Journal of Economic Geography **0** (2018) pp. 1–30 doi:10.1093/jeg/lby026

The geography of family differences and intergenerational mobility

*Ryan Gallagher**, *Robert Kaestner***,† *and Joseph Persky***

*Department of Economics, Northeastern Illinois University, 5500 North St. Louis Avenue, Chicago, IL 60625, USA
**Department of Economics, University of Illinois at Chicago, 601 S. Morgan St, Chicago, IL 60607, USA
†Correspondence to: *email* <kaestner.robert@gmail.com>

Abstract

A recent series of studies by the Equality of Opportunity Project has documented substantial geographical differences in intergenerational income mobility. These spatial differences are important because they suggest that place matters more than previously thought in determining economic well-being. In this article, we show that family characteristics vary widely across areas and simulations indicate that differences in these family characteristics can explain a substantial share of the variation in intergenerational income mobility across places documented by the Equality of the Opportunity Project. Additionally, we show that the characteristics of families that move differ substantially from families that do not move and that family characteristics differ by the type of move made, which raise questions about the external and internal validity of causal inferences based on the Equality of Opportunity Project's analysis of movers.

Keywords: Income inequality, migration, economic opportunity, intergenerational mobility
JEL classifications: I31, R23, R11
Date submitted: 11 May 2017 **Editorial decision:** 21 April 2018 **Date accepted:** 25 April 2018

1. Introduction

The USA is an incredibly diverse country consisting of a large number of places with distinctive physical characteristics, varied populations and different economic circumstances. A recent groundbreaking study by Chetty et al. (2014) has added to this list of differences. That study documents previously unknown, large geographical differences in intergenerational income mobility. For example, the authors reported that the income of a 30-year-old person from a low-income family who grew up in Cook County, IL (Chicago) is nearly 30% ($7, 420) lower than that of a person of the same age from a similarly low-income family who grew up in neighboring DuPage County. The present value of this future income difference is substantial—$167,000—assuming 40 years of working life and a 3% discount rate.

The large geographic differences in intergenerational income mobility documented by Chetty et al. (2014) are important because they raise the possibility that places, independently of the people that live there, matter in determining economic well-being. It is likely that the findings reported in Chetty et al. (2014) will become one of the key facts in the longstanding 'people versus place' debate in economic development (Kain and Persky, 1969; Bartik, 1991; Galster and Killen, 1995; Bartik, 2003; Kline and Moretti, 2014).

In a companion study, Chetty and Hendren (2017a, 2017b) extend the research of Chetty et al. (2014) to obtain estimates of the 'causal' effect of county (commuting zone) of childhood residence on intergenerational income mobility. This study used a novel approach that focused on families that move to obtain causal estimates. Chetty and Hendren (2017a) find that children who moved to a place with greater intergenerational income mobility have higher adult incomes than children (siblings) who did not move. According to Chetty and Hendren (2017b), if a child spends 20 years in a place with one standard deviation higher rank of income mobility, their earnings at age 26 will be 10% higher.

In this article, we contribute to the novel and important line of research begun by Chetty et al. (2014) and Chetty and Hendren (2017a, 2017b) on the influence of county (commuting zones) of childhood residence on adult incomes. We address two issues. The first is whether there are differences in characteristics of low-income families by place that may confound estimates of differences in intergenerational income mobility by place, and, if so, how large is the potential confounding. The second issue is the validity of the analysis of Chetty and Hendren (2017a, 2017b), which depends on whether there are differences between movers and nonmovers and within-movers that are potential threats to the internal and external validity of the findings of Chetty and Hendren (2017a, 2017b).

To accomplish our objectives, we use Census data to provide direct evidence of differences in characteristics of low-income families across places. This descriptive information is useful, if not essential, for assessing the potential confounding of family characteristics and place in explaining intergenerational income mobility. Then, using these measured differences in family characteristics across places, we simulate differences in expected intergenerational income mobility across places. The results of these simulations measure the extent to which differences in low-income family characteristics can explain place differences in mobility. A similar descriptive analysis is performed for differences in family characteristics between low-income movers and stayers, and across low-income families making different types of moves. Again these data are used to simulate expected intergeneration income mobility of movers and nonmovers by county.

We find that there are large differences in family characteristics (holding income constant) across places and that these differences are significantly correlated with differences in intergenerational income mobility. Simulations indicate that differences in a relatively small set of family characteristics across places can explain a substantial share of the variation in intergenerational income mobility across places documented by Chetty et al. (2014). For example, we find that differences in the income of adult children due to differences in mother's race, age, education, marital status and nativity explain 83–123% of the difference in intergenerational income mobility between places in the lowest and next lowest quintiles of absolute mobility, as predicted by Chetty et al.'s (2014) place-based estimates. The same limited set of characteristics explains 37–56% of the difference in intergenerational income mobility between the lowest and highest quintiles of absolute mobility in Chetty et al.'s (2014) place-based distribution of intergenerational income mobility.

We also find that there are substantial differences in family characteristics of movers and stayers. Whether based on a comparison with families in the origin or destination locations, families that move are more likely to have mothers who are more educated, married, white and younger than mothers of families that do not move. In addition, families that move are a more homogenous group, compared to families that choose not to move. Therefore, differences in family characteristics of movers explain much less of the differences in intergenerational income mobility across places (for a sample of movers). This is consistent with findings in Chetty and Hendren (2017a, 2017b), which uses a sample of movers, that suggests a causal effect of place on intergenerational income mobility.

However, the significant differences between families that do and do not move imply that estimates of place-based differences derived from movers are unlikely to generalize to most families because the vast majority of them are nonmovers. Moreover, we find that among movers, family characteristics that serve as favorable predictors of a child's future income, including time-varying characteristics such as home ownership and the number and ages of children, differ considerably by the type of move made, for example, from the lowest to highest quintiles of intergenerational income mobility. This finding is consistent with a behavioral model in which moves are based on idiosyncratic (unobserved) costs and benefits of moving, which further strengthens the likelihood that findings from an analysis of movers are unlikely to be broadly applicable (Heckman et al., 2006). Differences in time-varying characteristics (e.g., home ownership) among families by the type of move they make, for example, from places with the lowest intergenerational income mobility to places with the highest mobility, also raises the possibility that moving may be endogenous and estimates of the effect of place confounded by changing family characteristics. Altogether, this suggests the nontrivial likelihood that estimates of the effect of place on intergenerational income mobility using movers may not be internally valid given that movers making 'better' moves also tend to have more favorable predictors of future income.

The Equality of Opportunity Project has added novel and significant evidence to an exciting field of investigation. The purpose of this article is to further qualify the results and conclusions from this project with respect to 'people characteristics' that play into the process of upward mobility in a range of ways. None of what follows is meant to suggest that place does not matter. However, substantive reform cannot ignore the importance of these 'people characteristics' for the present-day realities of upward mobility.

2. What the Equality of Opportunity Project has shown

While the Chetty et al. (2014) study is innovative, it remains a descriptive analysis. The variation in intergenerational mobility documented therein does not reflect causality and is potentially confounded by differences between the families that live in these places. Chetty et al. (2014) were aware of the potential confounding issue:

> . . .[O]ur descriptive analysis does not shed light on whether the differences in outcomes across areas are due to the causal effect of neighborhoods or differences in the characteristics of people living in those neighborhoods.' (1559)

Unfortunately, their use of income tax records with little information on family-level characteristics does not allow them to conduct a meaningful analysis of the extent to which family-level characteristics possibly confound their estimates.

Chetty et al. (2014) did assess, in a limited way, the extent to which differences in the racial composition of families could explain differences in intergenerational income mobility between places. Specifically, Chetty et al. (2014) calculated intergenerational income mobility using their entire sample and then again using a sample consisting largely of non-Hispanic whites. The correlation between the two measures of intergenerational income mobility by place was quite high—0.91. This result is not surprising, however, because, as they report, non-Hispanic whites make up 68% of the entire sample. Therefore, the intergenerational income mobility of the limited sample would mechanically be highly correlated with the intergenerational income mobility of the full sample. In addition, for this particular analysis, approximately 20% of places (e.g., counties) were dropped presumably because they had no zip codes where at least 80% of the residents were non-Hispanic white, which was one of the criteria used to select the sample. In other words, places with relatively high concentrations of non-white (non-Hispanic white) people were omitted. This approach to assessing whether race is a confounding influence is quite indirect and does not rule out the possibility that the racial composition of families living in different places accounts for a nontrivial fraction of the geographic variation in intergenerational income mobility. In fact, Chetty et al. (2014) recommend undertaking the analysis that we conduct in this article: 'To distinguish between these two channels, we would ideally control for race at the individual level, essentially asking whether whites have lower rates of upward mobility in areas with a larger black population' (Chetty et al., 2014, 1605).

Using their tax data, Chetty et al. (2014) are able to explicitly examine one instance in which individual differences can confound place differences in intergenerational income mobility. This was the case of differences in family structure. Here the evidence suggests strongly that the place-based estimates of income mobility may be significantly confounded by family-level differences. Specifically, when intergenerational income mobility is recalculated using only children who grew up in two-parent families, the correlation between this measure of intergenerational income mobility by place and the baseline measure that used the entire sample was only 0.66. While still relatively large, the correlation between the two measures of intergenerational income mobility is far from perfect and indicative of a substantial amount of confounding of place effects by family characteristics.[1]

In a companion study, Chetty and Hendren (2017a, 2017b) address the issue of whether individual-level characteristics confound place effects of intergenerational income mobility by focusing on families that move. The motivation for this analysis is straightforward. If place matters, then moving to a place with greater income mobility should improve children's income mobility relative to those children who do not move, and the improvement should be larger if the longer the child spent in the better place. Indeed, this is exactly what Chetty and Hendren (2017a, 2017b) find—every year living in a place with 1 percentile higher intergenerational income mobility rank increases the

1 It should be noted that Chetty et al. (2014, 1604) did conduct an analysis that estimated associations between intergenerational income mobility in an area and several area-wide aggregate characteristics including racial composition and family structure. These estimates suggest that these family differences matter, and in fact, the fraction of an area's families headed by a single mother explained the most variation among the several variables examined. The other variables examined were commuting patterns, income inequality (Gini index), high school dropout rate and social capital index. In our analysis of the effect of household characteristics on children's incomes we control for such place variables.

child's rank in the income distribution by 0.04 percentage points. This result holds whether the comparison is to children in other families who moved at different ages or a comparison of children in the same family who were different ages when the family moved.

