School and Drugs: Closing the Gap

Evidence from a Randomized Trial in the US

BY NÚRIA RODRÍGUEZ-PLANAS *

IZA, IAE-CSIC, and Universitat Pompeu Fabra

First version: August, 2012

This version: December, 2012

Abstract

We present evidence on how The Quantum Opportunity Program (QOP hereafter) worked in the US. While the program was regarded as successful in the short-term, its educational results were modest and its effects on risky behaviors detrimental in the long-run. By exploiting the control group's self-reported drug use while in school, we evaluate whether the program worked best among those with high-predicted risk of problem behavior. We find QOP to be extremely successful among high-risk youths as it managed to curb their risky behaviors and, by doing so, it persistently improved high-school graduation by 14 percent and college enrollment by 21 percent. In contrast, QOP was unsuccessful among youths in the bottom-half of the risk distribution as it increased their engagement in risky behaviors (especially while in high-school). Evidence suggests that negative peer effects explain these results.

* Acknowledgments: The Quantum Opportunity Program experiment study was conducted under Contract No. K-5547-5-00-80-30 from the U.S. Department of Labor and the Ford Foundation. The views expressed herein do not necessarily reflect the policies or opinions of the U.S. Department of Labor. This paper has benefitted from comments from Antonio Cabrales, Lawrence Katz, Andrew Oswald, Aaron Sojourner, Christopher Taber, and two anonymous referees. Author's e-mail: rodriguez-planas@iza.org

I. Introduction

Poor academic performance and engagement in risky behaviors are two of the most serious problems facing youth today in the United States. Despite recent improvements, only 76 percent of young people graduated from high school in 2009 in the US, far from the OECD average of 82 percent. In addition, college completion among the 25- to 34-year-olds is also relatively low with the US ranking 15th among 34 OECD countries (*OECD in Figures*, 2011). At the same time, the US also underperforms in terms of 13- to 19-year-old engagement in risky behaviors as it ranks 15th among 30 OECD countries (OECD, 2009).[1] Estimates from the 2010 *National Survey on Drug Use and Health* reveal that as many as 10 and 14 percent of 12- to 17-year-olds reported illicit drug and alcohol use in the past 30 days, respectively. Consequently, a large part of the United States budget is devoted to improving the social and educational outcomes of youths. The U.S. Federal government alone allocated $14.5 billion to *Title I program,* the largest program under *No Child Left Behind,* and over $500 million to *The Safe and Drug Free Schools Program* in the 2009 fiscal year.[2] In addition, many policy makers, practitioners and researchers have sought to develop programs whose main objective is to curb adolescents' problem behaviors such as drug use, and school dropout.

According to psychologists, there is a bidirectional relationship between risky behaviors and poor academic performance (Grant *et al.*, 2011; Guttman *et al.*, 2002; and Symons *et al.*, 1997). At the same time, economists have found that engaging in a variety of risky behaviors leads to poor academic achievement (Yamada *et al.*, 1996; Cook and Moore, 2000) and vice-versa (Sander, 1995a, b, de Walque, 2004; Kenkel *et al.*, 2006; and Grimard and Parent, 2007; and Jensen and Lleras-Muney, 2012). Thus, an interesting and timely policy question given today's limited resources is to ask whether it is more effective to target adolescents at a high-risk of problem behavior or those on the margin of socially behaving. Empirical assessment of this question has

[1] This ranking is based on the following 3 indicators of risk taking: 15-year-olds who smoke regularly, 13- and 15-year-olds who report having been drunk on more than two occasions, and birthrates to females aged 15 to 19.

[2] These two programs aim at improving primary and secondary education of disadvantaged students and preventing violence and illegal use of tobacco, alcohol, and drugs by students, respectively.

proven difficult for the following two reasons. First, the way psychologists study risk taking relies on asking individuals to respond to hypothetical dilemmas under conditions designed to minimize emotional influences on decision making. However, as Steinberg, 2004, points out: "*…in the real world the risky, or potential risky, situations in which adolescents find themselves are anything but hypothetical….and they most likely occur under conditions of emotional arousal*." Second, non-experimental evaluations of remedial programs are likely to bias the results since the reasons for which individuals self-select into programs are likely to be correlated with the underlying determinants of their outcomes.

This paper avoids the latter problem by using data from a randomized experiment in which some low-performing high-school students from low-performing high-schools in the US were offered mentoring, educational services and financial rewards during the 4 years of high-school (plus one additional year in case students fell behind one grade) while others were not (the control group).[3] In addition, as the data collection included surveys conducted 5, 7 and 10 years after random assignment on four key domains—high-school, post-secondary education, labor market, and risky behaviors—, we are able to use the control group's self-reported drug use at age 19 to estimate predicted probabilities of problem behavior. We then evaluate whether the program was effective for two distinct groups defined by whether their predicted probabilities of drug use at age 19 was below or above the control group median. As young women may have more self-discipline (Duckworth and Seligman, 2006), be more likely to delay gratification (Silverman, 2003), or have lower discount rates than young men (Warner and Pleeter, 2001), implying a differential heterogeneity effect in drug use, the assignment to the top- or bottom-half of the risk distribution is performed separately by gender.[4]

The research design in this paper is based on comparison of youths with high- (or low-) predicted chance of drug use who were randomly assigned to a treatment and

[3] Enrollees who graduated from high-school on time received some mentoring and assistance in enrolling in post-secondary education or training between graduation and the end of the fifth year of the experiment.

[4] Because gender may lead to heterogeneity effects in programs effectiveness (Rodríguez-Planas, 2012b), the analysis of the effectiveness of QOP was also done by gender. Results by gender are similar to those presented in the main text below and available from the author upon request.

control group when they first entered high-school in the *Quantum Opportunity Program* (QOP thereafter) social experiment, which was implemented at 11 high-schools across 7 sites in the United States. The QOP goals were to help youths graduate from high school and enroll in post-secondary education or training, and reduce risky behaviors. Students entering 9th grade in September 1995, except for those with an 8th grade point average (GPA) above the 66th percentile, were randomly assigned *within each school* to a treatment group or a control group. It is thus important to note that the program targeted disadvantaged youths from low performing schools, and its targeted population thus had major difficult barriers to overcome, such as substance abuse problems or broken families. Over five years, the treatment group was offered substantial cash awards and an array of support services including mentoring by full-time social workers, academic tutoring and life skill activities. The control group was only eligible for standard youth programs offered in the community. The final sample for the QOP experiment consists of 1,069 students, 580 in the QOP group and 489 in the control group.[5]

While Rodríguez-Planas (*AEJ: Applied Economics 2012*) studies the same randomized evaluation, this author focuses on the average effects of the program and gender heterogeneity, leaving many unanswered but important questions. Why were the beneficial effects of QOP on high-school graduation short lived? Why didn't QOP reduce risky behaviors during high-school? Why did QOP increase engagement in risky behaviors 5 years after the end of the program? By focusing on QOP's differential effect by ex-ante risk of problem behavior, the current article unravels the mechanisms through which QOP may have worked (or failed to do so). In addition, the evidence from this article reveals that targeting resources may close the skills gap between those most at-risk and those at the margin.[6]

[5] The final sample includes a slightly larger number of youths in the treatment group (580) than in the control group (489) by design of the evaluation sample, where it was decided that each site would have either 50 or 100 youths in the treatment group (with the exception of the Washington D.C. site, which had 80 youths). The design of the evaluation sample is thoroughly explained in Rodríguez-Planas (2012a).

[6] The only subgroup analysis in Rodríguez-Planas, 2012, is by gender, The heterogeneity analyses of the Mathematica Policy Research Reports and the late 1980s pilot have focused on effects by gender, age, eight grade GPA or site.

More specifically, this study shows that QOP was extremely successful at improving educational and behavioral outcomes for those most at risk, that is, youths in the top-half of the predicted drug use distribution. For this group, QOP curbed risky behaviors (including substance abuse). The QOP mentors were social workers trained to identify and deal with the many structural barriers facing youths. They addressed any problems in any aspect of the lives of the treated youths, monitored their progress and advocated for them in matters pertaining to school, family, the juvenile justice system and college. The mentor was expected to model appropriate behavior and attitudes, set disciplinary standards, and be continually available. Our estimates suggest that mentors succeeded in addressing the most severe problems of the high-risk treated youths. Perhaps more importantly positive findings on substance abuse reduction among this group persist over time, up to 10 years after randomization. In addition, QOP also succeeded in improving their high-school and college outcomes both in the short-, medium- and long-run for this group. For instance, when youths were in their mid-twenties, QOP had increased high-school graduation for the high-risk group by 14 percent and college enrollment by 21 percent. In contrast, we find evidence that QOP was unsuccessful among youths in the bottom-half of the predicted drug-use distribution as it increased their engagement in risky behaviors (especially while in high-school) and had no effect on educational outcomes in the short-, medium-, or long-run. The evidence presented is suggestive that negative peer effects may explain the lack of beneficial effects among this group. Our results are robust to the use of alternative measures of problem behavior to estimate ex-ante high- and low-risk groups. In addition, we also discuss the consequences of differential attrition by treatment status.

Recently, several researchers have found that targeting disadvantaged children before formal schooling begins is most efficient as the cognitive ability and character of children 0- to 5-years-old is the most malleable (Currie, 2001; Krueger and Whitmore, 2001; Heckman, 2000; Garces *et al.*, 2002; and Carneiro and Heckman, 2003). This paper takes a different approach and contributes to the debate over whether investments in the later stages of a child's development have positive payoffs. While other rigorous studies have evaluated interventions targeted to youths at a

relatively late stage of schooling (Deming, 2011; Machin, McNally and Meghir, 2004; and Lavy and Schlosser, 2005; Angrist and Lavy, 2009; Rodríguez-Planas, 2012a, and Ludwig, 2012), very few explore the effectiveness of an intervention when targeting adolescents based on the predicted probability of problem behavior. The most similar study is that of Deming, 2011, who analyzes the impact of middle- or high-school choice on crime using a similar prediction based on ex-ante data to organize students by whether they have a high ex-ante predicted probability of arrest. Similarly, Angrist and Lavy, 2009, study whether targeting cash incentives based on the *predicted or fitted* chance of high-school matriculation certification is effective at improving school matriculation certification rates in Israeli low-performing schools. Another relevant study is that of Dynarski *et al.*, 2011, who use a similar approach to analyze the effect of early childhood investment (Project STAR) on college enrollment and degree completion by whether the child had low or high ex-ante probability of attending college. While Deming, 2011, and Dynarski *et al.*, 2011, find that the impacts are concentrated among high-risk youths, Angrist and Lavy, 2009, find that their intervention works best for girls with high predicted Bagrut rates. In contrast with these studies, this paper is the first to find that the program had unintended negative effects on low-risk youths, making a stronger point for targeting resources.[7]

Finally, this paper also contributes to the peer effects literature (see Epple and Romano, 2011, for a recent review). However, most of this literature has focused on analyzing the direct peer effect in education. In contrast, this paper explores the effects of peer's risky behavior on outcomes within a context of an evaluation of a mentoring program. Due to its nature, the program is likely to have facilitated peer bonding, which may have had detrimental effects on its participants. Our results complement findings that concentrations of high-risk youth increase the aggregate level of misbehavior (Cook and Ludwig, 2005; Imberman et al., 2011; Carrell and Hoekstra, 2010; and Deming, 2011).

[7] It is important to notice that while the samples in many early childhood interventions and other social programs (including QOP) were selected based on the principle that it may well be the most disadvantaged who can benefit the most from these interventions, the contribution of this paper is that, even *within* those populations, the question remains on whether it may be useful to target those at high-risk of engaging in risky behaviors.

This paper is organized as follows. The next section describes the program implementation, the evaluation framework, and the data. Section III analyzes the results and presents robustness checks. Section IV concludes.

