

The Long-Term Impact of Social Skills Training at School Entry:

A randomized controlled trial¹

Yann Algan (Sciences Po), Elizabeth Beasley (Sciences Po),

Frank Vitaro (University of Montréal), Richard E. Tremblay (UCD, University of Montréal)

February 2014

Abstract

Childhood behavioral skills are strong predictors of adult socio-economic success, but little is known about how to improve these skills, in particular among children the most at risk of poor adult outcomes. We use data from a long-term randomized evaluation of a childhood behavioral skills training program in Montreal to answer this question. We match detailed data on behavior from adolescence with administrative criminal and educational records and self-reported socio-economic outcomes. As adults, the subjects in the treated group are about 30% less likely to have a criminal record, 50% more likely to have a secondary diploma, are 16% more likely to be active fulltime in either work or school during ages 17-26, and 68% more likely to have ever belonged to a civic or social group. We distinguish the different potential channels through which this intervention operates, and present evidence that self-control and social trust are potentially important channels. One possibility is that these behavioral changes in early adolescence (ages 10-13) lead to improvements in school outcomes, particularly class assignment, in later adolescence (ages 14-17), which in turn lead to improved adult outcomes. Using conservative assumptions in a simple framework, we estimate that \$1 invested in this program yields \$4.5 in increased wages.

¹ We thank Sam Bowles, Ernst Fehr, Nicole Fortin, Larry Katz, Jens Ludwig, Philip Oreopoulos and Emmanuel Saez for very helpful comments. All the data from the MLES experiment are located at GRIP: Groupe de recherche sur l'inadaptation psychosociale chez l'enfant, Hopital Sainte-Justine, Montréal Canada. The research leading to these results has received funding from the European Research Council for Yann Algan under the European Community's Seventh Framework Program (FP7/2007-2013) / ERC grant agreement n° 240923. Elizabeth Beasley gratefully acknowledges the support of the Abdul Latif Jameel Poverty Action Lab (J-PAL).

1. Introduction

Can public policy improve outcomes for adults through interventions targeted at social skill development in childhood? This question is important for education policy in general, but it is critical for children with low levels of social skills who are likely to fail in school, to have important social problems and be locked into poverty during adulthood. Substantial evidence shows that non-cognitive skills, such as impulse control, social-information processing or social cooperation before school entry are strongly correlated with favorable school, economic and social outcomes (Tremblay et al., 1994, Vitaro et al. 2005, Moffit et al, 2011), and that formal education is important not only for training in cognitive skills but also for the development of non-cognitive skills (Jencks, 1979, Heckman and Rubinstein, 2001, Bowles et al. 2001, Segal, 2013, Dewey 1944, Algan, Cahuc and Shleifer 201; Tremblay, 2010). Recent assessments of preschool experiments aimed at improving social and cognitive skills, such as the Abecedarian project, the Perry Preschool program and Head Start, as well as evidence from Project STAR on teacher quality², also suggest that their long-term impact flows through some sort of non-academic or non-cognitive skills channels (see Almond and Currie 2010 for a review). This raises the question of what those non-cognitive skills are and how much of the impact comes through them. It also raises the question whether and how social skills training in early elementary school could have long term impacts for those children who poor social skills when they enter elementary school.

This paper contributes to this important literature by estimating the long-term impact of the Montreal Longitudinal Experimental Study (MLES), a randomized experiment of behavioral skills training to disruptive kindergarten boys. In the spring of 1984, the MLES evaluated 1037 boys at the end of kindergarten in schools in low socio-economic areas. From this original sample, 250 boys were targeted for the experiment based on teacher ratings of disruptive behavior. These 250 boys were randomly assigned to either participate in the social skills/parent training program or to be part of the control group.

The program was implemented over a two-year period, when the subjects were between seven and nine years old. The training program was intensive and involved different points of contact: the child himself, his parents, his peers and his teachers. The behavioral skills training program was inspired by the best practices known at the time the program was

² Project STAR was intended to measure the impact of reduced class size, but since the evaluation design included randomization into different classes, children were effectively randomized to different quality teachers. Using this information, Chetty et al (2011) find large long-term effects from teacher quality in kindergarten, and attribute some of this

planned. Over the two-year period 19 sessions were carried out by a team of professional childcare workers, a social worker and a psychologist. The sessions included disruptive boys and boys with good social skills to facilitate learning of the social skills by the disruptive boys. Sessions included social behavior/skills training (such as how to ask why, how to invite a bystander to play, or how to help) and self-control strategy lessons (like how to react to teasing or how to react when angry, how to correctly identify the intentions of others). The parents received a home-based training program in child rearing to promote the social skills of their children. The parent training program included reading material on parenting, teaching parents a) to monitor their children's behavior, b) to give their children positive reinforcement for prosocial behavior, c) to discipline effectively without using abusive punishment, d) to transfer the skills they learned to other problems in the family. The control group did not have access to this program but had access to all of the standard programs and resources available to the Montreal public school children in this period.

Over the last three decades the MLES collected detailed longitudinal data from these cohorts on the development of their non-cognitive and cognitive skills, and later adult outcomes. The outcomes include whether or not the subject received a high-school diploma from administrative data, criminal record at age 23 from administrative data, self-reported economic outcomes at ages 17-26 years, including labor market history, employment status and income, as well as self-reported social outcomes related to social capital such as group membership and civic activities.

This paper analyzes the long-term impact of this behavioral skills training on school achievement and adult, criminality and socio-economic outcomes at ages 17-26 years. We find that administrative records document a higher rate of secondary-school graduation for the treatment group (19 percentage points higher than the control group average of 32%), as well as a large (11 percentage points lower than the control group average of 33%) reduction in the probability of having a criminal record. We also provide some limited evidence on other adult outcomes, though these data suffer from some attrition. The treated group is more likely to be active fulltime in either work or school from ages 17-26 (12 percentage points higher than the control group average of 74%). As an adult, the treated group is more likely to belong to a civic or social group (22 percentage point increase on the control group average of 32%).

We then evaluate the channels producing the long-term impact of early childhood social skills training. We examine the relationship between treatment and a large set of intermediate variables concerning cognitive and non-cognitive skills at ages 10-17 years, which we group into early (10-13) and late (14-17) periods. The reason for this division is

data-driven, in that it is based on the potential interaction between behavior and school performance. We observed that divergence between the treatment and control groups in terms of school performance begins at age 14. In addition, this division parallels previous analysis of this program (Vitaro et al, 2012).

We first identify a Self Control channel, based largely on the types of variables used to identify the disruptive sample in 1984. Many of these variables concern behaviors that have to do with controlling ones behavior towards others or things outside of oneself (such as fighting, bullying, and destroying objects), and other variables concern behaviors that have to do with controlling ones behavior towards the task at hand (sitting still, remaining on task, focusing). We analyze a general Self Control channel, one based on the external Self Control variables, and one based on the internal Self Control variables. We find that all three channels are very important for adult outcomes, are significantly changed by the treatment in the early adolescent period, but the change persists into late adolescence only for external Self Control.

We also identify three different of social channels (Generalized Trust, Strong Social Ties, and Altruism), and a channel for Emotional Well-Being. The treatment program does not have any impact on Strong Social Ties, Altruism, or Emotional Well-Being. Our paper thus gives empirical support to the theoretical and observational literature on generalized trust in economics, psychology and sociology literature concerning generalized trust and morality, where behavior concerning people outside of one's family group important to socioeconomic success at an individual level and bilateral exchange more generally. This generalized trust and morality can be contrasted to strong social ties, which connect us to family and friends, with whom we share culture, behavior, and norms.

Finally, we examine school performance as a potential channel, focusing on grades and class assignment (i.e., whether the subject was held back or placed into a special education classroom). School performance channels are changed by treatment in late adolescence, but not in early adolescence.

The potential behavioral channels of impact are thus Self Control, Generalized Trust and school performance. We present some evidence that these changes in late adolescent school performance may be due, at least in part, to changes in behavior in the early adolescent period.

Our study contributes to the growing literature on the impact of education and early childhood development programs on adult outcomes. Substantial evidence shows that early childhood educational intervention, like the Abecedarian (Campbell et al. 2002) and Perry

Programs (Heckman et al., 2010), the Head Start program (Currie and Duncan, 1995; Ludwig and Phillips 2007) or the Project STAR (Krueger 1999, Chetty et al. 2011) leads to a variety of favourable long-term outcomes. Those outcomes include improved earnings and employment prospects, increases in educational attainment, and large drops in criminal and mortality rates. The Perry Preschool Project has been examined with particular attention given to constructing a social rate of return and to the potential mechanisms of impact (Heckman et al, 2010, 2012).³ The results on the benefits of the Perry Preschool Project, in particular, have been important to informing the current policy debate on government support for early childhood education in the United States.

The MLES project differs in three main dimensions from other experiments. First, this intervention is specifically targeted at children with very high deficit in social skills in kindergarten while the selection criterion of previous studies was mostly based on low IQ and low-income status of parents. There is a particular need for research on programs that focus on high-risk individuals (i.e., the “tail” of the distribution of behavior and ability), for equity reasons regarding the wellbeing of the worst off, but also for more practical reasons. Individuals at the end of the spectrum of anti-social behavior during their childhood are more likely to incur costs for society than those towards the middle of the spectrum. For example, untreated individuals in our sample, who were above the 70th percentile of the larger MLES sample in terms of disruptiveness, spent twice as many years receiving public transfers as the bulk of the larger sample (7% vs. 14%) and were more than twice as likely to have a criminal record (33% vs. 15%).

Second, the MLES project is centered exclusively on behavioral trainings. Other early childhood interventions analyzed in the literature were mainly based on additional academic support (such as extra teaching and quantity of schooling, new teaching methods, class size, teacher quality). Although some elements of new pedagogy were introduced,⁴ the non-cognitive skills were not the primary targets of those interventions and the boost in those skills came first as a surprise. The experiment that is closest to ours is Heller et al. (2012). The authors look at a randomized field experiment targeted at the development of social-

³ Our methodological approach differs from that of Heckman et al (2012). We do not attempt to estimate the amount of impact accounted for by each channel. This gives some additional freedom in constructing the channels as we allow the channels to be correlated, which we believe reflects the reality of behavior.

⁴ The Perry Preschool or Abecedarian mainly consisted of an additional quantity of schooling, with additional classroom sessions or home visits by teachers. The curriculum of the Perry Preschool Program also included some non-cognitive tasks with a plan-to-review sequence to develop a sense of responsibility and independence (see Schweinhart, Barnes and Weikart, 1983).

cognitive skills of children in Chicago Public Schools and find a significant effect of this intervention on schooling outcomes and violent crime arrests during the year of the intervention. This intervention is remarkable by the size of the sample (around 3000 children in grade 5-7th) and the social skills they look at (impulse control, emotional regulation, social information processing). However, Heller et al (2012) estimate the impact after one year while the longitudinal MLES follows an entire cohort and permits estimation of the long-term impact of social skills training on a comprehensive set of adult outcomes.

Third, the MLES provide comprehensive measures of non-cognitive skills during the entire childhood and adolescence. Many alternative longitudinal studies lacked detailed assessments of non-cognitive skills, taking them as a “black box” for any non-academic behaviors rewarded by the labor market (see the discussion by Bowles 2013). Several important studies have been able to examine channels of academic motivation and externalizing behavior (Chetty et al. 2011, Heckman et al. 2012). In this paper, access to a wide variety of non-cognitive measures allows us to identify many different channels of impact, and identify potential channels by ruling out others. In particular, we are able to differentiate social trust from altruism, gregarious behavior, school attitude, and well-being to show the ability to connect with and trust other people is the likely mechanism for improvements in economic outcomes and adult social capital (group membership), rather than, for example, improved networks (which would be more likely to be captured by gregariousness) (see Banfield 1956 and Coleman 1990 for seminal discussion of the various dimensions of social capital).

Our paper is also related to recent studies on the MLES project. Boisjoli et al. (2007), Tremblay et al. (1992, 1995, 2000), and Vitaro et al. (2001) have extensively analyzed the MLES effect on the development of aggressive behavior during childhood and on school achievement and crime during adolescence. Boisjoli et al. (2007) examined the impact on criminal records and secondary outcomes. Vitaro et al (2012) examined tested a series of potential pathways to explain impacts on criminal records. Our paper confirms the Boisjoli et al. (2007) results on education and crime and supplements the results in Vitaro et al. (2012) by examining additional channels and using a different methodological approach. We also provide an alternative way of structuring the data that identifies a channel – Generalized Trust – for which measurement instruments have not been developed. This channel seems to be potentially important for adult outcomes, and this finding suggests that creating better instruments specifically to test this trait would be an important line of research.

The paper is organized as follows. In section 2, we review the MLES design and address the potential issues surrounding the validity of the experiment. Section 3 estimates the adult economic and social impacts of the experiment. In section 4 we analyze the effect of the program on intermediate cognitive and non-cognitive skills during childhood. Section 5 estimates the rate of return of the MLES experiment. Section 6 concludes.

2. Experimental design and Data

The Montreal Longitudinal Experimental Study (MLES) began in 1984 and piloted a prevention program for disruptive children with three components: social skills training in a small group format involving a majority of pro-social peers, training in parenting skills during family visits, and teacher management skills (see Tremblay et al. 1992 for a detailed description of the experiment).

2.1 Selection

Kindergarten teachers in 53 schools in Montreal, Canada, were asked to rate the behavior of their male students at the end of the 1984 school year. These schools were located in areas of low socioeconomic status. Almost all (87%) of the teachers provided ratings for a total of 1,161 boys. To create a homogenous sample, only subjects whose parents were Canadian-born with French as a first language and 14 years or less of schooling were included in the longitudinal study, which reduced the number to 895 boys.

These assessments used the disruptiveness scale of the Social Behavior Questionnaire, measuring the frequency of physical aggressions, oppositional behavior and hyperactive behavior (Cronbach $\alpha = 0.93$). Boys with a score above the 70th percentile ($N = 250$) on this disruptiveness scale were considered to be at very high risk of later antisocial behavior (Vitaro et al., 2001). The sample for this paper consists of these 250 individuals.

These 250 subjects were randomly assigned to a treatment (69 boys) and a control group (181 boys) by drawing names from a box. The control group was initially divided into two groups, one of which was a no-contact group during the adolescent years. The no-contact group and the control group showed no difference in outcomes and so the two groups were collapsed into one control group. Some families (78 out of 250) from both the treatment and the control groups refused to participate in some elements of the study, particularly in the elements involving parent participation, but were included in the longitudinal data collection. The rate of non-participation was the same across groups. These subjects are included in the analysis as belonging to their initially assigned treatment groups (intention-to-treat analysis)

in order to avoid any bias due to differential. We use intention-to-treat analysis and use initial group assignment as the indicator for treatment. Figure 1 provides a visual depiction of the recruitment and selection process of the MLES.

2.2 Intervention Program

The intervention program was implemented over a 2-year period, from ages 7 to 9 years (Grades 2 and 3). The prevention program included three main components to improve children's non-behavioral skills, particularly social interactions and self-control: child training, parent training, and teacher training. The teacher training component was not fully implemented and so we focus on the child and parent training here. Appendix A provides detailed information on the program and contents of the training sessions.

2.2.1 Child social skills training

The first aspect of the intervention consisted of direct training on social skills to children. The experiment drew on non-randomized studies from psychology showing correlation between pro-social behavior and specific skills such as emotional regulation and impulse control, social-information processing and how to interpret other's intent (Cartledge and Milburn, 1980; Kettlewell and Kausch, 1983; Michelson et al., 1983; Schneider and Byrne, 1987; Weiss et al., 1992; Dodge 2003 and 2007).

The first aspect of the intervention consisted of social skills training with small groups of children. The content of the program was adapted from previous programs for children with behavior problems (Cartledge and Milburn, 1980; Kettlewell and Kausch, 1983; Michelson et al., 1983).

The social skills training sessions were conducted at school (outside the classroom), in groups of four to seven children, of which one or two would be the treatment subjects, and the rest would be boys identified by their teachers as pro-social. This arrangement was intended to provide positive role models for the treatment subjects and also avoided stigmatizing the treatment subjects. The sessions were held once a week for 45 minutes, during lunch or after school. During the first year, nine sessions of social behavior training were offered. Sessions included how to invite a bystander to play, how to ask "why", how to give a compliment, and how to help. The second year included 10 sessions of problem solving and self-control strategies (Camp et al 1977; Goldstein et al, 1980). Some stimulus situations for these sessions were how to react to teasing, how to react when angry, and what to do if other children refuse to play with you. For each situation, the children reviewed ways to define the

problem, identified the intentions of the other person, analyzed their feelings if they were in the role of the victim, suggested different action plans to solve the problem, anticipated their consequences, selected one action plan and, finally, reinforced themselves for their cognitive work. Verbal instructions, coaching, modeling, behavior rehearsal, and positive (verbal and material) reinforcement were all used. Children were encouraged to use their newly learned skills before the next training session. At the following meeting, the children were reinforced for having performed their new skills in the interim. Teachers and parents were informed through one-page letters of the new skills learned by the children during each session. They were encouraged to praise the child for using these new skills as often as possible.

For example, one session covered Self Control. The facilitator introduced the topic, and talked about situations where children are upset and might want to make an angry outburst, like a spilled glass of milk or a disappointment. The facilitator then modeled a situation: she has been playing tag, and she just got tagged and is now out. She's upset because she is the first person to be tagged out, and she's angry and disappointed. The facilitator demonstrates how children can respond in this situation: noticing clues in her body that she is going to lose control (clenching fists, feeling hot), she thinks about what happened to make her feel this way (she got tagged first, is worried other kids will laugh at her), she chooses a way to avoid making an angry outburst (count to ten, move away, say to herself "calm down", breathe), and then she acts and praises herself. The facilitator then invites children to perform additional role-plays based at school (one child bumps another's desk and their pen falls down), at home (someone suddenly turns off the TV because it's time for dinner) or while playing (a friend takes a ball that was dropped). Together, the group makes observations about what the actors are doing, how they are following the steps, and gives feedback. At the end of the session, the facilitator fills out a workbook with the children that explains how they can practice self-control until the next session ("homework").

