

THE IMPACTS OF NEIGHBORHOODS ON INTERGENERATIONAL MOBILITY II: COUNTY-LEVEL ESTIMATES*

RAJ CHETTY AND NATHANIEL HENDREN

We estimate the causal effect of each county in the United States on children's incomes in adulthood. We first estimate a fixed effects model that is identified by analyzing families who move across counties with children of different ages. We then use these fixed effect estimates to (i) quantify how much places matter for intergenerational mobility, (ii) construct forecasts of the causal effect of growing up in each county that can be used to guide families seeking to move to opportunity, and (iii) characterize which types of areas produce better outcomes. For children growing up in low-income families, each year of childhood exposure to a one standard deviation (std. dev.) better county increases income in adulthood by 0.5%. There is substantial variation in counties' causal effects even within metro areas. Counties with less concentrated poverty, less income inequality, better schools, a larger share of two-parent families, and lower crime rates tend to produce better outcomes for children in poor families. Boys' outcomes vary more across areas than girls' outcomes, and boys have especially negative outcomes in highly segregated areas. Areas that generate better outcomes have higher house prices on average, but our approach uncovers many "opportunity bargains"—places that generate good outcomes but are not very expensive. *JEL Codes:* J62, C00, R00.

*An earlier version of this article was circulated as the second part of the article, "The Impacts of Neighborhoods on Intergenerational Mobility: Childhood Exposure Effects and County Level Estimates." The opinions expressed in this article 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 article 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. We thank Gary Chamberlain, Maximilian Kasy, Lawrence Katz, Jesse Shapiro, and numerous seminar participants for helpful comments and discussions. Sarah Abraham, Alex Bell, Augustin Bergeron, Michael Droste, Niklas Flamang, Jamie Fogel, Robert Fluegge, Nikolaus Hildebrand, Alex Olssen, Jordan Richmond, Benjamin Scuderi, Priyanka Shende, and our other predoctoral fellows provided outstanding research assistance. This research was funded by the National Science Foundation, the Lab for Economic Applications and Policy at Harvard, Stanford University, and the Laura and John Arnold Foundation.

I. INTRODUCTION

How are children's economic opportunities shaped by the neighborhoods in which they grow up? In the first article in this series (Chetty and Hendren 2018a), we showed that neighborhoods have significant childhood exposure effects on children's life outcomes. Although those results establish that place matters for intergenerational mobility, they do not tell us which areas produce the best outcomes, nor do they identify the characteristics of neighborhoods that generate good outcomes—two key inputs necessary for developing place-focused policies to improve children's outcomes.

In this article, we build on the exposure-time design developed in our first article to estimate the causal effect of each county in the United States on children's incomes in adulthood. Formally, our first article identified one treatment effect—the average impact of exposure to an area where children have better outcomes—while this article pursues the more ambitious goal of identifying (approximately) 3,000 treatment effects, one for each county in the country.¹

We estimate counties' causal effects on children's ranks in the income distribution at age 26 using data from deidentified tax returns for all children born between 1980 and 1986.² We estimate each county's effect using a fixed effects regression model identified by analyzing families who move across counties, exploiting variation in children's ages when families move. To understand how the model is identified, consider families in the New York area. If children who move from Manhattan to Queens at younger ages earn more as adults, we can infer that growing up in Queens has a positive causal effect relative to growing up in Manhattan under the assumption that other determinants of children's

1. To maximize statistical precision, we characterize neighborhood (or "place") effects at the county level. We recognize that counties are much larger than the typical geographic units used to define "neighborhoods." In the presence of heterogeneity across local areas within counties, the county-level effects we estimate can be interpreted as weighted averages of the local area effects. In future work, the methods we develop here could be applied to estimate place effects at smaller geographies, such as census tracts.

2. We measure incomes at age 26 because children's mean ranks in each area tend to stabilize by age 26. For example, the (population-weighted) correlation between mean income ranks at age 26 and age 32 across commuting zones is 0.93 for children growing up in low-income (25th percentile) families and 0.77 for children growing up in high-income (75th percentile) families.

outcomes are unrelated to the age at which they move. Building on this logic, we use our sample of cross-county movers to regress children's income ranks at age 26 on fixed effects for each county interacted with the fraction of childhood spent in that county. We estimate the county fixed effects separately by parent income level, permitting the effects of each area to vary with parent income. We include origin-by-destination fixed effects when estimating this model, so that each county's effect is identified purely from variation in the age of children when families make a given move rather than variation in where families move.

The key assumption required to identify counties' causal effects using this research design is that children's potential outcomes are orthogonal to the age at which they move to a given county. This assumption is motivated by the evidence in our first article showing that the age at which children move to an area where permanent residents (nonmovers) have better or worse outcomes on average is orthogonal to their potential outcomes. However, it is a stronger requirement than the condition required to identify average exposure effects in our first article because it imposes 3,000 orthogonality conditions—one for each county—rather than a single orthogonality condition that must hold on average.

We assess the validity of this stronger identification assumption using two approaches. First, we show that controlling for parental income levels and marital status in the years before and after the move, which are strong predictors of children's outcomes, does not affect the estimates, supporting the view that our estimates are not confounded by selection on other determinants of children's outcomes. Second, we implement placebo tests by (i) estimating each area's fixed effect on teenage labor force participation rates at age 16 (a strong predictor of incomes in adulthood), using the subsample of families who move after age 16; and (ii) estimating each area's effect on income at age 26 using parents who move after their children turn 23, the point at which neighborhood exposure no longer appears to affect children's outcomes based on the evidence in our first article. These placebo fixed effect estimates are uncorrelated with our baseline estimates, supporting the assumption that the time at which parents move to a given county is orthogonal to their children's potential outcomes.

We use the estimates of counties' causal effects for three purposes. First, we quantify how much neighborhoods matter for children's incomes. We model the estimated county effects as the sum

of a latent causal effect and noise due to sampling error and estimate the signal variance of the latent causal effects. For a child with parents at the 25th percentile of the national income distribution, we find that spending one additional year of childhood in a one standard deviation better county (population-weighted) increases household income at age 26 by 0.17 percentile points, equivalent to an increase in mean income of approximately 0.5%. Extrapolating over 20 years of childhood, growing up in a one standard deviation better county from birth would increase a child's household income in adulthood by approximately 10%.³

Neighborhoods have similar effects in percentile rank or dollar terms for children of higher-income parents, but matter less in percentage terms because children in high-income families have higher mean incomes. For children with parents at the 75th percentile of the income distribution, the signal standard deviation of annual exposure effects across counties is 0.16 percentiles, which is approximately 0.3% of mean income. Importantly, areas that generate better outcomes for children in low-income families also generate slightly better outcomes on average for children in high-income families. This result suggests that the success of the poor does not have to come at the expense of the rich.

In the second part of our analysis, we construct forecasts of the causal effect of growing up in each county that can be used to guide families seeking to move to better areas. Formally, we construct forecasts that minimize the mean squared error (MSE) of the predicted impact of growing up in a given neighborhood relative to the true impact. Although the raw county fixed effects provide unbiased estimates of counties' causal effects, they do not themselves provide good forecasts because many of the estimates have substantial noise, leading to high MSE. In highly populated counties, such as Cook County (the city of Chicago), nearly 75% of the variance in the fixed effect estimates is signal; however, in most counties, more than half of the variance in the fixed effect estimates is due to noise from sampling variation.

3. We focus primarily on estimates of place effects on household income (including spousal income), but also report estimates using individual income below. The household income estimates are highly correlated with the individual income estimates for men, whose outcomes are typically used as a measure of economic opportunities in the literature on intergenerational mobility because the variance in earnings due to differences in labor force participation rates is smaller for men than women (e.g., Solon 1999).

To obtain forecasts that have lower MSE, we use a shrinkage estimator that incorporates data on the permanent residents' (nonmovers) outcomes in each area.⁴ The permanent residents' mean outcomes have very little sampling error but are imperfect forecasts of a county's causal effect because they combine causal effects with sorting. The best MSE-minimizing linear forecast of each county's causal effect is therefore a weighted average of the fixed effect estimate based on the movers and a prediction based on permanent residents' outcomes, with greater weight on the fixed effect estimate when it is more precisely estimated (i.e., in large counties).

Among the 100 most populated counties in the country, DuPage County, IL, is forecasted to generate the highest incomes for children growing up in low-income (25th percentile) families. Each additional year that a child in a low-income family spends in DuPage County instead of the average county in the United States raises his or her household income in adulthood by 0.80%. Growing up in DuPage County from birth—that is, having about 20 years of exposure to that environment—would raise such a child's income by 16%. In contrast, growing up in Cook County (one of the lowest-ranking counties in the United States) from birth reduces a child's income by approximately 13%. Hence, moving from Cook County (the city of Chicago) to DuPage County (the western suburbs) at birth would increase a child's income by about 30% on average.⁵

We find that neighborhoods matter more for boys than girls: the signal standard deviation of county-level effects is roughly

4. Our methodology contributes to a recent literature that builds on empirical Bayes methods dating to [Robbins \(1956\)](#) by using shrinkage estimators to reduce MSE (risk) when estimating a large number of parameters. For instance, [Angrist et al. \(2017\)](#) combine experimental and observational estimates to improve forecasts of school value added. Our methodology differs from theirs because we have unbiased (quasi-experimental) estimates of causal effects for every area, whereas Angrist et al. have unbiased (experimental) estimates of causal effects for a subset of schools. [Hull \(2017\)](#) develops methods to forecast hospital quality, permitting nonlinear and heterogeneous causal effects. [Abadie and Kasy \(2017\)](#) show how machine learning methods can be used to reduce risk, using the fixed effect estimates constructed in this article as an application.

5. Interestingly, many families involved in the well-known Gautreaux housing desegregation project moved from Cook County to DuPage County. Our results support the view that much of the gains experienced by the children of the families who moved as part of Gautreaux ([Rosenbaum 1995](#)) was due to the causal effect of exposure to better neighborhoods.

60% larger for boys than girls in low-income (25th percentile) families. The distribution has an especially thick lower tail for boys, as counties with high concentrations of urban poverty such as Baltimore City and Wayne County (Detroit) produce especially negative outcomes for boys.⁶

Our estimates of the causal effects of counties and commuting zones (CZs) are highly correlated with the observational statistics on intergenerational mobility reported in Chetty et al. (2014), as expected given the findings in Chetty and Hendren (2018a), but there are many significant differences. For example, children who grow up in low-income families in New York City have outcomes comparable to the national mean, but the causal effect of growing up in New York City—as revealed by analyzing individuals who move into and out of New York—is well below the national mean. One potential explanation for this pattern is that New York has a large share of immigrants, who tend to have high rates of upward mobility (Hilger 2016). More generally, this example illustrates the importance of estimating the causal effect of each area directly using movers (as we do in this article) rather than predicting neighborhood effects purely from permanent residents' outcomes.

In the third part of our analysis, we characterize the properties of CZs and counties that produce good outcomes for low-income children (i.e., generate high rates of upward mobility). Prior work has shown that in observational data, upward mobility is highly correlated with area characteristics, such as residential segregation, income inequality, social capital, and school quality, as well as demographic characteristics, such as the fraction of children being raised by single mothers and racial shares (Wilson 1987; Sampson, Morenoff, and Gannon-Rowley 2002; Chetty et al. 2014). However, it is unclear whether these correlations are driven by the causal effects of place or selection effects. For instance, is growing up in a less segregated area beneficial for a given child or do families who choose to live in less segregated areas simply have better unobservable characteristics?

We decompose the correlations documented in prior work into causal versus sorting components by correlating each characteristic with both our causal effect estimates and permanent

6. These gender differences are partly related to differences in rates of marriage. For example, the San Francisco area generates high individual incomes but relatively low household incomes for girls because growing up in San Francisco reduces the probability that a child gets married.

residents' outcomes, which combine both causal and selection effects. We find that many of the correlations between area-level characteristics and upward mobility are driven almost entirely by causal effects of place. For example, 80% of the association between segregation and upward mobility across CZs in observational data is driven by the causal effect of place; only 20% is due to sorting. Growing up in a CZ with a one standard deviation higher level of segregation from birth reduces the income of a child in a low-income (25th percentile) family by 5.2%.⁷ Urban areas, particularly those with concentrated poverty, generate particularly negative outcomes for low-income children. These findings support the view that growing up in an urban "ghetto" reduces children's prospects for upward mobility (Massey and Denton 1993; Cutler and Glaeser 1997).

Areas with greater income inequality—as measured by the Gini coefficient or top 1% income shares—also generate significantly worse outcomes for children in low-income families. Hence, the negative correlation between inequality and intergenerational mobility documented in prior work—coined the "Great Gatsby curve" by Krueger (2012)—is not simply driven by differences in genetics or other characteristics of populations in areas with different levels of inequality. Rather, putting a given child in an area with higher levels of inequality makes that child less likely to rise up in the income distribution. The negative correlation between the causal effects and top 1% shares contrasts with the findings of Chetty et al. (2014), who find no correlation between top 1% shares (upper-tail inequality) and rates of upward mobility in observational data. Our analysis of movers reveals that low-income families who live in areas with large top 1% shares (such as New York City) are positively selected, masking the negative association between top 1% shares and the causal effect of places on upward mobility in observational data.

We find strong correlations between areas' causal effects and output-based measures of school quality, such as test scores adjusted for parental income levels. We also find strong correlations between the causal effects and proxies for social capital, such as crime rates and Rupasingha and Goetz's (2008) summary index.

7. This result does not necessarily imply that reducing segregation in a given area will improve children's outcomes. Other factors associated with less segregation (e.g., better schools) could potentially be responsible for the gain a child obtains from moving to a less segregated area.

Selection plays a bigger role in explaining correlations between demographic characteristics and upward mobility in observational data. For example, the fraction of single mothers is the single strongest predictor of differences in upward mobility for permanent residents across areas. However, the fraction of single mothers (although still a significant predictor) is less highly correlated with CZs' causal effects on upward mobility than other factors, such as segregation. This is because nearly half of the association between permanent residents' outcomes and the fraction of single mothers is due to selection. Similarly, areas with a larger African American population have significantly lower rates of upward mobility in observational data. Roughly half of this association is also unrelated to the causal effect of place, consistent with Rothbaum (2016). Nevertheless, the correlation between the causal effects of place and the African American share remains substantial (-0.51 across CZs and -0.37 across counties within CZs). Place effects therefore amplify racial inequality: black children have worse economic outcomes because they grow up in worse neighborhoods.

Finally, we examine how much more one has to pay for housing to live in an area that generates better outcomes for one's children. Within CZs, counties that produce better outcomes for children have slightly higher rents, especially in highly segregated cities. However, rents explain less than 5% of the variance in counties' causal effects for families at the 25th percentile. This result suggests that current "small area" fair market rent proposals in housing voucher programs - which condition voucher payments on local neighborhood rents—may not maximize vouchers' effects on upward mobility, as many areas that are more expensive do not produce better outcomes. Moreover, it shows that some areas are "opportunity bargains"—counties within a labor market that offer good outcomes for children without higher rents.⁸ For example, in the New York metro area, Hudson County, NJ, offers much higher levels of upward mobility than does Manhattan or Queens, despite having comparable rents during the period we study.

To understand the source of these opportunity bargains, we divide our causal county effects into the component that projects

8. Of course, the areas that are "opportunity bargains" in rents may come with other disamenities, such as longer commutes to work, that might make them less desirable. Our point is simply that housing costs themselves are not necessarily a barrier to moving to opportunity.

onto observable area-level factors, such as poverty rates and school expenditures, and a residual “unobservable” component. We find that only the observable component is capitalized in rents, suggesting that the opportunity bargains may partly exist because families do not know which neighborhoods have the highest value added. This result underscores the importance of measuring neighborhood quality directly using children’s observed outcomes instead of using traditional proxies such as poverty rates.

The rest of this article is organized as follows. In [Section II](#), we summarize the data, focusing on differences relative to the sample used in our first article. In [Section III](#), we formalize our empirical objectives using a statistical model. [Section IV](#) reports the baseline fixed effect estimates and evaluates the validity of the key identification assumptions. [Section V](#) quantifies the magnitude of place effects, [Section VI](#) presents the CZ- and county-level MSE-minimizing forecasts, and [Section VII](#) examines the characteristics of places that generate better outcomes. [Section VIII](#) presents the results on housing costs and opportunity bargains. [Section IX](#) concludes. Supplementary results and details on estimation methodology are provided in an [Online Appendix](#). Estimates of CZs’ and counties’ causal effects are available on the Equality of Opportunity Project [website](#).

II. DATA AND SAMPLE DEFINITIONS

We use data from federal income tax records spanning 1996–2012. Our primary analysis sample is the same as the sample of movers in [Chetty and Hendren \(2018a, Section II\)](#), with three exceptions.

First, we limit the sample to children in the 1980–1986 birth cohorts because we measure children’s incomes at age 26. Measuring children’s incomes at age 26 strikes a balance between the competing goals of minimizing life cycle bias by measuring income at a sufficiently old age and having an adequate number of birth cohorts to implement our research design. Among permanent residents (parents who stay in the same CZ from 1996 to 2012) at the 25th percentile of the income distribution, the population-weighted correlation between children’s mean ranks at age 26 and age 32 across CZs is 0.93.⁹ This suggests that measuring

9. The key point here is that children’s average ranks in each area stabilize by age 26. At the individual level, children’s incomes stabilize later, around age 32

children's incomes at later ages would not affect our estimates of places' causal effects substantially.

Second, we focus on the subset of families who move across CZs or counties when their child is 23 or younger, motivated by our finding that childhood exposure effects persist until age 23. To simplify estimation, we focus on families who move exactly once during the sample period, dropping those who move multiple times. We divide the sample of one-time movers into two groups—those who move across CZs and those who move across counties within CZs—and analyze these two sets of movers separately.

Third, to maximize precision, we include all movers, not just those who move between large CZs in our analysis sample. However, we only report estimates of causal effects for CZs with populations above 25,000 and counties with populations above 10,000 in the 2000 census (excluding 0.36% of the population).¹⁰

We measure children's and parents' incomes at the household level using data from 1040 forms (for those who file tax returns) and W-2 forms (for nonfilers), which we label family (or household) income. We identify individuals' locations in each year using the ZIP code from which they filed their tax returns or to which their W-2 forms were mailed. We also measure a set of additional outcomes for children, such as individual income, college attendance, and marriage. All of these variables are defined in the same way as described in Section II of Chetty and Hendren (2018a).

Table I presents summary statistics for children in our primary sample who move across CZs (Panel A) and counties within CZs (Panel B). There are 1,397,260 children whose parents move once across CZs and 931,138 children whose parents move across counties within CZs in our primary sample. The sample

(Haider and Solon 2006, figure 2a). Intuitively, children's ranks change rapidly in their late 20s, but these idiosyncratic individual-level changes average out at the area level, so that areas where children have high ranks at 26 tend to have high ranks at age 32 as well.

10. In Chetty and Hendren (2018a), we limited our primary sample to CZs with populations above 250,000 to minimize attenuation bias in exposure effect estimates resulting from noise in permanent residents' outcomes. This attenuation bias does not arise here because we identify causal effects purely from the sample of movers, without projecting their outcomes onto permanent residents. This is why we impose lower population restrictions here, providing estimates for a larger set of CZs. Because our goal is to provide estimates of place effects at the county (rather than CZ) level, we also do not impose any restrictions on the distance of moves we examine.

