I Z A

IZA DP No. 4754

Longer-Term Impacts of Mentoring, Educational Services, and Incentives to Learn: Evidence from a Randomized Trial

Núria Rodríguez-Planas

February 2010

Longer-Term Impacts of Mentoring, Educational Services, and Incentives to Learn: Evidence from a Randomized Trial

Núria Rodríguez-Planas
Universitat Autònoma de Barcelona,
MOVE, FEDEA and IZA

ABSTRACT

Longer-Term Impacts of Mentoring, Educational Services, and Incentives to Learn: Evidence from a Randomized Trial[*]

This paper is the first to use a randomized trial in the US to analyze the short- and long-term educational and employment impacts of an after-school program, the Quantum Opportunity Program, that offered disadvantaged high-school youth: mentoring, educational services, and financial rewards with the objective to improve high-school graduation and post-secondary schooling enrollment. Average impacts reveal that the hefty beneficial educational outcomes quickly faded away. Heterogeneity matters. While encouraging results are found for the younger youth; detrimental long-lived outcomes for males suggest that extrinsic rewards may be crowding out intrinsic motivation. Evidence by sites' funding source, which led to implementation differences, supports this hypothesis.

Corresponding author:

Núria Rodríguez-Planas
Unitat de Fonaments de l'Anàlisi Econòmica
Edifici B
Universitat Autònoma de Barcelona
08193 Bellaterra
Spain
E-mail: Nuria.Rodriguez@uab.es

I. Introduction

In response to the observed increasing earnings differential between the most and the least educated workers since the late 1970s, policies aiming at improving high-school graduation and post-secondary education enrollment have recently received renewed attention from policy makers, practitioners and researchers. As a consequence, there has been a new wave of interventions whose main objective is to improve the school performance of disadvantaged youth.[1] Most of these interventions involve one or the combination of the following services: (i) a mentoring component; (ii) an educational component; and (iii) a financial incentive component; and they generally find promising results if not for all participants, for some particular subgroups.

While most of these studies 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. In addition, and probably because of the shorter-term focus, all of these studies look at educational outcomes as opposed to employment outcomes and wages.[2] 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 or do they quickly fade away? Do they translate into longer-term payoffs as measured by post-secondary schooling, employment, and earnings? These are the questions this paper aims to address for the first time in the United States.

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 educational and employment outcomes in the United States. QOP involved the combination of the following three services: a mentoring, an educational, and a financial incentives

---

[1] For experimental or quasi-experimental designed evaluations targeting primary schools, see Jacob and Lefgren, 2004; Banerjee *et al*., 2007; and Kremer *et al*., 2008; and James-Burdumy *et al*., 2008. For studies focusing on high-school youth, see Machin *et al*., 2004, and 2007; and Lavy and Schlosser, 2005; Angrist and Lavy, 2009; and Holmlund and Silva, 2009. And for evaluations aiming at improving college students' performance, see Bettinger and Long, 2005; Leuven *et al.*, 2003; Angrist *et al*., 2009, and Scrivener *et al.*, 2009. These studies complement the literature on employment and training programs for disadvantaged youth (Bloom *et al.*, 1996; Lalonde, 1995; Kemple and Willner, 2008; Schochet *et al.*, 2008).

[2] To the best of my knowledge, only three papers look at longer horizons as in the present paper: Leuven *et al.*, 2003, Bettinger and Long, 2005, and Angrist and Lavy, 2009, measure outcomes three, four and five years after the end of the intervention, respectively. Angrist *et al.*, 2009, focus on the impacts of three types of treatment (one involving financial rewards, another one involving educational services, and the third one combining the other two) on grades and academic standing of college students in Canada; Leuven *et al.*, 2003, study the impacts of financial rewards on academic achievement of college students in Amsterdam; and Angrist and Lavy, 2009, measure long-term impacts of ever attending college for high-school students who received merit awards for passing the Bagrut in Israel. Using a randomized assignment design, the current paper is the first to look at college completion, employment and earnings among the evaluations aiming at improving school performance.

