# Inequality of Educational Opportunity? Schools as Mediators of the Intergenerational Transmission of Income

Jesse Rothstein*

University of California, Berkeley, and NBER

October 2016

PRELIMINARY AND INCOMPLETE

**Abstract**

Chetty et al. (2014) document wide geographic variation in intergenerational income mobility, with children from low-income families achieving much better outcomes, relative to their neighbors from higher-income families, in Salt Lake City and Los Angeles than in Cincinnati. A plausible mechanism is school quality. I use data from several national surveys to investigate whether commuting zones (CZs) where parents' incomes are strongly related to children's incomes also exhibit strong relationships of parental income to measures of children's human capital accumulation, such as educational attainment, test scores, and non-cognitive skills. CZs with more income mobility indeed tend to have somewhat higher mobility as measured by children's educational attainment. They also tend to have lower returns to education. By contrast, CZ income mobility does not appear correlated with non-cognitive returns to parental income, and the correlation with children's achievement is small and roughly the same when achievement is measured in Kindergarten as when it is measured in high school. There is thus little evidence that differences in the quality of K-12 schooling are a key mechanism driving variation in intergenerational mobility. Access to college may play a larger role, but is also not a dominant factor. The largest part of the variation derives from direct impacts of parental income on children's income controlling for human capital. This

*E-mail: rothstein@berkeley.edu. I am grateful to Audrey Tiew, Leah Shiferaw, and Julien Lafortune for excellent research assistance, and to David Card and Avi Feller for helpful discussions. None are responsible for the contents.

points to job networks or the structure of the local labor market as plausible mechanisms.

# 1 Introduction

Social scientists have long looked to the intergenerational transmission of income – the strength of the association between an adult's income and that of his or her parents – as a key dimension of social inequality. The stronger is this association, the less likely it is that a child born into a disadvantaged family will succeed economically as an adult, and the further is a society from equality of opportunity among children. The salience of intergenerational transmission has grown with the rise in income inequality, which makes it harder for families of modest incomes to keep up in the educational investment arms race (Chetty et al., 2014). Indeed, Reardon (2011) has shown that the gap in test scores between students born to families in the top and bottom of the income distribution has grown in recent years as the gap in incomes has widened. Although we will not be able to observe the adult outcomes of recent cohorts of children for many years, Reardon's evidence at least suggests that economic mobility is likely to have declined.

The measurement of intergenerational income transmission is the subject of long literatures spanning a number of social science disciplines.[1] Nevertheless, little is understood about the channels by which this transmission occurs. Candidates include differences in parenting practices between high- and low-income families, differences in explicit investments in children's education, differences in access to educational or other public institutions, and labor market institutions (such as insider hiring, or spatial mismatch) that advantage children from high-income families regardless of their skills. While there is evidence in the literature for each of these explanations, we know little about their relative importance.

Important new research by Raj Chetty and co-authors (Chetty et al., 2014, hereafter "CHKS") offers a path forward. CHKS use data on the universe of U.S. tax filers to measure intergenerational income links more precisely than has been possible previously, and reveal

[1]Some literatures focus on other dimensions of intergenerational transmission, such as transmission of occupational status. As data on incomes has improved, and as inequality of incomes even within narrowly defined occupations has risen, the importance of income transmission *per se* has risen.

massive heterogeneity across space: The gap in adult earnings of children from high- vs. low-income families is nearly twice as large for children who grow up in Cincinnati as for those who grow up in Los Angeles.

This heterogeneity presents an opportunity: It should be possible to learn about the differences among areas that account for the differences in income transmission, and perhaps to use this understanding to craft policies that raise mobility in low opportunity areas at least to levels seen in cities that offer more opportunity. Most economic and educational institutions are similar across the U.S., so there is more prospect of understanding which remaining factors – and in particular which policy choices – contribute to differences in outcomes than there would be in cross-country studies (e.g., Bradbury et al., 2015).

This paper assesses whether geographic areas with high intergenerational transmission of income – strong associations between parental and child incomes – also show high degrees of transmission of parental income into children's educational achievement and attainment. We would expect the two to be strongly correlated across space if human capital accumulation is an important mechanism by which one generation's advantage is transmitted into the next generation; on the other hand, if parental income primarily helps children by, for example, buying them access to better labor market networks, then areas where poor children do relatively well in the labor market may not be areas where those children do relatively well in school.

I also investigate the ages at which gaps in child outcomes appear. In the simple case where skill is uni-dimensional and is reflected both in children's achievement and adults' earnings, the age profile of the gap in achievement between children from high- and low-income families is indicative of the ages at which the relevant mechanisms operate. In more complex (and more realistic) models, the interpretation is not so straightforward, but it would nevertheless be useful to understand when, and in which types of outcomes, gaps arise. This would point to institutional factors likely to contribute to intergenerational transmission of income, and provide useful directions for future research.[2] For example,

[2]Evidence on the developmental profile of family income gaps would also inform inform theories of child development such as Heckman's "skill-begets-skill" model (see, e.g., Carneiro and Heckman, 2003; Cunha and Heckman, 2007, 2008; Cunha et al., 2010 ), which posits that early investments are the key to closing gaps in eventual outcomes.

if in areas with high income transmission, gaps between high- and low-income children in test scores and other measures of child development are small at school entry but large at school exit, this would suggest that educational institutions are central to intergenerational transmission of advantage; on the other hand, if gaps are as large in Kindergarten as in adult outcomes, this would point away from schools and toward early childhood environments and services (e.g., prenatal and postnatal health care) as more likely contributory factors.

I rely on three panel surveys conducted by the National Center for Educational Statistics (NCES): the Education Longitudinal Survey (ELS), the Early Childhood Longitudinal Survey (ECLS), and the High School Longitudinal Survey (HSLS). Each is a representative national sample with information on parental income and children's achievement (test scores) at various ages. The ELS, which begins with older children, follows respondents long enough to measure attainment (years of schooling) and adult income as well. Importantly, versions of each sample are available that can be geocoded to commuting zones (CZs).

CHKS's measurement of income transmission for nearly all of the 741 commuting zones (CZs) in the United States was made possible by the availability of population income data, from tax records. The NCES samples are all much smaller, with about 15,000 respondents each, and do not permit the construction of income-achievement transmission measures for individual CZs. I show below, however, that this is not necessary in order to accomplish the more limited goal of measuring the across-CZ relationship between income-income transmission and income-achievement transmission. That is identified even with small numbers of observations from each CZ – essentially, one can pool information from many CZs with similar income-income transmission to identify the average income-achievement transmission among them, even when the latter is not reliably estimated for any individual CZ. I develop an estimator for this, based on a mixed (random coefficients) model for the relationship between parental income and children's achievement. This yields an estimate of the "reverse" regression of income-achievement transmission on the known income-income transmission, which can then be transformed into the "forward" regression of interest.

I find that the strength of intergenerational income transmission in a CZ is reasonably strongly related to the strength of transmission from parental income to children's educational attainment. This reproduces a similar result for college enrollment in CHKS. Income transmission is much less strongly related, however, to the transmission from parental income to children's achievement at the end of secondary school, and shows no relationship to various measures of children's non-cognitive skills (though there is some evidence that income transmission is related to *teachers' assessments* of non-cognitive skills). Moreover, the association between income-income transmission and income-achievement transmission is approximately as strong when achievement is measured early in elementary school as when it is measured in 12th grade. This is strongly suggestive that differential inequities in access to good elementary and secondary schools are not an important mechanism driving the across-CZ variation in income transmission. Educational attainment appears to play more of a role than does achievement, suggesting that access to higher education, rather than differences in the likelihood of having the skills needed to succeed in college, may be an important mechanism.

I also consider variation in the CZ-level labor market return to skill. Because children from low-income families complete less education (and acquire fewer skills) in every CZ than do children from higher-income families, differences in the return to skill could produce differences in income transmission even if the distribution of skill acquisition were the same everywhere. Indeed, I find that the return to education varies substantially across CZs, and is modestly correlated with the strength of income transmission in the CZ. This points to labor market institutions as a potentially important factor.

When I use a Mincer wage regression to decompose intergenerational transmission into impacts of parental income on child skill, the returns to child skill, and the residual component of children's income, I find that skill accumulation – including both achievement and attainment – accounts for only about 15-20% of the relative disadvantage of children from low-income families in high transmission CZs, while returns to skill account for about 30%. The remainder operates through the Mincer income residual. In other words, about half of the variation in intergenerational income transmission across CZs appears to be due to variation in the relationship between parental income and child income conditional on children's (observed) human capital.

It is important to emphasize that my results are purely observational; my estimates of the association between CZ-level income transmission and CZ-level transmission of parental income to children's test scores could be confounded by other CZ-level characteristics that are correlated with both.[3] Keeping this caveat in mind, my results indicate that human capital plays a relatively small role in the geographic variation in the intergenerational transmission of income. Much of this variation appears to reflect differences in adult incomes of children with similar skills, perhaps due to labor market institutions (e.g., unions, or other determinants of residual income inequality), differences in access to good jobs (due, perhaps, to labor market networks or socially stratified labor markets), or differences in marriage markets (which influence CHKS's intergenerational income transmission because children's income is measured at the family level). While this does not rule out an important role for educational interventions in raising mobility, it suggests that we should at least be considering interventions in these other domains as well.

## 2 Framework

CHKS use tax data to construct various measures of intergenerational income mobility. I focus on what they call "relative mobility," the relative advantage that a child from a high-income family has, relative to a child from a low-income family in the same CZ, in achieving a high income as an adult. Letting $p_{ic}$ represent the income of the parents of individual $i$ in CZ $c$, measured in national percentiles, and $y_{ic}$ the child's income, again in national percentiles, CHKS's preferred relative mobility measure is the slope coefficient $\theta_c$ from a CZ-level bivariate regression:

$$y_{ic} = \alpha_c + p_{ic}\theta_c + e_{ic}. \tag{1}$$

CHKS have sufficient data to estimate $\theta_c$ extremely precisely without pooling information across CZs. Thus, they estimate that $\theta_c = 0.43$ in Cincinnati, meaning that each percentile increase in parental income is associated with a 0.43 percentile increase in children's eventual income, on average, in that city, and that in Los Angeles $\theta_c = 0.23$, implying a relationship between parent and child income that is only a bit more than half as strong as in Cincinnati. Hereafter, I refer $\theta_c$ as the strength of *income transmission* in the CZ, deviating from CHKS's

[3] Reverse causality is also possible: For example, CZs with more equal labor markets may make it easier to attract high quality graduates into teaching, leading to a causal path from economic mobility to gaps in children's outcomes.

"relative mobility" name to make clear that higher values of the slope correspond to *less*, rather than more, mobility for low-income children relative to high-income children.

Across all CZs, CHKS find substantial variation in transmission: While the (unweighted) average CZ has a slope measure of 0.33 – indicating intergenerational mean reversion (in percentiles) of about two-thirds – the standard deviation is 0.065. They find that income transmission is positively correlated, across CZs, with the fraction black in the local population, with racial and economic segregation, and with income inequality. CHKS also examine correlations with various policy measures, such as proxies for school quality. They find that intergenerational income transmission is negatively correlated with average test scores and high school completion, as well as with school expenditures, but is essentially unrelated to average class size. But these are merely across-CZ correlations; CHKS are unable to investigate the role played by *differences* in access to school quality between high- and low-income students.