TABLE I
SUMMARY STATISTICS FOR MOVERS ANALYSIS SAMPLES

Variable	Mean (1)	Std. dev. (2)	Median (3)	Sample size (4)
Panel A: Between CZ movers				
Parent family income (\$)	89,029	353,465	56,700	1,397,260
Child family income at 26 (\$)	31,706	88,503	24,300	1,397,260
Child family income at 30 (\$)	45,890	99,172	33,100	459,952
Child individual income at 26 (\$)	23,731	79,083	19,900	1,397,260
Child married at 26 (%)	26.5	44.1	0.0	1,270,634
Panel B: County within CZ movers				
Parent family income (\$)	82,627	300,952	57,000	931,138
Child family income at 26 (\$)	32,304	62,314	25,000	931,138
Child family income at 30 (\$)	46,477	86,911	33,800	316,106
Child individual income at 26 (\$)	24,260	49,620	20,700	931,138
Child married at 26 (%)	25.9	43.8	0.0	842,547

Notes. This table presents summary statistics for the primary analysis samples used to estimate the causal effects of counties and CZs. The sample consists of individuals who (i) have a valid Social Security Number or Individual Taxpayer Identification Number, (ii) were born between 1980 and 1986, (iii) are U.S. citizens as of 2013, and (iv) were claimed as a child dependent at some point between 1996–2012. Panel A includes the subset of children satisfying these restrictions whose families moved exactly once across commuting zones between 1996 and 2012 before they turned 23. Panel B includes children whose families moved across counties within a CZ exactly once between 1996 and 2012 before they turned 23. Parent family income is the average pretax household income from 1996 to 2000, measured as AGI plus tax-exempt interest income and the non-taxable portion of Social Security and Disability (SSDI) benefits for tax-filers, and using information returns for nonfilers. Child family income is measured analogously at various ages, while child individual income is defined as the sum of individual W-2 wage earnings, unemployment insurance benefits, SSDI payments, and half of household self-employment income. See [Section II of Chetty and Hendren \(2018a\)](#) for additional details on sample and variable definitions. All dollar values are reported in 2012 dollars, deflated using the CPI-U.

characteristics are generally very similar to those reported in [Table I of Chetty and Hendren \(2018a\)](#), with a median family income of \$24,253 at age 26 for children in the CZ movers sample and \$24,993 in the county-within-CZ movers sample (in 2012 dollars).

III. EMPIRICAL FRAMEWORK

In this section, we define the estimands we seek to identify using a statistical model of neighborhood effects. We then describe the research design we use to identify these parameters and the key identification assumptions underlying our analysis. Finally, we discuss the empirical specification and estimation procedures we use to implement this research design.

III.A. Statistical Model

We estimate place effects using a statistical model motivated by the childhood exposure effects documented in [Chetty and](#)

[Hendren \(2018a\)](#). Let y_i denote a child's income (or other outcome) in adulthood, measured at age T . We model y_i as a function of three factors: the neighborhoods in which the child grows up, disruption costs of moving across neighborhoods, and all other non-neighborhood inputs, such as family environment and genetics.

Let $c(i, a)$ denote the place in which child i lives at age $a = 1, \dots, A$ of his childhood, where $A < T$. Let μ_c denote the causal effect of one additional year of exposure to place c on the child's outcome y_i .¹¹ Given the linear childhood exposure effects documented in [Chetty and Hendren \(2018a, Figure IV\)](#), we assume that the exposure effect μ_c is constant for ages $a \leq A$ and is 0 thereafter. Let κ denote the cost of moving from one neighborhood to another during childhood (e.g., due to a loss of connections to friends or other fixed costs of moving). Finally, let θ_i denote the impact of other factors, such as family inputs. The parameter θ_i captures both time-invariant inputs, such as genetic endowments, and the total amount of time-varying inputs, such as parental investments during childhood.

Combining the effects of neighborhoods, disruption effects of moving, and other factors, the child's outcome is given by

$$(1) \quad y_i = \sum_{a=1}^A [\mu_{c(i,a)} - \kappa \mathbf{1}\{c(i,a) \neq c(i,a-1)\}] + \theta_i.$$

The production function for y_i in [equation \(1\)](#) imposes three substantive restrictions that are relevant for our empirical analysis. First, it assumes that neighborhood effects μ_c do not vary across children (conditional on parent income).¹² Second,

11. In our empirical application, we permit place effects μ_{pc} to vary with parental income rank $p(i)$, but we suppress the parental income index in this section to simplify notation.

12. This constant treatment effects assumption is a common simplification in the literature. For instance, in work on firm effects and teacher effects ([Abowd, Kramarz, and Margolis 1999; Chetty, Friedman, and Rockoff 2014a](#)), analogous restrictions rule out worker-firm or teacher-student match effects. Our fixed effect estimates can be interpreted as mean place effects in the presence of heterogeneous treatment effects if such heterogeneity is orthogonal to individuals' transition rates across areas, that is, as long as there is no "essential" heterogeneity. Understanding how the present estimates of μ_{pc} can be interpreted in the presence of essential heterogeneity and estimating models that permit richer forms of heterogeneity are important directions for further work.

it assumes that place effects are additive and constant across ages, that is, that there are no complementarities between neighborhood effects across years. Third, it assumes that the disruption costs of moving κ do not vary across neighborhoods or the age of the child at the time of the move.¹³ We believe these restrictions are reasonable approximations given the findings of our first article, which show that childhood exposure effects are constant throughout childhood and symmetric when children move to areas with better or worse permanent resident outcomes on average (conditional on parent income). Nevertheless, we view estimating models that permit richer forms of heterogeneity in place effects as an important direction for future research.¹⁴

Our objective in this article is to identify $\vec{\mu} = \{\mu_c\}$, the causal exposure effect of spending a year of one's childhood in a given area (CZ or county) of the United States. One way to identify $\vec{\mu}$ would be to randomly assign children of different ages to different places and compare their outcomes, as in the Moving to Opportunity experiment (Chetty, Hendren, and Katz 2016). Because conducting such an experiment in all areas of the country is infeasible, we develop methods of identifying place effects in observational data.

III.B. Identifying Place Effects in Observational Data

Building on the approach in Chetty and Hendren (2018a), we identify $\vec{\mu}$ by exploiting variation in the timing of when children move across areas. To understand the intuition underlying our approach, consider a set of children who move from a given origin o (e.g., New York) to a given destination d (e.g., Boston). Suppose that children who make this move at different ages have comparable other inputs, θ_i . Then one can infer the causal effect of growing up in Boston relative to New York ($\mu_d - \mu_o$) by comparing

13. The model can be extended to allow the disruption cost to vary with the neighborhood to which the child moves, or to allow the disruption cost to vary with the age of the child at the time of the move. Neither of these extensions would affect our estimates. The key requirement for our approach to identifying $\{\mu_c\}$ is that the disruption costs do not vary in an age-dependent manner across neighborhoods.

14. We do not estimate such models here primarily because of a lack of adequate statistical power. As we will see, obtaining precise estimates of 6,000 treatment effects (one per county, interacted with parent income) with our sample of approximately 3 million movers is itself challenging. Estimating higher-dimensional models may be feasible as additional years of tax data become available in the United States or using longer administrative panels in other countries.

the outcomes of children who move at different ages; for instance, if those who move at younger ages have better outcomes, we learn that $\mu_d > \mu_o$.

Under the model in [equation \(1\)](#), we can combine information from all such pairwise comparisons to estimate each place's causal effect using the following fixed effects specification:

$$(2) \quad y_i = \alpha_{od} + \vec{e}_i \cdot \vec{\mu} + \epsilon_i,$$

where α_{od} denotes an origin-by-destination fixed effect and $\vec{e}_i = \{e_{ic}\}$ is a vector whose entries denote the number of years of exposure that child i has to place c before age A . In a sample of children who move exactly once before age A , e_{ic} is given by

$$e_{ic} = \begin{cases} A - m_i & \text{if } d(i) = c \\ m_i & \text{if } o(i) = c \\ 0 & \text{otherwise,} \end{cases}$$

where m_i denotes the age of the child at the time of the move.

[Equation \(2\)](#) is a reduced form of the model in [equation \(1\)](#) for one-time movers, where $\alpha_{od} = \bar{\theta}_{od} - \kappa$ captures the sum of the disruption effect and the mean value of other inputs θ_i for children who move from o to d , while $\epsilon_i = \theta_i - \bar{\theta}_{od}$ captures idiosyncratic variation in other inputs. By including origin-by-destination (α_{od}) fixed effects in [equation \(2\)](#), we identify $\vec{\mu}$ purely from variation in the timing of moves, rather than comparing outcomes across families that moved to or from different areas.¹⁵ The identification assumption required to obtain consistent estimates of $\vec{\mu}$ when estimating [equation \(2\)](#) using OLS is the following standard orthogonality condition.

ASSUMPTION 1. Conditional on α_{od} , exposure time to each place, \vec{e}_i , is orthogonal to other determinants of children's outcomes:

$$(3) \quad \text{Cov}(e_{ic}, \epsilon_i) = 0 \forall c.$$

15. The fixed effects $\vec{\mu}$ are identified up to the normalization that the average place effect is 0, $E[\mu_c] = 0$, because the matrix $\mathbf{E} = \{\vec{e}_i\}_i$ does not have full rank. Intuitively, using movers to identify place effects allows us to identify the effect of each place relative to the national average.

Assumption 1 requires that children with different exposure times to a set of places do not systematically differ in other inputs, $\epsilon_i = \theta_i - \bar{\theta}_{od}$, conditional on origin-by-destination fixed effects. This assumption is a stronger version of Assumption 1 in Chetty and Hendren (2018a), which required that exposure to better places—as measured by the outcomes of permanent residents—is not correlated with ϵ_i on average. Assumption 1 extends that assumption to require that the amount of exposure to every place satisfies such an orthogonality condition. This stronger assumption allows us to go beyond establishing that neighborhoods have causal effects on average and characterize precisely which areas produce the best outcomes. We provide evidence supporting Assumption 1 in Section IV.B after presenting our baseline results.

III.C. Empirical Implementation

In our empirical analysis, we generalize equation (2) to account for two features of the data that we omitted from the stylized model in Section III.A for simplicity.

First, we allow for the possibility that places may have different effects across parent income levels, as suggested by the maps of permanent residents' outcomes in Chetty and Hendren (2018a, Figure II). To do so, we first measure the percentile rank of the parents of child i , $p(i)$, based on their positions in the national distribution of parental household income for child i 's birth cohort. Chetty et al. (2014) show that within each area, children's expected income ranks are well approximated by a linear function of their parents' income rank $p(i)$. We therefore generalize equation (2) to allow place effects μ_c to vary linearly with parent income rank, $p(i)$. We denote the causal effect of place c at parent rank p by

$$(4) \quad \mu_{pc} = \mu_c^0 + \mu_c^1 p,$$

where μ_c^0 , the intercept, represents the causal effect of the place for children in the lowest-income families and μ_c^1 , the slope, captures how the causal effect varies with parent rank. For symmetry, we also allow the origin-by-destination fixed effect α_{od} in equation (2) to vary linearly with parent rank p to capture potential heterogeneity in selection effects by parent income.

Second, because we measure children's incomes at a fixed age, we measure their incomes in different calendar years. In particular, incomes at age 26 are measured between 2006 and

2012 for children in our primary sample (the 1980–1986 birth cohorts), a period with rapidly changing labor market conditions. We account for these fluctuations, allowing for differential shocks across areas and income group, by including a control function

$$(5) \quad g_{od}(p, s) = \psi_{od}^0 s + \psi_{od}^1 s^2 + \psi_{od}^2 s p + \psi_{od}^3 s^2 p,$$

when estimating [equation \(2\)](#), where s denotes the child's birth cohort (or, equivalently, the year in which the child's income is measured). We show in [Online Appendix A](#) that alternative parameterizations of the $g(p, s)$ control function yield very similar results.

Incorporating these two extensions of [equation \(2\)](#), our baseline estimating equation is

$$(6) \quad y_i = \alpha_{od} + \alpha_{od}^P p + \vec{e}_i \cdot \vec{\mu}_p + g_{od}(p_i, s_i) + \varepsilon_i,$$

where α_{od} is an origin-by-destination fixed effect, $\alpha_{od}^P p$ is an origin-by-destination fixed effect interacted with parent rank, and $\vec{\mu}_p = \{\mu_{pc}\}$ denotes the vector of causal place effects parameterized as in [equation \(4\)](#).

1. Estimation: Hierarchical Structure. Our goal is to use [equation \(6\)](#) to identify the causal effects of places at two geographic levels: CZs—aggregations of counties that represent local labor markets—and counties. Directly estimating the 741 CZ-level and 3,138 county-level fixed effects along with the incidental parameters in [equation \(6\)](#) is not feasible because of computational constraints. We therefore estimate $\vec{\mu}_p$ using a hierarchical structure, separately estimating the causal effects of CZs and counties within CZs.

We begin by estimating CZ effects using our sample of cross-CZ movers. For computational tractability, we use a two-step estimator, described in detail in [Online Appendix A](#). In the first step, we estimate [equation \(6\)](#) separately for each origin-destination (o, d) pair, which yields an estimate of the exposure effect for each origin relative to each destination, $\mu_{pod} = \mu_{pd} - \mu_{po}$, for each level of parental income, p .¹⁶ We then consider a fixed parental income level (e.g., $p = 25$) and regress the pairwise effects $\{\mu_{pod}\}$ on a

16. We restrict attention to the 11,216 $o - d$ pairs that have at least 25 observations, which account for 75% of moves across CZs in our sample. We have made the pairwise $o - d$ estimates publicly available in [Online Appendix Data Table 5](#) to facilitate future research using alternative models of neighborhood effects (e.g., models that permit heterogeneous match effects).

design matrix that consists of positive and negative indicators for each CZ to obtain an estimate of each CZ's fixed effect at percentile p (see [Online Appendix A](#) for the specification of this matrix). We weight each observation by the precision of the pairwise estimate in this regression. Finally, we normalize these estimates to have a population-weighted mean of 0 across CZs (using populations from the 2000 census), so that the fixed effects can be interpreted as the causal effect of the CZ relative to the average CZ in the country. In our baseline specifications, we estimate the standard errors of $\hat{\mu}_{pc}$ using bootstrap resampling of the microdata (see [Online Appendix B](#) for details). For alternative specifications, we report analytical standard errors obtained from the regression in the second step of the estimation procedure, which are very similar to the bootstrapped estimates in the baseline case, to simplify computation.

Identifying CZ effects using this two-step approach requires that moves into and out of each CZ are balanced across counties. Intuitively, if all movers into a CZ moved to one particular county, one would effectively identify the sum of the CZ and that county's effect rather than the mean effect across all counties within the CZ. In practice, moves are generally balanced relative to county populations: the correlation between gross flows into and out of counties within each CZ and county populations recorded in the 2000 census is 0.90.

Having identified the CZ-level effects, we estimate county effects within each CZ purely from moves across counties within CZs. Because there are only four counties on average within each CZ, we can directly estimate [equation \(6\)](#) separately for each CZ using movers across counties within that CZ.¹⁷ We normalize these estimates to have a population-weighted mean of 0 across counties within each CZ, so that the estimates can be interpreted as the causal effect of each county relative to the CZ mean. We obtain standard errors for these county-within-CZ fixed effects directly from the OLS regression in [equation \(6\)](#).

Finally, we construct county-level estimates by adding the CZ-level fixed effect estimate to the county-within-CZ fixed effect estimate. We use similar methods to estimate fixed effects for

17. Because the counties in each CZ are all in the same labor market, we do not permit the $\{\psi\}$ coefficients in the $g(p, s)$ cohort control function to vary with origin and destination when estimating the county-within-CZ models. This simplification reduces the number of parameters to be estimated significantly without affecting the results.

subgroups (e.g., for boys versus girls) and other outcomes (e.g., rates of marriage).

IV. FIXED EFFECT ESTIMATES

This section presents fixed effect estimates of CZ- and county-level causal effects. We first present baseline estimates using equation (6) and then discuss how we evaluate the key identification assumption (Assumption 1) underlying our design.

IV.A. Baseline Estimates

As in our first article, we measure children's incomes based on their percentile ranks in the income distribution. We define child i 's percentile rank y_i based on his or her position in the national distribution of incomes relative to all others in his or her birth cohort. We focus on children growing up with parents at either the 25th or 75th percentiles of the parent income distribution ($p = 25$ or $p = 75$). Given the linearity of the relationship between children's expected ranks and parent ranks, these estimates correspond to the mean rank outcomes of children in below-median ($p < 50$) and above-median ($p \geq 50$) income families, and fully summarize the conditional distribution of children's outcomes given parents' incomes in each area.

We report place effects on both children's own (individual) incomes and their household incomes (including spousal income). However, we focus primarily on the household income results because they provide a better measure of how areas affect children's economic opportunities, independent of variation in labor force participation rates. Prior work on intergenerational mobility has focused on the individual earnings of men as a way to sidestep the challenges in measurement that arise from differences in female labor force participation rates. In our data, we find that men's individual income ranks at age 30 are very highly correlated with both male and female household income ranks for permanent residents across CZs (with population-weighted correlations above 0.9 for children with parents at $p = 25$), but are not highly correlated with women's individual earnings ranks (correlation = 0.41). These correlations suggest that women's household incomes provide a better representation of the earnings levels that would

prevail if labor force participation rates were held constant than women's individual incomes.¹⁸

Figure I, Panel A plots the CZ fixed effect estimates on children's household incomes at age 26 given parents at $p = 25$, $\hat{\mu}_{25,c}$, versus the outcomes of children of permanent residents in each CZ, $\bar{y}_{25,c}$. CZs with more than 2.5 million residents are labeled, with the dashed vertical bars showing 95% confidence intervals for these estimates. We first discuss the variation in the fixed effects plotted on the y -axis and then turn to the relationship between this variation and the permanent residents' outcomes shown on the x -axis. As an example, we estimate that every year of exposure to Los Angeles decreases the expected income rank of a child growing up in a low-income family ($p = 25$) by 0.17 percentiles (std. err. = 0.04) relative to the average CZ in the United States. In contrast, every year of exposure to Cleveland, OH, increases a child's income rank by 0.12 percentiles (std. err. = 0.10) relative to the average CZ.

To interpret the magnitude of these effects, it is helpful to translate these percentile changes into dollar values.¹⁹ To do so, we regress the mean income levels of children of permanent residents in each CZ, \bar{y}_{pc}^* , on their mean income ranks, \bar{y}_{pc} , separately at each parent income percentile p , weighting by population. This regression yields a coefficient of \$818 for $p = 25$, implying that a

18. One may be concerned that measuring household income at age 26, as we do in our baseline analysis to maximize precision, could yield biased estimates of areas' impacts on individuals' permanent household income because of differences in ages at first marriage across areas. We approach this issue empirically by evaluating what the best predictor of place effects on household income at age 32 is given information available by age 26. We regress permanent residents' mean household income ranks at age 32 given parents at $p = 25$ on three predictors: permanent residents' mean household income rank at age 26, mean individual income rank at age 26, and college attendance rates from ages 18–23. All three measures have predictive power, but the coefficient on household rank is significantly larger than the other variables. Moreover, the predicted values from this regression have a correlation of 0.97 with mean household income ranks at age 26 across CZs. Hence, mean household income ranks at age 26 provide an accurate representation of place effects on household incomes at older ages.

19. A more direct method of estimating place effects on the level of income (in dollars) would be to estimate $\hat{\mu}_{pc}$ using income levels instead of ranks as the outcome. The estimates of $\sigma_{\mu_{pc}}$ obtained using income levels as the outcome are highly correlated with our baseline estimates using ranks (**Online Appendix** Table I, row 4), but have many more outliers because of outliers in income levels at the individual level. This is why we use rank outcomes, which yield more precise and stable results across specifications, for our primary analysis.