The Chetty and Hendren (2017a, 2017b) analysis provides credible evidence that places exert a causal effect on children's later life outcomes. Yet it is limited by a potential lack of external validity. A few pieces of evidence are relevant. First, as Chetty and Hendren (2017a) report, movers are different from stayers. Families that move have incomes that are approximately 12% higher than nonmovers. Secondly, out of 16.5 million possible movers, Chetty and Hendren (2017a) use only a small fraction (1.6 million, or 9%) of potential families selecting those that moved once and stayed in the same place for at least 2 years. Thirdly, the return to moving (the convergence of origin outcome to destination outcome) differs by the distance of the move and the number of moves, which suggests strongly that the types of families differ by the distance of the move and number of moves. In short, there appears to be considerable heterogeneity among movers, which suggests that the analysis limited to a fraction of movers may lack external validity even among movers. More importantly, there is a substantial literature on internal migration in the USA that demonstrates that movers and stayers differ significantly (Sjaastad, 1962; Greenwood, 1969, 1997; Mincer, 1978; Crowder and South, 2005; Kling et al., 2007; South et al., 2011; Molloy et al., 2011). This point is recognized by Chetty and Hendren (2017a):

> Our estimates of neighborhood exposure effects are based on households who choose to move to certain areas. The effects of moving a randomly selected household to a new area may differ, since households that choose to move to a given area may be more likely to benefit from that move. (6)[2]

Despite this important caveat, Chetty and Hendren (2017b) state:

> In the second part of the paper, 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. (3)

The claim that estimates obtained in the analysis can guide families seeking to move implies broad external validity that seems somewhat speculative given the authors acknowledgement of the potential lack of external validity and the points we noted earlier about differences between movers and stayers. Accordingly, the use of the results from the mover analysis to provide causal estimates of the effect of place more generally, as Chetty and Hendren (2017a, 2017b) do, is arguably going beyond the evidence of the study.

The remainder of the article is not meant to directly refute the conclusions from the analyses of the Equality of Opportunity Project. Rather it aims at plausibly establishing

2 The quote in the text is in a footnote in Chetty and Hendren (2017a). In an earlier version of the paper, the authors wrote something similar, but included it in the text: 'An important caveat in interpreting this estimate is that it is a local average treatment effect estimated based on house-holds who choose to move to certain areas. The mean exposure effect of moving a randomly selected household to a new area may differ, since households that choose to move to a given area may be more likely to benefit from that move than the average household in the population.' (Chetty and Hendren, 2015, 5). We note this difference because it illustrates an important change in emphasis.

the importance of a range of household characteristics over and above family income and county of childhood residence in determining the upward mobility of children. A second aim is to assess whether the same household characteristics are associated with the geographic mobility (i.e., moving) of low-income households, suggesting caution in using such mobility to 'test' the effects of place on mobility.

3. Data

The demographic data used for this study come from the 5% Public Use Microdata Sample (PUMS) of the 1990 and 2000 U.S. Decennial Censuses (Ruggles et al., 2015). For each Census, we selected observations from family units having at least one child aged 0–12 years old with their mother and/or father present.[3] Each family unit receives a single observation within its respective sample, where the mother's characteristics describe the family if a she is present; otherwise, the father's characteristics are used.[4]

The advantage of using PUMS data for the analysis is that they allow us to observe a set of family-level characteristics, such as race and education attainment, that are not available in the IRS data used by Chetty et al. (2014). Our starting point is to explore how several families' characteristics correlate with county-level measures of intergenerational income mobility provided by Chetty et al.[5] The finest geographic variable in the PUMS files is a family's public use microdata area, or PUMA.[6] For most (61%) families this identifies the county of residence. Where a PUMA crosses county lines, we assign its families to 'super-counties'. A super-county is constructed as the smallest possible group of contiguous counties that fully contains all overlapping PUMAs, but whose individual county components do not fully contain all of their overlapping PUMAs.[7] Going forward, we make no distinction between individual counties and super-counties, referring to each unit of geography as a 'super-county'.

The key metric of intergenerational income mobility highlighted by Chetty et al. (2014) is the absolute intergenerational income mobility (AIIM) of children whose

3 Chetty et al. (2014) used a sample of children born between 1980 and 1982 in their analysis, whose families can therefore be viewed as a subset of our larger sample of families. We do not limit our sample exclusively to families whose children would have belonged to the 1980–1982 birth cohort because we are not limited by data availability, as were Chetty et al. (2014). Because we are interested in documenting differences in family characteristics across space, we desire to use the largest, most representative sample of families for each area. We use the 2000 Census to assess whether the main findings from the 1990 Census are somehow unique. Families in the 2000 year sample are limited to those with children between the ages of 0 and 12.

4 A mother's characteristics are assigned to 96 (94)% of the observations in the 1990 (2000) samples.

5 Chetty et al. (2014) conduct analyses and construct measures of intergenerational income mobility for geographical units defined by commuting zone and counties. The choice of geography is not consequential, as their findings do not depend in a meaningful way on what level of geography is used. We use counties, although for some portion of our sample we need to combine observations into what we refer to as 'super-counties' (discussed below).

6 PUMAs have populations of at least 100,000 but, typically, no more than 200,000. PUMAs generally follow the borders of counties, groups of counties or census-defined places. Individual PUMAs do not cross state boundaries.

7 This approach yields 897 geographic units of analysis that we can match to microdata observations, 386 counties and 511 super-counties. Super-counties are built up from an average of 5.3 individual counties. Using this approach, for the 1990 sample we are able to assign 711,264 family observations to individual counties and 460,375 family observations to super-counties. For the 2000 sample, we assign 751,382 observations to 382 individual counties and 405,074 observations to 499 super-counties. Observations that we cannot assign to either a specific county or super-county are dropped from the samples.

parents' incomes fell within the 25th percentile of the national income distribution (for Years 1996–2000). Accordingly, we limit our samples to families whose parents' incomes fell within the 3rd decile of the national income distribution in each Census year. In each year, 1990 and 2000, we derive the distribution of family incomes using self-reported earnings data provided in the PUMS files. Family income is calculated as the sum of the mother's and, if present, the father's personal incomes. This approach mimics that used by Chetty et al. (2014). For convenience, we refer to our family income measure as 'nuclear family income'. For families whose super-county consists of only a single county, their county's AIIM is taken directly from the actual county-level AIIM estimates made available in Chetty et al.'s Online Data Appendix.[8] For families assigned to super-counties that are made up of multiple counties, their super-county's AIIM is calculated as the weighted average of the county-level AIIM estimates provided by Chetty et al.[9]

We then go on to measure several characteristics of families in our sample: race, Hispanic origin, educational attainment, marital status, immigration status and age. We also measured whether the family was a recent mover determined by comparing their super-county at the time of the Census survey (i.e., 1990 or 2000) with their super-county from 5 years prior, if it can be identified.[10] We assigned to each family its super-county's AIIM quintile, which simply measures its super-county's position along the AIIM distribution (across all super-counties). We chose these family attributes with exception of migration status because they overlap with data from the National Longitudinal Survey of Youth—1997 Cohort (NLSY97). We use the NLSY97 to construct measures of predicted adult incomes. We describe how we construct predicted income in more detail below. In addition to these family characteristics that overlap with information in the NLSY97, we measure several others: home ownership, status as a welfare recipient, number of own children in household, number of children aged 5 or younger and number of children ever born.

4. Analysis of family characteristics

4.1. Geography of demographic differences

The purpose of the analysis that follows is to assess the degree to which several family characteristics vary across super-county AIIM quintiles and the extent to which any variation in family-level characteristics can be used to explain the inter-quintile variation in AIIM. If the variation in AIIM across areas was purely a function of place-level characteristics and not family—or person-level characteristics, then we would expect little variation across areas in family characteristics. Conversely, if low-income families' characteristics differ across areas, then it may very well be just the families

8 This appendix is currently available at the Equality of Opportunity Project website, www.http://www.equality-of-opportunity.org/data/.

9 Weighted average AIIM scores for super-counties are constructed using person-level sample weights. Weighted averages do not differ substantially from non-weighted averages.

10 The 1990 and 2000 PUMS files provide information on a person's PUMA of residence 5 years prior. However, a family's migration status cannot be measured for a small number of cases due to differences in how PUMS records identify a family's current and previous PUMA. That is, for a small number of cases, it is not clear whether or not a family that indicated that it had moved actually crossed super-county boundaries.

themselves that explain an area's AIIM, either directly or through their influence on an area's institutional characteristics.[11]

To begin, we calculate the mean proportion of families in the kth AIIM quintile that have characteristic j, which we abbreviate by X_{jk}. In order to more easily conduct tests of statistical significance for differences in X_{jk} across the super-county quintiles we estimate the following regression model separately for the set of J family-level characteristics:

$$x_{ij} = \sum_{k=1}^{5} \beta_{jk} \text{AIIM}_{ik} + \epsilon_{ij}, j = 1, \ldots, J \tag{1}$$

where x_{ij} is a dichotomous 0-1 indicator equal to unity if family i has characteristic j, and AIIM_{ik} is also a dichotomous 0-1 indicator that is equal to unity if family i's super-county belongs to the kth quintile of the AIIM distribution. Because the five AIIM_{ik} variables included in the model are each mutually exclusive, the parameter β_{jk} in the j equation can be interpreted as the mean proportion of families residing in a kth-quintile super-county that have characteristic j.

Estimates of the mean family characteristics using the 1990, 5% PUMS file are provided in Table 1. Analogous estimates using the 2000 PUMS sample are provided in Table A1 of the Online Appendix. Within each table, individual columns are grouped into larger panels based on the broader demographic characteristic being described (e.g., race, educational attainment, etc.). Also, within each column, asterisks next to a share estimate indicates the degree to which that estimate is statistically different from the share estimated for the 3rd quintile of AIIM (reported in the middle row).

Estimates reported in Table 1 show clearly that there is substantial heterogeneity in family, demographic characteristics across super-counties of different AIIM status. Most notably, the racial composition of low-income families becomes increasingly black as AIIM declines, as does the share of low-income families reporting that the parent is not married. For example, about 36% of low-income families within bottom quintile super-counties are black, whereas blacks account for about only 4% of all families within top-quintile super-counties. Alone, this striking 9-fold difference in racial composition suggests that spatial differences in AIIM may be as much, if not more, about the characteristics of the low-income families themselves and their *individual* burdens, as it is about the actual places within which they reside.

Indeed, with the exception of age, the relationships observed in Table 1 indicate clear patterns of selection across AIIM quintiles based on race, ethnicity and family structure. For educational attainment, statistical differences are observed between the lowest and highest quintiles for the two endpoint categories of education (i.e., <HS and BA+), with low-income parents in bottom quintile super-counties 20% more likely to have not completed high school, relative to parents in the highest quintile super-counties. Similarly, parents in the bottom quintile super-counties are 23% less likely to have graduated from college. Families in the lowest AIIM quintile are approximately twice as likely as those in the top AIIM quintile to be headed by a never-married parent,

11 We note, however, that the families in our sample make up less than 10% of all families because our sample is limited to families with children and who are in the third decile of the income distribution. Therefore, the direct influence of these families on county (commuting zone) institutions, or policies, that influence intergenerational income mobility is likely quite small.