II. The Quantum Opportunity Program

A. Program Description

QOP engaged treated students in: (i) developmental activities aimed at developing their social and employment-readiness skills; (ii) community service activities aimed at developing their sense of community membership; and (iii) educational services designed at improving their academic performance. Examples of such activities are displayed in Appendix Table A.1. These activities were to be performed after school and during one half day on the weekend. The full treatment consisted of 750 hours of services per year (equally distributed among the three different activities), which (if achieved) would have represented around three-quarters of the hours required for in-school instruction per year.[8]

To encourage participation, QOP offered financial incentives to treated students. First, they received $1.25 for each hour devoted explicitly to educational activities, developmental activities (excluding recreational activities) and community service. Second, they were promised if they obtained a high school diploma or GED *and* enrolled in post-secondary education or training they would receive an amount equal to their total earnings. The treated youths received on average over $1,000 after high-school graduation and enrollment in post-secondary education. This aspect of the financial reward was partly motivated by evidence that a $1,000 reduction in college tuition increases college enrollment by about 4 percentage points (Susan Dynarski 2003).

[8]In 2000, the average number of instructional hours spent in public school by 15-year-olds was 990 hours (U.S. Department of Education 2005; Table 26-2).

Case managers with prior expertise in social services were hired for the five-year period.[9] Each had a caseload of only 15 to 25 treated students. The objective was to develop a highly personal, long-lasting connection with each youth that mirrored the relationship between a teenager and a nurturing, supportive older relative. Therefore, the case manager was instructed to focus on sustaining a strong relationship with the youth regardless of behavior, including whether the youth disengaged from the program, dropped out of school, became incarcerated, or moved out of the area. Case managers were also to manage the provision of supportive services to address personal, family and social barriers that might interfere with the treated student's ability to attend school and do well there.[10]

At almost $25,000 per enrollee for the whole intervention, QOP was an expensive program. By comparison, the operating costs of the likewise-expensive Job Corps were approximately $16,500 per participant in 1998 (Schochet, Burghardt, and McConnell 2008). Rodríguez-Planas estimates average effects of QOP and finds that the average beneficial effects of the program on high-school graduation were short lived, and that the program did *not* reduce risky behaviors during high-school. Although the program did manage to increase postsecondary education five years after the end of the program, it also increased engagement in risky behaviors. Thus, many researchers and policy makers question whether the outcomes of the program warrant its intensity and high costs.

B. Evaluation Framework

The estimates reported are intent-to-treat (ITT) effects that make no adjustments for remaining involved or service participation in QOP. They are computed from estimating the following equation:

$$Y_{ist} = \alpha + \beta_1 D_i^T + \beta_2 X_{is0} + \delta_s + \varepsilon_{ist} \quad (1)$$

[9] Most mentors stayed with the program for several years and many stayed for the entire five years of the experiment. Unfortunately, no information on sex, race or ethnicity of mentors was collected.

[10] These barriers could be addressed either directly by the case manager or by referral to a community resource, such as a substance abuse program or local agencies that provide housing, food, income support, or child care.

where Y_{ist} denotes an outcome variable for youth *i* from school *s* at time *t*, D_i^T is a dummy variable that takes value one if the youth belonged to the treatment group, and X_{is0} is a vector of controls including an indicator for being 14 years old when entering ninth grade, an indicator for being over age 14 when entering ninth grade, an indicator for being in the middle third of the eighth-grade GPA distribution, an indicator for being in the top third of the eight-grade GPA distribution, an indicator for being black and an indicator for being Hispanic.[11] δ_s are school dummies, and ε_{ist} is the error term. β_1 measures the treatment effect of the program's impact on outcome *Y*. Standard errors are clustered at the school level, which was the unit of randomization.[12]

The analysis focuses on several outcomes from four key families of outcomes, high-school, post-secondary education, labor market, and risky behaviors, measured at three different points in time. The results are presented separately for youths at the bottom- and top-half of the predicted drug use at age 19 distribution. To address the problem of multiple inference, we follow Kling, Liebman, and Katz 2007 and construct a summary index, Y*, as the unweighted average of all standardized outcomes within a family. That is, $Y^* = \frac{\sum_k Y_k^*}{k}$, where $Y_k^* = \frac{Y_k - \mu_k}{\sigma_k}$. Let Y_k be the k^{th} outcome of *K* variables within each family, with the sign for adverse outcomes (substance abuse, crime, welfare) reversed, so that a higher value of the normalized measure represents a more "beneficial" outcome. Standardization is performed using mean (μ_k) and standard deviation (σ_k) for the control group. The summary indices presented in this paper are constructed using the same specific outcome variables as in Rodríguez-Planas, 2012a.

[11] Since treatment was assigned randomly, the insertion of these covariates improves estimation precision without affecting the consistency of the parameter of interest. Precision is improved because the inclusion of the covariates accounts for chance differences between treated and control groups in the distribution of pre-random assignment characteristics, as well as for non-random attrition in the follow-up surveys.

[12] Results are robust to not clustering standard errors at the school level. In addition, we replicated our analysis using a logit model and find very similar results.

To classify youths by their risk of problem behavior we use a predictive regression that models the probability of problem behavior as a function of school dummies, individual covariates, and interactions between them. The predictive model is:

$$Y_{is} = \alpha_0 + X'_{is0}\beta + \delta_s + \varepsilon_{is} \quad (2)$$

where δ_s are as in equation (1), and Y_{is} a dummy variable that takes value one if the youth reported having used drugs in the past 30 days during the first survey, which was collected at age 19. To increase the prediction of equation (2), X_{is0} includes all the baseline variables in equation (1) with the exception of the GPA indicators, which now have been replaced by a continuous eighth-grade GPA variable. In addition, X_{is0} also includes interactions between all of the covariates among themselves and between the covariates and the school dummies. It is important to note that model (2) is first estimated using *only youths from the control group*. Appendix Table A.2 reports the coefficients estimated using equation (2). These estimates are strong predictors of future drug use. The pseudo R-squared from the regression is 0.24 and many regressors are statistically significant.[13]

We then estimated equation (1) using the following subgroup-classification scheme: we used the fitted values from model (2) to divide students into roughly equal-sized groups, those in the top and the bottom of the predicted-risk distribution. This classification was done by gender, as the predicted drug use during high-school of females was considerably lower than that of males.[14]

C. The Data and Descriptive Statistics

The data for this study come from baseline information collected prior to random assignment, program implementation and participation, high-school transcripts, math

[13] It is important to note that the model (2) in this version of the paper has all of the covariates interacted with each other, which was not done in the August 2012 Discussion Paper version. Doing so, considerably improves the prediction of the model as the pseudo-R squared has increased from about 0.07 to 0.24.

[14] For males, those with a predicted probability of drug use above 32 percent where classified in the top-half of the risk distribution, whereas the threshold for females was 11 percent.

and reading achievement tests conducted 4 years after random assignment, and three surveys conducted 5, 7 and 10 years after random assignment. Transcript data were collected 5 years after random assignment from *all* high-schools treated and control group individuals attended.

Table 1 reports means and differences in means by treatment status for our baseline variables.[15] Columns 1 to 3 refer to the whole sample. They show that the sample is roughly equally divided between boys and girls. The median age of students was 14 when random assignment took place. The sample consists mainly of minority students: about two thirds are African-American and about one fourth are Hispanic (black or nonblack). There are no significant differences by treatment status.

Columns 4 to 9 present treatment and control means and differences for those in the top- and bottom-half of the predicted drug use distribution. Among youths in the bottom-half of the predicted risk distribution we observe that treated youths are less likely to be males, while treated youths in the top-half of the predicted drug use distribution are more likely to be older than those in the control group. Note that if anything the imbalance among youths in the bottom-half of the risk distribution ought to introduce an upward bias in our estimates of QOP as earlier evidence shows that the program worked much better among females than males (Rodríguez-Planas, 2012a). In contrast, the imbalance in the top-half of the distribution may introduce a downward bias as earlier evidence shows that the program worked better among younger youths. Nonetheless, to account for these differences, estimation of equation (1) controls for all baseline characteristics as explained above.

Control response rates and treatment-control differences for the achievement test completion and the follow-up surveys are displayed in Table 2.[16] The response rate to the achievement tests and the first telephone survey was 80 percent for the control group. The response rates to the second and third telephone surveys were 69 and 72 percent for the control group, respectively. Attrition was higher among control

[15]As no baseline survey was collected, the pre-program information available is limited. However, the main baseline variable that is missing (compared to similar evaluations) is parent's education level. Fortunately, this variable is likely to be correlated with pre-program GPA, which we do have.

[16] All individuals from the treatment and control group were surveyed, regardless of whether they failed to respond in one of the earlier surveys.

group members than among treatment group members in all but the third telephone survey. The differential response rate between the treatment and control groups is 7 percentage points in the first telephone survey, 10 percentage points in the second telephone survey, and a statistically insignificant 3 percentage points in the third telephone survey. After dividing the sample by the predicted probability of drug use at age 19 we observe that most of the attrition differential between treated and control youths is driven by those in the top-half of the risk distribution. As these are the most difficult youths, higher attrition among them is expected. In addition, because those in the treatment group have a relationship with the mentor, they ought to also have been easier to contact in the follow-up surveys leading to a higher response rate relative to those in the control group. At the end of the results section, we discuss the consequences of differential attrition and non-response.[17]

The only available measure of program participation is the stipend or accrual accounts participants received. This is problematic as mentoring time did not count toward stipends or accrual account contributions. In addition, enrollees received "bonus hours" when they achieved a significant milestone, such as obtaining a B average or higher on his or her report card. These bonus hours cannot be distinguished from regular hours. Their inclusion overestimates time spent on program activities for some enrollees.

D. Implementation of QOP and Service Use

Overall, QOP was successfully implemented. Most case managers reported developing close mentoring relationships with the majority of the individuals assigned

[17]We were also able to use transcripts data and exploit the fact that survey non-respondents did not necessarily overlap across surveys to build two variables measuring high-school completion and GED diploma receipt for most of the population at each point in time. We recoded as high-school graduates those survey non-respondents who had reported in an earlier survey being a high school graduate or for whom their transcript data clearly confirmed that they had indeed graduated from high-school. Similarly, we recoded as high-school dropouts those first- (or second-) survey non-respondents who had later responded to the second or third surveys and who had reported having dropped out of high-school. In addition, in the few cases where we found inconsistencies across surveys or across survey and transcripts, we used the answer that was corroborated at least twice. Several robustness checks have been performed using alternative definitions of high-school graduation and the results are robust to those reported in the main tables. For these two variables, differences in the response rates are considerably smaller (although they remain statistically significant at the 5 percent level as shown in Table 2).

to them and they all provided access to services regardless of enrollees' behavior.[18] Although the educational, community service and development activity component fell short of the target of 750 hours per year, enrollees still invested a substantial amount of time in QOP activities during the first four years. The average of 886 hours for the whole group, including summers but not time spent with their mentors, corresponds to about 89 percent of an extra school year and more than half the average instruction time received by Job Corps participants, by far the most intensive education and training program for disadvantaged youths in the United States (Schochet *et al.*, 2008). The fact that QOP did not achieve its extremely ambitious target should not affect the external validity of this evaluation; if the program were implemented on a broader scale, it is likely that its implementation would be similar.

Enrollees spent an average of 61 hours per year on education, 76 hours on developmental activities, and 27 hours in community service.[19] Not surprisingly, the average time spent on QOP activities fell steadily from 279 hours in the first year of the experiment to 125 hours in the fourth year (see Appendix Table A.3). Youths who had participated in QOP activities early during the experiment and then stopped or decreased participation over time gave leaving school, working, and family responsibilities as the main reasons for doing so. It is important to highlight, however, that almost the totality of QOP youths engaged in QOP activities, as all but 1 percent of enrollees spent some time on QOP activities in the first year.