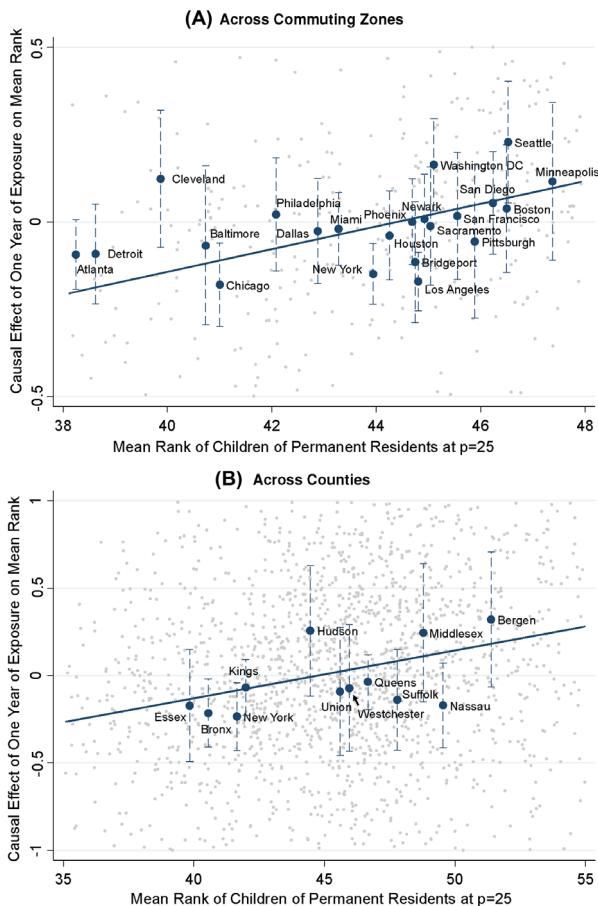


FIGURE I

Causal Effect Estimates versus Permanent Residents' Outcomes for Low-Income Families

Panel A plots causal effects of childhood exposure to each CZ, estimated using movers' outcomes, versus permanent residents' outcomes. The vertical axis shows estimates of $\hat{\mu}_{25,c}$, the causal effect of an additional year of childhood exposure to a CZ (relative to the average CZ) on the mean percentile rank at age 26 for children in families at income percentile $p = 25$. The horizontal axis plots $\bar{y}_{25,c}$, the mean ranks of children of permanent residents (nonmovers) at $p = 25$. CZs with populations above 2.5 million (based on the 2000 census) are labeled. Dashed vertical bars show 95% confidence intervals for $\hat{\mu}_{25,c}$. The solid line shows the conditional expectation of $\hat{\mu}_{25,c}$ given \bar{y}_{pc} , estimated using an OLS regression pooling all CZs and weighting by the precision of $\hat{\mu}_{25,c}$. Panel B replicates Panel A at the county level. Counties within the New York and Newark CZs that have populations above 500,000 are labeled. The sample in both figures consists of all children in the 1980–1986 birth cohorts who are U.S. citizens; see Section II for further details.

1 percentile increase in income translates to an additional \$818 at age 26 on average. The mean income of children with below-median income parents is \$26,091; therefore, a 1 percentile increase corresponds to approximately a $\frac{818}{26,091} = 3.14\%$ increase in income. The point estimates in [Figure I](#), Panel A therefore imply that one year of exposure to Cleveland instead of the average CZ would raise a child's income by $0.12 \times 3.14\% = 0.38\%$, whereas an additional year of exposure to Los Angeles instead of the average CZ would reduce a child's income by $0.17 \times 3.14\% = 0.53\%$.

If we assume that these exposure effects are constant throughout childhood in each area, these estimates would imply that children who move to Cleveland at birth and stay there for 20 years would earn $20 \times 0.38\% = 7.5\%$ more than if they had grown up in the average CZ. Conversely, spending 20 years of childhood in Los Angeles instead of the average CZ would reduce a child's income by about 10.7% relative to the average CZ.²⁰

[Figure I](#), Panel B presents analogous estimates for each county in the US, highlighting estimates for counties in the New York and Newark CZs that have populations above 500,000. For example, $\hat{\mu}_{25,c} = -0.23$ percentiles (std. err. = 0.10) in the Bronx, implying that growing up in the Bronx causes an income loss of approximately 0.72% per year of childhood exposure relative to the average county in the United States. In contrast, $\hat{\mu}_{25,c} = 0.25$ percentiles (std. err. = 0.19) in Hudson County, NJ, equivalent to an income gain of 0.79% per year of exposure.

The dispersion in the estimates on the y-axes on [Figure I](#), Panels A and B suggests that there may be substantial variation in the causal effects of places μ_{pc} , although part of the observed dispersion is driven simply by sampling error in our estimates $\hat{\mu}_{pc}$. We quantify the magnitude of the variation in μ_{pc} after accounting for the variation in $\hat{\mu}_{pc}$ that is due to sampling error in [Section V](#).

1. Comparison to Permanent Residents' Outcomes. It is instructive to compare the fixed effect estimates of μ_{pc} based

20. [Chetty and Hendren \(2018a, Figure IV\)](#) show that the effect of each additional year of exposure to a better area is roughly constant over the range of ages they are able to study (ages 9–23). To predict the causal impacts of growing up in an area from birth, we must assume that the exposure effects $\hat{\mu}_{25,c}$ remain constant even prior to age nine, a strong assumption that remains to be tested in future work. The estimates of the impact of growing up in a given area from birth reported here should therefore be interpreted as approximate values.

on movers to the outcomes of children of permanent residents (nonmovers) in each area. Under the model in [equation \(1\)](#), permanent residents' outcomes combine the causal effect of growing up in a given area with selection effects reflecting differences in family inputs across areas:

$$(7) \quad \bar{y}_{pc} = A\mu_{pc} + \bar{\theta}_{pc},$$

where $A\mu_{pc}$ is the cumulative effect of childhood exposure to place c and

$$\bar{\theta}_{pc} = E[\theta_i | p(i) = p, c(i, t) = c \forall t]$$

is the average of the other inputs θ_i obtained by children of permanent residents in location c .²¹

[Figure I](#) shows that the causal effect estimates $\hat{\mu}_{25,c}$ based on the sample of movers are highly correlated with permanent residents' outcomes $\bar{y}_{25,c}$, consistent with the findings in [Chetty and Hendren \(2018a\)](#). At the CZ level, regressing $\hat{\mu}_{25,c}$ on $\bar{y}_{25,c}$ yields a slope of $\gamma_{25} = \frac{dE[\mu_{25,c}|\bar{y}_{25,c}]}{d\bar{y}_{25,c}} = 0.032$ (std. err. 0.003), illustrated by the best-fit line in [Figure I](#), Panel A. That is, a year of exposure to a CZ where permanent residents' outcomes are 1 percentile higher increases a given child's outcomes by 0.032 percentiles.²² We find similar estimates at the county level and for children in high-income families ($p = 75$) ([Online Appendix](#) Figure I).

Although \bar{y}_{pc} is highly predictive of μ_{pc} on average, there are many differences between the causal effect estimates and

21. If neighborhood effects vary across areas within CZs or counties, as is likely to be the case, then differences in the geographical distribution of permanent residents relative to movers within an area c would also be incorporated into the selection term $\bar{\theta}_{pc}$.

22. This estimate of $\gamma_{25} = 0.032$ differs from the estimate of $\gamma \simeq 0.04$ reported in our first article because we impose less stringent population restrictions (requiring populations above 25,000 instead of 250,000), do not impose restrictions on the distance of moves, and focus on families at $p = 25$ rather than aggregating across all percentiles. In addition, we estimated γ in our first article by directly projecting movers' outcomes onto the outcomes of permanent residents. The advantage of that approach relative to the analysis in [Figure I](#) is that it directly uses permanent residents' outcomes as "goal posts" for movers' expected outcomes in each place. This allows us to implement placebo tests exploiting heterogeneity across subgroups and increases statistical power, allowing us to estimate exposure effects nonparametrically by age and to estimate specifications that control for family fixed effects or are identified from displacement shocks.

permanent residents' outcomes in certain cases. For example, children of permanent residents in Cleveland have worse outcomes than those in Los Angeles. The causal effect estimates based on movers, however, imply that Cleveland produces better outcomes for a given child than Los Angeles. We present a more systematic comparison between causal effects and permanent residents' outcomes in [Section VI.B](#).

2. Alternative Specifications. In [Online Appendix A](#), we assess the sensitivity of our estimates of $\hat{\mu}_{25,c}$ to several alternative specifications: (i) modeling heterogeneity in the impact of places across parental income levels using a quadratic function of parental income rank p instead of the linear specification used in [equation \(6\)](#); (ii) controlling for fluctuations across cohorts using alternative parameterizations relative to the specification in [equation \(5\)](#); (iii) measuring children's outcomes in levels instead of percentile ranks; and (iv) using income measures that adjust for local costs of living. All of these specifications yield fixed effect estimates that are very highly correlated with our baseline estimates ([Online Appendix Table I](#)).

IV.B. Validation of Research Design

The fixed effect estimates μ_{pc} obtained from [equation \(6\)](#) can only be interpreted as causal effects of areas under the identification assumption in [equation \(3\)](#), which requires that children's exposure to each area e_{ic} is orthogonal to other inputs θ_i , conditional on origin-by-destination fixed effects and parental income levels. In this section, we evaluate whether [equation \(3\)](#) holds using tests that build on the methods in [Section V](#) of [Chetty and Hendren \(2018a\)](#). We briefly summarize the results of these tests here; see [Online Appendix C](#) for details.

We organize our evaluation of [equation \(3\)](#) by partitioning θ_i 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.

1. Fixed Factors. Fixed factors $\bar{\theta}_i$ can create selection bias in estimates of μ_{pc} if $\bar{\theta}_i$ is correlated with the age at which child i moves to a given area c . In [Chetty and Hendren \(2018a\)](#), we showed that children who move to areas with better permanent resident outcomes \bar{y}_{pc} at younger ages do not have significantly different levels of $\bar{\theta}_i$ using specifications with family fixed effects.

Since \bar{y}_{pc} is very highly correlated with the causal effects of place μ_{pc} (Figure I), this finding implies that any selection biases in our estimates $\hat{\mu}_{pc}$ must arise from heterogeneity in $\bar{\theta}_i$ that is unrelated to a place's causal effect μ_{pc} . Intuitively, the concern that remains is that the deviations of $\hat{\mu}_{pc}$ from the permanent resident predictions $\gamma_p(\bar{y}_{pc} - \bar{y}_p)$ in Figure I might reflect selection bias rather than causal effects.

We test for such selection biases using two placebo tests. First, we examine the incomes (at age 26) of children who are older than 23 when their parents move. These children provide a natural placebo group because they are less likely to move with their parents and because our first article shows that neighborhoods no longer have exposure effects after age 23. Second, we implement a placebo test using teenage labor force participation (LFP) at age 16. Teenage LFP rates provide an informative pretreatment indicator because they are positively correlated with differences in children's incomes in adulthood across CZs. Both of these placebo estimates of place effects are uncorrelated with our baseline estimates of μ_{pc} (Online Appendix Table I, rows 7–8), indicating that families who move to a given area at different times do not differ systematically in their children's potential outcomes.

2. Time-Varying Factors. The second potential source of bias in our estimates of μ_{pc} are time-varying factors $\tilde{\theta}_i$ that are correlated with families' decisions to move, such as parents' incomes. In Chetty and Hendren (2018a), we showed that the changes in children's incomes when families move to areas with better permanent resident outcomes \bar{y}_{pc} are not driven by time-varying confounds using a set of placebo tests exploiting heterogeneity across subgroups. As above, this result implies that any remaining bias must arise from time-varying factors that are uncorrelated with places' causal effects μ_{pc} .

To assess the potential bias from such factors, we control for changes in parental income and marital status when estimating equation (6). The estimates of $\hat{\mu}_{25,c}$ obtained with these controls are nearly identical to the baseline estimates, with correlations above 0.97 (Online Appendix Figure II). Hence, any violation of our key identification assumption would have to arise from time-varying unobservables that are uncorrelated with both permanent residents' outcomes \bar{y}_{pc} (in the origin and destination) and with changes in income and marital status. We believe such violations of the identification condition are unlikely to be prevalent and

therefore view our baseline fixed effects $\{\hat{\mu}_{pc}\}$ as providing unbiased estimates of place effects.

In the next four sections, we use the fixed effect estimates to (i) quantify the magnitude of place effects, (ii) construct mean-squared-error-minimizing forecasts of the causal effect of growing up in each county, (iii) characterize the properties of areas that produce higher levels of upward mobility, and (iv) identify areas that produce good outcomes with low housing costs. The fixed effect estimates $\{\hat{\mu}_{pc}\}$ are provided in [Online Appendix Data Tables 3 and 4](#), and hence all of the results that follow can be reproduced using publicly available data.

V. MAGNITUDE OF PLACE EFFECTS

How much does the neighborhood in which a child grows up influence his or her outcomes in adulthood? In this section, we estimate the standard deviation of place effects ($\sigma_{\mu_{pc}}$) by decomposing the variation in the fixed effect estimates $\hat{\mu}_{pc}$ into the portion due to signal (differences in latent causal effects) versus noise (sampling error).

V.A. Methods

The raw standard deviation of place effect estimates $\sigma_{\hat{\mu}_{pc}}$ overstates the true (signal) standard deviation of place effects $\sigma_{\mu_{pc}}$ because part of the variation in the estimates $\hat{\mu}_{pc}$ is due to sampling error. To estimate $\sigma_{\mu_{pc}}$, we decompose the place effect estimates $\hat{\mu}_{pc}$ into the (latent) place effect μ_{pc} and sampling error η_{pc} :

$$(8) \quad \hat{\mu}_{pc} = \mu_{pc} + \eta_{pc},$$

where η_{pc} is orthogonal to μ_{pc} ($E[\eta_{pc}|\mu_{pc}] = 0$). This decomposition implies that we can estimate $\sigma_{\mu_{pc}}^2$ by subtracting the variance induced by sampling error, $\sigma_{\eta_{pc}}^2$, from the variance in the observed estimates, $\sigma_{\hat{\mu}_{pc}}^2$:

$$(9) \quad \hat{\sigma}_{\mu_{pc}}^2 = \sigma_{\hat{\mu}_{pc}}^2 - \sigma_{\eta_{pc}}^2.$$

We estimate the noise variance $\sigma_{\eta_{pc}}^2$ as the average squared standard error,

$$\sigma_{\eta_{pc}}^2 = E[s_{pc}^2],$$

where s_{pc} denotes the standard error of $\hat{\mu}_{pc}$, estimated using the methods discussed in [Section III.C](#), and the expectation is taken across areas. We compute the standard error of the signal standard deviation estimate $\hat{\sigma}_{\mu_{pc}}$ using an asymptotic approximation described in [Online Appendix D](#). We use precision weights $(\frac{1}{s_{pc}^2})$ when estimating all of these parameters to maximize efficiency.²³

V.B. Results

[Table II](#) reports the standard deviation of the raw fixed effects $\sigma_{\hat{\mu}_{pc}}$, the noise component $\sigma_{\eta_{pc}}$, and the latent causal effects $\sigma_{\mu_{pc}}$. We report estimates at the CZ level, county level, and across counties within CZs for children whose parents are at $p = 25$ and $p = 75$. In Panel A, we report estimates of the standard deviation of annual exposure effects on children's income ranks at age 26. Panel B rescales these estimates to present other metrics for the size of place effects.

1. Low-Income Families. We begin by discussing the magnitude of neighborhood effects for children who grow up in low-income families ($p = 25$). At the CZ level, the (precision-weighted) standard deviation of the raw fixed effects at $p = 25$ is $\sigma_{\hat{\mu}_{25,c}} = 0.25$, as reported in the first row in [Table II](#). A large fraction of this variation is due to noise: $\sigma_{\eta_{pc}} = \sqrt{E[s_{pc}^2]} = 0.21$. Subtracting the variance of the sampling error using [equation \(9\)](#) yields a signal standard deviation across CZs of $\sigma_{\mu_{pc}} = 0.13$. That is, living in a one standard deviation better CZ based on children's realized outcomes increases a given low-income child's expected rank by 0.13 per year of exposure.

Across counties at $p = 25$, we estimate $\sigma_{\hat{\mu}_{25,c}} = 0.43$, $\sigma_{\eta_{25,c}} = 0.40$, and $\sigma_{\mu_{pc}} = 0.17$. The county-level estimates exhibit more noise than the CZ-level estimates because sample sizes are smaller at the county level. The standard deviation of causal effects across counties $\sigma_{\mu_{pc}}$ is larger than that across CZs, which is expected because CZs are aggregations of counties. The estimates imply that the standard deviation of counties' causal effects

23. Precision weighting is efficient if the signal variance $\sigma_{\mu_{pc}}^2$ is constant (homoskedastic), but yields estimates that may vary with the sample in the presence of heteroskedasticity. The estimates reported below are very similar if we instead use population weights (which are sample-invariant), provided that we exclude estimates of $\hat{\mu}_{pc}$ with exceptionally high standard errors (e.g., above the 99th percentile of the distribution of s_{pc}).

TABLE II
MAGNITUDES OF PLACE EFFECTS

	Counties within CZs					
	Commuting zones			Counties		
	Below median $p = 25$	Above median $p = 75$	Below median $p = 25$	Above median $p = 75$	Below median $p = 25$	Above median $p = 75$
Panel A: Annual exposure effect estimates						
Std. dev. of raw exposure effects	0.248	0.243	0.484	0.435	0.357	0.361
Std. dev. of noise component	0.210	0.218	0.402	0.407	0.343	0.344
Signal std. dev. of exposure effects	0.132	0.107	0.165	0.155	0.099	0.112
Standard error of signal std. dev.	0.010	0.013	0.014	0.016	0.008	0.008
Panel B: Signal std. dev. of causal effects from birth: 20 years of exposure						
Percentile rank	2.65	2.14	3.31	3.09	1.98	2.23
Income in dollars	2,165	1,796	2,706	2,596	1,623	1,875
Percentage change in income	8.3	4.4	10.4	6.4	6.2	4.6
Panel C: Correlations with other outcomes						
Signal correlation of causal effects and permanent resident outcomes	0.799	0.909	0.698	0.492	0.577	0.044
Correlation between causal effects at $p = 25$ and $p = 75$	0.724	0.727	0.287	0.080		

Notes. This table quantifies the magnitude of place effects on children's household (i.e., family) incomes. The first pair of columns reports estimates of the variation in place effects across CZs, estimated using movers across CZs. The second pair reports estimates across counties, summing the fixed effect estimates obtained from the county-within-CZ movers sample and the CZ estimates; and the third pair reports the implied county-within-CZ estimates based on the difference between the estimates in the first two pairs of columns. In each case, we report estimates for families with below-median ($p = 25$) and above-median ($p = 75$) income. In Panel A, the first row presents the raw standard deviation of the fixed effect estimates across areas, weighting by precision ($\frac{1}{SE^2}$), where SE is the estimated standard error of the fixed effect. The second row presents the standard deviation of the sampling noise component (again weighted by precision, $\frac{1}{SE^2}$). The third row presents the estimated signal standard deviation, computed as the square root of the difference between the raw variance and the noise variance. The fourth row shows the standard error of the signal standard deviation estimate, computed using the asymptotic approximation in [Online Appendix B](#). Panel B quantifies the size of the effects of growing up in a one standard deviation better area from birth (20 years of childhood exposure). The first row shows the impact on a child's percentile rank, multiplying the signal standard deviation estimate in Panel A by 20. The next two rows rescale these percentile rank impacts into dollar increases and percentage increases in children's incomes using the methods described in [Section V.B](#). In Panel C, the first row reports the correlation between the causal effect estimates and the permanent resident outcomes (weighted by the precision of the fixed effect estimates), multiplied by the ratio of the raw standard deviation to the signal standard deviation to adjust for noise. The second row presents the correlation between the 25th and 75th percentile fixed effect estimates, weighting by the inverse of the sum of the variances of those estimates. These estimates are constructed by splitting the sample at the median family income to obtain estimates of the $p = 25$ and $p = 75$ fixed effects from independent samples, adjusting for noise by rescaling the raw correlation using the ratio of the raw standard deviations to the signal standard deviations.

within a given CZ is 0.10 on average, showing that there is nearly as much variation in children's outcomes across counties within CZs as there is across CZs.

To interpret the magnitude of these standard deviations, in Panel B, we rescale the annual exposure effects in three steps analogous to those used in [Section IV.A](#). First, we multiply the annual exposure effect estimates by 20 to obtain a rough estimate of the causal effect of growing up in a given area from birth.²⁴ Second, we translate the percentile changes into dollar values by multiplying the estimates by \$818, given our estimate above that each additional income rank translates to an additional \$818 of income at age 26 on average for children with parents at $p = 25$. Third, we translate the estimates to percentage impacts on income by dividing by the mean level of income at age 26 for children with below-median income parents (\$26,091).

Using this rescaling, our estimates imply that for a child with parents at $p = 25$, growing up in a one standard deviation better county from birth would increase his or her income at age 26 by 3.3 percentiles. This translates to an increase in income of \$2,700, which is a 10.4% increase in income (about 0.5% a year of childhood exposure). For comparison, a one standard deviation increase in parental income ranks is associated with an 7.1 percentile increase in children's ranks at age 26 in our sample. The causal effects of neighborhoods are thus nearly half as large as the association between parent and child income. As another benchmark, [Chetty, Friedman, and Rockoff \(2014b\)](#) estimate that being assigned to a one standard deviation better teacher (based on the teacher's test score value added) for a single school year raises income by 1.3%. Hence, growing up in a one standard deviation better county is roughly equivalent to having eight consecutive years of a one standard deviation higher value-added teacher. These comparisons show that neighborhood effects are an important determinant of children's outcomes, with an order of magnitude comparable to other potential interventions, such as changes in educational quality or family resources. However, like these other factors, the area in which a child grows up explains only a small portion of the total variance in children's outcomes: the standard deviation of county effects (3.3 percentiles) is only

24. As noted already, [Chetty and Hendren \(2018a\)](#) show that exposure effects are approximately constant between ages 9 and 23, suggesting that the appropriate scaling factor to estimate impacts from birth is between 14 and 23.