Moreover, both demographic and policy correlates are of limited value in identifying the channels responsible for differences across areas. The demographic correlates, for example, could indicate that segmented labor markets are an important factor, or they could reflect differences in the degree of stratification in the health or education systems, or differences in the pervasiveness of "poverty cultures." Another possibility is that local policies may be consequences, rather than causes, of either the area's demographic composition or its intergenerational transmission itself. For example, support for school spending may be higher in places with less economic stratification.

In this paper I analyze the channels by which income is transmitted across generations, with the goal of shedding light on the relevant mechanisms. For example, if school quality is a mechanism behind the geographic variation in income transmission, we would expect that CZs with high $\theta_c$ would also tend to be CZs in which the gap in educational outcomes between high- and low-income children is larger, while the gap in child incomes conditional on educational outcomes should be much smaller than the unconditional gap. Moreover, the timing with which any educational outcome gap emerges and grows over the child's development is informative about the particular mechanisms at work.

## 2.1 Test scores as mediators of intergenerational income effects

In this subsection I develop a simple model of intergenerational transmission, focusing on childhood outcomes as observable mediators of this process. For simplicity, I assume that child outcomes $s_{ict}$ (for "skills") for student $i$ in city $c$ are measured at two points, first at or around school entry ($t = 1$) and then again at school exit ($t = 2$). I also assume that skill (human capital) is uni-dimensional and measured perfectly at each stage. The framework can readily accommodate additional time points as well as multiple dimensions of child outcomes (e.g. achievement as well as non-cognitive skill).

Children's outcomes at period $t = 1$ depend on their parents' income, as mediated by local conditions and institutions (including such factors as health care and early childhood systems as well as local culture): $s_{ic1} = f_{1c}(p_{ic})$. Outcomes at period 2 depend on the earlier outcomes as well as on subsequent inputs that may themselves depend on parental income, again as mediated by local conditions (including school quality): $s_{ic2} = f_{2c}(s_{ic1}, p_{ic})$. Finally, the adult income of child $i$ depends on the child's skill in period 2. This is of course influenced by parental income, which may also have a direct effect on the child's income as well: $y_{ic} = g_c(s_{ic2}, p_{ic})$.[4]

Figure 1 displays this framework graphically. It shows that there are several channels by which parental income influences children's income. Algebraically, we can write the reduced-form relationship as:

$$y_{ic} \equiv h_c(p_{ic}) \equiv g_c(f_{2c}(f_{1c}(p_{ic}), p_{ic}), p_{ic}). \tag{2}$$

CHKS's relative mobility measure (i.e., income transmission) is the (linearized) slope of this relationship in the CZ:

$$\theta_c \equiv \frac{dh_c}{dp_{ic}} = \frac{\partial g_c}{\partial s_2}\frac{\partial f_{2c}}{\partial s_1}\frac{\partial f_{1c}}{\partial p_{ic}} + \frac{\partial g_c}{\partial s_2}\frac{\partial f_{2c}}{\partial p_{ic}} + \frac{\partial g_c}{\partial p_{ic}}. \tag{3}$$

The three terms here represent three different channels, and implicate different mechanisms. The first reflects impacts of parental income on children's period-1 skill, multiplied by the

[4] I assume here that early achievement $s_{ic1}$ affects labor market outcomes $y_{ic}$ only through later achievement $s_{ic2}$.

effect of period-1 skill on later outcomes; the second reflects impacts of parental income on skill in period 2 conditional on skill in period 1; and the third represents direct effects of parental income on children's income conditional on period-2 skill. A large role for the first would point to early childhood institutions and parenting practices as likely mechanisms behind income transmission; the second to educational institutions and parental investment in school-aged children; and the third to labor market institutions such as networks and pay norms.

It is useful to assume that each of the $f_1$, $f_2$, and $g$ functions is linear, with additive error terms deriving from inputs to skill accumulation that are orthogonal to parental income:

$$s_{ic1} = f_{1c}(p_{ic}) = \kappa_{1c} + p_{ic}\pi_{1c} + u_{ic1} \tag{4a}$$

$$s_{ic2} = f_{2c}(s_{ic1}, p_{ic}) = \kappa_{2c} + s_{ic1}\lambda_{2c} + p_{ic}\pi_{2c} + u_{ic2} \tag{4b}$$

$$y_{ic} = g_c(s_{ic2}, p_{ic}) = \kappa_{3c} + s_{ic2}\lambda_{3c} + p_{ic}\pi_{3c} + \epsilon_{ic}. \tag{4c}$$

Then we can write the reduced-form relationship between parental income and children's income as

$$\begin{aligned} h_c(p_{ic}) &= \kappa_{3c} + (\kappa_{2c} + (\kappa_{1c} + p_{ic}\pi_{1c} + u_{ic1})\lambda_{2c} + p_{ic}\pi_{2c} + u_{ic2})\lambda_{3c} + p_{ic}\pi_{3c} + \epsilon_{ic} \qquad (5) \\ &= (\kappa_{3c} + (\kappa_{2c} + \kappa_{1c}\lambda_{2c}))\lambda_{3c} + p_{ic}((\pi_{1c}\lambda_{2c} + \pi_{2c})\lambda_{3c} + \pi_{3c}) + ((u_{ic1}\lambda_{2c} + u_{ic2})\lambda_{3c} + \epsilon_{ic}), \end{aligned}$$

and income transmission as:[5]

$$\theta_c = \frac{dh_c}{dp_{ic}} = \lambda_{3c}\lambda_{2c}\pi_{1c} + \lambda_{3c}\pi_{2c} + \pi_{3c}. \tag{6}$$

With sufficient data, it would be possible to estimate each of the transmission coefficients $\Omega_c \equiv \{\pi_{1c}, \pi_{2c}, \pi_{3c}, \lambda_{2c}, \lambda_{3c}, \theta_c\}$ separately for each CZ. But this would require large representative samples in each CZ with measures not only of parental and child income (observed in CHKS's data) but also of children's intermediate developmental outcomes $s_{ic1}$ and $s_{ic2}$.

[5] CHKS present evidence that the parent income-child income relationship, in percentile ranks, is approximately linear in each CZ, as in (6).

In lieu of that, I focus on understanding the distribution of $\Omega_c$ across CZs, and in particular the covariance and correlation between $\theta_c$ and the other elements of $\Omega_c$.

CHKS present some analysis of this. They can measure college attendance in their tax data, and thus can measure the CZ-level transmission of parental income into college attendance. In my framework, college attendance can be seen as the post-schooling skill measure $s_{ic2}$, and the college transmission coefficient is thus $\theta'_c \equiv \lambda_{2c}\pi_{1c} + \pi_{2c}$. CHKS find that $\theta'_c$ is quite variable across CZs, just as is $\theta_c$, and that the two are highly correlated ($\rho = 0.68$). [6]

## 2.2 Exploiting and interpreting cross-CZ variation

Bradbury et al. (2015) estimate a system of equations similar to (4a) and (4b) at the national level. They find that reduced-form achievement gaps are roughly stable across ages (i.e., that $\pi_{1c}$ is of comparable magnitude to $\tilde{\pi}_{2c} \equiv \pi_{1c}\lambda_{2c} + \pi_{2c}$), but that there is a sizable income gap in later achievement conditional on earlier achievement (i.e., that $\pi_{2c}$ is not small). These are both possible because $\lambda_{2c}$ is relatively small – there is substantial mean reversion between earlier and later achievement. Bradbury et al. interpret the $\pi_{2c}$ result as evidence that post-Kindergarten investments account for an important share of the intergenerational transmission of parental income to children's achievement.[7]

But there are a number of complications with interpreting mobility measures computed from national samples. One is that the measured transmission of parental income to children's achievement is likely to be quite sensitive to the quality of the achievement measures. If, for example, a particular measure is directly related to parental income conditional on the child's actual skill at age $t < 3$, this will lead decompositions like that outlined above to overstate the importance of parental investments prior to $t$ and understate the impor-

[6]Although CHKS do not discuss this, the magnitude of the variation in $\theta'_c$ across CZs is not large enough to be an important mechanism for intergenerational income transmission. Note that $\theta_c = \lambda_{3c}\theta'_c + \pi_{3c}$, and that $\lambda_{3c}$ is the slope of children's adult income (measured in percentiles) with respect to college attendance. In data from the American Community Survey, pooling all CZs, those with some college or more have incomes about 19.2 percentile points higher than those without college, on average. (I discuss the sample here below.) Taking this as an estimate of $\lambda_{3c}$, a one standard deviation increase in $\theta'_c$, 0.0011, would drive only a 0.02 increase in $\theta_c$, or less than one-third of a standard deviation. This back-of-the-envelope calculation thus implies that the key mechanisms must operate through $\pi_{3c}$. My more detailed analyses with richer intermediate skill measures, below, confirm this conclusion.

[7]Bradbury et al. (2015) also compare results across four English-speaking countries, but measures are sufficiently different across measures to complicate interpretation.

tance of post-$t$ investments. This is not just a theoretical possibility. Many standardized tests, for example the SAT college entrance test, have been found to load too strongly onto family background relative to their information about students' human capital (see, e.g., Rothstein, 2004). Another concern is that differences in the measurement properties of the data sources used to construct each of the elements of the decomposition may confound the analysis. For example, data sources may differ in the degree of measurement error in family income (Rothstein and Wozny, 2013) or in the scaling of intermediate child outcome measures (Jacob and Rothstein, 2016; Bond and Lang, 2013; Nielsen, 2015).

Comparisons across CZs, using the same data sources and measures for each, can reduce this problem. So long as systematic or random measurement error or scaling problems are constant across CZs, they are unlikely to have much impact on between-CZ differences in these the transmission coefficients $\Omega_c$. Consider the decomposition of income mobility $\theta_c$ into the component reflecting achievement at school entry, $\pi_{1c}$, and a residual component reflecting post-school-entry investments. As noted above, this is $\theta_c = \lambda_{3c}\lambda_{2c}\pi_{1c} + \lambda_{3c}\pi_{2c} + \pi_{3c}$. I will assume for the time being that the direct effects of earlier achievement on later outcomes (i.e., $\lambda_{2c}$ and $\lambda_{3c}$) are constant across CZs, though I loosen this later.[8] In general, we would expect that institutions that create a high $\pi_{1c}$ – for example, the absence of high quality publicly provided childcare for pre-school-age children – would translate into a high value of $\theta_c$ (at least insofar as early achievement is important to eventual earnings – that is, if $\lambda_2\lambda_3$ is large). But the same may not be true in reverse: Institutions that raise $\theta_c$ may not raise $\pi_{1c}$, insofar as they operate on $\pi_{2c}$ or $\pi_{3c}$.

CHKS assess the importance of institutions to the transmission of inequality by comparing $\theta_c$ across CZs with different observed institutions. This observational analysis may be misleading relative to the causal effects of the particular institutions examined. This is of particular concern because the dependent variable $\theta_c$ is so far removed from the channels by which the institutions (e.g., primary school quality) operate.

