# When Does Early Disadvantage Stop Predicting Adult Outcomes? Evidence from a Randomized Intervention

Yann Algan[1], Abdel-Hamid Bello*[2], Sylvana Côté[3], and Maripier Isabelle[4]

[1]HEC Paris

[2]Université du Québec à Montréal

[3]Université de Montréal

[4]Université Laval

**Abstract**

We study who gains most from early childhood interventions and whether these programs can weaken the transmission of disadvantage across generations. Using a population-representative Quebec cohort linked to tax records, we show that inattention, disruptiveness, and prosociality at age six account for 13 percent of the intergenerational correlation in income. We then study a randomized program that taught social skills and self-control to disruptive kindergarten boys from low-income Montreal neighborhoods. Its long-run effect on earnings varies most with inattention at school entry, which strongly predicts lower adult income among control boys but not among treated boys, so that the most inattentive boys gain the most, mainly through fewer years without earnings. The program thereby weakens the link between parental income and the adult income predicted by early skills, contrary to a common reading of dynamic complementarities.

**JEL Classification:** J62, I24, I31, C93

*Corresponding author: bello.abdel-hamid@uqam.ca

# 1 Introduction

Who benefits most from an early-life intervention remains an open question among economists. Because skills formed at one stage make later investments more productive (Cunha and Heckman, 2007, 2008), children who start school with fewer behavioral difficulties may gain more from a given program. Since early skills are themselves shaped by family background (Carneiro and Heckman, 2003; Cunha and Heckman, 2007), such programs could end up favoring children who are already advantaged. Yet the long-run evidence from the best-known early interventions shows large gains for highly disadvantaged children (Heckman et al., 2013; Campbell et al., 2014; García et al., 2021). This evidence is hard to interpret because these studies rarely examine how effects vary with the specific skills the programs target, so we cannot tell whether gains go to the children who start furthest behind, or whether some early null results (Garces et al., 2002; Deming, 2009) hide this kind of heterogeneity in small samples.

The answer matters for policy, since a central debate in early childhood policy is whether programs should be universal or targeted at the most disadvantaged children (Heckman, 2006; Almond and Currie, 2011). If children with stronger initial skills gain more, as dynamic complementarities suggest (Cunha and Heckman, 2007), then targeting the most disadvantaged children involves a trade-off between equity and returns (Carneiro and Heckman, 2003). If instead the most disadvantaged children gain the most, as the long-run evidence from Perry and Abecedarian tentatively suggests (Heckman et al., 2013; García et al., 2021), targeting them serves both goals and supports proportional or even progressive allocation (Currie, 2011). Answering this question requires knowing how the effect of a program varies with children's behavior before it starts, and how that behavior relates to family background.

This paper addresses the question in two steps, starting with a population cohort and turning to a randomized experiment. First, we measure how much the socioemotional skills that early childhood programs target, namely inattention, disruptiveness, and prosociality, contribute to the intergenerational persistence of income. We use a prospective, population-representative cohort of children born in Quebec, whose teachers rated their behavior every year from kindergarten to age 13, and who are linked in adulthood, along with their parents, to administrative tax records (Rouquette et al., 2014). Following Blanden et al. (2007), we decompose the intergenerational income correlation into the part accounted for by each skill

and a residual direct channel.

Second, we ask whether an early intervention can weaken the link between behavioral disadvantage and adult income, and for whom, using the Montreal Longitudinal and Experimental Study (Tremblay et al., 1992). In 1984, teachers rated the disruptive behavior of 1,161 kindergarten boys in 53 schools in low-income neighborhoods of Montreal. Boys above the 70th percentile were randomly assigned either to a two-year program of social skills and self-control training, in which they learned, for example, how to invite another child to play, how to ask questions, and what to do when angry, or to a control group with access to standard school resources only. Less disruptive boys from the same schools, who were not part of the randomization, form a reference group. We link all three groups to tax records, which give us adult earnings and years with zero earnings in their thirties, as well as parental income during adolescence.

In the population cohort, inattention, disruptiveness, and prosociality at age six jointly account for 13 percent of the intergenerational income correlation, without controlling for later pathways such as education, cognitive ability, or criminal involvement. Children from higher-income families start school with less inattention and disruptiveness and more prosociality, and these skills in turn predict adult income. Early behavior is therefore one of the channels through which economic position is passed from parents to children, which raises the question of whether a program targeting these skills can change adult outcomes, and whether it does so equally for all children.

The average long-run effect of the intervention has already been documented by Algan et al. (2022), and we take it as our starting point. Relative to what their family background and behavior at age six would predict, treated boys earn 1.46 log points more than control boys between ages 30 and 36, a result that holds across small-sample inference methods (e.g., Young, 2019; Romano and Wolf, 2005). This large gain comes entirely from the extensive margin: treated boys are 9 percentage points less likely to have years with zero earnings in their thirties, and the income gap disappears among boys who always had positive earnings.

Our main result concerns who gains from the program: although boys were selected for disruptive behavior, the long-run effect of the intervention varies most with a different dimension, their inattention at age six. Among control boys, a one-standard-deviation increase in baseline inattention is associated with 0.71 log points lower adult income. The interaction between treatment and baseline inattention is 1.12 log points, large enough to offset this gra-

dient: among treated boys, the relationship between baseline inattention and adult income is small and imprecisely estimated. The boys who started with the most inattention therefore gained the most, the opposite of what a common reading of dynamic complementarities would predict. This interaction remains significant after correcting for multiple testing across the three skills, although the joint test of all three interactions is significant only at the 10 percent level. Like the average effect, this heterogeneity operates on the extensive margin, as the reduction in years with zero earnings is largest for the most inattentive boys.

Because teachers kept rating the boys during the program (ages 7 to 9) and after it (ages 10 to 15), we can also ask whether the intervention changed their behavioral trajectories, and we find little evidence that it did. The only exception is the persistence of inattention itself: after the program ends, baseline inattention predicts later inattention less strongly among treated boys than among control boys ($p = 0.05$). For the other skills, and for the links between skills, the differences between treated and control boys are small and imprecise. The program thus seems to have changed the economic consequences of early inattention more than the behavior itself.

Finally, we ask whether the intervention reduces the part of income persistence that runs through skills formed early in life. Because behavioral difficulties at school entry predict lower adult income and are more common in poorer families, the adult income predicted by a child's early skills should rise with parental income. We predict adult income from baseline skills separately for treated and control boys, and regress these predictions on parental income. Among control boys, the gradient is 0.50; among treated boys, it is close to zero ($-0.03$), and the difference is significant at the 10 percent level across all inference methods (worst-case $p = 0.07$), with the same pattern for years with zero earnings.

These findings contribute to three literatures, starting with the literature on early childhood interventions (Heckman et al., 2013; Campbell et al., 2014; García et al., 2021; Algan et al., 2022). Prior work documents average long-run effects of such programs; we show that the long-run effect on adult income varies with baseline inattention, the behavioral dimension most strongly associated with family disadvantage, and that the boys with the most inattention gain the most. Because treated and control boys grew up in the same neighborhoods and attended the same schools, these gains arise without any change in the disadvantaged environment in which the boys lived, which complements evidence on neighborhood effects (Chetty et al., 2016; Deutscher, 2020; Chyn and Katz, 2021).

Second, we contribute to work on the role of early-life traits in the intergenerational persistence of income (Björklund et al., 2017; Blanden et al., 2007, 2014). This literature shows that childhood characteristics account for a nontrivial share of the transmission of economic status, but it cannot say whether policy can weaken this channel, and our experimental evidence suggests that it can, since the intervention substantially reduces the association between parental income and the adult income predicted by early skills.

Third, we contribute to work on the technology of skill formation (Cunha and Heckman, 2007, 2008; Attanasio et al., 2020). We show that the long-run effects of an intervention can vary with early behavioral characteristics even when the measured trajectories of those characteristics do not diverge durably between treated and control children, a pattern that differs from the one usually emphasized in this literature.

In the rest of the paper, Section 2 presents the conceptual framework and empirical approach, and Section 3 describes the data. Section 4 presents the population evidence and Section 5 the experimental results. Section 6 examines other adolescent and adult outcomes, Section 7 discusses the implications for intergenerational income persistence, and Section 8 concludes.

# 2 Conceptual Framework and Methodology

## 2.1 Conceptual Framework

We use a simple framework to organize the relationship between family economic circumstances, early behavioral characteristics, and adult economic outcomes. The framework builds on the technology of skill formation (Cunha and Heckman, 2007, 2008), the long-run effects of childhood interventions (Heckman et al., 2013), and the decomposition of intergenerational income persistence (Blanden et al., 2007).

Let $p_i \equiv \log y_i^p$ denote parental log income, $Y_i \equiv \log y_i^c$ adult log income, and $\mathbf{s}_{i0} = (s_{1i0}, \ldots, s_{Ki0})'$ a vector of behavioral characteristics measured early in life. Let $T_i \in \{0,1\}$ denote treatment assignment.

**Family background and early behavioral characteristics.** For socioemotional and behavioral skills to contribute to the persistence of income across generations, they must

first be associated with family economic circumstances. We describe this relationship with the following equation:

$$\mathbf{s}_{i0} = \boldsymbol{\mu}_0 + \mathbf{B}p_i + \mathbf{u}_{i0}, \tag{1}$$

where $\boldsymbol{\mu}_0$ is a vector of intercepts, the vector $\mathbf{B}$ captures the socioeconomic gradient in each behavioral characteristic, and $\mathbf{u}_{i0}$ captures individual differences not associated with parental income. Each element $B_k$ therefore measures the association between parental income and behavioral characteristic $k$ early in life. For a behavioral difficulty, $B_k < 0$ indicates that the difficulty is more prevalent among children from lower-income families. For a positive behavioral characteristic, $B_k > 0$ indicates that it is more prevalent among children from higher-income families.

**Initial behavioral characteristics and adult outcomes.** Next, we assume that initial skills predict income in adulthood and that an intervention may change this relationship:

$$Y_i = a + \beta^d p_i + \boldsymbol{\phi}_C' \mathbf{s}_{i0} + \tau T_i + \boldsymbol{\theta}'(T_i \mathbf{s}_{i0}) + \varepsilon_i. \tag{2}$$

The vector $\boldsymbol{\phi}_C$ captures the relationship between initial behavioral characteristics and adult income in the absence of the intervention. Note that this relationship does not need to be causal, but summarizes how differences observed early in life are associated with differences in adult outcomes. Nevertheless, this relationship must exist for an intervention to matter by changing how baseline skills translate into adult income. The parameter $\tau$ captures a level effect of the intervention, while $\boldsymbol{\theta}$ allows its effect to vary with initial behavioral characteristics. For characteristic $k$,

$$\phi_k^T = \phi_k^C + \theta_k, \tag{3}$$

and hence

$$\theta_k = \phi_k^T - \phi_k^C. \tag{4}$$

Thus, when treatment status is randomly assigned, $\theta_k$ characterizes treatment-effect heterogeneity with respect to a predetermined behavioral characteristic. It does not identify

the causal effect of that characteristic itself.

Consider a behavioral difficulty for which $B_k < 0$ and $\phi_k^C < 0$. Children from lower-income families display more of the difficulty, and the difficulty predicts lower adult income. If $\theta_k > 0$, the negative relationship between that initial difficulty and adult income is attenuated under the intervention.[1] The intervention may therefore reduce the extent to which early behavioral disadvantage predicts poor adult economic outcomes. However, the opposite pattern is also possible. Dynamic complementarities in skill formation imply that existing skills may increase the productivity of subsequent investments. If children with stronger initial skills benefit more from the intervention, the initial behavioral gradient may instead become steeper.

**Implications for intergenerational persistence.** Combining equations (1) and (2) provides a simple implication for the behavioral component of intergenerational persistence. In the absence of the intervention, this component is

$$\beta_C^B = \boldsymbol{\phi}_C' \mathbf{B}, \tag{5}$$

whereas under the intervention it is

$$\beta_T^B = (\boldsymbol{\phi}_C + \boldsymbol{\theta})' \mathbf{B}. \tag{6}$$

It follows that

$$\beta_T^B - \beta_C^B = \boldsymbol{\theta}' \mathbf{B}. \tag{7}$$

This expression concerns the component of intergenerational persistence associated with the socioeconomic patterning of the behavioral characteristics included in the model. It is not, in general, the total difference in the parent–child income gradient.

For a behavioral difficulty with $B_k < 0$, a positive $\theta_k$ implies $\theta_k B_k < 0$. The intervention would therefore weaken the contribution of that dimension of early behavioral disadvantage to intergenerational persistence.

[1] The magnitude of the attenuation depends on the ratio $\theta_k / |\phi_k^C|$. When $\theta_k \approx |\phi_k^C|$, the negative control-group gradient is approximately eliminated among treated boys; when $\theta_k$ is small relative to $|\phi_k^C|$, the gradient is attenuated but remains negative. If $\theta_k > |\phi_k^C|$, the gradient changes sign among treated boys.

Importantly, this implication does not require $\phi_k^C$ to represent the causal effect of behavior. The empirical question is whether characteristics that are socially graded early in life and predictive of later disadvantage become less predictive of adult economic outcomes under the intervention.