11.4% of the unconditional standard deviation of children's ranks (which is 28.9 percentiles).²⁵

The standard deviation of the causal effects of places $\mu_{25,c}$ is smaller than the standard deviation of permanent residents' outcomes \bar{y}_{pc} , implying that the variation in permanent residents' outcomes across areas is partly due to selection. At the CZ level, $\sigma_{\bar{y}_{25,c}} = 3.3$ percentiles, while the causal effect of growing up in a one standard deviation better CZ from birth is 2.7 percentiles. The corresponding values at the county level are 4.2 and 3.3 percentiles. The correlation between $\bar{y}_{25,c}$ and $\mu_{25,c}$ is 0.80 across CZs and 0.58 across counties within CZs.²⁶ Hence, permanent residents' outcomes are quite informative about places' causal effects, but there are significant differences between $\mu_{25,c}$ and $\bar{y}_{25,c}$, especially at the county level. These results demonstrate that the differences between $\hat{\mu}_{25,c}$ and $\bar{y}_{25,c}$ in [Figure I](#) reflect not just sampling error, but differences between places' causal effects $\mu_{25,c}$ and permanent residents' outcomes that are driven by selection.

2. High-Income Families. Neighborhoods have similar effects in percentile rank and dollar terms for children of high-income ($p = 75$) parents, but matter less in percentage terms because children in high-income families have higher mean incomes. For children with parents at the 75th percentile of the income distribution, the signal standard deviation of place effects at the county level is 3.1 percentiles, which translates to a dollar change of \$2,600, or a 6.4% change in income. The place where a child grows up may matter less for children with higher-income parents because high-income families are able to insulate themselves from local conditions more effectively (e.g., by switching to private schools if public schools are weak).

25. In this sense, our estimates are consistent with the upper bound on neighborhood effects constructed by [Solon, Page, and Duncan \(2000\)](#) based on the correlation between neighbors' outcomes within an area. Solon, Page, and Duncan (2000, 390) estimate that a one standard deviation increase in neighborhood (defined as a PSID sampling cluster) quality is associated with at most a 0.32 standard deviation increase in years of education. Our estimates imply that a one standard deviation increase in county quality causes a $\frac{3.3}{28.9} = 0.11$ standard deviation increase in children's ranks.

26. We estimate these correlations as $\text{Corr}(\mu_{25,c}, \bar{y}_{25,c}) = \text{Corr}(\hat{\mu}_{25,c}, \bar{y}_{25,c}) \frac{SD(\hat{\mu}_{25,c})}{SD(\mu_{25,c})}$, where the ratio of standard deviations is obtained from [Table II](#) and adjusts for the attenuation in $\text{Corr}(\hat{\mu}_{25,c}, \bar{y}_{25,c})$ due to sampling error in $\hat{\mu}_{25,c}$.

For high-income families, the causal effects $\mu_{75,c}$ are highly correlated with permanent residents' outcomes $\bar{y}_{75,c}$ across CZs (correlation = 0.91). However, across counties within CZs, the correlation between $\mu_{75,c}$ and $\bar{y}_{75,c}$ falls to 0.04, suggesting that the observational variation in children's outcomes across areas within a given CZ is driven primarily by sorting rather than causal effects for affluent families.

Are the places that generate good outcomes for the poor the same as those that generate good outcomes for the rich? Across CZs, the signal correlation between $\mu_{25,c}$ and $\mu_{75,c}$ is 0.72.²⁷ Across counties within CZs, the correlation is 0.08. In short, there is no evidence that places that generate better outcomes for the poor generate worse outcomes for the rich; if anything, at broad geographies, places that are better for the poor are better for the rich, too.²⁸

3. Heterogeneity by Gender. Estimating fixed effects $\hat{\mu}_{pc}$ separately for male and female children, we find that the place where a child grows up matters more for boys than girls, especially in low-income families ([Online Appendix Table II](#)). Growing up in a one standard deviation better county from birth in a family at the 25th percentile increases boys' household income ranks by 5.5 percentiles (16.4%), compared with 3.5 percentiles (11.4%) for girls. The correlation between boys' and girls' place effects is 0.85 across counties, indicating that the places that are good for boys are generally good for girls as well. However, the variance of outcomes across areas is larger for boys, in particular because there are some areas with very negative outcomes for boys in poor families. We discuss these differences by gender in further detail in [Section VI](#).

27. When computing these correlations, we estimate $\mu_{25,c}$ using families with below-median income ($p < 50$) and $\mu_{75,c}$ using families with above-median income ($p > 50$) to obtain estimates from independent samples that are not spuriously correlated due to sampling error. We compute the signal correlation as $\rho = \frac{\text{Cov}(\mu_{25,c}, \mu_{75,c})}{\sigma_{\mu_{25,c}} \sigma_{\mu_{75,c}}} = \frac{\text{Cov}(\hat{\mu}_{25,c}, \hat{\mu}_{75,c})}{\sigma_{\hat{\mu}_{25,c}} \sigma_{\hat{\mu}_{75,c}}}$. We use precision weights ($\frac{1}{s_{\hat{\mu}_{pc}}^2}$) to estimate the signal standard deviations $\sigma_{\hat{\mu}_{pc}}$ and weight by the inverse of the sum of the standard errors squared, $\frac{1}{s_{\hat{\mu}_{25,c}}^2 + s_{\hat{\mu}_{75,c}}^2}$ when estimating $\text{Cov}(\hat{\mu}_{25,c}, \hat{\mu}_{75,c})$.

28. This cross-sectional correlation does not imply that policies that improve the outcomes of the poor will not affect the rich. Moreover, these correlations only show that better outcomes for the poor in a given CZ or county c are not associated with worse outcomes for the rich within the same CZ or county; they do not shed light on potential spillover effects on the outcomes of the rich in other areas.

VI. FORECASTS OF PLACE EFFECTS

Given that neighborhoods have substantial causal effects on children's outcomes, where should a family who wants to maximize their children's incomes live? In this section, we address this question by constructing forecasts of place effects that minimize the MSE of the true impact of growing up in a given area relative to the predicted impact.

VI.A. Methods

The fixed effect estimates based on movers $\hat{\mu}_{pc}$ provide unbiased but imprecise estimates of place effects, as illustrated by the wide confidence intervals for some of the estimates in Figure I. To obtain more precise forecasts of places' causal effects, we combine $\hat{\mu}_{pc}$ with information on permanent residents' outcomes \bar{y}_{pc} . Permanent residents' outcomes are estimated with essentially no sampling error, but are biased predictors of μ_{pc} because they combine causal effects with sorting. By shrinking our estimates of $\hat{\mu}_{pc}$ toward predictions based on \bar{y}_{pc} , we substantially reduce prediction errors, decreasing the estimator's variance at the expense of introducing some bias.²⁹

Formally, we construct forecasts μ_{pc}^f of each area's true causal effect μ_{pc} at a given level of parental income p that minimize the mean squared prediction error $\sum_c (\mu_{pc} - \mu_{pc}^f)^2$.³⁰ For simplicity, we restrict attention to linear predictors:

$$(10) \quad \mu_{pc}^f = \alpha + \rho_{1p}(s_{pc})\hat{\mu}_{pc} + \rho_{2p}(s_{pc})\bar{y}_{pc},$$

allowing the coefficients $\rho_{1p}(s_{pc})$ and $\rho_{2p}(s_{pc})$ to vary with the degree of sampling error s_{pc} in $\hat{\mu}_{pc}$.³¹

29. Including other predictors, such as racial demographics, poverty rates, or other observable neighborhood characteristics, in addition to permanent residents' outcomes yields very similar forecasts and does not reduce the MSE of the forecasts appreciably ([Online Appendix D](#) and [Online Appendix](#) Figure III).

30. We use a quadratic loss function as an analytically tractable specification that penalizes large errors more heavily, based on the reasoning that large errors in predicting place's effects are likely to have much larger utility costs to families seeking to move than small errors. The fixed effect estimates reported in the online data tables could be used to construct forecasts that achieve other objectives, such as maximizing the likelihood of moving to a neighborhood with a high causal effect.

31. In general, the MSE-minimizing forecast μ_{pc}^f would use the entire set of fixed effect estimates and their variance-covariance matrix $\{\hat{\mu}_{pc}, s_{pc}\}$. Intuitively, one would optimally incorporate information not only for place c but also other

We make two additional simplifying assumptions in constructing forecasts using [equation \(10\)](#). First, we assume that \bar{y}_{pc} is measured without error. Second, we model the true variance of place effects as homoskedastic, that is, we assume $Var(\mu_{pc})$ does not vary with c . The first assumption is purely an expositional simplification; incorporating sampling error in \bar{y}_{pc} yields forecasts that are correlated more than 0.99 with the baseline estimates. The second assumption is more substantive. We have found that permitting some forms of heteroskedasticity in μ_{pc} (e.g., by deciles of population size or sampling variance, s_{pc}) does not affect our estimates appreciably. Nevertheless, in future work, it would be useful to study whether more flexible models can yield further reductions in prediction errors using the estimates of $\hat{\mu}_{pc}$, s_{pc} , and \bar{y}_{pc} that are publicly available in [Online Appendix Data Tables 3 and 4](#).

1. Best Linear Predictors. For a given level of parental income p , the MSE-minimizing coefficients $\rho_{1p}(s_{pc})$ and $\rho_{2p}(s_{pc})$ in [equation \(10\)](#) are equivalent to those that would be obtained from a (hypothetical) OLS regression of μ_{pc} on $\hat{\mu}_{pc}$ and \bar{y}_{pc} , estimated with one observation per area (c) using the subset of areas whose fixed effect estimates have standard errors of a given level s_{pc} . We derive these coefficients using a partial regression approach in [Online Appendix D](#). The resulting MSE-minimizing forecast is given by

$$(11) \quad \mu_{pc}^f = \frac{\chi_p^2}{\chi_p^2 + s_{pc}^2} \hat{\mu}_{pc} + \frac{s_{pc}^2}{\chi_p^2 + s_{pc}^2} \gamma_p (\bar{y}_{pc} - \bar{y}_p),$$

where $\bar{y}_p = E[\bar{y}_{pc}]$ is the mean of \bar{y}_{pc} across areas, $\gamma_p = \frac{Cov(\hat{\mu}_{pc}, \bar{y}_{pc})}{Var(\bar{y}_{pc})}$ is the coefficient obtained from regressing $\hat{\mu}_{pc}$ on \bar{y}_{pc} , $\chi_p^2 = Var(\mu_{pc} - \gamma_p(\bar{y}_{pc} - \bar{y}_p)) = \sigma_{\mu_{pc}}^2(1 - \rho_{\mu_{pc}, \bar{y}_{pc}}^2)$ is the residual variance of place effects (across places c) after subtracting out the component explained by \bar{y}_{pc} , and s_{pc}^2 is the noise variance (squared standard error) of $\hat{\mu}_{pc}$.

[Equation \(11\)](#) shows that the best linear prediction of each county's causal effect is a weighted average of $\hat{\mu}_{pc}$ and $\gamma_p(\bar{y}_{pc} - \bar{y}_p)$,

places c' whose effects may be correlated with the effect of place c to predict μ_{pc} . We focus on the model in [equation \(10\)](#) here for simplicity, but note that more general forecasting models could be estimated in future work using the fixed effect estimates reported in the online data tables.

where the weights depend on the degree of signal (measured by χ_p^2) versus noise (measured by s_{pc}^2) in the fixed effect estimate. The weight on \bar{y}_{pc} falls as the variance in the latent causal effects that cannot be captured by permanent residents' outcomes (χ_p^2) rises.³² We estimate χ_p^2 by subtracting the average sampling variance across places ($E[s_{pc}^2]$) from the variance of the residuals obtained from regressing $\hat{\mu}_{pc}$ on \bar{y}_{pc} :

$$(12) \quad \chi_p^2 = \text{Var}(\hat{\mu}_{pc} - \gamma_p(\bar{y}_{pc} - \bar{y}_p)) - E[s_{pc}^2].$$

We estimate χ_p^2 , γ_p , and \bar{y}_p weighting by the precision of the fixed effect estimates ($\frac{1}{s_{pc}^2}$) to maximize efficiency.

2. Graphical Representation of Optimal Forecasts. Figure II presents graphical intuition for the construction of these optimal forecasts for a subset of the CZs shown in Figure I, Panel A. The circles plot the point estimates of each CZ's causal effect $\hat{\mu}_{25,c}$ at $p = 25$ versus the permanent residents' mean ranks $\bar{y}_{25,c}$. The predicted values from a regression of $\hat{\mu}_{25,c}$ on $\bar{y}_{25,c}$ in the full sample of all CZs, $\gamma_{25}(\bar{y}_{25,c} - \bar{y}_{25})$, are shown by the solid line. The optimal forecasts (shown by the diamonds) are a weighted average of $\hat{\mu}_{25,c}$ and $\gamma_{25}(\bar{y}_{25,c} - \bar{y}_{25})$.

The dashed vertical lines on this figure show one standard error confidence intervals (s_{pc}) for $\hat{\mu}_{25,c}$. In CZs where the standard error s_{pc} is smaller, the optimal forecast is closer to the estimate from movers $\hat{\mu}_{25,c}$ than the permanent resident prediction. For example, in Los Angeles, the optimal forecast $\mu_{25,c}^f = -0.13$ percentiles per year of exposure is obtained by placing 78% of the weight on $\hat{\mu}_{25,c}$ and 22% on $\gamma_{25}(\bar{y}_{25,c} - \bar{y}_{25})$. In smaller CZs, where the fixed effects estimates are less precise, the optimal forecast puts more weight on the predicted outcome based on the permanent residents. For example, in Providence, RI, the optimal forecast puts 70% of the weight on the permanent resident prediction.

3. Magnitude of Prediction Errors. The optimal forecasts differ from the true causal effect because of sampling error in $\hat{\mu}_{pc}$ and bias in permanent resident predictions, $\gamma_p(\bar{y}_{pc} - \bar{y}_p)$. The MSE of the prediction in equation (11) for place c with standard error s_{pc}

32. For CZs with populations below 25,000 and counties with populations below 10,000 in the 2000 census, we do not have an estimate of the causal effect $\hat{\mu}_{pc}$ given our sample restrictions. In these areas, we define the optimal forecast as $\gamma_p(\bar{y}_{pc} - \bar{y}_p)$, the forecast based purely on permanent residents' outcomes.

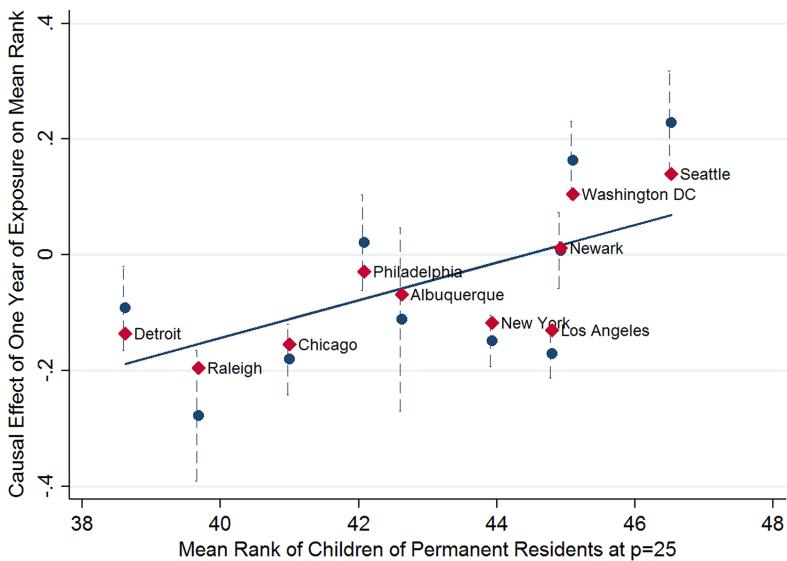


FIGURE II
Construction of Mean-Squared-Error-Minimizing Forecasts

This figure illustrates the construction of the MSE-minimizing forecasts of causal effects, μ_{pc}^f , for a selected set of CZs. The forecasts shown are for children with parents at income percentile $p = 25$. The circles plot the raw fixed effect estimates $\hat{\mu}_{25,c}$ (per year of childhood exposure) versus the mean ranks of children of permanent residents, as in Figure I, Panel A. The dashed vertical lines around these points represent ± 1 standard error of $\hat{\mu}_{25,c}$. The solid regression line shows $E[\mu_{pc} | \bar{y}_{pc}]$, the predicted causal effect of each CZ given the outcomes of its permanent residents, which is estimated from a population-weighted OLS regression of $\hat{\mu}_{25,c}$ on \bar{y}_{pc} using data for all CZs. The diamonds show the MSE-minimizing forecast, which is a weighted average of $\hat{\mu}_{25,c}$ (the circle) and $E[\mu_{pc} | \bar{y}_{pc}]$ (the prediction from the regression line), with greater weight placed on $\hat{\mu}_{25,c}$ when the standard error of $\hat{\mu}_{25,c}$ is smaller.

is

$$e_{pc}^2 = E[\mu_{pc} - \mu_{pc}^f]^2 = \frac{1}{\frac{1}{\chi_p^2} + \frac{1}{s_{pc}^2}}.$$

If either the sampling error or sorting bias goes to 0, the root mean squared error (RMSE) e_{pc} converges to 0 because the optimal forecast puts weight purely on the measure that provides the most accurate prediction. At the other extreme, if the sampling error s_{pc} gets very large, e_{pc} is bounded above by χ_p , the error in the permanent resident prediction. As a result, one obtains forecasts

that have much lower RMSE than forecasts based purely on $\hat{\mu}_{pc}$ (which would have $\text{RMSE} = s_{pc}$) especially in smaller CZs.³³

In addition to having much lower MSE than the raw fixed effects $\hat{\mu}_{pc}$, an attractive feature of μ_{pc}^f is that it is forecast unbiased: moving a child to a county with a one percentile higher forecasted effect increases that child's income in adulthood by one percentile on average. In this sense, the forecasts μ_{pc}^f provide unbiased predictions of the expected impacts of moving to a different area on children's outcomes.

VI.B. Forecasts for Commuting Zones

Figure III presents maps of the forecasted place effects μ_{pc}^f across CZs for children in below-median ($p = 25$) and above-median ($p = 75$) income families, with lighter colors (color version online) depicting areas that produce better outcomes.³⁴ **Table III** lists the forecasts for the 50 most highly populated CZs (which accounted for 55.5% of the U.S. population in 2000), sorted in descending order based on $\mu_{25,c}^f$, the forecasted effect for low-income families.

1. Estimates for Low-Income Families. Among the 50 largest CZs, Salt Lake City, UT, has the most positive forecasted causal effect for children in below-median income families. We predict that every additional year spent growing up in Salt Lake City will increase a child's income by 0.17 percentiles (RMSE = 0.07) relative to an average CZ. Rescaling the estimates as described in **Section V.B** into dollar impacts, this estimate implies that growing up in Salt Lake City from birth (assuming 20 years of exposure) would increase children's incomes at age 26 by 10.4% relative to growing up in the average CZ. Conversely, at the bottom of the list, every additional year spent growing up in New Orleans is predicted to reduce a child's income by 0.21 percentiles (RMSE = 0.07) relative to an average CZ. This estimate implies that growing

33. From a Bayesian perspective, under the simplifying assumption that the $\hat{\mu}_{pc}$ estimates are drawn independently across places, $\mu_{pc}^f = E[\mu_{pc} | \hat{\mu}_{pc}, \bar{y}_{pc}, s_{pc}]$ is the posterior expectation of each place's causal effect given a normal prior and likelihood function. The standard deviation of the posterior distribution is e_{pc} and the true parameter μ_{pc} lies within the credible interval $\mu_{pc}^f \pm 1.96e_{pc}$ with 95% probability.

