposure effects purely from events identified in newspapers.

ZIP code z and parent income decile q (averaging over all years in the sample, not just the year of the shock). We instrument for the difference in predicted outcomes in each family's destination relative to origin (Δ_{odps}) with $E[\Delta_{odps}|q,z]$ and estimate the linear specification in (7) using 2SLS to obtain IV estimates of exposure effects, γ_{IV}.

Figure VI presents the results of this analysis. To construct this figure, we take ZIP-year cells with above-median outflows ($k_{zt} > 1.17$) and divide them into (population-weighted) centiles based on the size of the shock k_{zt}.[33] The first point in Figure VI shows the 2SLS estimate of the annual exposure effect γ_{IV} using all observations with k_{zt} greater than its median value (1.17). The second point shows the estimate of γ_{IV} using all observations with k_{zt} at or above the 52nd percentile. The remaining points are constructed in the same way, increasing the threshold by 2 percentiles at each point, with the last point representing an estimate of γ_{IV} using data only from ZIP codes in the highest two percentiles of outflow rates. The dotted lines show a 95% confidence interval for the regression coefficients.

If the baseline estimates were driven entirely by selection, γ_{IV} would fall to 0 as we limit the sample to individuals who are more likely to have been induced to move because of an exogenous displacement shock. But the coefficients remain quite stable at$\gamma_{IV} \simeq 0.04$ even when we restrict to moves that occurred as part of large displacements. That is, when we focus on families who move to a better area for what are likely to be exogenous reasons, we continue to find that children who are younger at the time of the move earn more as adults.

These findings support the view that our baseline estimates of exposure effects capture the causal effects of neighborhoods rather than other unobserved factors that change when families move. Moreover, they indicate that the treatment effects of moving to a different area are similar for families who choose to move for idiosyncratic reasons and families who are exogenously displaced by an aggregate shock. This suggests that the effects identified in our baseline population of movers who choose to move have potential external validity to a broader set of families who may not otherwise be choosing to move.

V.D Outcome-Based Placebo Tests

As a second approach to account for potential time-varying observables, we implement placebo tests that exploit the heterogeneity in place effects across subgroups. We exploit variation along three dimensions: birth cohorts, quantiles of the income distribution, and child gender. The causal

[33] To ensure that large outflows are not driven by areas with small populations, we exclude ZIP-year cells with less than 10 children leaving in that year.

exposure effect model predicts precise convergence of a child's earnings to the place effect for his or her *own* subgroup. In contrast, we argue below that omitted variable and selection models would not generate such subgroup-specific convergence under plausible assumptions about parents' information sets and preferences. The heterogeneity of place effects thus gives us a rich set of overidentification restrictions to test whether neighborhoods have causal exposure effects.[34] We consider each of the three dimensions of heterogeneity in turn.

Birth Cohorts. Although place effects are generally very stable over time, outcomes in some areas (such as Oklahoma City, OK) have improved over time, while others (such as Sacramento, CA) have gotten worse.[35] Such changes in place effects could occur, for instance, because of changes in the quality of local schools or other area-level characteristics that affect children's outcomes. We exploit this heterogeneity across birth cohorts to test the exposure effect model.

Under the causal exposure effect model, when a child's family moves to destination d, the change in permanent residents' outcomes $\Delta_{odp,s(i)}$ for that child's own birth cohort $s(i)$ should predict his or her outcomes more strongly than the change in outcomes Δ_{odpt} for other cohorts $t \neq s(i)$. In contrast, it is unlikely that other time-varying unobservables θ_i will vary sharply across birth cohorts s in association with Δ_{odps} because the fluctuations across birth cohorts are realized only in adulthood and thus cannot be directly observed at the time of the move.[36] Therefore, by testing whether exposure effects are predicted by a child's own vs. surrounding cohorts, we can assess the importance of bias due to unobservables.

We implement this analysis by estimating the linear specification in (7), replacing the change in permanent residents' outcomes for the child's own cohort, $\Delta_{odps(i)}$, with analogous predictions for another nearby cohorts, Δ_{odpt} (see Online Appendix D for details). The series in red triangles in Figure VII plots $\tilde{\gamma}_t$ when we estimate (7) using the predicted outcome for a single cohort $t \in \{s(i)-4, s(i)+4\}$. The estimates of $\tilde{\gamma}_t$ are similar to our baseline estimate of $\gamma = 0.040$ for the leads and lags, consistent with the high degree of serial correlation in place effects. The series in blue circles plots analogous coefficients $\tilde{\gamma}_t$ when all the cohort-specific predictions from the four years before to the four years after the child's own cohort are included *simultaneously*. In this specification, the coefficients on the placebo exposure effects ($\tilde{\gamma}_t$ for $t \neq 0$) are all very close to zero

[34] In addition to being useful for identification, these results are also of direct interest in understanding the heterogeneity of place effects across subgroups.

[35] The autocorrelation of $\bar{y}_{pcs}$ with $\bar{y}_{pc,s-1}$ across children's birth cohorts is 0.95 at the 25th percentile of the parent income distribution.

[36] For instance, a family that moves with a 10 year old child will not observe $\bar{y}_{pds}$ for another 14 years (if income is measured at age 24).

and not statistically significant[37]. However, the exposure effect estimate for the child's own cohort, $t = 0$ remains at $\gamma = 0.04$ even when we control for the surrounding cohorts' predictions and is significantly different from the estimates of $\tilde{\gamma}_t$ for $t \neq 0$ ($p < 0.001$).

Since is it unlikely that a correlated shock – such as a change in wealth when the family moves – would covary precisely with cohort-level differences in place effects, the evidence in Figure VII strongly supports the view that the change in children's outcomes is driven by causal effects of exposure to a different place. Formally, assume that if unobservables θ_i are correlated with exposure to a given cohort $s(i)$'s place effect, they must also be correlated with exposure to the place effects of adjacent cohorts t:

$$Cov(\theta_i, m\Delta_{odp,s(i)}|X) > 0 \Rightarrow Cov(\theta_i, m\Delta_{odpt}|X, m\Delta_{odp,s(i)}) > 0, \tag{8}$$

where X represents the vector of fixed effects and other controls in (7). Under this assumption, the findings in Figure VII imply that our estimates of γ reflect causal neighborhood effects (which are cohort-specific) rather than omitted variables, which are *not* cohort-specific under (8).

Quantiles: Distributional Convergence. Places differ not only in children's mean outcomes, but also in the *distribution* of children's outcomes. For example, children who grow up in low-income families in Boston and San Francisco have comparable mean ranks, but children in San Francisco are more likely to end up in the tails of the income distribution than those in Boston. If neighborhoods have causal exposure effects, we would expect convergence in mover's outcomes not just at the mean but across the entire distribution in proportion to exposure time. In contrast, it is less plausible that omitted variables such as wealth shocks would perfectly replicate the distribution of outcomes of permanent residents in each CZ.[38] Therefore, testing for quantile-specific convergence can distinguish the causal exposure effect model from omitted variable explanations.

To implement these tests, we begin by constructing predictions of the probability of having an income in the upper or lower tail of the national income distribution at age 24 for children of permanent residents in each CZ c. We regress an indicator for a child being in the top or bottom 10% of the distribution on parent rank separately in each CZ using an equation analogous to (1), including a quadratic term in parental income to account for the nonlinearities in tail outcomes identified in Chetty et al. (2014). We then calculate the predicted probability of being below

[37] A joint test that all $\tilde{\gamma}_t = 0$ for all $t \neq s\,(i)$ yields a p-value of 0.251.

[38] Families are unlikely to be able to forecast their child's eventual quantile in the income distribution, making it difficult to sort precisely on quantile-specific neighborhood effects. Even with such knowledge, there is no ex-ante reason to expect unobserved shocks such as changes in wealth to have differential and potentially non-monotonic effects across quantiles, in precise proportion to the outcomes in the destination.

the 10th percentile, π^{10}_{pcs}, and above the 90th percentile π^{90}_{pcs} using the fitted values from these regressions, as in (2).[39]

In Table III, we estimate exposure effect models analogous to (7) using these distributional predictions instead of mean predictions. In Columns 1-3, the dependent variable is an indicator for having income in the top 10% of the income distribution. Column 1 replicates the baseline specification in (7), using $\Delta^{90}_{odps} = \pi^{90}_{pds} - \pi^{90}_{pos}$ instead of the mean prediction $\Delta_{odps} = \bar{y}_{pds} - \bar{y}_{pos}$.[40] We obtain an exposure effect estimate of $\gamma = 0.043$ per year in this specification. Column 2 uses the change in the predicted mean rank, Δ_{odps}, instead. Here, we obtain a significant, positive estimate of 0.022, as expected given the high degree of correlation in place effects across quantiles: places that push children into the top 10% also tend to improve mean outcomes. In Column 3, we include both the quantile prediction Δ^{90}_{odps} and the mean prediction Δ_{odps}, identifying the coefficients purely from differential variation across quantiles within CZs. The coefficient on the quantile prediction remains unchanged at approximately $\gamma = 0.04$, while the coefficient on the mean prediction is not significantly different from 0.[41]

Columns 4-6 of Table III replicate Columns 1-3, using an indicator for being in the bottom 10% as the dependent variable and the prediction for being in the bottom decile, Δ^{10}_{odps}, instead of Δ^{90}_{odps} as the key independent variable. As in the upper tail, children's probabilities of being in the lower tail of the income distribution are fully determined by the quantile-specific prediction rather than the mean prediction.

In sum, we find evidence of distributional convergence: controlling for mean outcomes, children's outcomes converge to predicted outcomes in the destination across the distribution in proportion to exposure time, at a rate of approximately 4% per year.[42] Since omitted variables such as wealth shocks would be unlikely to generate such distributional convergence, this finding again supports the view that the convergence in mover's outcomes is driven by causal effects of place. Formally, assume that if unobservables θ_i are correlated positively with exposure to place effects on upper or lower tail outcomes π^{q}_{pcs}, they must also be correlated with exposure to the place effects on mean

[39] Since more than 10% of children have 0 income at age 24, we define the lower-tail outcome as an indicator for being unemployed (measured by having no W-2).

[40] See Online Appendix D for the precise regression specifications.

[41] We present binned scatter plots verifying that children's outcomes are strongly predicted by quantile-specific predictions rather than mean predictions in Online Appendix C and Online Appendix Figure V

[42] The rate of convergence need not be identical across all quantiles of the income distribution because the prediction for permanent residents at each quantile π^{90}_{pcs} could reflect a different combination of causal effects and sorting. The key test is whether the prediction for the relevant quantile has more predictive power than predictions at the mean or other quantiles.

outcomes:

$$Cov(\theta_i, m\pi^q_{pcs}|X^q) > 0 \Rightarrow Cov(\theta_i, m\Delta_{odps}|X^q, m\pi^q_{pcs}) > 0. \quad (9)$$

Under this assumption, the findings in Table III imply that our estimates of γ reflect causal place effects (which are quantile-specific) rather than omitted variables, which are not quantile-specific under (9).

Gender. Finally, we conduct an analogous set of placebo tests exploiting heterogeneity in place effects by child gender. We begin by constructing gender-specific predictions of the mean outcomes of children of permanent residents by estimating (1) separately for male and female children, which we denote by $\bar{y}^m_{pcs}$ and $\bar{y}^f_{pcs}$. Places that are better for boys and generally better for girls as well: the (population-weighted) correlation of $\bar{y}^m_{pcs}$ and $\bar{y}^f_{pcs}$ across CZs is 0.9 at the median ($p = 50$).[43] We exploit the residual variation across genders to conduct placebo tests analogous to those above, based on the premise that unobservable shocks are unlikely to have gender-specific effects.

In Table IV, we estimate exposure effect models analogous to (7) with separate predictions by gender. Column 1 replicates (7) using the gender-specific prediction Δ^g_{odps} instead of the prediction that pools both genders. We continue to obtain an exposure effect estimate of $\gamma = 0.038$ per year in this specification. In Column 2, we use the prediction for the other gender Δ^{-g}_{odps} instead. Here, we obtain an estimate of 0.034, as expected given the high degree of correlation in place effects across genders. In Column 3, we include predictions for both genders, identifying the coefficients purely from differential variation across genders within CZs. In this specification, the coefficient on the own gender prediction is substantially larger than the other-gender prediction, which is close to zero (see Online Appendix Figure VII for non-parametric binned scatter plots corresponding to this regression).[44]

One may be concerned that families sort to different areas based on their child's gender, which – unlike the quantile and cohort-specific variation used above – is known at the time of the move. To address this concern, Columns 4-6 of Table V replicate Columns 1-3 including family fixed effects. The own-gender prediction remains a much stronger predictor of children's outcomes than the other-gender prediction even when we compare siblings' outcomes within families. Column 7

[43] Online Appendix Figure V presents a heat map of $\bar{y}^m_{pcs} - \bar{y}^f_{pcs}$. Some areas, such as Syracuse and Albany, NY are relatively better for males than females, while others, such as Milwaukee, WI are relatively better for females than males. In general, outcomes for boys are relatively worse than those for girls in areas with higher crime rates, a larger fraction of single parents, and greater inequality (Chetty et al. (2016)).

[44] It is not surprising that the other gender prediction remains positive, as the prediction for the other gender may be informative about a place's effect for children of a given gender due to measurement error. In general, finding a 0 effect on the "placebo" prediction is sufficient but not necessary to conclude that there is no sorting under an assumption analogous to (8).

shows that this remains the case when we restrict the sample to families that have at least one boy and one girl, for whom differential sorting by gender is infeasible.

The gender-specific convergence documented in Table IV supports the causal exposure effects model under an assumption analogous to (8), namely that the unobservable θ_i does not vary differentially across children of different genders within a family. This assumption requires that families who move to areas that are particularly good for boys do not invest systematically more in their sons relative to their daughters, a restriction that would hold if, for instance, families do not have systematically different preferences over their sons' and daughters' outcomes. Under this assumption, the gender-specific convergence in proportion to exposure time must reflect causal place effects.

V.E Summary

The results in this section show that various refinements of our baseline design – such as including family fixed effects or exploiting cohort- or gender-specific variation – all yield annual exposure effect estimates of $\gamma \simeq 0.04$. These findings have two important implications for identification of exposure effects.

First, our baseline design – which simply compares families who move with children of different ages in observational data – does not appear to be confounded by selection and omitted variable biases. We believe that such biases are small in our application for two reasons. First, the degree of age-dependent sorting across large geographies such as CZs and counties is limited, as families seeking better schools or environments for their children at certain ages presumably make more local moves (e.g., to a neighborhood school district). Second, children's outcomes conditional on parent income are not significantly correlated with mean parent incomes in an area (Chetty et al. 2014). As a result, moving to a better area for children is not systematically associated with parents finding better jobs, mitigating what might be the most important confounding factor.

Second, the findings above imply that any omitted variable θ_i that generates bias in our exposure effect estimates must: (1) operate within the family in proportion to exposure time (family fixed effects); (2) be orthogonal to changes in parental income and marital status (controls for observables); (3) persist in the presence of moves induced by displacement shocks (displacement shock analysis); and (4) precisely replicate permanent residents' outcomes by birth cohort, quantile, and gender in proportion to exposure time (outcome-based placebo tests). We believe that plausible omitted variables are unlikely to have all of these properties and therefore conclude that places have causal effects on children in proportion to the amount of time they spend growing up in the area.

VI Other Outcomes

In this section, we estimate neighborhood effects for other outcomes, including college attendance, marriage, teenage employment, and teenage birth. The results provide further evidence on the types of outcomes that are shaped by neighborhood environment and illustrate how neighborhoods affect behavior before children enter the labor market.

Figure VIII replicates Figure IVb for rates of college attendance and marriage. In Panel A, we replicate the baseline specification in equation 6 replacing Δ_{odps} with $\Delta^c_{odps} = C_{pds} - C_{pos}$, where C_{pcs} is the fraction of children who attend college at any point between ages 18 and 23 (among children of permanent residents in CZ c in birth cohort s with parental income rank p). In Panel B, we replace Δ_{odps} with $\Delta^m_{odps} = m_{pds} - m_{pos}$, where m_{pcs} is the fraction of children who are married at age 26.