component. The program's main objectives were to increase the likelihood that youth completed high-school and engaged in post-secondary education or training. The outcome variables involved high-school completion, post-secondary education or training enrollment and completion, and employment and earnings measured at three points in 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] Similar to recent studies, QOP appears to be quite effective in the short-run in terms of increasing the likelihood of high-school graduation and of attending college by 17.5% and 23%, respectively.[4][5] However, we find that these initial large gains are short-lived. All average positive effects of QOP on high-school completion and post-secondary education enrollment quickly fade away, and five years after the end of the program, QOP has no statistically significant educational or employment impacts on its enrollees. The reason is that control group members eventually caught up to the QOP group in terms of high school completion and post-secondary enrollment. These results suggest that, while QOP help youth take advantage of opportunities to get ahead in life, it is unclear as to whether helping disadvantaged youth get through the system quicker is helpful in the longer-run.

We find that heterogeneity matters. QOP's subgroup analysis reveals the following four

---

[3] See Maxfield *et al.*, 2003 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.

[4] Rigorous studies finding *average* beneficial short-term impacts of interventions aiming at improving educational performance of youth include Hahn, 1994; Hahn *et al.*, 1994; Machin, *et al.*, 2004; Jacob and Lefgren, 2004; Lavy and Schlosser, 2005; Bloom and Sommo, 2005; Banerjee *et al.*, 2007; Scrivener, *et al.*, 2009; and Kremer *et al.*, 2008. All these evaluations—except for Kremer *et al.*, 2008, which evaluated a merit awards program—, studied the effects of a variety of educational services offered to youth. Other rigorous evaluations find positive impacts for certain subgroups *only*—see Bettinger and Long, 2005; Machin *et al.*, 2007, for evaluations of educational services; Leuven *et al.*, 2003; and Angrist and Lavy, 2009, for evaluations of financial rewards; and Angrist *et al*., 2009, for a combination of services and financial rewards.

[5] Worth highlighting is the evaluation of the QOP pilot conducted between 1989 and 1993 in five sites with funding from the Ford Foundation (Hahn, 1994; Hahn *et al.*, 1994). While the results from the QOP pilot were slightly more promising than those from the large-scale evaluation demonstration presented in this paper, they were measured at most 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. Other important differences between the pilot and the demonstration included the sample size, which was smaller in the pilot, and 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.

interesting insights. First, QOP had long lasting beneficial educational outcomes for the younger enrollees. Second, QOP seemed to help female students by getting them through high school, and into post-secondary training (not necessarily college) quicker than their counterparts in the control group. However, these educational gains were short-lived and the youth from the control group eventually caught up to those in the treatment group. Five years after the end of the program female QOP enrollees had better employment outcomes than control group members, but the opposite was true for educational outcomes (measured as college attendance). Third, QOP had some adverse short-term impacts on males' academic performance. Although these adverse effects were short-lived and QOP enrollees ended up with higher GED receipt, the longer-term employment impacts continued to be detrimental five years after the end of the program. Finally, we found a differential impact by classifying sites according to their funding source, which led to differential implementation. In those sites that had: (i) high fidelity and expertise implementing the program, (ii) mentors working part-time and compensated by incentive payments based on enrollees' attendance, and (iii) periodic distribution of statements to enrollees with information on their accrual accounts, there was strong evidence of short-lived beneficial educational impacts. In contrast, in those sites with flexible implementation of the program, full-time mentors paid a flat salary, and no periodic distribution of accrual account's information to enrollees, there were long-lasting beneficial post-secondary educational outcomes.

The two latter findings suggest that financial rewards may replace powerful and enduring "intrinsic motivation" (the individual's desire to perform the task for its own sake) with short-lived "extrinsic motivation". If so, we may observe that there are detrimental or no short-term effects of the financial rewards or that the initial positive effects are short-lived and quickly fade away. These results are consistent with educators, sociologists and psychologists' beliefs that rewards can be counterproductive if they undermine intrinsic motivation. They are also in line with a substantial body of laboratory and field experimental evidence that indicates that contingent rewards can sometimes conflict with intrinsic motivation. For instance, Deci, 1975, found that those college students who were not paid to work for a certain time on an interesting puzzle, played with the puzzle significantly more in a later unrewarded "free-time" period and reported a greater interest in the task than paid subjects. This experiment was replicated for high-school students in tasks involving verbal skills (Kruglanski, Friedman and Zeevi, 1971), and for preschool children in activities involving drawing with new materials (Lepper, Greene and Nisbett, 1973). More recently,