**Post-treatment behavioral development.** One possible explanation for such a pattern is that the intervention changes the subsequent development of the behavioral characteristics on which children initially differ. Let $\mathbf{s}_{it}$ denote behavior measured at a later age $t$:

$$\mathbf{s}_{it} = \boldsymbol{\mu}_t + \boldsymbol{\Lambda}_t \mathbf{s}_{i0} + \boldsymbol{\alpha}_t T_i + \boldsymbol{\Delta}_t (T_i \mathbf{s}_{i0}) + \boldsymbol{\Gamma}_t p_i + \mathbf{u}_{it}. \tag{8}$$

The matrix $\boldsymbol{\Lambda}_t$ captures the conditional persistence of early behavioral characteristics. $\boldsymbol{\alpha}_t$ captures treatment-induced level differences, while $\boldsymbol{\Delta}_t$ captures whether the relationship between initial and subsequent behavior differs between treated and control boys.

The framework therefore allows treatment effects to differ with initial behavioral characteristics either because the intervention changes subsequent measured behavioral development or through other processes that need not generate persistent differences in those measured behaviors.

## 2.2 Testing the Framework

For ease of exposition, we retain the notation introduced in the conceptual framework throughout the empirical implementation. The framework generates three empirical objects of interest: (i) the socioeconomic gradient in early behavioral characteristics, $\mathbf{B}$; (ii) the relationship between these characteristics and adult economic outcomes, $\boldsymbol{\phi}$; and (iii) treatment-effect heterogeneity with respect to predetermined behavioral characteristics, $\boldsymbol{\theta}$. No single sample in our setting identifies all three objects. We therefore combine evidence from a population cohort and a randomized childhood intervention. The QLSKC allows us to establish the population relationships linking family background, childhood behavior, and adult income. The MLES experiment allows us to test whether the long-run effect of an intervention varies with behavioral characteristics measured before treatment assignment.

**Step 1 — Population gradients and the behavioral component of intergenerational persistence.** We first use QLSKC boys to estimate the socioeconomic gradients $\mathbf{B}$

for inattention, disruptiveness, and prosociality. We then estimate

$$Y_i = a_Q + \beta_Q p_i + \boldsymbol{\phi}_Q' \mathbf{s}_{i0} + \boldsymbol{\gamma}' \mathbf{X}_i + \varepsilon_i, \tag{9}$$

where, in our preferred specification, $Y_i$ is the log of average adult income, $p_i$ is the log of average parental income when the boy was aged 15–19, and $\mathbf{s}_{i0}$ is the vector of behavioral skills measured at age six. The vector $\mathbf{X}_i$ includes parental education, measured as the highest number of years of schooling completed by either parent, and parental ages at the child's birth. Following Blanden et al. (2007), we combine the estimates of $\mathbf{B}$ and $\boldsymbol{\phi}_Q$ to quantify the contribution of early behavioral characteristics to observed intergenerational income persistence. This decomposition is descriptive and is not given a causal interpretation.

We also use the population estimates to construct an external benchmark for adult outcomes in the experimental sample:

$$BAI_i^{QLSKC} = Y_i - \widehat{E}_{QLSKC}\left[Y_i \mid p_i, \mathbf{s}_{i0}\right]. \tag{10}$$

Benchmark-adjusted income measures the difference between realized adult income and the income predicted by the population relationship given family background and the behavioral profile observed before the intervention. We compare this outcome between treated and control boys in Step 2. Although the QLSKC provides a population benchmark, the relationship between family background and adult income may be nonlinear.[2] For this reason, we also consider an alternative benchmark using non-disruptive boys from the same schools and cohort. This provides a more local observational reference group drawn from the same institutional setting as the experimental sample, without requiring the population relationship estimated in the QLSKC to extrapolate linearly to the MLES population.

$$BAI_i^{REF} = Y_i - \widehat{E}_{REF}\left[Y_i \mid p_i, \mathbf{s}_{i0}\right]. \tag{11}$$

**Step 2 — Experimental starting point.** The average long-run effect of the intervention on adult income has been documented elsewhere (Algan et al., 2022). We therefore do not view the average treatment effect on adult income as a contribution of the present

[2]For example, Connolly and Haeck (2026) document a steeper relationship between parental income and children's adult income among children from lower-income families in Canada, a pattern that is also observed in the United States (Chetty et al., 2014).

paper. Instead, this established result provides the experimental starting point for our main question: whether the long-run effect of the intervention varies systematically with behavioral characteristics measured before treatment assignment.

For comparability with the population analysis, we first express adult income relative to the population benchmark and estimate:

$$BAI_i^{QLSKC} = \mu + \tau T_i + u_i. \tag{12}$$

The coefficient $\tau$ measures the average difference in benchmark-adjusted income between treated and control boys. It tells us how much more or less treated boys earn relative to what their parental income and age-six behavioral characteristics would predict, compared with control boys.

In principle, this difference may reflect differences in realized income or in the predicted benchmark. However, Table 1 shows no significant differences between treatment and control in the variables used to construct predicted income. We therefore interpret $\tau$ as capturing differences in realized adult income across treatment assignment, complementing the evidence in Algan et al. (2022). Our main contribution is to examine whether the treatment effect varies with the levels of inattention, disruptiveness, and prosociality displayed before the intervention, which we do in Step 3.

**Step 3 — Treatment-effect heterogeneity by initial behavioral characteristics.** Our focus is to test whether the intervention shifts adult income by changing how skills measured before the intervention predict outcomes later in life (see equation 7). To do so, we estimate regressions of the following form:

$$BAI_i^{QLSKC} = \mu + \tau T_i + \rho p_i + \boldsymbol{\phi}_C' \mathbf{s}_{i0} + \boldsymbol{\theta}'(T_i \mathbf{s}_{i0}) + u_i. \tag{13}$$

Our main specification includes inattention ($I$), disruptiveness ($D$), and prosociality ($P$) jointly. Since our main interest is to examine whether the interaction terms are jointly significant, our primary hypothesis is

$$H_0^{\theta} : \boldsymbol{\theta} = \mathbf{0}. \tag{14}$$

The joint significance of the interaction terms would indicate that the intervention changes

how behavioral characteristics measured before treatment predict adult outcomes. This interpretation is particularly relevant if these characteristics are predictive of adult outcomes in the control group. For this reason, we also test

$$H_0^{\phi} : \boldsymbol{\phi}_C = \mathbf{0}. \tag{15}$$

Although our main focus is on joint significance, we also examine which behavioral dimensions account for the joint pattern. To do so, we complement the full $I + D + P$ specification with three leave-one-skill-out specifications:

$$I + P, \qquad I + D, \qquad D + P. \tag{16}$$

For the joint hypotheses, we report robust Wald tests and Freedman and Lane (1983) residual-permutation tests.[3] For individual treatment-by-skill interactions, we additionally account for multiple testing across the three behavioral dimensions using the Romano and Wolf (2005) step-down procedure.

# 3 Data

Our analyses provide empirical evidence on how socioemotional skills measured in kindergarten contribute to the persistence of income across generations and how an early childhood intervention influences this relationship. To do so, we need data on socioemotional skills measured in kindergarten and measures of permanent income for both children and their parents. More importantly, we need experimental variation from an intervention targeting these skills, or at least some of their underlying dimensions.

We use two complementary datasets. First, we use the Quebec Longitudinal Study of Kindergarten Children (QLSKC), a population-based longitudinal cohort that collected repeated measures of children's behavioral skills and was subsequently linked to administrative

[3] We implement the Freedman–Lane procedure by first estimating the restricted model under the null hypothesis and retaining its fitted values and residuals. In each permutation, we randomly permute the restricted-model residuals, add them to the restricted fitted values to construct a pseudo-outcome, and re-estimate the unrestricted model. We then compare the observed joint Wald statistic with the distribution of the corresponding statistics across permutations. For tests of $H_0^{\theta}$, the restricted model excludes the treatment-by-skill interactions being jointly tested while retaining the treatment main effect and the baseline skill controls. For tests of $H_0^{\phi}$, it excludes the corresponding control-group skill gradients while retaining the treatment main effect and treatment-by-skill interactions. We use 2,000 permutations throughout.

tax records. These data allow us to document how early behavioral skills contribute to the persistence of income across generations. Second, we use the Montreal Longitudinal Experimental Study (MLES), a randomized childhood intervention linked to participants' and parents' administrative tax records, to explore how the intervention influences the relationship between early behavioral difficulties and adult income.

Both datasets were linked to federal tax files through a partnership with Statistics Canada (Findlay et al., 2018). Figure 1 summarizes the timing of data collection and administrative follow-up in the two samples.

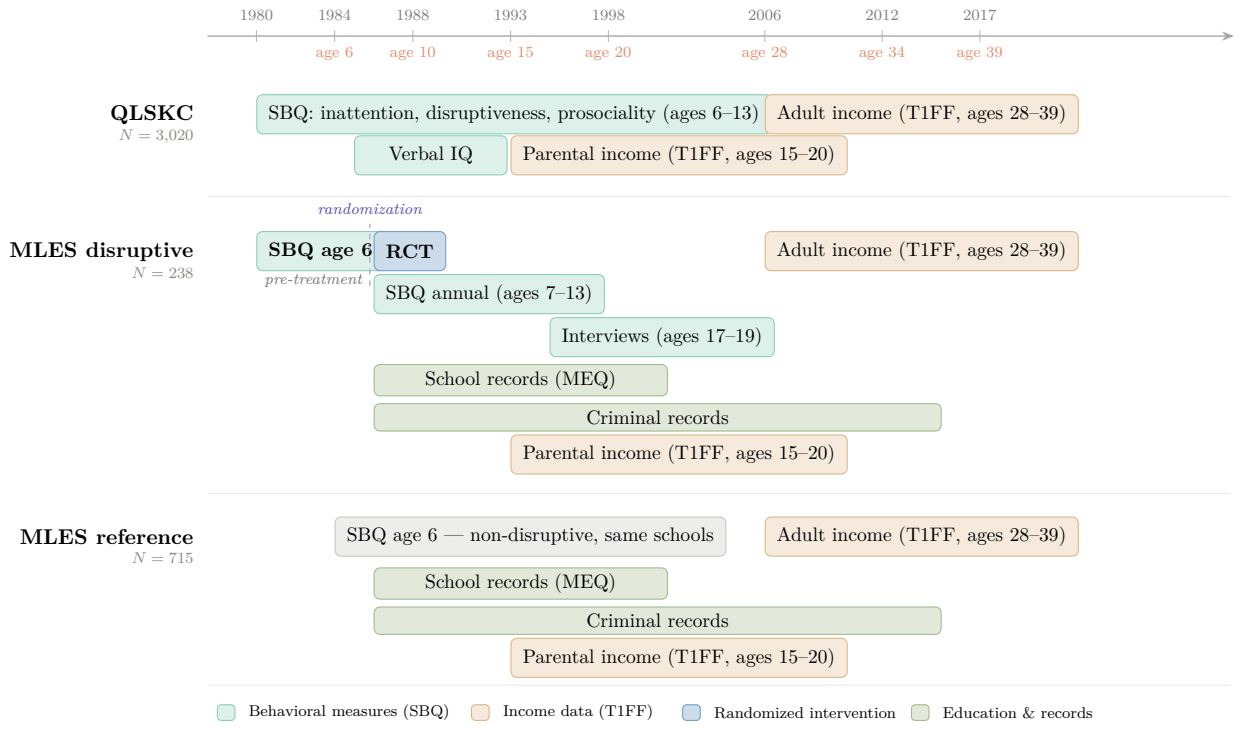



**Figure 1:** Study design, data sources, and follow-up periods

*Notes:* Each row shows the timing of data collection for one sample. The QLSKC follows a population-based cohort of Quebec children from kindergarten through adulthood. The MLES follows boys recruited in disadvantaged Montreal schools. Baseline behavioral characteristics are measured at age 6, before the intervention. The intervention was delivered between ages 7 and 9. Behavioral outcomes are subsequently observed through age 13, educational and criminal outcomes through age 24, and adult economic outcomes through administrative tax records. The non-disruptive reference group is drawn from the same schools and cohort as the MLES experimental sample but was not randomized.

## 3.1 The QLSKC Population Sample

The QLSKC is a prospective longitudinal study of 3,020 children born in 1980–1981 and enrolled in kindergarten across the province of Quebec. The cohort was selected using a strat-

ified design to ensure representativeness with respect to the socioeconomic and geographic composition of the province (Rouquette et al., 2014). Teachers completed the Social Behavior Questionnaire (SBQ) annually from ages 6 to 13, providing repeated measures of inattention, hyperactivity, opposition, aggression, and prosociality (Tremblay et al., 1987). At age 13, participants also completed a standardized verbal ability task (Lorge and Thorndike, 1954), which we use as a measure of cognitive ability in adolescence.

**Behavioral skill composites.** Hyperactivity, opposition, and aggression are closely related dimensions of externalizing behavior (Achenbach, 1991; Tremblay et al., 1987). We combine these three measures into a disruptiveness index defined as their average. We retain inattention and prosociality as separate dimensions, yielding three baseline behavioral constructs: inattention, disruptiveness, and prosociality.[4] These dimensions capture distinct aspects of childhood socioemotional functioning and exhibit different developmental trajectories and associations with later outcomes (Moffitt, 1993; Pingault et al., 2014).

For comparability across the population and experimental analyses, we standardize each of the three behavioral constructs using its mean and standard deviation among boys in the QLSKC. We apply these same population-based standardization parameters to the MLES sample.[5] Our main analyses use behavioral characteristics measured at age 6, before the intervention began in the MLES. We additionally use subsequent SBQ assessments through age 13. Because the intervention was delivered between ages 7 and 9, these repeated assessments allow us to describe the evolution of behavioral skills during and after the intervention and to examine whether the relationship between baseline behavior and subsequent behavior differs between boys assigned to treatment and control.