Appendix Table A.3 also shows the distribution of hours over activities and year by whether youths were in the bottom- or top-half of the predicted drug use distribution. It is important to highlight that youths in the top-half of the risk distribution spent slightly more hours in QOP activities than those in the bottom-half (albeit the difference is not statistically significant). In addition, although youths in the top-half of the risk distribution spent less hours during the first year than those in the bottom-half, the opposite is true during the fourth year, suggesting a more

[18] See Maxfield *et al.* 2003a, and Maxfield *et al.*, 2003b, for further description of program design and implementation.

[19] In the case of community services, the lower intake was due to enrollees' lack of interest in this type of activities and case managers' belief that enrollees had a greater need for other QOP services. Most sites decided to reallocate their resources away from community service to developmental and educational activities.

persistent involvement with QOP. Indeed we observe that while a higher percentage of them (44 percent) were no longer involved in QOP during the fourth year, of those who were still involved in QOP, the intensity of involvement was higher among those in the top-half of the risk distribution.

III. Results

Tables 3 through 6 display results for the summary indices followed by each specific outcome that were components of the index. Each table covers each of the key domains under analysis, risky behaviors, high-school, post-secondary education, and labor market, measured in the short-, medium- and long-run. To draw general conclusions about the results of the experiment, we begin with the summary indices for a given category and then we discuss the individual outcomes as the magnitudes of these separate outcomes are often easier to interpret than those of the summary indices. The first two columns present the control group outcome means for those in the bottom- or top-half of the fitted distribution of drug use. Separate ITT estimates for those in the bottom- or top-half of the fitted distribution of drug use are displayed in columns 3 and 4. It is important to keep in mind that a positive coefficient on the summary indices indicates a beneficial effect of the program, and a negative coefficient indicates a detrimental effect of QOP.

Control Group Means.— When youths were in their late-teens, we observe that average drug use for youths in the top-half of the predicted drug use distribution is 3.5 times higher than that of youths in the bottom-half (shown in columns 1 and 2 of Table 3). In addition, binge drinking and committing a crime is also higher for the former than the latter. As engagement in risky behaviors decreases with age (especially for youths in the top-half of the predicted drug use distribution), the differences between the two groups have considerably narrowed by the time youths are in their mid-twenties.

Focusing now on high-school outcomes (shown in Table 4), we also observe that, when youths were in their late teens, those in the top-half of the risk distribution

are about 4 percentage points less likely to have graduated from high-school than those in the bottom-half. This differential persists and increases over time: when youths are in their mid-twenties, those in the top-half are 7 percentage points less likely to have graduated from high-school than those in the bottom-half. Similarly, we observe in Table 6 that youths in the top-half of the risk distribution are, on average, less likely to be employed and they earn on average lower earnings than those in the top-half (at least in their mid-twenties). In contrast, we do not observe any differences between the two groups in terms of math and reading test scores when youths were in their late-teens. We also do not observe any major significant differences in terms of enrolling in college or post-secondary education (shown in Table 5).

Risky-Behavioral Outcomes.— QOP managed to curb risky behaviors among those most at-risk. When youths were in their late-teens, the ITT estimate for the summary index in column 4 of Table 3 indicates a positive effect of 0.103 (significant at the 5 percent level) for youths in the top-half of the predicted distribution. The individual variables indicate that QOP's positive effect on reducing risky behaviors for this group is driven by all of the individual estimates, albeit being measured with less precision. In fact, the only individual outcome that is statistically significant (at the 5 percent level) is illegal drug use indicating that QOP decreased substance abuse among the treated youths by 8.7 percentage points (or 21 percent). In contrast, the effect is -0.174 for the risky behaviors domain among those with low-predicted risk of drug use (significant at the 1 percent level) reveals that QOP increased risky behaviors among this group during their late-teens. This is driven by significant increases in the likelihood of binge drinking or using drugs in the past month, and committing a crime in the past year of 11.6, 22.3, and 8.5 percentage points, respectively.

By the time youths were in their early-twenties, the ITT estimates for the risky behaviors domain are considerably smaller and no longer statistically significant. Nonetheless, the estimate is positive for youths in the top-half of the predicted drug use distribution and negative for those in the bottom-half and we reject the null hypothesis that these estimates are equal at the 5 percent significance level. Moreover, it is important to note that the individual estimates continue to show large beneficial

effects of QOP in terms of reducing risky behaviors of similar size and significance as when youths were in their late-teens (even though, now, the control means are considerably lower). For instance, we observe that QOP decreased drug use and crime among this group by 10.2 and 5.3 percentage points, respectively--both estimates are statistically significant at the 5 percent level the former and 10 percent level the latter. The reason the summary index has decreased in size and precision for this group is that treated youths are more likely to be on welfare in their early-teens than those in the control group (6.6 percentage points more likely although the effect is not statistically significant). Several reasons may explain this, namely that treated youths may have more information about social programs, how to access them, or may have further trust in society than youths in the control group. For youths in the bottom-half of the risk distribution we observe that QOP's negative effect on risky behaviors has faded away by the time youths are in their early-twenties. Indeed, the coefficients for the individual estimates are considerably smaller and no longer statistically significant.

It is important to note that QOP was not only successful at decreasing substance abuse among those in the top-half of the risk distribution while the program lasted or soon after the program ended, but that this beneficial effect persisted up to 5 years *after* the end of the program. By the time youths are in their mid-twenties, we continue to observe that QOP reduced substance abuse among treated youths in the top-half of the risk distribution. Among this group, QOP decreased drug use by 5.9 percentage points (or 48 percent). This coefficient is significant at the 5 percent level. In contrast, we observe a significant and negative effect of QOP on the risky behaviors' domain among youths in the bottom-half of the risk distribution. This effect is driven by a 1.9 percentage points higher likelihood of committing a crime (significant at the 10 percent level). This is not the only intervention to find unanticipated adverse effect on crime in the long-run. Kling *et al.*, 2005 also find that the Moving to Opportunity (MTO) housing vouchers led to short-term reductions in violent crime arrests, but long-term *increases* in property crime arrests for males. Similarly, the randomized, experimental trial of the community-based treatment program *Cambridge-Somerville Youth Study*, whose objective was to prevent delinquency, found that youths in the treatment group were more likely in the long run

to be rearrested for crimes and have further negative impacts on physical and psychological health compared to youths in the control group (McCord, 1978 and 1992). The impacts were measured up to 30 years later using official state records.

In relation to earlier QOP evidence (Rodríguez-Planas, 2012a), this paper brings to light that the lack of average effects of QOP in terms of reducing risky behaviors when youths were in their late-teens is due to a differential effect of the program by whether youths had low- or high-predicted probability of drug use. This finding is similar to that of Deming, 2011, and Dynarsky *et al.*, 2011, in that it reveals that such type of programs work best among those most at-risk within an already disadvantaged population. However, in contrast with these other studies, this paper finds that QOP had unintended negative effects on low-risk youths, making a stronger point for targeting resources.

Sensitiveness of the Results to the Choice of Predicted Risk of Adverse Outcomes.— As engagement in risky behavior is self-reported (as opposed to being drawn from administrative records) a frequent concern in this literature is that certain individuals (those with higher propensity to engage in risky behaviors) may not respond or, if they do respond, they dishonestly report illegal behavior.[20] When estimating the predicted probabilities of drug use, this is a problem to the extent that control group youths do not respond or under-report substance abuse since only control youths are used to obtain the coefficients of model (2). Thus, if they under-report drug use, our predicted probabilities will reflect that for *both* the treatment and the control groups. When estimating the causal effect of QOP on the different outcomes of interest, the concern is that non-response or misreporting is correlated with treatment status and unobservable characteristics that are correlated with the outcomes of interest.

To minimize these problems our measures of risky behaviors were obtained from a paper survey in which students had quite some privacy when answering the questions (as opposed to a telephone survey where the student answers questions

[20] That said, it is difficult to have accurate administrative data on drug use, as the only way that it can be monitored is if the individual has been arrested and tested. Thus, in the case of substance abuse, self-reporting with paper survey is likely to be as good as it gets.

directly to the interviewer). Moreover, prior to responding the survey students were reassured that their answers would be treated confidentially. And to minimize such concerns, the paper surveys had a number id instead of the students' names on them to identify the students. Thus, the students were not necessarily aware that their survey would be matched to the rest of data collected from them. Nonetheless, as questions regarding drug use are sensitive questions, under- or misreporting may well have occurred. That said, it is important to note that item non-response for the drug use question was minimal (2.6 percent among control youths and 3.6 percent among treated youths).

To address this concern we explored the sensitivity of the results to using other adverse outcomes to predict risk, in particular outcomes that may either be easier to check (such as being a dropout, being arrested, or committing a crime at age 19), or outcomes that may not have such a strong stigma attached to them (such as binge drinking at age 19). Results are not sensitive to the choice of predictor question. Appendix Table A.4 presents findings when the estimated risk distribution is based on the likelihood of having committed a crime in the last year.[21] The results are very similar to the ones found above. We find that QOP was successful at curbing risky behaviors among youths in the top-half of the predicted crime distribution and that these positive effects persisted over time (for instance, in their mid-twenties, high-risk youths are 5 percentage points less likely to report using drugs than the control group--significant at the 10 percent level). Moreover, we also find that QOP had detrimental effects on youths in the bottom-half of the predicted crime distribution both when youths were in their late-teens (by increasing substance about and crime involvement) and in their mid-twenties (by increasing their odds of being arrested). Perhaps particularly relevant is that the individual outcome estimates measured in the medium- and long-run are very similar. Alternatively, at the end of this Section we address thoroughly the issue of survey non-response.

[21] Estimates using other predictor questions and for outcomes from other domains are available from the author upon request. Nonetheless they corroborate the findings presented in this paper.

High-School Outcomes.— Columns 3 and 4 of Table 4 bring to light another new result: QOP worked extremely well in terms of increasing high-school graduation for youths in the top-half of the predicted drug use distribution, and this result persist over time as youths age. Column 3 of Table 4 shows a small and non-significant ITT estimate of QOP on the high-school domain for youths with *low*-predicted drug use in the short-run that becomes negative although not statistically significant when youths are in their early- and late-twenties. In contrast, the ITT estimates of QOP for youths with *high*-predicted drug use during high-school show large, significant, and persistent ITT effects of QOP on the high-school domain. The estimates of 0.071, 0.114 and 0.126 in column 4 of Table 4 indicate that the mean effect of being in the treatment group for the high-school outcomes is one eighth to one fourteenth standard deviations for youths in the top-half of the predicted drug distribution. The individual outcomes reveal that these beneficial effects are driven by statistically significant increases in the incidence of graduating from high-school (by about 8 percentage points). As the high-school graduation rate in the control group among those in the top-half of the fitted distribution is 41 percent in the short-run, 54 percent in the medium-run and 58 percent in the long run, the treatment effect corresponds to a 20 percent increase when youths were in their late-teens, 15 percent increase when youths were in their early-twenties, and 14 percent increase when youths were in their mid-twenties. Because the control group means of those in the top-half of the distribution are 4, 6 and 7 percentage points lower than those in the bottom-half, it appears that QOP reduced treated youths differences in terms of high-school graduation by increasing the performance of those with a high predicted risk of drug use while in school towards (and beyond) the level of those with a lower predicted risk. While the size of these beneficial estimates is large compared to that of other evaluations, a recent study by Ludwig (2012) on the effects of social-cognitive skills on disadvantaged males grades 7 to 10 show similar large educational effects of the program.

