

Mentoring, Educational Services, and Economic Incentives Longer-term Evidence on Risky Behaviors from a Randomized Trial

Núria Rodríguez-Planas¹

Universitat Autònoma de Barcelona,
IZA and FEDEA

May 2010

Abstract

This paper is the first to use a randomized trial in the US to analyze the short- and long-term impacts of an afterschool program that offered disadvantaged high-school youth: mentoring, educational services, and financial rewards to attend program activities, complete high-school and enroll in post-secondary education on youths' engagement in risky behaviors, such as substance abuse, criminal activity, and teenage childbearing. Outcomes were measured at three different points in time, when youths were in their late-teens, and when they were in their early- and their late-twenties. Overall the program was unsuccessful at reducing risky behaviors. Heterogeneity matters in that perverse effects are concentrated among certain subgroups, such as males, older youths, and youths from sites where youths received higher amount of stipends. We claim that this evidence is consistent with different models of youths' behavioral response to economic incentives. In addition, beneficial effects found in those sites in which QOP youths represented a large fraction of the entering class of 9th graders provides hope for these type of programs when operated in small communities and supports the hypothesis of peer effects.

Key words: After-school program, short-, medium- and long-term effects, behavioral models, peer effects, criminal activity, teen childbearing and substance abuse.

JEL classification: C93, I21, I22, I28, J24

¹ The Quantum Opportunity Program demonstration study was conducted under Contract No. K-5547-5-00-80-30 from the U.S. Department of Labor or the Ford Foundation. The views expressed herein do not necessarily reflect the policies or opinions of the U.S. Department of Labor. This project would not have been possible without the leadership, support and contribution of Allen Schirm, Myles Maxfield, Elizabeth Stuart, Allison Mckie, and Laura Castner from Mathematica Policy Research, as well as Vida Maralani from UCLA, and Mary Vencill from Berkeley Policy Associates. I also welcome the useful comments and advice in the analysis and the exposition of the paper from Samuel Bentolila, Caterina Calsamiglia, David Frisvold, Michael Lechner, Jungmin Lee, Robert Moffit, Peter Mueser, Kenneth Troske, Sunde Uwe, Conny Wunsch, and participants at seminars at Saint Gallen, Switzerland; CEMFI, Spain; ZEW Mannheim, Germany, and the 2nd Annual Meeting of the Economics of Risky Behaviors in Atlanta. Author's e-mail: Nuria.Rodriguez@uab.es

I. Introduction

Many adolescents engage in a host of risky behaviors that might have important present and future ramifications. One of the frequently mentioned reasons for engaging in risky behaviors is the youth's weak non-cognitive skills, such as low self-esteem, lack of motivation and tenacity, poor trustworthiness, and little perseverance. From psychologists' perspective, building strong positive relationships with extra familial adults (mentors) promotes resiliency among at-risk youth because mentoring facilitates adolescents' capacity to benefit from the support of parents and other providers, and influences positively the youth's perceptions of self-worth and their beliefs about their competence as learners and their valuing of school (Rhodes, Grossman, and Resch, 2000). Rigorous studies on the effectiveness of mentoring programs provide evidence of their success in promoting better social, academic, and behavioral outcomes (Grossman and Tierney, 1998; DuBois *et al.*, 2002a, b; Jekielek *et al.*, 2002; Keating *et al.*, 2002; Herrera *et al.*, 2007; Karcher, 2005; DeWit *et al.*, 2006; and Rhodes, 2008).

Using a randomized experimental design, this study evaluates the short-, medium-and long-term effects of a five-year after-school program, the Quantum Opportunity Program (QOP hereafter), on risky behaviors in the United States. QOP involved the combination of the following three services: a mentoring, an educational, and a financial incentives component. Although the main objective of the program was to improve youths' educational achievement, QOP also had a highly relevant secondary objective of reducing risky behaviors such as substance abuse, crime, and teenage childbearing. This paper focuses on QOP's success in achieving its secondary objective.² The outcome variables were measured at three points in

² Rodriguez-Planas, 2010, analyzes QOP's impacts on education and employment outcomes. The author finds hefty beneficial average impacts when youth are in their late-teens that quickly faded away once youth reach their twenties. In addition, the results reveal that heterogeneity matters. While encouraging results are found for younger youth, and when the program is implemented in relatively small communities of 9th graders; detrimental long-lived

time: (i) during the fifth year of the demonstration while the students were still in, or just completing, high school; (ii) three years later when most sample members were about 21 or 22 years old; and (iii) five years after the end of the demonstration when most sample members were about 24 or 25 years old. To examine the effectiveness of this program, the Department of Labor and the Ford Foundation funded this demonstration in seven sites across the United States between 1995 and 2001.

This paper presents impact findings from this evaluation.³ The results are not very promising. The evidence suggests that, on average, QOP increased substance abuse during the fifth and last year of the program, and it rose criminal activity 5 years after the end of the program. Heterogeneity matters in that perverse effects are concentrated among certain subgroups, such as males, older youths, and youths from sites where youth received higher amount of stipends. These and additional empirical findings are hard to reconcile with simple economic models. We explore potential explanations, and, in our view, the most parsimonious explanations are those allowing for youths' behavioral response to economic incentives, including: (i) the hypothesis of deterrence (Becker, 1968); (ii) the perverse effects of raising the perception of riskiness discussed by O'Donoghue and Rabin, 2000, 2001; and (iii) youth's response to economic incentives (Nisbet and Vakil, 1972; Cook, 1981; Cook and Tauchen, 1982; Saffer and Grossman, 1987a, 1987b; Coate and Grossman, 1988; Kenkel, 1993; Cook and Moore, 1994, 2001; Grossman et al., 1987, 1994; Chaplounka and Grossman, 1996; Ruhm, 1996; Evans and Huang, 1998; Markowitz and Grossman, 1998; Gruber, 2001; Gruber and Zinman, 2001; Pacula *et al.*, 2001; Levine, 2001; Ruhm, 2005; Ruhm and Black, 2005; and Grossman,

outcomes are found for males, and when case managers are partially compensated by incentive payments and students receive more regular reminders of incentives.

³ See Maxfield *et al.*, 2003a and 2003b; Schirm *et al.*, 2003; Schirm and Rodríguez-Planas, 2004; and Schirm *et al.*, 2007 for detail description of program design and implementation as well as thorough analysis of the demonstration's impacts.

2005). In addition, beneficial effects found in those sites in which QOP youths represented a large fraction of the entering class of 9th graders provides hope for these type of programs when operated in small communities. This finding combined with the detrimental effects found in those sites in which QOP youths represented a small fraction of the entering class, suggest a strong relationship between peer effects and social pressure (as suggested by Mas and Moretti, 2009).

There are, at least, three novel contributions in this paper. First, it measures the effects of the intervention not only in the short-term, but also in the medium- or longer-term (up to five years after the end of the program). While most interventions aiming at improving disadvantaged youths' academic or emotional behavior analyze the effects of the interventions during or shortly after the students have been exposed to the program, the evidence on the medium- or longer-term impacts is very scarce.⁴ However, knowing the long-term impacts of these interventions is key to disentangle the following questions: Do the short-term changes generated by the intervention persist? Do they quickly fade away? Or do they reverse? Second, this paper measures behavioral outcomes in a variety of ways, including substance abuse, criminal activity, teenage childbearing, and several measures of family life and physical and mental well-being. Moreover, educational and employment outcomes, while not the focus of this paper, are also available and discussed when considered relevant (see Rodriguez-Planas, 2010, for thorough discussion on QOP's impacts on education and employment outcomes). To the best of my knowledge, most papers in the risky behaviors literature only observe (and, thus, focus) on

⁴ For instance, in the field of educational interventions whose main objective is to improve the school performance of disadvantaged youth, the following papers look at longer horizons as in the present paper: Leuven *et al.*, 2003, Bettinger and Long, 2005, Angrist and Lavy, 2009, Galárraga and Gertler, 2009, and Jensen and Lleras-Muney, 2010. Other well-know educational innovations that follow-up on the long-term outcomes of the subjects of the experiment include STEP (Grossman and Sipe, 1992); the Abecedarian Project, the Perry Preschool Program, and the Early Training Project (Barnett, 1992; and Anderson, 2008); and Project STAR (Krueger and Whitmore, 2001). Finally, Galárraga and Gertler, 2009, and Jensen and Lleras-Muney, 2010, analyze longer-term impacts of a conditional-cash-transfer program on risky behaviors in Mexico.

one or two types of behavior (such as substance abuse, criminal activity, or pregnancy), limiting the multi-angle analysis of risky behavior among youths.⁵ Finally, this paper adds to the growing interest in student incentive programs in primary and secondary education.⁶ While most of these studies have focused on the educational effects of financial rewards, very few have measured the effects on youths' risky behaviors.⁷ Most of the experimental or quasi-experimental interventions on the effects of financial rewards on risky behaviors have taken place in developing countries and they have, overall, found that they improve healthy behaviors (Gertler, 2004; Hoddinott and Skoufias, 2004; Schultz, 2004; Behrman and Hoddinott, 2005; Behrman *et al.*, 2005; Fernald *et al.* 2008a; Fernald *et al.* 2008b; Fernald *et al.* 2008c, and Garrálaga and Gertler, 2009). However, given that the social, environment, and cultural issues are drastically different between at-risk youths in developing and developed countries, it is not clear that the existent results will generalize to the developed world, in general, and to the United States, more specifically.

This paper is organized as follows. The next section describes the program implementation and presents the data. Section III explains the evaluation framework and analyzes the results, and Section IV discusses possible explanations for these results. Section V

⁵ Studies with information on a couple of youth risk-taking behavior include Rees *et al.*, 2001; Sen, 2002; Rashad, and Kaestner, 2004; Grossman and Markowitz, 2005; Cardoso and Verner, 2008; Garrálaga and Gertler, 2009; and Jensen and Lleras-Muney, 2010.

⁶ Such types of interventions are currently being offered in many countries. For instance, three high-schools are offering such types of incentives in Créteil, a suburb of Paris, France (El País, 2009). In the United States, the Education Innovation Laboratory at Harvard is in the process of implementing four evaluations offering financial rewards with the objective of improving academic performance of youth, such as Capital Gains in Washington DC, Spark and Million Motivation Campaign in New York, and The Paper Project in Chicago (Fryer, 2010). Other interventions involving incentive programs in primary and secondary education are also taking place in the Baltimore City Public School District (Ash, 2008); and in Texas (Jackson, 2007). Finally, other related international examples include Progresia in Mexico (Behrman *et al.*, 2000; and Schultz, 2004), Familias en Acción program in Colombia (Attanasio *et al.*, 2006), or the Education Maintenance Allowance in the United Kingdom (Dearden, Emmerson, Frayne, and Meghir, 2005).

⁷ Randomized evaluations studying the educational effect of financial rewards include Leuven *et al.*, 2003; Kremer *et al.*, 2008; Angrist *et al.*, 2009; Angrist and Lavy, 2009, and Rodriguez-Planas, 2010.

concludes with a discussion on alternative explanations and lessons learned. An additional appendix provides further information on the data and detailed empirical results.

II. The Quantum Opportunity Program Demonstration Project

Program Description

QOP differed from other interventions aiming at improving disadvantaged youth educational outcomes in at least two important ways. First, it was an intensive, long-term after-school program aiming to overcome the many serious challenges facing disadvantaged youth. It lasted five years and was offered year-round to low-achieving students from low performing high-schools entering in 9th grade in 1995 in the United States.⁸ It should therefore not come as a surprise that it was an expensive program. At almost \$25,000 per enrollee for the whole demonstration, QOP has been the most expensive Federal youth program offered. By comparison, the operating costs of the also-expensive Job Corps were approximately \$16,500 per participant in 1998 (Schochet *et al.*, 2008).

Second, QOP offered more comprehensive services than other programs. While most programs offer mentoring, educational services, or financial rewards, QOP offered all these services combined. As we shall discuss below, although the core of the QOP model was intensive case management and mentoring, its educational and development services, on the one hand, and the financial rewards, on the other, were similar in design and intensity as those (currently) implemented in other evaluations.

Case managers had small caseloads of only 15 to 25 youth and were to develop with each youth a highly personal, long-lasting connection that mirrored the relationship between a

⁸ Enrollees who graduated on time received some mentoring and assistance in enrolling in postsecondary education or training between graduation and the end of the fifth year of the demonstration.

teenager and a nurturing, supportive older relative. As such, the case manager would make every effort to sustain a strong relationship with the youth regardless of behavior or status, including if 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 for addressing all barriers to success that enrolled youth faced, whether related to school, family, or friends.⁹

In addition to case management, the program engaged youth in: (i) developmental activities that aimed to develop their social and employment-readiness skills; (ii) community service activities to develop a sense of community belonging, trustworthiness, and respect; and (iii) educational services to improve their academic performance. Examples of such types of activities are displayed in Table 1.

Finally, QOP also offered financial incentives to students. Such incentives had two components: an incentive on inputs and another one on outputs.¹⁰ First, youth received a stipend of \$1.25 for every hour devoted explicitly to educational activities, developmental activities (excluding recreational activities), and community service. This component is equivalent to providing incentives on inputs rather than outputs, which is more common in the student incentives literature.¹¹ Second, a matching amount was promised to the youth when he or she

⁹ 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.

¹⁰ As Fryer, 2010, explains, under certain assumptions, traditional price theory predicts that providing incentives based on output is socially optimal because each student decides which input from their production function to subsidize. Assuming that student's have superior knowledge about how they learn, it is socially optimal to allow them to allocate their time across inputs. However, if this assumption is violated, then it can be more effective to provide incentives for inputs. Fryer, 2010, finds that incentives for output did no increase achievement, while incentives for certain inputs did. He explains that the leading theory behind these findings is that students do not understand the educational production function and, thus, lack the know-how to translate their excitement about the incentive structure into measurable output.