Table 1. Distribution of parent characteristics within upward mobility quintiles, 1990 Census

Upward mobility quintile	Race			Hispanic	Educational attainment				Marital status			Foreign-born	Age
	White	Black	Other		<HS	HS only	Some college	BA+	Married	Div./ Sep./Wid.	Never married		
q1 (lowest)	0.589***	0.364***	0.047***	0.048***	0.248***	0.393	0.298***	0.061	0.569***	0.295***	0.136***	0.049***	31.9**
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)	(0.003)	(0.003)	(0.002)	(0.002)	(0.046)
q2	0.741***	0.162***	0.0988***	0.138***	0.263	0.397**	0.282	0.059	0.634*	0.275***	0.091	0.093***	31.6*
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)	(0.003)	(0.003)	(0.002)	(0.002)	(0.046)
q3	0.757	0.126	0.117	0.157	0.268	0.386	0.285	0.061	0.642	0.263	0.095	0.144	31.7
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.048)
q4	0.737***	0.072***	0.190***	0.264***	0.309***	0.365***	0.267***	0.059	0.662***	0.248***	0.090*	0.220***	32.0***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.049)
q5 (highest)	0.864***	0.041***	0.095***	0.120***	0.206***	0.415***	0.300***	0.079***	0.695***	0.233***	0.072***	0.110***	31.9**
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)	(0.002)	(0.004)	(0.003)	(0.002)	(0.002)	(0.057)

Notes: Each cell in this table reports the share of low-income families having the characteristic identified in the column's header, stratified by the upward mobility ranking of their place of residence. The patterns reported in this table indicate that low-income families in more upwardly mobile areas are whiter, and more likely to have a college educated or married parent present, relative to low-income families from less upwardly mobile areas. All values reported here are estimates of β_{jk}'s from Equation (1). All values are based on family-level observations ($N = 118, 857$) provided by the U.S. Census' 1990 PUMS file (5% sample). This sample is restricted to families with own children between the ages of 0 and 12 who have incomes within the 3rd decile of the national income distribution. The characteristics assigned to each family are based on those of the mother, if present, or those of the father, if the mother is not present. Families with no mother or father present are omitted from the sample as are families with multiple mothers or fathers present. All values are calculated using sample weights. A family's absolute upward mobility quintile is determined by assigning it to 1 of 897 counties or 'super' counties. Asterisks indicate the statistical significance of the difference between the value reported in that cell and the value reported in the cell that corresponds to areas with an absolute intergenerational mobility ranking of 3 (i.e., the middle row). ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.1$.

and half as likely to have a foreign-born parent. These significant differences in low-income family characteristics across places with different AIIM are notable because they are present even though all families in the sample are in the same 3rd decile of the national income distribution. It is clear that adjusting only for family income, as in Chetty et al. (2014), is not sufficient to make families comparable.

In Online Appendix Table A6, we present estimates similar to those in Table 1, but for family characteristics that do not overlap with information in the NLSY79. These estimates also show significant differences in family characteristics between counties. The share of families that own their home in the bottom AIIM quintiles is significantly lower (20%) than in the top AIIM quintile. The number of family members within a household is also significantly lower among families in the lowest AIIM quintile relative to those in the highest AIIM quintile. There are also differences in the share of families receiving welfare across super-counties, although these differences are not monotonically related to AIIM. The patterns in Online Appendix Table A6 serve to reinforce the conclusion we drew from Table 1—despite having roughly similar incomes, there are substantial differences in family characteristics between super-counties and these differences are correlated with the super-county AIIM.

It is important to note that the patterns observed here are similar if a greater number of quantiles are used to classify super-counties by their AIIM score. Although our use of quintiles may appear a bit coarse, with each quintile accounting for 20 percentiles along the AIIM distribution, we find that much narrower quantile bands yield similar results. To see this, Figure 1 plots the relationship between several family characteristics and super-county AIIM score after assigning each family's super-county to 1 of 25 quantiles, with each band now accounting for only 4 percentiles along the AIIM distribution. The patterns in these figures make clear the point that families with better predictors of children's future income are more likely to reside in super-counties with higher AIIM scores.

4.2. Simulation of adult incomes

The patterns observed in Table 1 suggest that, for low-income families, variation in family-level characteristics may explain a sizable share of the difference in AIIM across communities. If this share explained by measured characteristics is indeed large, then this raises questions about the interpretation of Chetty et al.'s (2014) and Chetty and Hendren's (2017a, 2017b) results. That is, the channels through which AIIM is determined may be more directly linked to an individual child's person- and family-level characteristics, and less so to a particular place's characteristics.

To assess how much of the variation in AIIM is due to the characteristics found in Table 1, we employ a two-step approach to gauge the share of AIIM that can be explained by low-income families' own characteristics without regard for where they live. In our first step, we use data from the NLSY97 Cohort to estimate the conditional (partial) correlations between an adult's (nuclear) family income and their mother's characteristics. The sample consists of adult ages 27–31 in 2011, which corresponds closely with the age of adults used in Chetty et al. (2014). Mother's characteristics are measured in 1997 when the children were between the ages of 12 and 17. We further limit the sample to adults (in 2011) whose family's income in 1997 was at or below the sample median of 1997 family incomes because we want to focus on children living in lower income households that are roughly comparable to the children in Chetty et al.

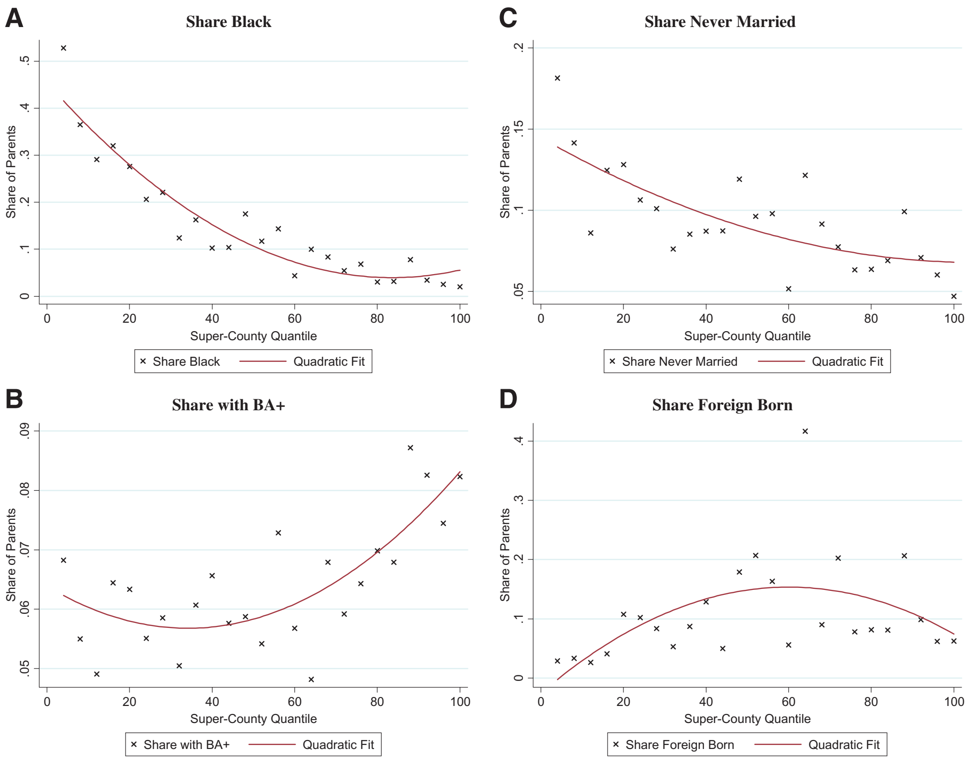

Figure 1. Distribution of parent characteristics by upward mobility quantile, 1990 Census (**A**) Share black (**B**) Share with BA+ (**C**) Share never married (**D**) Share foreign-born. Each figure displays the relationship between families' upward mobility quantile of residence (identified by their super-county of residence) and the fraction of parents sharing a specific characteristic. Each of the 25 upward mobility quantiles displayed in the figure reflect super-counties belonging to four upward mobility percentiles. For example, the 4th quantiles reflects super-counties belonging to the percentiles 1–4, and the 20th quantile reflects super-counties belonging to percentiles 17–20, etc.

(2014) from the 25th percentile of the income distribution. We estimate the following regression model:

$$\mathrm{ninc}_i^{11} = \delta_0 + \sum\nolimits_j \delta_j x_{ij}^{97} + \sum_a \gamma_a \mathrm{age}_{ia}^{11} + \epsilon_i^{11} \tag{2}$$

where ninc_i^{11} measures the person's nuclear family income in 2011 and x_{ij}^{97} measures their mother's jth characteristic in 1997. These maternal characteristics are, with two exceptions, the same as those used in Table 1 and include dummy variables for education (high school, some college, and BA or more), dummy variables for marital status (married and divorced/separated/widowed), mother's age, mother's age squared, a dummy variable for foreign born and dummy variables for race/ethnicity (white, black and Hispanic).[12] In some models, we also include the family's income in 1997 and

12 Please note that, as in Table 1, the dichotomous x_{ij}^{97} variables measuring specific classes of educational attainment, race, and marital status are mutually exclusive within the broader demographic

family income squared to adjust for income differences among the sample. Because of small sample sizes, we use a sample of families from the lower half of the income distribution instead of from the 3rd decile (or 25th percentile as in Chetty et al. 2014). The variables age_{ia}^{11} are a set of dichotomous indicators that identify person *i*'s age in 2011 (ages 27–31). We include this variable to control for differences in adult age that may influence income. Equation (2) makes no reference to a family's place of residence. It is estimated to identify family determinants of upward mobility. Estimates of Equation (2) are reported in columns 1 and 2 of Table 2 below.

Overall, the parameter estimates in Table 2 for each of these variables appear reasonable and align with intuition and previous evidence. Adults having had more-educated mothers during childhood earn more in adulthood than those with less-educated mothers. For example, depending on the specification, someone whose mother had a BA or more earn \$7600–\$9400 more than someone whose mother had less than a high school degree. Other estimates are similarly unsurprising. Adults whose mothers were married, or were not a racial minority, earn more than adults whose mothers were never married or who were non-white. Interestingly, adults whose mother was foreign-born earn more than adults whose mother was born in the USA.

We now turn to simulating expected 2011 nuclear family incomes for someone whose parent(s) lived in a given AIIM quintile ($\widehat{ninc_k}$), where the symbol '^' denotes an estimated value. Simulated income in AIIM quintiles are obtained from the following calculations:

$$\widehat{\text{ninc}}_1 = \delta_0 + \sum\nolimits_j \delta_j x_{1j}^{\text{Census}}$$

$$\ldots. \tag{3}$$

$$\widehat{\text{ninc}}_5 = \delta_0 + \sum\nolimits_j \delta_j x_{5j}^{\text{Census}}$$

The calculations in Equation (3) use the mean value of family characteristics in each quintile (e.g., x_{1j}^{Census}) obtained from the Census and the parameter estimates (δ_j) from the regression model of nuclear family income given by Equation (2). The difference in simulated income between quintiles k and m is given by:

$$\widehat{\Delta\text{ninc}}_{k,m} = \sum\nolimits_j \widehat{\delta}_j (x_{kj}^{\text{Census}} - x_{mj}^{\text{Census}}) \tag{4}$$

We focus on the difference in simulated income between AIIM quintiles and compare it with the differences in adult income between AIIM quintiles reported by Chetty and Hendren (2017a). Specifically, we calculate the share of the Chetty and colleagues difference in income between AIIM quintiles that can be explained by the differences given in Equation (4).