Gneezy and Rustichini, 2000, found that students who were collecting donations door-to-door actually visited fewer houses and collected less money when they were paid a small commission. Other experiments finding that people who were not paid at all exerted greater effort than those who were paid a small amount include Wilson, Hull and Johnson, 1981; Kohn, 1993; Frey and Jegen, 2000; and Heyman and Ariely, 2004, among others.[6]

This paper contributes to the literature in at least two important ways. First, the promising findings for the younger youth contribute to the debate over whether investments in the later stages of a child's development have positive payoffs. The results suggest that the earlier the intervention the better—consistent with many academics and practitioners' beliefs that early childhood interventions are preferred (Currie, 2001; Currie and Thomas, 2001; Krueger and Whitmore, 2001; Garces *et al.*, 2002; and Carneiro and Heckman, 2003). In addition, they provide some hope for high-school interventions—if they are implemented on the entering class of freshmen students—, and in contrast with most findings from service-oriented dropout-prevention programs that present quite discouraging results (Dynarski and Gleason, 2002). Recently two rigorous studies have also found positive returns for interventions offering educational services targeted to youth at a relatively late stage of schooling (Machin, McNally and Meghir, 2004; and Lavy and Schlosser, 2005). In contrast with the present paper, both of these studies use a quasi-experimental design, present only short-term impacts of the services offered, and focus on high-school students in Israel and the United Kingdom, respectively.

Second, this paper adds to the growing interest in student incentive programs in primary and secondary education.[7] While most of the recent evaluations have looked at the short-term effects of financial rewards (Kremer *et al.*, 2008; and Angrist *et al.*, 2009), very few have analyzed the longer-term impacts (Leuven *et al.*, 2003; and Angrist and Lavy, 2009), and none have looked at employment and earnings performance of youth. In addition, these earlier studies focus on primary

---

[6] See Bonner *et al.*, 1996; Herwig and Ortmann, 1998; Jenkins *et al.*, 1998; Camerer and Hogarth, 1999; Deci *et al.*, 1999; and Cameron *et al.*, 2001, for thorough reviews of the effects of financial incentives in experiments.

[7] 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, 2006). 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 Progresa 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).

students in Kenya (Kremer *et al.*, 2008); high-school students in Israel (Angrist and Lavy, 2009); and college students in Canada (Angrist *et al.*, 2009) and in The Netherlands (Leuven *et al.*, 2003). To the best of my knowledge, this is the first randomized evaluation in the United States to study the effects of financial rewards on disadvantaged high-school students' performance using a rich variety of outcomes. It is also the first real-world context randomized evaluation in the area of youth's educational outcomes to provide evidence that the short-term benefits in response to rewards may well fade away or even be detrimental—as feared by cognitive psychologists and behavioral economists. Earlier evidence has either found beneficial short-term effects of financial rewards (Kremer *et al.*, 2008);[8] or longer-term positive results for some subgroups but no effect for others (Angrist and Lavy, 2009; and Angrist *et al.*, 2009).[9] To the best of my knowledge, Leuven *et al.*, 2003, is the only real-world context study to find explicit evidence consistent with external rewards crowding out intrinsic motivation for the least able students both in the short-term and in the longer-run. Unfortunately, their small sample sizes limit statistical precision, complicating inference.[10] Does this mean that one should discard student incentive programs as a way to increase effort and performance of students? Clearly not. In fact the findings by age highlight that financial incentives do not generate adverse effects among the younger youth in the longer run, raising the need for further research using experimental designs to answer whether and how cash incentives work to motivate students.[11]

This paper is organized as follows. The next section describes the program implementation and the data. Section III provides a theoretical background motivating the intervention. Section IV explains the evaluation framework and analyzes the results. Section V concludes with

---

[8] Other studies using a quasi-experimental approach also find beneficial effects of financial rewards on academic performance (Ashworth et al., 2001; Dearden, Emmerson, Frayne, and Meghir, 2005; Dynarski, 2003 and 2008; and Cornwell Mustard and Sridhar, 2006). Similarly, Angrist *et al.,* 2002, and Angrist *et al.*, 2006, have also found that vouchers for private secondary school students conditional on their maintaining a satisfactory level of academic performance led to academic gains, one possible channel being the incentives associated with conditional renewal of scholarships.