¹¹ In Kremer *et al.*, 2008; and Angrist and Lavy, 2009; and most other studies on student incentives, the objective is to pass a test and students are paid if they complete the objective. Fryer, 2010, measures the effect of four different financial incentives on student achievement: two of them are "output" experiments (the ones in Chicago and New York city) and the other two (in Dallas and in Washington, DC) are "input" experiments.

earned a high school diploma or GED *and* enrolled in post-secondary education or training (including vocational training or military service).¹² This was clearly an incentive on output as the student received the economic incentive only if they graduated from high-school *and* enrolled in post-secondary education. As explained in the implementation sub-section below, by the end of the demonstration, this represented for most youths receiving between \$1,000 to \$3,000 after high-school graduation and enrollment in post-secondary education. Although some may question whether QOP's financial rewards were sufficiently large, numerous studies examining the impact of various types of tuition and financial aid policies on college-going show that students respond to changes in college cost (Leslie and Brinkman, 1988; Cornwell, Mustard, Cameron and Heckman, 1993; and Kane, 1998; Sridhar 2006; Dynarski, 2003; and Deming and Dynarski, 2009). A consensus estimate associates a \$1,000 change in college costs with an approximately 5 percentage point difference in college enrollment rates. Moreover, according to a recent study by Kane, 2007, there would be differential effects by race, being stronger for African American.

Target Population and Sample Selection

In the summer of 1995, QOP was implemented in eleven high schools across seven sites in the United States.¹³ In each of these schools, entering 9th grade students—except those with a GPA

¹² This is a similar design to the one currently applied in The Paper Project, which rewards high-school students for core class grades, in that half of the reward is given to the student immediately, the other half is distributed at graduation.

¹³ DOL awarded demonstration grants to implement this QOP model in five sites: Cleveland, Ohio; Fort Worth, Texas; Houston, Texas; Memphis, Tennessee; and Washington, DC. The Ford Foundation funded two sites: Philadelphia, Pennsylvania and Yakima, Washington. Six of the seven demonstration sites operated QOP between 1995 and 2000; the Washington, DC, site began one year later and operated the program through summer 2001. A local community-based organization (CBO) implemented the QOP model in each site. Each CBO teamed with one, two, or three high schools for a total of eleven high schools participating in the demonstration. Each program enrolled 50, 80, or 100 students.

from 8th-grade above the 66th percentile—, were randomly assigned to QOP or a control group.¹⁴ Youth assigned to the program group were enrolled in QOP. Youth assigned to the control group could not enroll in QOP, but could enroll in other youth programs offered in the community. Thus the counterfactual is *other* available programs that the study population would enroll in if QOP were not an option. Random assignment involved four steps as described in Table 2. The final sample for the QOP demonstration consists of 1,069 students, 580 in the QOP group and 489 in the control group.¹⁵ As expected, random assignment produced treatment and control groups whose distributions of characteristics prior to random assignment were similar. There were few significant differences by status on baseline (pre-random assignment) characteristics, as shown in Table 3 below. The only statistically significant difference was the proportion of youth in the middle third of the distribution, which was a bit larger for the control group.

As discussed with more detail at the end of Section IV, spill-over effects did not seem to be an issue in this intervention for the following two reasons. First, although QOP provided tutoring and computer-assisted instruction to its enrollees, it was not designed to influence the structure, policies, or operation of the high schools with which local QOP programs were associated. Second, QOP did not operate within the school or within school hours. Instead it was an after-school program, and its activities were scheduled *outside* the high-schools from 3 to 6 pm during weekdays, and for one half day over the weekend. That said, QOP could still have

¹⁴ The target group in the QOP demonstration was youth who met four eligibility criteria: (1) were beginning ninth grade in a high school selected for the QOP demonstration, generally schools with dropout rates of 40 percent or more; (2) were not repeating ninth grade; (3) were not so disabled that the school viewed participation in the program as inappropriate; and (4) had a grade point average (GPA) from the eighth grade below the 67th percentile among the students in the school meeting the first three requirements.

¹⁵ The reason the final sample includes 580 youths in the treatment group and 489 in the control group was *not* due to a failure in the randomisation process. Instead it was due to the design of the evaluation sample, in which 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). That said, within each school, eligible youths whose parents had given consent to participate in the program were randomly assigned either to the treatment or the control group.

stimulated control group members to work harder because they knew that some of their classmates were receiving additional help. In the results section, we discuss why we do not think this may have occurred.

The database

One of the highlights of the QOP demonstration was its intense data collection. The evaluation was designed to estimate the short-, medium-, and long-term impacts of the program by collecting survey data on youth's outcomes at four different points in time, as shown in Table 4. Most of the analysis is based on the paper survey conducted during the spring of 1999, and the two telephone surveys, conducted two and five years *after* the end of the program.

In addition, by the usual standards for observational evaluation studies, this data set is exceptionally rich and informative, as it contains information on various substance abuse, criminal activity, information on arrests, conviction, and time spent in jail, teenage pregnancy, family composition, social assistance dependency, self-reported health condition, resiliency factors, and educational and employment outcomes (see Maxfield *et al.*, 2003b; Schirm and Rodríguez-Planas, 2004; Schirm *et al.*, 2006; and Rodríguez-Planas, 2010, for thorough analysis of QOP's impacts on education, achievement tests, and employment outcomes). Finally, data on program implementation, participation, and costs, as well as (baseline) information on the youth (including their 8th-grade GPA) and high-school transcripts from *all* the high schools a sample member attended were obtained.

There are three drawbacks with the data at hand. First, due to data collection costs, no baseline survey was collected. As a consequence, the pre-program information available is reduced to the characteristics displayed in Table 3. Albeit the limited baseline information available, it is important to highlight that 8th-grade GPA, which is a good proxy for youth's

cognitive and non-cognitive skills, as well as unobserved ability, is available.¹⁶ Second, as these are (mainly) survey data, differential non-response between treatment and control group members could potentially bias the results.¹⁷ All of the estimates in the paper adjust for survey non-response to ensure that the sample of respondents reflects the experiences of all sample members, as explained in the Appendix and in Maxfield, *et al.*, 2003b; Schirm *et al.*, 2003; and Schirm *et al.*, 2006. In addition, a thorough sensitivity analysis on whether (and if so how) non-response may be affecting the results is also provided at the end of the Results' Section. Overall, I do not find evidence that differential non-response bias between treatment and control group is driving the results. Third, data on participation were recorded for the purpose of computing periodic stipend payments and accrual contributions for each enrollee, making its research use difficult for the following two reasons. First, given that mentoring time did not count toward stipends or accrual account contributions, data on time spent being mentored were not recorded. Second, bonus hours (given for achieving significant milestones) could not be distinguished from regular hours, and thus result in overestimates of the amount of time spent on program activities for some enrollees. While this prevents me from using participation hours to apply quasi-experimental methods to estimate the impacts by subgroups based on their predicted probability of participating, analysis of these data does provide some reliable information on how much these services were taken, and who was likely to be taking them, as discussed below.¹⁸

¹⁶ The main pre-program 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.

¹⁷ The survey effort is equiparable to that of other studies (Schochet *et al.*, 2008; Banerjee *et al.*, 2007; and Kremer *et al.*, 2008, among others). As such, the response rate to the in-person survey and the 1st telephone surveys was 84%. For each of the surveys, the response rate for the QOP group exceeded the response rate for the control group by 7 percentage points. The response rate to the 2nd and 3rd telephone survey was 75% and 76%—80% (77%) for the QOP group and 70% (74%) for the control group in the 2nd (3rd) telephone survey.

¹⁸ As explained in Section III.1, because the decision to participate in the different activities may be related to potential outcomes, all of the analysis is based on randomly-assigned intention to treat.

Implementation of QOP and Service Use

In many respects, the sites were successful in implementing QOP. As designed, QOP served youth who faced many barriers to academic success. In addition, the core component of QOP—case management and mentoring—was also well implemented across the sites.¹⁹ Case managers were hired for the whole length of the program and with prior expertise on social services.²⁰ Most of them reported developing close mentoring relationships with the majority of the youth assigned to them, and they all provided access to services regardless of an enrollee’s behavior or status (such as becoming incarcerated, moving to another community, or dropping out of high school) as originally planned by the program.

Although the educational, community services, and development activities component fell short of the targeted original design, the participation achieved was still a substantial investment of time—especially compared to other similar youth programs.²¹ The initially planned target consisted of 750 hours of services annually (equally distributed among the three different activities), which (if achieved) would have represented about three-quarters of the hours required for in-school instruction per year.²² As it was, the average amount of time (708 hours) enrollees spent on QOP activities during the first four years—including summers—corresponds to about 72% of an extra school year, still a substantial investment of time (in addition to the time spent with the mentor), as shown in Table 5.²³ Finally, the fact that QOP did not achieve its

¹⁹ For a thorough description on how well these and other features of the QOP model were implemented in the demonstration sites see Maxfield *et al.*, 2003a.

²⁰ Most case managers stayed with the program for several years, and many stayed for the entire five years of the demonstration. Unfortunately, no information on sex, race or ethnicity of mentors was collected.

²¹ For instance, the average participation in QOP activities (excluding mentoring) was more than half of the average instruction time received by Job Corps participants, the (by far) most intense education and training program for disadvantaged youths in the United States (Schochet *et al.*, 2008).

²² In 2000, the average number of instructional hours spent in public school by 15-year-old youth was 990 hours (U.S. Department of Education 2005; Table 26-2).

²³ Because QOP services in Period 5 differed substantially from those of the first four periods, we report trends over the first four periods. In Period 5, QOP offered enrollees who had graduated from high school only mentoring

extremely ambitious target is not a concern in terms of the external validity of this evaluation as if the program were to be implemented on a broader scale, it is likely that its implementation would not differ much from how it was implemented during the demonstration.

As explained there were two components of the financial incentives: the one that rewarded program participation, and the one that rewarded high-school completion (including obtaining a GED) *and* post-secondary enrollment. The enrollee stipends were well implemented and appeared to be an effective way to attract the enrollees to program activities in the first year or two of the demonstration. As enrollees aged and could earn much more per hour by working, case managers found that other incentives, such as recognition, attention, and prizes, could replace the stipends. As explained earlier, the financial incentive to outputs was quite successful in that, by the end of the demonstration, enrollee's accrual account balances ranged from a few hundred dollars to nearly \$10,000, with most being in the range of \$1,000 to \$3,000.²⁴ The size of this incentive is comparable to the ones currently being offered in ongoing evaluations, such as, Capital Gains, where the average student will earn \$750 per year; Spark, where 7th graders can earn up to \$500 per year; or The Paper Project, where the average student will earn \$800 per year (up to a maximum of \$2,000 per year).

Table 5 shows that enrollees spent an average of 76 hours per year on education, 77 hours on developmental activities, and 24 hours on community service (in addition to time they spent with their mentor).²⁵ Not surprisingly, the average time spent on QOP activities fell steadily from 247 hours in the first year of the demonstration to 103 hours in the fourth year. About 30% (20%) of those who had participated in QOP activities early during the demonstration and then

services and hours spent being mentored were not recorded.

²⁴ Final payments were made directly to the enrollee rather than to the postsecondary institution or to the enrollee's parents.

²⁵ 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 needed other QOP services more. Most sites decided to reallocate their resources away from community service to mentoring, case management, and educational activities.

stopped participating before the end of the fourth year reported to do so because they left high school (worked). Similarly, among the reasons given for reducing participation in QOP activities over time were having a job (40%), family responsibilities—including own child—(20%), and other after-school activities—such as sports—(13%).²⁶ It is important to highlight, however, that almost the totality of QOP youth engaged in QOP activities, as all but 1% of enrollees spent some time on QOP activities in the first year. As youth grew up, those not engaging in QOP activities in the fourth year amounted to 26%. Among the reasons given for not participating more in QOP activities were the lack of interest (25%) and their time commitment to a job (15%).

Analysis of baseline characteristics of QOP enrollees with higher and lower levels of participation reveals that those who attended more QOP activities during the demonstration tended to have higher grades at baseline, and be age 14 or younger upon entering the 9th grade than those with lower participation.²⁷ In addition, males were more likely to be among the heavy users (as measured by participating more than 1,500 hours during the demonstration) and the light users (as measured by participating 100 or fewer hours).

III. Results

III.1. Evaluation Framework

The estimates reported below are intention-to-treat effects that make no adjustments for remaining involved or service participation in QOP. They were estimated by computing differences in mean outcomes between the youth who were assigned to the QOP group

²⁶ This information was retrieved from the paper survey that was taken at the time youth were between 18 and 19 years old.

²⁷ Notice that caution is needed when trying to infer from these results as it is likely that bonus hours for good grades may well be concentrated among the more able youth, that is those with higher 8th-grade GPA.

(regardless of whether they remained involved in QOP and of how much they participated in QOP activities) and the youth in the control group. Weights were used in all calculations to adjust for the survey non-response and sample design, as explained in the Appendix. In addition, a thorough discussion on the sensitivity analyses performed is also available at the end of Section IV. More specifically, the results are robust to (i) excluding one site at the time (Tables A.2); (ii) using regression models to control for baseline characteristics (Table A.3); and (iii) non-response bias (Tables A.4 and A.5). Overall, the results are robust to the various sensitivity tests performed.