While the absolute dollar differences in Equation (4) are straightforward and useful, they may differ from Chetty and colleagues reported differences solely because Chetty and colleagues used the income distribution from income tax returns, which we do not

characteristic that they are describing. Thus, each of these parameter estimates should be interpreted relative to the excluded 'base' variable. For example, the estimated coefficients for the variables *Married* and *Divorced*/*Separated*/*Widowed* should be interpreted as measuring their correlation with 2011 nuclear family income relative to those who were never married.

Table 2. 2011 nuclear family income correlated with mother's 1997 characteristics, 1997 NLSY sample

Mother's characteristics in 1997	No county-level controls		County-level controls, 1990		County-level fixed effects	
	[1]	[2]	[3]	[4]	[5]	[6]
HS	5666***	4326***	5417***	4151**	4467***	3204*
	(1622)	(1629)	(1635)	(1643)	(1730)	(1736)
Some college	9682***	7679***	9179***	7318***	9716***	7741***
	(1929)	(1951)	(1950)	(1972)	(2068)	(2089)
BA or more	10068***	7609***	9912***	7577***	7655**	5179*
	(2843)	(2861)	(2879)	(2899)	(3070)	(3088)
Married	5494***	3149	5992***	3669*	5141**	2450
	(2096)	(2124)	(2128)	(2160)	(2265)	(2306)
Divorced/Separated/	3342	3203	3843*	3724*	2791	2406
Widowed	(2171)	(2157)	(2186)	(2174)	(2328)	(2313)
Mother's age	−522	−1058	−631	−1136	−951	−1424
	(972)	(974)	(970)	(971)	(1000)	(999)
Mother's age squared	6.36	12.82	7.87	13.91	11.80	17.35
	(11.71)	(11.72)	(11.68)	(11.70)	(12.02)	(12.01)
Foreign-born	6533***	6369***	4898*	4670*	6740***	6303**
	(2202)	(2188)	(2433)	(2420)	(2560)	(2545)
White	6443***	6617***	7265***	7526***	6199**	6719**
	(2233)	(2221)	(2279)	(2269)	(2682)	(2669)
Black	−9445***	−8026***	−9877***	−8846***	−7899**	−6861**
	(2588)	(2590)	(2788)	(2780)	(3335)	(3321)
Hispanic	−1177	68	−1999	−1075	−1194	−533
	(2154)	(2156)	(2455)	(2452)	(2785)	(2773)
Family income		206		246		319
(in thousands)		(229)		(233)		(249)
Family income squared		3.93		2.58		1.48
		(5.74)		(5.79)		(6.22)
Constant term	29,174	35,711*	44,690	42,552	41,712*	46,409*
	(20,079)	(19,986)	(34,998)	(34,848)	(24,444)	(24,316)

Notes: All values reflect estimates of Equation (2) using the NLSY's 1997 cohort. $N = 2{,}272$. The sample is limited to persons aged 27–31 in 2011 whose 1997 family income was at or below the sample median. Respondents' 2011 age indicators are included in each model. Estimates reported in columns [3] and [4] control for a range of 1990 county-level characteristics. (See text for details). Estimates reported in columns [5] and [6] control for county-level fixed effects. $^{***}p \le 0.01$, $^{**}p \le 0.05$, $^{*}p \le 0.1$.

have access to use. Our simulated income differences were obtained using the distribution of family characteristics in the 1990 Census and parameter estimates obtained from the NLSY97. Therefore, as an alternative to the absolute differences in income shown in Equation (4), we also constructed measures of the relative differences in simulated income between AIIM quintiles. To construct these relative differences, we set the simulated income in AIIM quintile 1 to $28,875. This is the 37th percentile of the nuclear family income distribution for the 1980–1982 birth cohort, taken the 2011 American Community Survey. We use the 37th percentile here because it corresponds to the income position predicted by Chetty et al. (2014) for this quintile, and it comes from the Census (ACS), which was used to calculate family characteristics. That is, the average AIIM score for our lowest quintile of super-counties is 37. For all other AIIM quintiles, k, we simulate expected 2011 nuclear family income as:

$$\widehat{\text{ninc}}_k = 28,875 + \sum_j \widehat{\delta}_j \left(x_{kj}^{\text{Census}} - x_{1j}^{\text{Census}} \right) \tag{5}$$

Note that the only difference between Equation (5) and Equation (3) is that we have used the actual income in AIIM quintile 1 in the 2011 Census to establish a baseline (AIIM quintile 1) income that we believe is more closely aligned with the income distribution from income tax records used by Chetty and colleagues. The relative differences in income between AIIM quintile k and AIIM quintile 1 are:

$$\frac{\widehat{\text{ninc}}_k}{\widehat{\text{ninc}}_1} = \frac{28,875 + \sum_j \widehat{\delta}_j \left(x_{kj}^{\text{Census}} - x_{1j}^{\text{Census}} \right)}{28,875} \tag{6}$$

We calculate the share of the Chetty and colleagues' *relative* difference in income between AIIM quintiles that can be explained by the differences given in Equation (6).

Simulations of 2011 nuclear family incomes of adult children from low-income families for each super-county quintile, $\widehat{\text{ninc}}_k$, are reported in columns 1 and 2 of Table 3's top panel. Here, differences between the values displayed in columns 1 and 2 simply reflect the differences between the coefficients reported in columns 1 and 2 of Table 2, respectively. Focusing on column 2 of Table 3, these values suggest that, on their own, the relatively limited number of 1990 family-level characteristics entering into the simulation predict a substantial difference between the 2011 nuclear family incomes of those who grew up in the least and most upwardly mobile super-counties (i.e., \$28,875 and \$34,121, respectively), with the greatest increase in simulated income occurring between the lowest and second-lowest upwardly mobile super-counties.

The bottom panel of Table 3 gauges the significance of these simulated income differences by reporting the share that they explain of the differences that Chetty et al.'s (2014) own AIIM indices would have predicted.[13] For example, focusing again on column 2, we see that, per our model, incomes simulated from 1990 family-level characteristics explain 114% of the AIIM-score-based predicted dollar difference in nuclear family incomes (and 92% of the percentage point change) between the lowest and second-lowest upwardly mobile super-counties. On the low end, simulated incomes predict just above 50% of the AIIM-score-based predicted dollar difference (about 40% of the percentage point change) between the lowest and highest upwardly mobile super-counties. Note that the explained share of variation in AIIM does not depend on the specification of the earnings model used in Table 2. Altogether, these shares are relatively large and suggest that much of the spatial variation in AIIM

13 To estimate the level of income for a super-county predicted by Chetty et al.'s own AIIM scores, we simply apply the average AIIM score for a super-county quintile to the 'child family' income distribution provided by Chetty et al. in their Online Data Appendix. For example, super-counties belonging to the lowest and highest AIIM quintile have average AIIM scores of 36.5 and 48.4, respectively, which translate into respective nuclear family incomes of approximately \$23,300 and \$33,520.

Table 3. Simulated nuclear family income of persons' ages 27–31 in 2011 within upward mobility quintiles based on 1990 parent characteristics

Chetty et al.'s absolute intergenerational mobility quintile (Mean AIIM in quintile)	Predicted child nuclear family income (using Chetty et al's distribution)	Simulated income in 2011 based on 1990 parent characteristics					
		No county-level controls		County-level controls, 1990		County-level fixed effects	
		[1]	[2]	[3]	[4]	[5]	[6]
q1 (36.5)	23,300	28,875	28,875	28,875	28,875	28,875	28,875
q2 (39.8)	25,940	32,115	31,875	32,212	32,030	31,783	31,589
q3 (42.2)	27,860	32,843	32,541	32,869	32,631	32,462	32,201
q4 (44.3)	29,640	33,311	33,038	33,144	32,912	32,873	32,566
q5 (48.4)	33,520	34,823	34,121	35,050	34,476	34,183	33,613
$Change							
q2 *less* q1	2640	3240	3000	3338	3155	2908	2714
q3 *less* q1	4560	3969	3667	3994	3756	3588	3326
q4 *less* q1	6340	4436	4164	4270	4038	3998	3691
q5 *less* q1	10,220	5949	5247	6176	5602	5309	4739
Share of Chetty et al.'s $change explained							
q2 *less* q1		1.23	1.14	1.26	1.20	1.10	1.03
q3 *less* q1		0.87	0.80	0.88	0.82	0.79	0.73
q4 *less* q1		0.70	0.66	0.67	0.64	0.63	0.58
q5 *less* q1		0.58	0.51	0.60	0.55	0.52	0.46
Share of Chetty et al.'s % change explained							
q2 *less* q1		0.99	0.92	1.02	0.96	0.89	0.83
q3 *less* q1		0.70	0.65	0.71	0.66	0.63	0.59
q4 *less* q1		0.56	0.53	0.54	0.51	0.51	0.47
q5 *less* q1		0.47	0.41	0.49	0.44	0.42	0.37

Notes: Each cell in the top panel reports the simulated income of adult children from low-income families stratified by the upward mobility ranking of their place of residence when they were young. Incomes predicted from Chetty et al.'s (2014) data are also provided. The bottom panel shows that simulated incomes explain a sizable share of the variation in adult children's incomes predicted using Chetty et al.'s data. Simulated incomes in the top panel are calculated by applying the coefficients reported in Table 2 to the values reported in Table 1. See Equation (5). All estimates in columns [1]–[6] of the first panel are anchored to the 37th percentile of the nuclear family income distribution of the 1980–1982 birth cohort ($28,875) calculated from the 2011 American Community Survey. See discussion in Section 4.2 for further details. The bottom panel reports the share of the dollar change or percentage change in AIIM-score-based predicted incomes that can be explained by the simulated incomes reported in the top panel. See Footnote 13 for a discussion of how AIIM-score-based predicted incomes are calculated.

observed by Chetty et al. (2014) is due to unobserved variation in family-level characteristics.