I do not attempt to measure institutional quality directly. Rather, I investigate whether CZs that have high $\theta_c$ – strong transmission of parental income to children's income – also

[8] This assumption would fail to hold if, for example, CZs differed in the relative quality of higher education options that were available to students with strong vs. weak high school records or in the labor market returns to ability. Indeed, I present evidence below that $\lambda_{3c}$ covaries across CZs with $\theta_c$.

tend to have high transmission into earlier outcomes, as measured by $\pi_{1c}$ and $\tilde{\pi}_{2c}$. As I discuss below, the available data do not permit me to measure $\pi_{1c}$ and $\tilde{\pi}_{2c}$ directly, but they do allow me to measure their associations with $\theta_c$. I report correlations of $\theta_c$ with $\pi_{1c}$ and $\tilde{\pi}_{2c}$, as well as with $\pi_{3c}$ and $\lambda_{3c}$, and coefficients from bivariate regressions of $\theta_c$ on each of these.

It is worth reiterating that these associations are only suggestive – if across CZs, institutions that promote high values of $\pi_{1c}$ are associated with institutions that promote high values of $\pi_{2c}\lambda_{3c} + \pi_{3c}$, $\pi_{1c}$ might appear to be strongly associated with $\theta_c$ even though the key channels for the transmission of inequality are via post-school-entry experiences. I can partially address this by decomposing $\theta_c$ into components reflecting end-of-school human capital ($s_{ic2}$), returns to human capital ($\lambda_{3c}$), and income residuals ($\epsilon_c$). This decomposition is presented in Section 7. In any event, the associations that I measure are more specific than the institution-mobility associations measured by CHKS, and will help to point to directions for further inquiry.

## 3 Data

The intergenerational income transmission coefficient at the CZ level, $\theta_c$, is a key element of my analysis. I draw this from CHKS, who refer to this as "relative mobility" and measure it as the coefficient of a regression, using data from CZ $c$, of the adult income of children born between 1980 and 1982 ($y_{ic}$) on the income of their parents ($p_{ic}$). Children's income is measured for their families (so includes spousal earnings as well as non-labor income) averaged over the years 2011 and 2012, when the children are between 29 and 32, and is scaled as percentile ranks within the national distribution for such children. Children are linked to parents who claimed them as dependents after 1996, and $p_{ic}$ is the average family Adjusted Gross Income (plus tax-exempt interest and non-taxable Social Security benefits) for those parents between 1996 and 2000, again converted to national percentile ranks within the population of parents. Column 1 of Table 1 presents unweighted summary statistics for the CZ-level regression coefficients. The average of 0.33 indicates that in the average CZ, each one percentile increase in parental income is associated with one-third of a percentile increase in children's income. Column 2 presents statistics for the transmission of parental income to children's college enrollment; as noted above, this is correlated 0.68 with the income transmission measure.

I also explore below two alternative measures of intergenerational income transmission. CHKS report estimates based on the birth cohorts of 1983-85. Children's incomes are still measured in 2011 and 2012, when these cohorts are 26-29 years old, so may not be reliable indicators of children's eventual labor market positions. Nevertheless, this measure (summarized in column 3 of Table 1) is correlated 0.84 across CZs with the measure for the earlier cohorts. Chetty and Hendren (2015) compute more plausibly causal estimates of CZ-level income transmission, based on children who move across CZs at different ages. These estimates, summarized in column 4, are measured relative to the average CZ, so have mean zero by construction. They are based on somewhat small samples and are noisy. Nevertheless, they are correlated 0.85 with CHKS's preferred estimates and 0.89 with the estimates from the later cohorts.

### 3.1 Samples

To measure the transmission of parental income to children's educational outcomes, I need data that contain each. For this, I rely on three nationally representative, longitudinal surveys conducted by the National Center for Education Statistics. Each covers a different birth cohort and age range.

My primary results are based on the Educational Longitudinal Study (ELS). This is a sample of just over 19,000 10th graders in 2002, corresponding roughly to the 1985-1986 birth cohorts. Respondents were surveyed in 2002 (10th grade), 2004 (12th grade), 2006 (two years after normal high school graduation), and 2012 (eight years after). Children are geocoded to commuting zones based on their residential zip codes in the base year, supplemented with later information if the base year zip code is missing. As child outcome measures, I use scores from math and reading assessments administered in the first two waves, college completion and educational attainment in years from the 2012 survey, and non-cognitive skill measures (discussed in Section 5.3) measured in the initial survey. For comparability with income measures, test scores are converted to percentiles. I also construct children's adult income, $y_{ic}$, as their self-reported 2011 family income (including spouses or partners), when children were 25 or 26 years old.

To examine earlier childhood outcomes, I use the Early Childhood Longitudinal Study, Kindergarten Cohort (ECLS-K). This survey sampled kindergarteners in 1998-9 and followed them through 8th grade in 2007. Child outcomes are math and reading scores, again converted to percentiles, and non-cognitive skill as assessed by teachers and by the students themselves in the 5th grade survey. Students are assigned to CZs based on their 8th grade residences.[9]

There are four limitations of the available samples for my purposes. First, none of the available surveys provides outcomes across the full range of ages, ranging from Kindergarten through labor market entry. Thus, mapping out the age profile of student outcomes requires comparing ECLS and ELS results for different students. It is not possible to measure directly the impact of parental income on later achievement, controlling for earlier achievement (i.e., $\pi_2$ in equation (4b)).

Second, the samples represent different birth cohorts. CHKS compute relative economic mobility measures for children born in 1980-1982 and 1983-85; as noted above, they are very highly correlated. The latter is nearly the same cohorts represented in the ELS, but the ECLS represents a later cohort, born around 1992-1993. To check whether differences between ELS-based results for older children and ECLS-based results for younger children are due to cohort rather than age differences,[10] I turn to a third survey, the High School Longitudinal Study (HSLS), which has a similar structure to the HSLS but represents children born in roughly 1994-1995, nearly the same cohort that is represented in the ECLS.[11]

Third, the parental income measures in the NCES surveys are limited.[12] In the ELS, parents report total family income in the base-year survey. This question is not asked

[9]Where 8th grade residences are unavailable, I use the location of the 8th grade school, then the 5th grade residence and school, then 3rd grade, and so on.

[10]Chetty et al. (2014) find that national aggregate relative mobility has been quite stable across a range of birth cohorts (born 1971-1993), but CZ-level measures might in principle vary across cohorts with little variation in the national aggregate.

[11]Post-secondary measures are not yet available for the HSLS sample, so I focus on test scores, measured in 9th and 11th grades.

[12]One possible solution, not yet pursued, would be to use broader measures of family resources, such as the SES indices in several surveys that combine parental education, occupation, and income, in place of income on its own.

in subsequent waves, so I cannot average over multiple years to better approximate the family's permanent income (Rothstein and Wozny, 2013; Mazumder, 2005) as in CHKS. Moreover, the parental income variable is binned into 13 categories. To construct a measure as comparable as possible to those used by CHKS, I assign each category to the midpoint of the national percentile range it represents. ECLS parents were asked their family incomes three separate times, in Kindergarten, 1st grade, and 3rd grade, and in all but the first the responses are again reported in 13 bins (the Kindergarten response is reported continuously). I convert the 1st and 3rd grade bins to dollars (assigning the midpoint), average across the three waves, and construct percentiles of the distribution of averages. In the HSLS, family income is reported in each of the first two survey rounds, without binning. I average these and construct percentiles.

Fourth, and most importantly, each of the surveys provides a national sample of only 10,000 - 20,000 observations. With 741 CZs in the country, this amounts to well under 100 observations per CZ in each survey. The surveys all use multi-stage sampling designs, typically using schools as one stage and then choosing relatively large samples of students within each school. This means that within-CZ heterogeneity is even more limited than even the small sample sizes imply. A consequence is that it is necessary to pool information across CZs in order to obtain any precision at all about the relationship between parental income and later outcomes (Gelman and Hill, 2006). This limits what I can measure: As I discuss below, I can estimate the distribution of $\pi_{1c}$ and $\tilde{\pi}_{2c}$ across CZs $c$, and their association with other CZ-level measures such as the CHKS relative mobility measure $\theta_c$, but I cannot estimate each city's $\pi_{1c}$ and $\tilde{\pi}_{2c}$ separately.

### 3.2 National estimates

Summary statistics for the three ECLS samples are reported in Table 2. Summary statistics are not reported for incomes or test scores – all analyses here convert each to a percentile within the relevant sample, with mean 50.0 and standard deviation 28.9.

Table 3 presents preliminary estimates of the national relationship between each of my primary outcome measures and parental income. The first panel presents results for math and reading scores in grades Kindergarten through 8 from the ECLS. Each percentile increase in parental income is associated with an increase in Kindergarten scores of 0.41 percentiles in math and 0.37 percentiles in reading. Each of these is essentially unchanged when CZ fixed effects are added, in columns 2 and 4. Coefficients rise very slightly as students progress through elementary and middle school; by 8th grade, the coefficients are 0.44 and 0.46.

Panel B presents results for grades 9 and 11 from the HSLS, which has only math scores. Coefficients are smaller here than in the ECLS.

Panel C presents results from the ELS, first for test scores in grades 10 and 12 and then for non-test outcomes. Test score coefficients are quite similar to those from the HSLS, indicating that each parental income percentile is associated with 0.35 - 0.38 test score percentiles, with a somewhat smaller within-CZ relationship. There is also an association between parental income and children's educational attainment: Each parental income percentile is associated with increases in college enrollment and completion of 0.26 and 0.49 percentage points, respectively, and with an additional 0.02 years of education on average. It is also associated with an additional 0.18 percentiles of children's income at age 25-26. This is somewhat lower than the average reported by CHKS, 0.33, a result that is plausibly attributable to a combination of the poor quality of the ELS income measures and, more importantly, the fact that the ELS measures children's income at a younger age than do CHKS.[13]

Table 4 presents additional analysis of the parental income - child income relationship in the ELS. Column 1 repeats the specification from the final row of Table 3. Column 2 divides parental income into the CZ-level sample mean and the deviation from that. The coefficient on the former is about double that of the latter.[14] Column 3 shows that each coefficient is robust to including CZ random effects. (In this column, and others labeled "RE," I do not

[13]CHKS report estimates of this regression across various ages at which children's income is measured in their Figure IIIA. They find that the slope coefficient (which corresponds to $\theta$ in the model above) rises sharply as children age through their 20s. When children are age 25, as in the final ELS wave, it is around 0.23.

[14]I have also estimated specifications that further decompose the deviation of parental income from the CZ mean into the deviation from the *school* mean and the difference between school and CZ means. The across-CZ and within-CZ, across-school coefficients are indistinguishable, and the within-school coefficient is much smaller. This is exactly what one would expect based on sorting into schools based on unobservables, school-based peer effects, and/or measurement error in individual family income. It indicates that there is nothing special about CZs in the production process, and in particular that the across-CZ association is not likely to be badly confounded by unobserved CZ factors correlated with both parental income and test scores.

use sampling weights.) Column 4 adds CZ fixed effects; I can no longer estimate the effect of CZ mean income, but the within-CZ parental income coefficient is the same as in earlier columns.