College Outcomes.— Moving now to Table 5, we again observe that QOP was successful among youths in the top-half of the predicted drug use distribution. All three ITT estimates of the college domain for this group are large and statistically significant at the 10 or 5 percent level (shown in column 4 of Table 5). The individual

outcomes reveal that the positive and statistically significant effect is driven by an 8.9 percentage points (or 35 percent) increase in the odds of attending college in the late-teens, and a 9.4 percentage points (or 29 percent) increase in the odds of attending college in the early-twenties. This effect persist over time: by the time youths are in their mid-twenties, QOP increased the likelihood of ever attending college of those in the top-half of the risk distribution by 8.3 percentage points (or 21 percent). As there are few differences in the control group means between those in the bottom- and top-half of the risk distribution, QOP did not narrow the gap between the two groups in terms of post-secondary educational outcomes, but put those in the top-half of the distribution ahead of those in the bottom-half instead.

Employment Outcomes.—Information on employment was first collected at the time of the 2nd survey, when youths were in their early twenties. For results measured at that time, we observe a negative and statistically significant effect at the 10 percent level effect for youths with low-predicted chance of using drugs during high-school. The ITT estimates of -0.207 in column 3 of Table 6 indicates that the mean effect of being in the treatment group for the six employment outcomes is one fifth of a standard deviation for those with low-predicted probability of using drugs. The individual estimates show that these treated youths were 17.3 percentage points *less* likely to have a job with health insurance than youths in the control group. This effect is significant at the 5 percent level. While the ITT estimate for the employment domain is not significant for youths in the top-half of the predicted drug use distribution, we find that treated youths within this group were 8.1 percentage points less likely (or 14 percent) less likely to have a full time job. These negative effects on employment may be explained by the fact that treated youths are more likely to be enrolled in post-secondary education (although effects on post-secondary attendance for those in the bottom-half of the predicted drug use distribution are not statistically significant as shown in column 3 of Table 5). By the time youths are in their mid-twenties, the negative ITT effect on the employment domain of youths in the bottom-half of the risk distribution is considerably smaller and no longer significant. For yours in the top-half of the risk distribution, all but one of the estimates are positive but they lack precision.

Potential Mechanisms and Peer Effects.— Unfortunately, the evaluation of QOP was not designed to test alternative mechanisms. However, the analysis thus far reveals that QOP worked very well for youths with a high-predicted risk of drug use as it decreased their engagement in risky behaviors (especially during high-school and shortly afterwards), and increased high-school graduation rate, and post-secondary educational involvement. Most importantly, the educational beneficial results persisted over time. Thus, it appears that QOP's mentors, who were social workers, did a good job at identifying youths most at-risk and addressing their non-cognitive and structural barriers, and by doing so their educational outcomes improved. To corroborate this, we find evidence that treated youths in the top-half of the distribution became more optimistic about life at the end of the program than those in the control group. For instance, they were 5.9 percentage points (or 10 percent) more likely to have a positive attitude towards life, or 4.7 percentage points (or 6 percent) more likely to disagree that bad things happen to people like them than the control group. Both estimates are statistically significant at the 10 percent level. In contrast, for those in the bottom-half of the distribution, QOP did not have any positive effect on their outlook of life.[22]

In contrast, QOP did not seem to work that well among treated youths in the bottom-half of the distribution as it had practically no effect on educational outcomes, and it increased engagement in risky behaviors while in high-school. Indeed, we find that among this group, treated youths in their late-teens were 7.5 percentage points (or 9 percent) *less* likely to disagree with the statement "feels committing a crime is always wrong" and 7.2 percentage points (or 10 percent) *less* likely to disagree with the statement "feels dropping out of school is wrong" than youths in the control group. So, what went wrong during implementation for youths with low-predicted chance of problem behavior? Negative peer effects are a plausible explanation. As discussed earlier, each mentor in QOP had assigned several mentees with whom he or she frequently worked at the same time as a group. In addition QOP offered cultural and

[22] Control means for the control group in the top-half of the risk distribution are 57 percent for having a positive attitude towards life and 79 percent for disagreeing that bad things happen to people like them. For those in the bottom half of the risk distribution, control group means are 57 percent and 79 percent, respectively.

recreational activities to treated youths to help them build strong relationships with mentors and peers. However, building strong cohesion within a group may have backfired. Evidence from psychologists and economists reveals that peer effects play an important role on getting youths involved in risky activities. This occurs because peers reinforce deviant conduct by responding with approval and attention (Dishion, *et al.*, 1999; Dishion, *et al.*, 1996; and Patterson, *et al.*, 2000).[23]

To test whether negative peer effects is a plausible story, we estimated the following peer effects variable. For each individual we estimated the proportion of students (excluding the individual himself) within the same school *s*, treatment status *T*, sex *g*, and race *r*, that reported doing drugs at age 19. Notice that we exploit the randomization in our evaluation design to isolate the source of variation in the distribution of peer-groups that is exogenous to an individual's own choices and abilities, and overcome the usual problems of selection in the peers' literature. To assess the credibility of this identification strategy, we find that there is considerable treatment status variation in the proportion of drug users within-cells. In addition, below we demonstrate that this within-cell variation does not tend to be related to variation in student background characteristics.

We then estimated equation (3) below:

$$(3) \quad Y_{ist} = \alpha + \beta_1 D_i^T + \beta_2 PDrugs_{sTgr1} + \beta_3 \left(D_i^T * PDrugs_{sTgr1} \right) + \beta_4 X_{is0} + \delta_s + \varepsilon_{ist}$$

where Y_{ist} denotes an outcome variable for youth *i* from school *s* at time *t*, D_i^T is a dummy variable that takes value one if the youth belonged to the treatment group, $PDrugs_{sTgr1}$ is the proportion of students in school *s*, treatment status *T*, gender status *g*, race status *r*, measured at age 19. X_{is0} is a vector of controls including an indicator for being 14 years old when entering ninth grade, an indicator for being over age 14 when entering ninth grade, an indicator for being in the middle third of the eighth-

[23] See Brook *et al.*, 1998; Kandel, 1985; Jessor *et al.*, 1980 for peers' influence on marijuana use; Norton *et al.*, 1998, Jensen and Lleras-Muney, 2010 for drinking and peer effects; and Case and Katz, 1991 for peer-group interactions and criminal activity.

grade GPA distribution, an indicator for being in the top third of the eight-grade GPA distribution, an indicator for being black and an indicator for being Hispanic. δ_s are school dummies, and ε_{ist} is the error term. Standard errors are clustered at the school level. β_1 measures the treatment effect of the program's impact on outcome *Y* if there were no peer effects. β_2 captures the effect of having a higher proportion of peers using drugs during high-school among the control group. β_3 captures the *differential* effect of having a higher proportion of peers using drugs during high-school among the treatment group relative to the control group. A negative β_3 is indicative of negative peer effects among the treated youths.

The key identifying assumption to yield valid estimates of the peer effects is that the variation in the proportion of drug users is random within school-treatment status-gender and race status. Appendix Table A.5 checks whether the proportion of drug users within a school/treatment status/gender/race cell is correlated with students' baseline characteristics. When conditioning by school effects, most of the associations between the proportion of drug users and background characteristics are eliminated. There are a few imbalances left. In any case we control for baseline characteristics in equation (3).

Table 7 displays these estimates. We present estimates from two specifications of equation (3): one with no controls or school fixed-effects (columns 1 to 3 and 7 to 9), and another one with all baseline controls and school fixed-effects (columns 4 to 6 and 10 to 12). It is important to note that statistical significance and sign of the β_3 estimates are, in general, robust to the two alternative specifications shown (with or without covariates and school fixed-effects). Our discussion focuses on estimates from the specification with all controls.

Among youths in the bottom-half of the risk distribution, we observe that there is a negative peer effect for youths in the treatment group. For instance, the coefficient β_3 (shown in column 6) is negative and statistically significant for the college outcome in the short-, medium-, and long-run, and the high-school outcome in the short-run, indicating that having peers who engaged in drug use during high-school decreased the beneficial effect of QOP for these domains. Once we account for this

peer effect, QOP improves the college outcomes in the short-, medium-, and long-run for youths in the bottom-half of the distribution and the high-school outcome in the short-run as β_1 in column 4 is now positive and significant. Similarly, the estimate of QOP on risky behaviors when youths were in their late-teens, β_1, is now positive (albeit not significant), while the coefficient on the interaction between QOP and the peers effect variable, β_3, is negative and statistically significant. Although the β_1 coefficients for the risky behaviors domain in the short-run is estimated with less precision, this sign reversal on the treatment dummy when peer effects are accounted for suggests that negative peer effects among treated youths in the bottom-half of the distribution are a part of the story behind these treated youths' increased engagement in risky behaviors during high-school.

Column (12) in Table 7 also shows that these negative peer effects on risky behaviors are particularly large among youths in the top-half of the risk distribution in the medium- and long-run. Thus, it seems that while mentors were able to curb risky behaviors among those at high-risk during high school, they did not succeed in completely changing behavior in the medium- and long-run. Notice that the potential reflection problem is less so in this case as youths have been away from their high-school peers for 3 to 5 years.

Alternative Explanations.— Below we explore alternative and potentially complementary explanations for these results. Such explanations must explain why QOP had differential impacts for youths in the bottom- and top-half of the risk distribution.

An alternative explanation is that as QOP gave stipends, youths may have used this additional income to purchase alcohol and drugs. There is indeed growing evidence showing that youths are very responsive to economic incentives, such as prices, when deciding to undertake risky behaviors (Nisbet and Vakil, 1972; Grossman *et al.*, 1994; Chaplouka and Wechler, 1996; Gruber, 2001; Pacula *et al.*, 2001). However, on average treated youths received $350 the first year and $156 the fourth year. Per week, this represents $6 and less than $3 in the first and fourth year, respectively. While $1.25 per hour may have encouraged participation, overall youths

received much less than minimum wage, which went from $4.25 in 1995 to $5.15 in 1999. Thus, it is unlikely that QOP's stipends alone would explain the observed substance abuse. Moreover, as the differences in participation across those in the top- and bottom-half of the distribution are small (as explained in Section I.D), it is difficult for this hypothesis to explain the differential effect across the two groups.

Another explanation is that QOP may have led to a substitution away from parents' attention, either because they trusted that the mentor was also watching over their children or because QOP unintentionally weakened the ties between parents and enrollees, breaking important social bonds and thus leading to the observed perverse effects. As this may have happened mainly among youths in the bottom-half of the risk distribution (as those in the top-half did not have those strong family ties to start with), this explanation could explain the differential effects of QOP by ex-ante predicted risk. To explore whether this could be a possible explanation we re-estimated equation (1) adding a dummy equal to 1 if the youth reported having an influential adult during high-school and the interaction of this variable with the treatment effect.[24] If the negative effect on risky behaviors for youths in the bottom-half of the distribution were caused by a substitution effect, we would find that this detrimental effect of QOP on risky behaviors would be partly explained by this interaction. We found no evidence of that. To the contrary, estimates suggest that, if anything, mentors decreased youths' engagement in risky behaviors.[25]

Attrition from the Follow-up Surveys.—As we saw in the data section, differential attrition between treated and control groups exists especially among youths in the top-half of the predicted drug use distribution. This section analyzes the sensitivity of the estimates to potential attrition bias. As most of the problem is within youths in the

[24] Clearly including having a dummy indicating presence of an influential adult while in high-school is endogenous, however, we do this only to see how much of the estimated causal detrimental effect of QOP could be driven by an influential adult substituting parents' attention. We find that if anything mentors decreased youths' engagement in risky behaviors.

[25] The coefficient on the QOP dummy in the risky behaviors domain when youths were in their late-teens becomes -0.316 (s.e. 0.089) and the interaction between QOP and having an influential adult is 0.208 (s.e. 0.129) for youths in the bottom-half of the distribution. For youths in the top-half of the distribution, the coefficient on the QOP dummy becomes 0.109 (s.e. 0.143) and the interaction term becomes 0.020 (0.180).

top-half of the predicted drug use distribution, the analysis focuses on this group.[26] To address the non-response problem, we impute outcomes for the missing observations. The most extreme approach is to impute the minimum (maximum) value of each variable in the observed treatment distribution to the non-responders in the treatment group, and the maximum (minimum) value of the observed control distribution to the non-responders in the control group. This approach renders the results uninformative. In all cases the bounds range from large negative to large positive and are always highly significant (see Appendix Table A.6).