34. The maps are colored by grouping CZs into (unweighted) deciles. The deciles are not symmetric around zero because μ_{pc}^f is normalized to have a population-weighted mean of zero and population density is negatively correlated with μ_{pc}^f .

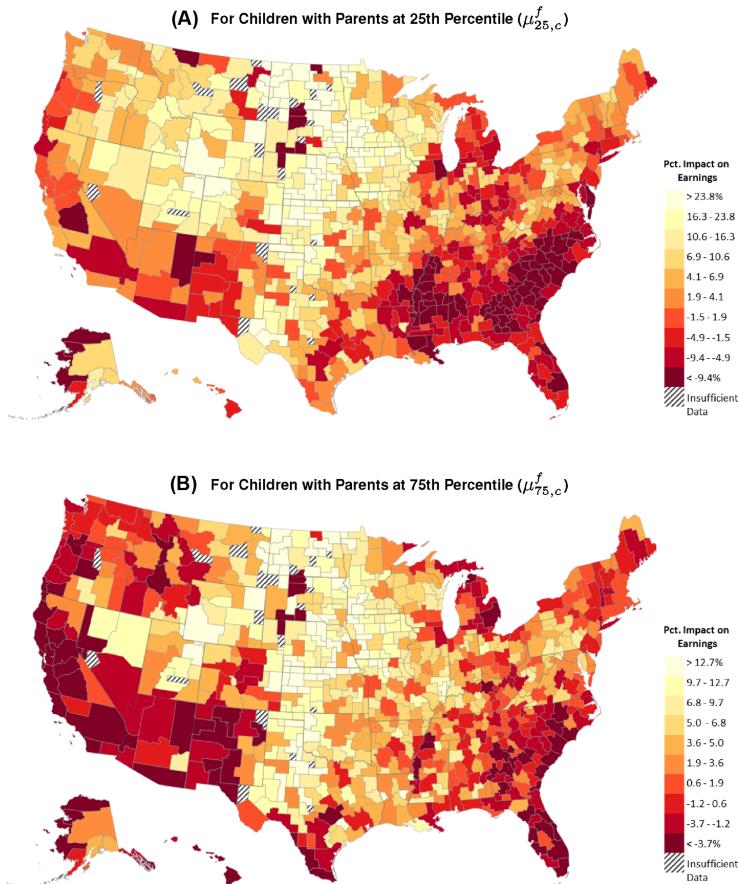


FIGURE III

Forecasts of Causal Effects on Children's Income by Commuting Zone

These maps show MSE-minimizing forecasts of each commuting zone's causal effect, μ_{pc}^f , on children's household income at age 26. Panel A shows estimates for children in below-median income families ($p = 25$), and Panel B shows estimates for children in above-median income families ($p = 75$). These forecasts are constructed using the methodology described in the notes to Figure II. Estimates are scaled to show the percentage change in income from growing up from birth (i.e., 20 years of childhood exposure) in a given CZ relative to the (population-weighted) average CZ in the country. CZs are grouped into (unweighted) deciles based on their causal effects, with lighter colors depicting areas with more positive causal effects. For example, growing up in CZs in the highest decile raises children's incomes at $p = 25$ by more than 23.8% relative to the average CZ, whereas growing up in the CZs in the lowest decile reduces their incomes by 9.4% relative to the average CZ. CZs with fewer than 250 permanent residents, for which we do not report permanent resident outcomes and therefore do not have forecasts of causal effects, are shaded with the striped pattern.

TABLE III

MSE-MINIMIZING FORECASTS OF CAUSAL EFFECTS FOR 50 LARGEST COMMUTING ZONES

Rank	Commuting ($p = 25$) zone	State	Below-median income parents ($p = 25$)				Above-median income parents ($p = 75$)			
			Impact on rank		% Impact from birth		Impact on rank		% Impact from birth	
			Forecast	RMSE	Forecast	Perm. res.	Forecast	RMSE	Forecast	Perm. res.
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	Salt Lake City	UT	0.17	0.07	10.4	13.2	0.11	0.04	4.4	4.8
2	Seattle	WA	0.14	0.06	8.8	4.3	-0.01	0.04	-0.4	-1.2
3	Washington DC	DC	0.10	0.05	6.6	1.4	0.06	0.03	2.5	0.2
4	Minneapolis	MN	0.10	0.07	6.4	6.0	0.08	0.04	3.2	3.2
5	Fort Worth	TX	0.06	0.06	3.6	0.8	0.05	0.04	2.0	1.5
6	San Diego	CA	0.06	0.05	3.5	3.7	-0.13	0.04	-5.4	-5.1
7	Boston	MA	0.06	0.06	3.5	4.3	0.03	0.04	1.4	1.3
8	Manchester	NH	0.05	0.07	3.2	5.9	0.02	0.04	1.0	1.0
9	San Jose	CA	0.05	0.06	3.0	4.8	-0.12	0.04	-4.9	-5.1
10	Las Vegas	NV	0.04	0.06	2.7	-0.1	-0.08	0.04	-3.2	-2.7
11	Denver	CO	0.04	0.07	2.6	2.6	-0.06	0.04	-2.5	-1.4
12	Portland	OR	0.04	0.07	2.4	0.7	-0.09	0.04	-3.8	-4.6
13	San Francisco	CA	0.03	0.06	1.8	2.3	-0.12	0.04	-4.9	-5.0
14	Pittsburgh	PA	0.01	0.06	0.8	3.0	0.10	0.04	4.3	4.0
15	Newark	NJ	0.01	0.05	0.7	1.1	0.06	0.03	2.4	1.5
16	Providence	RI	0.01	0.07	0.4	2.0	0.02	0.04	0.9	1.4
17	Sacramento	CA	0.01	0.06	0.4	1.3	-0.14	0.04	-5.9	-4.7
18	Phoenix	AZ	0.00	0.05	0.2	0.6	-0.02	0.04	-0.7	-1.7
19	Buffalo	NY	0.00	0.07	-0.2	-2.5	0.01	0.04	0.4	-0.2
20	Kansas City	MO	-0.01	0.07	-0.4	-2.3	0.02	0.04	0.8	1.1
21	Houston	TX	-0.03	0.05	-1.6	-0.3	0.01	0.04	0.3	-0.3
22	Miami	FL	-0.03	0.04	-1.6	-2.3	-0.20	0.04	-8.3	-7.8
23	Philadelphia	PA	-0.03	0.06	-1.8	-4.7	0.00	0.04	0.2	0.9
24	Grand Rapids	MI	-0.03	0.07	-2.0	-3.0	0.07	0.04	2.7	2.2
25	Dallas	TX	-0.04	0.06	-2.4	-3.1	-0.01	0.04	-0.4	-0.8
26	Cleveland	OH	-0.04	0.06	-2.7	-9.2	-0.03	0.04	-1.0	-1.7
27	Bridgeport	CT	-0.05	0.06	-2.9	0.7	0.03	0.04	1.2	1.5
28	Jacksonville	FL	-0.05	0.06	-3.0	-7.3	-0.07	0.04	-2.9	-2.7
29	Milwaukee	WI	-0.05	0.07	-3.0	-5.4	0.04	0.04	1.8	2.1
30	Dayton	OH	-0.06	0.07	-3.9	-6.4	0.02	0.04	0.6	0.4
31	Cincinnati	OH	-0.08	0.07	-5.2	-7.4	0.06	0.04	2.6	1.3
32	Columbus	OH	-0.09	0.07	-5.4	-7.1	0.01	0.04	0.3	-1.1
33	Nashville	TN	-0.09	0.07	-5.5	-6.4	-0.03	0.04	-1.1	-1.8
34	St. Louis	MO	-0.09	0.07	-5.6	-7.3	0.03	0.04	1.2	0.7
35	Austin	TX	-0.10	0.07	-6.1	-4.0	-0.10	0.04	-4.1	-3.4
36	Baltimore	MD	-0.10	0.07	-6.4	-7.5	0.07	0.04	2.8	2.1
37	San Antonio	TX	-0.11	0.06	-6.9	-3.5	-0.08	0.04	-3.2	-3.3
38	Tampa	FL	-0.11	0.05	-7.1	-4.5	-0.13	0.04	-5.3	-4.2
39	New York	NY	-0.12	0.04	-7.3	-0.9	-0.03	0.04	-1.3	-1.1
40	Indianapolis	IN	-0.12	0.07	-7.4	-8.5	-0.02	0.04	-0.8	0.1
41	Atlanta	GA	-0.12	0.04	-7.8	-12.5	-0.09	0.04	-3.9	-7.4
42	Los Angeles	CA	-0.13	0.04	-8.1	0.8	-0.23	0.03	-9.3	-7.4
43	Detroit	MI	-0.14	0.05	-8.5	-11.8	-0.13	0.04	-5.2	-6.7
44	Orlando	FL	-0.14	0.05	-8.5	-4.6	-0.14	0.04	-5.7	-3.9
45	Chicago	IL	-0.15	0.05	-9.7	-6.9	-0.03	0.03	-1.4	-1.3
46	Fresno	CA	-0.16	0.06	-10.3	-1.8	-0.12	0.04	-5.0	-4.3

TABLE III
CONTINUED

Rank (<i>p</i> = 25)	Commuting zone	State	Below-median income parents (<i>p</i> = 25)				Above-median income parents (<i>p</i> = 75)			
			Impact on rank		% Impact from birth		Impact on rank		% Impact from birth	
			Forecast	RMSE	Forecast	Perm. res.	Forecast	RMSE	Forecast	Perm. res.
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
47	Port St. Lucie	FL	-0.17	0.06	-10.9	-5.8	-0.20	0.04	-8.2	-6.2
48	Raleigh	NC	-0.19	0.06	-12.2	-9.6	-0.11	0.04	-4.7	-3.3
49	Charlotte	NC	-0.20	0.06	-12.8	-11.0	-0.08	0.04	-3.5	-2.2
50	New Orleans	LA	-0.21	0.06	-13.4	-7.8	-0.06	0.04	-2.5	-2.6

Notes. This table presents MSE-minimizing forecasts of causal effects on children's incomes in adulthood for the 50 most populous CZs. Columns (1)–(4) present estimates for children growing up in below-median income (*p* = 25) families. Column (1) reports estimates of the causal impact of spending an additional year of childhood in a given CZ relative to the (population-weighted) average CZ in the country on a child's household income rank at age 26. Column (2) reports the root mean squared error of this forecast. Column (3) rescales the estimates in column (1) to show the percentage change in income from growing up from birth (i.e., 20 years of childhood exposure) in a given CZ. Column (4) reports the same statistic, but for a forecast that is constructed solely based on the outcomes of permanent residents, excluding the data on movers. Columns (5)–(8) report the analogous statistics for above-median income (*p* = 75) families. The table is sorted based on CZs' causal impacts at *p* = 25.

up in New Orleans from birth would reduce a child's income by 13.4% relative to the average CZ and 23.8% relative to Salt Lake City.

Figure III shows that many of the places that produce the highest incomes at *p* = 25 are in the rural Midwest, which generate income gains (from birth) exceeding 23.8% relative to the average area. Certain parts of the Northeast and West Coast also generate very good outcomes, with gains above 10.6%. The Southeast produces some of the worst outcomes for children in low-income families, with income losses exceeding 9.4% relative to the average place. Parts of the industrial Midwest and other areas such as the large Native American reservations in South Dakota and Arizona generate very negative outcomes as well.

2. *Causal Effects versus Permanent Residents' Outcomes.* The geographical patterns of forecasted causal effects $\mu_{25,c}^f$ in Figure III are broadly similar to the geographical patterns of permanent residents' outcomes $\bar{y}_{25,c}$ in observational data (Chetty and Hendren 2018a, Figure II), but there are several notable differences. For instance, Los Angeles is above the national average in terms of its permanent residents' outcomes $\bar{y}_{25,c}$, but it is among the worst cities in terms of its causal effect on low-income children $\mu_{25,c}^f$ (Table III).

To quantify these differences more systematically, in the fourth column of **Table III** we show the forecast $\gamma_{25}(\bar{y}_{25,c} - \bar{y}_{25})$ that one would obtain if one were to use data only on permanent residents (rescaled into percentage impacts). In Los Angeles, $\mu_{25,c}^f = -8.1$, whereas the prediction based on permanent residents is +0.8%. Similarly, in New York, the causal impact is much more negative than one would predict based on permanent residents' outcomes. In Washington, DC, the pattern is reversed: permanent residents' outcomes are close to the national mean, but the forecasted causal effect is +6.6%, among the highest for large CZs. These differences between the causal forecasts μ_{pc}^f and the permanent residents' outcomes \bar{y}_{pc} can be interpreted as selection effects among permanent residents under our modeling assumptions, as shown in [equation \(7\)](#). We present a comprehensive analysis of the factors that drive the differences between μ_{pc}^f and \bar{y}_{pc} in Section VII.

Overall, the (population-weighted) correlation between $\mu_{25,c}^f$ and $\bar{y}_{25,c}$ is 0.69 among the 50 largest CZs shown in Table III and 0.89 across all CZs. The correlation between $\mu_{25,c}^f$ and $\bar{y}_{25,c}$ is higher in smaller CZs because the MSE-minimizing forecast in [equation \(11\)](#) puts more weight on $\bar{y}_{25,c}$ when the causal effect is estimated with less precision. In short, **Table III** shows that the permanent residents' outcomes used in our first article provide a very good starting point to predict places' causal effects, but combining that data with information on movers' outcomes as we do here yields much better predictions, especially in highly populated areas.

3. Estimates for High-Income Families. The right half of Table III and lower panel of [Figure III](#) show forecasts of place effects for children in above-median income ($p = 75$) families, $\mu_{75,c}^f$. As discussed in [Section V](#), the variation in causal effects as a percentage of income is much smaller for families at $p = 75$ than $p = 25$. The predicted impact of moving from the worst CZ (Los Angeles) to the best CZ (Salt Lake City) is 13.7% for children in above-median income families, roughly half the corresponding range for children in below-median income families.

The geographical variation in causal effects at $p = 75$ is weakly positively correlated with the variation at $p = 25$. Rural areas produce better outcomes at $p = 75$, particularly in the Midwest. The Southeast tends to produce worse outcomes, although parts of the South, such as Louisiana and Arkansas, generate considerably better outcomes for the rich than for the poor, thereby

amplifying inequality across generations. Perhaps most strikingly, much of the West Coast and parts of the Northeast have the lowest values of $\mu_{75.c}^f$. This result turns out to be driven by measuring children's incomes at the household rather than individual level, as we discuss next.

4. *Individual Income and Marriage Rates.* In our baseline analysis, we measure children's outcomes at the household level, summing the incomes of spouses for married couples and using own income for single individuals. [Online Appendix](#) Figure IV and [Appendix Tables III–IV](#) replicate [Figure III](#) and [Tables III–IV](#), measuring children's income at the individual level instead. The geographical patterns are broadly similar, with a (population-weighted) correlation between the household-income and individual-income estimates of 0.75 at $p = 25$ and 0.59 at $p = 75$. However, in certain areas—most notably in coastal California and the Northeast at $p = 75$ —the patterns differ sharply. These areas generate relatively high levels of individual income even though they have among the lowest levels of household income. For example, growing up in the San Francisco CZ from birth in a $p = 75$ family is predicted to increase individual income at age 26 by 0.7% but reduce household income by 4.9% relative to the average CZ. Conversely, Salt Lake City has much more positive impacts on household income than individual income.

Places' causal effects on individual and household income differ largely because they have different causal effects on children's rates of marriage. As shown in [Chetty and Hendren \(2018a, Figure VIIIb\)](#), places have linear childhood exposure effects on rates of marriage as well. We can therefore forecast each area's causal effect on marriage rates using the same approach as above, defining the outcome as an indicator for being married at age 26 instead of a child's income rank at age 26. The estimates are presented in [Online Appendix](#) Tables V and VI. Growing up in Salt Lake City from birth increases a given child's probability of being married by 10.8 percentage points at $p = 25$ and 15.8 percentage points at $p = 75$; the corresponding forecasts in San Francisco are –2.3 percentage points and –8.0 percentage points. More generally, the areas that produce the highest rates of marriage tend to produce higher levels of household income than individual income.

VI.C. Forecasts for Counties

[Table IV](#) presents forecasts of causal effects (on household income) for the 100 most populous counties, focusing on those in

TABLE IV

MSE-MINIMIZING FORECASTS OF CAUSAL EFFECTS FOR 100 LARGEST COUNTIES (TOP AND BOTTOM 25)

Rank (<i>p</i> = 25)	County	State	Below-median income parents (<i>p</i> = 25)				Above-median income parents (<i>p</i> = 75)			
			Impact on rank		% Impact from birth		Impact on rank		% Impact from birth	
			Forecast	RMSE	Forecast	Perm. res.	Forecast	RMSE	Forecast	Perm. res.
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)			
Panel A: Top 25 large counties										
1	Dupage	IL	0.26	0.09	16.0	10.6	0.08	0.08	3.1	1.8
2	Fairfax	VA	0.24	0.10	15.0	13.6	0.26	0.10	11.0	2.5
3	Snohomish	WA	0.22	0.10	14.0	6.6	0.06	0.09	2.4	0.3
4	Bergen	NJ	0.22	0.10	13.8	11.5	0.15	0.10	6.3	2.4
5	Bucks	PA	0.20	0.10	12.4	8.1	-0.02	0.10	-0.9	2.1
6	Norfolk	MA	0.18	0.10	11.5	11.8	0.15	0.10	6.2	2.0
7	Montgomery	PA	0.16	0.10	9.7	7.3	0.07	0.09	3.0	1.8
8	Montgomery	MD	0.15	0.10	9.5	7.8	0.00	0.10	0.1	0.0
9	King	WA	0.15	0.08	9.3	3.2	0.08	0.08	3.2	-1.3
10	Middlesex	NJ	0.15	0.10	9.1	7.0	0.01	0.10	0.5	2.0
11	Contra Costa	CA	0.14	0.09	8.8	2.1	-0.07	0.09	-2.9	-2.7
12	Middlesex	MA	0.12	0.09	7.7	8.5	0.01	0.09	0.5	1.0
13	Macomb	MI	0.11	0.09	7.0	-1.0	0.03	0.09	1.1	-1.8
14	Salt Lake	UT	0.10	0.10	6.2	9.7	0.02	0.09	0.7	2.1
15	Ventura	CA	0.10	0.10	6.2	5.0	-0.05	0.09	-2.3	-3.1
16	San Mateo	CA	0.08	0.10	5.3	6.2	-0.04	0.10	-1.5	-1.7
17	Worcester	MA	0.08	0.11	4.7	2.6	0.13	0.11	5.4	1.6
18	Monmouth	NJ	0.07	0.10	4.7	2.4	0.07	0.10	3.0	1.0
19	Honolulu	HI	0.07	0.10	4.6	2.1	-0.13	0.11	-5.4	-3.8
20	Hudson	NJ	0.07	0.10	4.2	-0.4	0.16	0.11	6.7	-0.1
21	Kern	CA	0.06	0.09	3.9	5.1	-0.06	0.11	-2.5	-0.1
22	Clark	NV	0.06	0.07	3.7	-0.7	-0.05	0.09	-1.9	-2.2
23	San Diego	CA	0.06	0.06	3.7	2.6	-0.14	0.06	-5.6	-3.4
24	Providence	RI	0.05	0.10	3.0	-0.4	-0.04	0.11	-1.8	0.5
25	San Francisco	CA	0.05	0.10	2.8	4.4	-0.18	0.10	-7.6	-4.9
Panel B: Bottom 25 large counties										
75	Jefferson	KY	-0.14	0.10	-8.6	-11.4	0.02	0.11	0.9	-2.4
76	Franklin	OH	-0.14	0.09	-8.6	-11.4	0.11	0.10	4.7	-2.4
77	San Bernardino	CA	-0.14	0.06	-8.8	0.2	-0.25	0.07	-10.1	-3.3
78	Davidson	TN	-0.14	0.10	-8.9	-13.9	-0.04	0.11	-1.5	-4.5
79	Pima	AZ	-0.14	0.08	-8.9	-4.7	-0.14	0.10	-5.7	-4.2
80	Montgomery	OH	-0.14	0.10	-8.9	-11.7	-0.02	0.12	-0.6	-1.9
81	Travis	TX	-0.15	0.09	-9.2	-7.8	-0.16	0.09	-6.6	-4.4
82	Essex	NJ	-0.15	0.10	-9.2	-8.4	0.07	0.10	3.0	-2.1
83	Bexar	TX	-0.15	0.09	-9.6	-4.8	-0.09	0.12	-3.8	-3.6
84	Milwaukee	WI	-0.16	0.10	-9.9	-12.0	-0.03	0.10	-1.1	-1.7
85	Riverside	CA	-0.16	0.07	-10.1	1.2	-0.25	0.08	-10.3	-3.4
86	Los Angeles	CA	-0.16	0.05	-10.3	-0.5	-0.25	0.05	-10.5	-6.0
87	Wake	NC	-0.17	0.10	-10.7	-6.7	-0.09	0.10	-3.9	-2.1
88	New York	NY	-0.17	0.08	-10.8	-5.3	-0.27	0.10	-11.4	-6.9
89	Fulton	GA	-0.17	0.08	-10.9	-15.8	0.02	0.08	1.0	-7.4
90	Bronx	NY	-0.17	0.08	-10.9	-7.2	-0.20	0.11	-8.3	-6.1
91	Wayne	MI	-0.18	0.08	-11.4	-14.3	-0.07	0.08	-3.0	-5.6
92	Orange	FL	-0.19	0.08	-12.1	-6.1	-0.09	0.09	-3.8	-3.6