When obtaining multiple impact estimates, there is a concern that some estimated impacts will be found to be significantly different from zero, even if there is actually no impact of QOP (a “Type 1” error). For simplicity, in the tables presented, we do not adjust for the multiple comparisons being done, but we checked whether the results were sensitive to such adjustments using two methods, the Bonferroni correction and a more powerful adjustment developed by Benjamini and Hochberg, 1995. Overall, results were not sensitive to such adjustments.

III.2. Measurement of Outcomes on Risky Behaviors, Physical and Mental Well-Being and Family Life

The analysis in this paper focuses on three types of risky behaviors: substance abuse, criminal activity, and teenage childbearing. In addition, several measures of family life and physical and mental well-being are also reported. The outcomes on risky behaviors have been measured at three different points in time: (i) during the fifth year of the demonstration while the students were still in, or just completing, high school; (ii) over seven years after the start of the program—or over two years after the end of the program; and (iii) ten (five) years after the start (end) of the demonstration. Measures of family life and physical and mental well-being were

also collected while youth were in their mid- and late-twenties. Treatment and control group means for the different outcomes are summarized in Appendix table A.1.

III.3. Impact Results for the Full Sample

QOP emphasized mentoring and offered developmental activities, in part, to reduce the likelihood that enrollees would engage in risky behaviors such as substance abuse, crime, and teenage childbearing. Unfortunately, we found that QOP generally did not achieve this objective. Instead, the evidence suggests that, on average, QOP increased substance abuse during the fifth and last year of the program, and it rose criminal activity 5 years after the end of the program.

Table 6 presents average estimates of the effect of QOP measured at three points in time. When most enrollees were in their late teens, QOP did not significantly reduce any self-reported risky behavior, and it significantly increased substance abuse (column 1). More specifically, in the 30 days before the in-person survey, QOP enrollees had significantly higher rates of drinking (40 percent of the QOP group and 33 percent of the control group) and illegal drug use (34 percent of the QOP group and 28 percent of the control group). By the time most sample members were in their early twenties, the only indication of a beneficial impact of QOP on risky behaviors was that QOP enrollees were less likely to use illegal drugs (column 2 of Table 6). However, results based on the same survey found that QOP did not significantly reduce the likelihood of binge drinking, committing a crime, being arrested or charged with a crime, or having a child before the age of 18.²⁸ By the time QOP enrollees were entering their mid-

²⁸ As explained below the insignificant beneficial effect of QOP on binge drinking when youth were in their early twenties is mainly driven by a suspiciously low probability among QOP youth from the Philadelphia site. Excluding this site leads to no impact on binge drinking (the impact drops to -1 percentage points and remains statically insignificant).

twenties, we found that QOP did not reduce substance abuse (including tobacco use), as shown in column 3 of Table 6. In contrast, we did find that QOP enrollees had higher rates of criminal activity and arrests than youth in the control group. QOP led to a statistically significant increase in the percentage of sample members who committed a crime in the 3 months prior to the survey (5 percent of the QOP group versus 2 percent of the control group) and the percentage of sample members who were arrested or charged with a crime in the 2 years prior to the survey (11 percent of the QOP group versus 5 percent of the control group). QOP also increased by 5 percentage points the percentage of committing a crime in the two years before the survey, but this impact was not statistically significant. Although the detrimental impacts on criminal activities become insignificant when we adjusted for random baseline differences in a regression (Appendix Table A.3), the coefficients were not reduced, and therefore, it is likely that the lack of precision is due to the small size of our sample (combined with the unlikely event of criminal activity). Subgroup analysis in the next section will enable us to explore further whether the detrimental effects of QOP held across different subgroups, or whether it is just a statistical artifact.

It is possible that the case management, mentoring, and development activities that were undertaken to reduce risky behaviors might have improved the family lives of enrollees by, for example, fostering the attitudes and interpersonal skills that promote better relationships with spouses, significant others, and children. We find, however, that QOP did not reduce the likelihood of having a child before age 18, being a single parent, or having a child with whom the enrollee is not living. We also find that QOP did not decrease the likelihood that an enrollee lives in a household that receives public assistance or that an enrollee has poor self-reported health.

III.4. Heterogeneity among Individuals

Although the program's average effects paint a gloomy picture, subgroup analysis may reveal where and for whom QOP had these perverse effects. Tables 7 through 10 show subgroup program impact estimates by: (i) age at 9th grade, (ii) gender, (iii) youth's 8th-grade GPA, and (iv) sites' funding source. Impacts for subgroups defined by youth socio-economic characteristics were estimated by comparing the average outcomes of QOP and control group members in the subgroup of interest. The discussion of the results presented below focus on those impacts that are significantly different from zero, unless otherwise noted.

Impacts by Age when Entering 9th Grade

Most of the youth in the QOP demonstration were 13 to 15 years old when the demonstration began. Youth aged 14 years old or younger (about two thirds of QOP enrollees) were those who had begin 9th grade on time, whereas youth older than 14 years old had been most likely previously held back in school (although not in the 9th grade). Table 7 considers the effects of heterogeneity with respect to age by dividing youth into two groups, based on whether they were older than 14 years old when they entered 9th grade or not. Below, we discuss the key findings.

When youth were in their late teens, we found that QOP was beneficial in terms of significantly reducing the likelihood of having a child before age 18 for the younger enrollees. As such, QOP decreased by 9 percentage points the fraction of younger enrollees who had a child. This impact was significantly different from both zero and the (insignificant) six-percentage-point increase in the fraction of older enrollees who had a child before age 18. However, as many of the younger enrollees were not yet 18 years old when they responded to the questionnaire, the beneficial result on younger enrollees needs to be taken with caution. In

effect, by the time of the 2nd survey, when youth are in their early twenties, any beneficial effect of QOP on preventing teenage childbearing has vanished.

The other beneficial effect of QOP observed during the fifth and last year of the program, a significant 11 percentage points reduction in the likelihood of being arrested in the 12 months prior to the survey for the older enrollees, also fades away by the time youth are in their early twenties. More concerning is that, by the time youth are in their mid-twenties, QOP enrollees are 9 percentage points more likely to have been arrested or charged in the two years prior to answering the 3rd telephone survey. This coefficient is statistically significantly different from zero, but not different from the also significant detrimental effect found among the younger enrollees. For the latter, we also find that they are 3 significant percentage points more likely to have been convicted or pled guilty.

Finally, QOP also seems to have had a detrimental impact for the older enrollees on their family life as it statistically increased by 12 and 11 percentage points the likelihood of being a single parent when youth were in their early-twenties and of having a child with whom they are not living with when youth were in their mid-twenties, respectively. Moreover, we also observe a greater dependence on welfare for older QOP enrollees as they were 15 significant percentage points more likely to receive welfare or food-stamps when they were in their early-twenties (much of this greater public assistance reliance is through the receipt of food-stamps). Although the significance of these coefficients disappears when youth are in their mid-twenties, they remain sizable.

Summarizing, QOP had detrimental effects on the criminal activity of both the younger and the older enrollees when these were in their mid-twenties that is 5 years after the end of the program. Moreover older QOP enrollees were more likely to be single parents, not live with their children, or receiving public assistance when they were in their early- or mid-twenties.

Impacts by Sex

Five years after the end of the program, when youth were in their mid-twenties, QOP clearly had a significantly differential effect on criminal activity by sex. As shown in Table 8, QOP significantly increased by 10 percentage points the likelihood of committing a crime in the 2 years prior to answering the 3rd telephone survey for males, while reducing it by a significant 4 percentage points for females. Similarly, QOP significantly increased by 12 percentage the likelihood of being arrested or charged over the same period for males (the effect on females is a not statistically significant reduction of 2 percentage points). It is worthwhile highlighting that when male enrollees were in their late teens, QOP reduced by a not statistically significant 8 percentage points the likelihood of being arrested or charged in the year prior to answering the survey. This finding seems to suggest that protecting difficult youth in their late teens (in terms of helping them deal with the judicial system and preventing them from being arrested or charged) may backfire with longer-term detrimental consequences as youth age.

Another relevant finding is that during the last year of the program, when youth were in their late teens, QOP had a perverse effect on males' likelihood of binge drinking and using illegal drugs in the month prior to the survey (of 7 percentage points increase each—although only the coefficient on binge drinking is statistically significant). For women, QOP increase the likelihood of binge drinking on 8 or more days in the past month when they were in their early-twenties. However, these detrimental effects do not persist and by the time youth are in their mid-twenties none of the coefficients is statistically significant.

Finally, although the effects of QOP on family life and welfare use seem more detrimental for males than females, the differences are small and not statistically significant from zero or from each other.

Impacts by Rank in the 8th-Grade Grade Distribution

Table 9 explores whether QOP was more effective for students who had lower GPAs prior to enrolling in QOP, or if instead it was more effective for those who had higher GPAs prior to enrolling in QOP. Grade distribution subgroups are defined by dividing each QOP schools' evaluation sample into thirds based on GPA in 8th-grade GPA.²⁹

QOP increased substance abuse during the last year of the program for enrollees in the bottom and the top of the distribution. For instance, QOP increased by 14 percentage points the likelihood of using illegal drugs for enrollees in the bottom of the distribution and by 8 percentage points the likelihood of binge drinking for youth in the top of the distribution (although the latter impact is not statistically significant). As observed earlier, these detrimental effects on substance abuse fade away by the time youth are in their early- and mid-twenties. In contrast, no detrimental effect on substance is found for enrollees in the middle of the distribution (most of the coefficients show a beneficial, although usually not statistically significant, effect of QOP in terms of reducing binge dinging and illegal drug use, but a not-significant positive effect on cigarette or tobacco use).

A puzzling finding is that QOP significantly increased criminal activity of enrollees in the middle of the distribution 5 years after the end of the program. QOP increased by 9 percentage points the likelihood that youth in the middle of the distribution committed a crime in the 3 months prior to the survey, and it increased by 7 percentage points their likelihood of being convicted or pleading guilty in the 2 years prior to the survey. Finally, QOP also increased by a

²⁹ Since a youth had to be in the bottom two-thirds of the 8th-grade distribution to be eligible for QOP, the middle third of the evaluation sample fell between roughly the 22nd and 44th percentiles in the grade distribution for all entering 9th graders.

significant 4 percentage points the likelihood of serving time in jail, prison or detention home over the 2 years prior to the survey.

Finally, QOP also had perverse effect on longer-term family life outcomes such as increasing by a significant 8 percentage points the likelihood of having a child with whom the youth is not living with for enrollees in the top two thirds of the distribution. QOP also increased the likelihood of welfare receipt for enrollees in the top third of the distribution by close to 10 percentage points (although the coefficient is no longer significant after adjustment for multiple comparisons).

Impacts by Funding Source

Finally, we estimated the impacts by classifying sites according to their source of funding—the Department of Labor (DOL) versus the Ford Foundation—, as shown in Table 10. This analysis led to compare Philadelphia and Yakima to the remaining five sites. Overall, we found that Ford-funded sites had more perverse effects than DOL-funded sites both in the short- and longer-term.

During the last year of the program, when youth were in their late teens, Ford-sites saw a strong detrimental effect on substance abuse (17 and 14 percentage points increase in binge drinking and using illegal drugs) and criminal activity, with a 16 percentage points increase on the likelihood of committing a crime. Although these negative effects do not subsist by the time of the 2nd telephone survey, when youth are in their early-twenties, both the detrimental effects of QOP on alcohol abuse and criminal activities re-emerged at the time of the 3rd telephone survey. As explained earlier (footnote 20), the insignificant beneficial effect of QOP on binge drinking when youth were in their early twenties is mainly driven by a suspiciously low probability among QOP youth from the Philadelphia site (only 5% of QOP youth reported binge

drinking in the Philadelphia site compared to a 28% average in the other six sites, and compared to a 19% and 23% average in the Philadelphia site when youth were in their late teens and in their mid-twenties, respectively.) The Ford-sites also observed a significant increase in the likelihood of having a child with whom the youth is not living with both in their early- and mid-twenties, and an increase of 10 percentage points in self-reported poor health at the time youth were in their mid-twenties.

In contrast, for DOL-sites, QOP had no detrimental effect on substance abuse. Moreover, although it did increase the likelihood of being arrested and convicted, the size of the coefficient on being arrested is one third smaller (although not significantly so) than that observed in the Ford-sites.

IV. What Went Wrong?

Something must have gone really wrong for QOP to end up having no beneficial results and some detrimental ones. On the one hand, QOP increased substance abuse during the last year of program participation (while youths were in their late teens); on the other hand, QOP also had longer-term consequences as it rose youths' criminal behavior and use of social assistance and it worsened their family life five years after the end of the program (when enrollees were in their early- to mid-twenties). Below, we explore several hypothesis that may help understand QOP's perverse effects.

Education and the Perverse Effects of Raising the Perception of Riskiness

An important component of QOP's educational services were its developmental activities, which included many life skills training activities designed to reduce the youths' likelihood of engaging in risky behaviors. A wide range of topics were covered in these life skills training sessions,

such as self-esteem, avoiding drug abuse, contraception, family planning, abstinence, male parenting roles, managing anger, gang prevention, or visits to prisons.