## 3.2 The MLES Experimental Sample

The Montreal Longitudinal-Experimental Study (MLES) followed a cohort of boys attending schools in low-income neighborhoods of Montreal. The initial sample consisted of 1,161 boys enrolled in kindergarten in low-income neighborhoods. The original study was restricted

[4]This grouping also maintains consistency with the MLES experimental design, in which eligibility for randomization was based on a measure of disruptive behavior constructed from related externalizing behaviors.

[5]Appendix Table A.1 compares the QLSKC and MLES samples using characteristics observed in both studies.

to white francophone boys whose parents had no more than 14 years of education. These eligibility criteria reflect the population targeted by the original intervention and necessarily limit the population to which the experimental estimates can be directly generalized.

Among the 1,037 boys satisfying the study eligibility criteria, 250 who scored above the 70th percentile of antisocial behavior on the SBQ were classified as disruptive. These boys were randomly assigned, through a simple lottery, to the preventive intervention (69 boys) or to the control group (181 boys). The intervention began when the boys were age 7 and continued for two years, until age 9. It combined child-focused social-skills training with a structured parent-training component grounded in social learning theory.

Of the 69 boys assigned to the intervention, 46 ultimately participated in the program, while 23 refused participation. Our analysis nevertheless retains boys according to their original randomized assignment, irrespective of actual program participation.[6] Detailed descriptions of the intervention design, implementation, and compliance are provided by Tremblay et al. (1991), Boisjoli et al. (2007), and Algan et al. (2022).

After applying the adult-income observation restrictions described below, the analytical sample used for the main income analysis contains 57 boys assigned to treatment and 144 boys assigned to control. Throughout the analysis, treatment status is defined by randomized assignment. Comparisons between these groups therefore identify intention-to-treat effects of assignment to the intervention.

Boys scoring below the disruptiveness threshold form a non-disruptive reference group drawn from the same schools and cohort ($N = 637$ in the analytical sample). We use this group as an alternative observational benchmark for the experimental sample. Comparisons between disruptive boys and this reference group are not randomized and are therefore not given a causal interpretation. Instead, the reference group provides a more local benchmark drawn from the same institutional setting as the experimental sample. As described in Section 2, we use it to assess whether our conclusions depend on applying the population relationship estimated in the QLSKC to the MLES sample.

[6] We follow previous studies of the MLES in defining treatment according to the original randomized assignment rather than actual program participation (Boisjoli et al., 2007; Algan et al., 2022). Accordingly, all experimental estimates in this paper should be interpreted as intention-to-treat (ITT) effects.

## 3.3 Administrative Income Data

Individual tax files are available from 1998 to 2017 for participants, and parental tax returns are available from 1991 to 2017. Following the intergenerational mobility literature (Chetty et al., 2014; Connolly et al., 2019), we construct adult income using income observed between ages 30 and 36 and parental socioeconomic status using household income observed when the child was aged 15 to 19.[7]Adult income captures labor-market earnings and includes income from employment and self-employment. It excludes government benefits and transfers. Parental income, by contrast, includes income from all sources to capture the economic resources available to the household during the child's adolescence.

Parental household income aggregates the income of both parents when present and is equivalized by dividing by the square root of the number of dependents. Following Algan et al. (2022), we convert nominal income to constant 2014 Canadian dollars assuming a constant annual inflation rate of 3 percent. Using multi-year averages for both generations reduces the influence of transitory income fluctuations on measures of long-run economic status (Solon, 1992; Connolly and Haeck, 2024).

To reduce measurement error from incomplete income histories, our main income measure requires at least three observed tax returns within the relevant age window. The main analysis uses ages 30–36. We also examine alternative age windows, 28–33 and 28–40, which trade off the age at which income is measured against the number of available observations; these results are reported in the appendix.

Requiring observed tax records may induce selection if income observation is related to treatment assignment or to baseline behavioral characteristics. We therefore examine whether adult-income observation varies systematically with treatment assignment, baseline behavior, or their interactions in the empirical analysis.

## 3.4 Balance in the Experimental Sample

Table 1 compares observed characteristics and data availability across experimental assignment. Individual standardized mean differences are small, and none of the individual

[7]This income window occurs after the intervention and parental income is therefore not strictly predetermined with respect to treatment assignment. In principle, the intervention could have affected subsequent parental income, for example through changes in parental behavior or labor supply. In practice, however, Table 1 shows virtually identical parental income across the treatment and control groups, providing no evidence of a differential effect of treatment assignment on parental income over this period.

randomization-inference tests rejects equality between treatment and control at the 5 percent level. The joint test of the pre-treatment characteristics reported in the table yields $F = 1.638$ ($p = 0.071$).

**Table 1:** Balance Table: Pre-Treatment Characteristics

| | Treatment | | Control | | | |
|---|---|---|---|---|---|---|
| Variable | $N$ | Mean | $N$ | Mean | SMD | RI $p$ |
| *Panel A: Family background* | | | | | | |
| Year of birth | 70 | 1977.73 | 170 | 1977.70 | 0.07 | 0.741 |
| Parental income (adj.) | 60 | 27,300 | 170 | 28,400 | -0.08 | 0.648 |
| Log parental income | 60 | 10.13 | 170 | 10.08 | 0.10 | 0.547 |
| Parental education (yrs) | 70 | 10.59 | 170 | 10.76 | -0.03 | 0.829 |
| Parental age at birth | 70 | 23.44 | 170 | 21.75 | 0.17 | 0.268 |
| Living with both parents | 70 | 0.33 | 170 | 0.28 | 0.10 | 0.454 |
| *Panel B: Age-6 behavioral skills (SBQ)* | | | | | | |
| Inattention | 70 | 3.59 | 170 | 3.34 | 0.10 | 0.518 |
| Hyperactivity | 60 | 2.40 | 170 | 2.30 | 0.07 | 0.666 |
| Opposition | 70 | 4.47 | 170 | 4.14 | 0.12 | 0.446 |
| Prosociality | 70 | 7.15 | 170 | 7.22 | -0.01 | 0.917 |
| Aggression | 70 | 9.00 | 170 | 9.93 | -0.12 | 0.418 |
| *Panel C: Cognitive ability and income observation* | | | | | | |
| Verbal IQ (age 13) | 60 | 7.54 | 170 | 7.46 | 0.02 | 0.895 |
| N valid tax years (30–36) | 60 | 6.34 | 150 | 6.31 | 0.02 | 0.906 |
| N valid tax years (28–33) | 50 | 4.58 | 130 | 4.52 | 0.10 | 0.609 |
| N valid tax years (28–40) | 60 | 8.00 | 160 | 7.82 | 0.10 | 0.541 |
| N valid parental tax years | 60 | 15.78 | 170 | 15.28 | 0.06 | 0.693 |
| Joint $F$-test (all variables) | | | | | | $F = 1.638$, $p = 0.071$ |

*Notes:* SMD denotes the standardized mean difference, $(\bar{x}_T - \bar{x}_C)/s_{\text{pooled}}$. RI $p$-values are based on 2,000 permutations of treatment assignment within the randomized treatment and control sample. Sample sizes vary across characteristics because of item-level missing data and are rounded to the nearest 10 in accordance with Statistics Canada disclosure requirements. The joint test regresses treatment assignment on the pre-treatment characteristics reported in the table.

# 4 Population Evidence: Behavioral Skills and Income Persistence

We begin by documenting the role of behavioral skills at school entry in intergenerational income persistence among Quebec boys, using the QLSKC population data. Following Blan-

den et al. (2007), we estimate for each skill the relationship with parental log income (the SES gradient, $\hat{\beta}_k$), the association with adult log income conditional on parental income (the income gradient, $\hat{\phi}_k$), and the contribution of each skill to the total intergenerational income elasticity ($C_k = \hat{\beta}_k \hat{\phi}_k$). These estimates are reported in Table 2.

It should be noted that we do not control for later-life characteristics such as education, labor-market attachment, or cognitive skills measured after school entry. This is intentional. Our interest here is in the overall association between behavioral characteristics at age six and adult income, which may partly reflect pathways operating through subsequent education and labor-market outcomes. Indeed, Blanden et al. (2007) show that the contribution of childhood noncognitive skills to intergenerational income persistence falls substantially once education and labor-market attachment are accounted for. We leave these potential pathways open at this stage. In the experimental sample, we subsequently examine whether the intervention changes how initial behavioral characteristics predict adult income and whether similar patterns emerge in other adult outcomes. [8]

[8] Appendix Table D.1 examines the relationship between childhood behavioral skills and intergenerational mobility more directly, using upward and downward income mobility as outcomes. Specifications based on skill quartiles allow for nonlinear associations and jointly control for the four behavioral dimensions. Higher inattention is associated with less upward and more downward mobility, while greater prosociality is associated with less downward mobility.

**Table 2:** Early Behavioral Skills and Intergenerational Income Persistence: Blanden et al. (2007) Decomposition

| | SES gradient $\hat{\beta}_k$ | Income gradient $\hat{\phi}_k$ | Contribution $\hat{\beta}_k\hat{\phi}_k$ | 95% CI | Share of IGE (%) |
|---|---|---|---|---|---|
| *QLSKC boys (primary sample, $N = 1{,}400$)* | | | | | |
| Inattention | $-0.124^{***}$ | $-0.156^{**}$ | 0.019 | [−0.001, 0.039] | 3.14 |
| | (0.042) | (0.057) | | | |
| Disruptiveness | $-0.128^{***}$ | $-0.257^{***}$ | 0.034 | [0.001, 0.067] | 5.52 |
| | (0.043) | (0.065) | | | |
| Prosociality | $0.148^{***}$ | $0.188^{***}$ | 0.028 | [0.009, 0.047] | 4.51 |
| | (0.043) | (0.059) | | | |
| All behavioral skills | | | 0.081 | [0.031, 0.131] | **13.17** |
| Direct (parental income) | | | 0.542 (0.096) | | 87.99 |
| **Total IGE** | | | **0.617** (0.097) | | **100.0** |

*Notes:* $\hat{\beta}_k$ is from a regression of standardized skill $k$ (age six) on log parental income plus controls. $\hat{\phi}_k$ is from a joint regression of adult log income on parental log income and all three skills. Skills standardized on the QLSKC boys sample. 95% confidence intervals on contributions $C_k = \hat{\beta}_k\hat{\phi}_k$ are computed by the delta method: $\widehat{\text{Var}}(C_k) \approx \hat{\phi}_k^2\,\widehat{\text{Var}}(\hat{\beta}_k) + \hat{\beta}_k^2\,\widehat{\text{Var}}(\hat{\phi}_k)$. The 95% CI on the total behavioral contribution uses the sum of delta-method variances (conservative). Standard errors in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$.

# 5 Experimental Evidence

In this section, we first examine whether the intervention improves adult income relative to what family background and early behavioral characteristics would predict. Building on the descriptive patterns documented in the population sample, we then examine whether the relationship between behavioral characteristics measured before the intervention and adult income differs between treated and control boys.

## 5.1 Average Treatment Effect

As emphasized in Section 2, we go beyond a simple comparison of intergenerational income elasticities across treatment groups. Instead, we ask whether the intervention improves

adult outcomes beyond what family background and early behavioral characteristics would predict.[9] The benchmark-adjusted income measure discussed in Section 2 serves this purpose.

Figure 2 plots the distribution of $BAI_i$ separately for treated boys, control boys, and the non-disruptive reference group. The figure suggests that the intervention shifted adult income above the levels predicted by parental income and age-six inattention, disruptiveness, and prosociality. First, negative BAI values are more common among control boys, while the distribution for treated boys is shifted to the right. Second, the distribution for treated boys overlaps substantially with that of the non-disruptive reference group, with a non-negligible share of treated boys achieving even higher BAI values. This second comparison, however, is descriptive. It suggests that the intervention may have helped treated boys achieve outcomes comparable to those of non-disruptive boys, but the reference group was not part of the random assignment.[10]

Table 3 reports estimates of the average treatment effect on benchmark-adjusted income (see Equation 12). Panel A presents the main estimates, while Panels B to D assess their robustness. Panel A shows that assignment to treatment increases benchmark-adjusted income by 1.46 log points. To put this magnitude in perspective, the observed difference in benchmark-adjusted income between non-disruptive reference boys and randomized control boys is 0.71 log points. The experimental treatment effect is therefore approximately twice as large as this descriptive difference. The magnitude of the treatment effect is substantial,

[9] A direct comparison of intergenerational income elasticities across experimental groups could itself provide evidence that the intervention altered intergenerational income persistence. Such a comparison, however, can be difficult to interpret. A difference in the estimated IGE captures a difference in the slope of the parent–child income relationship, but does not by itself reveal whether that difference reflects improved outcomes among children from disadvantaged families. In particular, greater mobility need not imply greater upward mobility: a flatter parent–child income relationship can also arise when some children attain lower adult incomes than would otherwise be expected given their family background. Bloome (2017), for example, shows that lower intergenerational persistence can coexist with greater economic insecurity and downward mobility. More generally, the literature documents important nonlinearities in the intergenerational income relationship, whose shape varies across settings (Chen et al., 2017). Our benchmark-adjusted approach therefore asks a different question: whether participants attain higher or lower adult income than predicted by their initial family and behavioral characteristics.

[10] As a complementary exercise, we also examine the intergenerational income relationship directly across the parental income distribution. Following the logic of Chetty and Hendren (2018), we divide parental income into four bins and use non-disruptive boys to construct a benchmark relating parental income to adult outcomes. Appendix Figure F.2 plots the distribution of deviations from this benchmark for disruptive boys assigned to treatment and control. The distribution for treated boys is shifted to the right relative to that of control boys. This suggests that treated boys tend to attain adult income positions above those predicted by parental income, and by more than their control-group counterparts. Unlike our main benchmark-adjusted income measure, however, this exercise conditions only on parental income and is constructed at the bin level. We therefore view it as complementary descriptive evidence rather than as our main experimental outcome.