2.2.2 Parent Social skills training

The second component of the MLES consisted of parent training in child rearing. This intervention was modeled on the Oregon Social Learning Center model (Patterson, 1982). General goals of the family intervention were to reduce coercive interactions, increase consistency of consequences for behaviors, and improve family conflict interactions. The program focused specifically on setting clear rules, supervising the child, positive reinforcement for pro-social behaviors, and problem solving skills. The family intervention was tailored to individual family needs and capacities. It included: (1) a reading program for

parents; (2) training parents to monitor their child's behavior; (3) training parents to positively reinforce their child's prosocial behavior; (4) training parents to punish effectively without being abusive; (5) training parents to manage family crises; and (6) helping parents to generalize what they have learned.

Work with the parents was planned to last for 2 school years with one session every 2 weeks. The sessions were intended to be weekly and last for one hour, with the frequency of the sessions decreasing over time. Parents participated in an average of 17.4 sessions, with a maximum of 47 sessions.

An example of the parental social skills training is given by the "Observation period". When focusing on encouraging pro-social behaviors, the facilitator would first ask the parents to undertake a period of observation. The parents would spend some time each day observing, and recording on a worksheet, the interactions between the child and parents (for example, "Mother asks for help setting table, child whines and grumbles, mother threatens punishment, child complies."). These observations would then form the basis for identifying patterns of behavior, and the facilitator could propose some strategies for addressing these behaviors. Role-playing might reinforce these strategies; particularly as playing the role of the child allows the parent a different point of view on their own actions. For example, in the case of pro-social behaviors, the facilitator would suggest letting the child know what is expected of him at the beginning of the day, so that there is no room for argument, and this interaction might be role-played. Finally, when the facilitator and parents had identified a few behaviors for work during the coming week, a "contract" between the parent and child might be drawn up, where the desired behaviors or tasks are listed ("Set the table", "Say hello when someone comes in" etc) and children would receive points each time the behavior manifested itself or the task was completed. Children could then use points for some reward: A game with dad after dinner, an extra half hour before bedtime, or some other special treat. In the next session, the facilitator would review the enforcement of the contract with the parents, and adjust goals and strategies as necessary.

A third component of the intervention consisted of information and support for teachers that focused on responding to the behavior of the at-risk boys, which was intended to reinforce the parent training. The intention of this third component was to improve teachers' management skills of behavior problems in the classroom and set up individualized behavior management programs for the participants. However, compliance with this aspect of the intervention was weak.

The training and support activities were carried out by two university-trained child-care

workers, one psychologist, and one social worker, both working full time. Each of those therapists had a caseload of 12 families. The team was coordinated by a fifth professional who worked on the project half time. The intervention program lasted 2 school years, from September 1985 to June 1987. Boys were 7 years of age when the intervention started and 9 years of age when it ended.

2.3 Timeline and Data

Figure 2 shows the timeline for data collection. Following the two-year intervention program between ages 7-9 (1985-1987), the MLES collected measures of cognitive and non-cognitive skills during the entire childhood and adolescence on a yearly basis (from age 10 to age 17, period 1988-1995). Those measures will be used to identify the channels of impact of the MLES experiment on various outcomes. The outcome variables are measured with administrative data and questionnaires. When participants were 21 and 26 years old (2001 and 2006), the MLES run adult questionnaire with detailed measures of the respondent economic and social situation. Around age 23-24 (year 2003), the MLES collected administrative data on criminal record and school achievement. We present the main outcome variables of interest below.

Educational achievement. Administrative information on whether each subject had received a secondary-school level degree was collected. This information was collected when the participants were around 23-24 years old (2003). This data is complete for 242 of the 250 treatment subjects.

Criminal record. Administrative data on the number of offenses, whether the offense was violent or nonviolent, whether a weapon was involved, and year of each offense were obtained. This information was collected when the participants were around 23-24 years old (2003) and includes adult criminal records prior to age 23. This data is complete for all subjects. We also include data on Juvenile criminal records.

Fulltime occupation. The MLES collected a wide range of adult economic outcomes. Subjects reported their activities (employment, school, recipient of government transfers) at age 21 and 26, and their recalled activities each year since age 17 (when they were 21). There are missing data at both age 21 and 26. We use these activities to construct a variable reflecting the percent of years reported years from 17-26 where the subject was occupied full time by either school or work or both. We also measure the wages at age 26 which we use as

one measure of cost-effectiveness in section 5.⁵ We also provide the impact on reported employment status and government transfers, but focus on the percent of time occupied fulltime as the principal economic outcome variable.

Adult Social Capital. We have one simple measure of adult social capital, whether or not the subject reported belonging to a civic or social group at age 21 or age 26.

Table 1 shows the summary statistics for the administrative data on adult criminal record and high-school diploma, and the self-reported data from the adult questionnaires at ages 21 and 26. Table 2 shows the summary statistics for the juvenile criminal records (aggregated and by type of crime).

2.4 Validity of the experiment design and baseline controls

The validity of the causal inference from the MLES rests on the quality of the randomization. To test for successful randomization, we test for balance in pre-determined variables between the treated and the control group.

Table 3 shows the baseline values of two groups for several critical variables measured prior to randomization. Among the 32 baseline variables, there are significant differences on 4 variables: initial anxiety measures, age of father at birth of subject, prestige of the mother's employment (all at the 10% level), and number of sisters (at the 1% level). The fact that there are some differences does not indicate that the selection process was non-random. It is not surprising to find imbalances for a handful of variables, given that the sample size is small and 32 variables were tested for differences and so we do not fear that the randomization protocol was violated. However, since these variables might impact the channels and outcomes we wish to examine, we control for them in the analysis that follows. Since some of the subjects are missing these control variables, we impute missing values for the control variables, using available baseline data, to maintain sample size.

Missing data is a problem for some of the outcomes and intermediate variables. One potential solution is dropping observations with missing data; the other is to impute the

⁵ We impute wages for the unemployed based on previous employment and secondary graduation. However, since there is a large number of unemployed, the results are sensitive to the assumptions used to impute wages for the unemployed (such as the replacement rate). For this reason, we do not focus on hourly wages as a principal outcome, though we do not exclude it from the index. Excluding it does not substantially change the results.

missing data points using data from other individuals in the sample to predict the missing data point. The former brings the risk of differential attrition and bias, as well as the possibility of substantially reducing sample size (which is already small). The second is not possible when the variable of interest is itself an outcome of interest.

We have imputed certain baseline measurements that are necessary to eliminate an imbalance between the groups (see above) but omit observations when the intermediate or outcome variable of interest is missing. Father's age, prestige of the mother's work, and number of sisters were imputed using all other available baseline information. We include a dummy equal to one when the value is imputed.

There are particular concerns with attrition in the self-reported variables (percent of years occupied fulltime and group membership). We give more information on the attrition on these variables and provide some evidence that the attrition is unlikely to substantially bias the treatment effect estimates in Appendix D.

3. The Effect of MLES on Adult Outcomes

To estimate the total impact of treatment on outcomes, our principal specification uses OLS with robust standard errors. Recall that the treatment and control groups are composed of the most disruptive subjects in the larger MLES sample. The data from the non-disruptive group contains useful information about the level of the adult outcomes in the non-disruptive group, the correlation of covariates to the outcomes, the relationship of the channels to the outcomes, and wage information used to calculate the replacement wage for the unemployed. For this reason, we retain observations that were not in the treatment and control groups, and our principal estimate comes from

$$y_i = \alpha + \beta T_i + \lambda S_i + \varepsilon_i \quad (1)$$

Where T_i is a dummy variable equal to one for the treatment group and zero for the control group and the non-disruptive group, and S_i is a dummy variable equal to one for the treatment and control groups, and zero for the non-disruptive group. λ may then be interpreted as the difference between the control group and the non-disruptive group, β as the difference between the treatment group and the control group, and $\lambda + \beta$ as the difference between the treatment group and the non-disruptive group. This format of estimation allows us to observe

how much of the difference between the disruptive group and the non-disruptive group the treatment has been able to redress.

In addition to the preferred specifications, we present results for a logit regression (when the outcome is binary),⁶ a specification that includes fixed effects at the level of the school in 1984, and a specification that includes clustered standard errors at the level of the school in 1984. Note that randomization was carried out at the individual level (within schools) and so neither fixed effects nor clustering are called for, but we present these results as robustness checks.

3.1 Crime

Figure 3 shows the percent of people in each group who have a criminal record of any kind. We find a large negative point estimate of the coefficient, significant at 10%: the treatment group is one-third (11 percentage points) less likely to have a criminal record than the control group to have a criminal record (column (1) of Table 4), controlling for the baseline variables discussed in section 2. These results are of a similar magnitude as those found in Boisjoli et al. (2007). This result is robust to the logit specification (column (2) of Table 4). When standard errors are clustered at the 1984 school level or when fixed effects are included (column (3) and column (4) of Table 4), the coefficient on treatment is no longer significant at 10% ($p=0.13$ and $p=0.11$, respectively) but note that the point estimate is not very different when fixed effects are included.

Our specification allows us to observe that this impact brings the treatment group a bit more than halfway back to the achievement of the non-disruptive group. Another way to put this is that it reduced the gap in outcomes between the non-disruptive and disruptive subjects by 60%.

The impact on juvenile crime, shown on aggregate and by type of crime in Table 2 and in Figure 4, is less clear. There is clearly a tendency towards less frequent offense by the treatment among almost all types of crime, but this difference is statistically significant only for the number of major theft offenses, and number of minor delinquencies, and then only when imbalance controls are included.

⁶ Since the standard means-based calculation of marginal effects for one variable depends on the values of the other covariates, changes in the marginal effects coefficient when covariates are added may be difficult to interpret. We present the coefficients rather than the marginal effects.

3.2 Secondary Diploma

Figure 5 shows the percent of people in each group who completed secondary school according to administrative records. Column (1) of Table 5 shows a significant and very large coefficient: subjects in the treatment group are 19 percentage points more likely to have received a secondary school diploma than subjects in the control group, significant at the 5% level (note that only 32% percent of control group subjects completed secondary school). As with the criminal record, these results are similar to those found in Boisjoli et al. (2007). The results are essentially unchanged with the logit specification and when including fixed effects or clustered standard errors at the 1984 school level (columns (2)-(4) of Table 5).

Our specification shows that the treatment reduced the gap between the non-disruptive and disruptive group by 80%.

3.3 Fulltime occupation

Figure 6 shows the distribution of the percent of reported years 17-26 where the subjects were occupied fulltime either by work or school. Column (1) of Table 6 shows that the treatment subjects were occupied 12 percentage points more than the control group (the control group average is 74%), significant at the 5% level. The results are essentially unchanged with clustered standard errors and fixed effects. Table 7 shows the other economic outcomes, and we observe significant effects on percent of years 17-26 working, and employment, full-time employment, and hourly wages (accounting for unemployment) at age 26. There is no significant impact on occupied or occupied fulltime at age 21, or on the percent of government transfers, though the point estimates are in the expected direction.

3.4 Social Capital

Figure 7 shows the percent of people in each group who were a member of a social group at age 21 or age 26. Table 8 shows that treatment group members were more than 20 percentage points more likely to belong to a social group, from 32% in the control group, significant at the 5% level (column (1) of Table 8). The results are essentially unchanged using a logit specification, clustered standard errors, and fixed effects.

4. Potential Channels of Impact

4.1 Identification of the Channels

This section estimates the effect of the MLES experiment on changing individual traits measured during adolescence, identified later as the channels of impact on adult outcomes. Measures relative to the different channels were collected annually from ages 10 and 17 years, as illustrated in Figure 2. Previous studies on this sample have found that subjects are on different trajectories of behavior, that is, subjects may be grouped as to whether their disruptiveness is increasing, decreasing, or stable over time with respect to their initial disruptiveness levels (Tremblay et al., 1991). We break this period into two and analyze data that are available from both periods: ages 10 to 13 (1988-1991), and ages 14 to 17 (1992-1995). The year 1992 was chosen as the break year because it is in this year that the treatment and control groups begin to diverge in whether or not they have repeated a year (see Figure 5), and repeating (being “held back”) has a large correlation with later life outcomes, as discussed below.

Because this paper does not aim to compare this population to standardized measures of the channels, we use individual questions from the psychological scales rather than the calculated scales themselves. This approach allows for the possibility that individual questions might cluster together effectively and also allows subject-reported and teacher-reported data to be used together when possible.

We begin by using the variables related to disruptiveness, which was initially used to identify the experimental sample using teacher-reported data. We combine these variables with subject-reported variables that are closely related and together form a channel related to Self Control. In addition, we note that Self Control has two aspects: external Self Control and internal Self Control (discussed below).

Our identification of further channels is based on exploratory factor analysis. Combining all data available and averaging over the years available, we use factor analysis to examine how the factors combine into groups potentially measuring the same latent variable. We include questions from several psychological inventories (Jesness and Wedge, 1983; Kovacs, 1983; March, 1990; Rosenberg, 1965; Lacourse et al., 2002; Tremblay et al., 1992). There is substantial correlation across channels (see Table 15), which demonstrates the underlying similarity of many of the behavioral traits.

Table 8 shows summary statistics for the construction of each channels: number of components, the Cronbach alpha index for internal consistency, the mean and number of observations for the non-disruptive, control, and treatment groups, and the p-value for the test

of equality of mean (t-test) and distribution (Kolmogorov-Smirnov test) between the treated and the control groups. Table 9 shows the correlation between the identified channels. Appendix C gives full information on the component variables for each channel.

Self Control (Figure 9) is the original disruptiveness scale combined with self reported variables. It includes information on whether the subject fights with others, damages property, is impolite, finds it hard to apply himself at school, gets upset when provoked (self-reported), is disobedient, fights, has poor concentration, and jumps from one thing to another (teacher reported).

Self Control – External (Figure 10) is the externalized component of Self Control. It includes variables on fighting, damaging property, reacting strongly to provocation (self-reported), and fighting and bullying (teacher reported).

Self Control – Internal (Figure 11) is the internalized component of Self Control. It includes variables on difficulty in applying himself in school, finding it hard to concentrate, finding it hard to find things to do (self-reported), and weak concentration, easily distracted, and jumps from one thing to another (teacher reported).

Generalized trust (Figure 12) measures generalized trust and weak social ties. It includes variables on whether the subject thinks the police are trustworthy, cares about what the teacher thinks of him, whether it is acceptable to steal from rich people or dishonest storeowners, and whether one is better off not to trust anyone (self-reported), assumes that a bump is intentional and gets angry (self and teacher reported), and is inconsiderate of others (teacher-reported).

Strong Social Ties (Figure 13) measures close relationships with family and friends. It includes variables on whether the subject trusts his best friend, gets advice from his best friend, shares his problem and personal information with his best friend, lies, cares about whether other people like him, spends time with his friend and talking to his friend, talks about the future with his parents and spends time talking with his parent (self-reported) and whether he tends to work alone and is liked by other children (teacher-reported).

Emotional Well-being (Figure 14) measures psychological well-being and self-esteem. It includes the components of the Rosenberg Self-Esteem scale, as well whether the subject thinks other people do things more easily than himself, people talk about him behind his back, he has trouble making decisions, he feels alone even when he's with someone, and that when

something doesn't work he is to blame (self-reported) as well as whether he gives up easily (teacher reported).

Altruism (Figure 15) measures active altruism. It is composed only of teacher-reported variables, and includes whether the subject tries to stop others from fighting, invites a child who is left out to play, helps injured children, volunteers to put things away, congratulates others, shows sympathy, helps children having trouble, helps ill children, or volunteers to clean up a spill.

School Performance Channels (Figure 8 and Figure 16) include yearly grades in Math and French, as well as the status of the subject's class each year: whether the subject had been held back or was in special education class.

To measure the channels, we used both principal component analysis and a simple average of z-scores for the non-missing variables available for a given observation. The estimated channels are quite similar, and the results do not differ substantially but there are fewer observations available. We use the z-score averages in the analysis that follows.

4.2 Potential channels of impact

Determining whether each channel explains part of the treatment effect requires both that the channel itself impacts the outcome and that the channel be changed by the treatment. Table 8 provides the coefficients on the channels in a regression on each outcome variable. All channels, with the exception of altruism, are consistently highly correlated to adult outcomes.⁷

In addition to the channels being significantly associated with both the treatment and the outcomes, we should find an impact on channels that are anticipated to be mediators – that is, channels that the program was designed to change (Maric et al., 2012; MacKinnon et al., 2004).

We test for the impact of treatment using a simple t-test, a Kolmogorov-Smirnov test for equality of distributions, and OLS regression with controls for imbalances. The results

⁷ The fact that altruism is more weakly related to adult outcomes than the other channels is somewhat surprising. One way in which altruism is different from the other channels is that it is the only channel that is measured exclusively with teacher responses. Thus the lack of correlation between altruism and concrete adult outcomes might be either that altruism is indeed not related to adult outcomes, or that teachers are not very perceptive or good at reporting altruistic tendencies.

from the different tests are generally similar with two exceptions.⁸ We find that the treatment has a significant impact on Self Control (both external and internal) and Generalized Trust in the first period (ages 10-13), and on Self Control (external), Generalized Trust, Grades, Held Back and Special Education in the second period (ages 14-17) (Table 10).

One puzzle is the increase in school performance. Since the treatment program included no training in cognitive skills (as opposed to other well-known programs such as Perry Preschool or Headstart), and no improvement on IQ tests was observed, we hypothesize that improvements in school outcomes are the consequence of changes in the behavioral channels (for example, that subjects with better self-control were less likely to be held back). Tables 11 and 12 present the coefficients for simple regressions of Period 1 (10-13 years) channels on the significantly impacted Period 2 (14-17) channels.

5. Rate of Return

An estimate of the overall benefits of the program would include the increase in productive of the treatment subjects (employment and wages), as well as the reduction in the cost to society from reduced criminal behavior. We are unable to estimate the overall rate of return, and so we examine the major social and economic outcomes separately.

5.1 Cost of the program

Unfortunately, we do not have information on the actual cost of the program. However, we do know the composition of the implementation team (1 full time social worker, 2 full time childcare specialists (BA level), 1 psychologist, and 1 half-time program administrator). We use reference salaries for these positions in Montreal in 2011, and we assume an additional 20% of salaries in administrative costs. We then calculate how much the program would cost, per person, to be implemented in 2011 (for two years). We calculate that the total cost per person would be around \$6,500 in 2011 US dollars.

5.2 Cost Efficiency: Criminality and Secondary Completion

8 The first exception is grades during ages 10-13, which are highly significantly related to treatment using the Kolmogorov-Smirnov test ($p=0.01$) but not using the t-test or OLS regression with controls ($p=0.33$ and $p=0.131$, respectively). This appears to be due to a clump of treatment students with low grades, which can be observed in Figure 12, and when these observations are omitted the OLS results are significant. The second exception is Generalized Trust from ages 14-17, which, on the contrary, is significant using a t-test or OLS regression with controls ($p=0.04$ and $p=0.003$, respectively) but not using the KS test ($p=0.22$).