A priori such type of educational activities ought to have reduced youths' risky behaviors, as raising perceptions of the severity of bad outcomes (such as, being pregnant, being jailed for committing a crime, or being addicted) is likely to have a direct and desired effect on youth's current and future behavior. Hence, it is hard to understand how education *per se* could have led to the observed detrimental impacts both right at the end of the program, and (especially) five years after the program had ended.³⁰

According to behavioral economists, raising perceptions of the likelihood of bad outcomes does not necessarily guarantee to reduce risk-taking behavior. For instance, O'Donoghue and Rabin, 2001, show that an increase in the perceived riskiness might either decrease or increase the person's level of the risky activity. As the authors explain, "*there are two possible reactions that a person might have to an increase in perceived riskiness. First, there is the intuitive reaction wherein the person reduces his indulgence so as to avoid the bad outcome whose likelihood has increased. But, second, there is a fatalistic reaction: the person instead might decide that, because she is not willing to choose very low indulgence, the bad outcome is now essentially unavoidable, and therefore she might as well increase indulgence*" (page 56). For instance, if an adolescent perceives that, once she uses drugs, she will surely become an addict, then she is likely to use drugs a great deal if she uses them at all. This fatalistic reaction relies on two key assumptions. First, it must be that, given the amount in which the person might plausibly want to indulge, the probability of harm done is not negligible. Second, the eventuality being risked must be all or nothing rather than cumulative—that is, if the

³⁰ For instance, Oettinger, 1999, uses data on individual enrollment in sex-education classes and finds that taking such classes increases the likelihood of becoming sexually active earlier but has only a weak effect on the likelihood of an earlier pregnancy.

bad thing happens once, then it either cannot happen again or will not cause much further harm if it does occur. According to O'Donoghue and Rabin, the fatalistic reaction identified above is likely to be important when a person suffers from an overly strong taste for immediate gratification or from projection bias—defined by Loewenstein *et al.*, 1999, as the capacity that people have to under appreciate the effects of changes in their states and hence falsely project their current consumption preferences onto their future preferences (see also Reed and van Leeuwen, 1998, on how individuals violate rational expectations by overprojecting from the current state.

Alternatively, these authors warn that an increase in the perceived riskiness of less risky behaviors can also lead people not to abstain, but rather to engage in riskier substitute behaviors. For instance, if an adolescent suddenly learns that even condoms do not protect fully against the risk of pregnancy, they may start having unprotected sex. Similarly preaching the dangers of marijuana use can cause more harm than good if it induces young people who use marijuana to exaggerate the degree to which their lives are ruined already and hence to underestimate the additional harm of cocaine or, worse yet, if it leads them to substitute cocaine for marijuana to begin with. Finally, Donoghue and Rabin show that when people make repeated risky choices, overindulgence due to self-control problems and projection bias become more pronounced.

Hence, according to behavioral economists, the perverse effects of raising perception of riskiness ought to be greater among those who are more likely to have already engaged in risky behaviors, younger teens (as they are both more impatient and more subject to peer pressure than older ones, as discussed by Lewis, 1981), or males (since they are known to have a stronger taste for immediate gratification, less self-discipline and higher discount rates than females, as found by Silverman, 2003, Duckworth and Seligman, 2006, and Warner and Pleeter, 2001, respectively). The findings from QOP present evidence consistent with this hypothesis, as males

are more likely to engage in criminal activity than females. Not finding stronger detrimental effects on substance abuse for younger enrollees versus older ones is not evidence against these models as older enrollees have already been held behind a grade prior to 9th grade, and therefore may have additional challenges than those who began 9th grade at the age of thirteen or fourteen. As these additional challenges may be correlated with having previously engaged in risky behavior or having higher discount rates, this may explain why detrimental effects among this group of enrollees is also found.

Mentoring and the Hypothesis of Deterrence

The purpose of case management was to assess both the unmet needs of enrollees and the barriers they faced and to fashion a service mix that best addressed those needs and barriers. Case managers addressed any problems in any aspect of the enrollee's life, monitored the enrollee's progress, and advocated for the enrollee in matters pertaining to school, family, the juvenile justice system, and college. The mentoring function involved a long-term personal relationship between an enrollee and his or her case manager that is similar to the relationship between a youth and a close and caring older relative. The case manager was expected to model appropriate behavior and attitudes, set disciplinary standards, and be continually available.

From psychologists' perspective, building strong positive relationships with extra familial adults (mentors) promotes resiliency among at-risk youth because mentoring facilitates adolescents' capacity to benefit from the support of parents and other providers, and influences positively the youth's perceptions of self-worth and their beliefs about their competence as learners and their valuing of school, and reduces their involvement with risky behaviors (Rhodes, Grossman, and Resch, 2000). Similarly, a willingness of adults (outside the family) to discipline

youths, provide positive role models and reduce the amounts of unsupervised youth activity are hypothesized to reduce crime (Wilson, 1987).³¹

In any case, it is hard to understand how case management and mentoring could have led to the observed detrimental effects. One possibility is that mentors ended up overprotecting youths in such a way that they reduced their costs of engaging in risky behaviors (in particular, but not exclusively, criminal activity). For instance, by acting as advocates of QOP enrollees and negotiating on behalf of them with the high school, the criminal justice and other public agencies when youth got in trouble, mentors mitigated the consequences of misbehaving. And by doing so, QOP may have had a perverse effect in that it prevented its enrollees from internalizing the full costs of engaging in such types of risky behaviors, leading to higher involvement in such type of activities in the future. This view is consistent with Becker's 1968 economic model of crime in which crime can be deterred through punishment, and with studies that have found that youths are responsive to sanctions (see for instance, Pacula, 1998 a; Chaloupka *et al.*, 1999 a and 1999 b; Levitt, 1998; and Levitt and Lochner, 2001, among others).

QOP's evidence supports this hypothesis of deterrence (that is, a behavioral response of potential criminals to the incentives they face). For instance, QOP enrollees were 16 and 12 statistically significant percentage points more likely than members from the control group to report having participated in a program that help them stay out of trouble and deal with police and the judicial system, respectively. In addition, QOP enrollees were 5 percentage points less likely to be arrested or charged over the 12 months prior to the survey conducted during the last year of the program. Albeit not statistically significant, the size of this coefficient is far from

³¹ Rigorous studies on the effectiveness of mentoring programs find that they have positive but modest effects on the young people that participate in them, and that the most disadvantaged or at-risk seem to benefit the most from them—see Dubois *et al.*, 2002, and Jekielek *et al.*, 2002, for thorough reviews on the effectiveness of mentoring programs; and Grossman and Tierney, 1998, for a random assignment evaluation of one of the most well known mentoring program in the United States, Big Brother/Big Sister.

negligible, and contrasts with the 6 percentage points significant increase in the likelihood of being arrested during the 2 years prior to the 3rd telephone survey, suggesting that reduced punishment while the program operated led to higher criminal activity in the future. As explained in the subgroup analysis section above, such pattern is even stronger (and statistically significantly so) for older youths and males.

QOP's findings also provide evidence in favor of the concept of deterrence and against incapacitation (that is, a mechanical reduction in crime that occurs because criminals are unable to commit crime while incarcerated). While QOP increased the likelihood of committing crime or being arrested over the 2 years prior to the survey conducted when enrollees were in their mid-twenties, it did not reduce the likelihood of serving time in jail over the same time period (this is particularly true for, males, enrollees from the middle of the distribution, and those from Ford-sites.)

Economic Incentives and Risky Behaviors

While stipends were intended to induce enrolled youth to participate in QOP activities, they may well have had a troublesome effect on substance abuse by providing income that could have been used to purchase alcohol and drugs. In other words, QOP may have indirectly financed the engagement in risky behaviors. There is a growing body of evidence showing that youths are very responsive to economic incentives, such as prices, when deciding to undertake risky behaviors (Nisbet and Vakil, 1972; Cook, 1981; Cook and Tauchen, 1982; Coate and Arluck, 1987; Saffer and Grossman, 1987a, 1987b; Coate and Grossman, 1988; Kenkel, 1993; Cook and Moore, 1993, 2000; Grossman et al., 1987, 1994; Chaplounka and Wechler, 1996; Ruhm, 1996; Evans and Huang, 1998; Markowitz and Grossman, 1998; Gruber, 2001; Gruber and Zinman, 2001; Pacula *et al.*, 2001; Levine, 2001; Ruhm, 2005; Ruhm and Black, 2005; and Grossman

2005). In addition, Bachman *et al.*, 1981 and 1988 find that both part-time employment and income are directly related to marijuana use.

Moreover, Gruber and Zinman, 2001, find that substance abuse is much more responsive to price for black youths and disadvantaged youths than for white teens and teens from higher socio-economic backgrounds. Similarly, Levine, 2001, also finds that teens' risk of pregnancy is much more responsive to prices, measured by labor market conditions, for black non-Hispanics than whites. These findings suggest a strong correlation between price sensitivity and socioeconomic status (and in particular lower income)—see Evans *et al.*, 1999, and Gruber and Zinman, 2001, for evidence corroborating this hypothesis—, and implies that perverse effects of economic incentives may be particularly concerning among QOP's targeted population, which is an extremely disadvantaged group.

Evidence consistent with youth responding to economic incentives is that QOP increased risky behavior specially in those sites where youth received higher amount of stipends. As such, QOP enrollees from Ford-sites were considerably and (statistically significantly so) more likely to engage in substance abuse and criminal activity than enrollees from DOL-sites. A major implementation difference between DOL and Ford sites was that while case managers in DOL-funded sites had a flat wage, those in Ford-funded sites were compensated by incentive payments based on program attendance. Not surprisingly this led to higher levels of enrollees' participation level in the Ford-sites than in DOL-sites, and hence, higher level of stipends in the former than the latter. For instance while on average, students' stipends amounted to \$3,099 and \$5,409 in Philadelphia and Yakima, respectively, over the five year period, in DOL-sites the average stipend ranged between \$973 (in Houston) and \$2,666 (in Cleveland).³²

³² Total expenditures per enrollee across sites vary between \$18,000 and \$49,000, with the two Ford sites having the highest expenditures with \$23,000 in Yakima and \$49,000 in Philadelphia.

Peer Effects and Spill-Over Effects

QOP also offered cultural activities and recreational activities (such as day trips, attendance to theater or concerts, summer camps, or music lessons) with the twofold objective of: (i) exposing enrollees to their own and other cultures and methods of self-expression; and (ii) helping enrollees build relationships with mentors and peers. The rationale for these activities was that strengthening peers' relationship would motivate enrollees to participate in the cognitive and non-cognitive educational activities. However, building strong peer effects may well have had a negative effect if it ended up facilitating or encouraging participation into risky behaviors.

It is clear from the existing literature that peer engagement in risky behaviors has an important effect on the susceptibility of youths to get involved in such type of activities (see Brook *et al.*, 1998; Kandel, 1985; Jessor *et al.*, 1980 for peers' influence on marijuana use; Norton *et al.*, 1998, for drinking and peer effects; and Case and Katz, 1991 for peer-group interactions and criminal activity; and Jensen and Lleras-Muney, 2009). Indeed, the assumption that peers are central to adolescent alcohol and drug use and criminal activity is reflected in the social-influence paradigm underlying many prevention programs (Bauman and Ennett, 1996).

Unfortunately, QOP was not designed to test for peer effects. Therefore it is very difficult to disentangle whether its detrimental effects on substance abuse and criminal activity were due to the encouragement of peer groups among its enrollees. In an attempt to identify possible peer effects, Table 11 displays QOP impacts by classifying sites according to whether QOP enrollees represented more than 38% or less than 17% of the entering 9th graders in that school.³³ This analysis led to compare Cleveland and Washington DC to the remaining five sites. A striking result is that QOP's detrimental impact on substance abuse is mainly driven by

³³ There were no schools in which QOP enrollees represented between 17% and 38% of the 9th-grade entering class.

schools in which QOP enrollees represented a small fraction of the entering class of 9th graders. Note that this result holds across time, suggesting that current indulgence leads to future indulgence.³⁴ In contrast, QOP had a beneficial impact of reducing substance abuse both in the short- and longer-term in those sites where QOP enrollees were a significant fraction of the entering class, suggesting that QOP can make a difference in relatively small communities (but that when the social network of the youth is large, QOP has a magnifying effect of the negative social-influence in the community). Indeed, in those sites where QOP enrollees represented less than 18% of the entering class of 9th graders, it appears that QOP ended up reinforcing the risky behaviors of its enrollees. It is important to remember that QOP targeted disadvantaged youth from low performing schools, and that its targeted population had major difficult barriers to overcome, such as, substance abuse problems, or broken families. These results seem to suggest that peer effect in risky behaviors is closely tied to the social pressure in the environment as negative effects are magnified in relatively large school, in which social pressure towards engaging in risky behaviors is relatively strong, while the opposite is true in relatively smaller schools. This is consistent with evidence from Mas and Moretti (forthcoming) that suggests that peer effect in productivity is entirely due to the social pressure in the environment.

The beneficial effects of QOP on reduction of risky behaviors in those sites where QOP enrollees were a large fraction of the entering 9th graders were not only long-lasting, but also led to significantly beneficial educational outcomes—for instance, QOP increased by 14 (11) percentage points the likelihood of graduating from high-school (ever having been enrolled in college) at the time youth were in their mid-twenties. This estimate was statistically significantly

³⁴ Again, results when youth were in the early-twenties are contaminated by the suspiciously low level of binge drinking among QOP enrollees in the Philadelphia site.

different from zero and from QOP's impacts in sites where QOP enrollees represented a small fraction of the population (a non-significant -5).

Finally, notice that the estimates from Table 11 indirectly test for spill-over effects. If these effects would have been an issue, we would have found that beneficial short-term impacts would have been smaller in those schools where QOP students represented a larger fraction of the GPA-eligible 9th graders in the school. Given that we find the opposite, this rules out spill-over effects from QOP to control group members.

Sensitivity Analysis

A concern may be that the results presented thus far are a statistical artifact due to peculiarities in a given site, random baseline differences across treatment and control group members, or non-response bias. Below, we explore these possibilities.