As a less extreme approach, we follow Kling and Liebman 2004, and impute to the lower (upper) bound the mean minus (plus) 0.25 standard deviations of the observed treatment distribution to the non-responders in the treatment group, the mean plus (minus) 0.25 standard deviations of the observed control distribution to non-responders in the control group. We also repeat this exercise using 0.1 (0.05) standard deviations. The wider bounds include zero for all our statistically significant results except for the positive effect on college and post-secondary outcomes when the youths were in their late teens and in their early-twenties; for high-school outcomes when youths were in their early- and mid-twenties; and for risky behaviors when the youths were in their late teens. Using the intermediate bounds ensures no sign reversals. The tightest bounds result in statistically significant positive effects on all (but two) estimates, even at the bounds in the case of all educational outcomes in the short-, medium- and long-run and risky behaviors in the short-run.[27]

Finally, since we were able to construct high-school completion and GED diploma receipt for a considerably more complete sample using transcript data (as explained in footnote 17 in Section I.D and shown in Table 2), we compare these estimates of QOP (discussed in the main text and shown in Table 4) to those estimated using the attrited sample. Appendix Table A.7 displays these results, and shows that differences across the two estimates are small, suggesting that if anything the differences between respondents and non-respondents are modest at most.

[26] Estimates for youths in the bottom-half of the distribution are available from the author upon request.

[27] The two exceptions are employment outcomes and risky behaviors when the youths were in mid-twenties.

Interpretation of Results and Targeting Implementation.—The high- and low-risk groups were constructed by splitting into separate samples by gender youths below or above the median predicted probability of drug use. This implies that the percentage of youths in the top-half of the predicted drug-use distribution varied widely across schools. In all but 3 schools, the percentage of high-risk youths ranged between 23 percent to 72 percent. In the other 3 schools the percentage was 17; 89; and 94 percent respectively. This variation is important as it reveals that the prediction in this paper is *not* picking up mainly school-specific treatment effects. Moreover, if targeting were to be implemented on a large scale in the population, one would expect that it would be as it is done in this paper as opposed to constraining risk groups to be a 50/50 split in *each* school and gender group; that is, the ideal policy would be to target those most at risk (even though some may concentrate in certain schools or neighborhoods) than to target those most at risk *within* a school.

Self-Selection into Social or Remedial Programs and Spillover Effects.—In many interventions (even when the treatment is randomly assigned), the eligible sample is self-selected as youths have to sign up for the program and then (if treatment is randomized) some get the program and others not. This is a concern because it is likely that these applicants (or their families) are trying to escape the negative influence of their peers in their neighborhood school (in the case of public school choice lotteries), or have higher perseverance, ability or social ambition, or lower discount rates than those who do not apply to the remedial program. If this is the case the concern is that youths in the control group may continue to pursue their objective of escaping the neighborhood or persevering, leading to biased causal estimates of the intervention. It is important to note that this concern is less likely to arise in the QOP evaluation as eligibility was not requested but instead assigned to *all* entering eight graders with a 8^{th}-grade GPA below the 66^{th} percentile in each school. Thus, QOP did not target the most driven youths from disadvantaged neighborhoods.

A related concern, though, is the existence of spill-over effects. Despite not operating within the school or within school hours, QOP may have affected the performance of youths in the control group as students spent time with each other and

consequently learnt from each other. Alternatively (and possibly complementary), QOP may have stimulated control group members to work harder as they knew that some of their classmates were receiving additional help. If QOP had positive spill-over effects, the estimates in this paper are likely to be underestimates of the true effect of the program, especially in the short-term as any potential spill-over effects are likely to be a greater problem when students are still together in school or have recently finished. Moreover, it is unclear how and why spillover effects may have a differential effect between youths in the top- and bottom-half of the predicted drug use distribution.

To explore whether spill-over effects are a concern, we compared outcomes of our control group to those with poor 8th-grade academic performance in the National Education Longitudinal Study (NELS) of 1988. As we found no evidence that our control group outperformed apparently similar youths, we conclude that this is most likely not an issue in this study. For instance, we find that the likelihood of receiving a high-school diploma (or GED) 5 to 6 years after scheduled high-school graduation was 62 percent (or 79 percent) for our control group, below, but not far from, the NELS 88 rate of 67 percent (or 79 percent) for similarly disadvantaged youths. Similarly, within 5 to 6 years of scheduled high-school graduation, 56 percent of control group members had engaged in some type of post-secondary education or training, not far from the NELS 88 rate of 58 percent for disadvantaged students.

III. Conclusion

This papers finds that the expensive and controversial QOP program was extremely successful at curbing risky behaviors and improving educational outcomes of youths with ex-ante high-predicted risk. These positive impacts persisted beyond high-school, 10 years after random assignment. In contrast, results indicate that QOP was unsuccessful among youths in the bottom-half of the predicted risk distribution as it increased their engagement in risky behaviors (especially while in high-school) and had no impact on educational outcomes. The evidence presented in this paper is

suggestive that negative impacts among the ex-ante low-risk group are driven by peer effects. Thus, this paper finds that targeting remedial programs and interventions to youths at high-risk of engaging in risky behaviors improves effectiveness for two reasons. First, resources are given to those who benefit the most from them (the high-risk group). Second, unintended detrimental peer effects are minimized as those at low-risk are not offered the program.

An important question is to what degree targeting intervention resources on the subgroup identified as the one who benefits the most in this paper is feasible in the real world? Because engagement in risky behaviors can be predicted using information that is available to the school district level, the implementation of giving priority to high-risk youth is possible. However, the concern is whether this allocation method may be controversial and even illegal as race was used as a predictor variable. Nonetheless other programs, such as free school lunch programs are based on income standards which are highly correlated with race. Moreover, the novel result in this paper is that well intended interventions may end up having unintended detrimental impacts for certain subgroups, which would justify that programs are tailored to youths' problems and needs.

REFERENCES

Angrist, Joshua D., and Victor Lavy. 2009. “The Effect of High School Matriculation Awards: Evidence from Randomized Trials.” *American Economic Review*, 99 (4): 1384-1414.

Carneiro, P., and J. Heckman, James J., 2003. "Human Capital Policy," IZA Discussion Papers 821, Institute for the Study of Labor (IZA).

Cook, Phillip J., and Michael J. Moore. “Alcohol.” In *Handbook of health economics*, vol. 1b, ed. Anthony Culyer and Joseph Newhouse, 629–74. Amsterdam: Elsevier.

Currie, Janet. 2001. “Early childhood education programs.” *Journal of Economic Perspectives* 15, no. 2:213–38.

de Walque, Damien (2007). "Does Education Affect Smoking Behaviors?: Evidence

Using the Vietnam Draft as an Instrument for College Education," *Journal of Health Economics*, 26(5), p. 877-895.

Deming, David. 2011. " Better Schools, Less Crime?" *The Quarterly Journal of Economics*, 126:2063-2115.

Duckworth, Angela, and Martin Seligman. 2006. "Self-Discipline Gives Girls the Edge: Gender in Self-Discipline, Grades, and Achievement Test Scores." *Journal of Educational Psychology* 98(1): 198-208.

Dynarski, Susan. 2003. "Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion." *American Economic Review* 93(1): 279–88.

Dynarski, Susan. 2008. "Building the Stock of College-Educated Labor." *Journal of Human Resources*, 43(3): 576-610.

Dynarski, Susan, Joshua Hyman and Diane Schanzenbach. 2011. "Experimental Evidence on the Effect of Childhood Investments on Postsecondary Attainment and Degree Completion." *NBER Working Paper* No. 17533

Epple, Dennis, and Richard Romano. 2011. "Peer Effects in Education: A Survey of the Theory and Evidence." In *Handbook of Social Economics*, Vol. 1B, ed. Jess Benhabib, Alberto Bisin, and Matthew Jackson, 1053–1163. Amsterdam: North-Holland, Elsevier.

Grant JE, Potenza MN, Krishnan-Sarin S, *2011*. "Stealing among high school students: prevalence and clinical correlates." *J Am Acad Psychiatry Law 39*:*44*–52.

Gutman LM, Sameroff AJ, Eccles JS. 2002. "The academic achievement of African American students during early adolescence: an examination of multiple risk, promotive, and protective factors." *Am. J. Community Psychol.* 30:367–99

Grimard, Franque & Parent, Daniel, 2007. "Education and smoking: Were Vietnam war draft avoiders also more likely to avoid smoking?," *Journal of Health Economics*, Elsevier, vol. 26(5), pages 896-926, September.

Heckman, James J., 2000. "Policies to foster human capital," *Research in Economics*, Elsevier, vol. 54(1), pages 3-56, March.

Jensen R., and A. Lleras-Muney, 2012. "Does Staying in School (and not Working) prevent Teen Drinking and Smoking?" ***Journal of Health Economics***, 31(4), p. 644-675.

Kenkel, D. S. 1993. "Drinking, driving, and deterrence: The effectiveness and social costs of alternative policies." *Journal of Law and Economics* 36:877–913.

Kling, Jeffrey, and Jeffrey Liebman. 2004. "Experimental Analysis of Neighborhood Effects on Youth." unpublished manuscript.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2005. "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment." *Quarterly Journal of Economics*, 120(1):87-130, 2005.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica*, 75 (1): 83-119.

Krueger A., and D. Whitmore, 2001. "Would Smaller Classes Help Close the Black-White Achievement Gap?," in John E. Chubb and Tom Loveless, eds., *Bridging the Achievement Gap*, Washington: Brookings Institution Press, 2002.

Lavy, Victor, and Analia Schlosser. 2005. "Targeted Remedial Education for Underperforming Teenagers: Costs and Benefits." *Journal of Labor Economics*, 23(4): 839–74.

Ludwig, Jens. 2012. "Improving Social-Cognitive Skills Among Youth." University of Chicago Crime Lab Research and Policy Brief.

Machin, Steven, Sandra McNally, and Costas Meghir, 2004. "Improving Pupil Performance in English Secondary Schools: Excellence in Cities," *Journal of the European Economic Association*, MIT Press, vol. 2(2-3), pages 396-405, 04/05.

Maxfield, Myles, Laura Castner, Vida Maralani, and Mary Vencill. 2003. "The Quantum Opportunity Program Demonstration: Implementation Findings." Washington, DC: Mathematica Policy Research, Inc.

Maxfield, Myles, Allen Schirm, and Núria Rodríguez-Planas. 2003. "The Quantum Opportunities Program Demonstration: Implementation and Short-Term Impacts." Mathematica Policy Research Report 8279-093. Washington, DC: Mathematica Policy Research, Inc.

McCord, Joan. 1978. "A thirty-year follow-up of treatment effects." *American Psychologist, 2*, 284-289.

McCord, Joan. 1992. "The Cambridge-Somerville Study: A pioneering longitudinal-experimental study of delinquency prevention." In J. McCord & R. E. Tremblay (Eds.), *Preventing antisocial behavior: Interventions from birth through adolescence* (pp. 196-206). New York: Guilford Press.

OECD, 2009. "Comparative Child Well-being across the OECD." Chapter 2 in *Doing Better for Children*, OECD.

OECD, 2011 Education in figures.

Rodríguez-Planas, N. (2012a) “Longer-Term Impacts of Mentoring, Educational Services, and Learning Incentives: Evidence from a Randomized Trial in the U.S.” *American Economic Journal: Applied Economics*, 4 (4), 121–139.