4.3. Simulation of adult incomes: separating people from place

A potential criticism of our simulations to this point is that the key family-level variables in the simulations may be correlated with county-level characteristics not included in the regression model of Equation (2). If this were the case the simulations would be attributing place-effects to the person-level variables. We address this possibility directly by including a set of county-level controls in Equation (2). Using confidential person-level county identifiers provided by the NLSY, we re-estimate Equation (2) in two ways:

$$\mathrm{ninc}_{ic}^{11} = \delta_0 + \sum\nolimits_j \delta_j x_{ij}^{97} + \sum_a \gamma_a \mathrm{age}_{ia}^{11} + \sum\nolimits_k \lambda_k z_{ck} + \epsilon_{ic}^{11} \tag{7}$$

$$\mathrm{ninc}_{ic}^{11} = \delta_0 + \sum\nolimits_j \delta_j x_{ij}^{97} + \sum_a \gamma_a \mathrm{age}_{ia}^{11} + D_c + \epsilon_{ic}^{11} \tag{8}$$

where z_{ck} in Equation (7) measures the kth attribute of county c and λ_k is the parameter estimated for the relation of that attribute to income mobility and D_c in Equation (8) is a county-level fixed effect.[14] The estimates of the coefficients on family characteristics from these regressions are reported in columns 3–6 of Table 2. These estimates are quite similar to those in columns 1 and 2, suggesting that county-level variables have only modest influence on future income after controlling for family characteristics.

Following the approach taken above we use the new δ's from Equations (7) and (8) to simulate 2011 nuclear family incomes. The results are presented in columns 3 through 6 in Table 3. Focusing on column 6, personal characteristics independently of the influence of county-level fixed effects explain 103% of Chetty et al.'s dollar difference (83% of their percentage change) between AIIM quintile 1 and AIIM quintile 2 and 46% of Chetty et al.'s dollar difference (37% of their percentage change) between AIIM quintile 1 and AIIM quintile 5. These figures are very similar to our estimates in column 2 with no county-level controls.

All told, county-level effects seem to have little confounding influence on the effect of family-level characteristics. If place matters, it is likely to be defined at a much finer geography. The neighborhood school or the socioeconomic characteristics on the block might be more appropriate than county. However, at that level of geography, personal characteristic and neighborhood characteristics may be extremely difficult to untangle.

5. Movers and nonmovers

We turn now to the question of movers. First, we assess if family-level characteristics vary by whether the family was a recent (within 5 years) mover out of a super-county, relative to nonmovers originating from the same super-county. Here, families are assigned the AIIM quintile of their previous (i.e., 'origin') super-county, AIIM_{ik}^{o}, which is determined by their super-county of residence 5 years prior. We then calculate the

14 Fourteen county-level variables are used for z_{ck}. These include the population share white, share black, share Asian, share Hispanic, share urban, share in poverty and share foreign born; the share of the population 25+ with less than a high-school degree, a high-school degree only, or a Bachelor's degree or higher; share of persons 15+ married, share divorced; share of families headed by a single mother; and the unemployment rate. With the exception of the 1990 unemployment rate, which is provided by the Bureau of Labor Statistics, all variables are calculated from the 1990 Decennial Census.

mean proportion of all movers *from* quintile k super-counties with characteristic j, x^{m}_{kj}, and the mean proportion of all nonmover families in quintile k counties with characteristic j,x^{nm}_{kj}.[15]

Table 4 takes a deeper look into the variation observed in Table 1 by highlighting demographic differences across families' location in the AIIM distribution *and* mover status, focusing on the AIIM quintile of the family's super-county of origin.[16] Here, in addition to the asterisks that indicate differences *within* a column, the 'a', 'b' or 'c' superscripts next to an estimate in the 'mover column' indicate how that estimate differs from the estimate in the 'nonmover' column within the same AIIM quintile.

Low-income families who moved out of the lowest AIIM super-counties, when compared with low-income nonmovers from the same super-counties, are 27% more likely to be white; 52% more likely to have a college educated parent and 20% more likely to be headed by a married couple. All of these family characteristics are positively associated with a child's future earnings (shown below). Similar differences characterize low-income movers and nonmovers from other quintiles, but, in general, movers and nonmovers tend to be more similar in the top quintile and less similar in the bottom quintile. There is much less 'selection' on family characteristics by AIIM among movers than nonmovers. Of course, this table reveals nothing about the type of move that a low-income family makes when choosing to leave an area with low AIIM scores. These families could be moving to areas with significantly better AIIM rankings, marginally better rankings, or simply be making 'lateral' moves across areas with relatively similar rankings.

To investigate this issue further, we limit our sample to low-income families originating in super-counties in the a lowest AIIM quintile and estimate the share of those families moving to a super-county Δk quintiles higher along the AIIM distribution that have characteristic j, $x^{m}_{\Delta k,j}$. Similarly, we also estimate the share of low-income families who choose not to move and have characteristic j, x^{nm}_{j}.[17]

---

15 Again, we use a regression framework to make significance testing easier. In particular, the following equation is estimated:

$$x_{ij} = \sum_{k=1}^{5} \beta^{m}_{kj}\mathrm{AIIM}^{o}_{ik} \times \mathrm{mover}_i + \sum_{k=1}^{5} \beta^{nm}_{kj}\mathrm{AIIM}^{o}_{ik} \times \mathrm{nonmover}_i + \varepsilon_{ij}, j = 1, \ldots, J \qquad (9)$$

where mover_i and $\mathrm{nonmover}_i$ are mutually exclusive 0-1 indicator variables set equal to unity if the family did or did not change super-counties within the 5 years prior to being surveyed, respectively. Here, the parameter β^{m}_{kj} measures the mean proportion of families who moved out of a kth-quintile super-county that had characteristic j. Conversely, β^{nm}_{kj} measures the share of families who stayed in the kth-quintile super-county that had characteristic j.

Alternative estimates of Equation (9) were made using the AIIM quintile of a family's current (i.e., 'destination') super-county, AIIM_{ik}. Here, the parameters β^{m}_{kj} would measure the mean proportion of families moving into a kth-quintile super-county that had characteristic j. These estimates are provided in Table A2 and A5 of the Appendix for 1990 and 2000, respectively.

16 Appendix Table A2 reports estimates using county of destination as comparison. The estimates are very similar to those reported in text.

17 For the reasons noted above, we estimate these shares from the following equation for only those families whose origins are in super-counties belonging to the 1st quintile of the AIIM distribution:

$$x_{ij} = \sum_{\Delta k=0}^{4} \beta^{m}_{\Delta k,j}\mathrm{AIIM}_{i,\Delta k} \times \mathrm{mover}_i + \beta^{nm}_{j} \times \mathrm{nonmover}_i + \varepsilon_{ij}, j = 1, \ldots, J \qquad (10)$$

where $\mathrm{AIIM}_{i,\Delta k}$ is a 0-1 indicator variable equal to unity if the family moved to a super-county Δk quintiles higher along the AIIM distribution relative to their super-county of origin (which falls within

Table 4. Distribution of parent characteristics within upward mobility quintiles by mover status in origin location, 1990 Census

Upward mobility quintile of origin	Race						Hispanic	
	White		Black		Other			
	Mover	Nonmover	Mover	Nonmover	Mover	Nonmover	Mover	Nonmover
q1 (lowest)	0.702^{***,a}	0.554***	0.251^{***,a}	0.399***	0.047***	0.047***	0.039^{***,c}	0.049***
	(0.006)	(0.003)	(0.005)	(0.002)	(0.004)	(0.002)	(0.005)	(0.002)
q2	0.805^{a}	0.724***	0.112^{a}	0.171***	0.083^{a}	0.105***	0.095^{*,a}	0.152
	(0.006)	(0.003)	(0.005)	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)
q3	0.808^{a}	0.746	0.103^{a}	0.136	0.089^{a}	0.118	0.108^{a}	0.163
	(0.006)	(0.003)	(0.005)	(0.003)	(0.004)	(0.002)	(0.005)	(0.002)
q4	0.790^{**,a}	0.716***	0.081^{***,b}	0.070***	0.128^{***,a}	0.214***	0.163^{***,a}	0.305***
	(0.006)	(0.003)	(0.005)	(0.003)	(0.004)	(0.002)	(0.005)	(0.003)
q5 (highest)	0.853^{***,c}	0.865***	0.058^{***,a}	0.039***	0.090	0.096***	0.105^{a}	0.129***
	(0.006)	(0.004)	(0.005)	(0.003)	(0.004)	(0.003)	(0.005)	(0.003)

Upward mobility quintile of origin	Educational attainment							
	<HS		HS only		Some college		BA+	
	Mover	Nonmover	Mover	Nonmover	Mover	Nonmover	Mover	Nonmover
q1 (lowest)	0.202^{*,a}	0.262***	0.373^{a}	0.396	0.342^{*,a}	0.288***	0.083^{a}	0.054
	(0.006)	(0.003)	(0.006)	(0.003)	(0.006)	(0.003)	(0.003)	(0.002)
q2	0.214^{a}	0.279	0.378^{a}	0.406*	0.325^{a}	0.265	0.082^{a}	0.049*
	(0.006)	(0.003)	(0.006)	(0.003)	(0.006)	(0.003)	(0.003)	(0.002)
q3	0.218^{a}	0.279	0.376^{a}	0.397	0.327^{a}	0.270	0.08^{a}	0.054
	(0.006)	(0.003)	(0.007)	(0.004)	(0.006)	(0.003)	(0.003)	(0.002)
q4	0.229^{a}	0.338***	0.361	0.362***	0.328^{a}	0.248***	0.082^{a}	0.053
	(0.006)	(0.003)	(0.007)	(0.004)	(0.006)	(0.003)	(0.003)	(0.002)
q5 (highest)	0.179^{***,a}	0.217***	0.360^{*,a}	0.428***	0.355^{***,a}	0.284***	0.107^{***,a}	0.071***
	(0.006)	(0.004)	(0.007)	(0.004)	(0.007)	(0.004)	(0.004)	(0.002)

(continued)

Table 4. Continued

Upward mobility quintile of origin	Marital status						Foreign born		Age	
	Married		Div./Sep./Wid.		Never married					
	Mover	Nonmover	Mover	Nonmover	Mover	Nonmover	Mover	Nonmover	Mover	Nonmover
q1 (lowest)	0.654***,a	0.547***	0.249a	0.304***	0.097***,a	0.149***	0.055***,c	0.047***	29.6**,a	32.5***
	(0.006)	(0.003)	(0.006)	(0.003)	(0.004)	(0.002)	(0.004)	(0.002)	(0.099)	(0.052)
q2	0.673a	0.621	0.258a	0.282**	0.069a	0.097	0.08***,a	0.095***	29.9a	32.2
	(0.006)	(0.003)	(0.006)	(0.003)	(0.004)	(0.002)	(0.004)	(0.002)	(0.098)	(0.053)
q3	0.684a	0.628	0.25a	0.272	0.066a	0.1	0.103a	0.151	29.9a	32.3
	(0.007)	(0.003)	(0.006)	(0.003)	(0.004)	(0.002)	(0.004)	(0.002)	(0.102)	(0.054)
q4	0.69a	0.654***	0.235*,b	0.249***	0.075a	0.097	0.139***,a	0.251***	30.1a	32.6***
	(0.006)	(0.004)	(0.006)	(0.003)	(0.004)	(0.002)	(0.004)	(0.002)	(0.1)	(0.056)
q5 (highest)	0.69	0.696***	0.237	0.232***	0.072	0.072***	0.107	0.113***	30.0a	32.5**
	(0.007)	(0.004)	(0.006)	(0.004)	(0.004)	(0.003)	(0.005)	(0.003)	(0.109)	(0.065)