First, we have checked robustness of the results to excluding one of the sites at a time. The results in Table A.2 show that overall the results for the whole sample are quite robust to eliminating one of the sites. There are a few exceptions worth mentioning. As discussed earlier, QOP's 6 percentage points beneficial impact on the reduction of binge drinking when youth were in their early twenties is exclusively driven by the Philadelphia site. Clearly this result is difficult to believe because, in this site, only 5% of QOP youth reported binge drinking at the time of the 2nd survey compared to a 28% average for QOP youths in the other six sites. Moreover, this estimate is also surprisingly and suspiciously low if we compare it to the percentage of youth reporting binge drinking in Philadelphia when they were in their late teens (19%) or in their mid-twenties (23%). The other outlier is the beneficial impact on cigarette smoking at the time of the 3rd survey, which is driven by a large negative estimate from Cleveland site. This time around, the result is due to unusually high smoking rates among the

youths in the control group in Cleveland—65% (56%) reported smoking (daily) in the past months, far from the average of 29% and 19% for the control youth in all the other six sites.

A second robustness check is to check whether the difference-of-means estimates presented above have been affected by purely random differences between the baseline characteristics of QOP enrollees and the baseline characteristics of members of the control group. Table A.3 presents regression-adjusted impact estimates. Again as discussed earlier although the statistical significance of the 3rd telephone survey impacts on criminal activity are no longer statistically significant, the coefficients remain large and overall the main findings and conclusions are unaffected.

Although the survey effort is equiparable to that of other studies (Schochet *et al.*, 2008; Banerjee *et al.*, 2007; and Kremer *et al.*, 2008; among others), a concern with the current paper would emerge if there were to be differential non-response bias between treatment and control group members.³⁵ More specifically, if non-respondents are more likely to be individuals with more difficult lives (and consequently worse outcomes), given that response rates are higher for the QOP members than for members of the control group, a concern is that the detrimental results are due to having a higher proportion of “difficult” youth responding in the QOP group relative to the control group. To explore the internal validity of the results presented in the text we have done the following robustness check. Table A.4 presents unweighted estimates that were derived by making the response rate for the QOP group equal to the response rate for the control group within each of the 11 QOP schools. That is, if the QOP group had a higher response rate, we treated enough QOP group respondents as nonrespondents to lower the implied

³⁵ The response rate to the in-person survey and the 1st telephone surveys was 84%. For each of the surveys, the response rate for the QOP group exceeded the response rate for the control group by 7 percentage points. The response rate to the 2nd and 3rd telephone survey was 75% and 76%—80% (77%) for the QOP group and 70% (74%) for the control group in the 2nd (3rd) telephone survey.

response rate to the level of the control group. The QOP group respondents who were treated as nonrespondents were the last ones to respond to the survey—as these were those most difficult to find and who had eventually responded because of our intense survey effort. Our preferred estimates displayed in the paper (in Table 6) are also displayed for comparison purposes. Overall the estimates from Table A.4 deliver practically the same results discussed earlier in the main text, suggesting that it is not the higher response rate among the treatment group youth that is driving the results.

V. Conclusion

This paper is the first to use a randomized trial in the US to analyze the short- and long-term impacts of an afterschool program that offered disadvantaged high-school youth: mentoring, educational services, and financial rewards to attend program activities, complete high-school and enroll in post-secondary education on youths' engagement in risky behaviors, such as substance abuse, criminal activity, and teenage childbearing. Outcomes were measured at three different points in time, when youths were in their late-teens, and when they were in their early- and their late-teens. Overall the program was unsuccessful at reducing risky behaviors. Heterogeneity matters in that perverse effects are concentrated among certain subgroups, such as males, older youths, and youths from sites where youth received higher amount of stipends. In contrast beneficial effects were found in those sites in which QOP youths represented a large fraction of the entering class of 9th graders provides hope for these type of programs when operated in small communities and supports the hypothesis of peer effects.

It is important to explain why the results of this demonstration differed from those of the evaluation of the QOP pilot, which was conducted between 1989 and 1993 in five sites with funding from the Ford Foundation (Hahn, 1999; Hahn *et al.*, 1999). In contrast with the large-scale

evaluation demonstration presented in this paper, the results from the QOP pilot were considerably more promising. However, they were measured *only* several months after participants should have graduated from high-school, and therefore it is unclear whether these findings would have persisted over time as youth grew older.³⁶ Other important differences between the pilot and the demonstration included: (i) the sample size, which was smaller in the pilot; and (ii) the targeted population, which, in the pilot, were low-income students (as opposed to academically low-performing students as in the demonstration). Finally, results from the Philadelphia site in the pilot were exceptional and frequently this site was the only one to produce statistically significant results.

In this paper, we present a number of explanations for these findings. In our view, the two leading explanations involve different models of youths' behavioral response to economic incentives, on the one hand, and peer effects and social pressure, on the other. However, this field experiment was not designed to test alternative theories, therefore other explanations are also possible. For instance, it may well be that by offering a mentor and having youth participate in many after school activities, QOP may have weakened the ties between enrollees and their parents, breaking important social bonds, and thus leading to the observed perverse effects. An alternative and related explanation is that because enrollees' parents trusted that another adult (the mentor) was also watching over their children, they ended up investing less time with their children and paying less attention to possible warning signs than parents of youths in the control group. In essence, QOP may have led to a substitution effect away from parents' attention, which could have explained these detrimental findings. Another explanation is that there may have been some substitution effect over time. If teenagers need to misbehave a certain extent over their adolescence, it may well be that QOP postponed some of that misbehavior (such as, criminal activity) and advanced other type of risky behavior (such as, substance abuse). Yet

³⁶ The only risky behavior outcome collected was teenage pregnancy.

another hypothesis for the gloomy results is that QOP may have led enrollees to be more aware of their relative disadvantaged situation in life, which may have brought upon them further disappointment, leading them to engage in diverse types of risky behaviors.

There are other weaknesses of this evaluation, such as the fact that our paper does help identify a link between the different types of risky behaviors (as it was not designed for this). In addition, alternative data collection methods could have been used to improve response rates and veracity of the information, such as audio computer-assisted self-interviews (A-CASI), as explained by Gutierrez and Torres-Pereda, 2009. Despite all these caveats, some lessons emerge. First, the findings in this paper highlight the relevance of measuring impacts of educational interventions at different points in time. Only by following youth over time, can we know whether the short-term impacts persist, how they evolve in the medium-term, and whether they translate into longer-run outcomes. Second, this paper highlights the relevance of a multi-angle analysis of educational interventions. In addition to the standard educational and employment outcomes, it is also extremely important to collect outcomes on risky behaviors, since these interventions may also influence such type of behaviors. Third, the discrepancy between the QOP results and those in developing countries suggests the need for further research using experimental designs to answer whether and how the different components (mentoring, education and cash incentives) work to motivate students.

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

<i>Activity</i>	<i>Examples of such types of activities</i>
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 services	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 postsecondary education or training.

TABLE 2
Random Assignment Protocol

Step 1	Generate a list of all eligible 9 th graders at each participating school.
Step 2	Because the number of eligible students was larger than the target sample size in all but two schools, the second step was to randomly select students who would participate in the evaluation (in either the program or control group) from among all eligible students. This minimized the burden of the evaluation on students, parents, and schools.
Step 3	Obtain consent for participation in the study from students' parents. We obtained consent from 98% of the study sample.
Step 4	Randomly assign students within each school to either the QOP group or the control group.

TABLE 3
Group mean baseline characteristics by treatment group
(Percentages)

<i>Pre-program Characteristics</i>	<i>QOP group</i>	<i>Control group</i>
Male	52	56
Age when entering 9 th grade		
< 14	11	11
14	53	57
> 14	36	31
Hispanic	26	26
Black	68	68
Rank based on 8 th grade GPA		
Bottom third	37	34
Middle third	31 [†]	36 [†]
Top third	32	30
Sample size	580	489

Note: † Significantly different from the mean for the other group at the 90% confidence level, two-tailed test

TABLE 4
Timing of QOP implementation and survey data collection

	<i>QOP demonstration implementation</i>			<i>Post-demonstration</i>		
	<i>Fall 1995</i>	<i>Spring 1999 On time graduation</i>	<i>November 1999- June 2000</i>	<i>September 2000</i>	<i>September 2002- April 2003</i>	<i>January- September 2005</i>
Youth entered 9 th grade		Paper survey on resiliency factors and Achievement tests in math and reading	1 st telephone survey	End of the program	2 nd telephone survey	3 rd telephone survey

Note: All events occurred one year later for the Washington DC site with the exception of the two post-demonstration surveys, which were collected at the same time in the DC site than in the other sites.

TABLE 5
Participation in QOP Activities

	<i>Cumulative Years 1 through 4</i>	<i>Year 1</i>	<i>Year 4</i>
Average Number of Hours	708	247	103
Average Hours on Educational Activities	305	110	40
Average Hours on Developmental Activities	306	105	41
Average Hours on Community Service Activities	97	32	22
No Hours of Participation (percent)	1	1	26
More Than 100 Hours (percent)	88	73	29
More Than 375 Hours (percent)	62	23	11
More Than 750 Hours (percent)	36	1	0
More Than 1,500 Hours (percent)	13	0	0

Source: QOP Demonstration Management Information System (MIS).

Note: Because QOP services in Period 5 differed substantially from those of the first four periods, I report trends 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 6
Short-, medium and long-term impacts of QOP
(Percentage points except where noted)

<i>Outcomes</i>	<i>When youth were in their:</i>		
	<i>Late teens</i>	<i>Early twenties</i>	<i>Mid-twenties</i>
Smoked cigarettes or used tobacco in past month	--	--	0
Smoked cigarettes or used tobacco daily in past month	--	--	-2
Drinking in the past 30 days	7**	--	--
Binge drinking in the past 30 days	4	-6	0
Binge drinking on 8 or more days in past month	2	2	3
Used any illegal drug in the past 30 days	7**	-6**	-0
Committed a crime in past 3 months	-1	-2	3*
Committed any crime in past 1 or 2 years (1)	3	--	5
Arrested or charged (2)	-5	-0	6**
Convicted or pled guilty in past 2 years	--	--	2
Served time in jail, prison, or detention home in past 2 years	--	--	1
Ever had sex	-5	--	--
Did not use condom last time	0	--	--
Taught about HIV/AIDS	0	--	--
Ever pregnant or get anyone pregnant	0	--	--
Have first child before age 18	-3	3	2
Currently living with natural children, but no spouse	--	3	1
Have child with whom not living	--	1	1
Currently receiving welfare or food-stamps	--	4	3
Currently receiving welfare	--	2	1
Currently receiving food-stamps	--	5	2
Self-reported health is fair, poor, or very poor	--	--	2
Physical or mental condition limited activities quite a lot	--	--	1
Or could not work because of these limitations			

Note: (1) In the last 12 months if short-term impacts; and in the past 2 years if long-term impacts. (2) Ever arrested or charged if short-term impacts; in the past 3 months if medium-term impacts; and in the last 2 years if long-term impacts.

TABLE 7

Short-, medium and long-term impacts of QOP, by age
(Percentage points except where noted)

	Greater than 14 years old			14 years old or less		
	Late teens	Early twenties	Mid-twenties	Late teens	Early twenties	Mid-twenties
Smoked cigarettes or used tobacco in past month	--	--	4	--	--	-4
Smoked cigarettes or used tobacco daily in past month	--	--	-2	--	--	-3
Binge drinking in the past 30 days	5	-13*	-7	4	-2	5
Binge drinking on 8 or more days in past month	--	4	2	--	2	3
Used any illegal drug in the past 30 days	8	-13*	-2	5	-3	-1
Committed a crime in past 3 months	--	-2	3	--	-2	2
Committed any crime in past 1 or 2 years (1)	-4	--	7	5	--	4
Arrested or charged (2)	-11*	2	9*	-3	-1	5**
Convicted or pled guilty in past 2 years	--	--	2	--	--	3*
Served time in jail, prison, or detention home in past 2 years	--	--	2	--	--	2
Have first child before age 18	6†	2	4	-9**††	3	-1
Currently living with natural children, but no spouse	--	12**	-5	--	1	3
Have child with whom not living	--	4	11*††	--	0	-4††
Currently receiving welfare or food-stamps	--	15**††	8	--	0††	0
Currently receiving welfare	--	6	8	--	1	-1
Currently receiving food-stamps	--	19***†††	5	--	-1†††	1
Self-reported health is fair, poor, or very poor	--	--	2	--	--	1
Physical or mental condition limited activities quite a lot Or could not work because of these limitations	--	--	0	--	--	-0

Note: (1) In the last 12 months if short-term impacts; and in the past 2 years if long-term impacts. (2) Ever arrested or charged if short-term impacts; in the past 3 months if medium-term impacts; and in the last 2 years if long-term impacts.