We find approximately linear childhood exposure effects until age 23 for both of these outcomes. Moving to an area with higher college attendance rates at a younger age increases a child's probability of attending college. Likewise, moving to an area where permanent residents are more likely to be married at age 26 at a younger age increases a child's probability of being married. The estimated annual exposure effect is$\gamma = 0.037$ for college attendance and $\gamma = 0.025$ for marriage.

In Figure IX, we analyze outcomes measured while children are teenagers. Panel D considers teen birth, defined as being listed as a parent on a birth certificate prior to age 20. We construct gender-specific predictions of teenage birth rates and plot the baseline specification in Equation (6), replacing Δ_{odps} with $\Delta^z_{odpsg} = z_{pdsg} - z_{posg}$, where z_{pcsg} is the fraction of children of permanent residents in CZ c with parental income p in cohort s and gender g who have a teenage birth. For both boys and girls, there are clear childhood exposure effects: moving at an earlier age to an area with a higher teen birth rate increases a child's probability of having a teenage birth. The gradient is especially steep between ages 13 and 18, suggesting that a child's neighborhood environment during adolescence may play a particularly important role in determining teen birth outcomes.

In Panels A-C of Figure IX, we analyze neighborhood effects on teenage employment rates. In these figures, the outcomes is an indicator for employment (based on having a W-2 form filed on one's behalf) at ages 16, 17, or 18. The key independent variable in each of these figures is constructed based on the employment rate of children of permanent residents at the corresponding age. For teen employment, we find discontinuous effects of moving just before employment is measured rather than continuous exposure effects. Children who move at age 15 to a CZ where more 16 year olds work are much more likely to work at age 16 than children who make the same move at age 17. Making the same move at earlier ages (before age 16) further increases the probability of working at age 16, but the exposure effect is small relative to the discrete jump at age 16 itself. Analogous discrete jumps are observed at ages 17 and 18 when one measures employment outcomes at ages 17 and 18 (Panels B and C).[45] These discrete jumps suggest that part of the effects of the neighborhoods may come from discrete experiences during childhood, such as summer jobs that are available in a given area at certain ages. These experiences may aggregate to produce the linear childhood exposure effects that shape outcomes in adulthood.

Although the mean earnings of individuals in an area are correlated with other outcomes such as college attendance and teenage birth rates, there is substantial residual variation in permanent resident's outcomes on each of these dimensions. For example, permanent residents' mean earnings rank at age 30 has a (population-weighted) correlation of 0.461 with college attendance rates, implying that more than 75% of the variance in college attendance rates is orthogonal to mean earnings ranks (Online Appendix Table VI). Hence, the finding that movers outcomes converge to those of permanent residents constitutes further evidence that neighborhoods have causal effects, as it would be unlikely that unobserved shocks would generate convergence on a spectrum of different outcomes.[46] Moreover, the fact that neighborhoods have causal effects on a wide variety of outcomes beyond earnings further suggests that the mechanism through which neighborhoods shape children's outcomes is not driven by labor market conditions but rather a set of environmental factors that shape behaviors throughout childhood.

VII Conclusion

This paper has shown that children's opportunities for economic mobility are shaped by the neighborhoods in which they grow up. Neighborhoods affect children's long-term outcomes through childhood exposure effects: every extra year a child spends growing up in an area where permanent residents' outcomes are higher increases his or her earnings. Movers' outcomes converge to those of permanent residents in the destination to which they move at a rate of approximately 4% per year of

[45] The magnitude of the $\{b_m\}$ coefficients in Panels B-D is approximately 0.8 at young ages and 0 after the age at which employment is measured. Under our identifying assumption of constant selection effects by age, this implies that movers pick up 80% of the differences in teenage employment rates across CZs observed for permanent residents.

[46] This logic is analogous to the tests for distributional convergence in Section V.D; here, we are exploiting other dimensions of the joint distribution of children's outcomes to test for distributional convergence.

childhood exposure. Extrapolating this annual exposure effect over 20 years of childhood, children who move to a new area at birth will pick up roughly 80% of the difference in permanent residents' outcomes between their origin and destination. Much of the variation in intergenerational mobility observed across areas thus appears to be driven by causal effects of place rather than heterogeneity in the types of people living in those places.

These results motivate place-based approaches to improving economic mobility, such as making investments to improve opportunity in areas that currently have low levels of mobility or helping families move to higher opportunity areas using targeted housing vouchers. Identifying specific policy solutions – i.e., the investments needed to improve mobility and the areas to which families should be encouraged to move – requires identifying the causal effect of each neighborhood and understanding what makes some areas produce better outcomes than others. The analysis in the present paper shows that differences in permanent residents' outcomes are predictive of neighborhoods' causal effects *on average*. However, it does not provide an unbiased estimate of the causal effect of each area on a given child's outcomes, as the outcomes of permanent residents in any given area could reflect an arbitrary mix of selection and causal effects. We construct estimates of the causal effect of growing up in each CZ and county in the U.S. and characterize the properties of areas that produce good outcomes in the second paper in this series.

References

Aaronson, D. (1998). Using sibling data to estimate the impact of neighborhoods on childrens educational outcomes. Journal of Human Resources 33(4), 915–46.

Andersson, F., J. Haltiwanger, M. Kutzbach, G. Palloni, H. Pollakowski, and D. H. Weinberg (2013). Childhood housing and adult earnings: A between-siblings analysis of housing vouchers and public housin. US Census Bureau Center for Economic Studies Paper No. CES-WP-13-48.

Basu, S. (2010). Age of entry effects on the education of immigrant children: A sibling study. Available at SSRN 1720573.

Census, U. (1999). 1999 Census Zip Code Crosswalk, web.archive.org/web/20120914150518/http:/www.census.gov/geo/www/tiger/zip1999.html Accessed July 2014.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014). Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. American Economic Review 104(9), 2633–79.

Chetty, R. and N. Hendren (2015). The impacts of neighborhoods on intergenerational mobility: Childhood exposure effects and county-level estimates. Working Paper.

Chetty, R. and N. Hendren (2016). The impacts of neighborhoods on intergenerational mobility ii: County-level estimates. Working Paper.

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. Quarterly Journal of Economics 129(4), 1553–1623.

Chetty, R., N. Hendren, F. Lin, J. Majerovitz, and B. Scuderi (2016). Gender gaps in childhood: Skills, behavior, and labor market preparedness childhood environment and gender gaps in adulthood. The American Economic Review 106(5), 282–288.

Cilke, J. (1998). A profile of non-filers. U.S. Department of the Treasury, Office of Tax Analysis Working Paper No. 78.

Clampet-Lundquist, S. and D. S. Massey (1993). Neighborhood effects on economic self sufficiency: A reconsideration of the moving to opportunity experiment. American Journal of Sociology 114.1 (2008): 107-143.

Crowder, K. and S. J. South (2011). Spatial and temporal dimensions of neighborhood effects on high school graduation. Social science research 40(1), 87–106.

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

HUD, U. (2011). USPS Zip Code Crosswalk Files, Accessed July 2014.

Jencks, C. and S. E. Mayer (1990). The social consequences of growing up in a poor neighborhood. Inner-city poverty in the United States 111, 186.

Katz, L. F., J. R. Kling, and J. B. Liebman (2001). Moving to opportunity in boston: Early results of a randomized mobility experiment. The Quarterly Journal of Economics 116(2), 607–654.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental analysis of neighborhood effects. Econometrica 75 (1): 83-119 75(1), 83–119.

Ludwig, J., G. J. Duncan, L. A. Gennetian, L. F. Katz, R. C. Kessler, J. R. Kling, and L. Sanbonmatsu (2013). Long-term neighborhood effects on low-income families: Evidence from moving to opportunity. American Economic Review Papers and Proceedings 103(3): 226-31.

Ludwig, J., J. B. Liebman, J. R. Kling, G. J. Duncan, L. F. Katz, R. C. Kessler, and L. Sanbonmatsu (2008). What can we learn about neighborhood effects from the moving to opportunity experiment? American Journal of Sociology 114, 144–188.

Massey, D. S. (1993). American apartheid: Segregation and the making of the underclass. Harvard University Press.

Oreopoulos, P. (2003). The long-run consequences of living in a poor neighborhood. Quarterly Journal of Economics 118(4), 1533–1175.

Plotnick, R. and S. Hoffman (1996). The effect of neighborhood characteristics on young adult outcomes: The effect of neighborhood characteristics on young adult outcomes: Alternative estimates. Institute for Research on Poverty Discussion Paper no. 1106-96.

Sampson, R. J. (2012). Great American City: Chicago and the Enduring Neighborhood Effect. Chicago: University of Chicago Press.

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.

Sharkey, P. and J. W. Faber (2014). Where, when, why, and for whom do residential contexts matter? moving away from the dichotomous understanding of neighborhood effects. Annual Review of Sociology 40(July): 559-79.

Solon, G. (1992). Intergenerational income mobility in the united states. American Economic Review 82(3), 393–408.

Tolbert, C. M. and M. Sizer (1996). U.S. commuting zones and labor market areas: A 1990 update. Economic Research Service Staff Paper 9614.

van den Berg, G. J., P. Lundborg, P. Nystedt, and D.-O. Rooth (2014). Crtical periods during childhood and adolescence. Journal of the European Economic Association 12(6), 1521–1557.

Wilson, W. J. (1987). The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy. Chicago: University of Chicago Press.

Wodtke, G. T. (2013). Duration and timing of exposure to neighborhood poverty and the risk of adolescent parenthood. Demography.

Wodtke, G. T., F. Elwert, and D. Harding (2012). Poor families, poor neighborhoods: How family poverty intensifies the impact of poor families, poor neighborhoods: How family poverty intensifies the impact of concentrated disadvantage on high school graduation. Population Studies Center Research Report 12-776.

Wodtke, G. T., D. J. Harding, and F. Elwert (2011). Neighborhood effects in temporal perspective the impact of long-term exposure to concentrated disadvantage on high school graduation. American Sociological Review 76(5), 713–736.

Online Appendix A. County-Level Estimates

Online Appendix Table I replicates our baseline specifications using counties as the geographic unit, as opposed to CZs. Broadly, the patterns at the county level reflect those at the CZ level, but with slightly attenuated values for γ. This attenuation is consistent with the presence of greater residential sorting at the county level. Online Appendix Figure VIII presents estimates of the permanent resident outcomes across counties for $p = 25$ and $p = 75$.

To construct these estimates, we measure $\bar{y}_{pcs}$ using county-level permanent residents and we consider two samples of 1-time county movers. Online Appendix Table IV presents the summary statistics of each of these samples. First, we consider a sample of 1-time movers who move at least 100 miles between counties with populations above 250,000, analogous to the same sample restrictions we impose on the 1-time CZ movers. Column (1) shows we obtain a baseline slope of 0.035, slightly lower than our baseline slope of 0.040 at the CZ level. The smaller slope is consistent with a slightly larger degree of residential sorting at the county, as opposed to the CZ level. Column (2) adds family fixed effects to the baseline specification in Column (1) and obtains an exposure slope of 0.033 (0.011), not significantly different from the baseline slope of 0.035. This suggests the quasi-experimental design is not confounded by dynamic sorting patterns operating at the county level within the CZ.

While our baseline analysis for CZ moves and for the county moves in Columns (1)-(2) focused on moves above 100 miles, Columns (3)-(7) in Table VI explore moves across counties within CZs including moves less than 100 miles. Column (3) replicates the baseline specification using moves across counties with populations at least 250,000, measuring outcomes of the children at age 24. Here, we obtain a slope of 0.022 (s.e. 0.003), significantly lower than the estimate of 0.035 we obtain for longer distance moves. This drop is consistent with the the child spending some time in the old location or the impact of border-effects from the coarse manner in which we measure neighborhood quality at the county level.

Column (4) measures the child's outcome (and predicted outcomes of permanent residents) at age 26 instead of age 24. Here, we obtain a similar but slightly higher slope of 0.032. Column (5) stacks the data across outcomes at age 24-32. Here, we obtain a more precisely estimated coefficient of 0.027 (s.e. 0.002). Column (6) adds family fixed effects to the specification in Column (5) and obtains a similar slope of 0.029 (s.e. 0.025). While our estimate remains stable, it is considerably more imprecise with the addition of family fixed effects across counties within CZs. Finally, Column (7) considers within-CZ moves across counties with populations of at least 10,000 as opposed to 250,000. Here, we obtain a similar but slightly attenuated coefficient of 0.024 relative to the 0.027 in column (5). This is consistent consistent with attenuation bias from using relatively imprecise estimates of the permanent resident outcomes in smaller places.

Online Appendix B. Additional Robustness Specifications

This Appendix presents additional robustness specifications, focusing on the role of the population and distance restrictions and the impact of more imprecision in the permanent resident outcomes on the family fixed effects estimates.

Population and Distance Restrictions. Our baseline analysis restricts to families who move between CZs with at least 250K people in both the origin and destination location and moves beyond 100 miles. Online Appendix Table I illustrates how our baseline results for the specification in equation (7) change when we vary the distance and population size restrictions.

Column (1) repeats the baseline specification in column (1) of Table II. Columns (2)-(4) include moves of all distances. Column (2) restricts to moves between origins and destinations with at least 50K people; column (3) restricts to moves to and from places with at least 250K people (our baseline population restriction); column (4) restricts to move to and from CZs with at least 500K people. Columns (5)-(7) repeat columns (2)-(4) imposing our baseline restriction that moves are farther than 100 miles. Hence, column (6) is identical to column (1) as it imposes our baseline distance restriction of 100 miles and a 250K population restriction. Columns (8)-(10) require that moves be farther than 200 miles.

Broadly, we find that using a more restrictive sample does not meaningfully affect our results; but including moves of smaller distances and between places with smaller populations tends to lead to slightly attenuated coefficients, γ. For example, relaxing the population restriction from 250K to 50K or removing our distance restriction drops our estimate of γ from 0.04 to 0.036 (s.e. 0.001), as shown in columns (3) and (5).

These patterns are to be expected because our permanent resident outcomes are estimated with sampling error. Because of spatial auto-correlation in permanent resident outcomes, one expects that more of the variation in Δ_{odps} for shorter distances reflects sampling variation as opposed to true differences in permanent resident outcomes. Similarly, including CZs with smaller populations naturally leads to less precise estimates of Δ_{odps} and hence an attenuated coefficient for γ. But, restricting to moves further than 100 miles or more restrictive population restrictions (e.g. 500K) does not meaningfully affect our results. This suggests sampling error in the permanent resident outcomes is not significantly affecting our baseline results.

Family Fixed Effects Specification. Sampling variation in permanent resident outcomes can more seriously affect the family fixed effect estimates. Online Appendix Table VII illustrates this by showing the impact of incorporating moves between smaller CZs. Columns (1) and (2) report the baseline specification and baseline family fixed effects specification that include a cohort-varying intercept and restrict to populations in the origin and destination location above 250,000 based on the 2000 Census. Columns (3)-(4) repeat the baseline specification dropping the cohort-varying intercept. As noted in the main text and in columns (5) and (9) of Table II, not allowing the intercept to vary by cohort attenuates the baseline coefficient to 0.036 and drops the coefficient in the family fixed effects specification to 0.031. The greater attenuation in the family fixed effects specification reflects the fact that γ is identified solely from comparisons across birth cohorts when using variation coming from siblings.

Columns (5)-(6) repeat the specification in Column (3)-(4) now extending the sample to include moves to places with populations above 100,000 in both the origin and destination. This attenuates the coefficient in Column (3) from 0.036 to 0.032 in Column (5). Adding family fixed effects now drops this coefficient to 0.024. As noted in the next row of Online Appendix Table VII, family fixed effects yields an estimate of γ that is 0.76 the size of the specification without the family fixed effects. Columns (7)-(8) incorporate moves to places with populations above 50,000, and the specifications in columns (9)-(10) remove the population restriction entirely. Fully removing the population restriction yields a coefficient for γ in equation (6) of 0.031, which is then attenuated to 0.020 using family fixed effects.