TABLE IV
CONTINUED

Rank (<i>p</i> = 25)	County	State	Below-median income parents (<i>p</i> = 25)				Above-median income parents (<i>p</i> = 75)			
			Impact on rank		% Impact from birth		Impact on rank		% Impact from birth	
			Forecast	RMSE	Forecast	Perm. res.	Forecast	RMSE	Forecast	Perm. res.
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
93	Cook	IL	-0.20	0.06	-12.8	-9.7	-0.03	0.05	-1.2	-2.6
94	Palm Beach	FL	-0.21	0.08	-13.0	-5.6	-0.31	0.10	-13.0	-4.6
95	Marion	IN	-0.21	0.10	-13.1	-13.3	-0.10	0.09	-4.2	-2.8
96	Shelby	TN	-0.21	0.09	-13.1	-16.8	0.03	0.10	1.2	-4.9
97	Fresno	CA	-0.22	0.09	-13.5	-3.7	-0.05	0.11	-2.1	-3.5
98	Hillsborough	FL	-0.22	0.09	-13.8	-4.3	-0.19	0.10	-7.9	-2.4
99	Baltimore City	MD	-0.22	0.09	-14.0	-15.5	-0.02	0.10	-0.7	-6.4
100	Mecklenburg	NC	-0.23	0.10	-14.5	-13.4	-0.09	0.10	-3.7	-3.8

Notes. This table presents MSE-minimizing forecasts of counties' causal effects on children's incomes in adulthood. The table reports estimates for counties that are in the top or bottom 25 among the 100 most populous counties based on their impacts on below-median income families (*p* = 25). Columns (1)–(4) present estimates for children growing up in below-median income (*p* = 25) families. Column (1) reports estimates of the causal impact of spending an additional year of childhood in a given county relative to the (population-weighted) average county in the country on a child's household income rank at age 26. Column (2) reports the root mean squared error of this forecast. Column (3) rescales the estimates in column (1) to show the percentage change in income from growing up from birth (i.e., 20 years of childhood exposure) in a given county. Column (4) reports the same statistic, but for a forecast that is constructed solely based on the outcomes of permanent residents, excluding the data on movers. Columns (5)–(8) report the analogous statistics for above-median income (*p* = 75) families. The table is sorted based on counties' causal impacts at *p* = 25.

the top and bottom 25 based on $\mu_{25,c}^f$. Analogous estimates for all counties in the United States are available in [Online Appendix Data Table 2](#).³⁵

1. *Estimates for Low-Income Families.* DuPage County (the western suburbs of Chicago) produces the best outcomes for children from below-median income families among the 100 largest counties. Growing up from birth in DuPage County would increase a child's income by 16.0% relative to the average county. The counties that produce the best outcomes are dispersed across the country: they include Fairfax County in Virginia, Snohomish County in Washington, and Bergen County in New Jersey. At the bottom of the list, Mecklenburg County (the city of Charlotte in

35. We construct these forecasts by applying [equation \(11\)](#) directly to the county fixed effect estimates $\hat{\mu}_{pc}$, defined as the sum of the CZ and county-within-CZ fixed effect estimates as discussed in [Section III.C](#). An alternative approach which permits different shrinkage factors at the county-within-CZ and CZ levels is to forecast the effect of each county relative to the average county within each CZ and then add these county-within-CZ forecasts to the CZ-level forecasts. In practice, these two methods deliver very similar results.

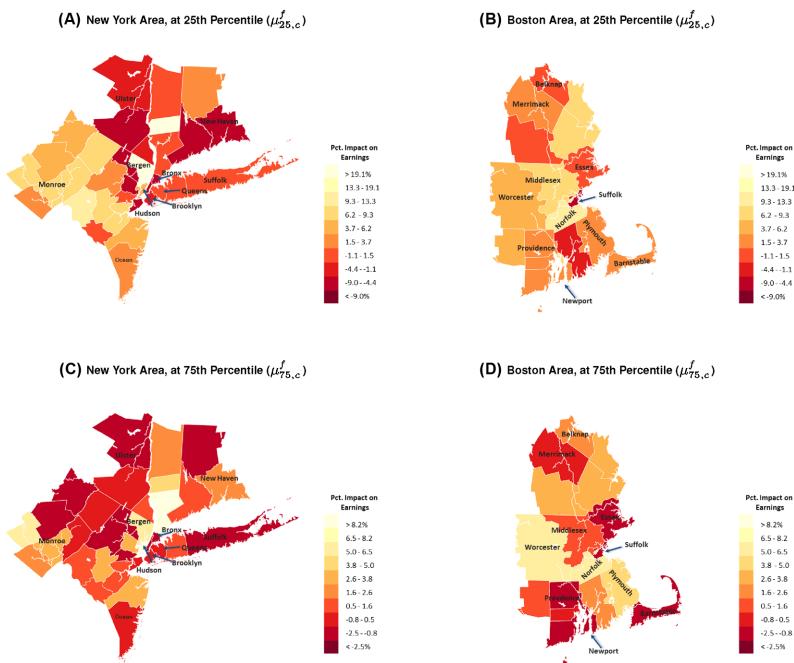


FIGURE IV

Forecasts of Causal Effects by County in the New York and Boston Areas

These maps show MSE-minimizing forecasts of each county's causal effect, μ_{pc}^f , on children's household income at age 26 for counties in the New York and Boston combined statistical areas. Panels A and B show estimates for children in below-median income families ($p = 25$), and Panels C and D show estimates for children in above-median income families ($p = 75$). These county-level forecasts are constructed using an approach analogous to that described in the notes to Figure II. Estimates are scaled to show the percentage change in income from growing up from birth (i.e., 20 years of childhood exposure) in a given county relative to the (population-weighted) average county in the country. Counties are grouped into deciles at the national level based on their causal effects, with lighter colors depicting areas with more positive causal effects.

North Carolina) generates the most negative outcomes, reducing children's incomes by 14.5%. Baltimore City in Maryland, Hillsborough County in Florida, and Fresno County in California also produce large income losses for children in low-income families.

Counties' causal effects vary substantially even within metro areas. Figure IV illustrates this variation by mapping $\mu_{25,c}^f$ for counties in the New York City and Boston combined statistical

areas (CSAs). The estimates imply that growing up in a low-income family in Manhattan from birth reduces children's incomes by 10.8% relative to the mean, whereas growing up across the Hudson river in Hudson County, NJ, increases children's incomes by 4.2% relative to the mean. Hence, moving from Manhattan to Hudson County at birth would increase children's incomes by 15.0%. Likewise, moving at birth from the city of Boston (Suffolk County) across the Charles River to Middlesex County would increase children's incomes by 13.8%. These maps illustrate a pattern observed in many cities, which is that city centers tend to produce worse outcomes, particularly for children in low-income families, than do suburbs.

One of the most striking examples of local area variation is the difference between DuPage County, the best county in the United States, and Cook County (the city of Chicago), which is the eighth worst county in the United States among the 100 largest counties. Moving from Chicago proper to the western suburbs of Chicago at birth would increase a child's household income by \$7,510 a year on average, a 28.8% increase. This comparison is of particular interest in light of a well-known 1976 U.S. Supreme Court ruling, *Hills v. Gautreaux*, which required that the Chicago Housing Authority provide residents living in high poverty housing projects in Cook County an opportunity to move to lower-poverty neighborhoods in the suburbs, many of which were in DuPage County. Observational studies comparing the outcomes of families who accepted offers to move to the suburbs to those who chose to remain in the city (e.g., Rosenbaum 1995) have found that children whose families moved to the suburbs had significantly better economic outcomes. The interpretation of these findings have been debated because the families who chose to move may have had better unobservables (θ_i in equation (1)). Our findings support the view that the gains observed for families who moved as part of Gautreaux reflect the causal impact of growing up in DuPage County instead of Cook County.

2. *Estimates for High-Income Families.* For children in above median income ($p = 75$) families, the best county in the United States is Fairfax, VA, which produces an income gain of 11.0%, and the worst county is Palm Beach, FL, which produces an income loss of 13.0%. Mirroring the CZ-level results, the degree of variation in outcomes across counties is generally smaller at $p = 75$ than $p = 25$, and there is a weak positive correlation between $\mu_{25,c}^f$ and $\mu_{75,c}^f$. For example, in Figure IV, the city of Boston produces

worse outcomes (even for children in rich families) than Middlesex County; similarly, in New York, Manhattan produces worse outcomes than Hudson County.

3. Heterogeneity by Gender. The places that generate the best outcomes for boys generally generate the best outcomes for girls as well ([Online Appendix](#) Figures V and VI), but there is a thick lower tail of counties that produce particularly negative outcomes for boys in low income families ([Online Appendix](#) Figure VII). For instance, growing up in Baltimore City from birth reduces household income by 27.9% for boys relative to the mean, but only 5.4% for girls ([Online Appendix](#) Tables VII–X). One explanation for why certain areas produce extremely low incomes for boys is that males in these areas are particularly likely to be incarcerated, and individuals who are incarcerated are included in our sample with zero or very low incomes.³⁶

VII. CHARACTERISTICS OF HIGH-OPPORTUNITY AREAS

What are the characteristics of places that produce high levels of upward mobility? Prior work has shown that in observational data, upward mobility is highly correlated with several factors: residential segregation, income inequality, the fraction of single mothers, social capital, school quality, and racial shares ([Sampson, Morenoff, and Gannon-Rowley 2002](#); [Chetty et al. 2014](#)). These correlations could reflect two very different phenomena. One possibility is that they are predictive of places' causal effects: for instance, growing up in a more segregated area may cause worse economic outcomes for a given child (lower μ_{pc}) ([Wilson 1987](#); [Sampson 2012](#)). Another possibility is that they capture sorting: the types of people who live in more segregated places may have different characteristics θ_i . Distinguishing between these two explanations is critical for understanding what types of areas produce the greatest economic opportunity, as opposed to simply attracting upwardly mobile people.

In this section, we decompose the correlations documented in prior work into causal versus sorting components by correlating local area characteristics with both our causal effect

36. [Chetty et al. \(2016\)](#) further explore heterogeneity in intergenerational mobility by gender and show that areas with concentrated poverty and crime have particularly negative impacts on the incomes and employment rates of boys growing up in low-income families.

estimates $\hat{\mu}_{pc}$, based on movers, and permanent residents' outcomes \bar{y}_{pc} , which combine both causal effects and selection effects. This analysis identifies the factors associated with the production of upward mobility; however, we caution that it does not show that these factors themselves have a direct causal effect on upward mobility.³⁷

We organize our analysis into three parts. First, we analyze correlations between the incomes of children in low-income families ($\mu_{25,c}$) and area characteristics that are defined at the group level: segregation, inequality, school quality, social capital. We then examine correlations between $\mu_{25,c}$ and demographic characteristics that are aggregates of variables defined at the individual level: the fraction of single mothers, immigrant shares, and racial shares. We find that correlations between the group-level characteristics and upward mobility are driven almost entirely by causal effects of place, whereas a substantial portion of the correlations between the individual-level demographic aggregates and upward mobility is due to selection. Finally, we examine correlations between these factors and outcomes for children growing up in high-income families ($\mu_{75,c}$). In the interest of space, we focus on a subset of correlations that illustrate the key results, shown in Figures V ($p = 25$) and VI ($p = 75$). A comprehensive set of correlations with 40 area-level characteristics is presented in Appendix Tables XI– XIV.

VII.A. Group-Level Characteristics

1. *Segregation.* We measure racial segregation across census tracts within each CZ using a Theil index H_c , which quantifies the extent to which the racial distribution in each Census tract differs from the overall racial distribution in the CZ (Chetty et al. 2014, equation 4).³⁸ We begin by examining the association between permanent residents' outcomes $\bar{y}_{25,c}$ and H_c , as in prior observational studies. The vertical tick mark in the first row of Figure V, Panel A shows that growing

37. For instance, we show later that more segregated places produce worse outcomes for low-income children. This does not necessarily imply that policies which reduce segregation will increase upward mobility, because more segregated places may have other characteristics that lead to worse outcomes (such as weaker schools).

38. The characteristics analyzed in this section are taken primarily from Chetty et al. (2014, Online Data Table 8); we provide detailed definitions and sources for all of the variables we use in Online Appendix Table XV of this article.

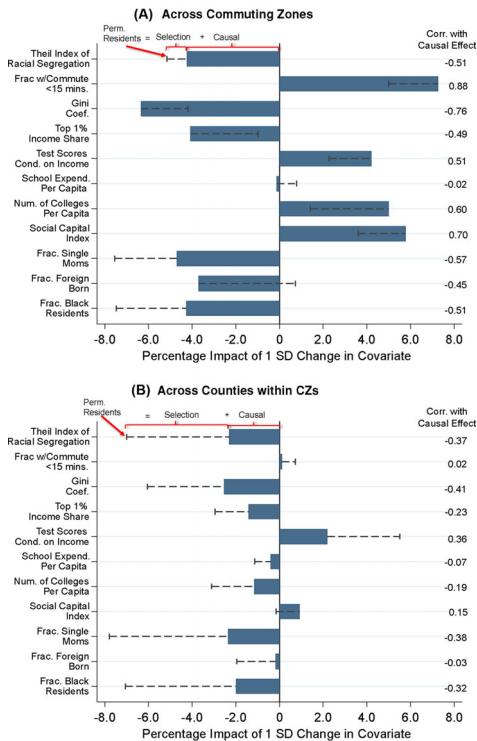


FIGURE V

Predictors of Place Effects for Children with Parents at 25th Percentile

Panel A plots coefficients from univariate OLS regressions of permanent resident outcomes $\bar{y}_{25,c}$ and causal effects $\hat{\mu}_{25,c}$ for below-median income families ($p = 25$) on various CZ-level characteristics (x_c), weighting by population. The characteristics are normalized to have a (population-weighted) mean zero and unit standard deviation across CZs. Both $\bar{y}_{25,c}$ and $\hat{\mu}_{25,c}$ are rescaled so the coefficients can be interpreted as impacts in percentage units using the approach described in Section V.B. The vertical tick marks plot coefficients from regressions of $\bar{y}_{25,c}$ on x_c . The solid bars plot coefficients from regressions of the causal effect of growing up in an area from birth (20 years of exposure), $20\hat{\mu}_{25,c}$, on x_c . The difference between the tick mark and the bar (depicted by the dashed horizontal line) therefore represents the coefficient from a regression of $\bar{y}_{25,c} - 20\hat{\mu}_{25,c}$ on the covariate x_c , which can be interpreted as the association between selection effects and the covariate. The numbers on the right report the correlations between $\hat{\mu}_{25,c}$ and x_c , which are obtained by dividing the coefficient from regressing $20\hat{\mu}_{25,c}$ on x_c by 20 times the standard deviation of $\hat{\mu}_{25,c}$ (from Table II). Panel B presents analogous estimates at the county-within-CZ level, standardizing the covariates to have (population-weighted) mean 0 and unit standard deviation across counties within CZs, and estimating regressions at the county level controlling for CZ fixed effects. Point estimates and standard errors for the characteristics shown in this figure are provided in Online Appendix Tables XI and XII, along with results for additional characteristics. Definitions of the covariates are provided in Online Appendix Table XV.

up in a CZ with one standard deviation higher segregation is associated with a 5.2% reduction in children's incomes for families at $p = 25$ ($\bar{y}_{25,c}$).³⁹ We estimate this 5.2% effect in three steps. First, we normalize H_c into standard deviation units by dividing the raw value of the index by its population-weighted standard deviation across CZs. Next, we regress $\bar{y}_{25,c}$ on the standardized value of H_c , weighting by CZ population. Finally, we multiply the regression coefficient by 3.14 to translate the percentile impact into percentage impacts on incomes, as in [Section V.B](#).

To determine how much of this 5.2% effect is due to causal effects of place versus sorting, we repeat the preceding exercise using the causal effect of growing up in an area from birth ($20 \times \hat{\mu}_{25,c}$) as the outcome.⁴⁰ We estimate that growing up in a one standard deviation more segregated CZ from birth reduces a given child's income by 4.2%, depicted by the solid bar in [Figure V](#), Panel A. Hence, $\frac{4.2}{5.2} = 80.8\%$ of the association between segregation and permanent residents' outcomes reflects the causal effect of place. Under our modeling assumption that place effects μ_{pc} do not vary between movers and permanent residents, the remaining 19.2% of the association—depicted by the dashed line in [Figure V](#), Panel A reflects sorting, that is, the association between $\bar{\theta}_{25,c} = \bar{y}_{25,c} - 20\hat{\mu}_{25,c}$ and H_c .⁴¹

This analysis shows that the majority of the association between segregation and upward mobility across CZs documented in prior work can be explained by the causal effect of place rather than sorting. The (population-weighted) correlation between the latent causal effect $\mu_{25,c}$ and H_c is -0.51 (shown on the right side of [Figure V](#), Panel A), implying that segregation explains a significant portion of the variation in children's outcomes across

39. Standard errors for all of the estimates shown in [Figures V](#) and [VI](#) are given in [Online Appendix](#) Tables XI and XII. All of the estimates discussed below are statistically significant with $p < .05$ unless otherwise noted.

40. We multiply the annual exposure effect estimates by 20 to obtain impacts from birth for the reasons described in [Section IV.A](#); using other plausible scaling factors (e.g., 15 or 23) yields qualitatively similar results.

41. If the underlying causal effects μ_{pc} vary across movers and permanent residents for example, because movers tend to live in different neighborhoods within CZs relative to permanent residents then $20\hat{\mu}_{25,c}$ would not capture the causal effect of place for permanent residents. In this case, the "selection" term $\bar{y}_{25,c} - 20\hat{\mu}_{25,c}$ would also capture the difference between place effects for permanent residents and movers.

CZs.⁴² We find similar results using measures of income segregation. The correlation between a Theil index of income segregation and $\mu_{25,c}$ is -0.57 (Online Appendix Table XI).

We also find a strong negative correlation with measures of sprawl, which is strongly associated with segregation. The correlation between the fraction of people with a commute time less than 15 minutes and $\mu_{25,c}$ is 0.88 . This is the single largest correlation we find across all 40 covariates we analyzed. Growing up in a CZ that is one standard deviation lower in the distribution of sprawl (as measured by commute times) would increase a given child's income by more than 7% on average.

In Figure V, Panel B, we replicate the analysis in Panel A across counties within CZs. We obtain these county-within-CZ estimates by estimating regressions analogous to those described above at the county level, including CZ fixed effects and normalizing each covariate to have a standard deviation of one across counties within CZs. Segregation remains a strong predictor of causal effects at the county level: the correlation between racial segregation and $\mu_{25,c}$ is -0.37 across counties within CZs. However, the sorting component is larger at the county level: more than two-thirds of the association between segregation and permanent residents' outcomes $\bar{y}_{25,c}$ is due to sorting. This finding is consistent with the intuition that families seeking better opportunities for their children are more likely to sort within rather than between labor markets.