* Estimate significantly different from zero at the 90% confidence level, two-tailed test; ** Estimate significantly different from zero at the 95% confidence level, two-tailed test; *** Estimate significantly different from zero at the 99% confidence level, two-tailed test, †, ††, ††† Significantly different from the impact for the other subgroups at the 90%, 95%, and 99% confidence level, two-tailed test

TABLE 8

Short-, medium and long-term impacts of QOP, by sex
(Percentage points except where noted)

<i>OUTCOMES</i>	<i>Males</i>			<i>Females</i>		
	<i>Late teens</i>	<i>Early twenties</i>	<i>Mid-twenties</i>	<i>Late teens</i>	<i>Early twenties</i>	<i>Mid-twenties</i>
Smoked cigarettes or used tobacco in past month	--	--	-3	--	--	4
Smoked cigarettes or used tobacco daily in past month	--	--	-6	--	--	2
Binge drinking in the past 30 days	7*	-12*	-5	0	-2	4
Binge drinking on 8 or more days in past month	--	0	1	--	4**	3
Used any illegal drug in the past 30 days	7	-8*	-4	7	-3	3
Committed a crime in past 3 months	--	-3	5*†	--	-1	-1†
Committed any crime in past 1 or 2 years (1)	5	--	10**†††	2	--	-4*†††
Arrested or charged (2)	-8	1	12***†††	0	-2	-2†††
Convicted or pled guilty in past 2 years	--	--	4*†	--	--	-0†
Served time in jail, prison, or detention home in past 2 years	--	--	3	--	--	-1
Have first child before age 18	-3	5	3	-5	0	0
Currently living with natural children, but no spouse	--	2	-3	--	3	3
Have child with whom not living	--	4	4	--	-1	-2
Currently receiving welfare or food-stamps	--	3	3	--	0	-1
Currently receiving welfare	--	1	2	--	-1	-1
Currently receiving food-stamps	--	4	3	--	2	-1
Self-reported health is fair, poor, or very poor	--	--	4	--	--	-2
Physical or mental condition limited activities quite a lot	--	--	0	--	--	1
Or could not work because of these limitations	--	--	--	--	--	--

Note: : (1) In the last 12 months if short-term impacts; and in the past 2 years if long-term impacts. (2) Ever arrested or charged if short-term impacts; in the past 3 months if medium-term impacts; and in the last 2 years if long-term impacts. * Estimate significantly different from zero at the 90% confidence level, two-tailed test; ** Estimate significantly different from zero at the 95% confidence level, two-tailed test; *** Estimate significantly different from zero at the 99% confidence level, two-tailed test, †, †† or ††† Significantly different from the impact for all other subgroups at the 90%, 95% or 99% confidence level, respectively, two-tailed test.

TABLE 9

Short-, medium and long-term impacts of QOP, by 8th grade GPA score
(Percentage points except where noted)

<i>OUTCOMES</i>	<i>Bottom third</i>			<i>Middle third</i>			<i>Top third</i>		
	<i>Late teens</i>	<i>Early twenties</i>	<i>Mid-twenties</i>	<i>Late teens</i>	<i>Early twenties</i>	<i>Mid-twenties</i>	<i>Late teens</i>	<i>Early twenties</i>	<i>Mid-twenties</i>
Smoked cigarettes or used tobacco in past month	--	--	-9†	--	--	6	--	--	3
Smoked cigarettes or used tobacco daily in past month	--	--	-6	--	--	4	--	--	-5
Binge drinking in the past 30 days	8	-7	-1	-4†	-4	-5	8*	-7	3
Binge drinking on 8 or more days in past month		4	4		2	-1	--	0	2
Used any illegal drug in the past 30 days	14**†	-10*	6†	-2†	-8*	-5	7	-1	-0
Committed a crime in past 3 months	--	2	1	--	-6*	9**††	--	-1	-2††
Committed any crime in past 1 or 2 years (1)	2	--	7	4	--	5	8	--	3
Arrested or charged (2)	-11*	2	5	0	-1	7**	1	-1	5
Convicted or pled guilty in past 2 years	--	--	-0	--	--	4*	--	--	3
Served time in jail, prison, or detention home in past 2 years	--	--	-3††	--	--	4**	--	--	3
Have first child before age 18	-4	6	2	-8*	3	-1	3	0	5
Currently living with natural children, but no spouse	--	2	-1	--	7	1	--	3	3
Have child with whom not living	--	3	4	--	8**††	-3	--	8**††	2
Currently receiving welfare or food-stamps	--	7	0	--	-2	3	--	10	9
Currently receiving welfare	--	3	-5	--	-4	-0	--	7	9*
Currently receiving food-stamps	--	8	2	--	1	3	--	8	8
Self-reported health is fair, poor, or very poor	--	--	4	--	--	0	--	--	-3
Physical or mental condition limited activities quite a lot or could not work because of these limitations	--	--	0	--	--	1	--	--	1

Note: : (1) In the last 12 months if short-term impacts; and in the past 2 years if long-term impacts. (2) Ever arrested or charged if short-term impacts; in the past 3 months if medium-term impacts; and in the last 2 years if long-term impacts. *, ** Estimate significantly different from zero at the 90% and 95% confidence level, respectively, two-tailed test; † Significantly different from the impact for all other subgroups at the 90% confidence level, two-tailed test; †† Significantly different from the impact for the other subgroup at the 95% confidence level, two-tailed test.

TABLE 10

Short-, medium and long-term impacts of QOP, by funding source
(Percentage points except where noted)

<i>OUTCOMES</i>	<i>Ford-sites</i>			<i>DOL-sites</i>		
	<i>Late teens</i>	<i>Early twenties</i>	<i>Mid-twenties</i>	<i>Late teens</i>	<i>Early twenties</i>	<i>Mid-twenties</i>
Smoked cigarettes or used tobacco in past month	--	--	9	--	--	-4
Smoked cigarettes or used tobacco daily in past month	--	--	3	--	--	-4
Binge drinking in the past 30 days	17***†	-13	8	-1†	-2	-3
Binge drinking on 8 or more days in past month	--	5	14**	--	0	-2†
Used any illegal drug in the past 30 days	14**	-8	-6	4	-5*	2
Committed a crime in past 3 months	--	3	7*	--	-3	1
Committed any crime in past 1 or 2 years (1)	16**†	--	7	-2†	--	4
Arrested or charged (2)	-3	-5	12*	-5	2	4*
Convicted or pled guilty in past 2 years	--	--	2	--	--	3*
Served time in jail, prison, or detention home in past 2 years	--	--	10	--	--	-2
Have first child before age 18	-14**†	1	1	1†	4	3
Currently living with natural children, but no spouse	--	-6	-7	--	7*	4
Have child with whom not living	--	12*	14*†	--	-3	-4†
Currently receiving welfare or food-stamps	--	9	-6	--	2	5*
Currently receiving welfare	--	4	-6	--	1	6
Currently receiving food-stamps	--	11	-9	--	2	4
Self-reported health is fair, poor, or very poor	--	--	10*†	--	--	-2†
Physical or mental condition limited activities quite a lot	--	--	-4	--	--	3
Or could not work because of these limitations						

Note: (1) In the last 12 months if short-term impacts; and in the past 2 years if long-term impacts. (2) Ever arrested or charged if short-term impacts; in the past 3 months if medium-term impacts; and in the last 2 years if long-term impacts. *, **, *** Estimate significantly different from zero at the 90%, 95%, and 99% confidence level, respectively, two-tailed test; †, †† or ††† Significantly different from the impact for all other subgroups at the 90%, 95% or 99% confidence level, respectively, two-tailed test.

TABLE 11

Short-, medium and long-term impacts of QOP, by relative weight of QOP group
(Percentage points except where noted)

OUTCOMES	QOP is a small fraction of 9 th graders			QOP is a large fraction of 9 th graders		
	Late teens	Early twenties	Mid-twenties	Late teens	Early twenties	Mid-twenties
Smoked cigarettes or used tobacco in past month	--	--	6**†	--	--	-15**†
Smoked cigarettes or used tobacco daily in past month	--	--	4*†	--	--	-16**†
Binge drinking in the past 30 days	10**†	-9*	5**†	-10*†	3	-12*†
Binge drinking on 8 or more days in past month	--	3	6**†	--	-1	-6*†
Used any illegal drug in the past 30 days	11**†	-2	-1	-4	-15	1
Committed a crime in past 3 months	--	0	4	--	-7	-1
Committed any crime in past 1 or 2 years (1)	4	--	4	-1	--	6
Arrested or charged (2)	-4	-1	6*	-6	3	6
Convicted or pled guilty in past 2 years	--	--	2	--	--	3
Served time in jail, prison, or detention home in past 2 years	--	--	0	--	--	4
Have first child before age 18	-4	3	5	-1	6	3
Currently living with natural children, but no spouse	--	-1	0	--	15**†	3
Have child with whom not living	--	1	2	--	-1	-2
Currently receiving welfare or food-stamps	--	7	5	--	-2	-4
Currently receiving welfare	--	3	0	--	0	4
Currently receiving food-stamps	--	8	4	--	-3	-4
Self-reported health is fair, poor, or very poor	--	--	2	--	--	1
Physical or mental condition limited activities quite a lot	--	--	0	--	--	3
Or could not work because of these limitations						

Note: (1) In the last 12 months if short-term impacts; and in the past 2 years if long-term impacts. (2) Ever arrested or charged if short-term impacts; in the past 3 months if medium-term impacts; and in the last 2 years if long-term impacts. *, **, *** Estimate significantly different from zero at the 90%, 95%, and 99% confidence level, respectively, two-tailed test; †, †† or ††† Significantly different from the impact for all other subgroups at the 90%, 95% or 99% confidence level, respectively, two-tailed test.

REFERENCES

- Anderson, M. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," Manuscript, Massachusetts Institute of Technology, Dept. of Economics.
- Angrist, Joshua D., and Victor Lavy. 2010. "The Effect of High School Matriculation Awards: Evidence from Randomized Trials." *American Economic Review*.
- Angrist, Joshua, Daniel Lang, and Philip Oreopoulos. Forthcoming. "Incentives and Services for College Achievement: Evidence from a Randomized Trial." *American Economic Journal: Applied Economics*.
- Ash, Katie. 2008. "Promises of Money meant to heighten Student Motivation." *Education Week*, February 13.
- Attanasio, Orazio, Emla Fitzsimons, Ana Gomez, Diana Lopez, Costas Meghir and Alice Mesnard. 2006. "Child education and work choices in the presence of a conditional cash transfer programme in rural Colombia." Working Paper W06/13. London: Institute for Fiscal Studies.
- Bachman, Jerald G., Lloyd D. Johnston, and Patrick M. O'Malley. 1981. "Smoking, drinking, and drug use among American high school seniors: Correlates and trends, 1975-1979." *American Journal of Public Health* 71, no. 1:59-69.
- Bachman, Jerald G., Lloyd D. Johnston, Patrick M. O'Malley, and Ronald H. Humphrey. 1988. "Explaining the recent decline in marijuana use: Differentiating the effects of perceived risks, disapproval, and general lifestyle factors." *Journal of Health and Social Behavior* 29:92-112.
- Banerjee, A., S. Cole, E. Duflo and L. Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India," *Quarterly Journal of Economics*, Vol. 122 (3), pp.1235-64.
- Bauman, K. E., and S. T. Ennett. 1996. "On the importance of peer influence for adolescent drug use: Commonly neglected considerations." *Addiction* 91, no. 2: 185-98.
- Barnett, W.S., 1992. "Benefit-cost analysis of preschool education: Findings from a 25-year follow-up." *American Journal of Orthopsychiatry*, 1993
- Becker, Gary. 1968. "Crime and punishment: An economic approach." *Journal of Political Economy* 76:169-217.
- Bettinger, Eric P., and Bridget Terry Long. 2005. "Addressing the Needs of Under-Prepared Students in Higher Education: Does College Remediation Work?" National Bureau of Economic Research Working Paper 11325.
- Behrman, J.R., Hoddinott, J., 2005. "Programme Evaluation with Unobserved Heterogeneity and Selective Implementation: The Mexican PROGRESA Impact on Child Nutrition." *Oxford Bulletin of Economics and Statistics* 67, 547-569.
- Behrman, Jere R., P. Sengupta, and P. Todd. 2000. "Final Report: The Impact of PROGRESA on Achievement Test Scores in the First year." Washington, DC:

- International Food Policy Research Institute, Food Consumption and Nutrition Division.
- Behrman, J.R., Sengupta, P., Todd, P., 2005. "Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment in Rural Mexico." *Economic Development and Cultural Change* 54, 237-275.
- Brook, Judith S., Patricia Cohen, and L. Jeager. 1998. "Developmental variations in factors related to initial and increased levels of adolescent drug involvement." *Journal of Genetic Psychology* 159, no. 2:179-94.
- Cardoso, Ana Rute, and Dorte Verner. 2008. "Youth Risk-Taking Behavior in Brazil: Drug Use and Teenage Pregnancy." World Bank, Policy Research Working Paper no. 4548.
- Case, A. and L. Katz. 1991. "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths." NBER Working Paper No. 3705. National Bureau of Economic Research, Cambridge, MA.
- Chaloupka, F. J., and H. Wechsler. 1996. "Binge drinking in college: The impact of price." *Contemporary Economic Policy* 14, no. 4:112-24.
- Coate, D., and M. Grossman. 1988. "Effects of alcoholic beverage prices and legal drinking ages on youth alcohol use." *Journal of Law and Economics* 31, no. 1:145-71.
- Cornwell, Christopher, David B. Mustard, and Deepa J. Sridhar. 2006. "The Enrollment Effects of Merit-Based Financial Aid: Evidence from Georgia's HOPE Program." *Journal of Labor Economics*, 24(4): 761-86.
- Cook, P. J. 1981. "The effect of liquor taxes on drinking, cirrhosis, and auto fatalities." In *Alcohol and public policy: Beyond the shadow of Prohibition*, ed. Mark Moore and Dean Gerstein. Washington, D.C.: National Academy of Sciences.
- Cook, Phillip J., and Michael J. Moore. 1993. "Drinking and schooling." *Journal of Health Economics* 12, no. 4:411-29.
- . 2000. "Alcohol." In *Handbook of health economics*, vol. 1b, ed. Anthony Culyer and Joseph Newhouse, 629-74. Amsterdam: Elsevier.
- Cook, P. J., and G. Tauchen. 1982. "The effect of liquor taxes on heavy drinking." *Bell Journal of Economics* 13, no. 2:379-90.
- Dearden, L., Emmerson, C., Frayne, C., and Meghir, C. 2005. "Education Subsidies and School Drop-Out Rates." Institute for Fiscal Studies Working Paper 05/11.
- Deming, D. and S. Dynarski. 2009. "Into college, out of poverty? Policies to increase the postsecondary attainment of the poor". NBER Working Paper 15387.
- DeWit, D. J., Lipman, E., Manzano-Munguia, M., Bisanz, J., Graham, K., Offord, D. R., O'Neill, E., Pepler, D., & Shaver, K. 2006. "Feasibility of a randomized controlled trial for evaluating the Effectiveness of the Big Brothers Big Sisters community match program at the national level." *Children and Youth Services Review*, 29, 383-404.
- DuBois, D.L., Holloway, B.E., Valentine, J.C. and Cooper, H. 2002. "Effectiveness of mentoring programs for youth: A meta-analytical review." *American Journal of Community Psychology*, 30 (2), 157-197.