Notes: Each cell in this table reports the share of low-income families having the characteristic identified in the column's header, stratified by their mobility status and the upward mobility ranking of their area of origin. Among some of the notable patterns reported in this table, it is observed that low-income families that moved out of the least upwardly mobile areas, when compared to low-income families that stayed, are 27% more likely to be white, 52% more likely to have a college educated parent and 20% more likely to be headed by a married couple. All values reported here are estimates of β_{kj}^{m}'s and $\beta_{k,\,j}^{nm}$'s from Equation (9). All values are based on family-level observations ($N = 118{,}209$) provided by the U.S. Census' 1990 PUMS file (5% sample). This sample is restricted to families with own children between the ages of 0 and 12 who have incomes within the 3rd decile of the national income distribution. The characteristics assigned to each family are based on those of the mother, if present, or those of the father if the mother is not present. Families with no mother or father present are omitted from the sample as are families with multiple mothers or fathers present. All values are calculated using sample weights. A family's absolute upward mobility quintile is determined by assigning it to 1 of 897 counties or 'super' counties. Movers in Table 4 are identified by comparing a family's current location to its location 5 years prior, as reported in the 1990 PUMS file. The absolute intergenerational mobility ranking of a mover's county or super county is determined by that family's *origin* location. That is, movers are defined as having recently moved *out of* that area. Statistics calculated after assigning movers to their destination location (i.e., their current location) are not substantially different from those reported here (see Online Appendix Table A2). Asterisks indicate the statistical significance of the difference between the value reported in that cell and the value reported in the cell that corresponds to areas with an absolute intergenerational mobility ranking of 3 (i.e., the middle row for that same column). ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.1$. For the 'mover columns', the letters reported in the superscripts indicate if there is a statistically significant difference between movers and nonmovers *within* areas having been assigned that ranking for absolute intergenerational mobility. [a]$p \leq 0.01$, [b]$p \leq 0.05$, [c]$p \leq 0.1$.

Share estimates of this type are reported in Table 5. The differences reported here, particularly between nonmovers and movers to counties with greater income mobility, are quite striking. For example, comparing low-income nonmovers to those who move to super-counties with the highest AIIM, it is clear that the latter group exhibits characteristics that are positively correlated with income. That is, compared to families that remain in their least upwardly mobile super-counties (i.e., they do not move), families moving from the lowest to the highest upwardly mobile super-counties are 35% more likely to be white; 117% more likely to have a college educated parent present and 18% more likely to be a married, two-parent family.

Perhaps even more significant among results in Table 5 is the fact that, among those families that have taken the initiative to move, there is considerable selection on observed characteristics by the type of move made. The heterogeneity in family characteristics is not just limited to that which exists between movers and nonmovers. It exists *within* movers as well. Low-income families that make 'lateral' moves in general possess less favorable predictors of income than those who relocate to the most upwardly mobile super-counties. For example, when compared to families who move to the most upwardly mobile areas, families making lateral moves are 15% less likely to be white, 25% less likely to have a college-educated parent, and 66% less likely to be foreign born, all of which are positive predictors of a child's future income. In Online Appendix Table A3, we provide a similar analysis, but for families who originate in AIIM quintile 5. Here too, we observe significant differences by type of move. For example, families that move from AIIM quintile 5 to AIIM quintile 1 are 10 times more likely to be black than families that move counties within AIIM quintile 5.

Taken together, these findings, along with those presented above, point to a significant amount of sorting between low-income families and their areas of residence, with families that possess the most 'income-favorable' attributes both residing in *and* moving to the most upwardly-mobile super-counties. The implication of these results is that the causal estimates of place on intergenerational income mobility in Chetty and Hendren (2017a, 2017b) are likely to lack external validity.[18]

In addition to the characteristics shown in Table 5, which are those used to simulate adult income, we also calculated means of several more family characteristics by mover status and the type of move. Table 6 reports these figures. A notable aspect of the family characteristics in Table 6 is that they are partly time-varying and changing between the time of move and the current period when the family is observed. The most striking result in Table 6 is the difference in home ownership between movers and nonmovers. Those who moved from the lowest AIIM quintile are approximately half as likely to own a home as those who remain in the lowest AIIM quintile. Families that move are also approximately 33% more likely to have children under age 5 than families that did not move. Similar differences between movers and nonmovers are

the 1st quintile). The parameter $\beta^{m}_{\Delta k,j}$ thus measures, among the families whose move yielded a Δk increase in their super-county's AIIM ranking, the share who exhibited characteristic j. Similarly, the parameter β^{nm}_{j} returns the share of nonmovers with characteristic j.

18 Chetty and Hendren (2017a) also conduct an analysis using a sample of moves likely driven by exogenous causes such as natural disasters and results from this are similar to those using their primary sample of movers.

Table 5. Distribution of parent characteristics by type of move made for families originating in the lowest quintile super-counties, 1990 Census

Mover type	Race			Hispanic	Educational attainment				Marital status			Foreign-born	Age
	White	Black	Other		<HS	HS only	Some college	BA+	Married	Div./Sep./Wid.	Never married		
No Move	0.554^{a}	0.399^{a}	0.047^{a}	0.049^{a}	0.262^{a}	0.396^{a}	0.288^{a}	0.054^{a}	0.547^{a}	0.304^{a}	0.149^{a}	0.047^{b}	32.5^{a}
(remain in q1)	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.004)	(0.003)	(0.002)	(0.004)	(0.003)	(0.003)	(0.002)	(0.055)
Move: q1 to q1	0.638***	0.338***	0.024***	0.012***	0.210***	0.360***	0.341***	0.088***	0.630***	0.261***	0.109***	0.034**	29.5***
	(0.012)	(0.011)	(0.005)	(0.005)	(0.010)	(0.012)	(0.011)	(0.006)	(0.012)	(0.011)	(0.008)	(0.005)	(0.178)
Move: q1 to q2	0.729^{***,a}	0.217^{***,a}	0.054^{a}	0.039^{*,a}	0.222***	0.380	0.331***	0.067^{**,b}	0.656***	0.245***	0.099***	0.058^{*,a}	29.5***
	(0.013)	(0.012)	(0.006)	(0.005)	(0.011)	(0.013)	(0.012)	(0.006)	(0.013)	(0.012)	(0.009)	(0.006)	(0.197)
Move: q1 to q3	0.758^{***,a}	0.182^{***,a}	0.059^{a}	0.069^{***,a}	0.182***	0.371	0.36***	0.087***	0.666^{***,c}	0.239***	0.094***	0.063^{**,a}	29.9***
	(0.016)	(0.016)	(0.007)	(0.007)	(0.014)	(0.016)	(0.015)	(0.008)	(0.016)	(0.015)	(0.011)	(0.007)	(0.249)
Move: q1 to q4	0.718^{***,a}	0.220^{***,a}	0.062^{*,a}	0.058^{a}	0.169^{***,b}	0.415^{b}	0.338***	0.077**	0.708^{***,a}	0.234***	0.058^{***,a}	0.072^{***,a}	29.1***
	(0.020)	(0.019)	(0.009)	(0.008)	(0.017)	(0.020)	(0.019)	(0.010)	(0.020)	(0.018)	(0.014)	(0.009)	(0.305)
Move: q1 to q5	0.750^{***,a}	0.171^{***,a}	0.080^{***,a}	0.071^{*,a}	0.187***	0.339**	0.357***	0.117^{***,b}	0.643***	0.260*	0.096***	0.1^{***,a}	30.5^{***,b}
	(0.027)	(0.027)	(0.012)	(0.012)	(0.024)	0.027)	(0.026)	(0.013)	(0.028)	(0.026)	(0.019)	(0.012)	(0.423)

Notes: Each cell in this table reports the share of low-income families that have the characteristic identified in the column's header, limited to only those low-income families that originated in the least upwardly mobile areas, stratified by the type of move made. Compared to those low-income families that do not move, low-income families that move from the least to the most upwardly mobile areas are 35% more likely to be white; 117% more likely to have a college educated parent present and 18% more likely to be a married, two-parent family. All values reported here are estimates of $\beta_{\Delta k,j}{}^{m}$'s and β_j^{nm}'s from Equation (10). All values are based on family-level observations ($N = 23{,}641$) provided by the U.S. Census' 1990 PUMS file (5% sample). This sample is restricted to families whose origin super-county belongs to the 1st quintile of the AIIM distribution. For families that move to different super-counties, the AIIM quintile of their destination super-county is used to determine their 'type' of move. See the caption of Table 1 for additional information on the sample. Asterisks indicate the statistical significance of the difference between the value reported in that cell and the value reported in the 'No Move' cell (i.e., the first row). $^{***}p \leq 0.01$, $^{**}p \leq 0.05$, $^{*}p \leq 0.1$. Letters reported in the superscripts indicate the statistical significance of the difference between the value reported in that cell and the value reported in the lateral move cell (i.e., q1 to q1). $^{a}p \leq 0.01$, $^{b}p \leq 0.05$, $^{c}p \leq 0.1$.

Table 6. Distribution of additional parent characteristics by type of move made for families originating in the lowest quintile super-counties, 1990 Census

	Own home	Welfare recipient	# of own children in household	# of children ever born	# of own children < 5
No move (remain in q1)	0.443^{a} (0.004)	0.069 (0.002)	2.068^{a} (0.008)	3.089^{a} (0.012)	0.592^{a} (0.005)
Move: q1 to q1	0.261^{***} (0.011)	0.067 (0.006)	1.872^{***} (0.026)	2.830^{***} (0.037)	0.787^{***} (0.017)
Move: q1 to q2	$0.299^{***,b}$ (0.013)	0.053^{**} (0.007)	1.884^{***} (0.029)	2.865^{***} (0.041)	0.808^{***} (0.019)
Move: q1 to q3	0.288^{***} (0.016)	0.082 (0.008)	1.939^{***} (0.036)	2.841^{***} (0.052)	0.787^{***} (0.025)
Move: q1 to q4	$0.212^{***,b}$ (0.020)	0.080 (0.010)	1.924^{***} (0.044)	2.927^{**} (0.064)	$0.893^{***,a}$ (0.030)
Move: q1 to q5	0.263^{***} (0.027)	0.059 (0.014)	2.064^{a} (0.062)	3.025^{b} (0.089)	0.830^{***} (0.042)

Notes: Each cell in this table reports the share of low-income families that have the characteristic identified in the column's header, limited to only those low-income families that originated in the least upwardly mobile areas, stratified by the type of move made. Compared to those low-income families that do not move, low-income families that move from the least to the most upwardly mobile areas are 35% more likely to be white; 117% more likely to have a college educated parent present and 18% more likely to be a married, two-parent family. All values reported here are estimates of $\beta_{\Delta k,j}{}^{m}$'s and β_j^{nm} 's from Equation (10). All values are based on family-level observations ($N = 23{,}641$) provided by the U.S. Census' 1990 PUMS file (5% sample). This sample is restricted to families whose origin super-county belongs to the 1st quintile of the AIIM distribution. For families that move to different super-counties, the AIIM quintile of their destination super-county is used to determine their 'type' of move. See the caption of Table 1 for additional information on the sample. Asterisks indicate the statistical significance of the difference between the value reported in that cell and the value reported in the 'No Move' cell (i.e., the first row). $^{***}p \le 0.01$, $^{**}p \le 0.05$, $^{*}p \le 0.1$. Letters reported in the superscripts indicate the statistical significance of the difference between the value reported in that cell and the value reported in the lateral move cell (i.e., q1 to q1). $^{a}p \le 0.01$, $^{b}p \le 0.05$, $^{c}p \le 0.1$.

observed for those who originated in the highest AIIM quintile (see Online Appendix Table A7).