As the table illustrates, incorporating less precise estimates for the permanent resident outcomes induces attenuation not only in the level of our baseline specification given by Equation (6), but also an even more severe attenuation in the family fixed effects specification. The family fixed effects specification yields a coefficient that is statistically indistinguishable from our baseline coefficient when including moves between large CZs and incorporating a cohort-varying intercept. Dropping the cohort-varying intercept continues to yield a coefficient in the family fixed effect specification that is 86% of the size of the baseline coefficient. But, incorporating moves from to places with more than 100K people drops this to 76%; and 71% if moves are included to places with more than 50K people, and 61% if the population restriction is removed. This reflects the fact that the variation in permanent resident outcomes across siblings reflects a greater sensitivity to measurement error from variation in permanent resident outcomes across cohorts. This motivates our primary focus on moves between large CZs so that we obtain sufficiently precise estimates of permanent resident outcomes.

Online Appendix C. Multiple Movers

This section presents the specification that incorporates families that move across CZs more than once. Let $d(j)$ denote the jth destination location and let m_i^j denote the age of the child when moving to destination. We consider families that move across CZs up to three times, which excludes the 3% of families who move across CZs more than 3 times. Let E_i denote the number of times the family moved, $E_i \in \{1, 2, 3\}$. To parameterize exposure to places when children are above versus below age 23, we define e_{ij} to be the number of years child i is exposed to the jth place below age 23. And, we let $e_{ij}^{>23}$ denote the number of years above age 23 that the individual is exposed to place j.

For each place j the child lives, we construct $\Delta_{od(j)ps}^j = \Delta_{od(j)ps} = \bar{y}_{pd(j)s} - \bar{y}_{pos}$ as the difference in the child's predicted outcome based on permanent residents in destination j and the mover's first observed (origin) CZ, o. We estimate a specification analogous to our baseline linear specification in equation (7) given by:

$$\begin{aligned} y_i \quad = \quad & \left[\sum_{j=1}^{3} \gamma_j e_{ij} \Delta_{od(j)ps}^j\right] + \sum_{s=1980}^{1988} I(s_i = s)(\alpha_s^1 + \alpha_s^2 \bar{y}_{pos} + \sum_{j=1}^{3} \left(\alpha_s^{3,j} \bar{y}_{pd(j)s} + \kappa_s^j \Delta_{od(j)ps}^j\right)) + \sum_{j=1}^{3} \delta^j e_{ij}^{>23} \Delta_{od(j)ps} + \sum_{j=1}^{3} \delta_0^j 1 \left\{e_{ij}^{>23} > 0\right\} \Delta_{od(j)ps} \\ & + \phi^1 \bar{y}_{pos} + \sum_{j=1}^{3} (\phi^{2,j} \bar{y}_{pd(j)s} + \zeta^{1,j} e_{ij}) + \sum_{j=1}^{3} \sum_{j'=1}^{3} I\left(e_{ij} = j'\right) \zeta^{5,j,j'} e_{ij} p_i + \sum_{j=1}^{3} I\left(E_i = j\right) \left(\zeta^{2,j} + \zeta^{3,j} e_{ij} + \zeta^{4,j} p_i + \sum_{s=1980}^{1988} \zeta^{6,s,j} \bar{y}_{pos}\right) + \epsilon_{3,i} \end{aligned} \quad (10)$$

Our primary coefficients of interest, γ_j, are the coefficients on the interaction of e_{ij} and $\Delta_{od(j)ps}^j$, which comprise $\sum_{j=1}^{3} \gamma_j e_{ij} \Delta_{od(j)ps}^j$. Analogous to our baseline specification, we also parameterize a level impact of $\Delta_{od(j)ps}^j$ that may vary by cohort, $\kappa_s^j \Delta_{odps}$, and we control for the outcomes of permanent residents in the origin, interacted with cohort, $\alpha_s \bar{y}_{pos}$. Moreover, we parameterize the impact of exposure above age 23 by including interactions of $e_{ij}^{>23}$ with $\Delta_{od(j)ps}$ and a level effect, $1\left\{e_{ij}^{>23} > 0\right\} \Delta_{od(j)ps}$. These two terms represent an extension of the terms $I(m_i > 23)(\delta + (23 - m_i)\delta')\Delta_{odps}$ in equation 7 to the case of multiple moves.

The second line includes a set of additional control variables analogous to those in equation 7, but differ in two respects. First, we simplify the controls by including linear terms in exposure, e_{ij}, as opposed to separate indicators for the age of the child at the time of the move (e.g. $I(m_i = m)$ in equation 7). Second, we generalize the controls to capture potential disruption effects of multiple moves. We include indicators for the number of total moves, along with its interaction with exposure time to each place, $\sum_{j=1}^{3} I(E_i = j)\zeta^{3,j} e_{ij}$, its interaction with parental income, $\sum_{j=1}^{3} I(E_i = j)\zeta^{4,j} p_i$, and its cohort-specific interaction with permanent resident outcomes, $\sum_{j=1}^{3} I(E_i = j) \sum_{s=1980}^{1988} \zeta^{6,s,j} \bar{y}_{pos}$.

Online Appendix Table III presents the results. We estimate a coefficient of 0.040 (s.e. 0.001) for the first destination (γ), 0.037 (s.e. 0.004) for the second destination (γ^2), and 0.031 (s.e. 0.006) for the third destination (γ^3), as shown in column 1. Column 2 presents results for γ^j that constrains the estimates to be constant for all j, yielding an estimate of 0.039 (s.e. 0.001), statistically indistinguishable from our baseline estimates using our sample of 1-time movers.

As noted in the text, these results support our interpretation of the results as a model of exposure, as opposed to a critical age model. To see how the multiple movers analysis implies this, it is helpful to consider an example of a critical age model. Suppose that moving to a better neighborhood improves a child's network of friends with a probability p that is declining with the age at move and that once one makes a new set of contacts, they last forever. In this model, neighborhood effects would decline with the age at move (as in Figure IV), but the duration of exposure to a better area would not matter for long-term outcomes. Conceptually, one cannot distinguish a critical age model from a duration of exposure model in a sample of one-time movers because the age at move is collinear with the duration of exposure.

However, the multiple movers specification breaks this collinearity between age at move and duration of exposure. The fact that outcomes remain proportional to the duration of exposure at the same rate even when duration is not collinear with the age at which the child moves shows that duration of exposure is the key determinant of children's earnings in adulthood.

Online Appendix D. Specification Details for Outcome-Based Placebo Tests

This section presents specification details and further results from the outcome-based placebo tests described in Section V.D.

Cohort Variation

We begin by describing the specification for the cohort variation test. We begin with the baseline linear specification in equation (7). For each cohort $t = s + a$ for $a \in \{-4, -3, -2, -1, 1, 2, 3, 4\}$, we add six variables to this regression. The first five variables replicate the linear parameterization of the exposure effect for cohort t by including: $1\{m_i \leq 23\} * \Delta_{odpt}$, $1\{m_i > 23\} * \Delta_{odpt}$, $1\{m_i \leq 23\} * m_i * \Delta_{odpt}$, $1\{m_i > 23\} * m_i * \Delta_{odpt}$, and $\bar{y}_{pot}$. These variables are populated for any child, i, in cohort $s(i)$ for which we observe permanent resident predictions, $\bar{y}_{pct}$. In the case we are unable to observe Δ_{odpt} and $\bar{y}_{pot}$ (e.g. if $t = 1979$ or $t = 1987$), we set each of these first five variables to zero and we include an indicator, $I_a = 1\{\text{cohort } s(i) + a \text{ is non-missing}\}$ for each $a \in \{-4, -3, -2, -1, 1, 2, 3, 4\}$. For example, since our data covers the 1980-86 cohorts, a child in the 1983 cohort would have non-zero entries for the permanent resident predictions for the three years surrounding 1983, but would have zero entries for the permanent resident predictions along with indicators $I_a = 1$ for $a = -4$ and $a = 4$. The resulting specification is

$$\begin{aligned}
y_i \quad = \quad & \sum_{s=1980}^{1988} I(s_i = s)(\alpha_s^1 + \alpha_s^2 \bar{y}_{pos}) + \sum_{m=9}^{30} I(m_i = m)(\zeta_m^1 + \zeta_m^2 p_i) + \sum_{s=1980}^{1987} \kappa_s^d I(s_i = s)\Delta_{odps} \\
& + I(m_i \leq 23)\left(b_0 + (23 - m_i)\gamma\right)\Delta_{odps} + I(m_i > 23)(\delta + (23 - m_i)\delta')\Delta_{odps} \\
& + \sum_{a \in \{-4,-3,-2,-1,1,2,3,4\}} I\{m_i \leq 23\}\left(b_0^a + (23 - m_i)\gamma_a\right)\Delta_{odp,s+a} + I(m_i > 23)(\delta_a + (23 - m_i)\delta_a')\Delta_{odp,s+a} \\
& + \sum_{a \in \{-4,-3,-2,-1,1,2,3,4\}} \alpha_a \bar{y}_{po,s+a} + \omega_a I_a + \varepsilon_{3i}
\end{aligned}$$

where the last line includes these additional controls. The estimates of γ_a from this specification are presented in Figure VII.

Quantiles: Distributional Convergence

For the distributional convergence analysis, we again begin with the baseline linear specification in equation (7). However, we now consider an outcome, y_i, to be an indicator for having income above the 90th percentile (or an indicator for being unemployed). Letting π^{90}_{pcs} denote the predicted probability of having income above the 90th percentile in CZ c for parental income level p and cohort s, we define $\Delta^{90}_{odps} = \pi^{90}_{pds} - \pi^{90}_{pos}$ to be the difference in permanent resident outcomes in the origin versus destination.

Column (1) of Table III reports the coefficient γ from the regression analogous to the baseline linear specification in equation (7) that everywhere replaces mean income outcomes with indicators for income above the 90th percentile:

$$\begin{aligned} y_i \quad = \quad & \sum_{s=1980}^{1988} I(s_i = s)(\alpha^1_s + \alpha^2_s \pi^{90}_{pos}) + \sum_{m=9}^{30} I(m_i = m)(\zeta^1_m + \zeta^2_m p_i) + \sum_{s=1980}^{1987} \kappa^d_s I(s_i = s)\Delta^{90}_{odps} \\ & + I(m_i \leq 23)\,(b_0 + (23 - m_i)\gamma)\,\Delta^{90}_{odps} + I(m_i > 23)(\delta + (23 - m_i)\delta')\Delta^{90}_{odps} + \varepsilon_{3i}, \end{aligned}$$

In column (3), we include five additional controls for the mean incomes of permanent residents: $1\{m_i \leq 23\} * \Delta_{odps}$, $1\{m_i > 23\} * \Delta_{odps}$, $1\{m_i \leq 23\} * m_i * \Delta_{odps}$, $1\{m_i > 23\} * m_i * \Delta_{odps}$, and $\bar{y}_{pos}$, yielding:

$$\begin{aligned} y_i \quad = \quad & \sum_{s=1980}^{1988} I(s_i = s)(\alpha^1_s + \alpha^2_s \pi^{90}_{pos}) + \sum_{m=9}^{30} I(m_i = m)(\zeta^1_m + \zeta^2_m p_i) + \sum_{s=1980}^{1987} \kappa^d_s I(s_i = s)\Delta^{90}_{odps} \\ & + I(m_i \leq 23)\,(b_0 + (23 - m_i)\gamma)\,\Delta^{90}_{odps} + I(m_i > 23)(\delta + (23 - m_i)\delta')\Delta^{90}_{odps} \\ & + I(m_i \leq 23)\,(b_{mean,0} + (23 - m_i)\gamma_{mean})\,\Delta_{odps} + I(m_i > 23)(\delta_{mean} + (23 - m_i)\delta'_{mean})\Delta_{odps} + \alpha^3 \bar{y}_{pos} + \varepsilon_{3i}, \end{aligned}$$

Column (2) present results from a specification that drops the terms in the second row of the specification for column (3). Here, we obtain a significant, positive estimate of 0.022, as expected given the high degree of correlation in place effects across quantiles: places that push children into the top 10% also tend to improve mean outcomes. Column (3) presents the main results from the specification above. We estimate $\gamma = 0.040$ (s.e. 0.003) and $\gamma_{mean} = 0.004$ (s.e. 0.003).

Online Appendix Figure V (Panel A) presents a visual illustration of these coefficients. We present a binned scatter plot of the probability a child is in the top 10%, y^{90}_i vs. the destination prediction π^{90}_{pds} and the mean rank prediction $\bar{y}_{pds}$ in the sample of children who move at or before age 13. The series in circles shows the non-parametric analog of a partial regression of a child's outcome on π^{90}_{pds}, controlling for the $\bar{y}_{pds}$ and the analogous predicted outcomes based on prior residents in the origin, π^{90}_{pos} and $\bar{y}_{pos}$. To construct this series, we regress both y^{90}_i and π^{90}_{pds} on the mean predicted income rank, $\bar{y}_{pds}$, and the analogous origin controls, π^{90}_{pos} and $\bar{y}_{pos}$, bin the π^{90}_{pcs} residuals into 20 equal-sized bins, and plot the mean residuals of y^{90}_i vs. the mean residuals of π^{90}_{pcs} within each bin. The series in triangles is constructed analogously, except that we plot residuals of y^{90}_i vs. residuals of $\bar{y}_{pcs}$, the predicted mean rank.

Online Appendix Figure V (Panel A) shows that children who move before age 13 to areas where children are more likely to be in the top 10% are much more likely to reach the upper tail themselves: a 1 percentile increase in π^{90}_{pcs} is associated with an 0.651 percentile increase in the movers' probability of reaching the top 10%, controlling for the mean rank outcomes of permanent residents in the origin and destination CZ along with the top 10% prediction in the origin CZ. In contrast, conditional on the probability of reaching the top 10%, variation in the mean predicted outcome has no impact at all on a child's probability of reaching the top 10% (slope of 0.030).

Columns (4)-(6) repeat the specifications in columns (1)-(3) but use an indicator for being employed, as opposed to having income above the 90th percentile. These results are also illustrated in Online Appendix Figure V (Panel B), which replicates Panel A using non-employment (roughly the bottom 10%) as the outcome instead of reaching the top 10%. Once again, we find that children's probabilities of reaching the lower tail are strictly related to the predicted probability of reaching the lower tail based on permanent residents' outcomes rather than the predicted mean outcome. The fact that mean predicted outcomes of permanent residents $\bar{y}_{pcs}$ have no predictive power implies that other omitted factors, which are not quantile-specific under (9), do not drive our findings.