[9] Angrist and Lavy, 2009, find beneficial lasting effects on college attendance among girls who were relatively close to success 5 years after the end of the program. Angrist *et al.*, 2009, find lasting improvements for college freshmen women in Canada when educational services, peer advising and financial services are combined. Using a regression discontinuity design, Garibaldi *et al.*, 2007, find that gradually increasing tuition payments in response to delayed completion had substantial effects on Italian college women.

[10] Leuven *et al.*, 2003, find positive (negative) long lasting effects on academic achievement for the academically strong (weak) students. Non-experimental studies have found that U.S. scholarships lead students to get better grades but to take less ambitious course loads (Binder, 2002, Cornwell *et al.*, 2002, Cornwell *et al.,* 2006).

[11] The Innovation Education Laboratory at Harvard University is currently implementing two evaluations, Capital Gains in Washington DC, and Spark in New York, that are offering financial rewards for school performance to primary students.

suggestions on how to improve program design and evaluation.

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.[12] 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 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.[13]

In addition to case management, the program engaged youth in: (i) developmental activities

---

[12] 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.

[13] 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.

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, youth received a stipend of \$1.25 for every hour devoted explicitly to educational activities, developmental activities (excluding recreational activities), and community service. A matching amount was promised to the youth when he or she earned a high school diploma or GED *and* enrolled in post-secondary education or training (including vocational training or military service).[14] The idea was to encourage the student to not only work hard in the present but also to recognize the larger goal of high-school graduation and post-secondary education enrollment. In addition, bonuses hours (50 extra hours, for example) were given when an enrollee achieved a significant milestone, such as earning a B average or higher on his or her report card. The extra hours resulted in an increased stipend payment and accrual account contribution.

Target Population and Sample Selection

In the summer of 1995, QOP was implemented in eleven high schools across seven sites in the United States.[15] In each of these schools, entering 9th grade students—except those with a GPA from 8th-grade above the 66th percentile—, were randomly assigned to QOP or a control group.[16] 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

---

[14] 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.

[15] 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 on 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.

[16] 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.

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. 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 in Section IV.3., 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 stimulated control group members to work harder because they knew that some of their classmates were receiving additional help. In the results section, we discussed 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 data from a series of three telephone surveys, two of which were 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 math and reading achievement tests, high-school completion status, engagement in post-secondary education and training, employment (including earnings and benefits), risky behaviors, and resiliency factors (see Maxfield *et al.*, 2003b; Schirm and Rodríguez-Planas, 2004, and Schirm *et al.*, 2006 for thorough analysis of QOP's impacts on resilience factors and risky behaviors). 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.[17] Second, as these are (mainly) survey data, differential non-response between treatment and control group members could potentially bias the results.[18] 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 in the Appendix. 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.[19]

Implementation of QOP and Service Use

In many respects, the sites were successful in implementing QOP. As designed, QOP served

---

[17] 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.

[18] 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.

[19] As explained in Section IV.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.

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.[20] Case managers were hired for the whole length of the program and with prior expertise on social services.[21] 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.[22]   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.[23] 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.[24] Finally, the fact that QOP did not achieve its 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.

Finally, 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. By the end of the demonstration, accrual account balances ranged from a few hundred dollars to nearly

---

[20] 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.

[21] 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.

[22] 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).

[23] 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).

[24] 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.

$10,000, with most being in the range of $1,000 to $3,000.[25] Final payments were made directly to the enrollee rather than to the postsecondary institution or to the enrollee's parents. The size of QOP's financial rewards 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).[26] 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 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%).[27] 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

---

[25] Although some may question how much of financial rewards may have helped QOP youth enroll into college, 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.

[26] 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.

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

than those with lower participation.[28] 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.   Theoretical Background

This section presents the theoretical background motivation of this intervention by looking at three mechanisms: mentoring, educational services, and financial rewards.