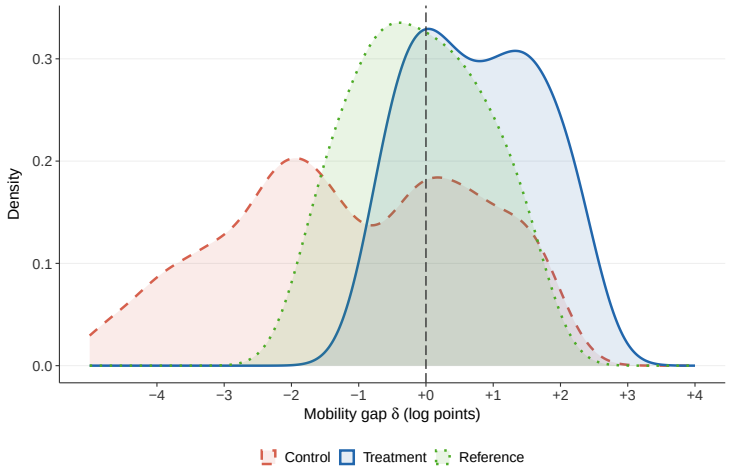



**Figure 2:** Distribution of Benchmark-Adjusted Adult Income by Group

*Notes:* Benchmark-adjusted income $BAI_i$ is defined in equation (10). Densities are estimated with a Gaussian kernel. Vertical lines indicate group means. The dashed line at zero indicates equality between realized adult income and the income predicted by the population benchmark. Reference: $N = 637$; Control: $N = 144$; Treatment: $N = 57$.

though it should be interpreted with caution. A difference of 1.46 log points corresponds to a large proportional increase, but as we show below, the estimated treatment effect is sensitive to the treatment of zero-earnings years.[11] This interpretation is also consistent with the pronounced left tail of the control-group distribution in Figure 2.[12]

Panel B shows that the result is robust to several alternative inference methods. The effect remains statistically significant under randomization inference ($p = 0.006$), which relies on the experimental assignment mechanism and is particularly useful in small samples. We obtain similar conclusions using the bootstrap percentile ($p = 0.002$) and studentized bootstrap ($p < 0.001$) methods. The worst-case maximum statistic (De Haan, 1981) yields a slightly higher $p$-value ($p = 0.058$). This test evaluates the sharp null of no individual-level treatment effect across subgroups of the parental income distribution and is more conservative by construction.[13]

[11] The percentage equivalent is calculated as $100 \times (e^{1.46} - 1) \approx 331\%$.

[12] We also consider income measured in percentile ranks. Although the estimated treatment effect points in the same direction, it is imprecisely estimated and is not statistically significant at the 10% level. Appendix Figure F.1 plots the difference between adult and parental income ranks. The distributions appear to differ mainly in the left tail, with little visible difference in the right tail. This pattern is consistent with the possibility that the intervention helped limit large downward movements rather than substantially increasing upward mobility.

[13] Our choice of inference methods follows the literature on small-sample randomized experiments. See, for example, García et al. (2023) in the context of the Perry Preschool experiment.

Panel C addresses two additional concerns. First, $BAI_i$ is constructed using parameters estimated in a first stage on the QLSKC sample. A two-stage bootstrap that propagates this first-stage uncertainty yields a standard error of 0.423, essentially identical to the analytic standard error of 0.428, indicating that first-stage estimation uncertainty is small relative to experimental sampling variation.[14] Second, the AIPW estimator yields an estimate of 1.494 (SE = 0.424, $p < 0.001$), close to the OLS estimate.

Panel D reports Lee (2009) bounds, which address the possibility that the result is driven by differential selection into the income sample across treatment and control groups. The lower bound is 1.413 and the upper bound is 1.571. Even at the lower end of the 90% confidence interval for the lower bound (0.489), the estimate remains positive.[15]

**Table 3:** Average Treatment Effect on Benchmark-Adjusted Adult Income

| | Estimate | $p$-value |
|---|---|---|
| *Panel A: Primary estimates* | | |
| Treatment − Control | 1.460*** | 0.001 |
| | (0.426) | |
| Reference − Control | 0.713** | 0.029 |
| | (0.327) | |
| *Panel B: Alternative inference (Treatment − Control)* | | |
| Randomization inference (2,000 permutations) | | 0.006 |
| Bootstrap percentile (1,000 draws) | | 0.002 |
| Bootstrap studentized (1,000 draws) | | < 0.001 |
| Worst-case max statistic (De Haan, 1981) | | 0.058 |
| *Panel C: Alternative estimators* | | |
| Two-stage bootstrap SE | (0.423) | 0.002 |
| SE ratio (two-stage / analytic) | 0.988 | |
| AIPW estimator | 1.494*** | < 0.001 |
| SE (AIPW) | (0.424) | |
| *Panel D: Lee (2009) attrition bounds* | | |
| Lower bound | 1.413 | |
| Upper bound | 1.571 | |
| 90% CI on lower bound | [0.489, 2.030] | |
| 90% CI on upper bound | [1.024, 2.342] | |
| Observations (Treatment / Control) | 57/144 | |
| Mean $BAI_i$ (Treatment / Control / Reference) | 0.401/ − 1.059/ − 0.345 | |

*Notes:* The outcome is benchmark-adjusted adult income ($BAI_i$), defined in equation (10). Panel A reports OLS estimates from equation (12), with standard errors in parentheses. Panel B reports alternative inference for the treatment-control contrast. Panel C reports two-stage bootstrap and AIPW estimates. Panel D reports Lee bounds under monotone selection, with 90% confidence intervals from 500 bootstrap draws. ***$p < 0.01$; **$p < 0.05$.

[14] Under randomization, the first-stage prediction error $(\hat{\boldsymbol{\phi}}_Q - \boldsymbol{\phi}_Q)'\mathbf{s}_{i0}$ is uncorrelated with $T_i$ and does not bias the estimated treatment effect.

[15] This is not surprising given the similar matching rates reported in Table A.1. Appendix Figure E.1 shows matching rates by age.

## 5.2 Does the Intervention Change How Age-Six Skills Predict Adult Outcomes?

The results above show that the MLES intervention helped treated boys attain adult income levels above those predicted by their family background and behavioral profile. Previous evaluations of the same intervention have documented positive effects on young-adulthood outcomes, including criminal activity, secondary-school completion, and labor-market outcomes (Algan et al., 2022).

We now ask whether the intervention also changes the relationship between behavioral characteristics measured before treatment and income in adulthood. To examine this, we estimate regressions of the following form:

$$BAI_i = \mu + \sum_{k=1}^{3} \phi_C^k s_{i0}^k + \sum_{k=1}^{3} \theta^k (s_{i0}^k \cdot T_i) + \tau T_i + \rho p_i + \nu_i, \tag{17}$$

where $\phi_C^k$ measures the association between baseline skill $k$ and benchmark-adjusted adult income among control boys, while $\theta^k$ measures the difference in this gradient between treated and control boys. The corresponding gradient among treated boys is $\phi_T^k = \phi_C^k + \theta^k$. Because treatment was randomly assigned at the individual level, we do not cluster standard errors. We report HC2 standard errors to account for heteroskedasticity.

Our primary hypothesis concerns the vector of treatment-by-skill interactions, $H_0 : \boldsymbol{\theta} = \mathbf{0}$, which tests whether treatment effects vary jointly with the behavioral characteristics measured at school entry. We also test $H_0 : \boldsymbol{\phi}^C = \mathbf{0}$ to assess whether these baseline behavioral characteristics are jointly associated with adult outcomes among control boys. Given the small experimental sample, we complement HC2-robust Wald tests with Freedman–Lane permutation inference. For the individual treatment-by-skill interactions, we additionally report Romano and Wolf (2005) step-down adjusted $p$-values across the three behavioral dimensions.

Table 4 presents the results. Panel A reports the individual skill gradients, treatment-by-skill interactions, and implied treated-group gradients. Panels B and C complement the joint tests with leave-one-skill-out specifications. In each case, we re-estimate the model after excluding one behavioral dimension and jointly test the coefficients on the remaining dimensions. This exercise allows us to assess whether the evidence from the joint tests is

concentrated in a particular baseline behavioral dimension. If the joint rejection is driven by a single dimension, the leave-one-skill-out tests will reject strongly when that dimension is retained and fail to reject when it is excluded.

Finally, we examine the sensitivity of the results to the construction of the adult economic outcome. Our primary specification uses benchmark-adjusted income constructed from the relationship estimated in the QLSKC population sample. We compare this specification with an alternative benchmark estimated among the non-disruptive boys attending the same schools as the MLES experimental participants, and with raw adult log income. The latter specification removes the benchmarking step altogether.

Our results indicate that treatment-effect heterogeneity is concentrated along the inattention dimension. Among disruptive control boys, greater inattention at school entry is associated with substantially lower adult economic outcomes, consistent with the population relationship documented above. For our primary outcome, a one-standard-deviation increase in baseline inattention is associated with a 0.711-log-point lower benchmark-adjusted adult income among control boys ($\hat{\phi}_I^C = -0.711$, SE = 0.350). In contrast, the treatment-by-inattention interaction is positive and large ($\hat{\theta}_I = 1.124$, SE = 0.413), implying a much flatter gradient among treated boys ($\hat{\phi}_I^T = 0.413$, SE = 0.223). Thus, the positive interaction does not imply that greater inattention improves adult outcomes among treated boys. Rather, the negative gradient observed among controls is substantially attenuated among treated boys, for whom the estimated gradient is positive but imprecisely estimated.

This pattern is robust to the choice of adult economic outcome. Using either the non-disruptive MLES boys as an alternative benchmark or raw adult log income yields a treatment-by-inattention interaction of 1.085 (SE = 0.399). The individual interaction is statistically significant under conventional and permutation inference and remains significant after Romano–Wolf adjustment across the three behavioral dimensions ($p_{\text{RW}} = 0.03$). By contrast, the treatment interactions with disruptiveness and prosociality are small and imprecisely estimated across all three outcome specifications.

The leave-one-skill-out tests in Panel B reinforce this interpretation. With all three behavioral dimensions included, the joint null of no treatment-effect heterogeneity yields permutation $p$-values of 0.08, 0.06, and 0.06 for QLSKC benchmark-adjusted income, reference-group benchmark-adjusted income, and raw log income, respectively. When inattention is excluded, however, there is no evidence of joint heterogeneity across disruptiveness and

prosociality: the corresponding permutation $p$-values rise to 0.67, 0.61, and 0.64. Conversely, specifications retaining inattention generally yield substantially smaller permutation $p$-values. Together with the individual estimates, these results indicate that the evidence of treatment-effect heterogeneity is concentrated along the inattention dimension.

Panel C examines the corresponding baseline skill gradients among control boys. For raw adult log income, the three behavioral dimensions are jointly associated with the outcome ($p_{\text{perm}} = 0.03$). Excluding inattention raises the permutation $p$-value to 0.16, whereas specifications retaining inattention yield $p$-values of 0.01. A similar, though weaker, pattern appears when the non-disruptive MLES boys are used to construct benchmark-adjusted income. For the QLSKC benchmark-adjusted outcome, by contrast, the joint null for the control-group skill gradients is not rejected in any specification. This difference is consistent with the construction of the outcome: the QLSKC benchmark already incorporates the population relationship between baseline behavioral skills and adult income, so benchmark adjustment removes part of the skill–income gradient observed in raw adult income.

**Understanding the effect size.** Whether we consider the average effect on income or the interaction between treatment and baseline inattention, the magnitude of our estimates is substantially larger than the earnings effects previously documented for the MLES by Algan et al. (2022). This difference is particularly noteworthy for the treatment-by-inattention interaction, which implies a 1.12 log-point increase for a one-standard-deviation increase in baseline inattention. One possible explanation comes from the lower tail of the income distribution. As shown in Figure 2, the largest differences between treated and control boys appear among individuals with very low income. We probe this interpretation in two ways. First, we re-estimate the heterogeneity specification (equation 13) using the percentage of observed years with zero earnings between ages 30 and 36 as the outcome. Second, we re-estimate the income specification among boys with no zero-earnings years over the same window. Table 5 reports both exercises using the same structure as Table 4.[16]

In Panel A, we report estimates of equation 13 using the percentage of observed years with zero earnings between ages 30 and 36 as the outcome. Among control boys, higher baseline inattention is associated with a greater proportion of years with zero earnings. A one-

[16] An alternative way to examine differences across the income distribution would be to estimate quantile treatment effects (e.g., Chernozhukov and Hansen, 2005). However, our small experimental sample limits the precision of such estimates, particularly at the lower quantiles.