The data in Table 6 strongly imply that moves are associated with changes in family circumstances with the change in home ownership being the largest and most intuitive. These changes in family circumstances raise the possibility that estimates in Chetty and Hendren (2017a, 2017b), which are based on a sample of movers, may be confounded by these changing family circumstances. Chetty and Hendren (2017a) assess this possibility by examining whether changes in family income and family structure (marital status) within 1 year of the move confounds estimates and concluded that such changes do not effect estimates. Nevertheless, the change in home ownership and

Table 7. Simulated nuclear family income of persons' ages 27–31 in 2011 within upward mobility quintiles and mover status based on 1990 parent characteristics

Chetty et al.'s absolute intergenerational mobility quintile (mean within quintile)	No county-level controls		County-level controls, 1990		County-level fixed effects	
	[1]		[2]		[3]	
	Nonmover	Mover	Nonmover	Mover	Nonmover	Mover
q1 (36.5)	28,875	32,561	28,875	32,854	28,875	32,589
q2 (39.8)	32,224	34,419	32,401	34,828	31,897	34,204
q3 (42.2)	32,997	34,627	33,111	34,993	32,617	34,395
q4 (44.3)	33,572	34,798	33,395	35,037	33,035	34,508
q5 (48.4)	34,738	35,608	35,148	36,036	34,174	35,280

Notes: Each cell in this table reports the simulated income of an adult child from a low-income family, stratified by their family's mobility status and the upward mobility ranking of their place of origin when they were young. Adult children from low-income families that moved out of the least upwardly mobile areas have predicted nuclear family incomes about 13% greater than those of nonmovers who stayed behind. Simulated incomes are calculated by applying the coefficients of columns [2], [4] and [6] reported in Table 2 to the values reported in Table 4. See notes to Table 4 for more details.

change in number of young children are potentially confounding influences and other unmeasured time-varying factors may have changed too. We note, however, that among movers, the variation in home ownership and number of young children does not vary significantly by the type of move.

Table 7 reports the simulated incomes of both movers and nonmovers within each AIIM quintile. The difference between movers and nonmovers is particularly striking within the least upwardly mobile areas. Adult children of low-income families that moved out of these areas have predicted nuclear family incomes about 13% greater than those of nonmovers who stay behind ($32,589 compared to $28,875), which, at the very least, suggests that the experiences of movers cannot be unconditionally extrapolated onto nonmovers, as these two groups are fundamentally different from one another. This is less the case, however, when comparing the simulated incomes of movers and nonmovers from the most upwardly mobile areas.

Notice that it would also be a mistake to take the experience of the children of low-income movers into a top quintile super-county and use that to estimate the true effect of their destination. This is because, as described in detail in Table 5, families that move from the least to the most upwardly- mobile super-counties tend to exhibit much more favorable income-predicting characteristics when compared to those families that choose not to move from the least upwardly mobile areas. To see this, Table 8 simulates the nuclear family income for children of low-income parents who once resided in the least upwardly mobile super-counties by type of move made.[19] Comparing Tables 3 and 8, we predict that children

19 For Table 8, income in each cell is simulated as $\widehat{ninc}_{\Delta k} = 28,875 + \sum_j \hat{\delta}_j \left(x^{m}_{\Delta k,j} - x^{nm}_{j} \right) \times \text{mover}_i$, where the values for $\hat{\delta}_j$ come from Table 2.

Table 8. Simulated nuclear family incomes by type of move made for persons' ages 27–31 in 2011 originating in the lowest quintile super-counties, 1990 Census

	Simulated income in 2011 based on 1990 parent characteristics		
	No county-level controls [1]	County-level controls, 1990 [2]	County-level fixed effects [3]
No move (remain in q1)	28,875	28,875	28,875
Move: q1 to q1	31,273	31,499	31,432
Move: q1 to q2	32,868	33,211	32,896
Move: q1 to q3	33,623	33,966	33,521
Move: q1 to q4	33,398	33,699	33,365
Move: q1 to q5	33,788	34,064	33,573

Notes: Each cell in this table reports the simulated income of an adult child from a low-income family stratified by the type of move their family made when they were young, limited only to those families that originated in the least upwardly mobile areas. Simulated incomes are substantially higher for those whose families move to the most upwardly mobile places. Simulated incomes are calculated by applying the coefficients of columns [2], [4] and [6] reported in Table 2 to the values reported in Table 5.

of low-income movers from the least to the most upwardly-mobile super-counties have expected nuclear incomes nearly identical to those for children of low-income parents currently residing in top quintile super-counties. This suggests that, among those low-income parents who started out in the least mobile super-counties, those that choose to move to the most upwardly mobile areas are very similar, in terms of their income-predicting characteristics, to the average parent residing in these areas (movers and nonmovers alike). However, the children of low-income parents who choose not to move from the least upwardly-mobile super-counties have expected incomes about 18% less than children of nonmover, low-income parents residing in the top quintile super-counties, suggesting that these two groups differ considerably in their income-predicting characteristics.

Lastly, turning to heterogeneity within movers, we note that those that move from the least upwardly mobile areas to the most upwardly mobile areas have a simulated income that is about 7% higher than that for those who make lateral moves from the least upwardly mobile areas (column 3). Again, this suggest that families making 'better' moves possess relatively favorable predictors of income and that the type of move made cannot be viewed as having been randomly assigned across families. This raises the possibility that the causal estimates of Chetty and Hendren (2017a, 2017b), which are derived from observing movers' outcomes, may be confounded by unobserved attributes that vary across families making different types of moves and, therefore, lack internal validity.

6. Conclusion

The descriptive, and previously unknown facts about the geographic variation in intergenerational mobility, documented by Chetty et al. (2014), and the causal estimates of the effect of place on intergenerational mobility reported in Chetty and Hendren (2017a, 2017b) are extremely important. They shine a light on a child's place of residence and the institutional features of those places as a potentially important source of lifetime well-being. However, given the evidence we presented, it seems premature to suggest that families should use estimates in Chetty and Hendren (2017b) to guide their choices about where to live, as the authors suggested. We find that much of the differences documented by Chetty et al. (2014) are arguably not place differences at all, but people differences. Indeed, a very limited set of people differences explain most of the place differences in intergenerational income mobility.

Specifically, we show that earnings predicted from a relatively few characteristics of low-income parental households generates simulated incomes for adult children that account for approximately 40–100% of the inter-quintile differences reported in Chetty et al. (2014). A large portion of the spatial pattern of upward mobility can be generated without reference to space. It seems reasonable to conclude that differences between places in intergenerational mobility would be even further reduced, perhaps to zero, with the addition of more family characteristics. We also show that low-income movers are a very different group than low-income nonmovers, which raises a question about the external validity of the more compelling causal estimates in Chetty and Hendren (2017a, 2017b). Indeed, there are also time-varying changes in family circumstances associated with moving that may bias estimates in Chetty and Hendren (2017a, 2017b).

The intuition that certain aspects of place matter for children's development and future success is strong, and perhaps most clearly reflected in families' locational decisions vis-à-vis school quality (Hoxby, 2003). However, the 'place' in this fundamental family decision is the school district, which may differ from the 'place' where parents work, and differ from the 'place' that sets public safety policy. A family may simultaneously access the institutions and amenities that affect children's well-being of several different, often geographically unique, 'places'. Notably, the research of Chetty et al. (2014) and Chetty and Hendren (2017a, 2017b) is not based on a well-specified conceptual model linking place to proximate causes of child development and adult well-being, for example, as in Galster and Killen (1995). Here is the main justification from Chetty et al. (2014):

> One way to conceptualize the choice of a geographical partition is using a hierarchical model in which children's outcomes depend on conditions in their immediate neighborhood (such as peers or resources in their city block), local community (such as the quality of schools in their county), and broader metro area (such as local labor market conditions). To fully characterize the geography of intergenerational mobility, one would ideally estimate all of the components of such a hierarchical model. (1586)
>
> As a first step toward this goal, we characterize intergenerational mobility at the level of commuting zones. CZs are aggregations of counties based on commuting patterns in the 1990 census.... CZs are designed to span the area in which people live and work, they provide a natural starting point as the coarsest partition of areas. (1586)

As noted by the Chetty et al. (2014), commuting zones (or counties), which rarely organize school districts, police departments, social services and other community influences that may affect children's development and their future success are distal causes of children's success.[20] Counties and commuting zones are most closely related

20 Chetty et al. (2014) argue that using the broader geographic areas for the analysis lessens concerns about sorting that could confound estimates. However, this concern with endogenous sorting still applies at the

to economic activity that may influence employment and wage opportunities that affect children's development and future success. There is also evidence that residential racial segregation at broader geographic levels adversely affect minority children's outcomes (Cutler and Glaeser, 1997; Card and Rothstein, 2009; Ananat, 2011).

Within any county or commuting zone there is often wide variation in school quality, public safety and other potential influences on child development and future success. Therefore, finding that intergenerational income mobility differs by commuting zone or county should be viewed skeptically from a causal perspective because the theoretical premises and plausibility of the investigation were not well established. There does not appear to be a prior literature suggesting that institutions, or policies, at the level of commuting zone, or county, would be particularly important to intergenerational income mobility. Of course, scientific inquiry sometimes makes discoveries incrementally, and the data, study and findings in Chetty et al. (2014) and Chetty and Hendren (2017a, 2017b) are novel.

There is considerable heterogeneity in family and neighborhood characteristics within counties that underscore the potential disconnect between a plausible conceptual model and the analysis of Chetty et al. (2014). To illustrate the extent of this variation, we selected the largest county in each of the five quintiles of AIIM. These counties are: Cook, IL (lowest quintile), Maricopa, AZ, Harris, TX, Los Angeles, CA and Orange, CA (highest quintile). For each county, we repeated the above exercise, but using PUMA as the geography of interest. We constructed the mean, predicted adult income for children of low-income families in each PUMA in those five counties. Table 9 reports the predicted adult incomes based on family characteristics.