In sum, our analysis strongly supports the hypothesis that growing up in a more segregated area—that is, in a neighborhood with concentrated poverty—is detrimental for disadvantaged youth. However, the mechanisms underlying our findings diverge from some of the theories posited in prior work. For instance, some studies have proposed that segregation is associated with worse outcomes for the poor because of spatial mismatch in access to jobs (Kain 1968; Kasarda 1989; Wilson 1996), an explanation that may appear particularly plausible given the strong correlation between upward mobility and commute times. However, the fact that our causal effect estimates $\hat{\mu}_{25,c}$ are identified from differences in childhood exposure is inconsistent with this

42. We estimate this correlation as $\text{Corr}(\mu_{25,c}, H_c) = \text{Corr}(\hat{\mu}_{25,c}, H_c) \frac{SD(\hat{\mu}_{25,c})}{SD(\mu_{25,c})}$, where the ratio of standard deviations is obtained from Table II and adjusts for the attenuation in $\text{Corr}(\hat{\mu}_{25,c}, H_c)$ due to sampling error in the raw fixed effect estimates $\hat{\mu}_{25,c}$.

theory. Our analysis shows that moving to a more sprawling, segregated city at an earlier age (e.g., 10 instead of 12) reduces a child's income in adulthood, demonstrating that these effects cannot be directly driven by a lack of access to jobs in adulthood. Moreover, we find a strong negative association between population density and $\mu_{25,c}$, showing that urban areas, which have more jobs, tend to be worse for upward mobility. Overall, these findings are more consistent with theories that emphasize peer effects or a lack of resources as an explanation for why growing up in a more segregated area reduces upward mobility.

2. *Income Inequality.* CZs and counties with greater income inequality produce significantly worse outcomes for children in low-income families. At the CZ level, the correlation between $\mu_{25,c}$ and the Gini coefficient is -0.76 . Growing up from birth in an area with a one standard deviation higher Gini coefficient reduces a given child's income by 6.4%. Interestingly, this effect is larger than the -4.2% effect observed for permanent residents. Under the assumption that the causal effect μ_{pc} is the same for movers and permanent residents, it follows that the sorting component $\bar{\theta}_{25,c}$ offsets the causal component $\mu_{25,c}$ in observational data. That is, residents of areas with high levels of income inequality tend to have better unobservables $\bar{\theta}_{25,c}$, leading prior observational studies (e.g., Chetty et al. 2014) to underestimate the association between inequality and upward mobility.

This offsetting pattern of selection and causal effects is particularly stark when we focus on upper-tail inequality, measured by the share of households in each CZ who are in the top 1% of the national income distribution. Growing up in an area with one standard deviation greater upper-tail inequality reduces the incomes of children in low income families by 4.1%, an estimate that is significant with $p < .01$ (Online Appendix Table XI). In contrast, the effect on $\bar{y}_{25,c}$ is only 1.0% and is not statistically significant. These findings are inconsistent with Chetty et al.'s (2014, p. 1557) hypothesis that "the factors that erode the middle class may hamper intergenerational mobility more than the factors that lead to income growth in the upper tail" based on their analysis of observational data. On the contrary, both upper-tail inequality and middle-class inequality are strongly negatively associated with causal effects on upward mobility; it is just that the effect of upper-tail inequality is masked by selection in observational data.

These results shed light on the determinants of the Great Gatsby curve, the widely noted negative correlation between

inequality and intergenerational mobility (Krueger 2012; Corak 2013). The fact that inequality is negatively correlated with $\mu_{25,c}$ implies that the Great Gatsby curve is not driven by differences in genetics or other characteristics of populations in areas with different levels of inequality. Rather, placing a given child in an area with higher levels of inequality makes that child less likely to rise up in the income distribution, showing that areas with greater income inequality generate less upward mobility.⁴³

3. Education. At both the CZ- and county-level, we find strong correlations between $\mu_{25,c}$ and output-based proxies for K-12 school quality, such as test scores and high school dropout rates controlling for parent income. We find weaker correlations with input-based measures of school quality, such as class size and expenditures per student. As with the other factors analyzed above, most of the association between proxies for school quality and permanent residents' outcomes $\bar{y}_{25,c}$ is due to the causal component $\mu_{25,c}$ rather than the sorting component $\bar{\theta}_{25,c}$.

Turning to higher education, we find that CZs with more colleges per capita tend to produce better outcomes $\mu_{25,c}$, with a correlation of 0.60. As with upper-tail inequality, this association is masked when one compares permanent residents' outcomes $\bar{y}_{25,c}$ across areas because of selection effects.

4. Social Capital. Growing up in a CZ with more social capital, as measured by the social capital index of Rupasingha and Goetz (2008), improves children's outcomes significantly (correlation = 0.70). This causal component accounts for virtually all of the correlation between social capital and permanent residents' outcomes observed in prior work (e.g., Chetty et al. 2014). We also find a significant negative association between $\mu_{25,c}$ and violent crime rates across CZs (Online Appendix Table XI). At the county-within-CZ level, we do not find a significant association between $\mu_{25,c}$ and the social capital index, but continue to find a significant relationship with violent crime rates.

Together, the factors discussed here explain 58% of the variance in $\mu_{25,c}$ across CZs and 24% of the variance across counties within CZs. These results imply that the places that produce good

43. Moreover, the gap in causal effects between children from low and high-income families ($\mu_{75,c} - \mu_{25,c}$) is larger in areas with greater inequality, especially among smaller CZs. That is, areas with greater inequality also produce greater intergenerational persistence of income.

outcomes share a common set of traits, increasing the likelihood that their successes may be replicable in other areas.⁴⁴

VII.B. Individual-Level Demographic Characteristics

We now turn to a set of characteristics that are aggregates of individual-level demographics.

1. Family Structure. In observational data, the strongest predictor of differences in rates of upward mobility across CZs is the fraction of single mothers (Chetty et al. 2014a). A one standard deviation increase in the fraction of single mothers is associated with a 7.6% reduction in the incomes of children of permanent residents at $p = 25$ ($\bar{y}_{25,c}$). However, when we examine areas' causal effects on upward mobility, we find that a one standard deviation increase in the fraction of single mothers reduces a given child's income ($\mu_{25,c}$) by only 4.7%. Hence, 38% of the association between single parenthood rates and upward mobility in observational data is explained by selection. Across counties within CZs, selection accounts for nearly 70% of the association between the fraction of single mothers and $\bar{y}_{25,c}$.

Selection may play a larger role in explaining the correlation between the fraction of single mothers and $\bar{y}_{25,c}$ than factors such as school quality and social capital because the fraction of single mothers is simply an aggregation of a household-level demographic characteristic. Insofar as such characteristics have direct effects on children's outcomes, they must mechanically capture selection effects, that is, differences in the types of families living in different areas. In contrast, school quality and the other area-level factors analyzed in the previous subsection do not have such a mechanical selection component.

Despite the importance of selection, the fraction of single mothers remains a strong predictor of the causal effect $\mu_{25,c}$, with a correlation of -0.57 across CZs and -0.38 across counties within CZs. However, it is no longer the strongest predictor of differences in upward mobility, because measures of segregation, inequality, and social capital are as or more highly correlated with $\mu_{25,c}$ than the fraction of single mothers.

44. The fact that much of the variance in places' causal effects can be explained by observables is noteworthy because efforts to explain causal effects in other settings based on ex ante observables, such as teachers' value added, have been much less successful (e.g., Chetty, Friedman, and Rockoff 2014a).

2. Immigrant Shares. Another important demographic characteristic that is strongly associated with upward mobility is immigrant status. The children of certain immigrant groups, such as Asians, have higher rates of upward mobility than the children of natives, perhaps because immigrant parents tend to have lower observed incomes relative to their latent ability. Consistent with this intuition, we find a strong positive association between the fraction of immigrants in an area (measured using census data) and the sorting component ($\bar{\theta}_{25,c} = \bar{y}_{25,c} - 20\mu_{25,c}$). That is, permanent residents in areas with large immigrant populations do better than one would expect based on our estimates of the causal effects of those places, consistent with the results for Los Angeles and New York shown in [Figure I](#).

Areas with larger immigrant shares also have more negative causal effects $\mu_{25,c}$ on average (correlation = -0.45), perhaps because they have other attributes (such as higher population density or greater inequality) that are negatively associated with $\mu_{25,c}$. Because the causal and selection effects work in opposite directions, immigrant shares are not significantly associated with permanent residents' outcomes $\bar{y}_{25,c}$, matching the observational findings of [Chetty et al. \(2014\)](#). These results echo the findings for single mothers in the sense that selection plays a key role in understanding the relationship between immigrant shares and upward mobility in observational data.

3. Racial Shares. The last demographic factor we consider is race. Areas with a larger share of black residents have much lower rates of upward mobility in observational data ([Chetty et al. 2014](#)). Across CZs, a one standard deviation increase in the black share is associated with a 7.5% reduction in $\bar{y}_{25,c}$ and a 4.3% reduction in $\mu_{25,c}$. As with the fraction of single mothers, this implies that about half of the association between black shares and upward mobility across CZs in observational data is driven by factors unrelated to the causal effects of CZs. These could include other factors that cause lower rates of upward mobility for blacks than whites, such as racial discrimination in the labor market, or differences in the types of low-income white families who live in CZs with large black populations.⁴⁵

45. Lacking data on race at the individual level, we cannot distinguish between these two explanations. [Rothbaum \(2016\)](#) uses SIPP-SSA linked data to show that upward mobility varies across racial groups within CZs and that controlling for

Despite these points, the black share remains a strong predictor of the causal effect $\mu_{25,c}$, with a correlation of -0.51 across CZs and -0.32 across counties within CZs. An important implication of this result is that African American children grow up in areas that tend to produce worse economic outcomes. Under our maintained assumption that place effects are not heterogeneous by race or other characteristics, our estimates of $\mu_{25,c}$ imply that black children grow up in counties that produce 5.3% lower incomes than nonblacks on average. This suggests that residential segregation by race thus amplifies racial inequality across generations.

VII.C. Predictors of Place Effects for High-Income Families

We turn now to the characteristics of areas that produce good outcomes for children in high-income families ($\mu_{75,c}$). Because $\mu_{25,c}$ and $\mu_{75,c}$ are highly positively correlated across CZs (Table II), one might expect the strongest correlates of $\mu_{25,c}$ to be highly correlated with $\mu_{75,c}$ as well. Indeed, we find that CZs with less residential segregation, higher quality education (as measured by test scores as well as class sizes), greater social capital, and less income inequality produce better outcomes for high-income children (Figure VI, Panel A). However, $\mu_{75,c}$ is not significantly correlated with black shares and single-parent shares, perhaps because these demographic factors are more reflective of the characteristics of low-income populations in each area.

Although $\mu_{25,c}$ and $\mu_{75,c}$ are positively correlated across CZs, $\mu_{25,c}$ and $\mu_{75,c}$ are essentially uncorrelated across counties within CZs. Correspondingly, the factors that strongly predict $\mu_{25,c}$ at the county level do not predict $\mu_{75,c}$. In general, the correlations between the causal effect $\mu_{75,c}$ and the factors we examine are quite small in magnitude and statistically insignificant at the county level (Figure VI, Online Appendix Table XIV). There are, however, significant correlations between permanent residents' outcomes in high-income families $\bar{y}_{75,c}$ and county-level characteristics, which are driven primarily by selection effects. For example, a one standard deviation increase in test scores (controlling for income) is associated with a 2.1% increase in permanent residents' incomes. However, the causal effect of growing up in a county with one standard deviation higher test scores is only 0.15%, implying that

race at the individual level reduces the degree of variation in intergenerational mobility across CZs.

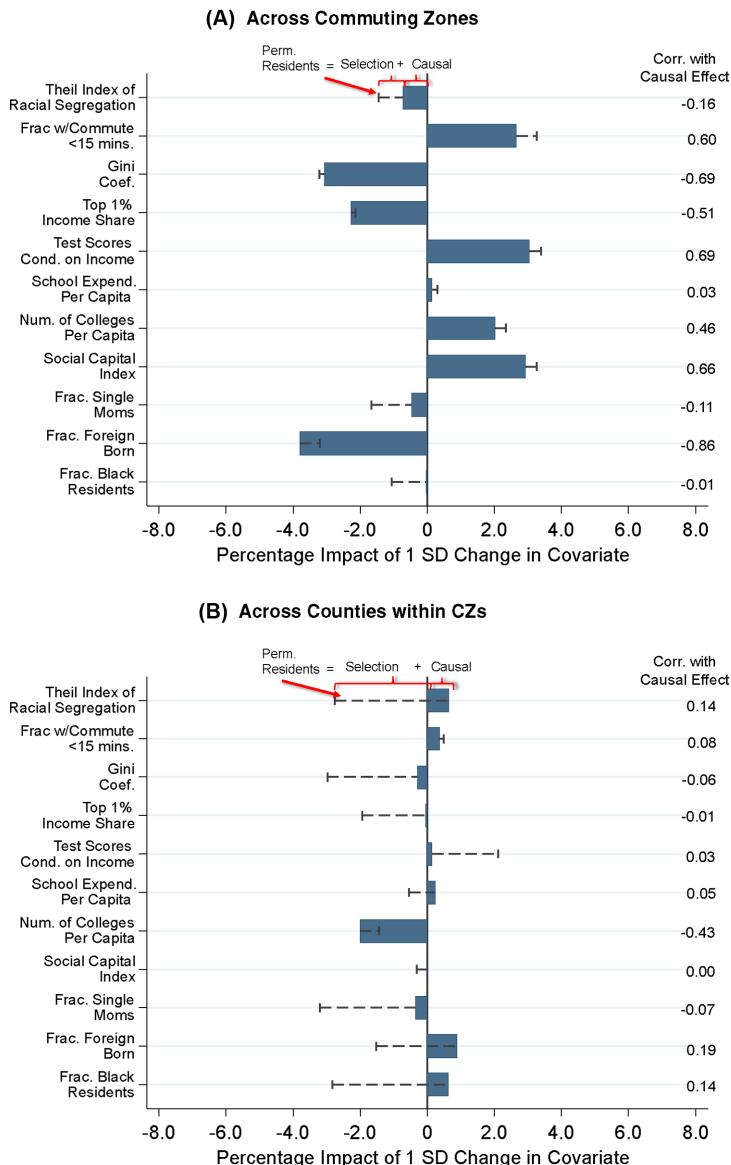


FIGURE VI

Predictors of Place Effects for Children with Parents at 75th Percentile

This figure replicates Figure V using children in above-median income families ($p = 75$) instead of below-median income families ($p = 25$); see notes to Figure V for details.

93% of the correlation observed for permanent residents is driven by selection. This finding suggests that for high-income families, places that have schools that are ostensibly of higher quality (as measured by test score performance) may not in fact produce better outcomes; they only appear to be better because they have a positively selected group of children.

Comparing the correlations at $p = 25$ and $p = 75$ shown in Figures V and VI, we see clearly that the types of areas that produce better outcomes for the poor generally produce better (or at least no worse) outcomes for the rich. Most notably, there is no evidence that more residentially integrated areas are harmful for children in high-income families. Segregation is negatively correlated with $\mu_{75,c}$ across CZs and uncorrelated with $\mu_{75,c}$ across counties within CZs.

VIII. HOUSING COSTS AND OPPORTUNITY BARGAINS

How much more does a family have to pay to live in an area that produces better outcomes for their children? In this section, we examine how opportunity for children is priced in the housing market.

VIII.A. Methods

Letting r_{pc} denote the average rent paid by families at percentile p in area c , we characterize the relationship between r_{pc} and μ_{pc} in two ways. First, we measure how much it costs on average to live in a place that produces higher incomes for children by estimating the conditional expectation of rents given an area's causal effect, $E[r_{pc}|\mu_{pc}]$. Second, we measure the extent to which a family could find a place that produces better outcomes for their children without paying more rent by estimating the variance in rents explained by places' causal effects, $R^2 = \frac{Var(E[r_{pc}|\mu_{pc}])}{Var(r_{pc})}$.⁴⁶

If areas' causal effects μ_{pc} were directly observed, these parameters could be estimated (under a linear approximation) using

46. We use rents in our baseline specifications instead of house prices because most low-income families rent and it is more straightforward to compare income gains for children to flow rental costs than house prices. We find qualitatively similar results using house prices, with a negative correlation between house prices and μ_{pc} across CZs and very small correlations across counties within CZs ([Online Appendix](#) Tables XI and XII).

the following OLS regression:

$$(13) \quad r_{pc} = \alpha + \beta_p \mu_{pc} + \zeta_{pc}.$$

Since μ_{pc} is not observed, we estimate the conditional expectation β_p by replacing μ_{pc} with $\mu_{pc}^r = \frac{\sigma_{\mu_{pc}}^2}{\sigma_{\mu_{pc}}^2 + s_{pc}^2} \hat{\mu}_{pc}$, the MSE-minimizing forecast of each place's causal effect using data purely on movers.⁴⁷ This regression yields an unbiased estimate of β_p under the identification assumption in equation (3).⁴⁸ Intuitively, the shrinkage factor $\frac{\sigma_{\mu_{pc}}^2}{\sigma_{\mu_{pc}}^2 + s_{pc}^2}$ adjusts for the fact that the raw causal effects $\hat{\mu}_{pc}$ are noisy estimates of μ_{pc} , leading to attenuation bias in β_p in a regression of r_{pc} on $\hat{\mu}_{pc}$.

Similarly, we estimate the variance in rents explained by μ_{pc} (R^2) using

$$R = \text{Corr}(r_{pc}, \mu_{pc}) = \frac{\text{Cov}(r_{pc}, \mu_{pc})}{\text{SD}(r_{pc})\text{SD}(\mu_{pc})} = \frac{\text{Cov}(r_{pc}, \hat{\mu}_{pc})}{\text{SD}(r_{pc})\text{SD}(\hat{\mu}_{pc})} \frac{\text{SD}(\hat{\mu}_{pc})}{\text{SD}(\mu_{pc})},$$

where $\frac{\text{SD}(\hat{\mu}_{pc})}{\text{SD}(\mu_{pc})}$ is computed using the total and signal standard deviations reported in Table II. Intuitively, the signal R^2 can be computed from the correlation between rents and $\hat{\mu}_{pc}$, again adjusting for attenuation due to noise in the causal effect estimates.

We scale μ_{pc}^r in terms of the percentage change in income per year of childhood exposure, as in Section V.B. We measure monthly rents using data from the 2000 census, defining the rent in each CZ or county as the mean of the median rent in each census tract, weighting by the number of families with children who have

47. When estimating equation (13) at the CZ level, we construct the shrinkage factor $\frac{\sigma_{\mu_{pc}}^2}{\sigma_{\mu_{pc}}^2 + s_{pc}^2}$ using the estimates of $\sigma_{\mu_{pc}}$ of 0.13 at $p = 25$ and 0.11 at $p = 75$ (Table II, columns (1) and (2)). When estimating equation (13) at the county-within-CZ level, we construct the shrinkage factor using the county-within-CZ signal standard deviation of 0.10 at $p = 25$ and 0.11 at $p = 75$ (Table II, columns (5) and (6)).

48. Formally, $\frac{\text{Cov}(r_{pc}, \mu_{pc}^r)}{\text{Var}(r_{pc}^r)} = \beta$ because $\frac{\text{Cov}(\mu_{pc}, \mu_{pc}^r)}{\text{Var}(\mu_{pc}^r)} = 1$ and $\text{Cov}(\zeta_{pc}, \mu_{pc}^r) = 0$. Unlike in Section VI, we do not use data on permanent residents when constructing the forecasts here because rents may be correlated with the selection component of permanent resident outcomes $\bar{\theta}_{pc}$, leading to a biased estimate of β_p . Indeed, we find that the association between rents and $\bar{y}_{25,c}$ is larger than the association between rents and $\hat{\mu}_{25,c}$ (Online Appendix Table XI), implying that families with positive unobservables tend to select into high-priced areas.