Mentoring

Frequently, the problem with academic failure is not necessarily due to the students' lack of cognitive skills, but instead due to (i) weak non-cognitive skills;[29] or (ii) personal, family, and social barriers that interfere with youth's ability to attend school and do well there (Finn, 1989).[30] 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). According to this view, the effects of mentoring ought to be stronger on younger youth and those more at-risk. This is so because younger children are more malleable and receptive than older ones, implying that activities aiming to improve their social and emotional skills ought to have a greater impact the sooner they are imparted to the child and the improved environment (Heckman, 2000; Currie, 2001). The mentoring view also expects greater gains among the most "at-risk" by virtue of individual or environmental circumstances because they are those for whom mentoring can make more of a difference in terms of emotional development (Rhodes, 1994; Hall, 2003). Finally, when strong personal relationships are built, the long-term positive effects of these relationships ought to be long lasting since the benefits from socialization and the integration into mainstream society will foster further personal and emotional development.[31]

---

[28] 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.
[29] By non-cognitive skills I refer to self-esteem, motivation, tenacity, trustworthiness, and perseverance.
[30] Examples of personal, family or social barriers are being a teenage mother, not having a family car (or unavailability of parents to drive students to school), being hungry, or having some drug abuse problem, among others.
[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

Human Capital Theory

The lack of academic preparation or poor study skills represents a major challenge for many teenagers' academic success and high-school completion. To improve their school performance and prevent them from dropping out of high-school, there are many academic support services offered to low-achieving at-risk youth based on the premise that augmenting instruction time ought to increase the students' cognitive skills and help them get back on track.[32] Just as in the case for mentoring programs (but for different reasons), the human capital theory predicts that the younger the individual the stronger the effect of the educational program. This is because younger individuals have a longer time to recoup their investment than older ones, for the same level of investment at each age, implying that the return to human capital will be higher for the former than the latter (Becker, 1964). However, in contrast with the mentoring view, the human capital model predicts a stronger effect of an educational program on the higher ability youth. This is so because human capital has a fundamental dynamic complementarity feature that implies that skills acquired early on make later learning easier (Heckman, Lochner and Taber, 1998a, and Heckman, 2000).

Whether the effects of academic support services are long-lived will depend on the nature of the knowledge acquired. A priori, the human capital model predicts that the learning ought to be long-lived. However, the improvement in knowledge will be short-lived if: (i) the learning is transitory—"declarative knowledge" (facts about the world) are much more likely to be quickly and subsequently forgotten than "procedural knowledge" (a repertoire of skills, rules and strategies for using declarative knowledge to solve problems); (ii) the learning is mechanical— for instance, it mainly involves improvement in test-taking techniques; or (iii) there is a Hawthorne effect—i.e., individuals do not learn but it seems like it because they improve their test performance due to an increase in effort and attention caused by their gratefulness towards the services received. Alternatively, the acquired knowledge may also fade away overtime if there is no continuous reinforcement and the learning that occurred gets swamped by the churning that takes place as youth grow older (Banerjee *et al.*, 2007).

---

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.

[32] Academic services take various forms, such as remedial instruction, summers schools, computer-assisted learning, and academic counseling, among others.

Financial Incentives Awards

Students may exert suboptimal study effort if they overly discount the future (Greene, 1986; Nurmi, 1991, Gruber, 2001), if they have time inconsistent preferences, such as hyperbolic discounting (Green *et al.*, 1994; Kirby, 1997; Laibson, 1997, Bettinger and Slonim, 2007), or if they perceive overly high education costs or overly low expected returns (Eckstein and Wolpin, 1999). In such cases, standard economic models predict that financial incentives ought to increase individual study effort and improve performance under the assumption that performance is positively related with effort (Smith and Walkers, 1993; Gibbons, 1997; and Lazear, 2000). Similarly, behaviorist psychology also predicts that rewards improve performance, albeit for different reasons, namely financial reward offered for an unpleasant activity will reduce the aversion towards that activity, leading to long-run positive effects on performance (Skinner, 1953). According to these theories, the effects of financial rewards ought to be long-lived, especially if they create a studying habit behavior, if they reduce distaste for studying, or if they increase human capital (and therefore lower the subsequent education costs).