Benefits due to reduced criminal behavior are difficult to estimate, as they include not only the cost of enforcement and, if applicable, incarceration, but also the cost to society of the crime committed. One common strategy is to use jury awards as an estimate of the monetary damage due to different types of crimes. We do not have data available on the particular crime committed. Moreover, we only have data available on crimes committed up until age 23, and so an estimate of cost-efficiency requires some assumptions about what is likely to happen after age 23, and it is unlikely that people commit crimes at the same rate as they age. We provide different estimates for cost-eficiency depending on the expected duration of the criminal career (Table 13), and in all cases assume that the rate of crime commission decreases by 10% per year and we discount future crimes at 3%. The cost per criminal record averted ranges between a high upper bound estimate of \$5,406 if subjects stopped committing any crime at age 23 (likely to be an unrealistic assumption) and \$1,767 if subjects continued committing crimes for 15 years (until age 38).

Benefits to increased secondary graduation are similarly difficult to estimate and so we provide an estimate of cost per additional secondary graduation. The cost of each additional secondary graduation is around \$40,000 in 2011 US dollars.

5.3 Rate of Return: Hourly wages

We calculate the increase in hourly wages from using several different strategies.⁹ First, we regress treatment on reported hourly wages at age 26. For those who are unemployed at age 26 (and thus have no reported wage), we impute the hourly wage using linear regression on baseline psychological characteristics and secondary school completion, obtaining the coefficients for the imputation from the entire cohort (including the non-disruptive subjects). Those with missing information on employment are not included in the calculation. The parameters we use for the estimation are that subjects work only part time (20 hours per week) and that their working life lasts 40 years (for this population, this implies that they work from ages 18 to 58). We use a discount rate of 3%. A sensitivity analysis (for different assumptions about working life and hours worked) is presented in Table 14.

Under these assumptions, we calculate a rate of return of 450%. This rate of return is quite high, implying that every dollar spent today yields \$4.5 in future benefits. This estimate

⁹ We do not use hourly wages as a principal outcome in this paper because of it is necessary to estimate a replacement wage for the unemployed, which makes these data very sensitive to assumptions about the replacement rate.

is in line with the high rate of return found in other cost-benefit analyses of childhood development programs.

6. Conclusion

We find that an early childhood program conducted in inner-city Montreal in the 1980s and focused on improving social skills had substantial long-term positive effects: crime rates decreased and secondary school graduation increased. We provide some evidence that social and economic outcomes improved, though there are concerns about attrition in the sample. While the sample size for this particular project is small, the similarity of the most basic version of our results (large positive impacts on adult outcomes from interventions in childhood) is congruent with the most basic version of the results obtained from other studies in multiple diverse settings, suggesting that external validity concerns are likely to be small.

The policy implication of this paper is that increased investment in childhood development programs is likely to be an efficient and profitable public policy, especially where such programs explicitly incorporate simple strategies to foster the development of non-cognitive skills, and in particular social skills and self-control. Outside of the public interest in the welfare of such individuals, such investment is likely to be particularly cost-effective when made for those children most at risk of unemployment and criminality as adults. We estimate a large rate of return to this program – \$1 invested yields at least \$4.5 in benefits - even under conservative assumptions, which is in line with other estimates of the profitability of early childhood development programs. We use a rich longitudinal dataset that includes subjective and teacher, parent and peer reported-data, as well as psychological inventories, to provide evidence on the channels of impact.

This paper also provides new micro evidence on the relationship of pro-social behavior to economic outcomes. First, we provide new evidence that pro-social behavior has an impact on social and economic success at an individual level. Second, we show that, among several possible types of social skills, it is generalized trust that is most relevant for adult success at the individual level, rather than altruism or strong social (family and close friends) ties.

This paper cannot disentangle the effect of changes in behavior that lead children to nonstandard class assignment and the effect of the nonstandard class assignment itself, but we provide evidence that whether or not nonstandard class assignment is harmful or beneficial to children may be an important question for future research.

References

- Y. Algan, P. Cahuc, Trust, Growth and Well-Being: New Evidence and Policy Implications. *Handbook of economic growth*, in Aghion and Durlauf (eds), Elsevier (2013).
- Y. Algan, P. Cahuc, A. Shleifer, Teaching practices and Social Capital. *American Journal of Economics: Applied*, (2013).
- D. Almond and J. Currie. Human Capital Development Before Age Five, *Handbook of Labor Economics*, Volume 4, Elsevier (2010).
- E. Banfield, *The Moral Basis of a Backward Society*. New York: Free Press (1958).
- G. Becker, W. Hubbard, K. Murphy, The Market for College Graduates and the Worldwide Boom in Higher Education of Women. *American Economic Review*, **100(2)**, 229-33 (2010).
- M. Bertrand, J. Pan, The Trouble with boys: social influences and the gender gap in disruptive behavior, NBER working paper, 17541 (2011)
- R. Boisjoli, F. Vitaro, E. Lacourse, E.D. Barker, R. Tremblay, Impact and clinical significance of a preventive intervention for disruptive boys: A 15-year follow-up. *British Journal of Psychiatry*, **191**, 415-419 (2007).
- S. Bowles, A (Science-Based) Poor Kids' manifesto, (Review of Heckman James, Giving Kids a Fair Chance: A strategy that works, Cambridge MIT), *Science*, **340**, 1044-1045 (2013).
- S. Bowles, H. Gintis, M. Osborne, The determinants of earnings: A behavioral approach. *Journal of Economic Literature*, **39**, 1137-1176 (2001).
- B.W. Camp, G.E. Blom, F. Hebert, W.J van Doorminck, Think aloud: A program for developing self-control in young aggressive boys. *Journal of Abnormal Child Psychology* **5**, 157-169 (1977).
- F.A. Campbell, C.T. Ramey, E. Pungello, J. Sparling S. Miller-Jonhson, Early childhood education: Young adult outcomes for the Abecderian project. *Applied Development Science*, **6**, 42-57 (2002).
- G. Cartledge, J. F. Milburn, *Teaching social skills to children. Innovative approaches* (Pergamon Press, New York, 1980)
- N. Castellanos-Ryan, J. Séguin, F.Vitaro, S.Parent, Richard E. Tremblay, A multimodal intervention for disruptive kindergarten children reduces substance-use across adolescence: a randomized control trial, *British Journal of Psychiatry*, forthcoming (2013).
- R. Chetty, J. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, D. Yagan, How does your kindergarten classroom affect your earnings? Evidence from project Star. *Quarterly Journal of Economics* **126(4)** (2011)

- D.R. Clasen, B.B Brown, The multidimensionality of peer pressure in adolescence. *Journal of Youth and Adolescence* **14**, 451-468 (1985).
- J. Coleman, *Foundations of Social Theory*. Cambridge, MA: Harvard University Press (1990).
- J. Currie, T. Duncan, Does Head Start make a difference? *American Economic Review*, **85**, 341-364 (1995).
- J. Currie, R. Khan, *The future of Children*, Princeton-Brookings, Washington DC (2012).
- J. Dewey, *Democracy and Education*, New York, NY: Dover (1944).
- K. Dodge, Godwin J., Conduct Problems Prevention Research Group, Social-information-processing patterns mediate the impact of preventive intervention on adolescent antisocial behavior. *Psychol Sci* **24(4)**, 456-65 (Apr 2013).
- K. Dodge, Processes in the Prevention of Crime and Delinquency. *Controlling Crime: Strategies and Tradeoffs*, P. Cook, J. Ludwig, J. McCrary (eds), University of Chicago Press (2011).
- K. Dodge, *Do social information-processing patterns mediate aggressive behavior?* (2003)
- E. Fehr, On the Economics and Biology of Trust, Presidential address at the 2008 meeting of the European Economic Association, *Journal of the European Economic Association*, **7 (2-3)**, 235-266 (2009).
- N. Fortin, P. Oreopoulos, S. Phipps, Leaving Boys Behind: Gender Disparities in High Academic Achievement, Working Paper Department of Economics UBC (2011).
- A. P. Goldstein, R. P. Sprafkin, N. J. Gershaw, P. Klein, The adolescent: Social skills training through structured learning. In G. Cartledge & J. F. Milburn (Eds.), *Teaching social skills to children: Innovative approaches* (Pergamon Press, New York, 1980) pp. 249-277.
- J. Heckman and Y. Rubinstein, The Importance of non-cognitive skills : lessons from the GED testing program, *American Economic Review*, **91**, 145-149 (2001).
- J.J. Heckman, S. H. Moon, R. Pinto, P. A. Savelyev, A. Q. Yavitz, Analyzing social experiments as implemented: A reexamination of the evidence from the High Scope Perry Preschool. *Quantitative Economics* **1(1)**, 1-46 (2010)
- J.J Heckman, S.H. Moon, R. Pinto, P. A. Savelyev, A. Yavitz, The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics* **94(1)**, 114-128 (2010).
- J.J. Heckman, L. Malofeeva, R. Pinto, P. Savelyev, Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, in press.
- S. Heller, H. Pollack, R. Ander and J. Ludwig, Improving social-cognitive skills among disadvantaged youth: A randomized field experiment, Working Paper Chicago (2013).

- B. Jacob, Where the Boys Aren't: Non-cognitive Skills, Returns to School and the Gender Gap in Higher Education. *Economics of Education Review*, **21(6)**, 589–98 (2002).
- C. Jencks, *Who gets ahead? The determinants of Economic Success in America*, (Basic Book, New York, 1979).
- C. F. Jesness, R. F. Wedge, *Classifying offenders: The Jesness inventory classification system* (California Department of Youth Authority, Program Research and Review Division, 1983).
- L. Katz, J. Ludwig, G. Duncan, L. Gennetian, R. Kessler, J. Kling, and L. Sanbonmatsu, Long-Term Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity," *American Economic Review*, 103, 226-31 (2013).
- L. Katz, J. R. Kling and J. B. Liebman, Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment, *Quarterly Journal of Economics*, 116, 607-654 (2001)
- J.R Kling, J. B. Liebman and L. F. Katz, Experimental Analysis of Neighborhood Effects, *Econometrica*, 75(1), 83-119 (2007)
- P. W. Kettlewell, D. F. Kaush, The generalization of the effects of a cognitive behavioral treatment program for aggressive children. *Journal of Abnormal Child Psychology* **11**, 101-114 (1983).
- M. Kovacs, *The Children's Depression Inventory: A self-rated depression scale for school-aged youngsters*. (University of Pittsburgh School of Medicine, Department of Psychiatry, Western Psychiatric Institute and Clinic, 1983).
- A. Krueger, Experimental Estimates of Education Production Functions. *Quarterly Journal of Economics*, 114, 497-532 (1999).
- E. Lacourse, S. Cote, D. S. Nagin, F. Vitaro, M. Brendgen, R. E. Tremblay, A longitudinal-experimental approach to testing theories of antisocial behavior development. *Development and Psychopathology* **14**, 909-924 (2002).
- Lipsey, M., "The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic review," *Victims and Offenders*, 4 (2009), 124-147.
- Lipsey, M, and Cullen, F. "The Effectiveness of Correctional Rehabilitation: A Review of Systematic Reviews," *Annual Review of Law and Social Science*, 3 (2007).
- J. Ludwig, D.L. Miller, Does Head Start improve children's life chances? Evidence from a regression-discontinuity design. *Quarterly Journal of Economics* **122(1)**, 159-208 (2007).
- M. Maric, R.W. Wiers, P. J. M. Prins, Ten ways to improve the use of statistical mediation analysis in the practice of child and adolescent treatment research, *Clin Child Fam Psychol Rev* **15**, 177-191 (2012).
- D.P. MacKinnon, C.M. Lockwood J. Williams, Confidence limits for the indirect effect: distribution of the product and resampling methods, *Multivariate Behavioral Research* **39(1)** 99-128 (2004).

- H.W. Marsh, A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review* **2(2)**, 77-172 (1990).
- D. Meichenbaum, *Cognitive-behavior modification: An integrative approach*. (Plenum Press, New York, 1977).
- L. Michelson, D.Sugai, R.Wood, A.E Kazdin, *Social skills assessment and training with children*. (Plenum Press, New York, 1983).
- G.R. Patterson, *Coercive family process*. Eugene, OR: Castalia. (1982).
- G.R. Patterson, J. B. Reid, R. R. Jones, R.W. Conger, *Families with aggressive children Volume 1 : A social learning approach to family intervention* (Castalia, Eugene, OR, 1975).
- R. Putnam, *Making Democracy Work*. Princeton, NJ: Princeton University Press (1993)
- R. Putnam *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster (2000).
- M. Rosenberg, Rosenberg self-esteem scale (RSE). *Acceptance and Commitment Therapy. Measures Package* **61** (1965).
- B. H. Schneider, B. M. Byrne, Individualizing social skills training for behavior disordered children. *Journal of Consulting and Clinical Psychology* **55**, 444-455 (1987).
- L.J Schweinhar , J. Montie, Z. Xiang, W. S. Barnett, C. R. Bel_eld, and M. Nores (2005) "Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40". Ypsilanti, MI: High/Scope Press.
- C. Segal, Misbehavior, Education and Labor Market Outcomes, *Journal of European Economic Association*, in press (2013)
- P. Szatmari, D. Offord, M. Boyle, Ontario Child Health Study: Prevalence of attention deficit disorder with hyperactivity. *Journal of Child Psychology and Psychiatry*. **30(2)**: 219-230 (1989).
- R. Tremblay, The development of aggressive behavior during childhood: what have we learned during the past century. *International Journal of Behavioral Development*, **24**, 129-141 (2000).
- R. Tremblay, F. Vitaro, L. Bertrand, M. Le Blanc, H. Beauchesne, H. Boileau, H., Parent and child training to prevent early onset of delinquency: The Montreal longitudinal experimental study. In J. McCord & R. E. Tremblay (Eds.), *Preventing antisocial behavior: Intervention from birth through adolescence* (Guilford Press, New York, 1992) pp. 117-138.
- R. E. Tremblay, F. Vitaro, C. Gagnon, A Prosocial Scale for the Preschool Behavior Questionnaire: Concurrent and Predictive Correlates. *International Journal of Behavioral Development* **15(2)** 227-245 (1992).

R. J. Tremblay, E. McCord, J. Boileau, P. Charlebois, C. Gagnon, and M. LeBlanc, Can disruptive boys be helped to become competent? *Psychiatry*, **54(2)** 2, 148-161 (1991).

R. E. Tremblay, L. Desmarais-Gervais, C. Gagnon, P. Charlebois, The Preschool Behavior Questionnaire: Stability of its Factor Structure Between Cultures, Sexes, Ages and Socioeconomic Classes. *International Journal of Behavioral Development* **10(4)** 467-484 (1987).

F. Vitaro, D. Larocque, M. Janosz, R. Tremblay, Negative social experiences and dropping out of school. *Educational Psychology* **21**, 401-415 (2001).

T.J Wansbeek, and E. Meijer. *Measurement error and latent variables in econometrics*. Vol. 37. (Elsevier, Amsterdam, 2000).

B. Weiss, K. A. Dodge, J. E. Bates, G. S. Pettit, Some consequences of early harsh discipline: Child aggression and a maladaptive social information processing style. *Child Development* **63**, 1321-1335 (1992).