TABLE V

ASSOCIATION BETWEEN RENTS AND PLACES' CAUSAL EFFECTS ON CHILDREN'S INCOMES FOR LOW-INCOME FAMILIES

	Dep. var.: Mean monthly rent for low-income families (\$)				
	Counties within CZs				
	CZs	All counties	Counties in 100 largest CZs	Observable component	Unobservable component
	(1)	(2)	(3)	(4)	(5)
Causal effect (1% increase in child's income)	-102.9 (82.91)	176.8** (65.50)	202.4** (64.93)		
Observable component				430.2*** (27.29)	
Unobservable component					46.18 (42.65)
Mean of dependent variable	797.02	811.49	884.67	876.94	876.94
Signal <i>R</i> -squared	0.180	0.013	0.051	0.301	0.000
Number of observations	595	2,367	694	673	673

Notes. This table shows estimates from OLS regressions of average monthly rents on areas' causal effects for children in below-median income ($p = 25$) families. Standard errors are shown in parentheses. Rents are measured using data from the 2000 census (in 2012 dollars), and are defined as the mean of median rents across census tracts, weighting by the number of below-median income families with children. Causal effects are MSE-minimizing forecasts constructed purely using data on movers (excluding permanent residents) and scaled in terms of percentage changes in children's incomes in adulthood per year of childhood exposure to a given CZ or county. All regressions are weighted by population based on the 2000 Census. In column (1), we regress rents on causal effects across CZs; the coefficient implies that CZs that generate 1% higher earnings for a single year of exposure have \$102.9 lower monthly rents on average. Columns (2)–(5) report estimates of regressions at the county level, including CZ fixed effects. Column (2) includes all CZs, while columns (3)–(5) restrict the sample to the 100 most highly populated CZs (with populations above 590,000). In column (4), the independent variable is the "observable" component of causal effects, the predicted values from a regression of the raw fixed effect estimates on the following area-level characteristics: the fraction of African American residents, the Theil index of racial segregation, the Gini index, the fraction of single parents, the social capital index, and expenditures on public schools per student. In column (5), the independent variable is the residual from the preceding regression, shrunk based on the signal-to-noise ratio as described in the text to account for sampling error. In each column, we also report the mean of the dependent variable (monthly rent for low-income families) and the signal *R*-squared, constructed as the square of the correlation between rents and the right-hand-side variable (after removing CZ fixed effects in columns (2)–(5)). The *R*-squared estimates in columns (1)–(3) and (5) are adjusted for noise using the signal to noise ratio as described in the text.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

below-median income for $p = 25$ and above-median income for $p = 75$ (see [Online Appendix](#) Table XV). We weight all regressions and correlations by population in the 2000 census.

VIII.B. Relationship Between Rents and Children's Outcomes

Column (1) of [Table V](#) reports estimates of [equation \(13\)](#) at the CZ level at $p = 25$; results at $p = 75$ are qualitatively similar and are presented in [Online Appendix](#) Table XVI. CZs that produce 1% higher incomes for children have \$103 lower monthly rents on average. This is consistent with our finding that rural areas, such as the Great Plains, tend to produce better outcomes than urban

areas. Of course, rural areas also tend to have fewer job opportunities and lower wage rates, so moving to such CZs may not be a plausible margin of choice for most families. Therefore, in the remainder of this section, we focus on variation across counties within a given CZ. Because CZs are constructed to approximate local labor markets (Tolbert and Sizer 1996), a household's location decision within a CZ aligns more closely with the conceptual exercise of determining the price of better outcomes for children while holding parents' job opportunities and wage rates fixed.

In column (2), we estimate the relationship between rents and children's outcomes across counties within a CZ by estimating equation (13) at the county level, including CZ fixed effects. On average, moving to a county that produces 1% higher income a year of exposure for children costs \$177 more in monthly rent. To interpret the magnitude of this coefficient, note that a 1% increase in income translates to approximately a \$4,900 increase in lifetime income for a child with parents at $p = 25$ in present value at age 10 (the middle of childhood) using a 3% discount rate (see Online Appendix D). Hence, a family with two children stands to gain approximately \$10,000 in the present value of future income by moving to a county that produces 1% better outcomes. This is four times larger than the $12 \times \$177 = \$2,124$ mean increase in annual rent associated with moving to a county that increases children's incomes by 1%.

The fraction of the variance in rents explained by $\mu_{25,c}$ (or, equivalently, the fraction of the variance in $\mu_{25,c}$ explained by rents) is $R^2 < 2\%$ across counties within a CZ. Even among the 100 most highly populated CZs (population $> 590,000$), where housing supply is likely to be most constrained, $\beta_{25} = \$202$ and $R^2 = 5\%$ (column (3) of Table V).⁴⁹

The relatively weak relationship between rents and children's outcomes suggests that policies that encourage low-income families to move to more expensive areas may not be sufficient to improve their children's outcomes. For example, current "small area" fair market rent proposals in housing voucher programs vary

49. Among these large CZs, the rent gradient is steeper in more sprawling, residentially segregated areas. In CZs with above-median commute times, $\beta_{25} = \$255$, compared with $\beta_{25} = -\$66$ in CZs with below-median commute times. The greater price of access to high-opportunity neighborhoods could potentially explain why more segregated, sprawling cities tend to generate worse outcomes for children in low-income families.

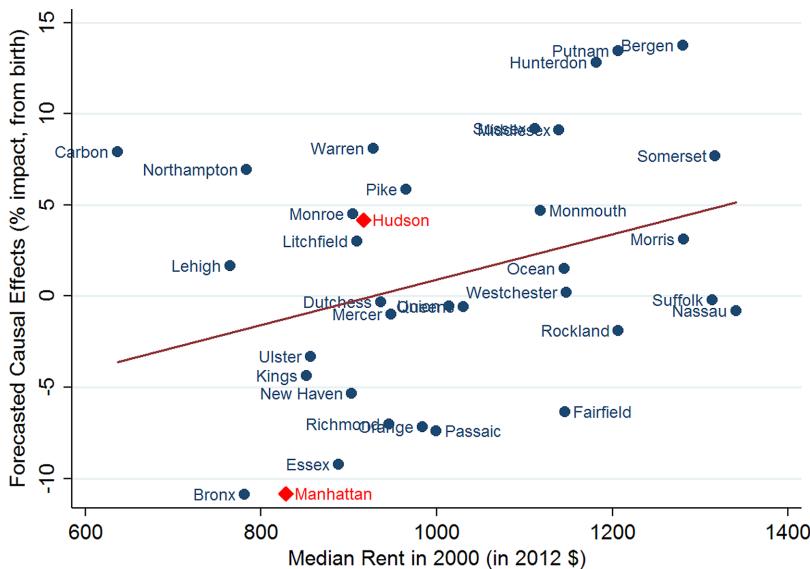


FIGURE VII
Opportunity Bargains in the New York Area

This figure plots the MSE-minimizing forecasts of causal effects, $\mu_{25,c}^f$, for counties in the New York combined statistical area (shown in Figure IV, Panel A) versus average monthly rents for low-income families. The forecasted causal effects $\mu_{25,c}^f$ are scaled to show the percentage change in income from growing up from birth (i.e., 20 years of childhood exposure) in a given county relative to the average county in the country. See notes to Figure IV, Panel A for further details on construction of $\mu_{25,c}^f$. Monthly rents are measured using data from the 2000 census (in 2012 dollars) and are defined as the mean of median rents across census tracts, weighting by the number of families with children who have below-median income. The solid line shows the best-fit line obtained from regressing $\mu_{25,c}^f$ on rents, weighting by county population. Points above the line are “opportunity bargains”—counties that produce particularly good outcomes for children relative to other counties with comparable rents.

voucher payments with neighborhoods' rents rather than their impacts on economic outcomes. Such policies may not maximize vouchers' effects on upward mobility, because many expensive areas do not have high levels of $\mu_{25,c}^f$. Indeed, our results imply that there are many “opportunity bargains”—counties within a labor market that offer good outcomes for children with relatively low rents.

Figure VII illustrates these findings by plotting our optimal forecasts $\mu_{25,c}^f$ (scaled as the percentage impact of growing up in a

given county from birth relative to the national mean) versus rents for counties in the New York City CSA. The substantial dispersion in this scatter plot around the mildly upward-sloping best-fit line illustrates the key empirical results of this section: on average, moving to a better area does not cost much more, and there is considerable residual variation in children's outcomes that is orthogonal to rents. For example, Hudson County offers much better outcomes for low-income children than does Manhattan, despite having comparable rents. More generally, there are many counties—in the upper left of the figure—that offer “opportunity bargains” within the New York area.

The existence of these opportunity bargains may be encouraging for families seeking to improve their children's prospects for upward mobility as well as for policy makers looking for affordable areas that produce good outcomes as models to emulate. However, the existence of such areas demands further explanation from the perspective of economic models of spatial equilibrium, as these areas seemingly offer arbitrage opportunities that have been left unexploited. We turn to explaining these empirical results in the next subsection.

VIII.C. Explaining Opportunity Bargains

Why are places' effects on children's future incomes not capitalized more fully into housing prices? We discuss three potential explanations for this result in this subsection.

First, the areas that appear to be opportunity bargains may have other disamenities that make them less desirable places to live. Our data suggest that such disamenities explain at least part of the residual variation in children's opportunities conditional on rents. The areas that generate the best outcomes for children in most cities are typically in the suburbs—which tend to have longer commutes and offer fewer urban amenities—as illustrated by the maps for New York and Boston in [Figure IV](#). However, data from the Moving to Opportunity (MTO) experiment suggest that such disamenities do not fully explain the existence of opportunity bargains. MTO families who received housing vouchers to move to low-poverty census tracts exhibited gains in both children's long-term outcomes and parents' subjective well-being and neighborhood satisfaction ([Ludwig et al. 2012; Chetty, Hendren, and Katz 2016](#)). These findings suggest that affordable neighborhoods that produce better outcomes for children in low-income

families are not necessarily less desirable to parents because of other disamenities.

A second potential explanation for why children's opportunities are not fully priced in the housing market is a lack of information. Families may not know which areas produce the best outcomes for children, particularly because there is a long delay in observing these outcomes. To evaluate this hypothesis, we divide place effects $\mu_{25,c}$ into an "observable" component that is related to attributes families can observe, such as school quality and poverty rates, and an "unobservable" component unrelated to these factors. We estimate these components by regressing $\hat{\mu}_{25,c}$ (at the county level) on the fraction of African American residents, the Theil index of racial segregation, the Gini index, the fraction of single mothers, the social capital index, and expenditures on public schools per student.⁵⁰ We then define the observable component as the predicted value from this regression and the unobservable component as the residual, which we shrink by its signal-to-noise ratio as above to adjust for attenuation bias. Thirty percent of the signal variance in $\mu_{25,c}$ is captured by the observable component; the remaining 70% is "unobservable" given these predictors.

In columns (4) and (5) of [Table V](#), we replicate the specification in column (3), regressing rents on the observed and unobserved components of place effects. We obtain an estimate of $\beta_{25} = \$430$ for the observed component. Hence, moving to a CZ that generates a 1% ($\$4,900$) increase in lifetime income for a child along the "observed" dimension—for instance, an area with better-funded schools and more two-parent families—costs $\$5,160$ in terms of annual rent on average. In contrast, there is no significant relationship between prices and the "unobservable" component (column (5)). These findings suggest that the causal effects $\mu_{25,c}$ may be underpriced partly because of a lack of information.⁵¹

50. We use these factors because they are highly predictive of children's outcomes, as shown in [Figure V](#), and are in principle easily observed by families. However, including additional variables from the set in [Figure V](#) does not affect the results. We estimate the regression at the county level with CZ fixed effects, restricting the sample to counties in the 100 most highly populated CZs and weighting by the county population.

51. One may be concerned that the unobservable component is highly transitory, in which case "opportunity bargains" could only be identified ex post after children become adults rather than when families are deciding where to live. Empirically, this does not turn out to be the case. The unobservable component has a precision-weighted signal correlation of 0.42 with permanent residents' mean

A third explanation for the existence of opportunity bargains is failures in optimization due to cognitive constraints or behavioral biases. Recent ethnographic studies show that low-income families frequently move under time pressure, either because they have been evicted (Desmond 2016) or because they are making a “reactive” move responding to a financial or health shock (DeLuca, Rosenblatt, and Wood 2013). In such circumstances, families often seek shelter as quickly as possible rather than weighing the benefits that may accrue to their children several years later if they choose a different neighborhood (DeLuca, Rosenblatt, and Wood 2013). Moreover, the decision of where to live has several features that may trigger well-established behavioral biases and induce suboptimal choice: delayed payoffs coupled with large initial up-front costs that compound present bias, a need to predict one’s preferences in a very different environment that may induce projection bias, and complex planning with scarce mental bandwidth (Chetty 2015, section IV.B).

We believe that these explanations—disamenities, a lack of information, and behavioral biases—are likely to play a role in explaining the existence of opportunity bargains. Understanding the relative importance of these theories is an important area for future research.

IX. CONCLUSION

This article has estimated the causal effect of each county in the United States on children’s outcomes in adulthood. Overall, the findings provide support for place-focused approaches to improving economic opportunity, by helping families move to opportunity and through place-based investments. The estimates show that there is substantial scope for households to move to areas within their labor market (CZ) that are opportunity bargains—places that produce better outcomes for children without paying higher rents. In addition, the areas that produce high levels of upward mobility share a common set of characteristics, such as less residential segregation and greater social capital, suggesting that their successes might be replicable in other areas.

ranks at $p = 25$. Permanent residents’ mean ranks in turn have a serial correlation exceeding 0.93 across cohorts, implying that one could reliably predict the unobservable component even while children are growing up based on observed outcomes for earlier cohorts.

There are two key areas for further research before one can apply these findings to make policy changes that improve children's outcomes. First, it would be useful to estimate places' causal effects at narrower geographies (e.g., census tracts) and for specific subgroups (e.g., by race and ethnicity) using the methods developed here. Such estimates would provide more granular data for families seeking to move to opportunity within their cities and for policy makers seeking to make targeted investments in neighborhoods that currently produce lower levels of upward mobility.

Second, it would be useful to understand the mechanisms through which some places produce better outcomes than others by isolating exogenous variation in the predictors of upward mobility identified here. For example, studying changes in local policies that have reduced residential segregation could shed light on whether segregation directly harms children in low-income families. To facilitate further investigation of these mechanisms, we have made all of the county- and CZ-level estimates of causal effects constructed in this study available on the Equality of Opportunity Project website.⁵²

STANFORD UNIVERSITY AND NATIONAL BUREAU OF ECONOMIC RESEARCH

HARVARD UNIVERSITY AND NATIONAL BUREAU OF ECONOMIC RESEARCH

SUPPLEMENTARY MATERIAL

An Online Appendix for this article can be found at *The Quarterly Journal of Economics online*. Code used to generate tables and figures in this article can be found in Chetty and Hendren (2018b), in the Harvard Dataverse, doi:10.7910/DVN/CEMFTJ.

52. In addition to the estimates discussed in this article, we provide estimates for other outcomes and subgroups, such as college attendance rates and estimates for children in one- versus two-parent households. For all outcomes, we provide estimates of both the raw fixed effect estimates $\hat{\mu}_{pc}$ and optimal forecasts μ_{pc}^f , as the appropriate measure of place effects will vary across applications. As a rule of thumb, those seeking to use the causal effect as a dependent variable should use $\hat{\mu}_{pc}$, while those seeking to use the causal effect as an independent variable should use μ_{pc}^f .

REFERENCES

- Abadie, Alberto, and Maximilian Kasy, "The Risk of Machine Learning," Working Paper, 2017.
- Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 67 (1999), 251–333.
- Angrist, Joshua D., Peter D. Hull, Parag A. Pathak, and Christopher R. Walters, "Leveraging Lotteries for School Value-Added: Testing and Estimation," *Quarterly Journal of Economics*, 132 (2017), 871–919.
- Chetty, Raj, "Behavioral Economics and Public Policy: A Pragmatic Perspective," *American Economic Review Papers and Proceedings*, 105 (2015), 1–33.
- Chetty, Raj, John N. Friedman, and Jonah E. Rockoff, "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates," *American Economic Review*, 104 (2014a), 2593–2632.
- , "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood," *American Economic Review*, 104 (2014b), 2633–2679.
- Chetty, Raj, and Nathaniel Hendren, "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects," *Quarterly Journal of Economics*, 133 (2018a), 1107–1162.
- , "Replication Code for: 'The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects' and 'The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates,'" 2018b. Harvard Dataverse, doi:10.7910/DVN/CEMFTJ.
- Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz, "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," *American Economic Review*, 106 (2016), 855–902.
- Chetty, R., N. Hendren, P. Kline, and E. Saez, "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States," *Quarterly Journal of Economics*, 129 (2014), 1553–1623.
- Chetty, Raj, Nathaniel Hendren, Frina Lin, Jeremy Majerovitz, and Benjamin Scuderi, "Childhood Environment and Gender Gaps in Adulthood," *American Economic Review Papers and Proceedings*, 106 (2016), 282–288.
- Corak, Miles, "Income Inequality, Equality of Opportunity, and Intergenerational Mobility," *Journal of Economic Perspectives*, 27 (2013), 79–102.
- Cutler, David M., and Edward L. Glaeser, "Are Ghettos Good or Bad?," *Quarterly Journal of Economics*, 112 (1997), 827–872.
- DeLuca, Stefanie, Peter Rosenblatt, and Holly Wood, "Why Poor People Move (and Where They Go): Residential Mobility, Selection and Stratification," Working Paper, 2013.
- Desmond, Matthew. *Evicted: Poverty and Profit in the American City* (New York: Crown, 2016).
- Haider, Steven, and Gary Solon, "Life-Cycle Variation in the Association between Current and Lifetime Earnings," *American Economic Review*, 96 (2006), 1308–1320.
- Hilger, Nathaniel, "Upward Mobility and Discrimination: The Case of Asian Americans," NBER Working Paper no. 22748, 2016.
- Hull, Peter, "Estimating Hospital Quality with Quasi-experimental Data," Working Paper, 2017.
- Kain, John F., "Housing Segregation, Negro Employment, and Metropolitan Decentralization," *Quarterly Journal of Economics*, 82 (1968), 175–197.
- Kasarda, John D., "Urban Industrial Transition and the Underclass," *Annals of the American Academy of Political and Social Science*, 501 (1989), 26–47.
- Krueger, Alan, "The Rise and Consequences of Inequality in the United States." Speech at the Center for American Progress, Washington, DC, January 12, 2012.
- Ludwig, Jens, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, and Lisa Sanbonmatsu, "Neighborhood Effects on

- the Long-Term Well-Being of Low-Income Adults," *Science*, 337 (2012), 1505–1510.
- Massey, Douglas S., and Nancy A. Denton, *American Apartheid: Segregation and the Making of the Underclass* (Cambridge, MA: Harvard University Press, 1993).
- Robbins, Herbert, "An Empirical Bayes Approach to Statistics," in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics* (Berkeley: University of California Press, 1956), 157–163.
- Rosenbaum, James E., "Changing the Geography of Opportunity by Expanding Residential Choice: Lessons from the Gautreaux Program," *Housing Policy Debate*, 6 (1995), 231–269.
- Rothbaum, Jonathan, "Sorting and Geographic Variation in Intergenerational Mobility," SEHSD Working Paper, 2016.
- Rupasingha, Anil, and Stephan J. Goetz, "US County-Level Social Capital Data, 1990–2005," Northeast Regional Center for Rural Development, 2008.
- Sampson, Robert J., *Great American City: Chicago and the Enduring Neighborhood Effect* (Chicago: University of Chicago Press, 2012).
- Sampson, Robert J., Jeffrey D. Morenoff, and Thomas Gannon-Rowley, "Assessing Neighborhood Effects: Social Processes and New Directions in Research," *Annual Review of Sociology*, 28 (2002), 443–478.
- Solon, Gary, "Intergenerational Mobility in the Labor Market," in *Handbook of Labor Economics*, volume 3, O. Ashenfelter and D. Card, eds. (Amsterdam: Elsevier, 1999), 1761–1800.
- Solon, Gary, Marianne E. Page, and Greg J. Duncan, "Correlations between Neighboring Children in Their Subsequent Educational Attainment," *Review of Economics and Statistics*, 82 (2000), 383–392.
- Tolbert, Charles M., and Molly Sizer, "U.S. Commuting Zones and Labor Market Areas: A 1990 update," *Economic Research Service Staff Paper*, 9614 (1996).
- Wilson, William J., *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy* (Chicago: University of Chicago Press, 1987).
- Wilson, William J., *When Work Disappears: The World of the New Urban Poor* (New York: Knopf, 1996).