Columns 5-9 explore heterogeneity in the within-CZ parental income coefficient. In columns 5-6 I add an interaction with the CHKS income transmission measure, first with CZ random effects and then with CZ fixed effects. The interaction coefficient, 0.51, indicates that the ELS estimate of parental income - child income transmission is higher in CZs that CHKS estimate have higher parent-child income transmission, as expected. If measures were perfectly comparable, one would expect this coefficient to be 1; again, the deviation from this likely reflects the young age at which ELS children's income is measured and the poor quality of this measure.

Column 7 presents a random coefficients model, developed in more detail below, that allows the parental income coefficient to vary randomly across CZs. The standard deviation of the random component of this coefficient is 0.016, and not statistically distinguishable from zero. The point estimate indicates that three-quarters of the across-CZ variation in income transmission, as measured in the ELS, is statistically explained by the CHKS transmission measure, and I cannot rule out that all of the ELS variation is explained. The results thus validate the use of the ELS to measure CZ variation in intergenerational transmission, despite the small sample and imperfect measures. Columns 8 and 9 repeat the model using alternative income transmission measures – the Chetty-Hendren (2015) causal measure in column 8, and the CHKS measure for younger cohorts in column 9. Results are quite similar to those in column 7.

# 4 Empirical framework: A random coefficients (mixed effects) model

The quantities of interest in my investigation are the role of children's developmental outcomes, $s_{ic1}$ and $s_{ic2}$, in mediating the transmission of parental income $p_{ic}$ to children's income $y_{ic}$. This calls for a traditional mediation analysis, which would involve including $s_{ic1}$ and/or $s_{ic2}$ as controls in the basic intergenerational transmission regression (1). But these permit only a national-level mediation analysis[15]; the ELS sample is not large enough to permit estimation of these regressions at the CZ level, and none of the samples that are large enough contain all of $p_{ic}$, $y_{ic}$, and $s_{ic}$.

A fallback approach might be to estimate the decomposition, (6). This would require CZ-level measures of each of the components of $\Omega_c$, potentially from different samples. Even this is not possible, however, as there is no sample containing useful measures of child skills and parental income that is large enough to permit this.

Instead, I set my sights on a more achievable target, the variance-covariance matrix of $\Omega_c$. Feasible empirical models, estimable with the available samples, can be used to identify the "reverse" regressions of right-side elements of (6) on the left-side variable $\theta_c$. The coefficients and residual variances of these regressions, each of which is identified, can then be used to infer $V(\Omega_c)$ and, in turn, the correlations of $\theta_c$ with the other transmission coefficients.

Consider the transmission of parental income into some child developmental outcome $w_{ic}$:

$$w_{ic} = \kappa_c + p_{ic}\pi_c + u_{ic}. \tag{7}$$

For example, when the child outcome is the test score at school entry, this is equation (4a). Now consider the "reverse" projection of $\pi_c$, the transmission of parental income to the child's outcome, onto the intergenerational income transmission coefficient $\theta_c$:

$$\pi_1 = \gamma + \theta_c\beta + \eta_c, \tag{8}$$

where $\beta = {}^{cov(\theta_c,\pi_c)}/_{V(\theta_c)}$ is the across-CZ linear projection coefficient and $\eta_c$ is orthogonal to $\theta_c$. (I focus on identifying observational relationships; I do not give $\beta$ a causal interpretation.) Substituting (8) into (7), we obtain

$$w_{ic} = \kappa_c + p_{ic}\left(\gamma + \theta_c\beta + \eta_c\right) + u_{ic} \tag{9}$$

I estimate three types of regressions based on (9). First, Table 3, above, presented

[15] At the national level, when I add controls for educational attainment and 12th grade math scores to the specification from Table 4, column 4, the parental income coefficient falls from 0.17 to 0.08, indicating that a bit more than half of income transmission is mediated by human capital.

regressions like (9) for various child outcomes with $\beta$ and $V(\eta_c)$ constrained to zero. Second, I estimate simple regressions of $s_{ic1}$ on $p_{ic}$ and its interaction with $\theta_c$ (which, recall, is measured with high precision by CHKS):

$$w_{ic} = \kappa_c + p_{ic}\gamma + (p_{ic}\theta_c)\,\beta + e_{ic}, \tag{10}$$

where the error term is $e_{ic} \equiv p_{ic}\eta_c + u_{ic}$ and standard errors account for clustering at the CZ level. (I explore various specifications for the CZ-level effect $\kappa_c$, and find that OLS, random effects, and fixed effects specifications are all quite similar.) The interaction coefficient identifies the projection slope $\beta$. The models in Table 4, columns 5 and 6, are examples of this specification.

Specification (10) allows estimation of $\beta$, but does not provide an estimate of $V(\eta_c)$, which is needed to compute $V(\pi_c)$ and thus the correlation between $\pi_c$ and $\theta_c$. (Because $\theta_c$ is observed, it is straightforward to compute $V(\theta_c)$ and thus to recover from $\beta$ an estimate of the covariance between $\theta_c$ and $\pi_c$.) Thus, my third specification models the role of $\eta_c$ directly. Note that (9) is a random coefficients model, and that the quantity of interest is the across-CZ variance of the random component of the $p_{ic}$ coefficient:

$$w_{ic} = \kappa_c + p_{ic}\,(\gamma + \theta_c\beta + \eta_c) + u_{ic}. \tag{11}$$

If we assume that $(\kappa_c, \eta_c)$ and $u_{ic}$ are each normally distributed and i.i.d., this model (also known as a "mixed" model, with fixed parameters $\gamma$ and $\beta$ and random effects variance-covariance matrix $V(\kappa_c, \eta_c)$) can be estimated by maximum likelihood.[16] Common implementations of mixed models impose restrictions on the covariance between $\kappa_c$ and $\eta_c$, but this is not necessary for identification. Identification does require, however, that we assume that $\kappa_c$ and $\eta_c$ are orthogonal to both $\theta_c$ and the CZ-level average of $p_{ic}$. This assumption is the same as the caveat mentioned above: I can identify the observational regression of $\pi_c$

[16] Gelman and Hill (2006) discuss the estimation of models like this, which are referred to variously as mixed, hierarchical, random coefficient, or multi-level models. In economics, it is common to estimate models like (11) in two steps: First, $w_{ic}$ is regressed on $p_i$ separately for each CZ $c$, to estimate $\pi_c$, and the resulting coefficients are then regressed on $\theta_c$ in a second step. This approach is unsuitable when the samples in each CZ are so small; the parameters of (11) are much better estimated in the mixed model, which is able to obtain much better precision by pooling information from across CZs.

on $\theta_c$ and vice versa, but have no basis for the exclusion restriction that would be needed to interpret either as causal.

There is no fully satisfactory way to handle sampling weights in mixed models. Accordingly, I estimate these models without weights. Fortunately, when I estimate simpler models (e.g., fixed effects models without random coefficients), estimates are nearly identical with and without weights, so this limitation is not likely to dramatically affect my results.[17]

Recall from above that I found a stronger relationship between parental income and children's income across than within CZs. One expects (and indeed I find) similar results for other children's developmental outcomes. Of interest here is the within-CZ relationship with parental income and how this varies across CZs. It is not computationally feasible to absorb $\kappa_c$ via CZ fixed effects. As an alternative, to ensure that my estimates are identified from within-CZ relationship, I divide $p_{ic}$ into its CZ-level mean $\bar{p}_c$ and its deviation from that, $p_{ic} - \bar{p}_c$. It is the latter that is allowed to interact with $\theta_c$ and to have a random coefficient in (11); a main effect for $\bar{p}_c$ is included, but it is not interacted with $\theta_c$. Deviation of $p_{ic}$ from the CZ mean ensures that any correlation between $\kappa_c$ and $p_{ic}$ does not confound the relationships of interest. Similarly, I de-mean $\theta_c$ before interacting with $p_{ic} - \bar{p}_c$ to permit interpretation of the $p_{ic} - \bar{p}_c$ main effect coefficient as reflecting the relationship in the average CZ. The full mixed model is thus:

$$w_{ic} = \kappa_c + \bar{p}_c\phi + (p_{ic} - \bar{p}_c)\gamma + (p_{ic} - \bar{p}_c)\left(\theta_c - \bar{\theta}\right)\beta + (p_{ic} - \bar{p}_c)\eta_c + u_{ic} \tag{12}$$

With estimates of $\beta$ and $V(\eta_c)$ from the mixed model, the variance of the parental income - child outcome transmission coefficient $\pi_{1c}$ can be computed as $V(\pi_c) = V(\theta_c)\beta^2 + V(\eta_c)$. The correlation between the CHKS income-income transmission measure $\theta_c$ and income-child outcome transmission $\pi_c$ is then $\beta\sqrt{V(\theta_c)/V(\pi_c)}$. These statistics are reported for the ELS parent income-child income relationship in the bottom rows of Table 4, Column 7. This indicates that the standard deviation of $\pi_c$ is 0.033, with 77% of this variation coming from $\theta_c\beta$ rather than from $\eta_c$, which has a standard deviation of only 0.016. When I compute the

[17] In principle, it is possible to estimate mixed models while allowing for clustering of the error terms $u_{ic}$. In practice, however, I have encountered numerical convergence problems in this case. Thus, in some cases I report non-clustered versions of the mixed models. This is indicated in table notes.

correlation between the CHKS income transmission measure and the transmission implied by the ELS sample, it is $\rho = 0.875$. This high correlation is not surprising, of course, since in this case $\pi_c$ and $\theta_c$ differ only because the underlying measures differ slightly. Indeed, the p-value from of the hypothesis that $V(\eta_c) = 0$ is 0.57, indicating that I can't rule out the possibility that *all* of the across-CZ variation in $\pi_c$ is explained by $\theta_c$.[18] Of more interest is the correlation between income-income transmission $\theta_c$ and the transmission of parental income to children's developmental outcomes.

# 5 Results: The transmission of parental income to children's human capital outcomes across CZs

In this section, I present results for the relationship between CHKS's parent income-child income transmission measure and measures from the ELS, ECLS, and HSLS of the transmission from parental income to children's human capital outcomes. I begin by examining educational attainment, then end-of-year test scores, then consider variation across the life course in transmission to test scores, and finally examine non-cognitive scores.

## 5.1 Transmission to children's educational attainment

Table 5 presents results from specifications like those in Table 4, except this time using measures of children's eventual educational attainment – years of education by age 26, in Panel A, and an indicator for college graduation by that age, in Panel B. Not surprisingly, parental income is strongly related to both, about twice as much so across than within CZs. The within-CZ relationship is invariant to the inclusion of random or fixed CZ effects. In Column 4, when I interact parental income with CHKS's income transmission measure, the coefficient is positive – indicating that CZs with stronger transmission to child income also have stronger transmission to educational attainment – and significant.

Column 5 presents the full mixed model specification. The standard deviation of the unexplained component of the parental income coefficient is 0.002 in Panel A and 0.08 in

[18] I use a likelihood ratio test rather than a Wald test based on the point estimate and standard error of $\sigma_\eta$ because the likelihood is specified in terms of $\ln(\sigma_\eta)$ to ensure that any real value yields a valid likelihood. Confidence intervals are not symmetric around $\hat{\sigma}_\eta$, particularly near $\sigma_\eta = 0$.