APPENDIX A. MAIN TABLES AND FIGURES

Figure 1. Illustration of the Experimental Design of the MLES

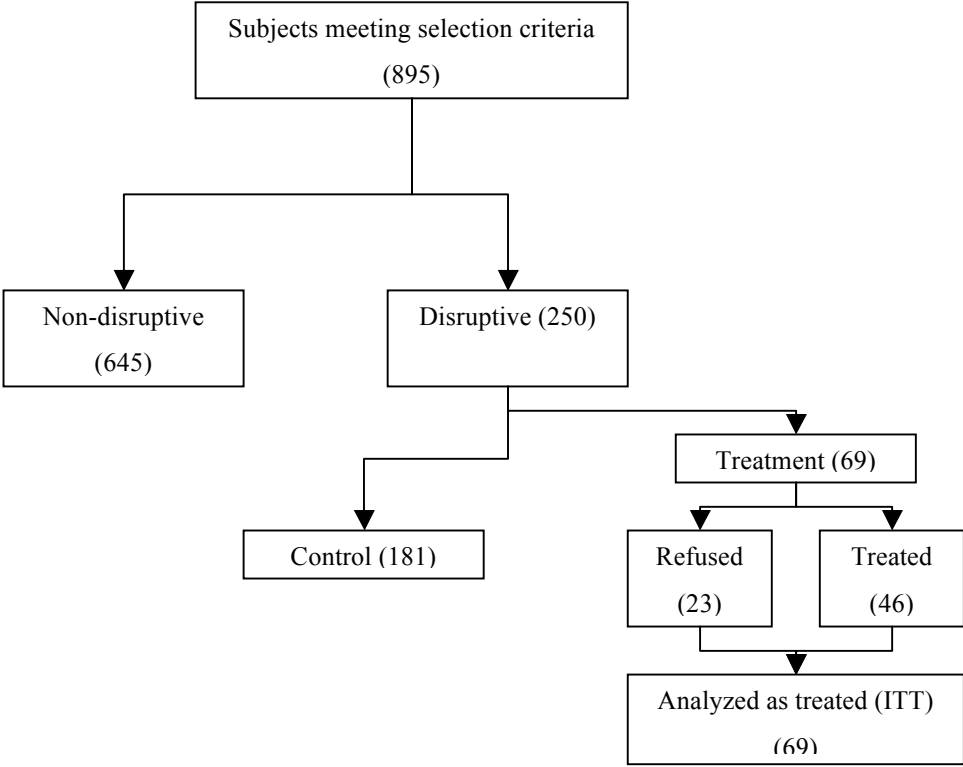


Figure 2. Timeline of the Experimental Study

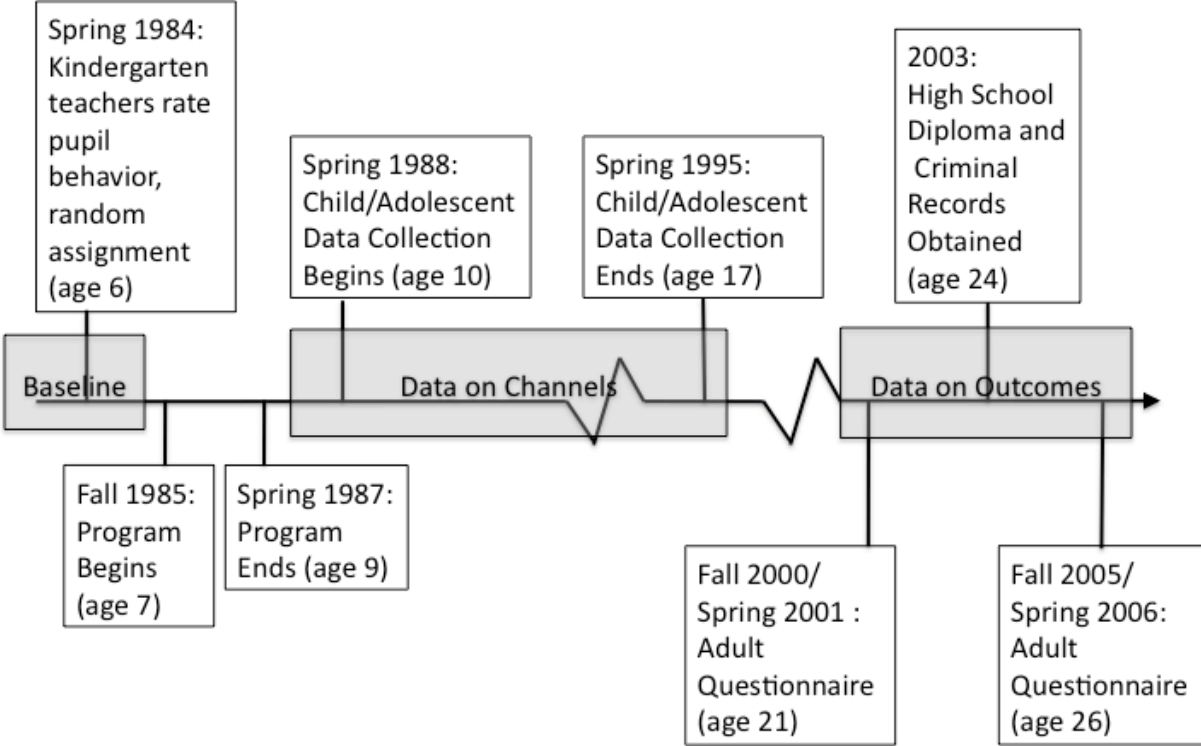
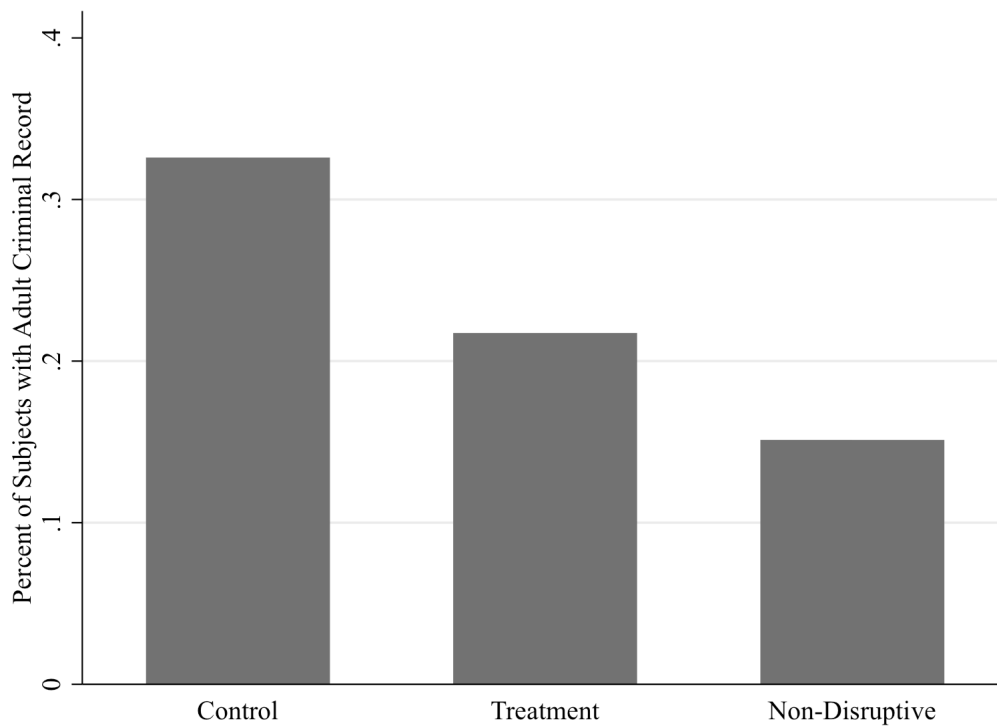
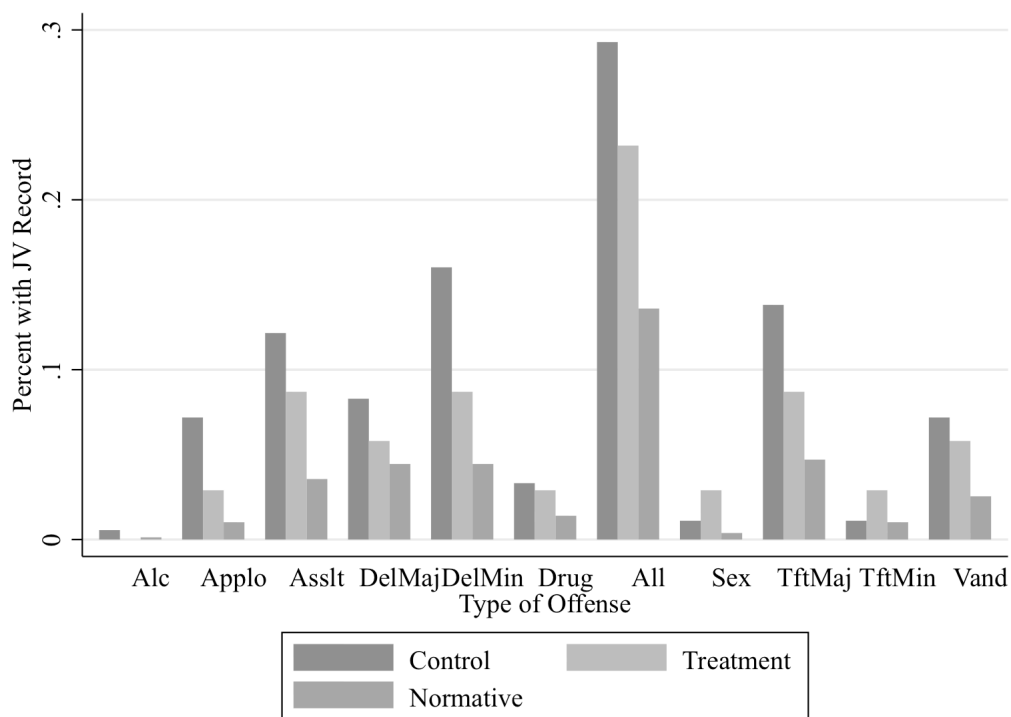


Figure 3. Percent of subjects with adult criminal record



Source: Administrative data. P-value of difference between treatment and control: 0.09.

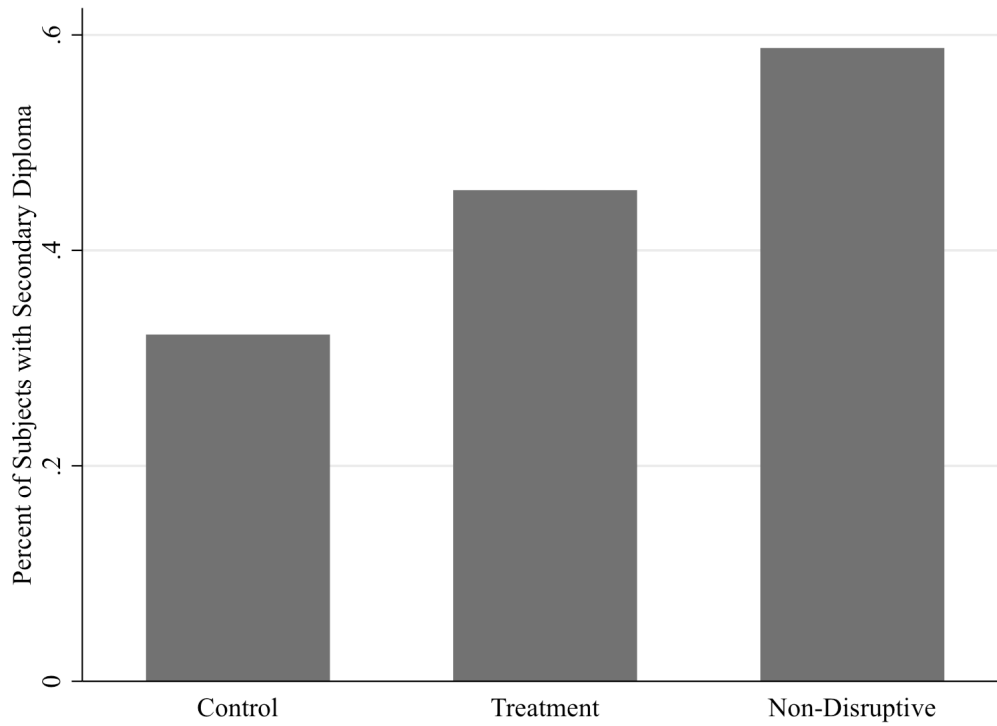
Figure 4: Juvenile Crime



§

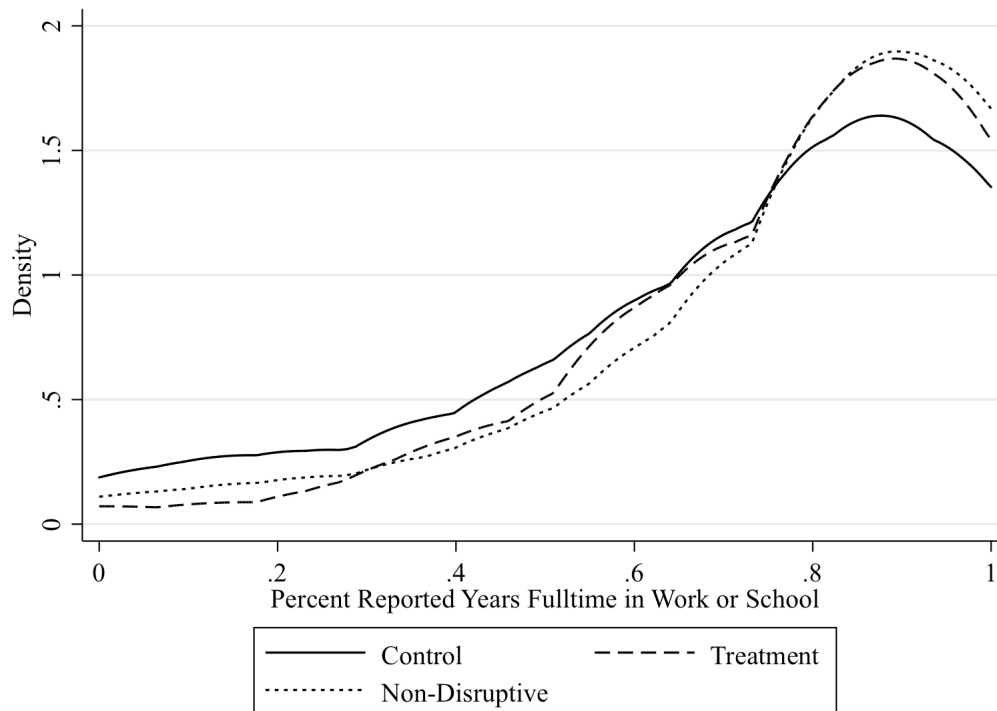
Source: Administrative Data. Difference between treatment and control significant ($p < 0.10$) with controls only for Delinquency (minor).

Figure 5. Rate of obtaining a Secondary School Diploma



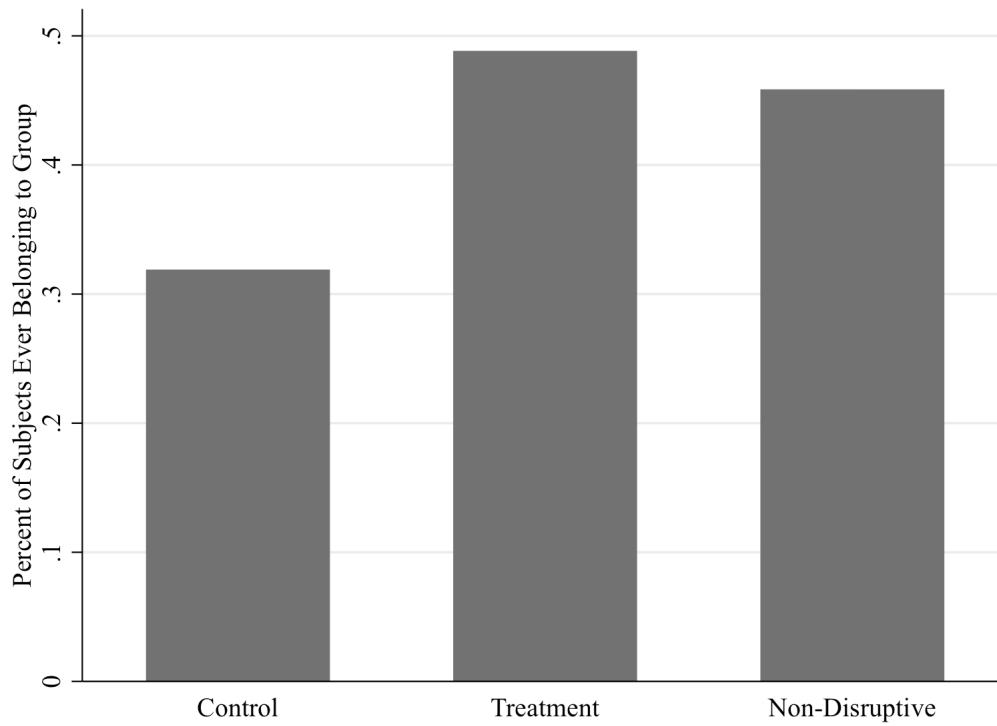
Source: Administrative data. . P-value of difference between treatment and control: 0.05.

Figure 6. Percent of reported years from age 17-26 occupied full time by school or working



Source: Self-reported, 2001 and 2006. . P-value of difference between treatment and control: 0.15.

Figure 7. Percent of subjects reporting membership in any group in 2001 or 2006.



Source: Self-reported, 2001 and 2006. P-value of difference between treatment and control:0.05.

Figure 8. Percent of subjects held back and in special education by year

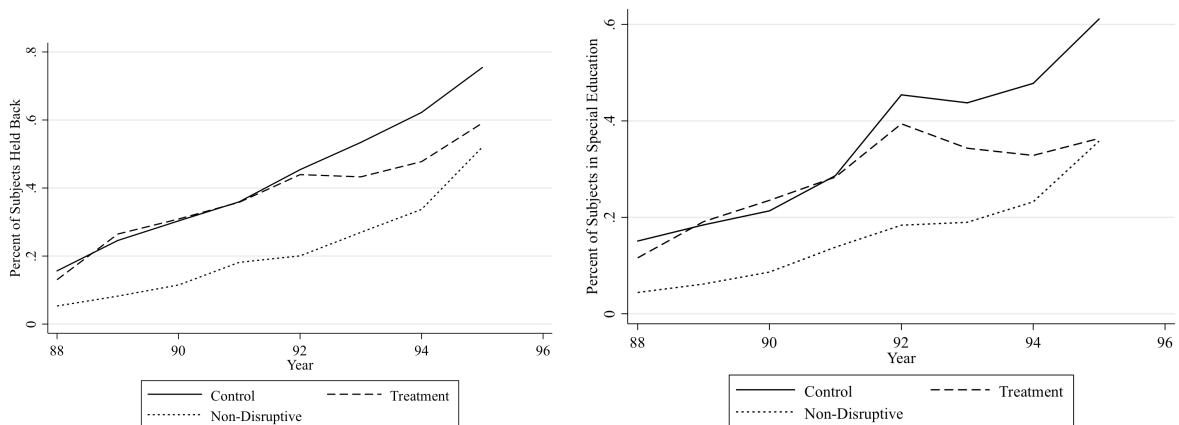


Figure 9. Self Control (All)

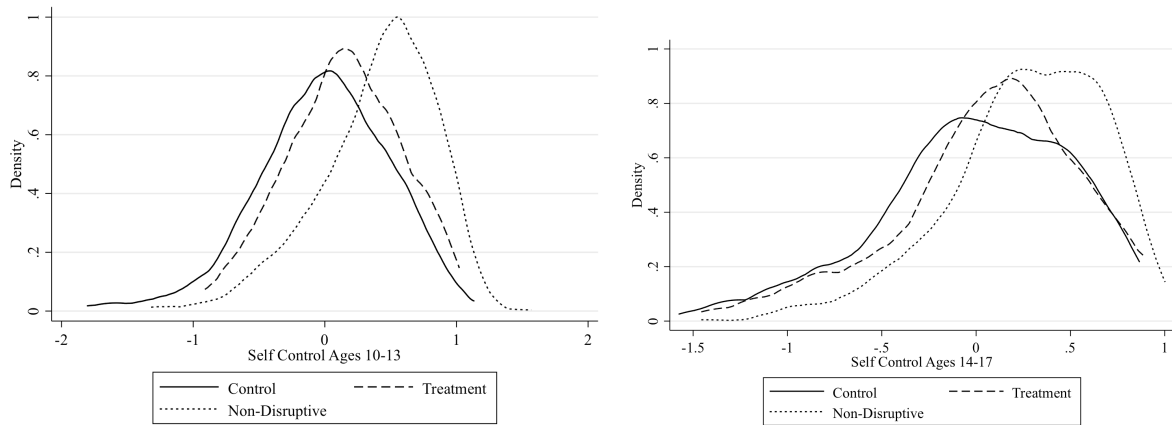


Figure 10. Self Control (External)