Rodríguez-Planas, N. (2012b) “Mentoring, Educational Services, and Financial Incentives: What Do We Know about Them?” *Evaluation and Program Planning*, 35, 481–490.

Sander, William. (1995a). "Schooling and Quitting Smoking," *Review of Economics and Statistics*, 77, p. 191-199.

Sander, William. (1995b). "Schooling and Smoking," *Economics of Education Review*, 14, p. 23-33.

Schochet, Peter, John Burghardt, and Sheena McConnell. 2008. “Does Job Corps Work? Impact Findings from the National Job Corps Study.” *American Economic Review*, 98:5, 1864-1886.

Silverman, Irwin. 2003. “Gender Differences in the Delay of Gratification: A Meta-Analysis.” *Sex Roles*, 49(9-10): 451-463.

Steinberg, L. 2004. "Risk taking in adolescence: What changes, and why?" *Annals of the New York Academy of Sciences, 1021,* 51–58.

U.S. Department of Education, National Center for Education Statistics. *The Condition of Education 2005*. NCES 2005-094. Washington, DC: U.S. Department of Education.

Warner, John, and Saul Pleeter. 2001. “The Personal Discount Rate: Evidence from Military Downsizing Programs.” *The American Economic Review,* 91(1): 33-53.

Yamada, Tetsuji, Michael Kendix, and Tadashi Yamada. 1996. The impact of alcohol consumption and marijuana use on high school graduation. *Health Economics* 5, no. 1:77–92.

TABLE 1. Descriptive Statistics By Predicted Drug Use at Age 19

				Low-risk of drug use at 19			*High-risk of drug use at 19*		
	Treatment means	*Control means*	*Treatment - Control*	*Treatment means*	*Control means*	*Treatment - Control*	*Treatment means*	*Control means*	*Treatment - Control*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*
Baseline Covariates									
Male	0.522	0.558	-0.040	0.492	0.557	-0.056*	0.544	0.559	-0.032
			(0.030)			(0.029			(0.038)
Age when entering 9th grade									
< 14	0.107	0.110	-0.000	0.112	0.090	0.029	0.102	0.131	-0.017
			(0.019)			(0.020)			(0.019)
14	0.533	0.575	-0.040	0.505	0.533	-0.030	0.561	0.616	-0.039
			(0.030)			(0.050)			(0.033)
> 14	0.360	0.315	0.041	0.383	0.377	0.001	0.337	0.253	0.055**
			(0.028)			(0.049)			(0.022)
Hispanic	0.262	0.257	0.005	0.369	0.315	0.025	0.151	0.200	-0.020
			(0.016)			(0.015)			(0.021)
Black	0.683	0.679	0.004	0. 590	0. 615	0.007	0.779	0.743	0.0047
			(0.013)			(0.008)			(0.016)
Rank based on 8th grade GPA									
Bottom third	0.365	0.329	0.037	0.397	0.324	0.062	0.333	0.335	-0.000
			(0.029)			(0.037)			(0.029)
Middle third	0.307	0.352	-0.044	0.295	0.365	-0.065	0.319	0.339	-0.016
			(0.029)			(0.039)			(0.038)
Top third	0.328	0.319	0.008	0.308	0.311	0.002	0.347	0.327	0.017
			(0.029)			(0.040)			(0.045)
Sample size	*580*	*489*	*1,069*	*295*	*244*	*539*	*285*	*245*	*530*

Note: Robust standard errors clustered at the school level are reported in parentheses in columns 3, 6 and 9. Columns 3, 6 and 9 presents the coefficient on the Treatment dummy from a regression model with the Treatment dummy plus school dummies.

*, **, *** Estimate significantly different from zero at the 90%, 95%, or 99% confidence level.

TABLE 2. Treatment Effect on Response Rates By Predicted Drug Use at Age 19

				Low-risk of drug use at 19			*High-risk of drug use at 19*		
	Control means	*Treatment - Control*		*CONTROL MEANS*	*TREATMENT-CONTROL*		*CONTROL MEANS*	*TREATMENT-CONTROL*	
OUTCOMES		*School dummies*	*Full set of controls*		*School dummies*	*Full set of controls*		*School dummies*	*Full set of controls*
Achievement tests	0.800	0.080***	0.081***	0.811	0.069**	0.064*	0.788	0.097**	0.100***
		(0.023)	(0.022)		(0.028)	(0.029)		(0.032)	(0.031)
First telephone survey	0.795	0.071***	0.072***	0.811	0.050	0.047	0.780	0.095***	0.097***
		(0.023)	(0.023)		(0.032)	(0.033)		(0.033)	(0.033)
Second telephone survey	0.685	0.096***	0.097***	0.689	0.064	0.062	0.682	0.129***	0.123***
		(0.027)	(0.0273)		(0.039)	(0.040)		(0.038)	(0.037)
Third telephone survey	0.724	0.032	0.0326	0.725	0.018	0.016	0.722	0.047	0.046
		(0.027)	(0.0270)		(0.038)	(0.039)		(0.038)	(0.038)
High-school and GED recodes									
First survey	0.926	0.031**	0.032**	0.934	0.024	0.022	0.918	0.036*	0.035
		(0.015)	(0.015)		(0.020)	(0.020)		(0.021)	(0.022)
Second survey	0.856	0.046**	0.049**	0.865	0.034	0.036	0.849	0.059**	0.057**
		(0.020)	(0.020)		(0.029)	(0.029)		(0.029)	(0.028)
Third survey	0.836	0.046**	0.050**	0.848	0.014	0.017	0.825	0.078**	0.079***
		(0.021)	(0.021)		(0.030)	(0.031)		(0.030)	(0.0301)

Notes: Robust standard errors clustered at the school level are reported in parentheses. "Full set of controls" includes school dummies, an indicator for being male, an indicator for being 14 years old when entering ninth grade, an indicator for being over age 14 when entering ninth grade, an indicator for being in the middle third of the eighth-grade GPA distribution, an indicator for being in the top third of the eight-grade GPA distribution, an indicator for being black and an indicator for being Hispanic.

*, **, *** Estimate significantly different from zero at the 90%, 95%, or 99% confidence level.

TABLE 3. Treatment Effect on Risky Behaviors By Predicted Drug Use at Age 19

	Control Means		*Treatment-Control*	
	Low-risk	*High-risk*	*Low-risk*	*High-risk*
OUTCOMES	*(1)*	*(2)*	*(3)*	*(4)*
Youths Were in their Late Teens				
Summary measure			**-0.174***†††**	**0.103**†††**
			(0.035)	**(0.046)**
Binge drinking in the past 30 days	0.141	0.237	0.116**†††	-0.043†††
			(0.051)	(0.048)
Used any illegal drug in the past 30 days	0.121	0.419	0.223***†††	-0.087**†††
			(0.033)	(0.037)
Committed a crime in past 12 Months	0.233	0.335	0.085**††	-0.038††
			(0.038)	(0.039)
Ever arrested or charged	0.258	0.295	0.0009	-0.062
			(0.020)	(0.045)
Have first child before age 18	0.244	0.236	-0.021	-0.031
			(0.027)	(0.028)
Youths Were in their Early Twenties				
Summary measure			**-0.049††**	**0.053††**
			(0.041)	**(0.044)**
Binge drinking in the past 30 days	0.242	0.287	-0.013	-0.050
			(0.042)	(0.047)
Used any illegal drug in the past 30 days	0.096	0.144	0.022†	-0.102**†
			(0.033)	(0.039)
Committed a crime in past 3 months	0.048	0.029	0.032††	-0.053*††
			(0.029)	(0.026)
Arrested or charged in past 3 months	0.042	0.069	0.014	-0.0028
			(0.026)	(0.019)
Have first child before age 18	0.151	0.167	0.0213	0.037
			(0.0300)	(0.044)
Currently on welfare	0.170	0.249	0.005	0.066
			(0.028)	(0.038)
Youths Were in their Mid-Twenties				
Summary measure			**-0.11*** **	**-0.039**
			(0.030)	**(0.055)**
Binge drinking in the past 30 days	0.290	0.296	0.058	-0.030
			(0.035)	(0.084)
Used any illegal drug in the past 30 days	0.1097	0.124	0.046†	-0.059**†
			(0.037)	(0.022)
Committed a crime in past 3 months	0.011	0.016	0.019*	0.0127
			(0.010)	(0.017)
Arrested or charged in past 2 years	0.034	0.059	0.036	0.050
			(0.023)	(0.031)
Have first child before age 18	0.158	0.156	0.005	0.023
			(0.031)	(0.031)
Currently on welfare	0.192	0.253	0.038	0.0413
			(0.049)	(0.038)

Notes: The table reports estimates of treatment effects on the dependent variables indicated in row headings. Robust standard errors clustered at the school level are reported in parentheses. "Full set of controls" includes school dummies, an indicator for being male, an indicator for being 14 years old when entering ninth grade, an indicator for being over age 14 when entering ninth grade, an indicator for being in the middle third of the eighth-grade GPA distribution, an indicator for being in the top third of the eight-grade GPA distribution, an indicator for being black and an indicator for being Hispanic.

*, ** , *** Estimate significantly different from zero at the 90%, 95% level, or 99% level.

†, ††, ††† indicates that the difference of the estimated effects between youths in the bottom- and top-half of the predicted drug use distribution is significant at the 90%, 95%, and 99% level.

TABLE 4. Treatment Effect on High-School Outcomes By Predicted Drug Use at Age 19

	Control Means		*Treatment-Control*	
	Low-risk	*High-risk*	*Low-risk*	*High-risk*
OUTCOMES	*(1)*	*(2)*	*(3)*	*(4)*
Youths Were in their Late Teens				
Summary measure			**0.027**	**0.071***
			(0.032)	**(0.039)**
Obtained high-school diploma	0.456	0.413	0.008	0.082*
			(0.038)	(0.041)
Obtained a GED	0.065	0.087	-0.008	-0.006
			(0.027)	(0.028)
Obtained HS diploma or GED or still in HS	0.658	0.613	-0.012	0.0853*
			(0.042)	(0.040)
Math test scores (percentile)	40.199	40.892	0.850	-0.306
	[7.818]	[6.473]	(0.585)	(0.570)
Reading test scores (percentile)	42.313	43.472	0.679	-0.041
	[7.818]	[7.216]	(0.637)	(0.582)
Youths Were in their Early Twenties				
Summary measure			**-0.043†**	**0.114*†**
			(0.044)	**(0.053)**
Obtained high-school diploma	0.597	0.538	-0.038††	0.088**††
			(0.045)	(0.031)
Obtained a GED	0.120	0.153	-0.000	0.016
			(0.033)	(0.036)
Youths Were in their Mid-Twenties				
Summary measure			**-0.092†††**	**0.126***†††**
			(0.053)	**(0.035)**
Obtained high-school diploma	0.652	0.579	-0.045††	0.079**††
			(0.048)	(0.034)
Obtained a GED	0.174	0.158	-0.035	0.033
			(0.036)	(0.0367)

Notes: Standard deviation in brackets. See notes on Table 3.