Gender Variation

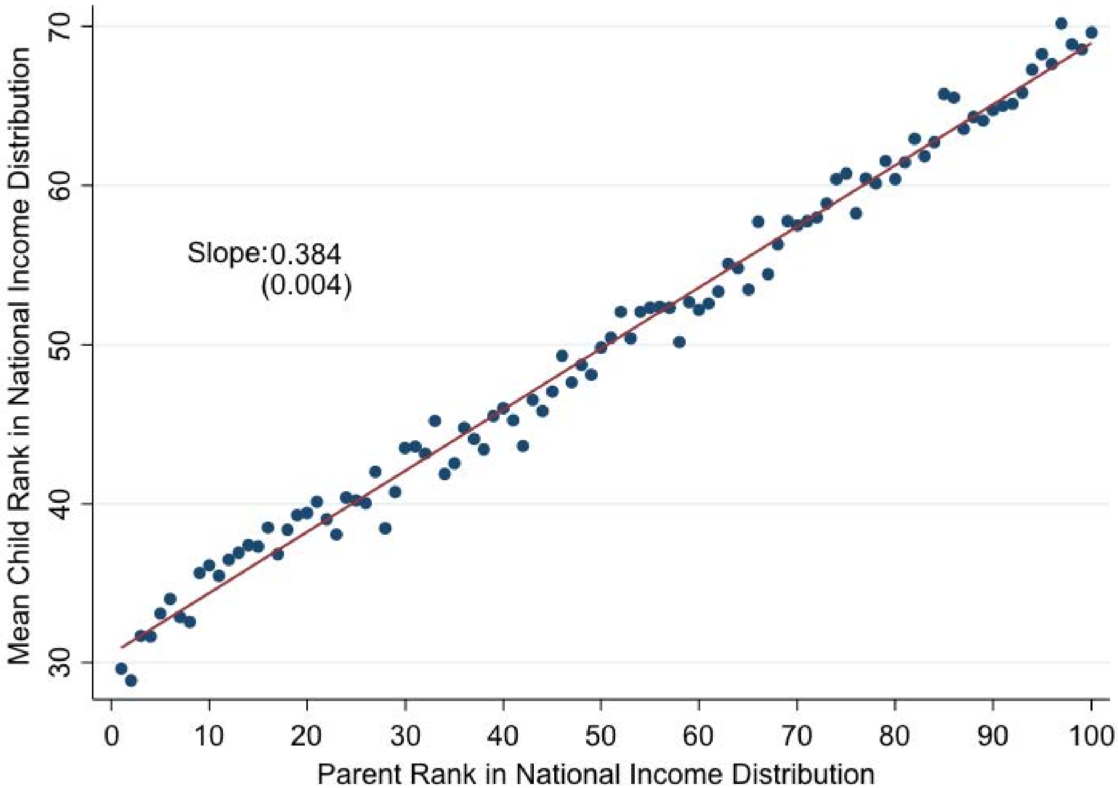
We begin by constructing gender-specific predictions of the mean outcomes of children of permanent residents by estimating (1) separately for male and female children, which we denote by $\bar{y}^m_{pcs}$ and $\bar{y}^f_{pcs}$. For a child of gender $g \in \{m, f\}$, we define $\Delta^g_{odps} = \bar{y}^g_{pds} - \bar{y}^g_{pos}$. The primary specification in column (3) of Table IV adds five variables to the baseline specification in equation (7): $1\{m_i \leq 23\} * \Delta_{odps}$, $1\{m_i > 23\} * \Delta_{odps}$, $1\{m_i \leq 23\} * m_i * \Delta_{odps}$, $1\{m_i > 23\} * m_i * \Delta_{odps}$, and $\bar{y}_{pos}$. This yields the specification:

$$\begin{aligned} y_i \;=\; & \sum_{s=1980}^{1988} I(s_i = s)(\alpha^1_s + \alpha^2_s \bar{y}^g_{pos}) + \sum_{m=9}^{30} I(m_i = m)(\zeta^1_m + \zeta^2_m p_i) + \sum_{s=1980}^{1987} \kappa^d_s I(s_i = s)\Delta^g_{odps} \\ & + I(m_i \leq 23)\left(b_0 + (23 - m_i)\gamma\right)\Delta^g_{odps} + I(m_i > 23)(\delta + (23 - m_i)\delta')\Delta^g_{odps} \\ & + I(m_i \leq 23)\left(b_{other,0} + (23 - m_i)\gamma_{other}\right)\Delta^{-g}_{odps} + I(m_i > 23)(\delta_{other} + (23 - m_i)\delta'_{other})\Delta^{-g}_{odps} + \alpha^3 \bar{y}^{-g}_{pos} + \varepsilon_{3i}, \end{aligned}$$

where g is the child's gender and $-g$ is the opposing gender. We obtain a coefficient of $\gamma = 0.031$ (s.e. 0.003) and $\gamma_{other} = 0.009$ (s.e. 0.003).

To further illustrate these differential patterns, Online Appendix Figure VII presents a binned scatter plot of children's ranks vs. the difference in the destination and origin prediction, Δ^g_{odps}, for their own gender (circles) and the prediction Δ^{-g}_{odps} for the other gender (triangles) in the sample of children who move at or before age 13. Each series shows the non-parametric analog of a partial regression of a child's outcome on the prediction for a given gender, controlling for the other-gender prediction. To construct the series in circles, we regress both y_i and Δ^g_{odps} on Δ^{-g}_{odps} and origin by parent income decile by cohort by gender fixed effects. We then bin the Δ^g_{odps} residuals into 20 equal-sized bins, and plot the mean residuals of y_i vs. the mean residuals of Δ^g_{odps} within each bin. The series in triangles is constructed analogously, except that we plot residuals of y_i vs. residuals of Δ^{-g}_{odps}, the prediction for the *other* gender. The figure shows that children who move before age 13 to areas where children of their own gender have better outcomes do much better themselves: a 1 percentile increase in the mean rank $\bar{y}^g_{pds}$ for $g = g(i)$ is associated with a 0.523 percentile increase in the movers' mean rank. In contrast, conditional on the own-gender prediction, variation in the prediction for the other gender is associated with only a 0.144 percentile increase in the movers' mean rank.

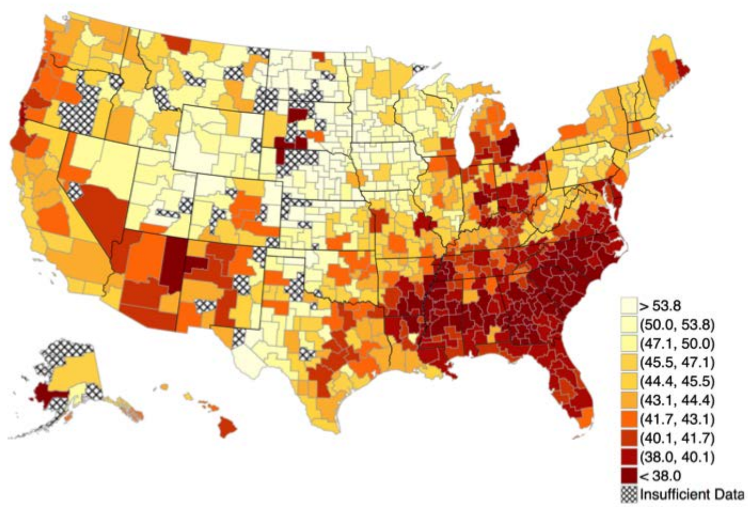
FIGURE I: Mean Child Income Rank at Age 30 Vs. Parent Income Rank for Children Raised in Chicago



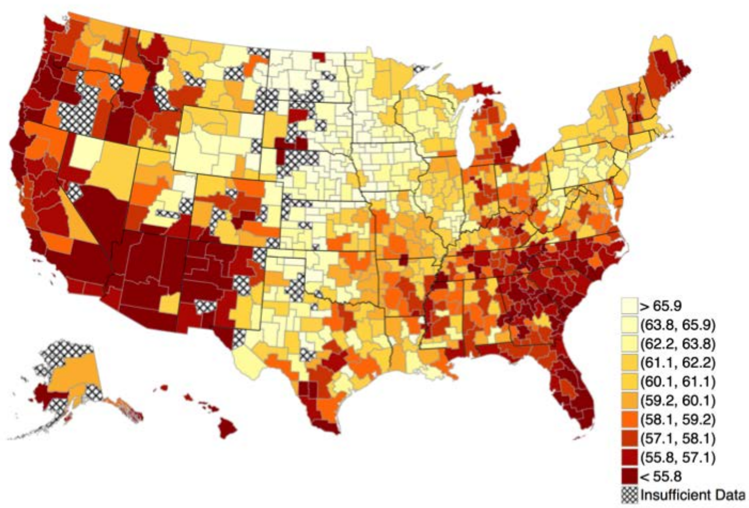
Notes: This figure presents a non-parametric binned scatter plot of the relationship between mean child income ranks and parent income ranks for all children raised in Chicago. The figure measures income of the children at age 30 using the 1982 cohort. Child income is family income at age 30, and parent income is mean family income from 1996-2000. We define a child's rank as her family income percentile rank relative to other children in her birth cohort and his parents' rank as their family income percentile rank relative to other parents of children in the core sample. The ranks are constructed for the full geographic sample, but the graph illustrates the relationship for the sub-sample of families who report living in Chicago for all years of our sample, 1996-2012. The figure then plots the mean child percentile rank at age 30 within each parental percentile rank bin. The slope and best-fit line are estimated using an OLS regression on the micro data. Standard errors are reported in parentheses.

FIGURE II: Predicted Income Rank at Age 30 - Permanent Residents

A. For Children with Parent at the 25^{th} Percentile

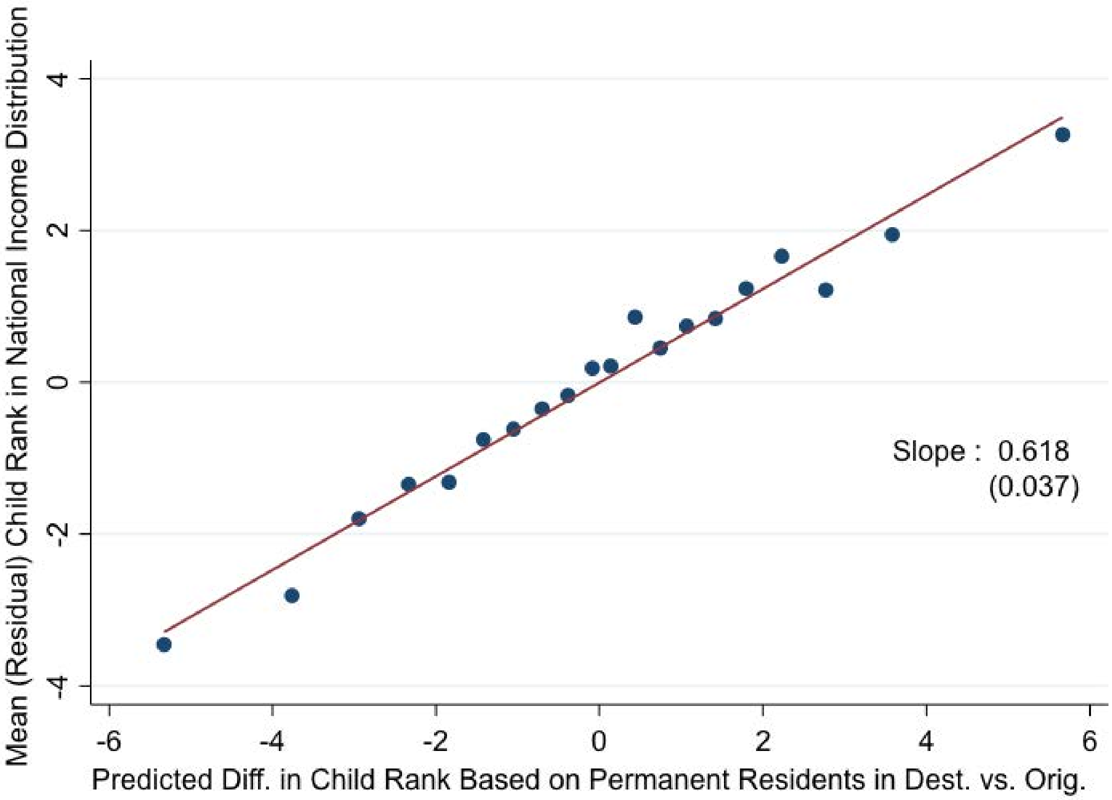


B. For Children with Parent at the 75^{th} Percentile



Notes: These figures illustrate the geographic variation in child income rank outcomes at age 30 amongst our sample of permanent residents born in the 1980-82 cohorts, across commuting zones (CZs) in the U.S. Panel A reports the expected rank for children whose parental income is at the 25th percentile of the income distribution of parents, and Panel B reports the expected rank for children whose parental income is at the 75th percentile. Both figures use the baseline family income definitions for parents and children. The figure restricts to the subset of parents who stay in the commuting zone throughout our sample period (1996-2012) (but does not restrict based on the geographic location of the child at age 30). To construct this figure, within each cohort we regress child income rank on a constant and parent income rank in each CZ, exploiting the linearity property shown in Figure I. Panel A then reports the predicted child rank outcome for parents at the 25th percentile of the family income distribution (~$30K per year), pooling across cohorts 1980-82. Similarly, Panel B reports the predicted child rank outcome for parents at the 75th percentile of the family income distribution (~ $97K per year).

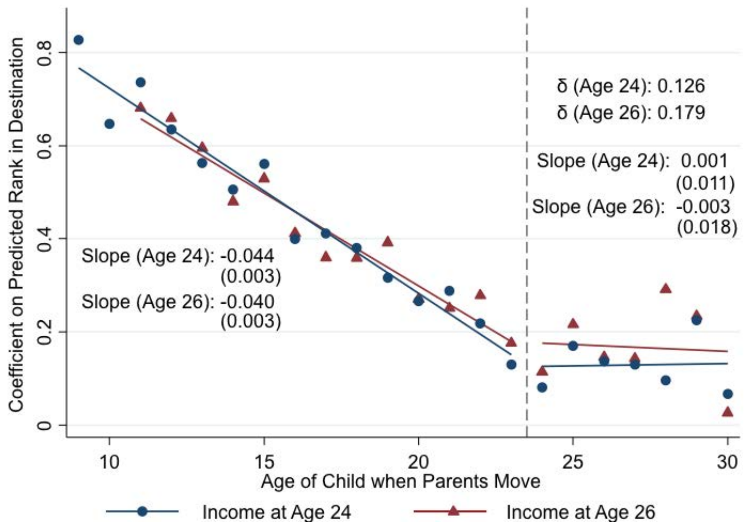
FIGURE III: Movers' Outcomes at Age 24 vs. Predicted Outcomes Based on Residents in Destination Moves at Age 13



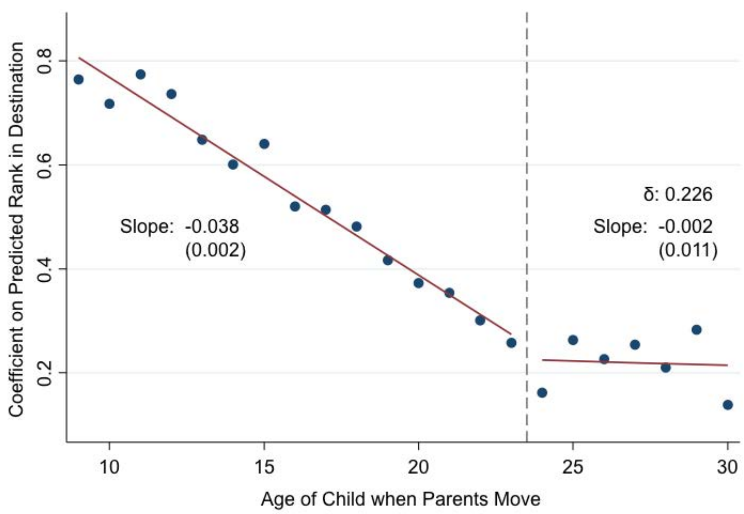
Notes: This figure presents a non-parametric illustration of the b_{13} coefficient in equation (4). The sample includes all children in 1-time moving households whose parents moved when the child was 13 years old. Child income is measured when the child is age 24. The figure is constructed by first partialing out the fixed effects (the interaction of (a) origin CZ, (b) the child's age at the parental move, (c) cohort, and (d) parental income deciles): we regress the difference in the destination versus origin prediction, Δ_{odps}, on the fixed effects and the child rank outcome on the fixed effects. The figure then plots the relationship between these residuals from each of these regression. We construct 20 equal sized bins of the residuals from the destination regression and, in each bin, plot the mean of the residuals from the child rank regression.