Panel B, in each case somewhat larger than the explained component (which has standard deviation of 0.025*0.065=0.0016 in Panel A and 0.75*0.065=0.05 in Panel B). This may indicate that there is some independent variation in income-education transmission not explained by income-income transmission. This was clear already from CHKS's results, which found a correlation of 0.68 between income-income and income-college enrollment transmission across CZs. I estimate that the correlation between CHKS's income-income transmission and the ELS-based income-attainment transmission is about 0.5 for each of the two attainment measures.[19] This is not too far below the CHKS estimate, and the difference is likely explained by the divergence between the ELS-based and tax-based income transmission coefficients seen earlier. Results are similar for the income transmission measure computed from the 1983-5 cohorts (column 7), but somewhat weaker for the Chetty-Hendren mobility-based measure (column 6). Overall I interpret the educational attainment results as further validation of the strategy of using the ELS panel data to measure the mechanisms behind intergenerational income transmission.

## 5.2 Transmission to children's test scores

In Table 6, I shift focus from educational attainment to achievement, examining 12th grade math scores in the ELS. On average, each percentile of parental income is associated with about 0.35 percentiles of children's math scores within CZs and 0.69 percentiles across CZs. (As before, there is little distinction between between-CZ and within-CZ, across-school variation, but the association between income and achievement is only about half as strong within schools as between.) When I interact family income with CZ-level income transmission, in column 3, the coefficient is 0.36. This is comparable in magnitude to the parental income main effect, as in earlier analyses.

In the mixed model in Column 4, the variance of the random component of the income coefficient is quite large, accounting for nearly 90% of the total variance of $\pi_c$, and highly statistically significant. This stands in marked contrast to the earlier results, where $V(\eta_c)$ was trivial or zero and accounted for only a small share of the overall variance. Accordingly, when I compute the correlation between income-test score transmission $\pi_c$ and income-

[19] I cannot reject the null hypothesis that $\sigma_\eta = 0$ and thus that the correlation equals 1.

income transmission $\theta_c$, this is only 0.32. This is hard to reconcile with the hypothesis that test scores, or the knowledge and skill that they represent, are a key mechanism determining intergenerational income transmission, since there is evidently substantial variation in test score outcomes across CZs that does not translate into corresponding variation in income transmission. I explore this argument more formally below, in Section 7. Columns 5 and 6 of Table 6 show that the result is robust to using either of the alternative income transmission measures.

Table 7 presents mixed model estimates for each of the available test scores from the ECLS, ELS, and HSLS, using the CHKS preferred income transmission measure. The coefficient on the interaction between parental income and CZ-level income transmission ($\beta$ in equation 12) is significant for three of the four estimates for high school math scores, but not for the one available high school reading score or for any of the elementary scores in either subject. By contrast, the random component of the parental income coefficient is statistically significant and has substantial variance in each specification, accounting for 90% or more of the overall variance of $\pi_c$. (The one exception is the HSLS 11th grade scores, where it accounts for only 81%.) Thus, the estimated correlation between income-income transmission and income-test score transmission is quite weak across specifications, ranging from 0.08 to 0.44 and generally around 0.2-0.3.

The pattern of results in Table 7 has several implications. First, there is some indication from the $\hat{\beta}$ estimates (column 1) that the relative importance of parental income to student test scores in high-income-transmission CZs grows between elementary and high school, consistent with the hypothesis that differential access to school quality is a mechanism contributing to differential income transmission. This is based largely on the comparison between the ECLS survey for elementary and middle school and the HSLS and ELS for later grades, however, so is sensitive to measurement explanations for differences in results. (Moreover, it is driven more by differences in the standard errors across surveys than by differences in point estimates, which are quite similar.) In any event, the growth in these coefficients is quite small. Second, there is substantial heterogeneity across CZs in the transmission of parental income to children's test scores that is not associated with CZ-level income transmission (column 2), indicating that the institutions or other CZ characteristics that contribute to test score transmission differ from those determining income transmission. Put somewhat differently, there is only a weak correlation across CZs between income-income and income-test score transmission (column 5), so different influences must be at work. Finally, results are quite similar for the HSLS as for the ELS, suggesting that cohort differences are unable to explain the weak relationship of income-income and income-test score transmission in the HSLS and ECLS.

### 5.3 Transmission to non-cognitive skills

The results above suggest that CZs in which the transmission of parental income to children's income is stronger tend also to be CZs in which there is relatively strong transmission of parental income to children's college attainment, but not, by and large, CZs in which parental income matters more to children's test scores, either early or late in schooling careers. This is suggestive that learning in school is not a key channel determining the across-CZ variation in income transmission. One possibility not yet considered, however, is that schools do matter, but that math and reading test scores do not capture their impacts. A growing literature in recent years has documented the importance of non-cognitive skills as a component of the learning process. Both the ECLS and the ELS contain batteries of questions aimed at identifying children's non-cognitive skills, and I use these to assess whether high-income-transmission CZs tend to be CZs with large gaps in non-cognitive skills between children from high- and low-income families.

Each row of Table 8 presents results for one non-cognitive skill measure from the mixed model specification. I present results in three panels. Panel A presents results from the ELS 10th grade survey, in which students were asked batteries of questions aimed at identifying their instrumental motivation, their general effort and persistence, their sense of control, and their sense of self-efficacy in math and in reading. Across most of these, there is statistically significant variation across CZs in the parental income slope (column 2). But this variation is largely independent of the variation in the CHKS income transmission measure (column 1), so the two are only weakly and inconsistently correlated – the strongest correlation is for self efficacy in math, but the correlation for self efficacy in reading is nearly as large and of the opposite sign (column 5). When I average all five non-cognitive skill measures (after converting each to a z-score), I find essentially no relationship between income-non-cognitive skill transmission and income-income transmission. Thus, there is no indication here that non-cognitive skills are a key channel driving income transmission.

Panel B presents results from the student questionnaire administered as part of the ECLS 5th grade wave. The ECLS measures somewhat different non-cognitive skills – interest and competence in several academic and social domains, as well as internalizing and externalizing problem behaviors – but the results are similar: There is variation in the income slope across CZs, but it is not meaningfully related to income transmission (and in the one case in which it is statistically significant, it has the "wrong" sign).

Panel C presents results for non-cognitive skill measures constructed from the *teacher* survey component of the ECLS 5th grade wave. Results are quite different here: In four of the six cases, the parental income-CZ income transmission interaction is positive, large, and statistically significant, and correlations between income-income and income-non cognitive skill transmission are around 0.5. It is not clear how to account for the discrepancy between these results and those from the student self-reports in Panel B – even when the concepts overlap (e.g., for externalizing problem behaviors), results are quite different. This may indicate that high-transmission CZs tend to be CZs in which teachers are more biased in their assessments of low-income children, but this is quite speculative.

# 6 Results: Returns to education

Another channel for income transmission is via the *return* to skill. I have focused in the previous section on fairly small differences across CZs in the relative skill accumulation of children from low- and high-income families. While there is variation in this, it is small relative to the average: In all CZs, children from high-income families accumulate more skills than do children from low-income families. Thus, if some CZ labor markets offer higher returns to those skills, that would contribute to income gaps between children from high- and low-income families.

To investigate this, I turn to data from the American Community Survey (ACS) 2010, 2011, and 2012 one-year public use microdata samples. For maximum comparability to

CHKS's results, I focus on the 253,852 individuals in these samples born between 1980 and 1982. As before, I estimate mixed models, in this case allowing the return to education (specified as percentiles of income per year of completed education) to vary both with the CHKS income transmission measure and independently across CZs.

Results are presented in Table 9. Column 1 shows that each year of education is associated with 9.4 percentiles of income across CZs, and 3.9 percentiles within CZs. Column 3 presents a simple interacted model with CZ random effects but fixed coefficients. It indicates that the return to education is larger in high-income-transmission CZs – as before, the interaction coefficient is comparable in magnitude to the income main effects (though recall that the transmission measure has standard deviation 0.065, so most CZs have returns to education within about 10% of the average shown in columns 1-2). Column 4 presents the mixed model. There is also substantial, statistically significant variation across CZs conditional on the CHKS income transmission measure, which accounts for only one-quarter of the total across-CZ variation in the return to education. The basic pattern is, as before, robust to using either of the alternative transmission measures (columns 5-6).

# 7 Decomposing the variation in CZ-level income-income transmission

The results thus far indicate that CZs with relatively strong intergenerational income transmission tend to have stronger relationships between parental income and children's educational attainment, only slightly stronger associations between parental income and children's test scores, and higher returns to education. The correlations thus far do not resolve, however, whether these relationships are strong enough to account for the variation in income transmission. In this section, I explore decompositions of the across-CZ variation in the income-income relationship into components attributable to (a) children's skill accumulation by the end of school; (b) returns to skills; and (c) relationships between parents' incomes and the component of children's incomes that is not attributable to observed human capital.

Suppose that $z_{ic}$ is a scalar measure of the human capital of child $i$ from CZ $c$. We can project children's incomes onto this separately for each CZ:

$$y_{ic} = z_{ic}\psi_c + \nu_{ic}, \tag{13}$$

where $\psi_c$ is the return to human capital in city $c$ and $\nu_{ic}$ is the income residual. Thus, the relationship between children's incomes and parents' incomes in CZ $c$ is:

$$\frac{dy_{ic}}{dp_{ic}} = \frac{\partial z_{ic}}{\partial p_{ic}}\psi_c + \frac{\partial \nu_{ic}}{\partial p_{ic}}. \tag{14}$$

This is the definition of $\theta_c$, the income transmission measure. I label it $\theta_c^{ELS}$ to indicate that this relationship may differ somewhat based on the sample used to compute it. Now consider how this varies with the CHKS transmission measure, $\theta_c$:

$$\frac{d\theta_c^{ELS}}{d\theta_c} = \frac{d^2 y_{ic}}{dp_{ic}d\theta_c} = \frac{\partial^2 z_{ic}}{\partial p_{ic}\partial\theta_c}\psi_c + \frac{\partial z_{ic}}{\partial p_{ic}}\frac{\partial \psi_c}{\partial\theta_c} + \frac{\partial^2 \nu_{ic}}{\partial p_{ic}\partial\theta_c}. \tag{15}$$

The left side of this equation was estimated in columns 5-7 of Table 4 as 0.51. The right side has three components. The first term represents variation in human capital accumulation gaps between high- and low-income families. This might occur, for example, if high-$\theta_c$ CZs offer less equal school quality to children from different family backgrounds. The second term reflects covariance of the CZ-level return to skill with CZ-level income transmission, scaled by the overall average difference in skill accumulation between high- and low-income families. For example, high-$\theta_c$ CZs may have fewer unions or just a more unequal income distribution. The third term reflects differences in the transmission of parental income to children's incomes holding skills constant. This might be large if high-$\theta_c$ CZs have more stratified labor (or marriage) markets or employment networks that allow high-income parents to ensure good outcomes for their children regardless of the children's skills.