However, the positive effect of financial rewards on behavior has been seriously questioned both in economics and psychology. The economists' argument is that this prediction may fail because factors other than money and effort may enter into the decision of the agent. Such factors include the signaling consequences of accepting very small compensations for a task, social norms that prescribe a behavior independently of financial rewards, or reciprocity, among others (see Titmuss, 1970, and 1971; Frey, 1994; Frey and Oberholzer-Gee, 1997; Kohn, 1993, and Benabou and Tirole, 2003, among others). The cognitive psychology school also challenged the conclusion of behaviorist psychologists by arguing that financial incentives can interfere with enduring "intrinsic motivation" with short-lived "extrinsic motivation" that ultimately reduces effort and achievement (Kruglansky *et al.*, 1973; Deci *et al.*, 1973; Lepper and Greene, 1978; and Kohn, 1999). If so, we may observe that there are no short-term effects of the financial rewards or that the initial positive effects are short-lived and quickly fade away, eventually becoming negative. In contrast with the earlier models, these theories do *not* predict positive long-run effects of the financial rewards, as *weak reinforcers* in the short term may have *hidden costs*, in that they become negative reinforcers once they are withdrawn, and can even lead to detrimental long-run effects.

IV.   Results

IV.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 (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 in the Appendix. More specifically, the results are robust to (i) non-response bias (Tables A.2 and A.3); and (ii) using regression models to control for baseline characteristics (Table A.4). Overall, the results are robust the various sensitivity tests performed.

IV.2. Measurement of Performance Outcomes

The analysis in this paper focuses in three types of outcomes: those that measure youth's high-school performance, those that measure youth's post-secondary education, and those that measure youth's employment.[33] The outcomes 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. Control group means for the different outcomes are summarized in Appendix table A.1.

IV.2. Impact Results for the Full Sample

Table 6 presents average estimates of the effect of QOP measured at three points in time. Similar to recent studies, QOP appears to be quite effective in the short-run (as shown in column 1). Estimates of high-school completion and enrollment in post-secondary education measured during the last year of the demonstration, show that QOP increased by 7 percentage points the mean high-school graduation rate and by 6 percentage points the mean likelihood of attending

---

[33] Achievement test scores are expressed as percentiles in the distribution of scores for tenth graders in the United States. Post-secondary education includes two- and four-year college, vocational or technical school, and the armed forces. Earnings are coded as zero if the person is reported not working. This measure of earnings is one of *realized* earnings and is frequently used in the literature, despite being a crude measure of productivity—since earnings are only observed for employed individuals.

(or accepted in) college or post-secondary education, implying an improvement of 17.5% and 23%, respectively. Beyond the statistically significant impacts on high-school graduation and post-secondary education enrollment, QOP did not significantly improve achievement test scores, grades, or credits earned (not shown). Although QOP might not have raised grades if QOP enrollees were taking more challenging courses than the youth in the control group, it ought to have increased standardized test scores if it had an impact on improvement of youth's achievement. These results are not far from those found by Lavy and Schlosser, 2005, who find an increase in the probability of earning a matriculation certificate of 12 percentage points (corresponding to a 22% improvement). As in QOP, Lavy and Schlosser, 2005, found no effects on achievements.[34]

An important consideration in the evaluation of educational interventions is whether or not the changes generated by the interventions persist over time and last beyond the period in which the intervention is administered. To investigate this question, column 2 in Table 6 reports the estimates two years after the end of the program (and three after the youth's scheduled high-school graduation time), and show that the initial large gains are short-lived.[35] All average positive effects of QOP on high-school completion and post-secondary education enrollment have faded away three years later. The results on high-school graduation are explained by a larger number of control group members earning a high-school diploma between the 1st and 2nd telephone survey. While the likelihood of earning such a degree for QOP enrollees has increased by 10 percentage points (to 57%) during this time, the control group's percentage increased by 16 percentage points (to 56%).

At the time of the 2nd telephone survey, some advantage remains for QOP youth as they are 7 percentage points (9 percentage points) more likely than youth from the control group to have ever attended college (post-secondary education), implying a 23% (17%) improvement. At the same time, QOP has a negative effect on employment, as participants were 7 percentage points less likely to have a job two years after the end of the program (representing a 10% reduction in the likelihood of working). This negative impact on employment may be due to the usual locking-in effects of training programs—as QOP participants are 4 (not statistically significant)

---

[34] Although they did find that program participants gained on average two additional credits without lowering their average store.