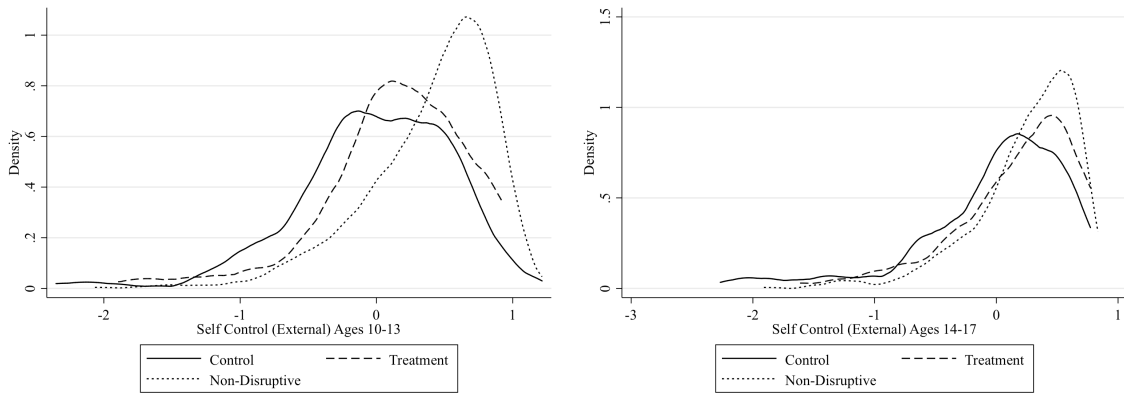


Figure 11. Self Control (Internal)

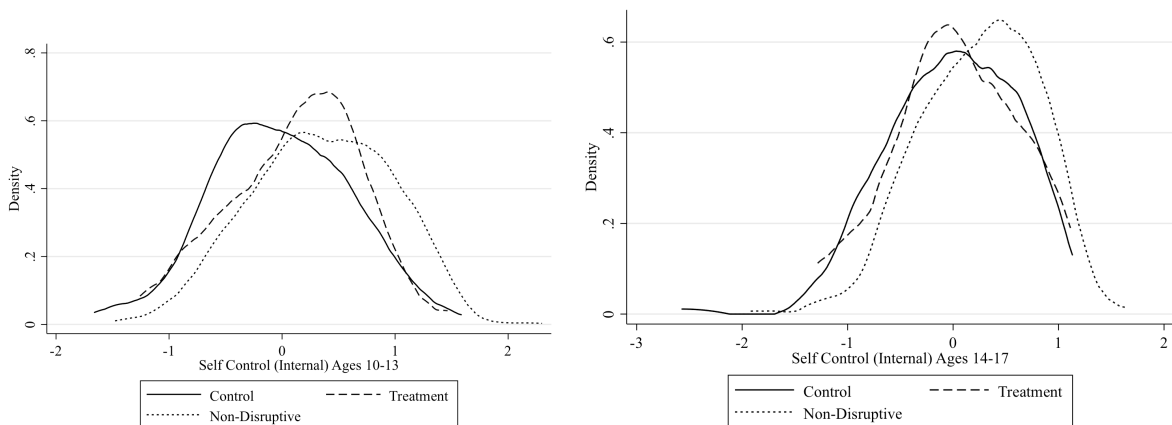
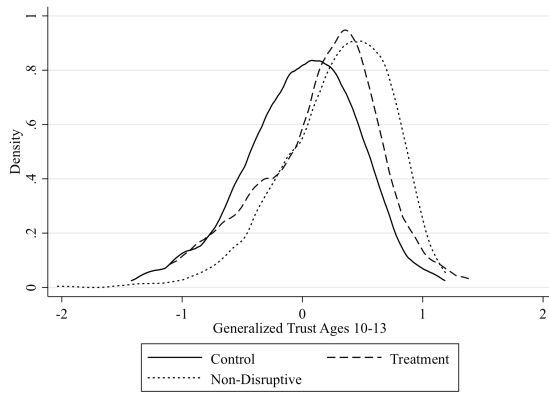
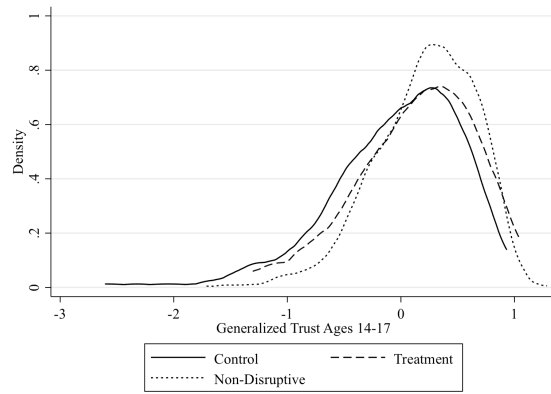


Figure 12. Generalized Trust

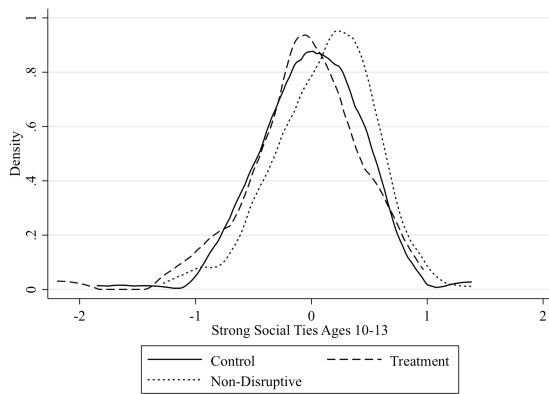


P-value of KS test for treatment vs. control: 0.00

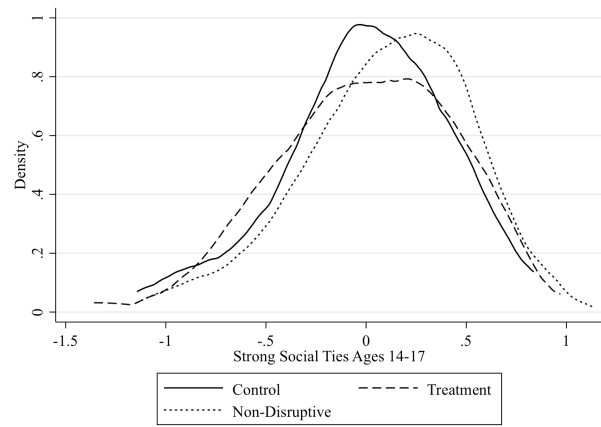


P-value of KS test for treatment vs. control: 0.22

Figure 13: Strong Social Ties

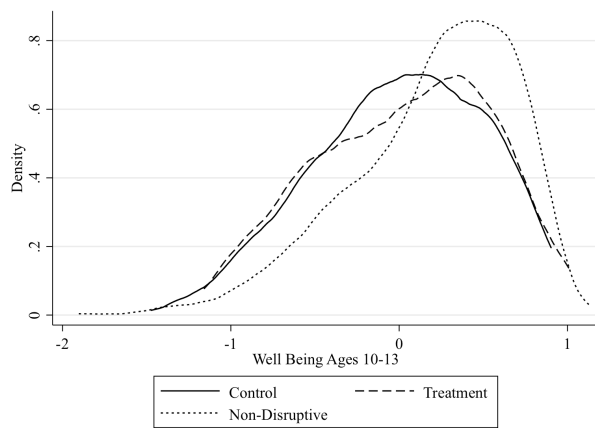


P-value of KS test for treatment vs. control: 0.55

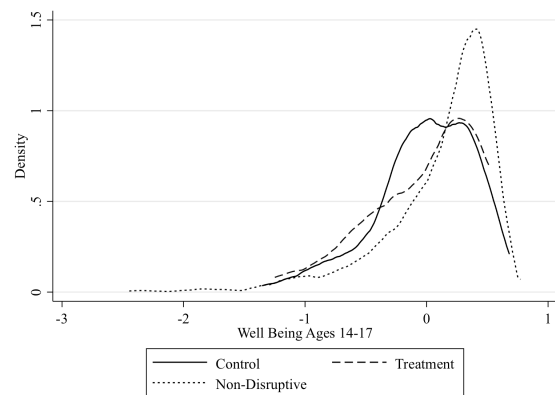


P-value of KS test for treatment vs. control: 0.96

Figure 14. Well-being

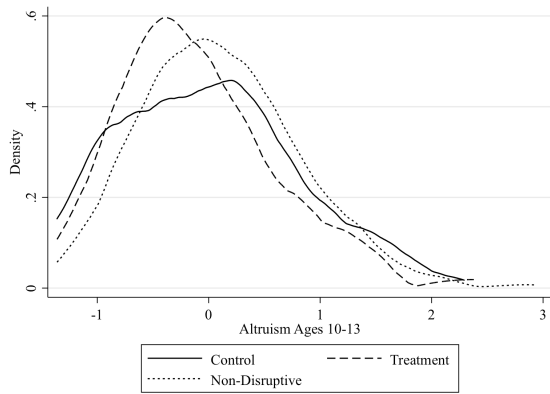


P-value of KS test for treatment vs. control: 0.82

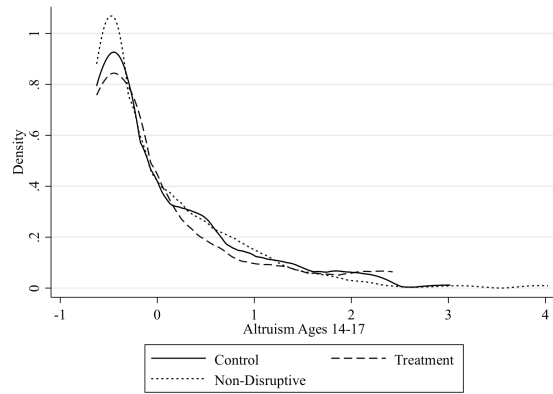


P-value of KS test for treatment vs. control: 0.39

Figure 15. Altruism

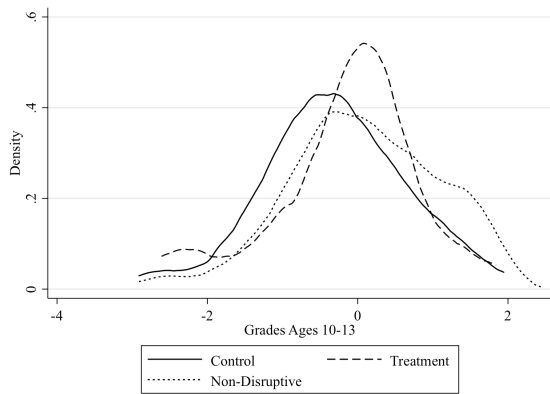


P-value of KS test for treatment vs. control: 0.27

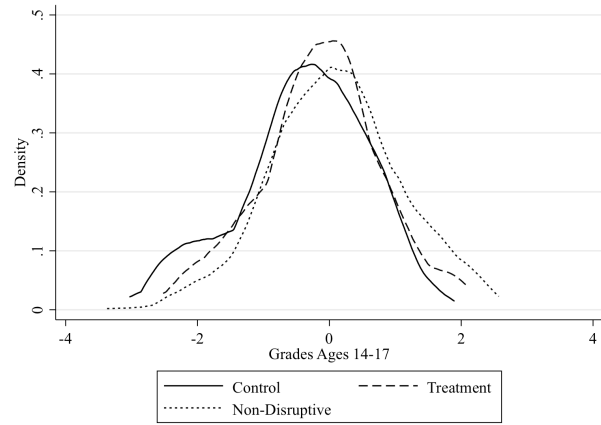


P-value of KS test for treatment vs. control: 0.89

Figure 16. Grades



P-value of KS test for treatment vs. control: 0.01



P-value of KS test for treatment vs. control: 0.09

Table 1. Outcome Variable Summary Statistics

	Non-disruptive population		Disruptive				
	mean	n	Treatment		Control		
			mean	sd	n	mean	sd
Has a criminal record at age 23	0.15	69	0.22	0.42	181	0.33	0.47
Completed Secondary School	0.59	68	0.46	0.50	174	0.32	0.47
% years 17-26 occupied fulltime	0.8	39	0.81	0.23	113	0.73	0.27
Ever member of a group	0.46	43	0.49	0.51	116	0.32	0.47

Education data from the Ministry of Education. Crime data from the Ministry of Public Security of Quebec. Economic and Social outcomes from adult questionnaires in 2001 and 2006.

Table 2. Summary Statistics on Juvenile Crime

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Non-disruptive mean	Non-disruptive N	Non-disruptive sd	Control mean	Control N	Control sd	Treatment mean	Treatment N	Treatment sd	p-value (T-C) t-test	p-value (T-C) OLS
Number of assaults	0.08	787	0.49	0.33	181	1.39	0.30	69	1.42	0.91	0.88
Assault (0/1)	0.04	787	0.19	0.12	181	0.33	0.09	69	0.28	0.44	0.34
Number of vandalism	0.05	787	0.36	0.20	181	0.89	0.16	69	0.78	0.47	0.74
Vandalism (0/1)	0.03	787	0.16	0.07	181	0.26	0.06	69	0.24	0.70	0.61
Number of minor theft	0.01	787	0.13	0.02	181	0.24	0.07	69	0.50	0.28	0.40
Minor theft (0/1)	0.01	787	0.10	0.01	181	0.11	0.03	69	0.17	0.31	0.42
Number of major theft	0.21	787	1.63	0.93	181	3.51	0.32	69	1.22	0.16	0.04
Major theft (0/1)	0.05	787	0.21	0.14	181	0.35	0.09	69	0.28	0.27	0.20
Number of major delinq	0.09	787	0.47	0.22	181	0.95	0.17	69	1.00	0.73	0.68
Major delinq (0/1)	0.04	787	0.21	0.08	181	0.28	0.06	69	0.24	0.51	0.39
Number of minor deliq	0.16	787	1.49	0.63	181	2.14	0.19	69	0.90	0.10	0.04
Minor delinq (0/1)	0.04	787	0.21	0.16	181	0.37	0.09	69	0.28	0.14	0.09
Number of sex crimes	0.01	787	0.13	0.02	181	0.17	0.04	69	0.27	0.34	0.42
Sex crimes (0/1)	0.00	787	0.06	0.01	181	0.11	0.03	69	0.17	0.31	0.39
Number of Apploi	0.03	787	0.37	0.19	181	0.86	0.20	69	1.29	0.92	0.98
Apploi (0/1)	0.01	787	0.10	0.07	181	0.26	0.03	69	0.17	0.20	0.13
Number of Alcohol	0.00	787	0.07	0.01	181	0.15	0.00	69	0.00	0.54	0.27
Alcohol (0/1)	0.00	787	0.04	0.01	181	0.07	0.00	69	0.00	0.54	0.27
Number of Drug	0.04	787	0.36	0.07	181	0.49	0.15	69	0.86	0.37	0.48
Drug (0/1)	0.01	787	0.12	0.03	181	0.18	0.03	69	0.17	0.88	0.88
Total JV Crimes	2.35	787	11.38	6.37	181	17.60	5.10	69	16.25	0.61	0.65
JV crimes (0/1)	0.14	787	0.34	0.29	181	0.46	0.23	69	0.43	0.34	0.27

p-value for OLS taken from regression including controls for imbalance in treatment groups.

Table 3. Baseline Characteristics and Randomization Check

	Non-disruptive population mean	Disruptive						Difference	
		Control			Treatment			C-T	p-value
		mean	N	sd	mean	N	sd		
Age	6.00	6.03	181	0.30	5.97	69	0.29	0.05	0.20
Attended Pre-school	0.16	0.21	181	0.41	0.19	69	0.40	0.02	0.71
Birth order	1.66	1.56	181	0.82	1.50	68	0.72	0.06	0.61
Number brothers	0.56	0.5	180	0.69	0.47	68	0.68	0.03	0.76
Number sisters	0.53	0.33	180	0.52	0.57	68	0.68	-0.24	0.00
Number half brothers	0.04	0.03	181	0.21	0.00	68	0	0.03	0.19
Number half sisters	0.04	0.04	181	0.32	0.02	68	0.12	0.02	0.55
Lives with both parents	0.59	0.53	181	0.50	0.46	69	0.50	0.06	0.39
Age of mother	25.69	23.99	180	4.18	24.01	68	4.71	-0.02	0.97
Age of father	28.67	26.90	161	5.34	28.28	56	5.33	-1.38	0.10
Mother education	10.67	9.97	180	2.23	9.90	68	2.28	0.07	0.83
Father education	10.81	9.70	160	2.45	9.93	60	2.42	-0.24	0.52
Number of children in HH	1.14	0.97	181	0.90	1.07	68	0.80	-0.10	0.42
Adversity index	0.30	0.43	181	0.24	0.43	68	0.27	-0.00	0.96
Mother works	1.61	1.73	177	0.45	1.78	68	0.42	-0.05	0.42
Mother fulltime	1.70	1.69	48	0.47	1.57	14	0.51	0.12	0.43
Father Works	1.12	1.21	148	0.41	1.20	49	0.41	0.01	0.86
Father fulltime	1.97	1.94	115	0.24	1.92	38	0.27	0.02	0.70
Mother prestige	39.35	36.03	161	11.02	33.16	60	10.13	2.87	0.08
Father prestige	40.74	35.19	156	9.58	35.22	53	9.83	-0.03	0.99
Initial Aggression	4.00	14.51	181	4.78	14.62	69	4.58	-0.11	0.86
Initial Anxiety	2.65	3.55	181	2.73	4.26	69	2.82	-0.71	0.07
Initial Opposition	1.63	5.62	181	2.19	5.81	69	1.93	-0.19	0.53
Initial Prosociality	8.21	6.52	181	4.79	6.99	69	4.51	-0.47	0.49
Initial Combativeness	0.82	3.53	181	1.59	3.48	69	1.54	0.05	0.83
Initial Inattention	2.23	4.19	181	2.35	4.19	69	2.18	0.01	0.99
Initial Hyperactivity	0.98	2.79	180	1.21	2.96	68	1.19	-0.16	0.35
Initial Antisociality	0.84	0.99	181	1.11	1.21	68	1.23	-0.21	0.20

Data from baseline data collection, 1984.

Table 4: Impact on Criminal Record

	(1) Criminal Record OLS	(2) Criminal Record Logit	(3) Criminal Record OLS Cluster	(4) Criminal Record OLS Fixed
Treatment	-0.114* (0.0615)	-0.620* (0.351)	-0.114 (0.0746)	-0.0953 (0.0605)
Sample	0.161*** (0.0380)	0.939*** (0.201)	0.161*** (0.0373)	0.164*** (0.0380)
Constant	0.424*** (0.0762)	0.325 (0.569)	0.424*** (0.0932)	0.455*** (0.0801)
Observations	1,037	1,037	1,037	1,037
R-squared	0.054		0.054	0.074

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The dependent variable is whether or not subjects have a criminal record at age 23.