TABLE 5. Treatment Effect on Post-Secondary Education By Predicted Drug Use at Age 19

	Control Means		*Treatment-Control*	
	Low-risk	*High-risk*	*Low-risk*	*High-risk*
OUTCOMES	*(1)*	*(2)*	*(3)*	*(4)*
Youths Were in their Late Teens				
Summary measure			**0.072**	**0.155***
			(0.075)	**(0.084)**
Attending or accepted in	0.279	0.257	0.023	0.089**
College			(0.026)	(0.034)
Attending college	0.204	0.173	0.015	0.052
			(0.025)	(0.033)
Attending postsecondary	0.255	0.280	0.049	0.059
Education			(0.037)	(0.047)
Youths Were in their Early Twenties				
Summary measure			**0.057**	**0.181****
			(0.10)	**(0.062)**
Ever in college	0.347	0.329	0.037	0.094**
			(0.055)	(0.041)
Number of semesters	1.004	0.905	0.001	0.132
in College	[1.921]	[1.652]	(0.185)	(0.144)
Ever in postsecondary	0.545	0.554	0.044	0.136***
Education			(0.058)	(0.036)
Youths Were in their Mid-Twenties				
Summary measure			**0.0339**	**0.125****
			(0.0440)	**(0.046)**
Obtained a bachelor's	0.011	0.028	0.007	0.014
Degree			(0.013)	(0.022)
Obtained a bachelor's or	0.068	0.073	-0.017	0.0042
associate degree			(0.019)	(0.020)
Number of semesters in	1.56	1.67	-0.007	0.305*
College	[3.80]	[2.96]	(0.174)	(0.157)
Completed 2 years of	0.274	0.328	0.058*	0.083**
college or training			(0.028)	(0.028)
Ever in college	0.363	0.390	0.007	0.083**
			(0.035)	(0.034)
Ever in post-secondary	0.551	0.565	0.041	0.101**
Education			(0.039)	(0.036)

Notes: See notes on Table 3. Post-secondary education includes two- and four-year college, vocational or technical school, and the armed forces.

TABLE 6. Treatment Effect on Employment Outcomes By Predicted Drug Use at Age 19

	Control Means		*Treatment-Control*	
	Low-risk	*High-risk*	*Low-risk*	*High-risk*
OUTCOMES	*(1)*	*(2)*	*(3)*	*(4)*
Youths Were in their Early Twenties				
Summary measure			**-0.207***	**-0.0154**
			(0.114)	**(0.069)**
Has a job	0.760	0.699	-0.102	0.012
			(0.064)	(0.030)
Has a full-time job	0.584	0.566	-0.106	-0.081**
			(0.065)	(0.034)
Has a job with health insurance	0.463	0.344	-0.173**†††	0.033†††
Has a job			(0.064)	(0.036)
Attending postsecondary education	0.838	0.777	-0.064	0.046
or working			(0.046)	(0.047)
Usual hours worked per week at main	28.789	27.56	-3.911	-1.006
Job	[19.154]	[20.814]	(2.817)	(1.668)
Hourly wage at main job	7.614	7.524	-0.975	0.344
	[16.498]	[16.016]	(0.979)	(0.528)
Youths Were in their Mid-Twenties				
Summary measure			**-0.045**	**0.0469**
			(0.074)	**(0.0557)**
Has a job	0.746	0.667	-0.006	0.00362
			(0.045)	(0.0370)
Has a full-time job	0.604	0.500	-0.044	0.0519
			(0.055)	(0.0442)
Has a job with health insurance	0.537	0.445	-0.064†	0.0386†
			(0.036)	(0.0297)
Hourly wage at main job (dollars)	10.46	8.25	-0.604	-0.409
	[18.02]	[13.55]	(1.308)	(1.057)
Usual hours worked per week at main	29.48	25.56	-0.290	0.582
job	[20.44]	[20.74]	(2.063)	(1.439)
Total earnings in past 12 Months	14,156	12,732	381.5	1,208
(dollars)	[14,120]	[12,459]	(1,836)	(1,324)

Notes: See notes on Table 3. No information on employment was collected when youths were in their late-teens. Earnings are coded as zero if the person is reported not working. This measure of earnings is one of *realized* earnings and is frequently used in the literature, despite being a crude measure of productivity—since earnings are only observed for employed individuals.

TABLE 7. Peer Effects By Predicted Drug Use at Age 19

	Low-risk						*High-risk*					
OUTCOMES	*QOP*	*PEER*	*QOP*PEER*	*QOP*	*PEER*	*QOP*PEER*	*QOP*	*PEER*	*QOP*PEER*	*QOP*	*PEER*	*QOP*PEER*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*	*(12)*
Youths Were in their Late Teens												
High-school	0.294***	0.717*	-1.271***	0.203*	0.613	-0.884*	0.00626	-0.574**	0.160	-0.0594	-0.031	0.441
	(0.110)	(0.381)	(0.464)	(0.103)	(0.415)	(0.456)	(0.112)	(0.257)	(0.362)	(0.113)	(0.273)	(0.367)
College and Post-Secondary Education	0.448***	0.612	-1.730**	0.411**	0.729	-1.472*	0.118	-1.229***	0.0713	-0.0414	-0.734	0.682
	(0.170)	(0.593)	(0.698)	(0.182)	(0.712)	(0.768)	(0.213)	(0.394)	(0.605)	(0.225)	(0.488)	(0.651)
Risky Behaviors	0.138	-0.647*	-0.829*	0.161	-0.384	-1.019**	-0.0153	-1.144***	0.440	-0.0397	-0.979***	0.548
	(0.108)	(0.389)	(0.489)	(0.098)	(0.487)	(0.488)	(0.119)	(0.300)	(0.410)	(0.124)	(0.330)	(0.417)
Youths Were in their Early Twenties												
High-school	0.148*	0.248	-0.783***	0.069	0.0533	-0.422	0.096	-0.344***	0.0361	0.035	-0.066	0.271*
	(0.068)	(0.231)	(0.187)	(0.090)	(0.276)	(0.323)	(0.091)	(0.107)	(0.219)	(0.076)	(0.225)	(0.132)
College and Post-Secondary Education	0.410***	0.858	-1.745*	0.402***	0.726	-1.530*	0.153	-0.877	0.0726	-0.037	-0.387	0.782
	(0.122)	(0.881)	(0.854)	(0.113)	(0.913)	(0.774)	(0.222)	(0.522)	(0.745)	(0.255)	(0.506)	(0.806)
Employment	-0.309**	0.610	0.120	-0.216	0.447	-0.157	0.243	0.804*	-1.058	0.294	0.571*	-1.114
	(0.135)	(0.615)	(0.550)	(0.146)	(0.578)	(0.538)	(0.180)	(0.430)	(0.604)	(0.180)	(0.263)	(0.671)
Risky Behaviors	0.052	-0.282	-0.200	0.0624	-0.160	-0.327	0.216*	0.390	-0.640	0.287*	0.475	-0.847*
	(0.087)	(0.335)	(0.336)	(0.085)	(0.410)	(0.349)	(0.107)	(0.299)	(0.361)	(0.138)	(0.282)	(0.463)
Youths Were in their Mid-Twenties												
High-school	0.003	0.110	-0.369	-0.054	-0.109	-0.094	0.121	-0.227	-0.030	0.080	0.145	0.140
	(0.093)	(0.267)	(0.355)	(0.115)	(0.283)	(0.449)	(0.112)	(0.210)	(0.336)	(0.103)	(0.292)	(0.303)
College and Post-Secondary Education	0.352***	0.566	-1.376***	0.311**	0.671*	-1.279**	0.212	-0.643	-0.215	0.045	-0.0228	0.275
	(0.085)	(0.462)	(0.375)	(0.129)	(0.350)	(0.487)	(0.172)	(0.511)	(0.634)	(0.161)	(0.633)	(0.597)
Employment	0.103	0.829	-0.960	0.176*	-0.175	-0.720*	0.0156	-0.0233	0.0265	0.016	-0.004	0.106
	(0.109)	(0.631)	(0.628)	(0.0932)	(0.621)	(0.333)	(0.153)	(0.498)	(0.487)	(0.188)	(0.486)	(0.635)
Risky Behaviors	-0.064	0.130	-0.182	-0.084	0.560	-0.331	0.206**	0.0485	-0.903***	0.260**	0.332	-1.054**
	(0.077)	(0.355)	(0.405)	(0.079)	(0.421)	(0.443)	(0.0670)	(0.254)	(0.199)	(0.086)	(0.213)	(0.345)
Covariates and School Fixed Effects		No			Yes			No			Yes	

Notes: See notes on Table 3. In addition to the covariates described in Table 3, the specifications from this Table include a peers' variable and a variable interacting peers' effects and the treatment dummy. The peers' variable is measured as follows: for each individual we estimated the proportion of students within the same school, treatment status, sex, and race, that reported doing drugs at age 19.

Online Appendix:

School and Drugs: Closing the Gap

Evidence from a Randomized Trial in the United States

By Núria Rodríguez-Planas

TABLE A.1. QOP's Developmental Activities, Community Service Activities and Educational Services

Activity	Examples of such types of activities
Developmental	Life skills activities/ discussion topics (such as, family planning, nutrition, personal hygiene, managing anger, avoiding drug behaviors, among others); pre-employment training; cultural activities; and recreational activities.
Community service	Visits to the residents of a local nursing home, or volunteering at a local food bank.
Educational services	Academic assessment, development of individualized education plans, one-on-one tutoring, and computer-assisted instruction in specific coursework as well as basic reading and mathematics. Making the youth aware of, and helping them plan for, college and other post-secondary education or training.

TABLE A.2. Determinants of Drug Use During High-School in the Control Group, by Gender

Covariates		*Covariates*	
age_13	19.17***	age_14_6	-24.74***
	(1.938)		(2.350)
age_14	4.073*	age_14_7	-15.24***
	(2.303)		(3.833)
grades	-0.0372*	age_14_8	-7.198**
	(0.0218)		(3.351)
sex_age_14	-3.281**	age_14_10	-20.23***
	(1.367)		(2.978)
sex_ihispanic	1.526*	age_15_1	-4.515***
	(0.910)		(0.416)
iblack_age_13	-34.71***	age_15_2	-19.60***
	(4.781)		(4.593)
ihspanic_age_13	-31.93***	age_15_3	-18.84***
	(1.505)		(4.573)
ihspanic_age_14	-0.792	age_15_4	-19.52***
	(2.633)		(4.414)
grades_age_13	-0.0922***	age_15_6	-24.33***
	(0.0303)		(2.096)
grades_ihspanic	0.0333***	age_15_7	-19.84***
	(0.0114)		(4.739)
grades_school5	-0.0630*	age_15_8	-8.719*
	(0.0343)		(4.547)
grades_school6	-0.0141***	age_15_10	-22.16***
	(0.00445)		(4.913)
iblack_ihspanic	-13.61***	black_2	-14.64***
	(1.516)		(1.554)
sex_2	2.662***	black_3	17.80***
	(0.916)		(3.104)
sex_4	1.748*	black_4	18.48***
	(1.057)		(2.450)
sex_5	1.744*	black_7	17.51***
	(1.044)		(3.320)
		black_8	5.741*
sex_6	-1.882***		(3.049)
	(0.344)	black_9	3.988
sex_7	2.178**		(3.523)
	(0.852)	black_10	16.92***
sex_9	2.671**		(3.344)
	(1.342)	hisp_6	27.22***
age_14_2	-18.19***		(3.047)
	(3.435)	hisp_10	18.57***
age_14_3	-21.14***		(1.742)
	(3.054)	school_2	31.36***
age_14_4	-19.14***		(2.567)
	(3.202)		
Observations	381		
Pseudo R-squared	0.2438		

Notes: The table reports logit estimates. Robust standard errors clustered at the school level are reported in parentheses. The estimates in this table are constructed using the sample of control group youths only. School dummies, baseline characteristics and interactions between all of these covariates and the covariates and school fixed effects were included in the specification. Only significant coefficients are displayed.