[35] Medium-term high-school diploma or GED estimates differ somewhat from those in our initial 2004 report (Schirm and Rodríguez-Planas, 2004). The estimates reported here have been updated with more complete and accurate data collected after the 2nd telephone survey.

percentage points more likely to be attending post-secondary education.[36] An alternative interpretation is that these detrimental medium-term results are consistent with external financial rewards crowding out intrinsic motivation. Further analysis on longer-term impacts and subgroup analysis will help interpreting these average results.

Column 3 of Table 6 reports the longer-term average impacts of QOP. Clearly, QOP had no long-term effect on education or employment outcomes, suggesting that the medium-term detrimental effects of QOP on employment were likely to be locking-in effects. Worth highlighting are the positive (albeit not significant) effects of QOP on the likelihood of ever attending college or post-secondary education (a 4 and 6 percentage points, implying a 12% and 11% improvement) and the detrimental (non-significant) effects on total earnings in the past 12 months (a 4% decrease).

The departure of the long-term findings from the short- and medium-term ones, with no impacts on post-secondary attainment, seems to be primarily due to a larger number of control group members than QOP enrollees engaging in postsecondary education in the two years between the 2nd and 3rd telephone surveys. While the percentage of QOP enrollees who were ever engaged in any post-secondary education or training increased by only 1 percentage point (to 62 percent) during this time, the control group percentage increased by 3 percentage points (to 56 percent). Similar differences are seen for college enrollment, with the gap between the QOP enrollees and the control group narrowing by 3 percentage points between the two surveys. The changes seen over time are perhaps not surprising given that many students do not engage in post-secondary education directly after high school. For example, over 10% of 1999-2000 bachelor's degree recipients enrolled in college more than 2 years after high school graduation (U.S. Department of Education 2003).

What do we learn from the short-, medium- and long-term average impacts of QOP? It seems that QOP affected the timing of completing high-school and participating in post-secondary education. However, the control group members had caught up to the QOP group in terms of high-school graduation by the 2nd telephone survey and in terms of post-secondary enrollment by the 3rd telephone survey. These results are consistent with short-term evidence on resiliency factors, which show that youth in QOP reported to be 7 percentage points (19%) more likely (and significantly so) to have an influential adult helping them to "take advantage of

---

[36] For a discussion on locking-in effects in the training literature, see Kluve 2006.

opportunities to get ahead" in life. However, it is unclear as to whether helping at-risk youth finish school and begin post-secondary education earlier is helpful in the longer-run (as shown by the lack of educational and employment results five years after the end of the program). They suggest that for disadvantaged youth taking a bit longer to graduate from high school and to enroll in post-secondary education may well have beneficial effects, in line with findings from Cascio and Lewis, 2006, that an additional year of formal schooling among minority teenagers went a long way in building helpful skills for the labor market.[37]

An alternative and complementary explanation is that QOP may have reduced intrinsic motivation and that these adverse effects could persist after the incentive was removed. Using self-reported attitudes and behaviours, we found evidence of detrimental effects on several measures of intrinsic motivation, as such, we found that QOP did not significantly improve enrollees' outlook on the future and a clear sense of right and wrong, but it did significantly increase by 6% the fraction of youth dissatisfied with their physical appearance at the end of the demonstration. In addition, QOP significantly increased the likelihood of risky behaviours both in the short-term (by increasing drinking and the use of illegal drugs) and in the long-run (by increasing the likelihood of being arrested or committing a crime)—as shown in Appendix table A.6.[38] While it is difficult to disentangle the mechanisms through which QOP operated, subgroup analysis in the next section will enable us to explore further the hypothesis of the detrimental effects of financial rewards.

Spill-Over Effects

As previously discussed in Section II, we did not expect for this intervention to have spill-over effects because it was an after-school program that operated outside the school and outside of school hours. That said, potential spill-over effects could have bias downward the results if the intervention had had beneficial effects on the members of the control group.[39] To explore the

---

[37] It is important to keep in mind that QOP was offered to disadvantaged youth from low-performing high-schools. The consequences of falling behind a grade and remaining one additional year in high-school may well be very different in this population than for the average student in the U.S.

[38] The significant benefitial effect of QOP on binge drinking when youth were in their early twenties is mainly driven by a surprisingly and extremely low probability among QOP youth from the Philadelphia site (only 5% of QOP youth reported binge drinking in the Philadelphia compared to a 28% average in the other six sites. Moreover, in the Philadelphia site the percentage of youth reporting binge drinking when they were in their late teens and in their mid-twenties was 19% and 23%, respectively).