FIGURE IV: Exposure Effect Estimates for Children's Income Rank in Adulthood

A. Income at Age 24 and 26



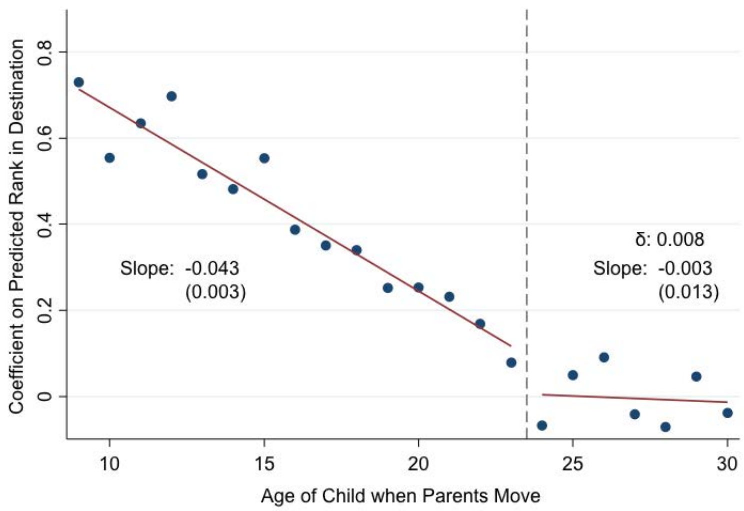
B. Income at Age 24



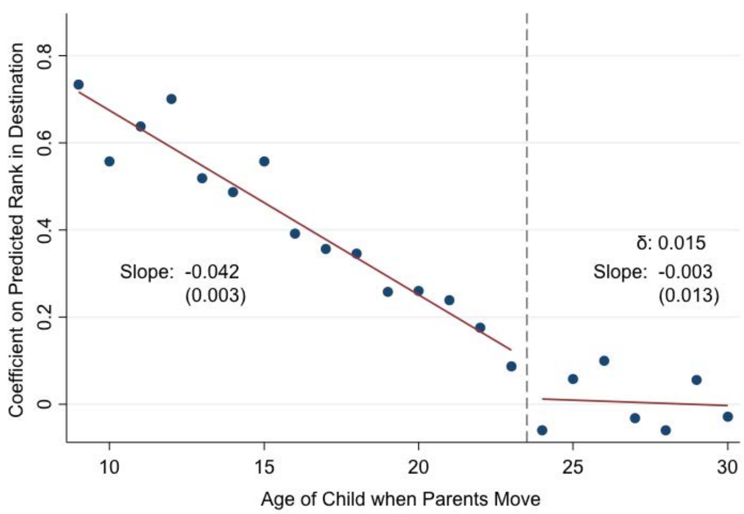
Notes: Panel A presents estimates of the coefficients, $\{b_m\}$, from the semi-parametric specification in equation (5) for various child income measurements at different ages. The sample includes all children in 1-time moving households. Child income is measured when the child is age 24, and 26. We estimate these coefficients by regressing the child's family income rank on the difference in the predicted family income rank based on prior residents in the destination location relative to the origin location (computed using the linear regression illustrated in Figure I) interacted with each age of the child at the time of the move. We include the set of fixed effects for origin by parent income decile by cohort by the child's age at the time of the move (as in Figure III). Panel B presents estimates from the specification in equation (6). This specification drops the large set of fixed effects and instead includes (a) dummies for the child's age at the time of the move, (b) parental rank (within the child's cohort) interacted with child age dummies, and (c) cohort dummies and predicted outcomes in the destination and origin interacted with cohort dummies. Panels A and B report slopes and intercepts from a regression of the b_m coefficients on m separately for $m \leq 23$ and $m > 23$. We compute δ as the predicted value of the line at age 23 using the b_m estimates for $m > 23$.

FIGURE V: Exposure Effect Estimates for Children's Income Rank in Adulthood with Controls for Observables

A. Family Fixed Effects

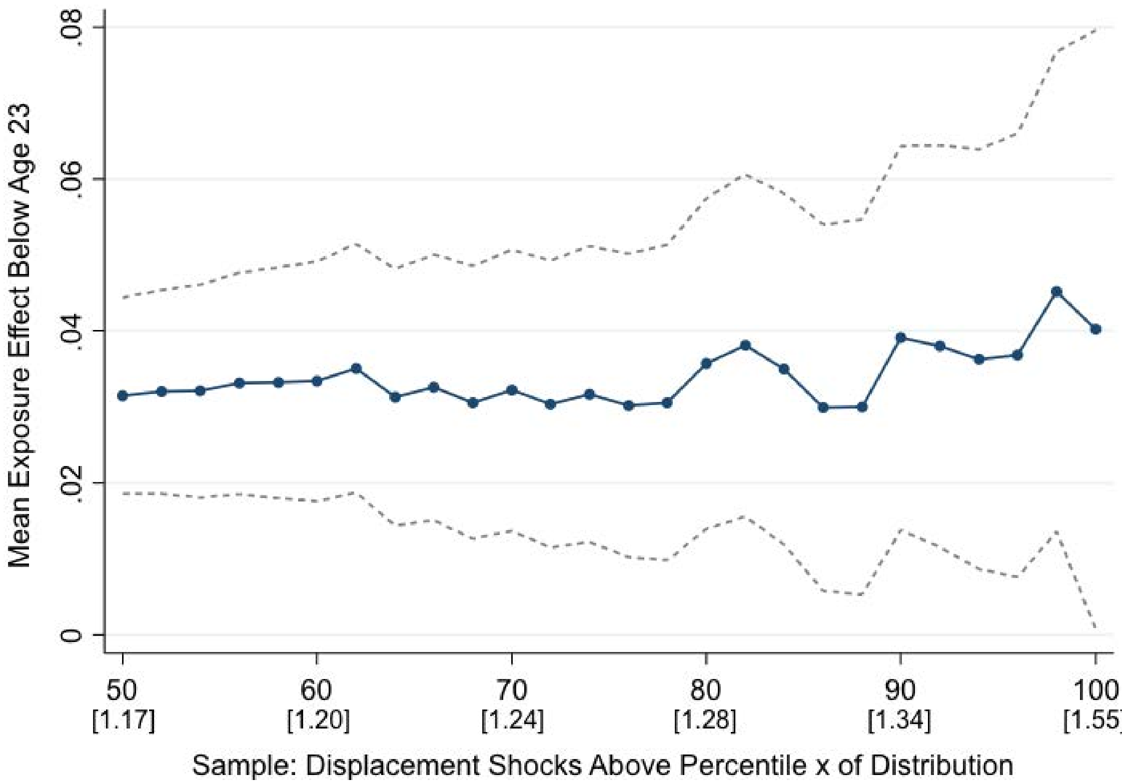


B. Family Fixed Effects and Time Varying Controls



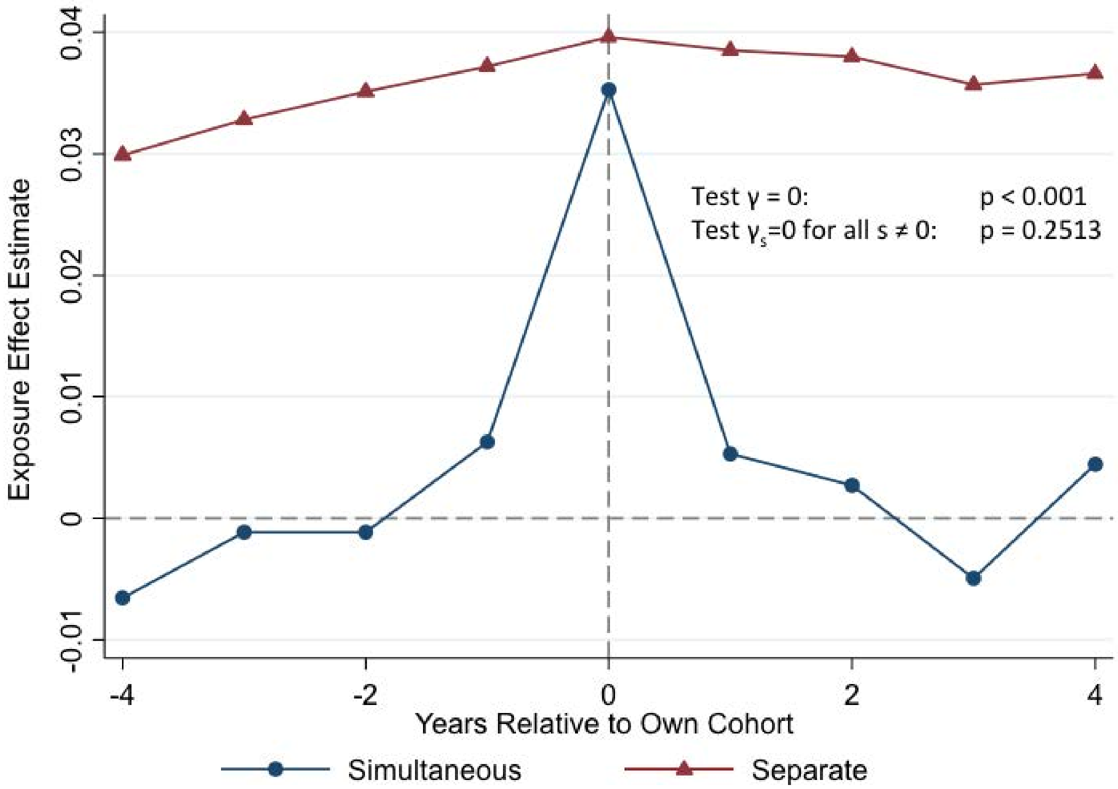
Notes: This figure presents estimates of the coefficients, $\{b_m\}$, in specifications that add family fixed effects (Panel A) and both family fixed effects and controls for changes in marital status and parental income (Panel B). Panel A presents estimates of b_m from the baseline specification in equation (6) with the addition of family fixed effects. Panel B adds family fixed effects along with a set of controls for income rank changes and marital status changes around the time of the move. To do so, we construct parental income ranks by cohort by year of outcome measurement. We interact the differences in parental ranks in the year before versus after the move with an indicator for the child age at the time of the parental move (for ages below 24) and an interaction with an indicator for child age greater than 23 at the time of the parental move. We also construct a set of indicators for marital status changes. We define marital status indicators for the year before the move and the year after the move and construct indicators for being always married, getting divorced, or being never married (getting married is the omitted category). We include these variables and their linear interactions with the child age at the time of the parental move (for ages below 24) and an interaction with an indicator for child age greater than 23 at the time of the parental move. As in Figure IV, we report slopes and intercepts from a regression of the b_m coefficients on m separately for $m \leq 23$ and $m > 23$. We compute δ as the predicted value of the line at age 23 using the b_m estimates for $m > 23$.

FIGURE VI: Displacement Shocks IV Exposure Effects Estimates



Notes: This figure presents estimates of the exposure time slope for a subsample of moves restricted to zipcode-by-year observations with large outflows, instrumenting for the change in predicted outcomes based on prior residents, Δ_{odps}, with the average change in predicted outcomes for the given origin. More specifically, for each zipcode in our sample of children in the 1980-1993 cohorts, we calculate the number whose parents leave the (5-digit) zipcode in each zipcode, z, in year t, m_{zt}. Then, we compute the average number of people who leave in a given year across our 1997-2012 sample window, $\bar{m}_z$. We then divide the outflow in a zipcode-year observation, m_{zt}, by the mean outflow for the zipcode to construct our measure of the displacement shock, $d = \frac{m_{zt}}{\bar{m}_{zt}}$. The horizontal axis presents the results for varying quantile thresholds of d ranging from the median to the 95th percentile. The corresponding mean value of d for the sample is presented in brackets. For each zipcode, we compute the mean value of Δ_{odps} for each parental income decile (pooling across all years and all movers in the zipcode). Throughout, we restrict to zipcode-years with at least 10 observations. Then, for each sample with values of d above the threshold, we estimate γ in equation (7). We plot the estimate of γ as a function of the threshold.

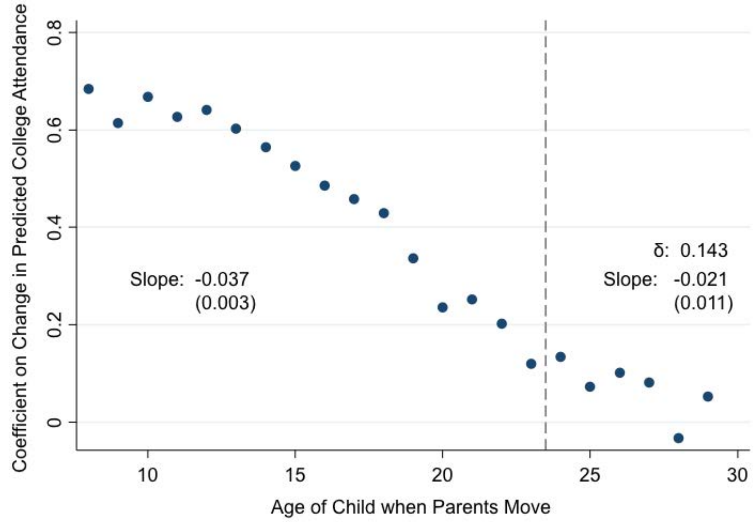
FIGURE VII: Exposure Effects Based on Cross-Cohort Variation, with Cohort-Varying Intercepts



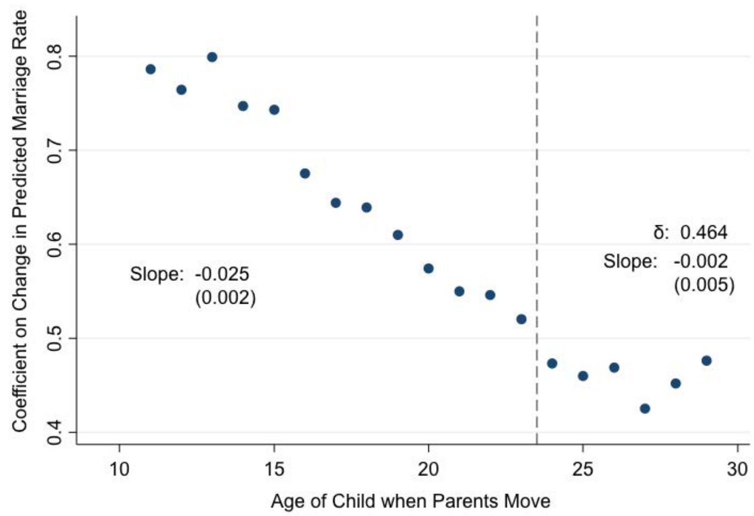
Notes: This figure presents estimates of the exposure time slope using own and placebo cohort place predictions. The sample includes all children in 1-time moving households whose parents moved when the child was less than or equal to 23 years old. The series in red triangles plots estimates of 9 separate regressions using place predictions for child in cohort c as if s/he were in cohort $c+k$, where k ranges between -4 and 4. By construction, the estimate for $k=0$ corresponds to the baseline slope of 0.040, illustrated in Figure IV (Panel B). Regressions include the predicted outcomes based on prior residents in the origin and destination (for cohort $c+k$), and the interactions of the child's age at the time of the move with the predicted outcomes in the origin and destination based on prior residents (for cohort $c+k$). To be consistent with the baseline specifications, regressions also include dummy indicators for true cohort and its interaction with the predicted outcomes in the origin location. The blue series reports coefficients from a single regression that includes all variables in each of the regressions for $k=-4,...,4$ and plots the coefficient on the interaction of the child's age at the time of the move with the predicted outcome based on prior residents in the destination location in cohort $c+k$. The figure presents the p-value for the test that $\gamma=0$ in the simultaneous specification ($p<0.001$), along with the p-value for the test that all other γ_a of than the main coefficient, γ, are equal to zero: $\gamma_a=0$ for $a\neq 0$. This generates a p-value of 0.2513, indicating we cannot reject the hypothesis that the other cohort predictions do not have any explanatory power for the child's outcome conditional on the child's own cohort predictions.