Column (1) presents simple OLS results. Column (2) presents the results of the Logit specification. Column (3) presents results from the OLS specification when standard errors are clustered at the school (1984) level. Column (4) presents results from the OLS specification when fixed effects for school in 1984 are included.

All specifications control for group imbalances.

Table 5: Impact on Secondary Completion

	(1) Secondary Completion OLS	(2) Secondary Completion Logit	(3) Secondary Completion OLS Cluster	(4) Secondary Completion OLS Fixed
Treatment	0.189*** (0.0682)	0.864*** (0.313)	0.189** (0.0763)	0.188*** (0.0682)
Sample	-0.231*** (0.0396)	-1.045*** (0.190)	-0.231*** (0.0379)	-0.241*** (0.0404)
Constant	0.237** (0.0946)	-1.271*** (0.446)	0.237** (0.104)	0.251*** (0.0972)
Observations	1,016	1,016	1,016	1,016
R-squared	0.139		0.139	0.156

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The dependent variable is whether or not subjects completed secondary school.

Column (1) presents simple OLS results. Column (2) presents the results of the Logit specification. Column (3) presents results from the OLS specification when standard errors are clustered at the school (1984) level. Column (4) presents results from the OLS specification when fixed effects for school in 1984 are included.

All specifications control for group imbalances.

Table 6: Impact on Fulltime Occupation

	(1) Occupied Fulltime	(2) Occupied Fulltime Cluster	(3) Occupied Fulltime Fixed
Treatment	0.118** (0.0462)	0.118** (0.0477)	0.121*** (0.0465)
Sample	-0.0131 (0.0314)	-0.0131 (0.0283)	-0.00922 (0.0311)
Constant	0.973*** (0.0692)	0.973*** (0.0581)	0.995*** (0.0692)
Observations	743	743	743
R-squared	0.075	0.075	0.099

Robust standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The dependent variable the percent of reported years from age 17-26 that the subject reported fulltime occupation in either work or school.

Column (1) presents simple OLS results. Column (2) presents results from the OLS specification when standard errors are clustered at the school (1984) level. Column (3) presents results from the OLS specification when fixed effects for school in 1984 are included.

All specifications control for group imbalances and variables related to attrition from the adult questionnaires.

Table 7. Additional Economic Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7) Percent of years receiving public transfers, age 17-26
	Active in work or school age 21	Active full time age 21	Employed age 26	Employed fulltime age 26	Hourly wage age 26	Percent of years working age 17-26	
Treatment	0.0923 (0.185)	0.233 (0.167)	0.348** (0.158)	0.418*** (0.159)	0.395** (0.192)	0.384** (0.154)	-0.229 (0.151)
Sample	-0.0638 (0.112)	-0.0148 (0.112)	-0.0641 (0.157)	-0.0322 (0.159)	-0.0594 (0.154)	-0.0217 (0.124)	0.142 (0.112)
Constant	0.558** (0.241)	0.477* (0.274)	0.0279 (0.293)	0.156 (0.347)	0.556 (0.361)	1.131*** (0.275)	-0.473** (0.208)
Observations	670	670	535	535	535	743	743
R-squared	0.068	0.050	0.049	0.026	0.109	0.047	0.120

Robust standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 8: Impact on Social Group Membership

	(1) Group Member OLS	(2) Group Member Logit	(3) Group Member OLS Cluster	(4) Group Member OLS Fixed
Treatment	0.216** (0.0907)	0.927** (0.387)	0.216*** (0.0757)	0.190** (0.0912)
Sample	-0.109* (0.0566)	-0.473* (0.250)	-0.109** (0.0476)	-0.105* (0.0575)
Constant	0.503*** (0.149)	0.0234 (0.619)	0.503*** (0.148)	0.486*** (0.153)
Observations	750	750	750	750
R-squared	0.039		0.039	0.048

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The dependent variable is group membership, a binary variable equal to 1 if the subject reported participating in any kind of civic group at ages 21 or 26.

Column (1) presents simple OLS results. Column (2) presents the results of the Logit specification. Column (3) presents results from the OLS specification when standard errors are clustered at the school (1984) level. Column (4) presents results from the OLS specification when fixed effects for school in 1984 are included.

All specifications control for group imbalances and variables related to attrition from the adult questionnaires.

Table 9. Description of Channels

Ages 10-13										
Channel	Num.	Alpha	Non-disruptive		Control		Treatment		p-values (T-C)	
	Var		mean	N	mean	N	mean	N	t-test	KS test
Self Control	23	0.87	0.40	780	-0.01	180	0.14	68	0.03	0.06
Self Control (Ext.)	13	0.83	0.40	780	-0.01	180	0.14	68	0.05	0.04
Self Control (Int.)	6	0.73	0.35	780	-0.01	180	0.11	68	0.18	0.09
Generalized Trust	12	0.61	0.29	775	0.00	177	0.15	66	0.02	0.00
Strong Social Ties	13	0.64	0.12	780	0.00	180	-0.07	68	0.28	0.55
Well-being	18	0.82	0.20	754	0.00	169	0.03	63	0.68	0.82
Altruism	10	0.92	0.11	778	0.00	180	-0.11	68	0.33	0.27
Grades	2	0.85	0.09	734	-0.31	158	-0.17	63	0.33	0.01
Held Back	1		-0.14	783	0.38	181	0.37	69	0.96	1.00
Special Ed	1		-0.12	783	0.37	181	0.35	69	0.95	1.00

Ages 14-17										
Channel	Num.	Alpha	Non-disruptive		Control		Treatment		p-values (T-C)	
	Var		mean	N	mean	N	mean	N	t-test	KS test
Self Control	23	0.86	0.27	706	0.00	153	0.07	60	0.32	0.25
Self Control (Ext.)	13	0.86	0.25	706	-0.01	153	0.18	60	0.04	0.01
Self Control (Int.)	6	0.67	0.25	704	0.00	151	0.03	59	0.76	0.99
Generalized Trust	12	0.68	0.22	706	-0.04	153	0.14	60	0.04	0.22
Strong Social Ties	13	0.65	0.11	706	0.00	153	0.00	60	0.99	0.96
Well-being	18	0.75	0.12	686	-0.01	144	-0.02	58	0.98	0.39
Altruism	10	0.90	-0.01	687	0.00	143	-0.04	56	0.74	0.89
Grades	2	0.77	0.12	707	-0.36	155	-0.08	60	0.04	0.09
Held Back	1		-0.14	781	0.49	180	0.26	69	0.12	0.17
Special Ed	1		-0.14	779	0.50	179	0.24	69	0.11	0.22

Channels are the z-score averages of the component variables. All p-values are two sided and robust to inclusion of control variables.

Table 10. Treatment Impact on Channels

Period 1 (Ages 10-13)										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Self Control	Self Control (Ext.)	Self Control (Int.)	Gen. Trust	Strong Social Ties	Well being	Altruism	Grades	Held Back	Special Ed
Treatment	0.155** (0.0653)	0.147* (0.0783)	0.154* (0.0870)	0.177** (0.0700)	-0.0203 (0.0664)	0.0636 (0.0745)	-0.120 (0.106)	0.220 (0.145)	-0.133 (0.176)	-0.108 (0.188)
Sample	-0.404*** (0.0399)	-0.408*** (0.0452)	-0.331*** (0.0519)	-0.275*** (0.0397)	-0.0757** (0.0373)	-0.180*** (0.0444)	-0.0763 (0.0656)	-0.328*** (0.0840)	0.477*** (0.0966)	0.425*** (0.0980)
Constant	-0.000463 (0.0928)	0.0256 (0.0952)	-0.112 (0.130)	-0.00197 (0.0905)	0.0418 (0.0866)	-0.0361 (0.106)	-0.0406 (0.140)	-0.585*** (0.199)	-0.168 (0.208)	-0.00431 (0.217)
Obs.	1,028	1,028	1,028	1,018	1,028	986	1,026	955	1,033	1,033
R-squared	0.160	0.131	0.103	0.087	0.090	0.067	0.018	0.099	0.122	0.084

Period 2 (Ages 14-17)										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Self Control	Self Control (Ext.)	Self Control (Int.)	Gen. Trust	Strong Social Ties	Well being	Altruism	Grades	Held Back	Special Ed
Treatment	0.0647 (0.0761)	0.168** (0.0810)	0.0333 (0.0903)	0.173** (0.0812)	0.0345 (0.0653)	0.0185 (0.0664)	-0.0889 (0.124)	0.326** (0.140)	-0.374** (0.146)	-0.397** (0.160)
Sample	-0.267*** (0.0438)	-0.263*** (0.0519)	-0.236*** (0.0552)	-0.253*** (0.0509)	-0.0752** (0.0378)	-0.115*** (0.0399)	0.0175 (0.0717)	-0.423*** (0.0865)	0.546*** (0.0808)	0.574*** (0.0898)
Constant	-0.0243 (0.0937)	-0.0117 (0.0959)	-0.0403 (0.124)	0.0182 (0.0972)	0.109 (0.0922)	0.107 (0.103)	-0.209 (0.190)	-0.673*** (0.204)	0.425** (0.189)	0.150 (0.199)
Obs.	919	919	914	919	919	888	886	920	1,030	1,027
R-squared	0.091	0.084	0.057	0.061	0.051	0.045	0.013	0.116	0.172	0.140

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 11. Coefficients of Channels on Outcomes (Ages 10-13)

	(1)	(2)	(3)	(4)
	Completed Secondary	Criminal Record	Percent of years 17- 26 occupied fulltime	Group Membership
Self Control	0.495*** (0.0246)	-0.272*** (0.0265)	0.164*** (0.0218)	0.108*** (0.0410)
Delinquency	0.324*** (0.0261)	-0.233*** (0.0251)	0.116*** (0.0218)	0.0287 (0.0382)
Concentration	0.410*** (0.0183)	-0.137*** (0.0182)	0.125*** (0.0147)	0.140*** (0.0297)
Generalized Trust	0.348*** (0.0296)	-0.199*** (0.0276)	0.140*** (0.0225)	0.0825** (0.0410)
Strong Social Ties	0.259*** (0.0342)	-0.0535** (0.0269)	0.112*** (0.0254)	0.140*** (0.0440)
Emotional Well Being	0.420*** (0.0262)	-0.133*** (0.0252)	0.122*** (0.0200)	0.134*** (0.0365)
Altruism	0.0115 (0.0215)	-0.0293* (0.0172)	0.0337*** (0.0122)	0.0406 (0.0251)
Grades	0.234*** (0.0120)	-0.0522*** (0.0127)	0.0608*** (0.00995)	0.0690*** (0.0192)
Held Back	-0.212*** (0.0105)	0.0724*** (0.0142)	-0.0787*** (0.0130)	-0.0477** (0.0197)
Special Education	-0.184*** (0.0101)	0.0771*** (0.0147)	-0.0643*** (0.0138)	-0.0378* (0.0203)

Each cell is the coefficient of an independent regression of the channel on the dependent variable with no controls.

Table 12. Coefficients of Channels on Outcomes (Ages 14-17)

	(1)	(2)	(3)	(4)
	Completed Secondary	Criminal Record	Percent of years 17- 26 occupied fulltime	Group Membership
Self Control	0.441*** (0.0302)	-0.291*** -0.0299	0.101*** (0.0265)	0.0582 (0.0448)
Delinquency	0.290*** (0.0316)	-0.263*** -0.0317	0.140*** (0.0255)	0.0319 (0.0423)
Concentration	0.356*** (0.0232)	-0.146*** -0.0224	0.0928*** (0.0167)	0.0637* (0.0336)
Generalized Trust	0.264*** (0.0300)	-0.209*** -0.0294	0.0915*** (0.0247)	0.0935** (0.0396)
Strong Social Ties	0.234*** (0.0384)	-0.0694** -0.0318	0.0874*** (0.0260)	0.125*** (0.0450)
Emotional Well Being	0.281*** (0.0381)	-0.106*** -0.0309	0.0959*** (0.0228)	0.0944** (0.0412)
Altruism	-0.0832*** (0.0221)	0.00797 -0.0183	-0.0126 (0.0138)	-0.0304 (0.0249)
Grades	0.270*** (0.0104)	-0.0822*** -0.0124	0.0676*** (0.00954)	0.0399** (0.0196)
Held Back	-0.349*** (0.00961)	0.109*** -0.0126	-0.0894*** (0.0106)	-0.0929*** (0.0185)
Special Education	-0.268*** (0.0100)	0.114*** -0.0138	-0.0912*** (0.0123)	-0.0747*** (0.0191)

Each cell is the coefficient of an independent regression of the channel on the dependent variable with no controls.

Table 13. Cost per adult crime averted

	Duration of criminal activity after age 23			
	0 years	5 years	10 years	15 years
Cost per crime averted	5,406 \$	3,374 \$	2,320 \$	1,767 \$

We assume that the rate of criminal activity decreases by 10% each year. The period up until age 23 is six years long. Annual discount rate is 3%.

Table 14. Rate of Return from Increased Hourly Wages

Years of Work Hours per Week	20		30		40	
	20	40	20	40	20	40
Rate of Return in wages	254%	608%	366%	833%	450%	1000%

Wages are self-reported at age 26. For unemployed subjects (who have no wage information) wage is imputed using baseline characteristics, schooling, and a replacement rate of 0.63. Annual discount rate is 3%.

APPENDIX B: SUPPLEMENTARY INFORMATION ON THE INTERVENTION PROGRAM

The activity program is covered in more detail in Bertrand and Reclus-Prince (1988), from which this section draws heavily.¹⁰ The training sessions for the children lasted around 45 minutes and were conducted as follows.

1. The previous session's activity would be reviewed, the group would talk about whether they had used the previous session's behavior, and any homework would be discussed. Children would be praised for having tried to use the skill, and the group would try to identify the benefits of that behavior.
2. The activity for the session would be introduced with a brief explanation from the facilitator, who would ask children to talk about situation in life that are relevant to the behavior. The facilitator would highlight the results of the behavior. The behavior would be modeled, generally through role playing, where the children and/or the teacher would take different parts. Several role plays might be carried out to demonstrate the behavior.
3. The facilitator would guide a practice session, where children would take turns acting in different role plays that demonstrate the behavior, and would alternate roles. During and after the role plays, the facilitator would give feedback and articulate how the behavior was being demonstrated.
4. The facilitator would generalize the behavior by talking about its benefits and how it can make them (the children) feel better and make other people feel better. The facilitator might assign homework (for example, making a drawing of the behavior being discussed). The facilitator would try to work with children to anticipate situations that would call for the behavior, and plan to respond appropriately in those situations. The facilitator would also try to communicate the session's topic to the parents and teachers, who could reinforce that topic at home and in the classroom.
5. The facilitator would also work to reinforce the learned behaviors during the sessions, for example by using a scoring system such as the one below reported in Figure B.1.

¹⁰ Bertrand and Reclus-Prince (1988) also provide references used in development of the program: Bessell and Palomares (1972), Cartledge and Milburn (1980), Schneider and Byrne (1984), Michelson et al (1983) and Shure and Spivack (1984).

Figure B.1: Example of activity program

Behaviour	Duration of the activity			
	First 10-min. period	Second 10-min. period	Third 10-min. period	Fourth 10-min. period
Looking at the person who is talking (1 point)				
Staying seated (1 point)				
Waiting your turn to talk (1 point)				
TOTAL				
Criteria: <ul style="list-style-type: none"> - Individual performance: 10 points - Choice of rewards: sticker, eraser, felt marker, etc. - Group performance (5 children): 50 points - Choice of rewards: group game, free time, etc. 				

B.1) First Year: Social Skills

The trainings during first year focused on social skills, and in the second year self-control (described below). Nine behaviors were covered in the social skills training sessions:

- Making Contact (eye contact, smiling, approaching another person)
- Speaking Nicely (friendly body language, tone of voice, explaining what we like)
- Gentle Physical Contact (body language, touching with friendly intent, briefly)
- Helping (approaching someone who needs help, offering to help, waiting for the answer)
- Including and Inviting (body language, saying the person’s name, inviting)
- Doing things together (body language, proposing how to do things, seeing whether the other person agrees)
- Saying No (body language, tone of voice, without anger, giving a reason)
- Asking Why (body language, letting the other person finish speaking, asking the question)
- Saying “You’re bothering me” (body language, saying the person’s name, saying what’s wrong, asking the person to stop)

For example, in the “Speaking Nicely” session, the facilitator would begin by asking children if they tried the previous session’s behavior (Making Contact), and the results. The facilitator would then explain what “Speaking Nicely” means and why it is useful, and the children would try to give examples of speaking nicely. The role play for modeling would involve a friend who arrives at school wearing a new sweater, and the way to speak nicely in this situation is to go up to the friend and make eye contact, smile, use a pleasant tone of voice, and explain that you like his sweater. Other role plays would include congratulating a friend who got a good grade, encouraging a partner who made a mistake in a game, or thanking parents for doing something nice for. After several role plays, the facilitator would review how important words and body language are to being understood, and assign homework and give out letters for parents and teachers.

In the “Saying No” session, the facilitator would ask if the children had practiced the previous session’s behavior (Doing Things Together), and discuss how it went. Then the facilitator would introduce the topic of “Saying No”, and explain that we are allowed to say no if we don’t want to do something, and that we can do so clearly and without getting angry. The children would then discuss times that they had said no, or times that they wanted to say no but did not feel comfortable doing so. The first role play is about a friend who wants to borrow a bicycle. To say no, the children are told to look at their friend, speak firmly and without anger, and to give a reason for saying no. Other role plays are when a child is not interested in playing but wants to be alone, when a sibling wants to share candy or wants to play. The facilitator reviews that it is good to say honestly what we want or don’t want, and clearly explain our reasons, so that the other person can better understand why we are saying no, and accept our answer. Then homework is assigned and letters to parents and teachers are distributed.

Figure B.2: Example of Take-home sheet for parents from Social Skills trainings

To the parents of _____

Today and yesterday we worked on

acceptable ways of saying “no”

To help your child practise this method, we have given him a drawing.
The drawing is to remind him to practise today's method at home.

YOU CAN HELP HIM PRACTISE

- ⇒ by talking to him about what he did with us
- ⇒ by helping him practise the method when the opportunity arises.