**Table 4:** Treatment-Effect Heterogeneity by Baseline Behavioral Skills

| | Benchmark-adjusted income ($BAI_i$) | | | Reference benchmark | | Log income | |
|---|---|---|---|---|---|---|---|
| Skill | $\hat{\phi}_k^C$ | $\hat{\theta}_k$ | $\hat{\phi}_k^T$ | $\hat{\theta}_k$ | $p_{\text{RW}}$ | $\hat{\theta}_k$ | $p_{\text{RW}}$ |
| *Panel A: Individual skill estimates* | | | | | | | |
| Inattention | −0.711 | 1.124 | 0.413 | 1.085 | 0.03 | 1.085 | 0.03 |
| | (0.350) | (0.413) | (0.223) | (0.399) | | (0.399) | |
| Disruptiveness | 0.197 | −0.022 | 0.175 | 0.010 | 0.97 | 0.010 | 0.97 |
| | (0.158) | (0.224) | (0.160) | (0.223) | | (0.223) | |
| Prosociality | 0.243 | −0.131 | 0.112 | −0.142 | 0.92 | −0.142 | 0.92 |
| | (0.286) | (0.378) | (0.247) | (0.376) | | (0.376) | |

| | $BAI_i$ | | Reference benchmark | | Log income | |
|---|---|---|---|---|---|---|
| Skills | Wald | Perm. | Wald | Perm. | Wald | Perm. |
| *Panel B: Joint tests of treatment interactions* | | | | | | |
| I + D + P | 0.06 | 0.08 | 0.05 | 0.06 | 0.05 | 0.06 |
| I + P | 0.04 | 0.04 | 0.03 | 0.04 | 0.03 | 0.04 |
| I + D | 0.12 | 0.12 | 0.03 | 0.04 | 0.05 | 0.07 |
| D + P | 0.65 | 0.67 | 0.61 | 0.61 | 0.62 | 0.64 |
| *Panel C: Joint tests of control-group gradients* | | | | | | |
| I + D + P | 0.20 | 0.20 | 0.09 | 0.10 | 0.02 | 0.03 |
| I + P | 0.24 | 0.25 | 0.04 | 0.04 | 0.01 | 0.01 |
| I + D | 0.12 | 0.12 | 0.05 | 0.05 | 0.01 | 0.01 |
| D + P | 0.71 | 0.71 | 0.46 | 0.47 | 0.14 | 0.16 |

*Notes:* Panel A reports estimates from equation (13). $\hat{\phi}_k^C$ and $\hat{\phi}_k^T$ denote control- and treated-group skill gradients; $\hat{\theta}_k$ is their difference. HC2 standard errors are in parentheses. RW $p$-values use the Romano–Wolf step-down correction across three skill interactions within each outcome. Panel B tests the joint significance of treatment interactions; Panel C tests control-group gradients. Wald tests use HC2 standard errors; permutation tests use 2,000 Freedman–Lane permutations. The reference benchmark uses the non-disruptive group. I, D, and P denote inattention, disruptiveness, and prosociality.

standard-deviation increase in baseline inattention is associated with a 7.37 percentage point increase in this proportion (SE = 3.42). The interaction between treatment and baseline inattention is negative (−9.61 percentage points, SE = 4.06), suggesting that the intervention weakens this relationship. The interaction is statistically significant under randomization inference ($p = 0.02$) and remains significant at the 10% level after correcting for multiple hypothesis testing ($p = 0.08$).

In Panel B, we re-estimate the specification from Table 4, restricting the sample to boys with positive earnings in every observed year between ages 30 and 36. We find no statistically significant treatment interactions for any of the baseline behavioral skills. The interaction with inattention is small (0.04, RI $p = 0.81$). Together, these findings suggest that years with zero earnings may account for an important part of the estimated heterogeneity in adult income. However, because the sample restriction is based on post-treatment earnings, these estimates are descriptive and do not establish how much of the income effect operates through years without earnings.

**Table 5:** Understanding the Effect Size: Zero Earnings and Adult Income

| | Control gradient | | Treatment interaction | | | |
|---|---|---|---|---|---|---|
| Skill | $\phi_C$ | SE | $\theta$ | SE | RI $p$ | RW $p$ |
| *Panel A: Percentage of Years with Zero Earnings* | | | | | | |
| Inattention | 7.372** | (3.420) | -9.606** | (4.057) | 0.02 | 0.08 |
| Disruptiveness | -0.144 | (1.540) | 0.201 | (2.219) | 0.93 | 0.99 |
| Prosociality | -3.382 | (3.040) | 0.296 | (3.983) | 0.94 | 0.99 |
| Joint test: all interactions | | | | | 0.15 | — |
| *Panel B: Log Income, Positive Earnings at All Ages 30–36* | | | | | | |
| Inattention | 0.11 | (0.07) | 0.04 | (0.14) | 0.81 | 0.87 |
| Disruptiveness | 0.24*** | (0.03) | -0.04 | (0.06) | 0.49 | 0.87 |
| Prosociality | -0.15*** | (0.05) | -0.06 | (0.11) | 0.59 | 0.87 |
| Joint test: all interactions | | | | | 0.86 | 0.86 |
| Observations (Treatment / Control) | | | 43 / 91 | | | |

*Notes:* Panel A uses the percentage of observed years with zero employment earnings between ages 30 and 36 as the outcome. Panel B uses adult log income and restricts the sample to individuals with positive earnings in every observed year. Baseline behavioral skills are standardized. $\phi_C$ is the control-group gradient and $\theta$ is the treatment-by-skill interaction. Standard errors are in parentheses. RI denotes randomization inference; RW denotes Romano–Wolf adjusted $p$-values. Panel B conditions on post-treatment earnings and is interpreted as descriptive evidence rather than a causal intensive-margin effect. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

### 5.2.1 Does the Intervention Change Behavioral Skill Trajectories in Childhood?

Having documented that the intervention changes the relationship between baseline inattention and adult earnings, we now examine whether it also changes the persistence and cross-persistence of behavioral skills during and after the intervention. As Cunha and Heckman (2007) show, skills acquired at one stage influence the formation of skills at later stages. In our setting, this means that baseline behavioral skills may predict both their own subsequent development and the development of other behavioral skills. The intervention may change these relationships.

To examine this possibility, we estimate the following specification for each behavioral skill $k$ and period $t$[17]

$$s_{it}^k = \alpha_t^k + \sum_{j\in\{I,D,P\}} \beta_t^{kj} s_{i6}^j + \tau_t^k T_i + \sum_{j\in\{I,D,P\}} \delta_t^{kj}(s_{i6}^j \times T_i) + \gamma_t^k p_i + \varepsilon_{it}^k. \tag{18}$$

Here, $I$, $D$, and $P$ denote inattention, disruptiveness, and prosociality. The coefficients $\beta_t^{kk}$ capture own-skill persistence among control boys, while $\beta_t^{kj}$, for $j \neq k$, capture cross-persistence. The corresponding interaction coefficients $\delta_t^{kj}$ measure how treatment changes these relationships. We estimate the specification separately for the intervention period (ages 7–9) and the post-intervention period (ages 10–15).

Figure 3 reports the estimated treatment interactions for own-skill persistence and cross-persistence during and after the intervention. The most notable result concerns the persistence of inattention. During the intervention, the interaction between treatment and baseline inattention is negative but not statistically significant (RI $p = 0.50$). After the intervention, the interaction remains negative and becomes statistically significant (RI $p = 0.05$). This suggests that the intervention weakens the relationship between initial inattention and subsequent inattention, even after the intervention has ended. For the other behavioral dimensions, including cross-persistence between skills, we find no statistically significant treatment interactions. Overall, the results provide some evidence that the intervention changes the persistence of initial inattention, although the estimates for other skill relationships are imprecise.

[17]Using annual teacher ratings, we compute the average of each behavioral skill index separately for the intervention period (ages 7–9) and the post-intervention period (ages 10–15).

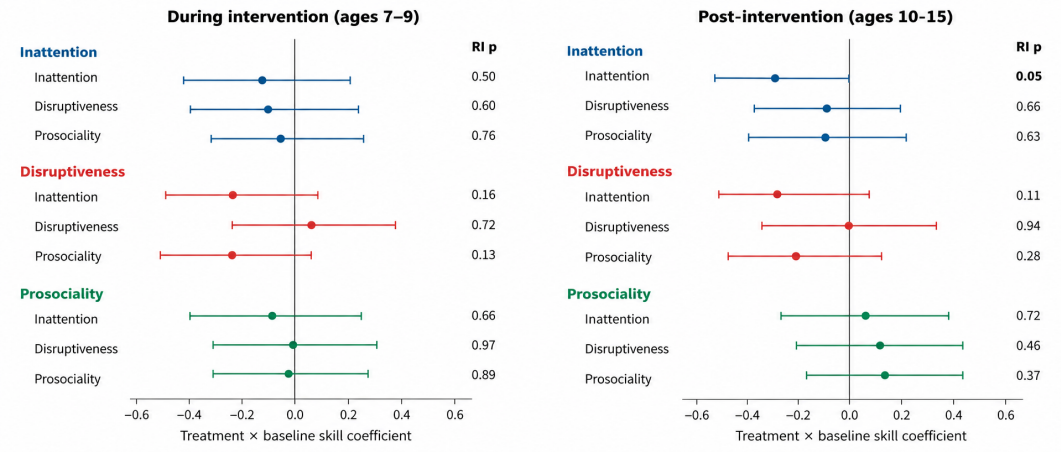



**Figure 3:** Treatment Effects on Skill Persistence and Cross-Persistence

*Notes:* The figure reports treatment-by-baseline-skill interactions from equation (18). The left and right panels correspond to ages 7–9 and 10–15, respectively. Rows identify baseline skills and group labels identify subsequent skills. Diagonal combinations capture persistence; off-diagonal combinations capture cross-persistence. Points indicate estimates, bars indicate 95% confidence intervals, and RI denotes randomization inference.

# 6 Adult Outcome Correlates

The preceding results show that the effect of the intervention on benchmark-adjusted income varies with behavioral characteristics measured at school entry. These estimates do not reveal why this heterogeneity emerges. Rather than attempting to identify causal mechanisms, we revisit other adolescent and adult outcomes previously studied in the MLES. We use these outcomes to provide descriptive context for the benchmark-adjusted income results. We conduct two complementary exercises. First, we examine whether these outcomes are associated with benchmark-adjusted income, conditional on behavioral characteristics at school entry. This corresponds to the $b$-path in a conventional mediation decomposition, but we use it here only as a descriptive association. Second, we examine whether treatment effects on these outcomes vary with the same baseline behavioral characteristics that generate heterogeneity in benchmark-adjusted income.

**Associations with benchmark-adjusted income.** Figure 4 reports the estimated $b$-path for each outcome. Several outcomes are strongly associated with benchmark-adjusted income. Postsecondary education is positively associated with BAI, as is having ever been in a couple relationship between ages 30 and 36. Trust measured at ages 17–19 is also positively

associated with BAI. By contrast, the number of years with zero income between ages 30 and 36 is strongly negatively associated with BAI.

The remaining associations are generally smaller. Grades and concentration at ages 17–19 are positively associated with BAI, although their estimates are less precise. Aggression at age 13 is negatively associated with BAI. We find weaker associations for verbal IQ, delinquency, self-esteem, criminal records, inattention and prosociality at age 13, and altruism.

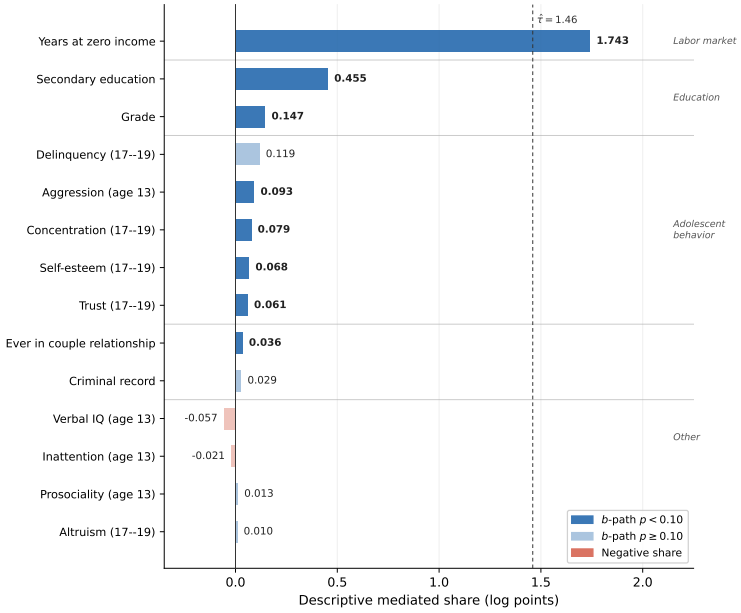



**Figure 4:** Adolescent and Adult Correlates of Benchmark-Adjusted Income

*Notes:* The figure reports estimated associations between each adolescent or adult outcome and benchmark-adjusted income, conditional on behavioral characteristics measured at school entry. Horizontal lines report 95 percent confidence intervals. These estimates are descriptive and should not be interpreted as causal mediation effects.

**Heterogeneity in treatment effects on other outcomes.** We next examine whether treatment effects on these outcomes vary with behavioral characteristics measured at school entry. For each outcome, we estimate the same treatment-by-baseline-skill interactions used in the benchmark-adjusted income analysis. The full set of estimates is reported in Appendix Table B.1. The most notable pattern concerns subsequent labor-market attachment. Among control boys, a one-standard-deviation increase in baseline inattention is associated with 0.546 additional years of zero income between ages 30 and 36. The treatment-by-inattention interaction is −0.813 years, indicating that this gradient is substantially weaker among boys

assigned to treatment. The corresponding randomization-inference $p$-value is 0.019. A similar pattern appears for baseline disruptiveness, for which the treatment interaction is $-0.294$ years.

There is also some evidence of heterogeneity in postsecondary education. The treatment-by-disruptiveness interaction is 0.052, while the treatment-by-inattention interaction is 0.070. The latter estimate is imprecise. Across the remaining outcomes, however, treatment-effect heterogeneity is neither large nor systematic.