FIGURE VIII: Exposure Effect Estimates for College Attendance (18-23) and Marriage at Age 26

A. College Attendance (Age 18-23)

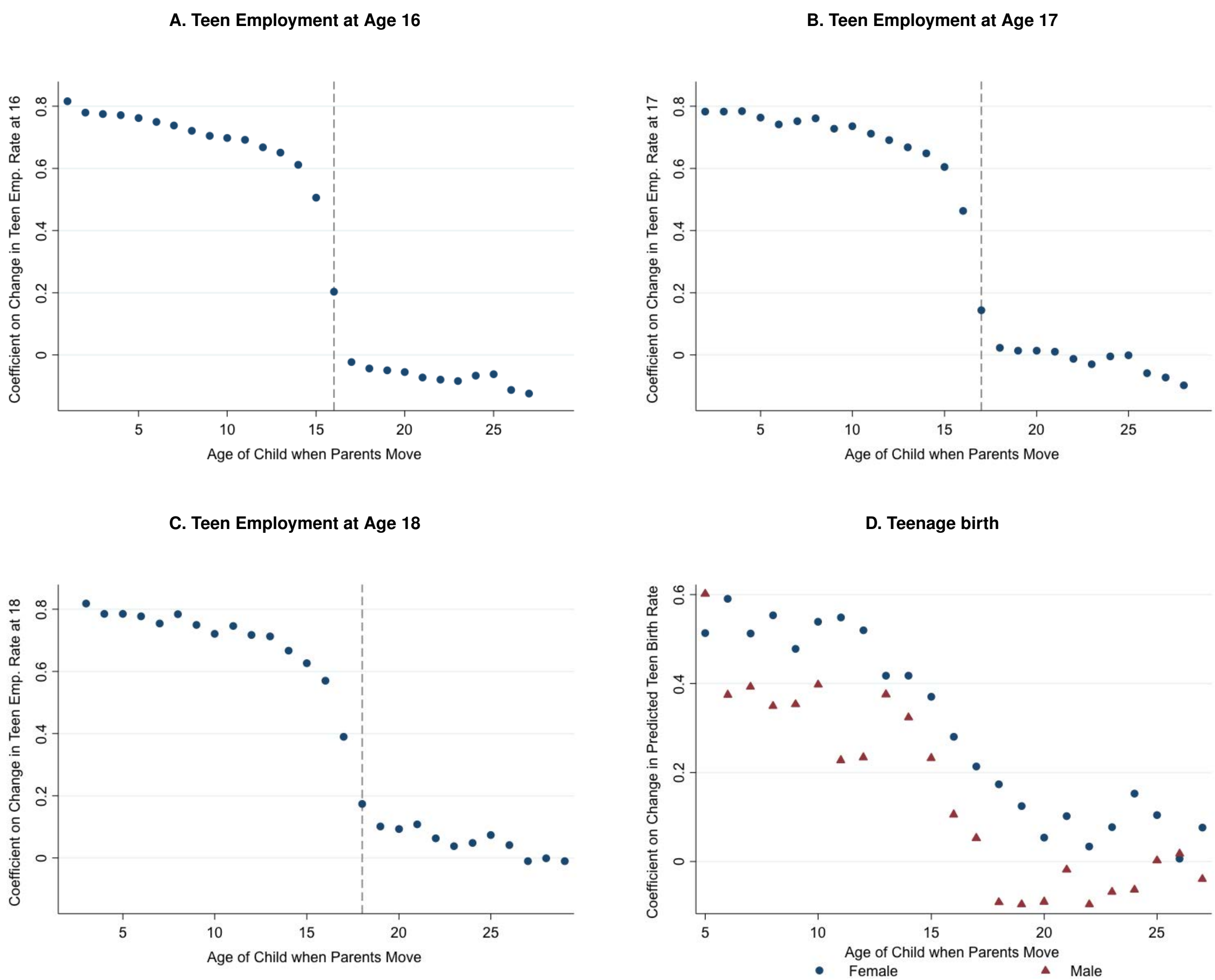


B. Marriage (Age 26)



Notes: This figure presents exposure effect estimates for college and marriage outcomes. In Panel A, we replicate the baseline specification in equation (6) replacing the child's outcomes with an indicator for college attendance at any age between 18-23. We construct separate analogous predicted outcomes based on the prior residents in each CZ for each outcome. We define college attendance as the existence of a 1098-T form (indicating college enrollment) when the child is 18-23 years old and restrict the sample to observations we observe for years 18-23. Because we observe college attendance in years 1999-2012, we obtain estimates for ages at move of 8-29. In Panel B, we replicate the baseline specification in equation (6) replacing the child's outcomes with an indicator for being married at age 26 using the child's filing status at age 26.

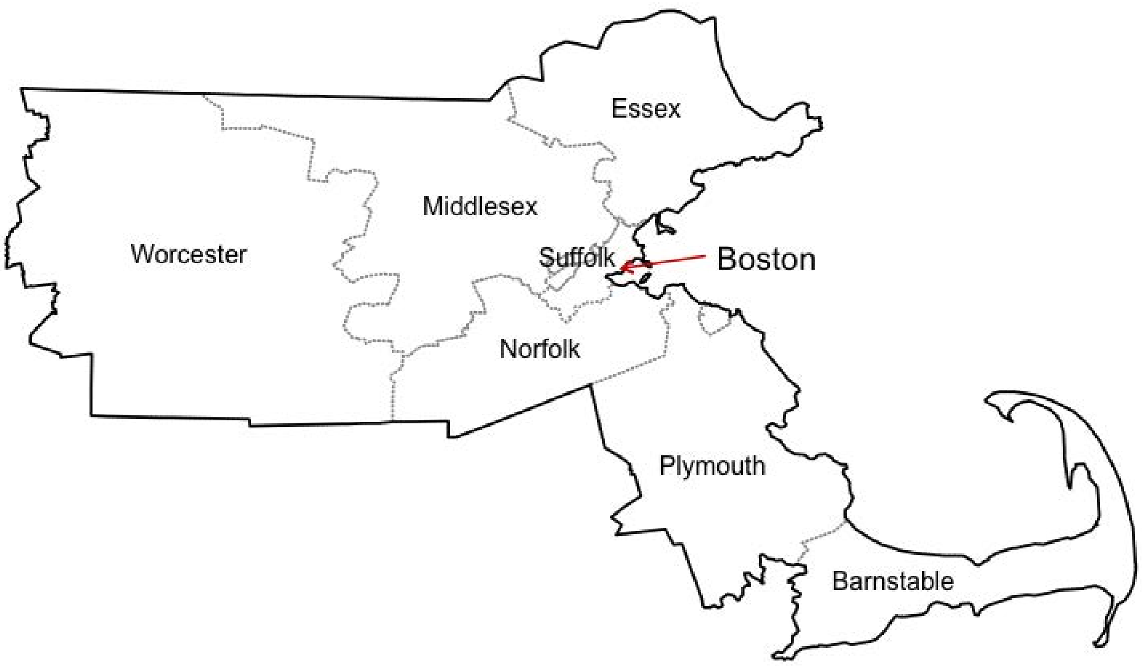
FIGURE IX: Exposure Effect Estimates for Teen Outcomes



Notes: This figure presents exposure effect estimates for teen outcomes. Panels A-C replicate the baseline specification in Figure IV (Panel B), but replaces the child's outcomes with an indicator for working at age 16-18 (defined as the existence of a W-2 during the year in which the child turned age a). Panel D presents estimates from the baseline specification using teen birth as the outcome. We define teenage birth as having a birth in the calendar year prior to turning age 20 using birth certificate records from social security administration records, and estimate the model separately for males and females.

ONLINE APPENDIX FIGURE I

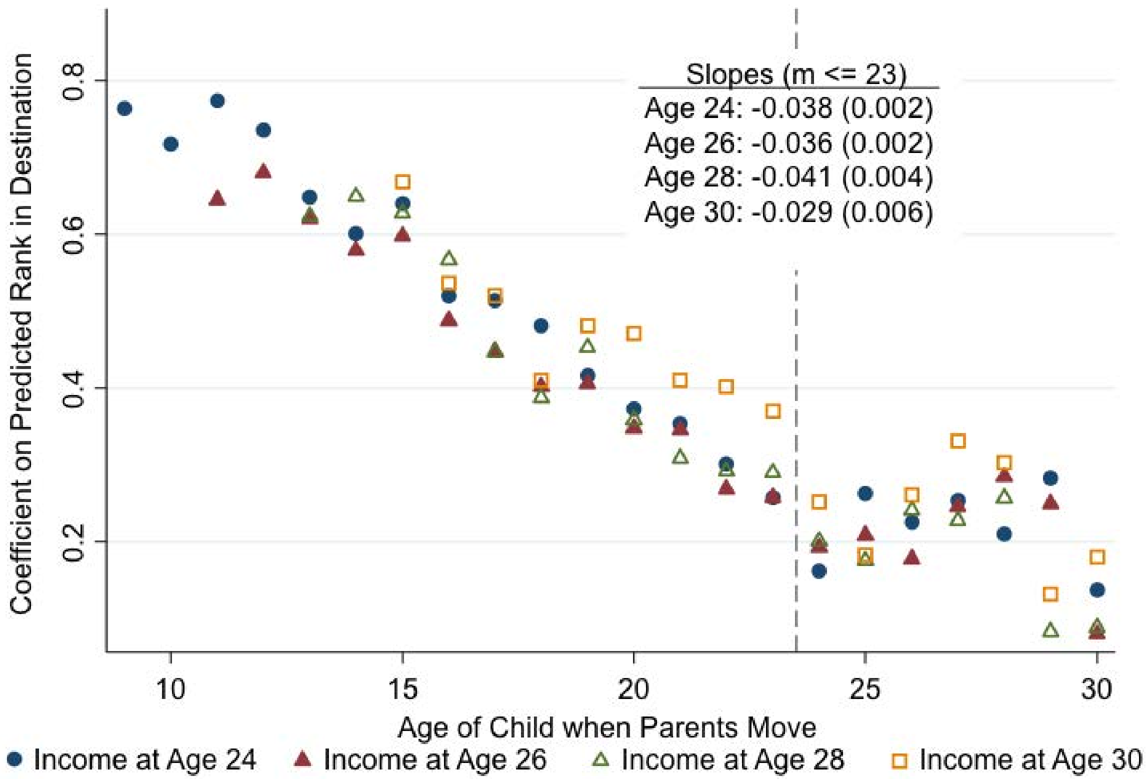
Map of Boston CZ



Notes: This figure presents a county map of the Boston commuting zone.

ONLINE APPENDIX FIGURE II

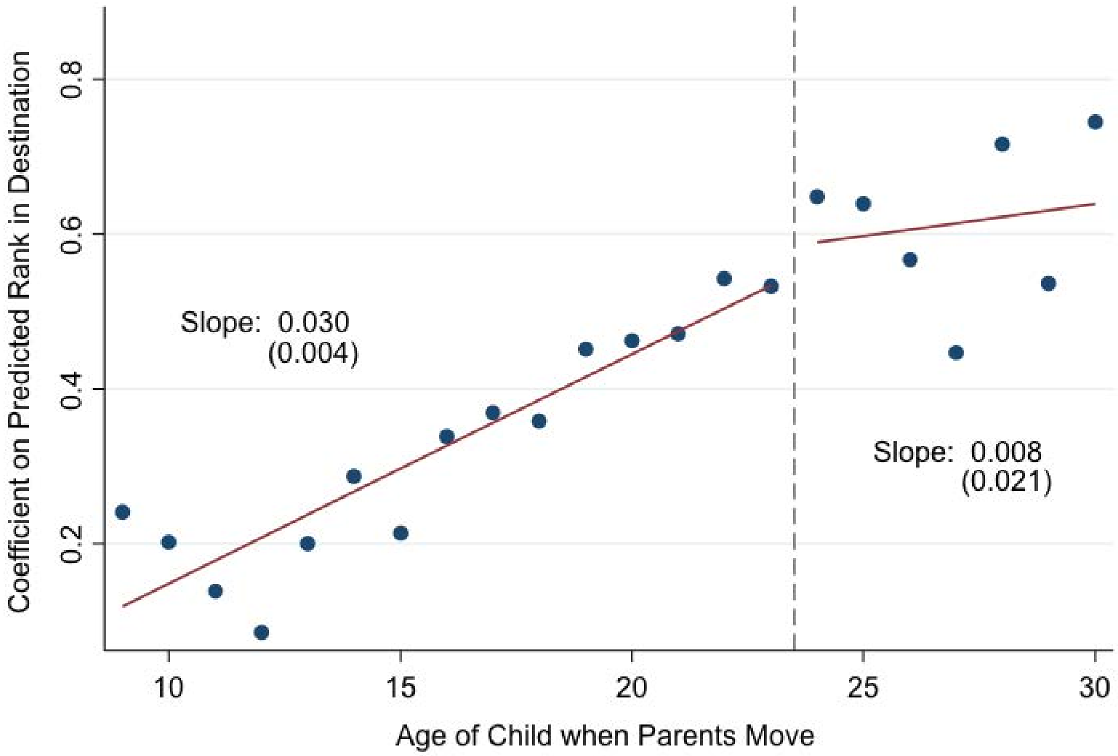
Exposure Effect Estimates at Age 24, 26, 28, and 30



Notes: This figure replicates our baseline specification in equation (6), shown in Figure IVb, using incomes measured at age 24, 26, 38, and 30. The figure reports the slopes from a regression of the b_m coefficients on m for $m \leq 23$, with standard errors in parentheses.

ONLINE APPENDIX FIGURE III

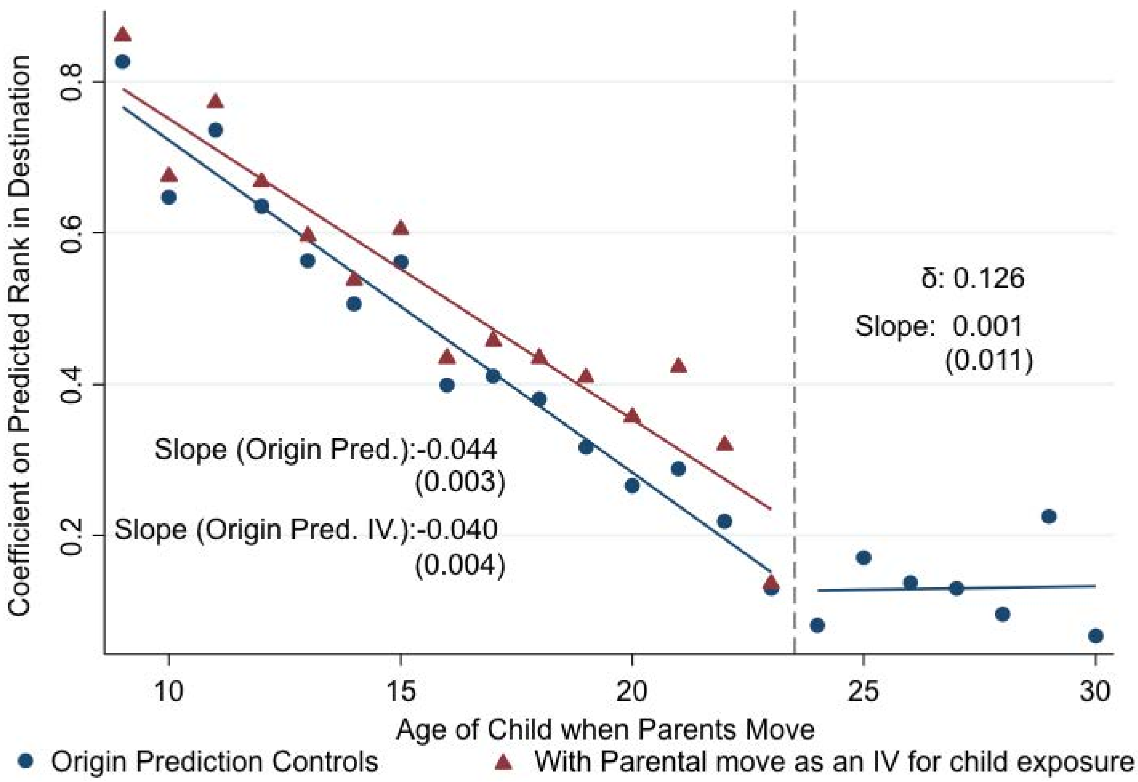
Exposure Effect Estimates using Origin Variation



Notes: This figure presents estimates of b_m in equation (5) that replaces the α_{qosm} fixed effects in equation (5) with α_{qdsm} fixed effects that control for the destination, d, instead of the origin, o, so that the slope is identified from variation in the origin exposure. As in Figure IVa, the figure reports the estimated slopes from a regression on the dots on the figure.