As Table 9 suggests, there is considerable variation in family characteristics and predicted adult incomes within each of the five counties except for Orange County, CA. In Cook County, IL, predicted adult incomes (net of county fixed effects) range from \$19,803 to \$36,498 and there are several PUMAs in Cook County, IL, which is in the lowest quintile of AIIM, with predicted adult incomes greater than the predicted adult income of the lowest PUMAs in Orange County, CA, which is in the top quintile of AIIM. Similarly, in Los Angeles County, CA predicted adult incomes range from \$26,766 to \$38,967. The variation in family characteristics and predicted adult incomes within counties matches intuition that there is considerable neighborhood segregation by race, education and family structure within counties (holding income constant). There is also considerable variation in amenities and public goods within counties. For example, there are 23 independent school districts in Harris County, TX and 46 municipal police departments in Los Angeles County, CA.

The variation documented in Table 9 also bears directly on the exploratory analyses of Chetty et al. (2014) that attempt to identify factors that explain geographic variation in AIIM. Chetty et al. (2014) obtained associations between AIIM and racial segregation, income segregation (inequality), school quality, commuting patterns and family structure. However, what does average school quality measure in Harris County, TX when there are 23 independent school districts? Similarly, what do commuting patterns measure in Los Angeles County, CA? With the type of within county

broader level of geography, as we demonstrate, and the analysis in Chetty et al. (2014) is purely descriptive, as acknowledged by the authors. Thus, the justification for using the larger geographical units is not strong.

Table 9. Simulated incomes by PUMA reported for the largest county within each AIIM quintile

PUMA rank	County				
	Cook, IL	Maricopa, AZ	Harris, TX	Los Angeles, CA	Orange, CA
1	19,803	28,335	26,423	26,766	35,083
2	20,047	29,299	30,064	27,417	35,578
3	20,232	30,330	30,096	27,647	35,584
4	21,048	30,679	30,700	27,883	35,609
5	21,363	30,962	30, 885	28,624	35,870
6	23,551	31,299	31,428	29,000	36,006
7	24,531	31,486	32,320	30,598	36,510
8	26,972	32,005	33,085	31,234	36,877
9	27,588	32,262	33, 312	31,371	36,983
10	28,286	32,575	33,430	31,396	37,447
11	28,875	32,703	33,644	31,971	37,856
12	29,994	33,117	33,747	32,143	38,207
13	30,095	33,288	34,129	32,474	38,832
14	30,248	33,379	34,322	32,555	38,850
15	30,692	33,550	34,472	32,888	
16	30,755	34,141	35,359	32,998	
17	30,798		35,705	33,380	
18	30,842		36,056	33,418	
19	31,243		36,300	33,454	
20	31,882		36,397	33,656	
21	32,420		36,471	33,696	
22	32,753		36,787	33,708	
23	32,868		36,958	33,736	
24	33,589		37,224	33,768	
25	34,389		38,284	33,913	
26	34,998			33,934	
27	35,162			33,965	
28	35,269			34,000	
29	35,508			34,012	
30	35,609			34,099	
31	35,663			34,231	
32	36,154			34,292	
33	36,498			34,309	
34				34,395	
35				34,623	
36				34,690	
37				34,761	
38				34,792	
39				35,020	
40				35,029	
41				35,183	
42				35,661	
43				35,794	
44				36,066	
45				36,172	
46				36,250	
47				36,483	
48				36,551	
49				36,587	

(continued)

Table 9. Continued

PUMA rank	County				
	Cook, IL	Maricopa, AZ	Harris, TX	Los Angeles, CA	Orange, CA
50				36,605	
51				36,637	
52				36,796	
53				36,894	
54				36,991	
55				37,197	
56				37,555	
57				38,224	
58				38,967	

Notes: Simulated incomes for each PUMA are calculated using methods similar to those applied in Table 3, column [2]. Here, the share of families within the 3rd income decile sharing a particular characteristic is measured for each PUMA, not for a broad AIIM quintile. Coefficients from Column [2] of Table 2 are combined with these values. Each county is anchored to its position along the nuclear family income distribution taken from the 2011 ACS as predicted by its AIIM score taken from Chetty et al. (2014).

(commuting zone) variation that is common, the average characteristic of a county (commuting zone) is a poor measure of the underlying causal mechanism that affects AIIM. Notably, the results of this exploratory analysis in Chetty et al. (2014) suggested that family structure and commuting patterns explain most of the variation in AIIM. While commuting patterns may reflect some place-based policy that affects child development, although which policies is not obvious, family structure is clearly not caused by place-based policies. Therefore, it is notable that this family characteristic explains most of the variation in intergenerational income mobility, consistent with the findings we showed earlier.

Overall, the lack of a plausible conceptual model linking commuting zones, or counties, to proximate causes of child development and adult success is an *a priori* reason to be skeptical of the causal possibilities of the Chetty et al. (2014) line of inquiry.[21] A legitimate question is whether the 'facts' presented by Chetty et al. (2014) should be something future research investigates. While Chetty and Hendren (2017a, 2017b) provide credible evidence of causal effects of commuting zones on intergenerational mobility, the external validity of this evidence is debatable and there is evidence that the internal validity of these estimates may not hold. Families that move are different and there is no way of knowing whether similar moves by stayers would result in the same consequences (Cartwright, 2011, 2012, 2013). The arguably weak theoretical premise of the Chetty et al. (2014) study combined with the substantial evidence of

21 An arguably more promising approach to these questions is suggested in Chetty et al. (2016). In this study, the authors examine whether neighborhoods affected adult well-being among participants in the Moving to Opportunity randomized experiment. In our view, and in a large literature (e.g., Wilson, 1996; Rosenbaum et al., 2002; Kling et al., 2007), the geography of neighborhoods is much more compelling unit of analysis conceptually than the geography of counties. However, external validity of the findings may again be an issue, as the experiment was conducted in only a few cities, only 40–48% of the children in 'winning' families actually took up the offered vouchers for moving to better neighborhoods, and movers were different from nonmovers (Kling et al., 2007; Chetty et al., 2016).

significant differences in family characteristics between counties and between movers and nonmovers that we presented raises questions about the usefulness and interpretation of the evidence of the research of Chetty and colleagues.

Supplementary material

Supplementary data for this paper are available at *Journal of Economic Geography* online.

References

Ananat, E. O. (2011) The wrong side(s) of the tracts: the causal effects of racial segregation on urban poverty and inequality. *American Economic Journal: Applied Economics*, 3: 34–66.

Bartik, T. J. (1991) Boon or boondoggle? The debate over state and local economic development policies. *Who Benefits from State and Local Economic Development Policies*?, pp. 1–16. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Bartik, T. J. (2003) Local economic development policies. Upjohn Institute Working Paper No. 03-91. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. https://doi.org/10.17848/wp03-91 [Accessed 14 May 2018].

Card, D., Rothstein, J. (2009) Racial segregation and the black-white test score gap. *Journal of Public Economics*, 91: 2158–2184.

Cartwright, N. (2011) Evidence, external validity and explanatory relevance. In J. M. Gregory (ed.) *Philosophy of Science Matters: The Philosophy of Peter Achinstein*, pp. 15–28. New York: Oxford University Press.

Cartwright, N. (2012) RCT's, evidence and predicting policy effectiveness. In H. Kincaid (ed.) *The Oxford Handbook of the Philosophy of the Social Sciences*, pp. 298–318. New York: Oxford University Press..

Cartwright, N. (2013) Knowing what we are talking about: why evidence doesn't always travel. *Evidence and Policy* 9: 97–112.

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

Chetty, R., Hendren, N. (2015) The impacts of neighborhoods on intergenerational mobility: childhood exposure effects and county-level estimates. Unpublished paper, May 2015.

Chetty, R., Hendren, N. (2017a) The effects of neighborhoods on intergenerational mobility I: childhood exposure effects. *Quarterly Journal of Economics*. In press, http://www.equality-of-opportunity.org/assets/documents/movers_paper1.pdf

Chetty, R., Hendren, N. (2017b) The effects of neighborhoods on intergenerational mobility II: county-level estimates. *Quarterly Journal of Economics*. In press, http://www.equality-of-opportunity.org/assets/documents/movers_paper2.pdf

Chetty, R., Hendren, N., Katz, L. (2016) The effects of exposure to better neighborhoods on children: new evidence from the moving to opportunity experiment. *American Economic Review*, 106: 855–902.

Crowder, K., South, S. J. (2005) Race, class, and changing patterns of migration between poor and nonpoor neighborhoods. *American Journal of Sociology* 110: 1715–1763.

Cutler, D. M., Glaeser, E. L. (1997) Are ghettos bad? *Quarterly Journal of Economics*, 112: 827–872.

Galster, G. C., Killen, S. P. (1995) The geography of metropolitan opportunity: a reconnaissance and conceptual framework. *Housing Policy Debate*, 6: 1, 7–43

Greenwood, M. J. (1969) An analysis of the determinants of geographic mobility in the United States. *The Review of Economics and Statistics*, 51: 189–194.

Greenwood, M. J. (1997) Internal migration in developed countries. In M. R. Rosenzweig and O. Stark (eds) *Handbook of Population and Family Economics*. Vol 1B. New York: Elsevier.

Heckman, J. J., Urzua, S., Vytlacil. E. (2006) Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics*, 88: 389–432.
Hoxby, C. M. (ed.) (2003) *The Economics of School Choice*. Chicago: University of Chicago Press.
Kain, J. F., Persky, J. J. (1969) Alternatives to the gilded ghetto. *The Public Interest* (Winter), 74–87.
Kline, P., Moretti, E. (2014) People, places and public policy: some simple welfare economics of local economic development programs. *Annual Review of Economics*, 6: 629–662.
Kling, J. R., Liebman, J. B., Katz, L. F. (2007) Experimental analysis of neighborhood effects. *Econometrica*, 75: 83–119.
Mincer, J. (1978) Family migration decisions. *Journal of Political Economy*, 86: 749–773.
Molloy, R., Smith C. L., Wozniak, A. (2011) Internal migration in the United States. *Journal of Economic Perspectives*, 25: 173–196.
Rosenbaum, J., Reynolds, L., Deluca, S. (2002) How do places matter? The geography of opportunity, self-efficacy and a look inside the black box of residential mobility. *Housing Studies*, 17: 71–82.
Ruggles, S., Genadek, K., Goeken, R., Grover, J., Sobek, M. (2015) *Integrated Public Use Microdata Series: Version 6.0 [Dataset]*. Minneapolis: University of Minnesota. http://doi.org/10.18128/D010.V6.0.
Sjaastad L. A. (1962) The costs and returns of human migration. *Journal of Political Economy*, 70: 80–93.
South, S. J., Pais, J., Crowder, K. (2011) Metropolitan influences on migration into poor and nonpoor neighborhoods. *Social Science Research* 40: 950–964.
Wilson, W. J. (1996) *When Work Disappears*. New York: Vintage.