These patterns should be interpreted cautiously. None of the individual treatment-by-skill interactions survives adjustment for the full set of multiple comparisons. In particular, the Romano–Wolf adjusted $p$-value for the treatment-by-inattention interaction in years with zero income is 0.206. We therefore do not interpret these estimates as evidence for a particular causal mechanism. Instead, they provide descriptive context for the main experimental result: the effect of the intervention on benchmark-adjusted income is concentrated among boys with higher inattention at school entry, and there is suggestive evidence that this same group experienced stronger labor-market attachment in adulthood.

# 7 Implications for Intergenerational Persistence

The core idea of our framework is that skills measured in kindergarten contribute to the persistence of income across generations for two reasons. First, there is a social gradient in the formation of these skills. Second, these skills play a non-trivial role in adult income.[18] In other words, expected adult income conditional on childhood skills should be highly correlated with parental income, with children expected to earn more coming from higher-income families. In regression terms, we expect the following relationship:

$$E[Y_i \mid S_i] = \alpha + \rho P_i + \varepsilon_i, \tag{19}$$

where $S_i$ denotes skills measured in kindergarten and $P_i$ denotes log parental income. Our framework predicts that $\rho > 0$. Because we show above that the intervention reduces the importance of these skills, or at least inattention, for adult income, it is reasonable to expect a weaker relationship between parental income and expected adult income conditional

[18] This is the logic behind decompositions of intergenerational persistence that account for childhood characteristics, such as Blanden et al. (2007).

on childhood skills for treated boys. That is $\rho_T < \rho_C$, where $\rho_T$ and $\rho_C$ denote the parental-income gradients in expected adult income conditional on childhood skills under treatment and control, respectively.[19]

To explore this pattern, we estimate the relationship between childhood skills and adult outcomes, allowing this relationship to differ by treatment status:

$$Y_i = \alpha + \tau T_i + \boldsymbol{\beta}' S_i + \boldsymbol{\theta}'(T_i S_i) + \varepsilon_i, \tag{20}$$

where $T_i$ indicates treatment assignment. We then use the estimated coefficients to predict each individual's adult outcome under treatment and control, holding childhood skills fixed.

Finally, we regress these predicted outcomes on log parental income separately for each treatment status:

$$\widehat{Y}_i(t) = a_t + \rho_t P_i + u_{it}, \qquad t \in \{0,1\}. \tag{21}$$

We compare the resulting gradients by computing $\Delta\rho = \rho_T - \rho_C$. We report the results of this analysis in Table 6. Since our main results seem to be driven by the extensive margin—the share of years with zero earnings—we report the results for log adult earnings in columns (1)–(2) and the percentage of years with zero earnings in columns (3)–(4).

The estimates in bold represent the differences in the parental-income gradients of predicted adult outcomes between treated and control boys. While we find a strong association between parental income and expected adult income conditional on baseline skills for control boys, this association is close to zero for treated boys. The difference in the gradients is statistically significant at the 10 percent level across the inference methods (maximum $p$-value = 0.07), whether we use only inattention to predict outcomes (Column 1) or all three skills (Column 2). The results on percentage of years with zero income go in the same direction: parental income negatively predicts the expected percentage of years with zero earnings among control boys, and this gradient is close to zero among treated boys.

Combining the evidence presented above with these results suggests that the intervention

[19]For example, consider a boy with a standardized inattention score of 1 at age 6. His expected adult income under treatment would be $\alpha + \tau + \beta + \theta$, while a control boy with the same inattention score would have an expected income of $\alpha + \beta$. Even with balance in baseline inattention between treatment and control groups, these expected outcomes need not be equal. Randomization ensures that baseline characteristics are balanced on average, but the intervention can change how these characteristics relate to adult income.

reduces the contribution of kindergarten skills to the persistence of income across generations.

**Table 6:** Intervention-Induced Changes in the Skill-Mediated Intergenerational Gradient

| | Log adult earnings | | % zero-earnings years | |
|---|---|---|---|---|
| | Inattention only | All three skills | Inattention only | All three skills |
| | (1) | (2) | (3) | (4) |
| Projected gradient: Control | 0.501 | 0.567 | −4.213 | −4.617 |
| | (0.207) | (0.244) | (2.034) | (2.258) |
| Projected gradient: Treatment | −0.031 | −0.035 | 0.498 | 0.626 |
| | (0.146) | (0.193) | (1.381) | (1.867) |
| **Change in projected gradient** | **−0.530** | **−0.603** | **4.711** | **5.237** |
| | (0.251) | (0.313) | (2.481) | (3.025) |
| Randomization $p$-value | 0.020 | 0.030 | 0.020 | 0.050 |
| Bootstrap percentile $p$-value | 0.030 | 0.060 | 0.060 | 0.090 |
| Bootstrap studentized $p$-value | 0.070 | 0.070 | 0.080 | 0.090 |
| Worst-case $p$-value | 0.070 | 0.070 | 0.080 | 0.090 |
| Treatment × Inattention | 1.112 | 1.153 | −9.849 | −10.466 |
| | (0.425) | (0.406) | (4.276) | (4.143) |
| Worst-case $p$-value | 0.020 | 0.010 | 0.030 | 0.010 |

*Notes:*The table reports model-implied changes in the parental-income gradient of adult outcomes operating through baseline behavioral skills. The outcome is log adult earnings in columns (1)–(2) and the percentage of years with zero earnings in columns (3)–(4). Columns (1) and (3) include inattention and its interaction with treatment; columns (2) and (4) additionally include disruptiveness, prosociality, and their interactions with treatment. Adult outcomes are predicted under treatment and control from equation (20), which does not condition on parental income, and the predictions are then regressed on log parental income. Because the outcome is raw log earnings and the specification excludes parental income, the treatment-by-inattention interaction differs slightly from the estimates in Table 4. The change in the projected gradient is the difference between treatment and control gradients. Bootstrap standard errors are reported in parentheses. The worst-case $p$-value is the maximum of the three reported inference-based $p$-values. These projections do not identify changes in the overall intergenerational income elasticity.

# 8 Conclusion

Using rare data that link children's behavior at school entry to their own and their parents' tax records, we show that prosociality, disruptiveness, and inattention at age six account for a nontrivial share of the persistence of income across generations, about 13% of the intergenerational income correlation in a population-representative Quebec cohort. We further show that an intervention targeting some of these traits raised adult income and weakened the link between early inattention and income in adulthood.

The intervention randomly assigned disruptive kindergarten boys from low-income Montreal neighborhoods to two years of social skills and self-control training. Relative to what their family background and behavior at age six would predict, treated boys earn 1.46 log points more than control boys between ages 30 and 36, extending the average effects documented by Algan et al. (2022). Among control boys, a one-standard-deviation increase in inattention at age six is associated with 0.71 log points lower adult income, whereas among treated boys this relationship is largely offset and imprecisely estimated. Both the income gain and this pattern vanish when we restrict the sample to boys with positive earnings in every year between ages 30 and 36, which suggests that the intervention worked through labor market attachment, by keeping the most inattentive boys from spending years without any earnings.

The boys who were most inattentive at school entry thus gained the most, which runs against the prediction, often drawn from models of dynamic complementarities, that children with stronger initial skills benefit more from early investments, and this is all the more striking because the program selected boys for disruptive behavior rather than inattention. Because inattention is socially graded, the intervention also weakened the link between parental income and the adult income implied by early skills, a gradient that is strong among control boys and close to zero among treated boys. Teacher assessments from ages six to thirteen show little lasting divergence in behavior between the two groups, so the program seems to have changed the economic consequences of early inattention more than the behavior itself.

These results come with two limitations that call for caution. The experimental sample is small, so the heterogeneity estimates are imprecise: the interaction with inattention survives correction for multiple testing, but the joint test across the three skills is significant only at the 10 percent level, and the confidence interval for the total implied change in intergener-

ational persistence includes zero. The program was also limited to white francophone boys from disadvantaged Montreal schools in 1984, and whether similar effects would arise for girls, in other populations, or in today's schools remains an open question.

Even with these caveats, our findings show that early disadvantage can strongly predict adult outcomes without being fixed. Programs that reach the children who start furthest behind can raise their adult income and, at the same time, weaken one of the channels through which economic position is passed from parents to children, a part of their value that average treatment effects alone do not capture.

# References


Achenbach, Thomas M. (1991) *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*, Burlington, VT: Department of Psychiatry, University of Vermont.

Algan, Yann, Elizabeth Beasley, Sylvana Côté, Jungwee Park, Richard E. Tremblay, and Frank Vitaro (2022) "The Impact of Childhood Social Skills and Self-Control Training on Economic and Noneconomic Outcomes: Evidence from a Randomized Experiment Using Administrative Data," *American Economic Review*, 112 (8), 2553–2579, 10.1257/aer.20200224.

Almond, Douglas and Janet Currie (2011) "Killing Me Softly: The Fetal Origins Hypothesis," *Journal of Economic Perspectives*, 25 (3), 153–172.

Attanasio, Orazio, Sarah Cattan, Emla Fitzsimons, Costas Meghir, and Marta Rubio-Codina (2020) "Estimating the production function for human capital: results from a randomized controlled trial in Colombia," *American Economic Review*, 110 (1), 48–85.

Björklund, Anders, Markus Jäntti, and Martin Nybom (2017) "The contribution of early-life versus labour market factors to intergenerational income persistence: a comparison of the UK and Sweden."

Blanden, Jo, Paul Gregg, and Lindsey Macmillan (2007) "Accounting for intergenerational income persistence: noncognitive skills, ability and education," *The Economic Journal*, 117 (519), C43–C60.

Blanden, Jo, Robert Haveman, Timothy Smeeding, and Kathryn Wilson (2014) "Intergenerational Mobility in the U nited S tates and G reat B ritain: A Comparative Study of Parent–Child Pathways," *Review of Income and Wealth*, 60 (3), 425–449.

Bloome, Deirdre (2017) "Childhood family structure and intergenerational income mobility in the United States," *Demography*, 54 (2), 541–569.

Boisjoli, Rachel, Frank Vitaro, Éric Lacourse, Edward D. Barker, and Richard E. Tremblay (2007) "Impact and Clinical Significance of a Preventive Intervention for Disruptive Boys: 15-Year Follow-Up," *British Journal of Psychiatry*, 191 (5), 415–419, 10.1192/bjp.bp.106.030007.

Campbell, Frances, Gabriella Conti, James J Heckman, Seong Hyeok Moon, Rodrigo Pinto, Elizabeth Pungello, and Yi Pan (2014) "Early childhood investments substantially boost adult health," *Science*, 343 (6178), 1478–1485.

Carneiro, Pedro and James J Heckman (2003) "Human Capital Policy," in Heckman, James J and Alan B Krueger eds. *Inequality in America: What Role for Human Capital Policies?*, 77–239, Cambridge, MA: MIT Press.

Chen, Wen-Hao, Yuri Ostrovsky, and Patrizio Piraino (2017) "Lifecycle variation, errors-in-variables bias and nonlinearities in intergenerational income transmission: new evidence from Canada," *Labour Economics*, 44, 1–12.

Chernozhukov, Victor and Christian Hansen (2005) "An IV model of quantile treatment effects," *Econometrica*, 73 (1), 245–261.

Chetty, Raj and Nathaniel Hendren (2018) "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects," *Quarterly Journal of Economics*, 133 (3), 1107–1162, 10.1093/qje/qjy007.

Chetty, Raj, Nathaniel Hendren, and Lawrence F Katz (2016) "The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment," *American Economic Review*, 106 (4), 855–902.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez (2014) "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States," *Quarterly Journal of Economics*, 129 (4), 1553–1623, 10.1093/qje/qju022.

Chyn, Eric and Lawrence F Katz (2021) "Neighborhoods matter: Assessing the evidence for place effects," *Journal of Economic Perspectives*, 35 (4), 197–222.

Connolly, Marie, Miles Corak, and Catherine Haeck (2019) "Intergenerational Mobility between and within Canada and the United States," *Journal of Labor Economics*, 37 (S2), S595–S641.

Connolly, Marie and Catherine Haeck (2024) "Intergenerational income mobility trends in Canada," *Canadian Journal of Economics/Revue canadienne d'économique*, 57 (1), 5–26.

——— (2026) "The Great Mobility Puzzle Resolved," *Canadian Public Policy*, 52 (1), 1–21.

Cunha, Flavio and James Heckman (2007) "The technology of skill formation," *American economic review*, 97 (2), 31–47.

Cunha, Flavio and James J Heckman (2008) "Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation," *Journal of human resources*, 43 (4), 738–782.

Currie, Janet (2011) "Inequality at Birth: Some Causes and Consequences," *American Economic Review: Papers & Proceedings*, 101 (3), 1–22.

De Haan, Laurens (1981) "Estimation of the minimum of a function using order statistics," *Journal of the American Statistical Association*, 76 (374), 467–469.

Deming, David (2009) "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start," *American Economic Journal: Applied Economics*, 1 (3), 111–134, 10.1257/app.1.3.111.

Deutscher, Nathan (2020) "Place, peers, and the teenage years: long-run neighborhood effects in Australia," *American Economic Journal: Applied Economics*, 12 (2), 220–249.

Findlay, Leanne C, Elizabeth Beasley, Jungwee Park, Dafna E Kohen, Yann Algan, Frank Vitaro, and Richard E Tremblay (2018) "Longitudinal child data: what can be gained by linking administrative data and cohort data?" *International Journal of Population Data Science*, 3 (1), 451.

Freedman, David and David Lane (1983) "A Nonstochastic Interpretation of Reported Significance Levels," *Journal of Business & Economic Statistics*, 1 (4), 292–298.