To implement this decomposition, I need a scalar index of human capital. I form this by estimating a regression of children's income percentiles on their educational attainment and their 12th grade math scores, with CZ fixed effects, in the ELS sample. My human capital index is the predicted value from this regression (neglecting the fixed effects). Note that this method of forming $z_{ic}$ normalizes $\psi = 1$ in the average city.

Column 1 of Table 9 repeats the basic random effects regression of children's income percentiles on parents' income and its interaction with CZ-level income transmission, from Table 4, Column 5. (Here, unlike there, I include a main effect for CZ income transmission, but this has little impact on the results.) The interaction coefficient, again, is 0.51.

In Column 2, I repeat the specification but use the child skill index as the dependent variable. The interaction coefficient here represents the first term of the decomposition (15). It is estimated at 0.09, implying that skill accumulation accounts for only 19% of the differences in ELS income transmission between cities with low and high values of the CHKS transmission measure.

Column 3 explores the role of returns to skill. Here, the dependent variable is the child's income percentile, but explanatory variables are the child's skill index and its interaction with the CZ-level income transmission. This interaction coefficient estimates $\frac{\partial \psi_c}{\partial \theta_c}$; the second term of the decomposition (15) can then be obtained by multiplying it by the average effect of parental income on children's skill, which by column 2 is 0.09. Thus, the second term is 0.16, indicating that differences in returns to skill account for 31% of the variation in income transmission.

Finally, column 4 returns to the specifications from columns 1 and 2, but here the dependent variable is the residual from the column 3 regression (representing $\nu_{ic}$ in (13)). The interaction coefficient here is 0.23, indicating that 46% of the variation in income transmission is attributable to differences in the transmission of parental income to child income controlling for the child's observable skills and for CZ-level differences in the returns to these skills.

# 8 Conclusion

Chetty et al.'s (2014) pathbreaking work showed that there is dramatic variation in intergenerational income mobility across geographic areas within the United States. This raises the intriguing possibility that we can identify policies that account for this variation and, by exporting these policies from high- to low-mobility areas, move closer to equality of opportunity.

CHKS presented suggestive correlations that indicated that school quality might be an important contributing factor. This paper has investigated this suggestion further, by asking whether high- and low-income children's academic outcomes are more equal in areas where their adult economic outcomes are more equal – that is, in areas with more intergenerational mobility. I find that there is statistically significant variation across commuting zones in the gradients of educational attainment, academic achievement, and non-cognitive skills with respect to parental income. Intergenerational income transmission is reasonably strongly correlated with the educational attainment gradient and with the labor market return to education, but does not covary strongly with either academic achievement or non-cognitive skill gradients (with the exception of gradients computed from teacher reports of children's non-cognitive skills).

I find that less than one-fifth of the across-CZ variation in intergenerational income mobility is attributable to differences in children's skill accumulation. Just under one-third is attributable to differences in the returns to skill. The remaining nearly half of the variation is attributable to differences in the return to parental income holding skills (and the returns to skills) constant.

Although this evidence is observational rather than causal, it strongly suggests that differences in elementary and secondary school quality are not an important determinant of variation in income mobility. (This is not to say that school quality is not important for other reasons, of course, or even that it does not contribute to overall mobility in a way that is roughly constant across CZs.) There appears to be more of a role for access to higher education in driving economic mobility, though even here the contribution is not large relative to the overall variation. Further investigation into the determinants of local intergenerational mobility should focus on differences in the returns to education and in the returns to family income conditional on children's human capital. Plausible factors driving the former might include institutions determining local income inequality, such as state income taxation and union density. The latter might reflect variation in the importance of labor market networks or in spatial or social stratification of the labor market. Further work will be needed to determine which of these, if any, play important roles.

## References

Bond, T. N. and K. Lang (2013). The evolution of the black-white test score gap in grades K–3: The fragility of results. *Review of Economics and Statistics 95*(5), 1468–1479.

Bradbury, B., M. Corak, J. Waldfogel, and E. Washbrook (2015). *Too many children left behind: The US achievement gap in comparative perspective.* Russell Sage Foundation.

Carneiro, P. and J. J. Heckman (2003). Human capital policy. In B. M. Friedman (Ed.), *Inequality in America: What Role for Human Capital Policies?* MIT Press.

Chetty, R. and N. Hendren (2015). The impacts of neighborhoods on intergenerational mobility: Childhood exposure effects and county-level estimates. Unpublished manuscript, Harvard University and NBER.

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

Chetty, R., N. Hendren, P. Kline, E. Saez, and N. Turner (2014). Is the United States still a land of opportunity? Recent trends in intergenerational mobility. *The American Economic Review 104*(5), 141–147.

Cunha, F. and J. Heckman (2007). The economics of human development: The technology of skill formation. *The American Economic Review 97*(2), 31–47.

Cunha, F. and J. J. Heckman (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources 43*(4), 738–782.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

Gelman, A. and J. Hill (2006). *Data analysis using regression and multilevel/hierarchical models.* Cambridge University Press.

Jacob, B. and J. Rothstein (2016). The measurement of student ability in modern assessment systems. *Journal of Economic Perspectives*. Forthcoming.

Mazumder, B. (2005). Fortunate sons: New estimates of intergenerational mobility in the united states using social security earnings data. *Review of Economics and Statistics 87*(2), 235–255.

Nielsen, E. R. (2015). Achievement gap estimates and deviations from cardinal comparability. Finance and Economics Discussion Series 2015-04, Board of Governors of the Federal Reserve System.

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan and R. J. Murnane (Eds.), *Whither Opportunity? Rising Inequality and the Uncertain Life Chances of Low-Income Children*, Chapter 5, pp. 91–116. New York: Russell Sage Foundation Press.

Rothstein, J. (2004). College performance predictions and the SAT. *Journal of Econometrics 121*(1), 297–317.

Rothstein, J. and N. Wozny (2013). Permanent income and the black-white test score gap. *Journal of Human Resources 48*(3), 510–544.

**Figure 1: Academic achievement as mediator of the effect of parental income on children's income**

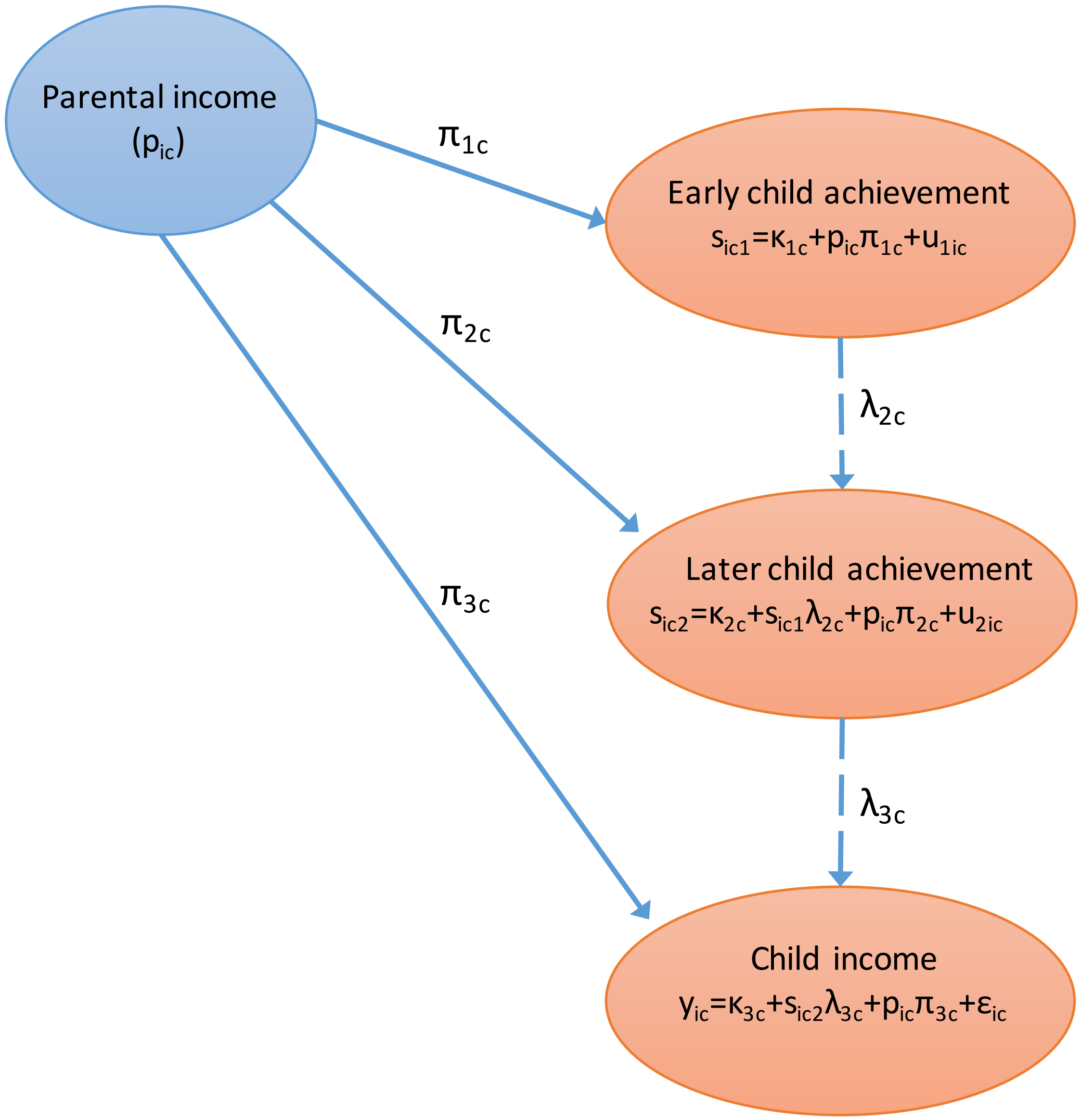

**Table 1. Measures of income transmission at the CZ level from tax data.**

| | Slope of: | | | |
|---|---|---|---|---|
| | child income percentile (0-100) | child college enrollment (0/1) | child income percentile (0-100) | child income percentile (0-100) |
| | with respect to parent income percentile (0-100) | | | |
| | (1) | (2) | (3) | (4) |
| Birth cohorts | 1980-1982 | 1980-1982 | 1983-1985 | 1980-1991 |
| Identification | observational | observational | observational | causal (movers) |
| | | | | |
| N | 718 | 709 | 708 | 718 |
| Mean | 0.33 | 0.0069 | 0.32 | 0.01 |
| Standard deviation | 0.06 | 0.0011 | 0.07 | 0.07 |
| 10th percentile | 0.24 | 0.0053 | 0.21 | -0.08 |
| 90th percentile | 0.40 | 0.0080 | 0.39 | 0.09 |
| Correlations | | | | |
| (1) | 1 | | | |
| (2) | 0.68 | 1 | | |
| (3) | 0.84 | 0.62 | 1 | |
| (4) | 0.85 | 0.61 | 0.89 | 1 |

Notes: Data from Chetty et al. (2014) (columns 1-3) and Chetty and Hendren (2015) (column 4). Summary statistics are computed across commuting zones, unweighted. These differ slightly from those reported by Chetty et al. (2014), which are weighted. Causal, movers-based measures in column 4 are computed relative to the average CZ, so have (weighted) mean zero.