- Duckworth, Angela Lee, and Martin P. 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.
- El País, 2009. "Chequera Contra el Fracaso Escolar." Reportaje de J.A. Aunión, 14 October, 2009.
- Evans, W.N. and Huang, L. 1997. "Cigarettes taxes and teen smoking: new evidence from panels of repeated cross-sections." Mimeo, University of Maryland.
- Evans, William, Jeanne Ringel, and Diana Stech. 1999. "Tobacco taxes and public policy to discourage smoking." In *Tax policy and the economy 13*, ed. James Poterba. Cambridge, Mass.: MIT Press.
- Fernald, L.C., Gertler, P.J., Hou, X., 2008a. "Cash component of conditional cash transfer program is associated with higher body mass index and blood pressure in adults." *Journal of Nutrition* 138, 2250-7.
- Fernald, L.C., Gertler, P.J., Neufeld, L.M., 2008b. "Role of cash in conditional cash transfer programmes for child health, growth, and development: an analysis of Mexico's Oportunidades." *Lancet* 371, 828-37.
- Fernald, L.C., Hou, X., Gertler, P.J., 2008c. "Oportunidades program participation and body mass index, blood pressure, and self-reported health in Mexican adults." *Review of Chronic Diseases* 5, A81.
- Fryer, Roland. 2010. "Financial Incentives and Student Achievement: Evidence from Randomized Trials." Department of Economics, Harvard University Working Paper.
- Galárraga, Omar and Paul J. Gertler. 2009. "Conditional Cash & Adolescent Risk Behaviors: Evidence from Urban Mexico," mimeo.
- Gertler, P. 2004. "Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment." *American Economic Review*.
- Grossman, Michael. 2005. "Education and non-market outcomes," NBER Working Paper 11582.
- Grossman, M., D. Coate, and G. M. Arluck. 1987. "Price sensitivity of alcoholic beverages in the United States: Youth alcohol consumption." In *Control issues in alcohol abuse prevention: Strategies for states and communities*, ed. H. Holder. Greenwich, Conn.: JAI.
- Grossman, M., F. J. Chaloupka, H. Saffer, and A. Laixuthai. 1994. "Effects of alcohol price policy on youth: A summary of economic research." *Journal of Research on Adolescence* 4:347–64.
- Grossman, Michael and Sara Markowitz. 2005. "I did what last night?! Adolescent risky sexual behaviors and substance use." *Eastern Economic Journal*, 31(3): 383–405.
- Grossman, J. B., and J. P. Tierney. 1998. "Does Mentoring Work? An Impact Study of the Big Brothers Big Sisters program," *Evaluation Review*, 22(3), pp. 402–425.

- Grossman J.B and C.L Sipe. 1992. "Summer Training and Education Program (STEP): Report on long-term impacts." Public Private Ventures, Philadelphia, PA.
- Gruber, Jonathan. 2001. *Risky Behavior Among Youth: An Economic Analysis*. Chicago: University of Chicago Press.
- Gruber, Jonathan and Jonathan Zinman. 2001. "Youth Smoking in the United States: Evidence and Implications," in, *Risky Behavior among Youths: An Economic Analysis*, p. 69-120. Chicago: University of Chicago Press.
- Gutierrez, J.P., Torres-Pereda, P., 2009. "Acceptability and reliability of an adolescent risk behavior questionnaire administered with audio and computer support." *PanAmerican Journal of Public Health* 25, 418-22.
- Hahn, A. 1999. "Extending the Time of Learning," in Douglas J. Besharov, ed., *America's Disconnected Youth*, Washington, D.C.: Child Welfare League of America, Inc.
- Hahn, A., T. Leavitt and P. Aaron. 1999. *Evaluation of the Quantum Opportunities Program*, Heller Graduate School, Center for Human Resources, Brandeis University,
- Herrera, C., Grossman, J. B., Kauh, T. J., Feldman, A. F., & McMaken, J. (with Jucovy, L. Z.). 2007. "Making a difference in schools: The big brothers big sisters school-based mentoring impact study." Philadelphia, PA: Public/Private Ventures.
- Hoddinott, J., Skoufias, E. 2004. "The Impact of PROGRESA on Food Consumption." *Economic Development and Cultural Change* 53, 37-61.
- Jackson, C. K. 2007. "A Little Now for a Lot Later: A Look at a Texas Advanced Placement Incentive Program."
- Jekielek, S.M., Moore, K.A., Hair, E.C. and Scarupa, H.J. 2002. *Mentoring: A promising strategy for youth development*. (Child Trends Research Brief.) Washington, DC: Child Trends.
- Jensen, R., and A. Lleras-Muney, 2009. "Does Staying in School Prevent Teen Smoking and Drinking?" CCPR-044-09.
- Jessor, R. 1991. "Risk behaviour in adolescence: a psychosocial framework for understanding and action", *Journal of Adolescent Health*, vol. 21, pp. 597-605.
- Kandel, Denise B. 1985. On processes of peer influences in adolescent drug use: A development perspective. *Advances in Alcohol and Substance Abuse* 4, no. 3:139-63.
- Kane, Thomas J. 2007. "Evaluating the Impact of the DC Tuition Assistance Grant Program." *Journal of Human Resources*.
- Kane, Thomas J. 1998. "Savings Incentives for Higher Education." *National Tax Journal*, Vol. 51, No. 3, pp. 609-620.
- Karcher, M. J. 2005. "The effects of developmental mentoring and high school mentors' attendance on their younger mentees' self-esteem, social skills, and connectedness." *Psychology in the Schools*, 42, 65-77.
- Keating, L. M., Tomashina, M. A., Foster, S., and Allesandri, M. 2002. "The effects of a mentoring program on at-risk youth." *Adolescence*, 37, 717-734.

- Kenkel, D. S. 1993. "Drinking, driving, and deterrence: The effectiveness and social costs of alternative policies." *Journal of Law and Economics* 36:877–913.
- Kremer, Michael, Edward Miguel, and Rebecca Thornton. Forthcoming. "Incentives to Learn." *The Review of Economics and Statistics*.
- Krueger, Alan B. and Diane M. Whitmore. 2002. "Would Smaller Classes Help Close the Black-White Achievement Gap?" In *Bridging the Achievement Gap*, ed. John E. Chubb and Tom Loveless, 11-46. Washington, D.C.: The Brookings Institution.
- Leslie, L. L., and Brinkman, P. T. 1988. *The Economic Value of Higher Education*. San Francisco: Jossey-Bass.
- Leuven, E., H. Oosterbeek, and B. van der Klaauw. 2003. "The Effect of Financial Rewards on Students' Achievement: Evidence from a Randomized Experiment." CEPR Discussion Paper 3921.
- Levine, P. (2001). "The sexual activity and birth-control use of American teenagers", in (J. Gruber, ed.), *Risky Behaviour Among Youth: An Economic Analysis*, pp. 167–218, Chicago: University of Chicago Press.
- Lewis, C. C. 1981. "How adolescents approach decisions: Changes over grades seven to twelve and policy implications." *Child Development* 52:538–44.
- Loewenstein, G., T. O'Donoghue, and M. Rabin. 1999. "Projection bias in predicting future utility." Mimeo.
- Markowitz, S., and M. Grossman. 1998. "Alcohol regulation and domestic violence towards children." *Contemporary Economic Policy* 16, no. 3:309–20.
- Mas, Alexandre, and Enrico Moretti. 2009. "Peers at Work." *American Economic Review*, 99(1): 112–45.
- Maxfield, Myles, Laura Castner, Vida Maralani, and Mary Vencill. 2003a. "The Quantum Opportunity Program Demonstration: Implementation Findings." Washington, DC: Mathematica Policy Research, Inc.
- Maxfield, Myles, Allen Schirm, and Nuria Rodriguez-Planas. 2003b. "The Quantum Opportunities Program Demonstration: Implementation and Short-Term Impacts." Mathematica Policy Research Report 8279-093. Washington, DC: Mathematica Policy Research, Inc.
- Nisbet, Charles T., and Firouz Vakil. 1972. "Some estimates of price and expenditure elasticities of demand for marijuana among U.C.L.A. students." *Review of Economics and Statistics* 54, no. 4:473–75.
- Norton, E. C., R. C. Lindrooth, and S. T. Ennett. 1998. "Controlling for the endogeneity of peer substance use on adolescent alcohol and tobacco use." *Health Economics* 7, no. 5:439–53.
- O'Donoghue, T., Rabin, M., 2000. "Risky Behavior Among Youths: Some Issues from Behavioral Economics." In: Gruber J (ed.) *Youthful Risky Behavior: An Economic Perspective*. University of Chicago Press, Chicago.

- O'Donoghue, T. and Rabin, M. 2001. "Risky behaviour among youths: some issues from behavioural economics", in (J. Gruber, ed.), *Risky Behaviour Among Youth: An Economic Analysis*, pp. 29–68, Chicago: University of Chicago Press.
- Oettinger, G.S., 1999. "The effects of sex education on teen sexual activity and teen pregnancy." *Journal of Political Economy* 107 (3), 606–644.
- Pacula, R.L.; Grossman, M; Chaloupka, F.J.; P. O'Malley; L.D. Johnston and M.C. Farrelly. 2001. "Marijuana and Youth" in Jonathan Gruber's *An Economic Analysis of Risky Behavior Among Youths*. University of Chicago Press.
- Rashad, Inas and Robert Kaestner. 2004. "Teenage sex, drugs and alcohol use: problems identifying the cause of risky behaviors." *Journal of Health Economics*, 23: 493–503.
- Read, D., and B. van Leeuwen. 1998. Predicting hunger: The effects of appetite and delay on choice. *Organizational Behavior and Human Decision Processes* 76:189–205.
- Rees, Daniel I., Laura M. Argys, and Susan L. Averett. 2001. "New evidence on the relationship between substance use and adolescent sexual behavior." *Journal of Health Economics*, 20: 835–845.
- Rhodes, Jean. 1994. "Older and Wiser: Mentoring Relationships in Childhood and Adolescence," *The Journal of primary prevention*, v 14 n 3, Spring.
- Rhodes Jean. 2008. "Improving Youth Mentoring Interventions Through Research-based Practice." *American Journal of Community Psychology* 41:35–42
- Rhodes, Jean E., Jean B. Grossman, and Nancy L. Resch. 2000. "Agents of Change: Pathways through Which Mentoring Relationships Influence Adolescents' Academic Adjustment." *Child Development*, November/December, Volume 71, Number 6, Pages 1662–1671.
- Rodríguez-Planas, Núria. 2010. "Longer-term Impacts of Mentoring, Educational Services, and Incentives to Learn: Evidence from a Randomized Trial in the US". IZA discussion paper No. 4754, February 2010
- Ruhm, C. 1996. "Alcohol policies and highway vehicle fatalities." *Journal of Health Economics* 15:435–54.
- Ruhm, Christopher J. 2005. "Healthy Living in Hard Times," *Journal of Health Economics*, 24(2), p. 341-63.
- Ruhm, Christopher J. and William E. Black. 2002. "Does Drinking Really Decrease in Bad Times?," *Journal of Health Economics*, 21(4), p. 659-78.
- Saffer, H., and M. Grossman. 1987a. Beer taxes, the legal drinking age, and youth motor vehicle fatalities. *Journal of Legal Studies* 16:351–74.
- . 1987b. Drinking age laws and highway mortality rates: Cause and effect. *Economic Inquiry* 25, no. 3:403–17.
- Sen, A.K., 1999. *Development as Freedom*. Alfred A. Knopf, New York.
- Schirm, Allen, Nuria Rodriguez-Planas, Myles Maxfield, and Christina Tuttle. 2003. "The Quantum Opportunity Program Demonstration: Short-Term Impacts." Washington, DC: Mathematica Policy Research, Inc.
- Schirm, Allen, and Nuria Rodriguez-Planas. 2004. "The Quantum Opportunity Program

- Demonstration: Short-Term Impacts.” Washington, DC: Mathematica Policy Research, Inc.
- Schirm, Allen, Elizabeth Stuart, and Allison McKie. 2007. "The Quantum Opportunity Program Demonstration. Final Impacts.” Washington, DC: Mathematica Policy Research, Inc.
- 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.
- Schultz, T. Paul. 2004. “School Subsidies for the Poor: Evaluating the Mexican Progress Poverty Program.” *Journal of Development Economics*, 74(2): 199-250.
- Silverman, Irwin W. 2003. “Gender Differences in the Delay of Gratification: A Meta-Analysis.” *Sex Roles*, 49(9-10): 451-463.
- 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 T., and Saul Pleeter. 2001. “The Personal Discount Rate: Evidence from Military Downsizing Programs.” *The American Economic Review*, 91(1): 33-53.
- Wilson, William Julius. 1987. *The truly disadvantaged: The inner city, the underclass, and public policy*. Chicago: University of Chicago Press.