Garces, Eliana, Duncan Thomas, and Janet Currie (2002) "Longer-term effects of Head Start," *American economic review*, 92 (4), 999–1012.

García, Jorge Luis, James J. Heckman, Duncan Leaf, and María José Prados (2021) "Quantifying the Life-Cycle Benefits of a Prototypical Early Childhood Program," *Journal of Political Economy*, 129 (11), 2922–3015, 10.1086/714426.

García, Jorge Luis, James J Heckman, and Victor Ronda (2023) "The lasting effects of early-childhood education on promoting the skills and social mobility of disadvantaged African Americans and their children," *Journal of Political Economy*, 131 (6), 1477–1506.

Heckman, James J (2006) "Skill Formation and the Economics of Investing in Disadvantaged Children," *Science*, 312 (5782), 1900–1902.

Heckman, James, Rodrigo Pinto, and Peter Savelyev (2013) "Understanding the mechanisms through which an influential early childhood program boosted adult outcomes," *American economic review*, 103 (6), 2052–2086.

Lee, David S. (2009) "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *The Review of Economic Studies*, 76 (3), 1071–1102, 10.1111/j.1467-937X.2009.00536.x.

Lorge, Irving and Robert Ladd Thorndike (1954) *The Lorge-Thorndike Intelligence Tests...*: Houghton, Mifflin.

Moffitt, Terrie E (1993) "Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy.," *Psychological review*, 100 (4), 674.

Pingault, J–B, Sylvana M Côté, Frank Vitaro, Benoît Falissard, Clement Genolini, and Richard E Tremblay (2014) "The developmental course of childhood inattention symptoms uniquely predicts educational attainment: a 16-year longitudinal study," *Psychiatry Research*, 219 (3), 707–709, 10.1016/j.psychres.2014.06.022.

Romano, Joseph P and Michael Wolf (2005) "Exact and approximate stepdown methods for multiple hypothesis testing," *Journal of the American Statistical Association*, 100 (469), 94–108.

Rouquette, Alexandra, Sylvana M Côté, Laura E Pryor, René Carbonneau, Frank Vitaro, and Richard E Tremblay (2014) "Cohort profile: The Quebec longitudinal study of kindergarten children (QLSKC)," *International journal of epidemiology*, 43 (1), 23–33.

Solon, Gary (1992) "Intergenerational Income Mobility in the United States," *American Economic Review*, 82 (3), 393–408.

Tremblay, Richard E, Lyse Desmarais-Gervais, Claude Gagnon, and Pierre Charlebois (1987) "The Preschool Behaviour Questionnaire: Stability of its factor structure between cultures, sexes, ages and socioeconomic classes," *International Journal of Behavioral Development*, 10 (4), 467–484.

Tremblay, Richard E, Joan McCord, Helene Boileau, Pierre Charlebois, Claude Gagnon, Marc Le Blanc, and Serge Larivée (1991) "Can disruptive boys be helped to become competent?" *Psychiatry*, 54 (2), 148–161.

Tremblay, Richard E., Frank Vitaro, Lucie Bertrand, Marc Leblanc, Hélène Beauchesne, Hélène Boileau, and Lucille David (1992) "Parent and Child Training to Prevent Early Onset of Delinquency: The Montreal Longitudinal-Experimental Study," in McCord, Joan and Richard E. Tremblay eds. *Preventing Antisocial Behavior: Interventions from Birth Through Adolescence*, 117–138, New York: Guilford Press.

Young, Alwyn (2019) "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results," *Quarterly Journal of Economics*, 134 (2), 557–598.

# A

**Table A.1:** Descriptive Statistics: QLSKC Boys and MLES Samples

| | QLSKC Boys | | MLES Control | | MLES Treatment | |
|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD |
| *Panel A: Family background* | | | | | | |
| Year of birth | 1980.42 | 0.70 | 1977.70 | 0.46 | 1977.73 | 0.45 |
| Parental income (adj., 000 CAD) | 37.49 | 25.53 | 28.40 | 18.50 | 27.30 | 11.30 |
| Log parental income | 10.37 | 0.62 | 10.08 | 0.62 | 10.13 | 0.42 |
| Parental education (years) | 12.65 | 3.15 | 10.76 | 5.70 | 10.59 | 4.80 |
| Parental age at birth | 26.39 | 4.46 | 21.75 | 10.60 | 23.44 | 9.20 |
| *Panel B: Age-6 behavioral skills (SBQ, raw scores)* | | | | | | |
| Inattention | 2.12 | 2.30 | 3.34 | 2.47 | 3.59 | 2.63 |
| Hyperactivity | 1.13 | 1.38 | 2.30 | 1.47 | 2.40 | 1.44 |
| Opposition | 2.00 | 2.44 | 4.14 | 2.88 | 4.47 | 2.86 |
| Prosociality | 7.04 | 4.53 | 7.22 | 4.54 | 7.15 | 4.62 |
| Aggression | 1.21 | 2.61 | 9.93 | 7.72 | 9.00 | 7.50 |
| *Panel C: Cognitive ability and income observation* | | | | | | |
| Verbal IQ (age 13) | 9.46 | 2.22 | 7.46 | 3.66 | 7.54 | 3.95 |
| N valid tax years (30–36) | 6.78 | 0.72 | 6.31 | 1.23 | 6.34 | 1.29 |
| *Panel D: Sample size* | | | | | | |
| *N* | 1,400 | | 144 | | 57 | |

*Notes:* MLES sample sizes are rounded to the nearest ten per Statistics Canada disclosure requirements. QLSKC statistics are for the boys-only analytical sample used to estimate the population mobility technology. Parental income is the average adjusted family income when the child was aged 15–19, expressed in 2014 constant Canadian dollars. MLES boys are drawn from low-income Montreal neighborhoods and were selected on scoring above the 70th percentile of antisocial behavior at age six, which explains the substantially higher behavioral difficulty scores and lower family income relative to the QLSKC population sample. Behavioral skills are reported as raw SBQ scores; in all regression analyses they are standardized using QLSKC boys sample means and standard deviations. The large difference in aggression scores reflects the selection criterion: MLES boys were chosen for their high disruptiveness at school entry.

# B

**Table B.1:** Adult Outcome Correlates: Full Regression Results

| | | Control group | | | Treatment difference ($\hat{\theta}_{km}$) | | | $p$-values on $\hat{\theta}_{km}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Outcome | Skill | $\hat{\phi}^C_{km}$ | SE | $p$ | $\hat{\theta}_{km}$ | SE | $p_{\text{OLS}}$ | $p_{\text{RI}}$ | $p_{\text{RW}}$ | Sig.† |
| *Panel A: Labor market attachment* | | | | | | | | | | |
| Years at zero income | Inattention | 0.546 | 0.209 | 0.010 | −0.813 | 0.249 | 0.001 | 0.019 | 0.206 | ✓ |
| | Disruptiveness | 0.163 | 0.084 | 0.054 | −0.294 | 0.092 | 0.002 | 0.022 | 0.239 | ✓ |
| | Prosociality | −0.142 | 0.181 | 0.434 | −0.111 | 0.235 | 0.638 | 0.695 | 0.995 | |
| *Panel B: Educational attainment* | | | | | | | | | | |
| Secondary education | Inattention | −0.069 | 0.033 | 0.037 | 0.070 | 0.058 | 0.234 | 0.226 | 0.887 | |
| | Disruptiveness | 0.008 | 0.015 | 0.586 | 0.052 | 0.023 | 0.022 | 0.020 | 0.210 | ✓ |
| | Prosociality | −0.036 | 0.037 | 0.327 | 0.117 | 0.070 | 0.096 | 0.081 | 0.592 | ✓ |
| Grade | Inattention | −0.210 | 0.073 | 0.005 | 0.115 | 0.135 | 0.394 | 0.374 | 0.975 | |
| | Disruptiveness | 0.030 | 0.031 | 0.343 | 0.015 | 0.047 | 0.746 | 0.749 | 0.986 | |
| | Prosociality | 0.037 | 0.084 | 0.666 | 0.148 | 0.128 | 0.247 | 0.297 | 0.957 | |
| *Panel C: Adolescent behavior (ages 17–19)* | | | | | | | | | | |
| Delinquency | Inattention | 0.032 | 0.051 | 0.537 | 0.057 | 0.072 | 0.430 | 0.451 | 0.989 | |
| | Disruptiveness | −0.034 | 0.024 | 0.154 | 0.066 | 0.027 | 0.014 | 0.051 | 0.396 | ✓ |
| | Prosociality | −0.006 | 0.038 | 0.867 | 0.079 | 0.071 | 0.265 | 0.242 | 0.920 | |
| Concentration | Inattention | −0.010 | 0.051 | 0.852 | 0.058 | 0.082 | 0.477 | 0.528 | 0.989 | |
| | Disruptiveness | −0.006 | 0.024 | 0.815 | 0.023 | 0.035 | 0.512 | 0.542 | 0.986 | |
| | Prosociality | 0.065 | 0.048 | 0.176 | 0.041 | 0.090 | 0.650 | 0.659 | 0.995 | |
| Self-esteem | Inattention | −0.032 | 0.038 | 0.397 | −0.012 | 0.068 | 0.859 | 0.863 | 0.989 | |
| | Disruptiveness | −0.020 | 0.015 | 0.199 | 0.026 | 0.023 | 0.267 | 0.282 | 0.943 | |
| | Prosociality | 0.030 | 0.041 | 0.461 | 0.061 | 0.071 | 0.387 | 0.448 | 0.985 | |
| Trust | Inattention | −0.015 | 0.036 | 0.681 | −0.039 | 0.064 | 0.538 | 0.545 | 0.989 | |
| | Disruptiveness | 0.014 | 0.016 | 0.379 | −0.020 | 0.026 | 0.437 | 0.418 | 0.972 | |
| | Prosociality | −0.021 | 0.033 | 0.536 | 0.097 | 0.068 | 0.160 | 0.136 | 0.773 | |
| Altruism | Inattention | 0.017 | 0.066 | 0.793 | −0.055 | 0.107 | 0.606 | 0.620 | 0.989 | |
| | Disruptiveness | −0.023 | 0.026 | 0.382 | 0.022 | 0.043 | 0.607 | 0.572 | 0.986 | |
| | Prosociality | 0.059 | 0.062 | 0.348 | −0.006 | 0.153 | 0.967 | 0.966 | 0.995 | |
| *Panel D: Adolescent skills (age 13)* | | | | | | | | | | |
| Inattention (age 13) | Inattention | 0.126 | 0.189 | 0.506 | −0.261 | 0.291 | 0.372 | 0.422 | 0.985 | |
| | Disruptiveness | −0.063 | 0.080 | 0.429 | 0.004 | 0.137 | 0.975 | 0.967 | 0.995 | |
| | Prosociality | −0.452 | 0.166 | 0.007 | 0.401 | 0.360 | 0.268 | 0.236 | 0.911 | |
| Aggression (age 13) | Inattention | −0.069 | 0.442 | 0.877 | −1.265 | 0.718 | 0.080 | 0.110 | 0.691 | |
| | Disruptiveness | −0.233 | 0.217 | 0.284 | 0.024 | 0.307 | 0.937 | 0.935 | 0.995 | |
| | Prosociality | −0.531 | 0.365 | 0.148 | 0.265 | 1.038 | 0.799 | 0.734 | 0.995 | |
| Verbal IQ (age 13) | Inattention | −0.602 | 0.256 | 0.019 | −0.294 | 0.490 | 0.549 | 0.549 | 0.989 | |
| | Disruptiveness | −0.412 | 0.121 | 0.001 | −0.103 | 0.226 | 0.649 | 0.606 | 0.986 | |
| | Prosociality | 0.299 | 0.232 | 0.198 | −0.207 | 0.524 | 0.693 | 0.650 | 0.995 | |
| *Panel E: Social outcomes* | | | | | | | | | | |
| Criminal record | Inattention | 0.017 | 0.033 | 0.606 | −0.076 | 0.054 | 0.160 | 0.184 | 0.872 | |
| | Disruptiveness | −0.005 | 0.014 | 0.718 | −0.014 | 0.019 | 0.468 | 0.517 | 0.986 | |
| | Prosociality | 0.012 | 0.031 | 0.704 | −0.042 | 0.056 | 0.458 | 0.477 | 0.985 | |
| Couple relationship | Inattention | −0.015 | 0.036 | 0.682 | −0.042 | 0.065 | 0.515 | 0.503 | 0.989 | |
| | Disruptiveness | −0.001 | 0.017 | 0.955 | 0.021 | 0.022 | 0.354 | 0.380 | 0.972 | |
| | Prosociality | 0.008 | 0.036 | 0.828 | 0.039 | 0.057 | 0.496 | 0.537 | 0.986 | |

# C

**Table C.1:** Early Behavioral Skills and Intergenerational Income Persistence: Full QLSKC Sample (Appendix)

| | SES gradient $\hat{\beta}_k$ | Income gradient $\hat{\phi}_k$ | Contribution $\hat{\beta}_k\hat{\phi}_k$ | 95% CI | Share of IGE (%) |
|---|---|---|---|---|---|
| *QLSKC full sample (boys and girls)* | | | | | |
| Inattention | $-0.125^{***}$ | $-0.145^{***}$ | 0.018 | [0.002, 0.034] | 3.06 |
| | (0.030) | (0.050) | | | |
| Disruptiveness | $-0.078^{***}$ | $-0.284^{***}$ | 0.022 | [0.003, 0.041] | 3.74 |
| | (0.025) | (0.059) | | | |
| Prosociality | $0.127^{***}$ | $0.187^{***}$ | 0.024 | [0.008, 0.040] | 4.01 |
| | (0.030) | (0.045) | | | |
| All behavioral skills | | | 0.064 | [0.021, 0.107] | **10.81** |
| Direct (parental income) | | | 0.532 (0.073) | | 89.86 |
| **Total IGE** | | | **0.593** (0.073) | | **100.0** |

*Notes:* Same methodology as Table 2 estimated on the full QLSKC sample of boys and girls, with a control for gender. 95% confidence intervals on contributions computed by the delta method. Standard errors in parentheses. $^{***}p < 0.01$.