ONLINE APPENDIX FIGURE IV

Exposure Effect Estimates using Parental Move as an Instrument for Child Exposure

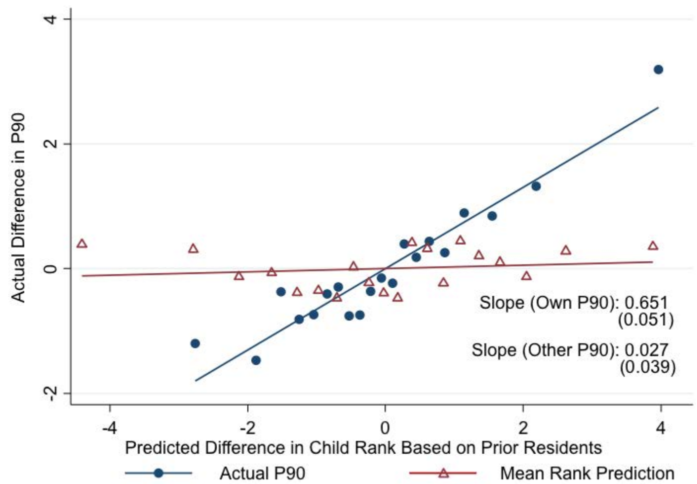


Notes: This figure presents estimates of the coefficients b_m adjusted for the probability that the child follows the parent to the destination. Formally, we construct the fraction of children who follow their parents when the parents move when the child is m years old, ϕ_m, as the fraction of children who either (a) file a tax return in the destination, (b) have a form W-2 mailing address in the destination location, or (c) attend a college (based on 1098-T filings by institutions) in the destination location. The figure plots the series of $b_m^{IV} = \frac{b_m - \delta}{\phi_m} + \delta$, where $\delta = 0.125$ is the estimated selection effect shown in Figure IVa.

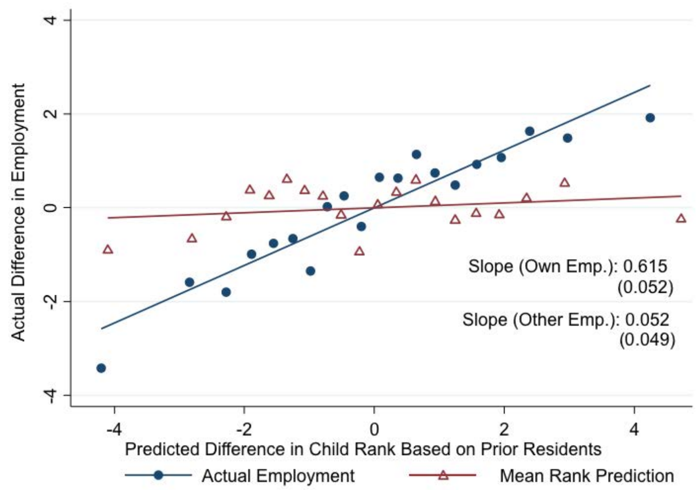
ONLINE APPENDIX FIGURE V

Movers' Outcomes vs. Predicted Employment and Probability of Reaching top 10% in Destination

A. Probability of Reaching Top 10%



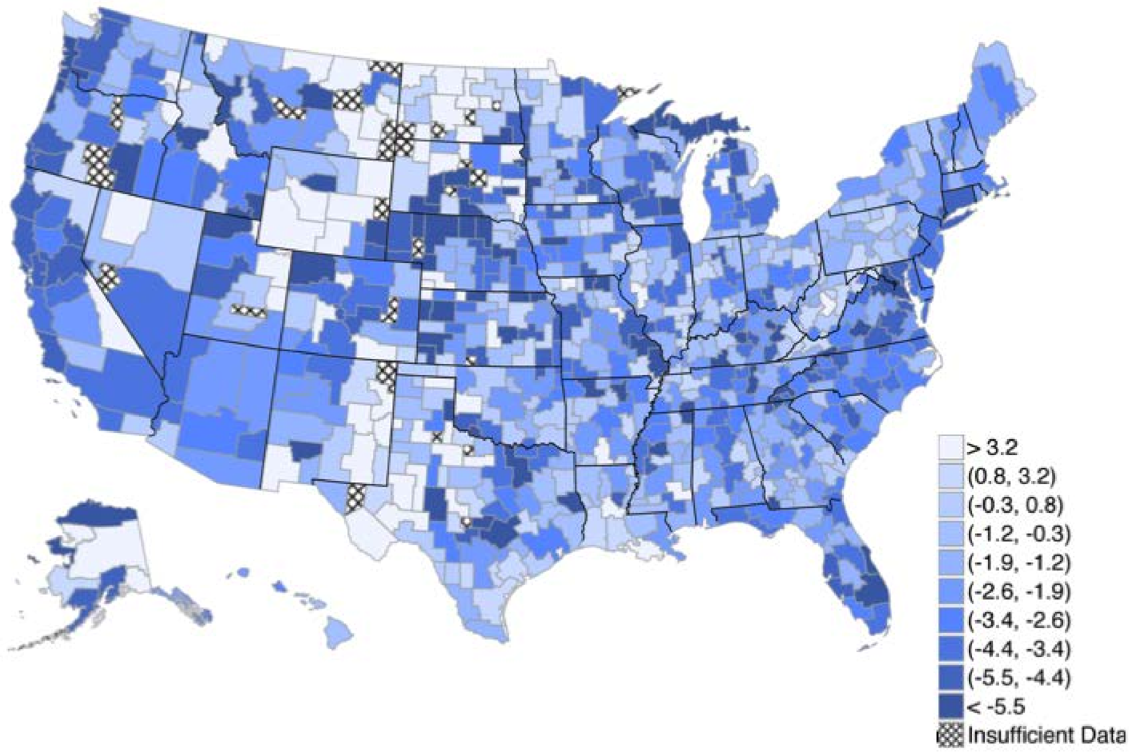
B. Employment



Notes: This figure presents binned scatter plots analogous to Figure III, but with the outcome being the event that the child reaches the top 10% of the income distribution at age 24 (Panel A) and the event that the child is employed at age 24 (Panel B), controlling for the mean rank predictions. In Panel A, we construct the event that the child is in the top 10% of the national (cohort-specific) income distribution. Using permanent parental residents in each CZ, we compute the fraction of children in the top 10% of the national cohort-specific income distribution. The blue series presents a non-parametric representation of the relationship between the event the child is in the top 10% and the predicted chance that the child is in the top 10% based on the prior residents in the destination CZ, controlling for the predicted chance the child is in the top 10% based on prior residents in the origin CZ and placebo controls for the predicted mean child rank in the origin and destination locations. Analogous to the binned scatter plots above, we partial out these controls, bin the residuals for the regression of the destination location into 20 equal bins, and plot the mean residual of the child outcome in each bin. For the red series, we instead plot the placebo relationship between the child being in the top 10% and the predicted mean rank of the child in the destination, controlling for the mean rank predictions in the origin and the top 10% predictions in both the origin and destination. In Panel B, we define employed as filing a W-2 at some point during the age of 24. We then repeat the above process replacing the event the child is in the top 10% with the event that the child is employed.

ONLINE APPENDIX FIGURE VI

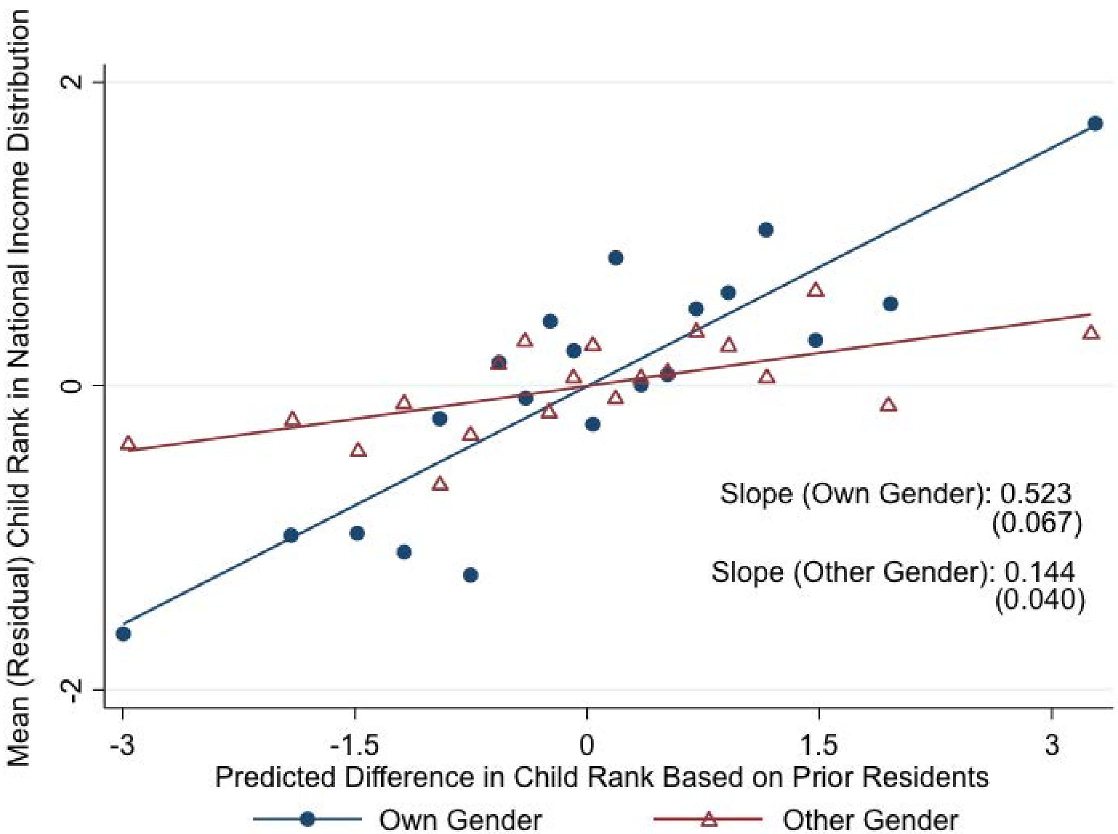
Map of Difference in Gender Outcomes, $\bar{y}^m_{pcs} - \bar{y}^f_{pcs}$, Evaluated at the 25th Percentile of Parental Income,



Notes: This figure presents estimates of the difference in male versus female outcomes of permanent residents, $\bar{y}^m_{pcs} - \bar{y}^f_{pcs}$ by CZ, c, for income at age 24. To estimate $\bar{y}^m_{pcs}$ and $\bar{y}^f_{pcs}$, we estimate linear regressions of child rank on parent income rank for each CZ on separate male and female samples, pooling cohorts 1980-1988.

ONLINE APPENDIX FIGURE VII

Movers' Outcomes vs. Gender-Specific Predicted Outcomes in Destination

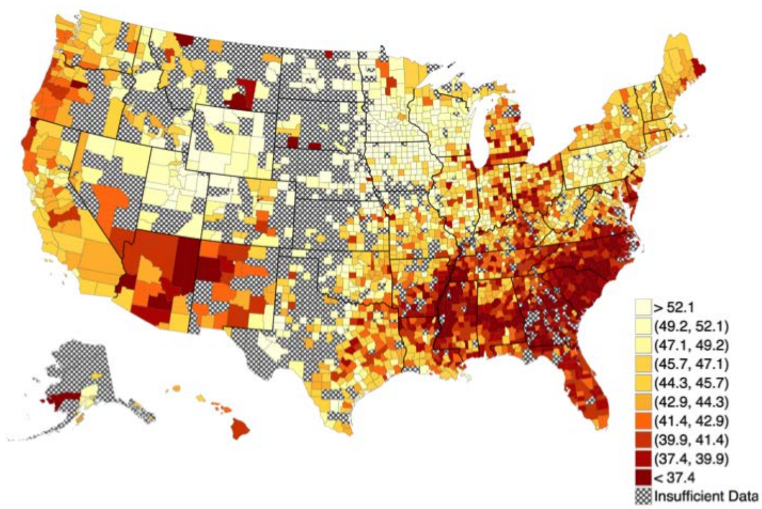


Notes: This figures presents binned scatter plots analogous to Figure III, but using gender-specific predicted outcomes based on prior residents. The blue series provides a non-parametric representation of the relationship between the child's own gender place prediction and the child's outcome; the red series provides a non-parametric representation of the relationship between the other (placebo) gender place predictions for the child's outcome, controlling for the own gender prediction. The sample includes all children in 1-time moving households whose parents moved when the child was less than or equal to 13 years old. Child income is measured when the child is age 26. For the blue circle series, we regress the own gender destination prediction for the child's outcome on the other gender destination prediction, other gender origin prediction, and own gender origin prediction. Similarly, we regress the child's income rank on the other gender destination prediction, other gender origin prediction, and own gender origin prediction. The figure then plots the relationship between the residuals from these regressions with sample means added to center the graphs. We construct 20 equal sized bins of the residuals from the destination regression and, in each bin, plot mean of the residuals from the child rank regression. For the red series, we repeat this process but using the placebo (other) gender predictions. We regress the other gender destination prediction for the child's outcome on the own gender destination prediction, other gender origin prediction, and own gender origin prediction. Similarly, we regress the child's income rank on the own gender destination prediction, other gender origin prediction, and own gender origin prediction. The red triangle series then plots the relationship between these residuals from these regressions with sample means added to center the graphs.

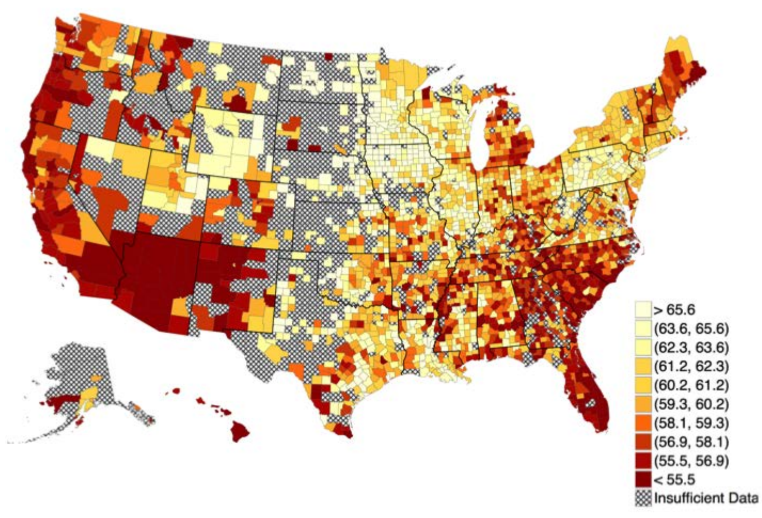
ONLINE APPENDIX FIGURE VIII

County-Level Predicted Income Rank at Age 30 - Permanent Residents

A. For Children with Parent at the 25^{th} Percentile)



B. For Children with Parent at the 75^{th} Percentile



Notes: These figures present the estimated $\bar{y}_{pcs}$ by CZ and County for $p = 25$ and $p = 75$.

TABLE I
Summary Statistics for CZ Permanent Residents and Movers

Variable	Mean (1)	Std. Dev. (2)	Median (3)	Sample Size (4)
Non-Movers				
Parent Income	89,909	357,194	61,300	19,499,662
Child family income at 24	24,731	140,200	19,600	19,499,662
Child family income at 26	33,723	161,423	26,100	14,894,662
Child family income at 30	48,912	138,512	35,600	6,081,738
Child individual income at 24	20,331	139,697	17,200	19,499,662
College attendence (18-23)	0.70	0.46	1.00	17,602,702
College quality (18-23)	31,833	13,466	31,600	17,602,702
Teen Birth (13-19)	0.11	0.32	0.00	9,670,225
Teen employment at age 16	0.28	0.45	0.00	19,499,662
Number of movers				
1 time	2,511,213			
2 times	1,478,177			
3 times	583,405			
4+ times	544,489			
Total	5,117,284			
1 time -3 times Movers				
Parent Income	89,526	372,354	53,500	4,572,795
Child family income at 24	23,431	57,448	18,100	4,572,795
Child family income at 26	31,535	98,091	23,700	3,424,020
Child family income at 30	46,116	106,599	32,300	1,364,025
Child individual income at 24	19,003	51,373	15,500	4,572,795
College attendence (18-23)	0.658	0.474	1.000	4,145,753
College quality (18-23)	30,439	13,212	30,000	4,145,753
Teen Birth (13-19)	0.130	0.337	0.000	2,266,897
Teen employment at age 16	0.280	0.449	0.000	4,572,795
One-time Movers				
Parent Income	97,064	369,971	58,700	1,553,021
Child family income at 24	23,867	56,564	18,600	1,553,021
Child family income at 26	32,419	108,431	24,500	1,160,278
Child family income at 30	47,882	117,450	33,600	460,457
Child individual income at 24	19,462	48,452	16,000	1,553,021
College attendence (18-23)	0.692	0.462	1.000	1,409,007
College quality (18-23)	31,418	13,489	31,100	1,409,007
Teen Birth (13-19)	0.114	0.317	0.000	769,717
Teen employment at age 16	0.277	0.448	0.000	1,553,021

Notes: The table presents summary statistics for the samples used in the CZ-level analyses. We split the summary statistics into the permanent residents ("non-movers") whose parents do not move across CZs throughout our sample window (1996-2012) and movers. Section II provides details on variable and sample definitions.