APPENDIX

(not necessarily for publication)

A.I. Weights

Overall impacts of QOP were estimated using student-, school-, and site-specific weights.³⁷ Weights for each student were used to adjust for survey nonresponse and ensure that the sample of respondents reflects the experiences of all sample members.³⁸ The impacts for each school were calculated as the weighted difference in the outcomes of members of the QOP and control groups. The impacts for each site were calculated as a weighted average across schools using weights that reflected the proportion of QOP slots in each school. This weighting was selected because we believe that each program would have allocated slots across schools in the same way they did in the demonstration if they had been part of an ongoing, national program.³⁹ Finally, to obtain the overall demonstration impacts, the site-specific effects were averaged, with each site weighted equally. The equal weighting of sites was based on our best guess that if QOP were implemented as an ongoing, national program, each site would have roughly equal numbers of QOP slots.

³⁷ For thorough description on how the weighted averages were calculated see Maxfield *et al.*, 2003 a; Schirm *et al.*, 2004; and Schirm *et al.*, 2006.

³⁸ Non-response weights were estimated using response propensity scores for the treatment and the comparison group, separately. The predictors used in the response propensity scores included school dummies, baseline characteristics, interactions between the previous school and baseline characteristics and between any two baseline characteristics, and outcomes measured in any of the earlier surveys.

³⁹ Weighting each site in proportion to the number of students in the study did not lead to different study conclusions (Schirm *et al.* 2006).

TABLE A.1
Control group means
(Percentage points except where noted)

OUTCOMES	<i>Late teens</i>	<i>Early twenties</i>	<i>Mid-twenties</i>
Smoked cigarettes or used tobacco in past month	--	--	34
Smoked cigarettes or used tobacco daily in past month	--	--	24
Drinking in the past 30 days	33	--	--
Binge drinking in the past 30 days	20	31	31
Binge drinking on 8 or more days in past month	5	5	6
Used any illegal drug in the past 30 days	28	18	13
Committed a crime in past 3 months	11	9	2
Committed any crime in past 1 or 2 years (1)	28	--	11
Arrested or charged (2)	29	5	5
Convicted or pled guilty in past 2 years	--	--	3
Served time in jail, prison, or detention home in past 2 years	--	--	2
Ever had sex	83	--	--
Did not use condom last time	28	--	--
Taught about HIV/AIDS	94	--	--
Ever pregnant or get anyone pregnant	33	--	--
Have first child before age 18	26	15	16
Currently living with natural children, but no spouse	--	23	31
Have child with whom not living	--	13	17
Currently receiving welfare or food-stamps	--	20	24
Currently receiving welfare	--	13	14
Currently receiving food-stamps	--	17	24
Self-reported health is fair, poor, or very poor	--	--	8
Physical or mental condition limited activities quite a lot	--	--	7
Or could not work because of these limitations	--	--	7

Note: : (1) In the last 12 months if short-term impacts; and in the past 2 years if long-term impacts. (2) Ever arrested or charged if short-term impacts; in the past 3 months if medium-term impacts; and in the last 2 years if long-term impacts.

TABLE A.2
Short-, medium and long-term impacts of QOP, by 8th grade GPA score
(Percentage points except where noted)

<i>OUTCOMES</i>	<i>Impacts dropping the site listed below:</i>							<i>All sites included</i>
	<i>Fort Worth</i>	<i>Cleveland</i>	<i>DC</i>	<i>Houston</i>	<i>Memphis</i>	<i>Philadelphia</i>	<i>Yakima</i>	
			<i>Late teens</i>					
Binge drinking in the past 30 days	7	8	6	4	2	2	2	4
Used any illegal drug in the past 30 days	6**	8**	8**	6**	6**	6**	5*	7**
Committed any crime in past 1 year	4	6	3	3	5	1	1	3
Ever arrested or charged	-5	-6	-4	-5	-4	-6	-4	-5
Have one or more own children	-3	-4	-3	-4	-5	-1	-2	-3
			<i>Early twenties</i>					
Binge drinking in the past 30 days	-5	-5	-9	-5	-7	-1	-8	-6
Binge drinking on 8 or more days in past month	2	3	1	1	3	1	1	2
Used any illegal drug in the past 30 days	-7**	-5*	-4*	-8**	-7**	-6**	-6**	-6**
Committed a crime in past 3 months	-2	-1	-1	-2	-2	-2	-3	-2
Arrested or charged in the past 3 months	0	-1	-0	-1	-1	1	0	0
Had first child before age 18	4	5	1	4	3	4	4	3
Currently living with natural children, but no spouse	5	2	1	4	2	5	5	3
Have child with whom not living	1	1	1	2	2	-1	-1	1
Currently receiving welfare or food-stamps	5	5	6	4	3	4	3	4
Currently receiving welfare	2	3	2	2	2	3	1	2
Currently receiving food-stamps	6	5	7	4	4	5	3	5
			<i>Mid-twenties</i>					
Smoked cigarettes or used tobacco in past month	0	5	-0	-2	1	-2	-1	0
Smoked cigarettes or used tobacco daily in past month	-3	3	-1	-3	-2	-3	-2	-2
Binge drinking in the past 30 days	2	3	1	-2	-1	1	-3	0
Binge drinking on 8 or more days in past month	3	3	5**	3	3	2	-0	3
Used any illegal drug in the past 30 days	-1	-1	-1	-2	-0	1	-1	0
Committed a crime in past 3 months	3*	3*	4**	3*	2	2	2	3*
Committed any crime in past 2 years	5	3	6*	5	4	4	4	5
Arrested or charged in the past 2 years	7***	6**	7***	7***	7***	5**	6**	6**
Convicted or pled guilty in past 2 years	2	2	2	3	2	2	3	2
Served time in jail, prison, or detention home in past 2 years	1	1	1	2	1	1	2	1
Self-reported health is fair, poor, or very poor	3	2	1	1	3	0	-0	2

Physical or mental condition limited activities quite a lot or could not work because of these limitations	0	-0	1	1	-0	2	1	1
Had first child before age 18	4	5	2	5	3	4	3	2
Currently living with natural children, but no spouse	2	2	-0	2	-3	4	1	1
Have child with whom not living	2	-0	3	3	3	-1	-1	1
Currently receiving welfare or food-stamps	2	6	1	0	1	7	1	3
Currently receiving welfare	1	1	1	1	1	5	2	1
Currently receiving food-stamps	1	6	0	-0	0	6	0	2

Note: : *, ** Estimate significantly different from zero at the 90% and 95% confidence level, respectively, two-tailed test.

TABLE A.3
Short-, medium and long-term regression-adjusted impacts of QOP
(Percentage points except where noted)

<i>OUTCOMES</i>	<i>IMPACTS</i>		
	<i>Late teens</i>	<i>Early twenties</i>	<i>Mid-twenties</i>
High-school performance			
Earned high-school diploma ^a	7**	-1	3
Earned high-school diploma or GED ^a	5	2	4
Earned a bachelor's or associate's degree			
Attending college	3	2	-3
Attending post-secondary education	5	5	-0
Ever in college	--	8*	5
Ever in post-secondary education	--	10**	7*
Has a job			
Total earnings in the past 12 months (dollars)	--	-6	-0
Hourly earnings (dollars)	--	--	-211
Substance abuse			
Smoked cigarettes or used tobacco in past month	--	--	-0
Smoked cigarettes or used tobacco daily in past month	--	--	-2
Binge drinking in the past 30 days	3	-5	0
Binge drinking on 8 or more days in past month	--	2	3
Used any illegal drug in the past 30 days	7**	-6**	-1
Criminal Activity			
Committed a crime in past 3 months	4	-2	2
Committed any crime in past 1 or 2 years (1)	--	--	4
Arrested or charged (2)	-3	-1	6
Convicted or pled guilty in past 2 years	--	--	2
Served time in jail, prison, or detention home in past 2 years	--	--	1
Family Life and Welfare Use			
Have first child before age 18	-4	3	2
Currently living with natural children, but no spouse	--	3	0
Have child with whom not living	--	1	1
Currently receiving welfare or food-stamps	--	3	2
Currently receiving welfare	--	1	1
Currently receiving food-stamps	--	4	1
Self-reported health is fair, poor, or very poor	--	--	1
Physical or mental condition limited activities quite a lot or could not work because of these limitations	--	--	0

Note: : (1) In the last 12 months if short-term impacts; and in the past 2 years if long-term impacts. (2) Ever arrested or charged if short-term impacts; in the past 3 months if medium-term impacts; and in the last 2 years if long-term impacts.

TABLE A.4
Impacts Using Alternative Approaches to Adjusting for Non-Response
(Percentage points)

<i>OUTCOMES</i>	<i>Early twenties</i>		<i>Mid-twenties</i>	
	<i>Preferred Estimates</i> (1)	<i>Same non-response rate</i> (2)	<i>Preferred Estimates</i> (3)	<i>Same non-response rate</i> (4)
High-school performance				
Earned high-school diploma ^a	0	1	0	2
Earned high-school diploma or GED ^a	2	2	2	2
Post-secondary training				
Earned a bachelor's or associate's degree	--	--	-1	1
Attending college	1	2	-4	-3
Attending post-secondary education	4	5	-1	0
Ever in college	7*	8**	4	4
Ever in post-secondary education	9**	9**	6	7*
Employment				
Has a job	-7*	-5	-1	-0
Total earnings in the past 12 months (dollars)	--	--	-522	-349
Hourly earnings (dollars)	--	--	-1.20	-0.95
Substance abuse				
Smoked cigarettes or used tobacco in past month	--	--	0	1
Smoked cigarettes or used tobacco daily in past month	--	--	-2	-2
Binge drinking in the past 30 days	-6	-6*	0	1
Binge drinking on 8 or more days in past month	2	1	3	3
Used any illegal drug in the past 30 days	-6**	-5**	-0	-1
Criminal Activity				
Committed a crime in past 3 months	-2	-2	3*	1
Committed any crime in past 1 or 2 years (1)	--	--	5	1
Arrested or charged (2)	-0	0	6**	5**
Convicted or pled guilty in past 2 years	--	--	2	2
Served time in jail, prison, or detention home in past 2 years	--	--	1	1
Family Life and Welfare Use				
Have first child before age 18	3	4	2	2
Currently living with natural children, but no spouse	3	6*	1	2
Have child with whom not living	1	-2	1	-0
Currently receiving welfare or food-stamps	4	5	3	4
Currently receiving welfare	2	3	1	3
Currently receiving food-stamps	5	5	2	3
Self-reported health is fair, poor, or very poor	--	--	2	1
Physical or mental condition limited activities quite a lot or could not work because of these limitations	--	--	1	-0
Sample size	788	670	793	710

Note: Estimates were obtained using weights to adjust for differences between respondents and nonrespondents. Estimates in columns (2) and (4) were derived by making the response rate for the QOP group equal to the response rate for the control group for each of the 11 schools. That is, if the QOP group had a higher response rate, we treated enough QOP group respondents as nonrespondents to lower the implied response rate to the level of the control group. The QOP group respondents that were treated as nonrespondents were the last ones to respond to the survey. : (1) In the last 12 months if short-term impacts; and in the past 2 years if long-term impacts. (2) Ever arrested or charged if short-term impacts; in the past 3 months if medium-term impacts; and in the last 2 years if long-term impacts.

*, ** Estimate significantly different from zero at the 90% or 95% confidence level, two-tailed test.

^aHigh school diploma or GED outcomes have been complemented with high-school transcript information.

TABLE A.5
Impacts Not Adjusting for Non-Response
(Percentage points)

	<i>Early twenties</i>		<i>Mid-twenties</i>	
	<i>Preferred Estimates (1)</i>	<i>Not adjusted for non- response (2)</i>	<i>Preferred Estimates (3)</i>	<i>Not adjusted for non- response (4)</i>
OUTCOMES				
High-school performance				
Earned high-school diploma ^a	0	1	0	2
Earned high-school diploma or GED ^a	2	2	2	3
Post-secondary training				
Earned a bachelor's or associate's degree	--	--	-1	-0
Attending college	1	1	-4	-3
Attending post-secondary education	4	5	-1	-0
Ever in college	7*	6*	4	4
Ever in post-secondary education	9**	9**	6	7*
Employment				
Has a job	-7*	-6*	-1	1
Total earnings in the past 12 months (dollars)	--	--	-522	-66
Hourly earnings (dollars)	--	--	-1.20	-1.24
Substance abuse				
Smoked cigarettes or used tobacco in past month	--	--	0	-1
Smoked cigarettes or used tobacco daily in past month	--	--	-2	-2
Binge drinking in the past 30 days	-6	-7*	0	0
Binge drinking on 8 or more days in past month	2	1	3	3
Used any illegal drug in the past 30 days	-6**	-6**	-0	-1
Criminal Activity				
Committed a crime in past 3 months	-2	-2	3*	2
Committed any crime in past 1 or 2 years (1)	--	--	5	2
Arrested or charged (2)	-0	0	6**	5**
Convicted or pled guilty in past 2 years	--	--	2	2
Served time in jail, prison, or detention home in past 2 years	--	--	1	1
Family Life and Welfare Use				
Have first child before age 18	3	4	2	2
Currently living with natural children, but no spouse	3	5	1	3
Have child with whom not living	1	-1	1	-0
Currently receiving welfare or food-stamps	4	5	3	3
Currently receiving welfare	2	2	1	2
Currently receiving food-stamps	5	5*	2	2
Self-reported health is fair, poor, or very poor	--	--	2	1
Physical or mental condition limited activities quite a lot or could not work because of these limitations	--	--	1	-1
Sample size	788	670	793	710

Note: Estimates in columns (2) and (4) are not adjusted for non-response.

*, ** Estimate significantly different from zero at the 90% or 95% confidence level, two-tailed test.

^aHigh school diploma or GED outcomes have been complemented with high-school transcript information.