**Table 2. Summary statistics for NCES samples**

| | Early Childhood Longitudinal Study (ECLS) | High School Longitudinal Study (HSLS) | Educational Longitudinal Study (ELS) |
|---|---|---|---|
| Birth year | 1992-1993 | 1994-1995 | 1985-1986 |
| N | 19,942 | 21,444 | 15,244 |
| # of CZs | 365 | 295 | 312 |
| Demographics | | | |
| Female | 0.48 | 0.50 | 0.50 |
| Black | 0.18 | 0.17 | 0.16 |
| Hispanic | 0.19 | 0.22 | 0.16 |
| Asian | 0.03 | 0.03 | 0.05 |
| Other non-white | 0.02 | 0.08 | 0.03 |
| Test scores available for grades | K,1,2,3,5,8 | 9,11 | 10,12 |
| Post-high school outcomes (from 2012 follow-up survey) | | | |
| College enrollment | | | 0.84 |
| College completion | | | 0.33 |
| Years of education | | | 14.0 |
| | | | (1.8) |
| Income at age 26 | | | 42,198 |
| | | | (39,979) |

Note: Sample sizes and demographics are computed for the base-year sample for each survey, and use sampling weights. Standard deviations reported in parentheses.

**Table 3. Cross-sectional regressions of child outcomes on parental income**

| | Math | | | Reading | | |
|---|---|---|---|---|---|---|
| | | | N | | | N |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| CZ FEs | N | Y | | N | Y | |
| ECLS-K | | | | | | |
| K (spring) | 0.41 | 0.40 | 19,186 | 0.37 | 0.36 | 18,498 |
| | (0.01) | (0.01) | | (0.01) | (0.01) | |
| G1 (spring) | 0.40 | 0.41 | 16,368 | 0.38 | 0.37 | 16,077 |
| | (0.01) | (0.01) | | (0.01) | (0.01) | |
| G3 | 0.44 | 0.42 | 14,180 | 0.45 | 0.43 | 14,086 |
| | (0.01) | (0.01) | | (0.01) | (0.01) | |
| G5 | 0.45 | 0.43 | 11,141 | 0.45 | 0.43 | 11,133 |
| | (0.02) | (0.02) | | (0.02) | (0.02) | |
| G8 | 0.44 | 0.42 | 9,213 | 0.46 | 0.44 | 9,154 |
| | (0.02) | (0.02) | | (0.02) | (0.02) | |
| HSLS | | | | | | |
| G9 | 0.36 | 0.32 | 20,164 | | | |
| | (0.01) | (0.01) | | | | |
| G11 | 0.34 | 0.31 | 20,462 | | | |
| | (0.01) | (0.01) | | | | |
| ELS | | | | | | |
| G10 | 0.37 | 0.34 | 15,244 | 0.35 | 0.32 | 15,244 |
| | (0.01) | (0.01) | | (0.01) | (0.01) | |
| G12 | 0.38 | 0.35 | 13,648 | | | |
| | (0.01) | (0.01) | | | | |
| College enrollment (*100) | 0.26 | 0.24 | 13,250 | | | |
| | (0.01) | (0.01) | | | | |
| College completion (*100) | 0.49 | 0.45 | 13,250 | | | |
| | (0.02) | (0.02) | | | | |
| Years of education | 0.020 | 0.019 | 13,250 | | | |
| | (0.001) | (0.001) | | | | |
| Income at 26 | 0.18 | 0.17 | 11,510 | | | |
| | (0.01) | (0.01) | | | | |

Notes: Each entry represents the coefficient from a separate regression of the child's test score (or other outcome) on family income, in columns 2 and 5 with CZ fixed effects. Test scores, family incomes, and child incomes are measured in percentile units, scaled 0-100. College enrollment and completion are binary, with coefficients multiplied by 100; years of education is measured in natural units. Regressions are weighted using inverse sampling probability weights. Standard errors are clustered on the CZ.

**Table 4. Parent income - child income relationships in the ELS**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| Parental income | 0.18 | | | | | | | | |
| | (0.01) | | | | | | | | |
| Parental income - CZ mean | | 0.16 | 0.17 | 0.17 | 0.16 | 0.16 | 0.16 | 0.16 | 0.16 |
| | | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| CZ mean parental income | | 0.34 | 0.35 | | 0.35 | | 0.34 | 0.34 | 0.34 |
| | | (0.04) | (0.04) | | (0.04) | | (0.04) | (0.04) | (0.04) |
| (Parental income - CZ mean) * | | | | | 0.51 | 0.51 | 0.52 | 0.42 | 0.49 |
| CZ income transmission | | | | | (0.16) | (0.21) | (0.17) | (0.18) | (0.16) |
| Income transmission measure | | | | | CHKS 80-82 | | | CH | CHKS 83-5 |
| CZ controls | None | None | RE | FE | RE | FE | RE | RE | RE |
| SD of parental income random coefficient (η) | | | | | | | 0.016 | 0.010 | 0.020 |
| | | | | | | | (0.015) | (0.016) | (0.016) |
| SD of total parental income coefficient (π) | | | | | | | 0.033 | 0.024 | 0.035 |
| Share of variance attributable to CZ income transmission | | | | | | | 77% | 83% | 68% |
| p-value, share of variance = 100% (LR test) | | | | | | | 0.57 | 0.80 | 0.44 |
| Corr(tax data income transmission, ELS π) | | | | | | | 0.875 | 0.910 | 0.828 |

Notes: Dependent variable in each column is the child's family income at age 25, in percentile units (0-100). Parental income is also measured in percentiles. Standard errors in columns 1-6 are clustered at the CZ level; column 7 SEs assume i.i.d. errors. CZ income transmission in columns 5-7 is the observational measure for the 1980-82 birth cohorts from Chetty et al. (2014); in column 8, it is the causal estimate from Chetty and Hendren (2015); in column 9 it is the Chetty et al. (2014) observational measure for the 1983-85 birth cohorts. Each is demeaned before interacting with parental income. Sampling weights are used in columns 1, 2, 4, and 6 but not in the random effects specifications in columns 3, 5, and 7-9. Standard errors are clustered on the CZ in columns 1-6, but not in the mixed model specifications in 7-9. p-values in columns 7-9 are for likelihood ratio tests of the mixed models against random effects models with fixed coefficients (as in column 6), in which all variation in parental income coefficients is attributable to income transmission and the two are therefore perfectly correlated.

**Table 5. Parental income - children's educational attainment relationships in the ELS**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| *Panel A: Dependent variable is years of education at age 26* | | | | | | | |
| Parental income | 0.021 | | | | | | |
| | (0.001) | | | | | | |
| Parental income - CZ mean | | 0.019 | 0.019 | 0.019 | 0.019 | 0.019 | 0.019 |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| CZ mean parental income | | 0.038 | | 0.041 | 0.042 | 0.042 | 0.042 |
| | | (0.003) | | (0.003) | (0.002) | (0.003) | (0.003) |
| (Parental income - CZ mean) * | | | | 0.023 | 0.025 | 0.016 | 0.025 |
| CZ income transmission | | | | (0.011) | (0.011) | (0.011) | (0.011) |
| Income transmission measure | | | | CHKS 80-82 | | CH | CHKS 83-5 |
| CZ controls | | None | FE | RE | RE | RE | RE |
| SD of parental income random coefficient (η) | | | | | 0.002 | 0.003 | 0.002 |
| | | | | | (0.001) | (0.001) | (0.001) |
| SD of total parental income coefficient (π) | | | | | 0.003 | 0.003 | 0.003 |
| Share of variance attributable to CZ income transmission | | | | | 28% | 10% | 29% |
| p-value, share of variance = 100% (LR test) | | | | | 0.36 | 0.27 | 0.37 |
| Corr(tax data income transmission, ELS π) | | | | | 0.53 | 0.32 | 0.54 |
| *Panel B: Dependent variable is college graduation by age 26* | | | | | | | |
| Parental income (/100) | 0.51 | | | | | | |
| | (0.01) | | | | | | |
| Parental income - CZ mean (/100) | | 0.44 | 0.45 | 0.45 | 0.45 | 0.45 | 0.45 |
| | | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) |
| CZ mean parental income/100 | | 0.92 | | 1.02 | 1.01 | 1.01 | 1.01 |
| | | (0.07) | | (0.06) | (0.06) | (0.06) | (0.06) |
| (Parental income - CZ mean) * | | | | 0.65 | 0.75 | 0.47 | 0.68 |
| CZ income transmission | | | | (0.30) | (0.30) | (0.30) | (0.28) |
| Income transmission measure | | | | CHKS 80-82 | | CH | CHKS 83-5 |
| CZ controls | | None | FE | RE | RE | RE | RE |
| SD of parental income (/100) random coefficient (η) | | | | | 0.08 | 0.08 | 0.08 |
| | | | | | (0.03) | (0.03) | (0.03) |
| SD of total parental income (/100) coefficient (π) | | | | | 0.09 | 0.09 | 0.09 |
| Share of variance attributable to CZ income transmission | | | | | 24% | 8% | 21% |
| p-value, share of variance = 100% (LR test) | | | | | 0.15 | 0.10 | 0.15 |
| Corr(tax data income transmission, ELS π) | | | | | 0.49 | 0.29 | 0.46 |

Notes: Dependent variable in panel A is years of education as of the 2012 (age 26) follow-up survey; in Panel B, it is an indicator for a four-year college degree. Parental income is measured in percentile units (0-100), but this is scaled to 0-1 in Panel B. Standard errors are clustered at the CZ level; estimates in columns 1-3 are weighted but those in columns 4-7 are unweighted. See notes to Table 4 for description of income transmission measures. p-values in columns 5-7 are for likelihood ratio tests of the mixed models against random effects models with fixed coefficients (as in column 4).

**Table 6. Parental income and children's 12th grade math achievement in the ELS**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Parental income - CZ mean | 0.35 | 0.35 | 0.34 | 0.33 | 0.33 | 0.33 |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| CZ mean parental income | 0.69 | | 0.71 | 0.71 | 0.71 | 0.71 |
| | (0.04) | | (0.04) | (0.04) | (0.04) | (0.04) |
| (Parental income - CZ mean) * | | | 0.36 | 0.41 | 0.18 | 0.26 |
| CZ income transmission | | | (0.15) | (0.19) | (0.18) | (0.16) |
| Income transmission measure | | | CHKS 80-82 | | CH | CHKS 83-5 |
| CZ controls | None | FE | RE | RE | RE | RE |
| SD of parental income random coefficient ($\eta$) | | | | 0.07 | 0.07 | 0.07 |
| | | | | (0.01) | (0.02) | (0.02) |
| SD of total parental income coefficient ($\pi$) | | | | 0.07 | 0.07 | 0.07 |
| Share of variance attributable to CZ income transmission | | | | 10% | 2% | 5% |
| p-value, share of variance = 100% | | | | <0.01 | <0.01 | <0.01 |
| Corr(tax data income transmission, ELS $\pi$) | | | | 0.32 | 0.13 | 0.22 |

Notes: Dependent variable is the 12th grade math score, converted to percentile units (0-100). Parental income is also measured in percentile units. CZ income transmission in columns 4-5 is the observational measure for the 1980-82 birth cohorts from Chetty et al. (2014); in column 6, it is the causal estimate from Chetty and Hendren (2015); in column 7 it is the Chetty et al. (2014) observational measure for the 1983-85 birth cohorts. Each is demeaned before interacting with parental income. Columns 1-2 are weighted, while 3-6 are unweighted; in each column, standard errors are clustered at the CZ level. p-values in columns 4-6 are for likelihood ratio tests of the mixed models against random effects models with fixed coefficients (as in column 3), in which all variation in parental income coefficients is attributable to income transmission and the two are therefore perfectly correlated.