YOU CAN HELP HIM USE THE METHOD MORE OFTEN

- ⇒ by telling him that it's nice to see someone look at the person he is talking to
- ⇒ by telling him that it makes you happy when he says “no” calmly to his brother or sister
- ⇒ by telling him that you appreciate it when he politely explains to a friend why he is refusing something.

We greatly appreciate your child's participation in our program.

B.2) *Second Year: Self-Control*

In the second year, the sessions focused on the theme of self-control. The principal was to provide a roadmap for children to decide on an appropriate response in a given situation. Four steps are reinforced over the ten sessions:

1. Identify what is happening and feelings associated with it
2. Think about possibilities to say and do
3. Choose an appropriate solution
4. Do the solution, and feel good about it

These steps are reinforced in each session, where the child would say the steps aloud and add the activity-specific details. The specific topics of the ten sessions were

- Paying attention and listening
- Obtaining information before acting
- Following rules and instructions
- Controlling myself
- What to do when angry
- What to do when feeling left out
- What to do when feeling like hitting
- What to do when teased
- Asking forgiveness
- Expressing appreciation

For example, in the unit “I feel like hitting... what do I do?” the facilitator would begin by asking the children if they had use the skills used in the previous session (Controlling Myself). Then the facilitator would introduce the topic of hitting, talking about all the times when one feels like hitting or pushing, and why. The facilitator would explain the appropriate behavior, that is when one wants to hit, one must stop, pause, tell oneself “I won’t hit, calm down” and choose another way to act that will not hurt the other person. Children would be asked to give examples from their own life. The first role played, by the facilitator, would involve someone who is bumped into and hurt by another student who was rushing along. Following the steps above, the children are told to

1. Identify what is happening and feelings associated with it: He bumped into me. Was it on purpose? I feel angry, and I want to hit him.

2. Think about possibilities to say and do: I could hit him, I could fall down on the floor and cry, I could go tell my teacher, I could yell at him. I think that usually that classmate is not mean to me. I could do or think about something else. I could move away from him.
3. Choose an appropriate solution: I will stop, tell myself that I won't hit him, and choose another way of dealing with the situation. I will tell him clearly but without being angry how I feel.
4. Do the solution, and feel good about it: I tell him how I feel, that he hurt me, and that I want him to be more careful next time. I tell myself, good job for not getting into a fight! I am proud of myself.

Additional role plays are performed with the children involve, all demonstrating the four steps (identify, think, choose, do), and the principle of stop, pause, say "I won't hit" and choose another action. Other role plays include someone breaking a toy, aggression from others, teasing, or having a foot stepped on accidentally. The facilitator then gives feedback, reinforcing the four steps and the principal of stop. Then the facilitator helps each child prepare their homework, which involves identifying opportunities to practice the behavior of not hitting when you want to.

Figure B.3: Example of take home sheet for self-control training

Activity No. 3

TO FOLLOW RULES AND INSTRUCTIONS

- **I listen**
- **I check whether I have understood**
- **I do what is asked**

I practise the activity by

- **Looking at the person who is speaking**
- **Speaking when it is my turn to speak**
- **Staying seated and remaining calm**
- **Eating without making a mess**

I will practise again tonight at suppertime

My signature _____

Parent's signature _____

Facilitator's comment:

APPENDIX C: ADDITIONAL TABLES FOR CHANNEL COMPOSITION

Table 15. Correlation between channels

	Ages 10-13									
	Self Control	Self Control (Ext.)	Self Control (Int.)	Gen. Trust	Strong Social Ties	Well-being	Altruism	Grades	Held Back	Special Ed
Self Control	1.000									
Self Control (Ext.)	0.908	1.000								
Self Control (Int.)	0.739	0.419	1.000							
Generalized Trust	0.655	0.644	0.420	1.000						
Strong Social Ties	0.180	0.067	0.296	0.265	1.000					
Wellbeing	0.527	0.337	0.664	0.429	0.330	1.000				
Altruism	0.167	0.144	0.119	0.186	0.126	0.028	1.000			
Grades	0.427	0.228	0.585	0.301	0.232	0.433	0.161	1.000		
Held Back	-0.327	-0.241	-0.327	-0.269	-0.258	-0.352	0.053	-0.266	1.000	
Special Ed	-0.326	-0.255	-0.295	-0.253	-0.247	-0.291	0.039	-0.243	0.890	1.000

	Ages 14-17									
	Self Control	Self Control (Ext.)	Self Control (Int.)	Gen. Trust	Strong Social Ties	Well-being	Altruism	Grades	Held Back	Special Ed
Self Control	1.000									
Self Control (Ext.)	0.881	1.000								
Self Control (Int.)	0.736	0.404	1.000							
Generalized Trust	0.665	0.659	0.412	1.000						
Strong Social Ties	0.096	0.018	0.164	0.105	1.000					
Wellbeing	0.386	0.233	0.497	0.314	0.230	1.000				
Altruism	0.052	0.080	-0.018	0.107	-0.034	-0.084	1.000			
Grades	0.423	0.268	0.475	0.293	0.139	0.227	0.046	1.000		
Held Back	-0.393	-0.265	-0.402	-0.282	-0.286	-0.328	0.205	-0.487	1.000	
Special Ed	-0.378	-0.269	-0.341	-0.295	-0.284	-0.326	0.220	-0.364	0.851	1.000

Table 16. Components of Self Control Channel

		Ages 10-13					
Variable	Source	Non-disruptive		Control		Treatment	
		mean	N	mean	N	mean	N
Fights	Subject	-0.09	769	0.29	173	0.33	66
Mistreats others	Subject	-0.02	712	0.12	157	-0.12	63
Steals	Subject	-0.04	769	0.24	173	-0.20	66
Difficult to apply self at school	Subject	-0.03	764	0.07	171	0.20	66
Tries hard at school	Subject	-0.07	764	0.27	171	0.08	66
Force others to do things	Subject	-0.05	764	0.20	171	0.02	66
Damage other's property	Subject	-0.08	727	0.25	165	0.27	64
Impolite to Teacher	Subject	-0.12	764	0.43	171	0.26	66
Disturbs Class	Subject	-0.13	764	0.45	171	0.30	66
Vandalizes School	Subject	-0.06	768	0.31	173	-0.08	66
Make people angry	Subject	-0.06	754	0.27	169	-0.03	63
Hard to concentrate	Subject	-0.04	754	0.17	169	-0.02	63
React strongly to insult	Subject	-0.07	754	0.31	169	-0.02	63
Hard to find things to do	Subject	-0.03	754	0.07	169	0.17	63
Crazy if provoked	Subject	-0.06	754	0.25	169	0.04	63
Do things I know I shouldn't	Subject	-0.04	754	0.21	169	-0.04	63
Destroys own things	Teacher	-0.15	778	0.51	180	0.39	68
Fights	Teacher	-0.17	778	0.60	180	0.39	68
Disobedient	Teacher	-0.20	778	0.69	180	0.45	68
Weak concentration	Teacher	-0.14	778	0.48	180	0.32	68
Bullies	Teacher	-0.19	778	0.68	180	0.43	68
Easily distracted	Teacher	-0.13	778	0.44	180	0.30	68
Hits/bites/kicks	Teacher	-0.17	778	0.57	180	0.40	68
Jumps from one thing to another	Teacher	-0.12	709	0.49	160	0.12	63

		Ages 14-17					
Variable	Source	Non-disruptive		Control		Treatment	
		mean	N	mean	N	mean	N
Fights	Subject	-0.07	700	0.28	150	0.15	59
Mistreats others	Subject	-0.06	700	0.26	150	0.00	59
Steals	Subject	-0.05	700	0.27	149	-0.07	59
Difficult to apply self at school	Subject	-0.01	698	0.05	147	-0.06	57
Tries hard at school	Subject	-0.05	698	0.21	147	0.02	58
Force others to do things	Subject	-0.07	700	0.29	150	0.14	59
Damage other's property	Subject	-0.06	698	0.27	147	0.04	59
Impolite to Teacher	Subject	-0.09	698	0.36	147	0.15	59
Disturbs Class	Subject	-0.05	698	0.22	147	0.02	58
Vandalizes School	Subject	-0.05	698	0.27	147	-0.10	59
Make people angry	Subject	-0.02	671	0.17	138	-0.13	55
Hard to concentrate	Subject	-0.02	672	0.14	138	-0.09	55
React strongly to insult	Subject	0.01	673	-0.01	138	-0.07	55
Hard to find things to do	Subject	0.00	672	-0.03	137	0.11	55
Crazy if provoked	Subject	-0.05	670	0.20	138	0.06	54
Do things I know I shouldn't	Subject	-0.01	672	0.07	138	-0.03	55
Destroys own things	Teacher	-0.13	687	0.34	143	0.71	56
Fights	Teacher	-0.11	687	0.33	142	0.53	56
Disobedient	Teacher	-0.12	687	0.38	142	0.52	56
Weak concentration	Teacher	-0.11	687	0.35	142	0.45	56
Bullies	Teacher	-0.10	687	0.37	143	0.33	56
Easily distracted	Teacher	-0.12	687	0.40	143	0.46	56
Hits/bites/kicks	Teacher	-0.09	686	0.27	142	0.39	56
Jumps from one thing to another	Teacher	-0.09	687	0.33	142	0.27	56

Table 17. Components of Self Control (External) Channel

Variable	Source	Non-disruptive		Ages 10-13 Control		Treatment	
		mean	N	mean	N	mean	N
Fights	Subject	-0.09	769	0.29	173	0.33	66
Mistreats others	Subject	-0.02	712	0.12	157	-0.12	63
Steals	Subject	-0.04	769	0.24	173	-0.20	66
Force others to do things	Subject	-0.05	764	0.20	171	0.02	66
Damage other's property	Subject	-0.08	727	0.25	165	0.27	64
Impolite to Teacher	Subject	-0.12	764	0.43	171	0.26	66
Disturbs Class	Subject	-0.13	764	0.45	171	0.30	66
Vandalizes School	Subject	-0.06	768	0.31	173	-0.08	66
Make people angry	Subject	-0.06	754	0.27	169	-0.03	63
React strongly to insult	Subject	-0.07	754	0.31	169	-0.02	63
Crazy if provoked	Subject	-0.06	754	0.25	169	0.04	63
Fights	Teacher	-0.17	778	0.60	180	0.39	68
Bullies	Teacher	-0.19	778	0.68	180	0.43	68
Hits/bites/kicks	Teacher	-0.17	778	0.57	180	0.40	68

Variable	Source	Non-disruptive		Ages 14-17 Control		Treatment	
		mean	N	mean	N	mean	N
Fights	Subject	-0.07	700	0.28	150	0.15	59
Mistreats others	Subject	-0.06	700	0.26	150	0.00	59
Steals	Subject	-0.05	700	0.27	149	-0.07	59
Force others to do things	Subject	-0.07	700	0.29	150	0.14	59
Damage other's property	Subject	-0.06	698	0.27	147	0.04	59
Impolite to Teacher	Subject	-0.09	698	0.36	147	0.15	59
Disturbs Class	Subject	-0.05	698	0.22	147	0.02	58
Vandalizes School	Subject	-0.05	698	0.27	147	-0.10	59
Make people angry	Subject	-0.02	671	0.17	138	-0.13	55
React strongly to insult	Subject	0.01	673	-0.01	138	-0.07	55
Crazy if provoked	Subject	-0.05	670	0.20	138	0.06	54
Fights	Teacher	-0.11	687	0.33	142	0.53	56
Bullies	Teacher	-0.10	687	0.37	143	0.33	56
Hits/bites/kicks	Teacher	-0.09	686	0.27	142	0.39	56

Table 18. Components of Self Control (Internal) Channel

Variable	Source	Non-disruptive		Ages 10-13 Control		Treatment	
		mean	N	mean	N	mean	N
Difficult to apply self at school	Subject	-0.03	764	0.07	171	0.20	66
Tries hard at school	Subject	-0.07	764	0.27	171	0.08	66
Hard to concentrate	Subject	-0.04	754	0.17	169	-0.02	63
Hard to find things to do	Subject	-0.03	754	0.07	169	0.17	63
Weak concentration	Teacher	-0.14	778	0.48	180	0.32	68
Easily distracted	Teacher	-0.13	778	0.44	180	0.30	68
Jumps from one thing to another	Teacher	-0.12	709	0.49	160	0.12	63

Variable	Source	Non-disruptive		Ages 14-17 Control		Treatment	
		mean	N	mean	N	mean	N
Difficult to apply self at school	Subject	-0.01	698	0.05	147	-0.06	57
Tries hard at school	Subject	-0.05	698	0.21	147	0.02	58
Hard to concentrate	Subject	-0.02	672	0.14	138	-0.09	55
Hard to find things to do	Subject	0.00	672	-0.03	137	0.11	55
Weak concentration	Teacher	-0.11	687	0.35	142	0.45	56
Easily distracted	Teacher	-0.12	687	0.40	143	0.46	56
Jumps from one thing to another	Teacher	-0.09	687	0.33	142	0.27	56

Table 19: Components of Generalized Trust Channel

Variable	Source	Ages 10-13					
		Non-disruptive		Control		Treatment	
		mean	N	mean	N	Mean	N
Trust Police	Subject	0.04	764	-0.24	171	0.15	66
Thinks a bump is on purpose	Subject	0.05	712	-0.21	157	-0.08	63
Cares about what teacher thinks	Subject	0.01	764	-0.05	170	0.05	66
Angry at a bump	Teacher	0.13	708	-0.50	160	-0.23	63
Inconsiderate	Teacher	0.17	777	-0.59	180	-0.37	68
If police don't like you...	Subject	0.07	754	-0.23	169	-0.20	63
Better to trust no one	Subject	0.06	754	-0.14	169	-0.32	63
People like to play the boss	Subject	0.06	754	-0.23	169	-0.06	63
People say one thing and do another	Subject	0.06	754	-0.26	169	-0.01	63
OK to steal from dishonest store	Subject	0.05	754	-0.20	169	-0.09	63
Tempted to lie when in trouble	Subject	0.05	754	-0.25	169	0.08	63
OK to steal from someone rich	Subject	0.02	753	-0.12	169	0.02	63

Variable	Source	Age 14-17					
		Non-disruptive		Control		Treatment	
		mean	N	mean	N	Mean	N
Trust Police	Subject	0.05	700	-0.30	149	0.10	59
Thinks a bump is on purpose	Subject	0.06	700	-0.22	150	-0.12	59
Cares about what teacher thinks	Subject	0.02	698	-0.08	147	-0.02	59
Angry at a bump	Teacher	0.14	685	-0.49	143	-0.49	56
Inconsiderate	Teacher	0.10	686	-0.34	143	-0.40	56
If police don't like you...	Subject	0.04	672	-0.27	137	0.14	55
Better to trust no one	Subject	0.05	672	-0.27	136	0.05	55
People like to play the boss	Subject	0.00	673	-0.04	138	0.05	54
People say one thing and do another	Subject	0.04	672	-0.21	136	-0.02	54
OK to steal from dishonest store	Subject	0.01	672	-0.05	136	-0.03	54
Tempted to lie when in trouble	Subject	0.00	673	-0.10	138	0.26	55
OK to steal from someone rich	Subject	0.03	673	-0.12	138	-0.11	55

Standardized Values.

Table 20: Components of Strong Social Ties Channel

Variable	Source	All Years					
		Non-disruptive		Control		Treatment	
		mean	N	mean	N	Mean	N
Trusts best friend	Subject	0.06	737	-0.16	166	-0.25	64
Gets advice best friend	Subject	0.02	737	-0.04	166	-0.07	64
Talks pers best friend	Subject	0.01	737	-0.05	166	0.02	64
Talks probs best friend	Subject	0.01	737	0.01	166	-0.10	64
Care if other people like me	Subject	0.01	759	-0.07	171	0.03	64
I never lie	Subject	0.00	759	0.04	171	-0.11	64
Better not to talk to anyone	Subject	0.08	759	-0.25	171	-0.30	64
Time spent with friend	Subject	0.02	772	-0.08	173	-0.04	66
Time talking to friend	Subject	0.02	770	-0.08	172	-0.07	66
Talk about future with parent	Subject	0.02	772	-0.03	173	-0.15	66
Time talking with parent	Subject	0.03	770	-0.12	172	-0.06	66
Tends to work alone	Teacher	0.05	778	-0.16	180	-0.11	68
Not liked by other children	Teacher	0.20	778	-0.65	180	-0.53	68

Variable	Source	Age 10-13					
		Non-disruptive		Control		Treatment	
		mean	N	mean	N	Mean	N
Trusts best friend	Subject	0.06	718	-0.14	162	-0.31	64
Gets advice best friend	Subject	0.02	718	-0.02	162	-0.20	64
Talks pers best friend	Subject	-0.02	718	0.07	162	0.01	64
Talks probs best friend	Subject	-0.01	718	0.10	162	-0.09	64
Care if other people like me	Subject	0.02	754	-0.04	169	-0.07	63
I never lie	Subject	0.00	754	0.04	169	-0.11	63
Better not to talk to anyone	Subject	0.07	754	-0.21	169	-0.23	63
Time spent with friend	Subject	0.03	764	-0.10	171	-0.04	66
Time talking to friend	Subject	0.01	763	0.01	171	-0.13	66
Talk about future with parent	Subject	0.03	764	-0.05	171	-0.16	66
Time talking with parent	Subject	0.02	764	-0.09	171	-0.06	66
Tends to work alone	Teacher	0.04	778	-0.14	180	-0.08	68
Not liked by other children	Teacher	0.18	778	-0.60	180	-0.50	68

Variable	Source	Age 14-17					
		Non-disruptive		Control		Treatment	
		mean	N	mean	N	Mean	N
Trusts best friend	Subject	0.05	700	-0.18	150	-0.15	59
Gets advice best friend	Subject	0.00	700	-0.05	150	0.08	59
Talks pers best friend	Subject	0.02	700	-0.08	150	0.00	59
Talks probs best friend	Subject	0.02	700	-0.03	150	-0.12	59
Care if other people like me	Subject	0.04	686	-0.12	142	-0.21	56
I never lie	Subject	0.01	671	-0.12	138	0.22	55
Better not to talk to anyone	Subject	-0.01	671	0.05	137	-0.06	54
Time spent with friend	Subject	0.07	672	-0.18	138	-0.38	55
Time talking to friend	Subject	0.13	686	-0.46	143	-0.45	56
Talk about future with parent	Subject	0.00	675	-0.04	146	0.10	57
Time talking with parent	Subject	0.00	700	0.03	150	-0.04	59
Tends to work alone	Teacher	0.02	699	-0.06	150	-0.14	59
Not liked by other children	Teacher	0.00	655	0.00	137	-0.06	56