TABLE II
Exposure Effect Estimates

	Baseline Spec.							Family FE		
Specification:	Pooled	Age ≤ 23	Age ≤ 18	Claimed Sample	No Cohort Controls	Individual Income	Child CZ Fixed Effects	Baseline	No Cohort Controls	Inc/Mar. Controls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Exposure Slope	0.040	0.041	0.041	0.031	0.036	0.041	0.031	0.044	0.031	0.043
	(0.002)	(0.002)	(0.006)	(0.005)	(0.002)	(0.002)	(0.002)	(0.008)	(0.005)	(0.008)
Controls										
Cohort-Varying Intercept	X	X	X	X		X	X	X		X
Family FE								X	X	X
Child Income Definition	Family	Family	Family	Family	Family	Individual	Family	Family	Family	Family
Num of Obs.	1,553,021	1,287,773	687,323	604,602	1,553,021	1,553,021	1,473,218	1,553,021	1,553,021	1,553,021

Notes: Table II reports the coefficients on the child's age at the time of the parental move interacted with the difference in the predicted outcomes based on prior residents in the destination relative to the origin. Coefficients are multipled by -1 to correspond to exposure to destination. We allow separate lines for child age <= 23 and child age > 23 at the time of the parental move. Column (1) reports the coefficient γ in equation (6). Column (2) restricts the sample to those with age <= 23 at the time of the move. Column (3) restricts the sample to those below age 18. Column (4) further restricts to the sample of children who are claimed as a dependent on a 1040 in the destination CZ in the years subsequent to the move. Column (5) drops the cohort interactions with the predicted outcomes of permanent residents in the origin and destination location and instead includes one control for the predicted outcomes of those in the origin location. Column (6) presents the baseline specification (equation 6) using individual income for both the outcome and predicted outcomes in the origin and destination. Column (7) adds the child's CZ in adulthood (2012) as a fixed effect to the baseline linear specification in equation (6). This results in a slightly smaller sample size because of missing Zip Code information for some children in 2012 (e.g. a child with no earnings or taxable income). Column (8) adds family fixed effects to the baseline specification in equation (6). Column (9) adds family fixed effects to the specification in Column (5) that does not include cohort-varying intercepts. Column (10) considers the baseline specification in equation (6) and adds both family fixed effects and controls for changes in marital status and parental income in the year before versus after the move, along with their interactions with the age of the child at the time of the move.

TABLE III
Distributional Convergence

	Child Rank in top 10%			Child Employed		
	(1)	(2)	(3)	(4)	(5)	(6)
Distributional Prediction	0.043		0.040	0.046		0.045
	(0.002)		(0.003)	(0.003)		(0.004)
Mean Rank Prediction (Placebo)		0.022	0.004		0.021	0.000
		(0.002)	(0.003)		(0.002)	(0.003)
Num. of Obs.	1,553,021	1,553,021	1,553,021	1,553,021	1,553,021	1,553,021

Notes: Table III presents estimates of the exposure time relationships for the outcome of being in the top 10% of the cohort-specific income distribution at age 24 and being employed. We define employment as an indicator for filing a W-2 at some point during the year in which the child is age 24. Analogous to these outcomes, we construct predicted outcomes using permenent residents in each CZ. Column (1) presents the estimated exposure time slope using a top 10% indicator as the dependent variable and predicted outcomes based on permanent residents in the origin and destination CZ. Column (2) continues to use the indicator of being in the top 10% as the dependent variable, but uses the mean rank predictions from the baseline regressions as the origin and destination predictions. Column (3) combines all variables in specifications (1) and (2). Column (4) presents the estimated exposure time slope using an indicator of being employed as the dependent variable and predicted outcomes based on permanent residents in the origin and destination CZ. Column (5) retains the employment indicator as the dependent variable but replaces the predicted outcomes in the origin and destination with the mean rank predictions from the baseline regressions. Column (6) combines all variables in specifications (4) and (5).

TABLE IV
Gender Placebos

	No Family Fixed Effects			Family Fixed Effects			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Own Gender Prediction	0.038		0.031	0.031		0.027	0.0308
	(0.002)		(0.003)	(0.006)		(0.006)	(0.007)
Other Gender Prediction (Placebo)		0.034	0.009		0.017	0.017	0.0116
		(0.002)	(0.003)		(0.006)	(0.006)	(0.007)
Family Fixed Effects				X	X	X	X
Sample		Full Sample			Full Sample		2-Gender HH
Num. of Obs.	1,552,898	1,552,898	1,552,898	1,552,898	1,552,898	1,552,898	490964

Notes: Table presents estimates of the exposure time relationships using gender-specific predictions based on prior residents. The outcome is child rank when the child is 24 years old. Column (1) presents estimates for the baseline specification replacing the predicted outcomes based on prior residents in the origin and destination with gender-specific predictions. Column (2) replaces own-gender predicted outcomes with predicted outcomes in the origin and destination based on the other gender. Column (3) combines all variables in the specification in (1) and (2). Columns (4)-(6) repeat the specifications in (1)-(3) with the addition of family fixed effects. Column (7) repeats the specification in (6) but restricts to households with at least two children and at least one of each gender.

Appendix Table I

Population and Distance Restrictions

	Baseline	No Distance			100 Miles (Baseline)			200 Miles		
		Pop > 50K	Pop > 250K	Pop > 500K	Pop > 50K	Pop > 250K	Pop > 500K	Pop > 50K	Pop > 250K	Pop > 500K
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Exposure Slope	0.040	0.033	0.036	0.038	0.036	0.040	0.041	0.038	0.040	0.042
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Num of Obs.	1,553,021	3,002,272	2,143,525	1,562,615	2,062,277	1,553,021	1,165,153	1,655,105	1,288,816	991,657

Notes: This table presents estimates of the baseline specification in equation (6) varying the sample restriction. Column (1) presents the baseline sample restricting to populations in the origin and destination CZ of greater than 250,000 people based on the 2000 Census and requiring a distance of move > 100 miles between zipcode centroids. Columns (2)-(10) vary these distance and population restrictions.

Appendix Table II
Heterogeneity in Exposure Effects

		Parental Income		Moves	
	Baseline	Above Median Income	Below Median Income	Positive Moves	Negative Moves
	(1)	(2)	(3)	(4)	(5)
Exposure Slope	0.040	0.047	0.031	0.030	0.040
	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)
Num of Obs.	1,553,021	803,189	749,832	783,936	769,085

Notes: This table presents estimates of the heterogeneity in the baseline exposure time estimates (Column (1) of Table II) for various subsamples. Column (1) reports the baseline coefficient. Column (2) (Column (3)) restricts to moves by parents with above (below) median income (median defined as parent rank = 0.5; note there are more observations of 1x movers with parent rank > 0.5, reflecting the fact that the likelihood of moving is increasing in parental income). Column (4) (Column (5)) restricts to moves in which the predicted outcomes based on prior residents in the destination are higher (lower) than in the origin.

Appendix Table III
Multiple Moves

	Separate Coeffs (1)	Constrained Coeff (2)
1st Destination	0.040	
	(0.001)	
2nd Destination	0.037	
	(0.004)	
3rd Destination	0.031	
	(0.006)	
Constrained Specification		0.039
		(0.001)
Num of Obs.	4,374,418	4,374,418

Notes: This table reports results using our expanded sample of 1-3 time movers. Column (1) presents estimates for the exposure effect of the 1st, 2nd, and 3rd destination by adding exposure effect coefficients corresponding to each move, and using the sample of 1-3-time movers, as opposed to the 1-time movers sample. Column (2) presents the estimates of the exposure effect restricting the coefficient to be the same across each move.

Appendix Table IV
County Exposure Effect Estimates

Specification:	Baseline Spec.		Within CZ Moves				
	Baseline	Family FE	Age 24	Age 26	Age ≥ 24	Family FE	Small Ctys
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exposure Slope	0.035	0.033	0.022	0.032	0.027	0.029	0.024
	(0.003)	(0.011)	(0.003)	(0.004)	(0.003)	(0.025)	(0.002)
Num of Obs.	654,491	654,491	617,502	457,140	2,900,311	2,900,311	7,311,431

Notes: Appendix Table IV reports exposure effect coefficients in equation (6), analogous to those presented in Table II, using county-level predictions for the sample of 1-time county movers. Column (1) presents the baseline specification analogous to Column (1) of Table II, replacing CZ-level predictions with county-level predictions based on prior residents. We restrict the sample to moves of at least 100 miles and require the county-level population to be at least 250,000 in the origin and destination county. Column (2) adds family fixed effects to the specification in Column (1). Columns (3)-(7) drop the distance restriction and consider the set of within-CZ county moves (between counties with populations of at least 250,000). Column (3) replicates the baseline specification. Column (4) replicates the baseline specification using income at age 26 as the outcome, analogous to the outcomes considered in Section V. Column (5) presents the pooled estimate that stacks all outcomes for ages 24 and above (multiple observations per person). Column (6) adds family-by-age of outcome fixed effects to the specification in Column (5). Column (7) expands the sample in Column (5) to include moves between all CZs with populations above 10,000).

Appendix Table V
Summary Statistics for County Permanent Residents and Movers

Variable	Mean (1)	Std. Dev. (2)	Median (3)	Sample Size (4)
Non-Movers				
Parent Income	81,932	320,026	54,800	37,689,238
Child family income at 24	25,066	136,016	19,900	19,956,828
Child family income at 26	34,091	157,537	26,600	15,364,222
Child family income at 30	48,941	133,264	36,200	6,355,414
Child individual income at 24	20,686	202,833	17,300	20,069,124
College attendence (18-23)	0.703	0.457	1.000	20,418,691
College quality (18-23)	31,608	13,207	31,400	20,418,691
Teen Birth (13-19)	0.107	0.309	0.000	14,503,588
Teen employment at age 16	0.276	0.447	0.000	37,464,779
One-time Movers Across CZ Sample				
Parent Income	94,738	400,685	55,100	1,498,319
Child family income at 24	23,815	72,306	18,200	654,491
Child family income at 26	32,532	139,563	24,300	483,407
Child family income at 30	48,834	110,619	33,500	188,801
Child individual income at 24	20,247	61,185	16,000	654,491
College attendence (18-23)	0.717	0.451	1.000	690,207
College quality (18-23)	32,171	14,001	31,900	690,207
Teen Birth (13-19)	0.103	0.304	0.000	524,194
Teen employment at age 16	0.233	0.423	0.000	1,498,319
One-time Movers Within CZ Sample				
Parent Income	84,850	356,758	48,900	1,425,096
Child family income at 24	24,006	68,559	18,300	617,502
Child family income at 26	32,993	75,520	24,500	457,140
Child family income at 30	49,974	108,248	33,500	179,856
Child individual income at 24	20,844	56,639	16,500	617,502
College attendence (18-23)	0.719	0.450	1.000	650,045
College quality (18-23)	32,883	14,086	33,200	650,045
Teen Birth (13-19)	0.095	0.293	0.000	496,122
Teen employment at age 16	0.245	0.430	0.000	1,425,096

Notes: The table presents summary statistics for county movers sample discussed in Online Appendix A.

Appendix Table VI

Correlations of Permanent Resident Outcomes Across CZs

	Child Rank at Age 24	Child Rank at Age 26	Child Rank at Age 30	Child Indiv. Inc. Rank at Age 26	Child Rank at Age 24 in Top 10%	Child Employed at Age 24	Child Rank at Age 24 for Females	Child Rank at Age 24 for Males	College Attendance Age 18-23	Married at age 26
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Below Median Income Parents (p25)										
Child Rank at Age 24	1.000									
Child Rank at Age 26	0.970	1.000								
Child Rank at Age 30	0.867	0.956	1.000							
Child Indiv. Inc. Rank at Age 26	0.666	0.774	0.864	1.000						
Child Rank at Age 24 in Top 10%	0.852	0.827	0.708	0.386	1.000					
Child Employed at Age 24	0.707	0.670	0.592	0.436	0.586	1.000				
Child Rank at Age 24 for Females	0.932	0.973	0.948	0.830	0.743	0.594	1.000			
Child Rank at Age 24 for Males	0.961	0.979	0.920	0.686	0.865	0.710	0.905	1.000		
College Attendance Age 18-23	0.149	0.294	0.461	0.628	-0.077	0.052	0.401	0.183	1.000	
Married at age 26	0.644	0.554	0.380	-0.076	0.758	0.486	0.439	0.636	-0.354	1.000
Above Median Income Parents (p75)										
Child Rank at Age 24	1.000									
Child Rank at Age 26	0.942	1.000								
Child Rank at Age 30	0.734	0.871	1.000							
Child Indiv. Inc. Rank at Age 26	0.476	0.555	0.794	1.000						
Child Rank at Age 24 in Top 10%	0.793	0.807	0.535	0.093	1.000					
Child Employed at Age 24	0.666	0.645	0.587	0.503	0.304	1.000				
Child Rank at Age 24 for Females	0.912	0.977	0.831	0.496	0.791	0.603	1.000			
Child Rank at Age 24 for Males	0.935	0.984	0.873	0.583	0.792	0.657	0.923	1.000		
College Attendance Age 18-23	-0.211	-0.062	0.242	0.300	-0.300	0.000	-0.042	-0.078	1.000	
Married at age 26	0.531	0.507	0.113	-0.423	0.749	0.213	0.549	0.457	-0.374	1.000

Notes: This table presents correlations of permanent resident outcomes across CZs for a range of outcomes. All correlations weighted by population in the 2000 Census.

Appendix Table VII

Family Fixed Effect Specifications using Less Precise Permanent Resident Forecasts

	Baseline		Pop > 250K		Pop > 100K		Pop > 50K		Pop > 0	
	No Fam FE	Fam FE	No Fam FE	Fam FE	No Fam FE	Fam FE	No Fam FE	Fam FE	No Fam FE	Fam FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Exposure Slope	0.040	0.044	0.036	0.031	0.032	0.024	0.032	0.023	0.031	0.020
	(0.002)	(0.008)	(0.002)	(0.005)	(0.001)	(0.004)	(0.001)	(0.003)	(0.001)	(0.003)
Family FE Attenuation		1.10		0.86		0.76		0.71		0.63
Family FE		X		X		X		X		X
Cohort-Varying Intercept	X	X								
Population Restriction	>250K	>250K	>250K	>250K	>100K	>100K	>50K	>50K	> 0	> 0
Num of Obs.	1553021	1553021	1553021	1553021	1914623	1914623	2062277	2062277	2162194	2162194

Notes: This table presents robustness specifications for the baseline specification (Column 1 of Table II) and family fixed effect specification (Column 8 of Table II) that remove the cohort-varying intercept and adopt alternative population restrictions in the origin and destination CZs. Columns 3-4 present the baseline population restriction that requires at least 250,000 residents in both the origin and destination, but drops the cohort varying intercept. Columns 5-6 consider the specification in Columns 3 and 4 but extend the sample to include moves in which the origin and destination CZ has at least 100,000 residents. Columns 7-8 expand the population restriction to 50,000 residents, and Columns 9-10 drop the population restriction entirely.