**Table 7. Variation in parental income - child achievement relationships across grades, cohorts, and subjects**

| | | Parental income * CZ income transmission | SD of parental income random coefficient ($\eta$) | SD of total parental income coefficient ($\pi$) | Share of variance explained by parental income transmission | Corr(CHKS income transmission, $\pi$) |
|---|---|---|---|---|---|---|
| | | (1) | (2) | (3) | (4) | (5) |
| *Panel A: Math scores* | | | | | | |
| ECLS-K | K (spring) | 0.14 | 0.08 | 0.08 | 0.01 | 0.11 |
| | | (0.23) | (0.01) | | | |
| ECLS-K | G1 (spring) | 0.21 | 0.06 | 0.06 | 0.04 | 0.20 |
| | | (0.22) | (0.01) | | | |
| ECLS-K | G3 | 0.35 | 0.06 | 0.07 | 0.10 | 0.32 |
| | | (0.21) | (0.02) | | | |
| ECLS-K | G5 | 0.35 | 0.06 | 0.07 | 0.10 | 0.32 |
| | | (0.20) | (0.01) | | | |
| ECLS-K | G8 | 0.20 | 0.05 | 0.05 | 0.06 | 0.25 |
| | | (0.20) | (0.02) | | | |
| HSLS | G9 | 0.30 | 0.05 | 0.05 | 0.10 | 0.31 |
| | | (0.17) | (0.01) | | | |
| HSLS | G11 | 0.62 | 0.07 | 0.07 | 0.19 | 0.44 |
| | | (0.17) | (0.01) | | | |
| ELS | G10 | 0.34 | 0.06 | 0.07 | 0.08 | 0.29 |
| | | (0.17) | (0.01) | | | |
| ELS | G12 | 0.41 | 0.07 | 0.07 | 0.10 | 0.32 |
| | | (0.17) | (0.02) | | | |
| *Panel B: Reading scores* | | | | | | |
| ECLS-K | K (spring) | 0.29 | 0.08 | 0.08 | 0.05 | 0.22 |
| | | (0.26) | (0.01) | | | |
| ECLS-K | G1 (spring) | 0.07 | 0.06 | 0.06 | 0.01 | 0.08 |
| | | (0.24) | (0.01) | | | |
| ECLS-K | G3 | 0.13 | 0.08 | 0.08 | 0.01 | 0.10 |
| | | (0.22) | (0.01) | | | |
| ECLS-K | G5 | 0.36 | 0.09 | 0.09 | 0.06 | 0.24 |
| | | (0.24) | (0.01) | | | |
| ECLS-K | G8 | 0.16 | 0.07 | 0.07 | 0.02 | 0.14 |
| | | (0.21) | (0.02) | | | |
| ELS | G10 | 0.24 | 0.08 | 0.08 | 0.03 | 0.17 |
| | | (0.18) | (0.01) | | | |

Notes: Each row presents a single mixed model regression. Parental income is de-meaned by CZ; income transmission measure is the Chetty et al. (2014) measure for the 1980-82 cohorts. Estimates are unweighted; standard errors are clustered at the CZ level.

**Table 8. Parental income and children's non-cognitive skills in the ELS**

| | Parental income * CZ income transmission | SD of parental income random coefficient (η) | SD of total parental income coefficient (π) | Share of variance explained by parental income transmission | Corr(CHKS income transmission, π) |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| *Panel A: ELS (10th grade)* | | | | | |
| Instrumental motivation | 0.14 | 0.02 | 0.02 | 0.11 | 0.33 |
| | (0.15) | (0.01) | | | |
| General effort and persistence | 0.07 | 0.06 | 0.06 | 0.01 | 0.07 |
| | (0.20) | (0.02) | | | |
| General control beliefs | -0.14 | 0.05 | 0.05 | 0.02 | -0.16 |
| | (0.18) | (0.02) | | | |
| Self-efficacy - Math | 0.21 | 0.03 | 0.03 | 0.18 | 0.43 |
| | (0.14) | (0.01) | | | |
| Self-efficacy - Reading | -0.42 | 0.07 | 0.07 | 0.12 | -0.35 |
| | (0.22) | (0.02) | | | |
| Index of five measures | -0.04 | 0.05 | 0.05 | 0.00 | -0.05 |
| | (0.16) | (0.02) | | | |
| *Panel B: ECLS-K 5th grade student survey* | | | | | |
| Perceived interest / competence in reading | -0.19 | 0.05 | 0.05 | 0.05 | -0.22 |
| | (0.20) | (0.01) | | | |
| Perceived interest / competence in math | 0.08 | 0.04 | 0.04 | 0.02 | 0.13 |
| | (0.18) | (0.02) | | | |
| Perceived interest / competence in all school subjects | 0.11 | 0.05 | 0.06 | 0.01 | 0.12 |
| | (0.20) | (0.01) | | | |
| Perceived interest / competence in peer relations | -0.09 | 0.04 | 0.04 | 0.02 | -0.13 |
| | (0.17) | (0.02) | | | |
| Externalizing problem behaviors | 0.04 | 0.01 | 0.01 | 0.06 | 0.25 |
| | (0.12) | (0.01) | | | |
| Internalizing problem behaviors | -0.36 | 0.05 | 0.06 | 0.14 | -0.38 |
| | (0.18) | (0.01) | | | |
| Index of six measures | -0.31 | 0.03 | 0.04 | 0.27 | -0.52 |
| | (0.18) | (0.02) | | | |

*Table continued on next page*

**Table 8 (continued)**

| | Parental income * CZ income transmission | SD of parental income random coefficient (η) | SD of total parental income coefficient (π) | Share of variance explained by parental income transmission | Corr(CHKS income transmission, π) |
|---|---|---|---|---|---|
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** |
| *Panel C: ECLS-K 5th grade teacher survey* | | | | | |
| Approaches to learning | 0.56 | 0.06 | 0.07 | 0.24 | 0.49 |
| | (0.21) | (0.02) | | | |
| Self-control | 0.71 | 0.06 | 0.08 | 0.32 | 0.56 |
| | (0.21) | (0.02) | | | |
| Interpersonal skills | 0.21 | 0.05 | 0.05 | 0.05 | 0.23 |
| | (0.20) | (0.02) | | | |
| Peer relations | 0.57 | 0.06 | 0.07 | 0.28 | 0.53 |
| | (0.20) | (0.02) | | | |
| Externalizing problem behaviors | 0.33 | 0.03 | 0.04 | 0.31 | 0.56 |
| | (0.14) | (0.01) | | | |
| Internalizing problem behaviors | -0.04 | 0.07 | 0.07 | 0.00 | -0.04 |
| | (0.19) | (0.01) | | | |
| Index of six measures | 0.60 | 0.07 | 0.08 | 0.22 | 0.47 |
| | (0.20) | (0.02) | | | |

Notes: Each row presents a single mixed model regression, estimated without sampling weights. Parental income is de-meaned by CZ; income transmission measure is the Chetty et al. (2014) measure for the 1980-82 cohorts. Standard errors are clustered at the CZ level except in the models for peer relations (both ECLS student and teacher surveys) and externalizing (ECLS teacher survey only).

**Table 9. Returns to education in American Community Survey (ACS) data**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Years of education - CZ mean | 3.94 | 3.98 | 4.02 | 3.84 | 3.85 | 3.81 |
| | (0.06) | (0.05) | (0.04) | (0.04) | (0.04) | (0.04) |
| CZ mean years of education | 9.35 | | 5.59 | 5.28 | 5.29 | 5.27 |
| | (0.87) | | (0.27) | (0.28) | (0.28) | (0.28) |
| (Years of education - CZ mean) * | | | 4.91 | 5.66 | 4.42 | 5.25 |
| CZ income transmission | | | (0.66) | (0.77) | (0.78) | (0.73) |
| Income transmission measure | | | CHKS 80-82 | | CH | CHKS 83-85 |
| CZ effects | None | FE | RE | RE | RE | RE |
| SD of education random coefficient ($\eta$) | | | | 0.55 | 0.59 | 0.57 |
| | | | | (0.04) | (0.04) | (0.04) |
| SD of total education coefficient ($\pi$) | | | | 0.64 | 0.63 | 0.65 |
| Share of variance attributable to CZ income transmission | | | | 25% | 13% | 23% |
| p-value, share of variance = 100% (LR test) | | | | <0.01 | <0.01 | <0.01 |
| Corr(tax data income transmission, ACS $\pi$) | | | | 0.50 | 0.36 | 0.48 |

Notes: Sample consists of individuals in the ACS 2010-2012 one-year public use microdata samples who were born between 1980 and 1982 (N=253,852). Dependent variable in each specification is the child's family income in percentile units (0-100). p-values in columns 4-6 are for likelihood ratio tests of the mixed models against random effects models with fixed coefficients (as in column 3), in which all variation in parental income coefficients is attributable to income transmission and the two are therefore perfectly correlated.

**Table 10. Decomposition of the variation in intergenerational income transmission**

| | Actual transmission | Skills | Return to skills | Residual |
|---|---|---|---|---|
| | | *Dependent variable* | | |
| | Child income | Child skill index | Child income | Child income residual |
| | (1) | (2) | (3) | (4) |
| Parental income - CZ mean | 0.16 | 0.09 | | 0.07 |
| | (0.01) | (0.00) | | (0.01) |
| CZ mean parental income | 0.35 | 0.19 | | 0.04 |
| | (0.04) | (0.01) | | (0.02) |
| CZ income transmission | 1.97 | -1.61 | 0.13 | -2.47 |
| | (7.64) | (1.69) | (7.59) | (3.73) |
| (Parental income - CZ mean) * CZ income transmission | 0.51 | 0.09 | | 0.23 |
| | (0.17) | (0.04) | | (0.17) |
| Skill index - CZ mean | | | 0.96 | |
| | | | (0.04) | |
| CZ mean skill index | | | 1.47 | |
| | | | (0.15) | |
| (Skill index - CZ mean) * CZ income transmission | | | 1.75 | |
| | | | (0.68) | |
| Coefficient of skills index on parent income (scale factor) | | | 0.09 | |
| Scaled component | | 0.09 | 0.16 | 0.23 |
| Share of total parental income * CZ income transmission | 100% | 19% | 31% | 46% |

Notes: Skill index is the predicted value from a regression of children's income (in percentiles) on their educational attainment and 12th grade math scores. Each specification includes CZ random effects, is unweighted, and clustered on the CZ.