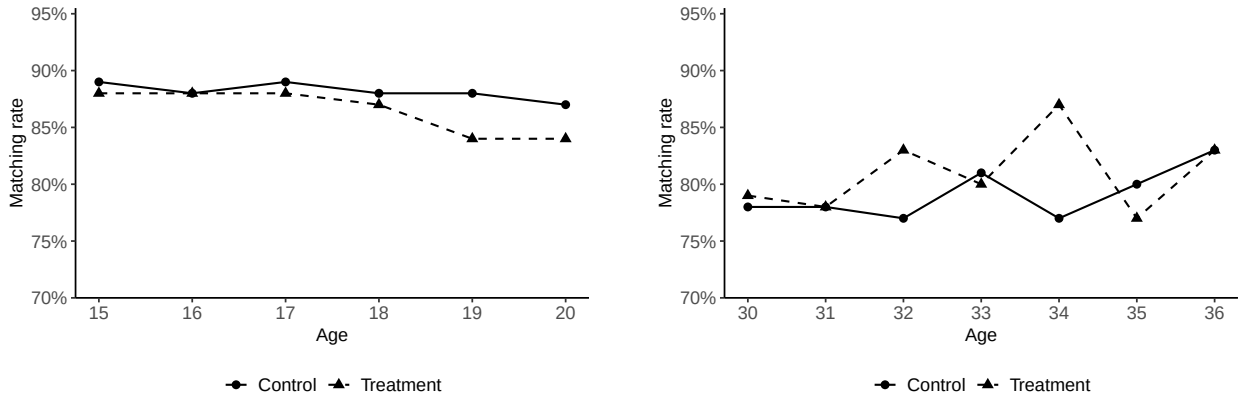
**Table D.1:** Linear Probability Models: Childhood Skills and Intergenerational Mobility

| | Upward Mobility (1) | Downward Mobility (2) |
|---|---|---|
| Inattention Q2 | −0.033(0.028) | 0.023(0.029) |
| Inattention Q3 | −0.0209(0.032) | 0.036(0.033) |
| Inattention Q4 | −0.158(0.035)*** | 0.1152(0.036)*** |
| Prosociality Q2 | −.0102(0.0274) | −0.026(0.028) |
| Prosociality Q3 | 0.0025(0.028) | −0.065(0.029)*** |
| Prosociality Q4 | 0.0115(0.0302) | −0.066(0.0307)*** |
| Opposition Q2 | −0.0403(0.0288) | 0.022(0.029) |
| Opposition Q3 | −0.024(0.0328) | −0.023(0.033) |
| Opposition Q4 | −0175(0.0392) | −0.0151(0.040) |
| Hyperactivity Q2 | 0.0162(0.028) | −0.022(0.029) |
| Hyperactivity Q3 | −0.002(0.033) | −0.005(0.033) |
| Hyperactivity Q4 | 0.0241(0.039) | 0.0005(0.0394) |
| Controls | Yes | Yes |
| Observations | 2354 | 2354 |
| $R^2$ | 0.0657 | 0.0652 |

*Notes*: Each column reports coefficients from an OLS linear probability model. The dependent variable equals 1 for upward (downward) mobility. The omitted category for each skill is the first quartile (lowest difficulty / highest prosociality). Controls include gender, parental education, parental age at birth, parental occupational prestige, birth month, and birth year. Robust standard errors in parentheses.

# D Correlation between skills in childhood and intergenerational income mobility

# E Income data matching rate by treatment status




**(a)** Parental income matching rate by child age

**(b)** Adult income matching rate by age

**Figure E.1:** Tax Data Matching Rates by Treatment Status

*Notes:* Panel A reports the share of randomized MLES participants matched to parental income tax records at each child age from 15 to 20. Panel B reports the share of randomized MLES participants matched to their own income tax records at each age from 30 to 36. Matching rates are calculated relative to the original randomized samples of 181 control and 69 treatment participants.

# F Additional Figures and Tables

## F.1 Changes in Income Rank Between Generations

Figure F.1 plots the distribution of changes in income rank between generations, measured as the difference between adult and parental income percentile ranks:

$$\Delta R_i = R_i^{adult} - R_i^{parent}. \tag{22}$$

Positive values indicate upward movements in the income distribution, while negative values indicate downward movements.

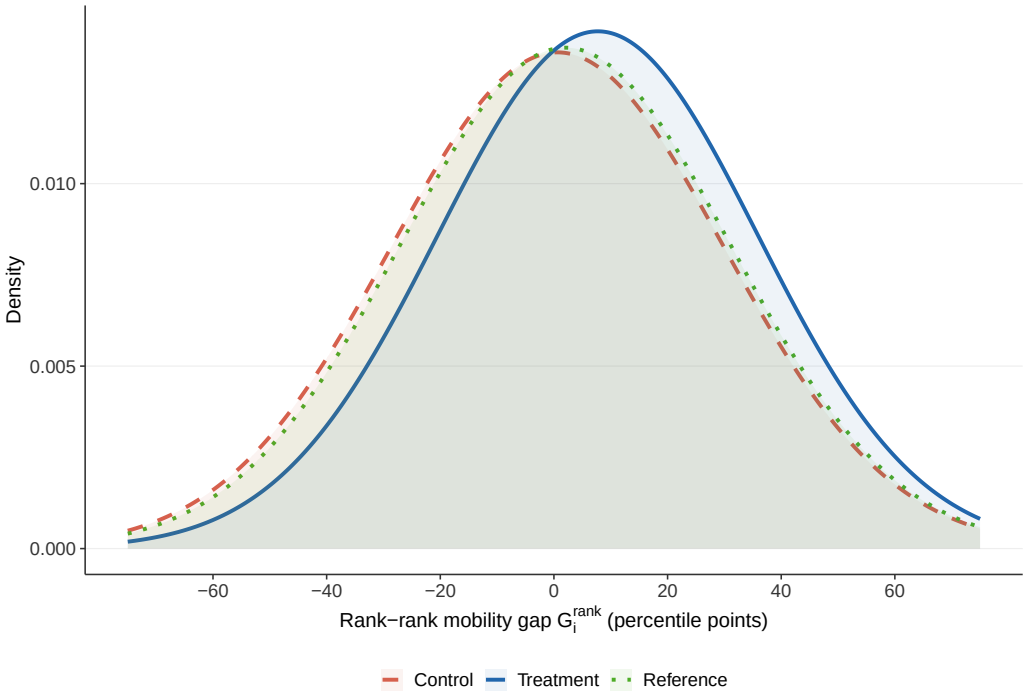



**Figure F.1:** Distribution of Changes in Income Rank Between Generations

*Notes:* The figure plots the distribution of the difference between adult and parental income percentile ranks separately for treatment and control boys. The distributions appear to differ mainly in the left tail, with little visible difference in the right tail. These comparisons are descriptive and should not be interpreted as quantile treatment effects.

## F.2 Log–Log (IGE) Specification

We re-estimate the benchmark mobility curve on the reference group using the log–log specification:

$$\log(y^{\text{child}}) = \alpha + \beta \log(y^{\text{parent}}).$$

Residuals are defined as $\Delta = \log(y^{\text{child}}) - \widehat{\log(y^{\text{child}})}$, with leave-one-bin-out predictions for the reference group.

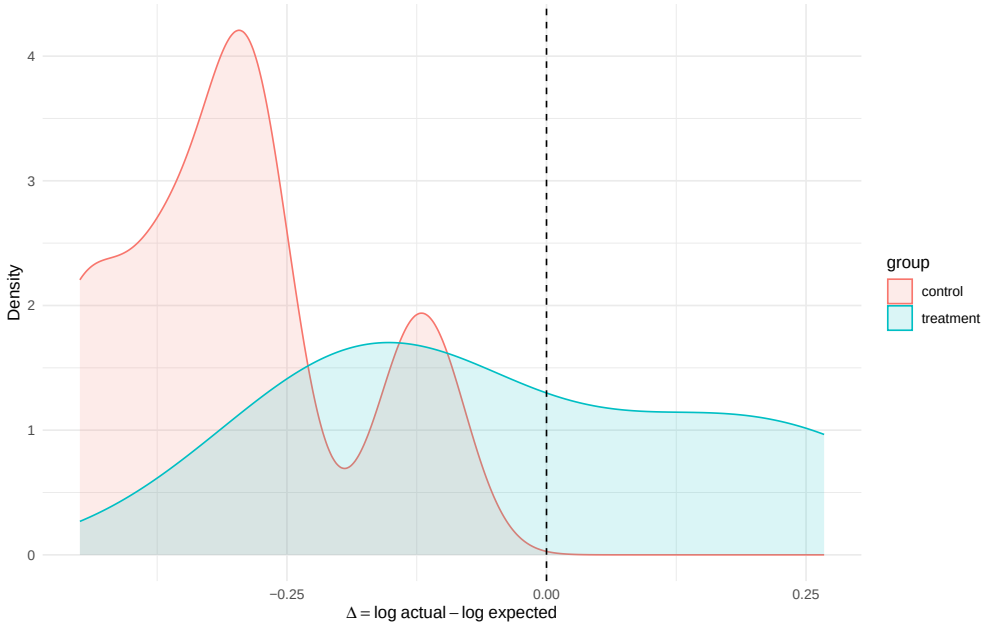



**Figure F.2:** Residual distributions under the log–log specification.

*Notes:* Residuals are in log points. A value of 0.10 corresponds to approximately 10.5% above the benchmark. Results mirror the baseline: treated boys outperform controls relative to the reference mobility curve.

## F.3 Midpoint Definition of Parental Rank Bins

Instead of using the upper bound of each parental income bin as the regressor, we assign the bin midpoint as $y^{\text{parent}}$. We re-fit the benchmark on the reference group and recompute residuals.

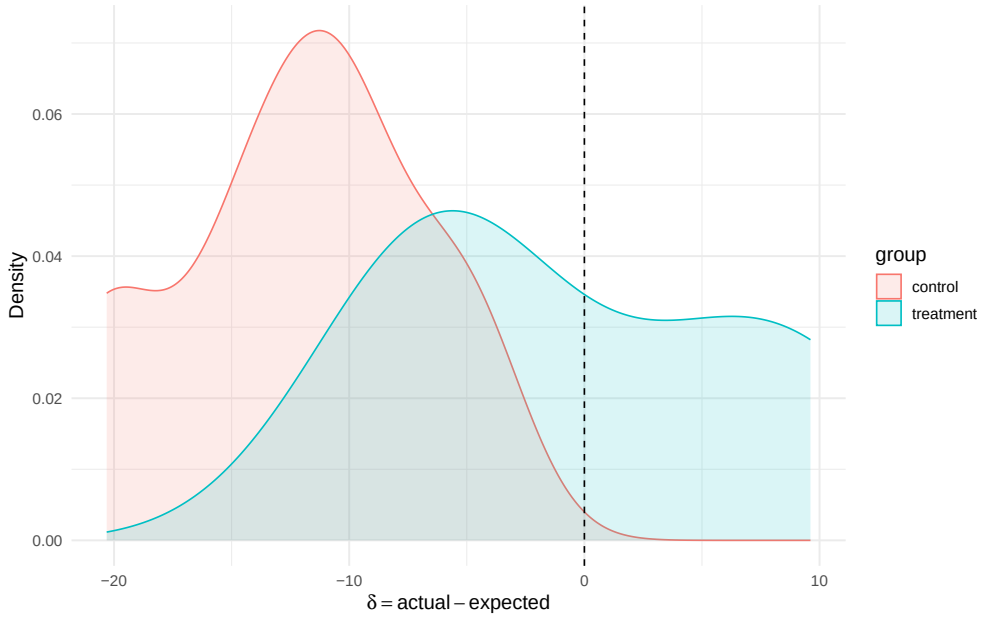



**Figure F.3:** Residual distributions using parental-rank midpoints.

*Notes:* Same outcome definitions and sample as the baseline. Using midpoints yields virtually identical conclusions to the baseline upper-bound specification.

## F.4 Sensitivity to outliers

**Table F.1:** Sensitivity to Individual Observations: Leave-One-Out Estimates

| Outcome | Model | Min. estimate | Max. estimate | Min. $p$-value | Max. $p$-value |
|---|---|---|---|---|---|
| Benchmark-adjusted income | ATE | 1.336 | 1.596 | $< 0.001$ | 0.001 |
| | HTE | 0.935 | 1.242 | 0.002 | 0.010 |
| Log adult income | ATE | 1.275 | 1.502 | $< 0.001$ | 0.002 |
| | HTE | 0.904 | 1.206 | 0.002 | 0.010 |

*Notes:* Each specification is re-estimated after excluding one observation at a time. The table reports the minimum and maximum coefficient estimates and $p$-values across all leave-one-out regressions. ATE refers to the treatment coefficient in the average treatment effect specification. HTE refers to the interaction between treatment and baseline inattention in the heterogeneous treatment effect specification. Reported $p$-values are based on heteroskedasticity-robust standard errors.