Standardized Values

Table 21: Components of Emotional Well-Being Channel

		All Years					
Variable	Source	Non-disruptive		Control		Treatment	
		mean	N	mean	N	Mean	N
Other people do things more easily than I	Subject	0.04	759	-0.14	171	-0.06	64
People talk about me behind my back	Subject	0.06	759	-0.18	171	-0.25	64
No one understands what I feel	Subject	0.06	759	-0.20	171	-0.12	64
I feel worried	Subject	0.03	759	-0.10	171	-0.12	64
Have trouble deciding	Subject	0.01	759	-0.07	171	0.05	64
Feel alone even when I'm with someone	Subject	0.08	759	-0.18	171	-0.42	64
Something doesn't work	Subject	0.07	759	-0.25	171	-0.16	64
Satisfied with myself	Subject	0.03	727	-0.05	163	-0.25	63
I'm not worth anything	Subject	0.03	727	-0.09	163	-0.04	63
I have good qualities	Subject	0.06	727	-0.21	163	-0.16	63
I do things well	Subject	0.02	727	-0.02	162	-0.14	63
Not much to be proud of	Subject	0.06	727	-0.19	163	-0.14	63
I feel useless	Subject	0.02	727	-0.11	163	0.00	63
I'm a valuable person	Subject	0.06	727	-0.18	163	-0.24	63
I should have more self-respect	Subject	0.04	727	-0.12	162	-0.14	63
I feel like a failure	Subject	0.06	726	-0.20	163	-0.18	63
I have a positive attitude to myself	Subject	0.06	726	-0.23	162	-0.12	63
Gives up easily	Teacher	0.16	778	-0.53	180	-0.43	68

		Ages 10-13					
Variable	Source	Non-disruptive		Control		Treatment	
		mean	N	mean	N	Mean	N
Other people do things more easily than I	Subject	0.03	754	-0.11	169	-0.04	63
People talk about me behind my back	Subject	0.06	754	-0.19	169	-0.16	63
No one understands what I feel	Subject	0.06	754	-0.23	169	-0.06	63
I feel worried	Subject	0.03	754	-0.10	169	-0.07	63
Have trouble deciding	Subject	0.00	754	-0.08	169	0.16	63
Feel alone even when I'm with someone	Subject	0.07	754	-0.14	169	-0.45	63
Something doesn't work	Subject	0.07	754	-0.24	169	-0.15	63
Satisfied with myself	Subject	0.03	719	-0.06	161	-0.26	62
I'm not worth anything	Subject	0.02	719	-0.06	161	-0.10	62
I have good qualities	Subject	0.06	719	-0.19	161	-0.18	62
I do things well	Subject	0.02	719	-0.04	160	-0.18	62
Not much to be proud of	Subject	0.07	719	-0.24	161	-0.14	62
I feel useless	Subject	0.03	719	-0.13	161	0.03	62
I'm a valuable person	Subject	0.05	719	-0.13	161	-0.23	62
I should have more self-respect	Subject	0.02	719	-0.08	160	-0.03	62
I feel like a failure	Subject	0.06	718	-0.18	161	-0.21	62
I have a positive attitude to myself	Subject	0.07	718	-0.25	160	-0.20	62
Gives up easily	Teacher	0.15	778	-0.50	180	-0.34	68

		Ages 14-17					
Variable	Source	Non-disruptive		Control		Treatment	
		mean	N	mean	N	Mean	N
Other people do things more easily than I	Subject	0.02	671	-0.09	138	-0.02	55
People talk about me behind my back	Subject	0.05	670	-0.12	137	-0.28	55
No one understands what I feel	Subject	0.04	669	-0.08	138	-0.32	54
I feel worried	Subject	0.03	673	-0.07	138	-0.21	55

Have trouble deciding	Subject	0.02	670	-0.02	137	-0.19	54
Feel alone even when I'm with someone	Subject	0.05	671	-0.22	138	-0.03	54
Something doesn't work	Subject	0.07	672	-0.27	138	-0.18	55
Satisfied with myself	Subject	0.00	615	-0.03	124	0.08	50
I'm not worth anything	Subject	0.01	613	-0.07	124	-0.01	50
I have good qualities	Subject	0.02	615	-0.09	123	-0.05	50
I do things well	Subject	-0.02	615	0.03	124	0.11	50
Not much to be proud of	Subject	0.00	615	0.08	123	-0.14	50
I feel useless	Subject	0.00	615	0.06	124	-0.09	50
I'm a valuable person	Subject	0.05	616	-0.27	124	0.04	50
I should have more self-respect	Subject	0.05	618	-0.12	124	-0.28	50
I feel like a failure	Subject	0.02	618	-0.04	124	-0.11	50
I have a positive attitude to myself	Subject	-0.01	617	-0.02	124	0.20	50
Gives up easily	Teacher	0.13	687	-0.43	143	-0.54	56

Table 22: Components of Altruism Channel

Variable	Source	All Years					
		Non-disruptive		Control		Treatment	
		mean	N	mean	N	Mean	N
Tries to stop a fight	Teacher	0.0252	778	-0.0674	180	-0.11	68
Invites a child who is left out to play	Teacher	0.0196	778	-0.0103	179	-0.197	68
Helps an injured child	Teacher	-0.0123	778	0.0516	180	0.00376	68
Volunteers to put things away	Teacher	0.0136	778	-0.0247	180	-0.0909	68
Congratulates others	Teacher	0.0571	778	-0.165	180	-0.216	68
Shows sympathy	Teacher	0.0523	778	-0.138	180	-0.233	68
Helps a child having trouble	Teacher	0.0274	778	-0.0448	180	-0.195	68
Helps an ill child	Teacher	0.00309	778	0.048	180	-0.162	68
Helps a crying child	Teacher	0.0188	778	-0.0367	180	-0.118	68
Volunteers to clean a spill	Teacher	-0.0319	778	0.154	180	-0.0419	68
Variable	Source	Ages 10-13					
		Non-disruptive		Control		Treatment	
		mean	N	mean	N	Mean	N
Tries to stop a fight	Teacher	0.0472	778	-0.121	180	-0.221	68
Invites a child who is left out to play	Teacher	0.0405	778	-0.0756	179	-0.264	68
Helps an injured child	Teacher	-0.000581	778	0.0159	180	-0.0354	68
Volunteers to put things away	Teacher	0.0288	778	-0.0971	180	-0.0725	68
Congratulates others	Teacher	0.0685	778	-0.196	180	-0.264	68
Shows sympathy	Teacher	0.0628	778	-0.146	180	-0.332	68
Helps a child having trouble	Teacher	0.0433	778	-0.105	180	-0.219	68
Helps an ill child	Teacher	0.0253	778	-0.0299	180	-0.21	68
Helps a crying child	Teacher	0.0338	778	-0.0803	180	-0.174	68
Volunteers to clean a spill	Teacher	-0.00884	777	0.0781	180	-0.106	68
Variable	Source	Ages 14-17					
		Non-disruptive		Control		Treatment	
		mean	N	mean	N	Mean	N
Tries to stop a fight	Teacher	-0.0143	687	0.0141	142	0.14	56
Invites a child who is left out to play	Teacher	-0.0172	685	0.0846	143	-0.00535	56
Helps an injured child	Teacher	-0.00477	684	0.00692	142	0.0408	56
Volunteers to put things away	Teacher	0.00659	686	0.0191	142	-0.129	56
Congratulates others	Teacher	0.0136	686	-0.0463	142	-0.0499	56
Shows sympathy	Teacher	0.0232	686	-0.113	142	0.0024	56
Helps a child having trouble	Teacher	0.0106	687	0.0334	143	-0.216	56
Helps an ill child	Teacher	-0.0154	684	0.1	142	-0.0667	56
Helps a crying child	Teacher	0.00707	685	-0.0121	142	-0.0558	56
Volunteers to clean a spill	Teacher	-0.0322	686	0.105	143	0.128	56

Standardized Values

Table 23: School Outcomes

	All Years					
	Non-disruptive		Control		Treatment	
	mean	N	mean	N	Mean	N
Math	0.0864	755	-0.357	169	-0.0745	65
French	0.0846	755	-0.337	168	-0.112	65
Held Back	-0.144	783	0.496	181	0.335	69
Special Ed	-0.138	783	0.478	181	0.311	69
	Ages 10-13					
	Non-disruptive		Control		Treatment	
	mean	N	mean	N	Mean	N
Math	0.0681	732	-0.265	159	-0.121	63
French	0.0734	733	-0.259	157	-0.209	63
Held Back	-0.127	783	0.402	181	0.392	69
Special Ed	-0.116	783	0.365	181	0.354	69
	Ages 14-17					
	Non-disruptive		Control		Treatment	
	mean	N	mean	N	Mean	N
Math	0.0971	700	-0.371	154	-0.18	60
French	0.0829	705	-0.365	153	-0.0424	60
Held Back	-0.136	781	0.492	180	0.259	69
Special Ed	-0.138	779	0.506	179	0.241	69

Standardized Values

APPENDIX D: ATTRITION AND COMPLIANCE

D.1 Compliance

Compliance was not complete. Of those assigned to the control group, 70% declined to participate in the project, and of those assigned to the treatment group, 67% declined (p-value of difference is 0.65).

D.2 Attrition

From the original 250 subjects, roughly 44% did not respond at the age 21 questionnaire, and 59% did not respond at age 26. This attrition comes both from those that did not respond to the questionnaires, and those that did not complete that question on the questionnaire. Proportionally more treatment than control subjects are missing, although the difference is not statistically significant (p-value= 0.37).

This level of attrition raises concern that our estimate of the impact on treatment may be biased. It is relatively easy to propose reasons that attrition might not be random and uncorrelated to treatment, and not clear which direction the resulting bias would be. For example, suppose that the higher level of attrition in the treatment group is due to higher job mobility in the treatment group due to greater economic success: then the most successful of the treatment group might have moved away from Montreal, which would result in a downward bias of the impact estimate. On the other hand, suppose that the treatment subjects who were the least successful felt more ashamed of their lack of success than the control subjects, because they felt that the University of Montréal had invested in their development, and so they were less likely to respond to the questionnaire – this would result in an upward bias on the impact estimate.

We test whether attrition is related to 25 variables collected prior to the program: both parents living at home; age of mother and father at birth of subject; education of mother and father; work status of mother and father; number of children in the household; adversity index of household; age of subject in 1984; rank in birth order of child; number of brothers, sisters, half-brothers, and half-sisters; whether or not the subject went to pre-school; and initial scores on aggression, anxiety, opposition, pro-sociality, combativeness, inattention, hyperactivity, and anti-social behavior. We construct three dummy variables equal to 1 if data on employment is missing and 0 otherwise (one for age 21, one for age 26, and one for percentage of time worked full time). For each baseline variable to be tested, we construct an interaction term between treatment and the baseline variable (in order to capture any possibility of differential attrition between the groups along that variable). We conduct 26 logit regressions for each year (25 baseline variables plus treatment alone), where the dummy for missing M_i is regressed on treatment T_i , the interaction term ($T_i * X_i$) and the baseline variable X_i , as follows:

$$M_i = \alpha + \beta T_i + \gamma(T_i * X_i) + \delta X_i + \varepsilon_i$$

We do not combine all of the variables into one regression because of the small sample

size. Of greatest concern are the coefficients β and γ , because if they are significant they indicate that attrition is not orthogonal to treatment (random) along that baseline variable, and so at a minimum, that baseline variable should be controlled for in the regression specifications.

Table 39 shows all the significant channels. Father's age at subject's birth and schooling, are negatively correlated to attrition, but the correlation is the same in treatment and control groups. Mother's employment is weakly positively correlated to attrition in the treatment group for employment at age 26. Number of sisters is weakly positively correlated to attrition on the percent full time variable, but only in the treatment group. Note that the number of sisters is one variable that is not equal across treatment and control groups and is already controlled for due to that imbalance. Father's employment is very significantly correlated with attrition and treatment status at age 21 and 26. The combination of coefficients implies that, once father's employment is taken into account, treatment subjects are more likely to be missing than control subjects, and control subjects are more likely to be missing if their father was employed. Note that the checks on randomization showed no correlation (near zero and non significant) of father employment with treatment status. Once aggression or fighting is controlled for, the treatment group is more likely than the control group to be missing for employment at age 21 (note that aggression was not correlated with assignment to treatment). Pro-sociality is positively correlation with attrition only for the treatment group at age 26 (neither group at age 21). Hyperactivity is correlated with attrition and treatment group at age 21. Once hyperactivity is controlled for, the treatment group was more likely to be missing, but this difference decreased with hyperactivity. Note that the checks on randomization showed no correlation between hyperactivity and treatment assignment (near zero and non-significant).

Following these results, we include father's employment, mother's employment, pro-sociality, aggression, fighting and hyperactivity as controls in our preferred specification, along with the variables that were significant in the random assignment check (father's age at subject's birth, prestige of mother's work, initial anxiety, and number of sisters).

In order to get an idea of the possible direction and magnitude of the bias, we estimate the bias introduced by the same level of attrition into the estimate of the program impact on secondary school completion¹¹. Administrative data on secondary school completion is available for nearly the entire sample (242 out of 250 subjects, or 97%). We use the following procedure. We create a false secondary school variable that takes the value of missing if the subject is missing data for employment at age 21 (or 26) and the value of the true secondary school variable if they have data for employment at age 21 (or 26). In this way we mimic the level of attrition in the adult questionnaire data in the secondary school data (we call this the "attrited" dataset). We then estimate the impact of the program on secondary school

11 It is sometimes possible to create bounds on the estimate by calculating the coefficient under two hypothetical cases, one where all of the missing from the treatment group are artificially assigned the "bad" outcome and all of the missing from the control group are artificially assigned the "good" outcome to obtain the lower bound, and the converse to obtain the higher bound. This is not useful in this case because the high rate of attrition would result in extremely wide and uninformative bounds.

completion using our preferred paper specification. We control for all variables related to treatment assignment and that are related to attrition across treatment groups. We compare this to an estimate of the impact of the program using data from the entire sample (the “true” dataset). The validity of this comparison rests on the assumption that the direction of the bias induced in the secondary school completion data is likely to be the same as the direction of the potential bias in the employment data.

The first three columns of Table 38 show this test for attrition at age 21. The point estimates of the attrited (column 1) and true (column 2) datasets are very close in value, suggesting little attrition bias. The second three columns of Table 38 show this test for attrition at age 26. The point estimate for the attrited dataset for age 26 (column 3) is 0.133, while the point estimate for the true dataset is 0.164. There is a bias due to attrition (the point estimate in the attrited sample is 23% lower than in the true sample) but note that it is a downward bias. If the estimate of impact for secondary school completion and group membership and percent of fulltime activity would be biased in the same way by attrition, the estimates used in this paper are likely to be an underestimate of the impact of this program on the questionnaire outcomes.

Table 24. Attrition for Questionnaire data

	Missing questionnaire data			
	Age 21		Age 26	
	Number	%	Number	%
Control (N=181): C	76	42	104	57
Treatment (N=69): T	34	49	44	64
Total (N=250)	101	44	148	59
p-value of difference: C-T	0.30		0.37	

Table 25. Test for Attrition Bias

	Dependent Variable: Secondary School Completion				
	True Sample	Age 21		Age 26	
		Attrited Sample	Attrition Bias	Attrited Sample	Attrition Bias
	(1)	(2)	(3)	(4)	(5)
Treatment (0/1)	0.164** (0.0697)	0.171 (0.105)		0.133 (0.128)	
			-0.007		-0.031
Constant	0.379 (0.288)	0.402 (0.432)		0.827 (0.564)	
Observations	242	139		101	
R-squared	0.156	0.119		0.155	

*Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1*
 Column (1) gives the treatment impact on secondary school graduation estimated for the true (entire) sample. Column (2) gives the treatment impact estimated for a sample where attrition matches the attrition on employment data for age 21. Column (3) gives the treatment impact estimated for a sample where attrition matches the attrition on employment data for age 26. Columns (3) and (5) give the difference in the treatment impact estimates for (1) and (2) and (1) and (4), respectively.

Table 26. Correlation of treatment and attrition for questionnaires at age 21 and 26

	Dependent Variable: Dummy for missing questionnaire at age													
	Age 21	Age 26	Age 21	Age 26	Age 21	Age 26	Age 21	Age 26	Age 21	Age 26	Age 21	Age 26	Age 21	Age 26
Treat (T)	0.129 (0.287)	0.265 (0.293)	1.516 (1.365)	-0.270 (1.468)	-0.625 (0.475)	-0.288 (0.483)	0.344 (1.311)	-2.122 (1.291)	2.949*** (1.067)	2.959*** (1.117)	-0.0875 (0.532)	-0.641 (0.525)	2.129*** (0.789)	1.039 (0.790)
T*Father School			-0.135 (0.134)	0.0589 (0.141)										
Father School			-0.0561 (0.0654)	-0.179*** (0.0666)										
T* Number of Sisters					0.707* (0.370)	0.516 (0.386)								
Number of Sisters					-0.332* (0.170)	-0.114 (0.167)								
T* Mother's Employment							-0.150 (0.714)	1.346* (0.715)						
Mother's Employment							0.654* (0.364)	0.277 (0.341)						
T* Father's Employment									-1.933** (0.836)	-2.010** (0.884)				
Father's Employment									1.476*** (0.421)	1.604*** (0.491)				
T* Initial pro-social score											0.0296 (0.0638)	0.137** (0.0690)		
Initial Prosocial Score											0.0161 (0.0313)	0.000979 (0.0319)		
T* Initial hyperactivity Sco													-0.677*** (0.254)	-0.267 (0.246)
Initial Hyperactivity Score													0.0479 (0.130)	-0.0255 (0.126)
Constant	-0.392*** (0.152)	0.301** (0.151)	0.0308 (0.652)	1.945*** (0.666)	-0.0795 (0.219)	0.413* (0.224)	-1.541** (0.661)	-0.194 (0.606)	-2.441*** (0.555)	-1.743*** (0.588)	-0.497* (0.256)	0.294 (0.257)	-0.516 (0.397)	0.385 (0.384)
Observations	250	250	220	220	249	249	245	245	197	197	250	250	248	248

Robust standard errors in parentheses , *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
 Logit regressions where dependent variable is missing data at age 21 or 26,