TABLE A.3. Participation in QOP Activities By Predicted Drug Use at Age 19

				Low-risk			*High-risk*		
	Cumulative Years 1 through 4	*Year 1*	*Year 4*	*Cumulative Years 1 through 4*	*Year 1*	*Year 4*	*Cumulative Years 1 through 4*	*Year 1*	*Year 4*
Average Number of Hours	886	279	125	867	286	113	906	272	138
Average Hours on Educational Activities	245	97	26	245	100	24	246	93	29
Average Hours on Developmental Activities	302	114	27	299	113	24	305	114	30
Average Hours on Community Service Activities	109	37	14	107	37	12	111	37	17
No Hours of Participation (percent)	1	1	40	1	1	36	1	1	44
More Than 100 Hours (percent)	88	74	22	87	73	18	76	76	25
More Than 375 Hours (percent)	60	25	9	57	26	8	25	25	12
More Than 750 Hours (percent)	37	7	4	34	8	3	6	6	5
More Than 1,500 Hours (percent)	17	0	2	17	0	1	0	0	2
Total Amount Earned (dollars)	$1,122	$349	$156	$1,098	$358	$141	$1,146	$340	$172

Note: Because QOP services in Period 5 differed substantially from those of the first four periods, we report hours over the first four periods. In Period 5, QOP offered enrollees who had graduated from high school only mentoring services, and hours spent being mentored were not recorded.

TABLE A.4. Treatment Effect on Risky Behaviors By Predicted Crime at Age 19

	Control Group Means		*Treatment-Control*	
	Low-risk	*High-risk*	*Low-risk*	*High-risk*
OUTCOMES	*(1)*	*(2)*	*(3)*	*(4)*
Youths Were in their Late Teens				
Summary measure			**-0.164***†††**	**0.101*†††**
			(0.038)	**(0.049)**
Binge drinking in the past 30 days	0.106	0.277	0.133**††	-0.048††
			(0.051)	(0.046)
Used any illegal drug in the past 30 days	0.218	0.328	0.083**	-0.049
			(0.031)	(0.029)
Committed a crime in past 12 Months	0.142	0.439	0.164***†††	-0.121***†††
			(0.036)	(0.024)
Ever arrested or charged	0.205	0.353	0.029†	-0.099*†
			(0.034)	(0.053)
Have first child before age 18	0.243	0.237	-0.041	-0.018
			(0.032)	(0.027)
Youths Were in their Early Twenties				
Summary measure			**0.019**	**0.013**
			(0.048)	**(0.056)**
Binge drinking in the past 30 days	0.279	0.308	-0.027	-0.068
			(0.057)	(0.045)
Used any illegal drug in the past 30 days	0.133	0.188	-0.041	-0.051
			(0.036)	(0.050)
Committed a crime in past 3 months	0.052	0.112	0.008	-0.017
			(0.015)	(0.027)
Arrested or charged in past 3 months	0.040	0.044	0.009	-0.002
			(0.026)	(0.020)
Have first child before age 18	0.140	0.158	0.012	0.034
			(0.028)	(0.041)
Currently on welfare	0.227	0.138	-0.018†	0.072**†
			(0.036)	(0.031)
Youths Were in their Mid-Twenties				
Summary measure			**-0.106***	**-0.019**
			(0.050)	**(0.057)**
Binge drinking in the past 30 days	0.325	0.245	-0.005	0.046
			(0.062)	(0.086)
Used any illegal drug in the past 30 days	0.088	0.160	0.038	-0.050*
			(0.034)	(0.027)
Committed a crime in past 3 months	0.0154	0.025	0.008	0.028
			(0.010)	(0.017)
Arrested or charged in past 2 years	0.041	0.064	0.054**	0.032
			(0.017)	(0.031)
Have first child before age 18	0.160	0.166	-0.009	0.036
			(0.025)	(0.029)
Currently on welfare	0.201	0.244	0.054	0.0026
			(0.038)	(0.058)

Notes: See notes on Table 3.

TABLE A.5. Balancing Tests for the Proportion of Drug Users at Age 19 By Predicted Drug Use at Age 19

	MEANS				*TREATMENT: PROPORTION OF DRUG USERS*	
	Low-risk		*High-risk*		*Low-risk*	*High-risk*
OUTCOMES	Drug users	Others	Drug users	Others	*Full set of controls*	*Full set of controls*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Rank based on 8th grade						
Middle third	0.274	0.323	0.344	0.316	-0.050	0.146
					(0.047)	(0.21)
Top third	0.321	0.326	0.270	0.405	-0.006	-0.579
			[-0.135***]		(0.049)	(0.416)
Male	0.660	0.463	0.626	0.491	0.197***	1.993**
	[0.197***]		[0.135**]		(0.044)	(0.780)
Age when entering 9th grade						
14	0.500	0.555	0.540	0.636	-0.055	-0.680*
			[-0.096**]		(0.063)	(0.341)
> 14	0.406	0.323	0.337	0.242	0.082	0.799**
			[0.096**]		(0.053)	(0.348)
Black	0.512	0.617	0.773	0.773	-0.098	-0.314
					(0.071)	(0.360)
Hispanic	0.425	0.335	0.160	0.178	0.089	0.072
					(0.09)	(0.267)
Sample size	106	337	163	269	443	269

Notes: Columns 1 to 4 report means for youths reporting drug use during high-school and those who did not. In brackets statistically significant differences in means between youths reporting drug use and those who do not report drug use while in high-school are reported. Columns 5 and 6 report school FE estimates with robust standard errors clustered at the school level from separate regressions of the relevant variables on the proportion of drug users. The sample in columns 5 and 6 exclude those who reported drug use.

TABLE A.6. Mean Standardized Treatment Effects Under Varying Missing Data Assumptions, High-Risk Group

	Treatment-Control Lower Bounds					*Treatment-Control Upper Bounds*			
OUTCOMES	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*
Youths Were in their Late Teens									
High-school	-0.314***	-0.002	0.040	0.054*	**0.071***	0.082**	0.096**	0.137***	0.480***
	(0.035)	(0.028)	(0.029)	(0.029)	**(0.039)**	(0.030)	(0.030)	(0.031)	(0.042)
College and Post-Secondary Education	-0.295**	0.069	0.131*	0.152**	**0.155***	0.193**	0.213***	0.273***	0.545***
	(0.103)	(0.070)	(0.067)	(0.066)	**(0.084)**	(0.064)	(0.063)	(0.062)	(0.060)
Risky Behaviors	0.488***	0.221***	0.162***	0.142***	**0.103**†††**	0.101**	0.081*	0.020	-0.289***
	(0.053)	(0.047)	(0.045)	(0.044)	**(0.046)**	(0.043)	(0.042)	(0.041)	(0.038)
Youths Were in their Early Twenties									
High-school	-0.253***	0.032	0.084*	0.101**	**0.114*†**	0.135***	0.152***	0.203***	0.453***
	(0.050)	(0.045)	(0.044)	(0.043)	**(0.053)**	(0.042)	(0.042)	(0.042)	(0.050)
College and Post-Secondary Education	-0.464***	0.034	0.130***	0.162***	**0.181****	0.226***	0.258***	0.350***	0.971***
	(0.046)	(0.040)	(0.040)	(0.040)	**(0.062)**	(0.040)	(0.040)	(0.041)	(0.078)
Employment	-0.616***	-0.160**	-0.071	-0.041	**-0.0154**	0.019	0.0493	0.138**	0.644***
	(0.048)	(0.052)	(0.052)	(0.053)	**(0.069)**	(0.053)	(0.053)	(0.054)	(0.051)
Risky Behaviors	0.634***	0.206***	0.115**	0.084**	**0.053††**	0.022	-0.009	-0.101**	-0.714***
	(0.040)	(0.038)	(0.037)	(0.037)	**(0.044)**	(0.037)	(0.037)	(0.037)	(0.089)
Youths Were in their Mid-Twenties									
High-school	-0.222***	0.052*	0.100***	0.116***	**0.126***†††	0.148***	0.163***	0.210***	0.430***
	(0.0532)	(0.027)	(0.026)	(0.027)	**(0.035)**	(0.029)	(0.030)	(0.035)	(0.065)
College and Post-Secondary Education	-0.500***	-0.011	0.0850*	0.117**	**0.125****	0.181***	0.213***	0.306***	1.237***
	(0.0557)	(0.041)	(0.038)	(0.038)	**(0.046)**	(0.037)	(0.037)	(0.038)	(0.147)
Employment	-0.637***	-0.110*	-0.007	0.0271	**0.0469**	0.096*	0.131**	0.233***	1.101***
	(0.065)	(0.052)	(0.050)	(0.050)	**(0.0557)**	(0.049)	(0.049)	(0.049)	(0.101)
Risky Behaviors	0.575***	0.119**	0.023	-0.010	**-0.039**	-0.075	-0.107**	-0.203***	-0.971***
	(0.060)	(0.052)	(0.048)	(0.047)	**(0.055)**	(0.046)	(0.046)	(0.045)	(0.112)

Notes: Sample size is 1,069 observations when bounds are used to estimate the impact. Column (5) presents estimates discussed in the main text and presented in Tables 3 to 6. Columns 1 and 9 present the lower and upper bounds obtained under the worst-case scenario. For the lower (upper) bound, we impute the minimum (maximum) value of each variable in the observed treatment distribution to the non-responders in the treatment group, and the maximum (minimum) value of the observed control distribution to the non-responders in the control group. Columns 2 and 8 impute to the lower (upper) bound the mean minus (plus) 0.25 standard deviations of the observed treatment distribution to the non-responders in the treatment group, the mean plus (minus) 0.25 standard deviations of the observed control distribution to non-responders in the control group. Columns 3 and 7 (and columns 4 and 6) repeat the previous scenario but with a 0.1 (0.05) standard deviations.

*, **, *** Estimate significantly different from zero at the 90%, 95%, or 99% confidence level.

TABLE A.7. Treatment Effect on High-School Outcomes Using Complete and Attrited Samples

	COMPLETE SAMPLE		*ATTRITED SAMPLE*	
	Low-risk	*High-risk*	*Low-risk*	*High-risk*
OUTCOMES	*Full set of controls*	*Full set of controls*	*Full set of controls*	*Full set of controls*
	(2)	*(3)*	*(5)*	*(6)*
Youths Were in their Late Teens				
Obtained high-school	0.008	0.082*	-0.005	0.076*
Diploma	(0.038)	(0.041)	(0.045)	(0.043)
Obtained a GED	-0.008	-0.006	-0.005	-0.011
	(0.027)	(0.028)	(0.022)	(0.027)
Sample size[a]	511	497	453	439
Youths Were in their Early Twenties				
Obtained high-school	-0.038††	0.088**††	-0.025††	0.076††
Diploma	(0.045)	(0.031)	(0.046)	(0.046)
Obtained a GED	-0.000	0.016	-0.015	0.023
	(0.033)	(0.036)	(0.032)	(0.037)
Sample size	476	466	390	396
Youths Were in their Mid-Twenties				
Obtained high-school	-0.045††	0.079**††	-0.026††	0.100**††
Diploma	(0.048)	(0.034)	(0.046)	(0.045)
Obtained a GED	-0.035	0.033	-0.059	0.019
	(0.036)	(0.0367)	(0.036)	(0.036)
Sample size	463	456	399	393

Notes: The table reports estimates of treatment effects on the dependent variables indicated in row headings. Robust standard errors clustered at the school level are reported in parentheses. "Full set of controls" includes school dummies, an indicator for being male, an indicator for being 14 years old when entering ninth grade, an indicator for being over age 14 when entering ninth grade, an indicator for being in the middle third of the eighth-grade GPA distribution, an indicator for being in the top third of the eight-grade GPA distribution, an indicator for being black and an indicator for being Hispanic.

[a] Sample size differs from that of Table 4 in the main paper because when estimating summary indices if an individual has a valid response to at least one component measure of an index, then any missing values for other component measures are imputed at the random assignment group mean (as in Jeffrey Kling, Jeffrey Liebman, and Lawrence Katz 2007).

*, **, *** Estimate significantly different from zero at the 90%, 95%, or 99% confidence level.

†, ††, ††† indicates that the difference of the estimated effects between youths in the bottom- and top-half of the predicted drug use distribution is significant at the 90%, 95%, and 99% level.