[39] Any potential spill-over effects are likely to be a greater problem in the short-term, while students have just recently finished school or are in their last year, than in the longer-term where students are likely to have moved on with their lives.

extent of spill-over effects, we first compared control group outcome means to national averages for youth who had exhibited poor 8th-grade academic performance using data from the National Education Longitudinal Study of 1988 that has followed a cohort of students who where 9th graders in the 1997-1998 school year. If there were any spill-over effects, our control group means may have outperformed the national averages. However, we did not find any evidence of this. For instance, the likelihood of receiving a high-school diploma (or GED) 5 to 6 years after scheduled high school graduation was 60% (or 75%) for our control group, below, but not far from, the national average of 67% (or 79%) for similarly disadvantaged youth. Similarly, within 5 to 6 years of scheduled high school graduation, 56% of control group members had engaged in some type of post-secondary education or training, not far from the national average of 58% for disadvantaged students (U.S. Department of Education 2002b). Second, we looked at whether the beneficial short-term impacts were smaller in those schools where QOP students represented a larger fraction of the GPA-eligible 9th graders in the school.[40] We did not find any increasing or decreasing monotonic pattern between the QOP impacts in the schools and the relative size of the treatment group within the entering 9th-grade school population, suggesting that the spill-over effects are not important.

IV.3. Heterogeneity among Individuals

So far, the paper has focused on the average effect of the program. This section explores whether the impacts vary with the socio-economic characteristics of its participants. Tables 7 through 9 show subgroup program impact estimates by age at 9th grade, gender, and sites' funding source, respectively.[41] Impacts for subgroups defined by youth 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

---

[40] While in two schools, about half of the GPA-eligible youth were assigned either to the treatment or the control group; in the other seven schools, the ratio of QOP students to the GPA-eligible students ranged between one tenth and a third. As GPA-eligible youth were those in the bottom two-thirds of the 8th-grade distribution, QOP students represented between 6% and one third of the schools' 9th grade entering cohort.

[41] In the Appendix section A.IV, impact estimates for subgroups defined by youth's 8th-grade GPA are also presented and discussed.

20

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. The estimates show that QOP was effective for the younger enrollees, while no significant effects were found for the older ones. Although this differential result by age holds all along, the short-term impacts are not statistically significantly different from each other. In contrast, there is a statistically significant differential treatment effect by age in the medium- and long-run.

In the short-run, QOP increased by 12 percentage points the likelihood that younger enrollees graduated from high-school. Although this effect is slightly reduced in the longer-term, it remains sizeable, persistent, and statistically significant. Five years after the end of the program, younger QOP participants were 7 percentage points more likely to have a high-school degree and 6 percentage points more likely to have a high-school degree or a GED than younger members of the control group.[42] QOP also had a long-lived positive impact on post-secondary education attendance of younger enrollees as it increased their likelihood of attending college or post-secondary education by 7 percentage points in either case in the short-term, and it increased the likelihood of having ever attended college by 12 percentage points two years after the end of the program, and of having ever attended post-secondary education by 10 percentage points five years after the end of the program.

Although the medium-term employment impacts of QOP on younger enrollees are negative and statistically significant—QOP reduced the likelihood of being employed and of having a good job (defined as a job with health insurance) by 9 and 14 percentage points, respectively, the fact that they come hand in hand with positive and sizeable (non-significant) impacts of QOP on college attendance seems to suggest that the employment effects of QOP are due to locking-in effects of training (as opposed to detrimental effects of QOP that could have been driven by extrinsic incentives crowding out intrinsic motivation). This interpretation is confirmed by the longer-term impacts. Although none of the effects of QOP on employment for the younger enrollees are statistically significant, they are all positive reflecting a sharp shift from the effects observed three years earlier and opening the possibility that the returns from higher education

---

[42] However, no effects were found on short-term achievement tests, and GPA.

achievements are beginning to feed in the employment and earnings outcomes.[43]

For older enrollees, QOP did not significantly impact high-school completion or post-secondary education or training. While QOP increased the short-term likelihood of attending post-secondary education or working by 10 percentage points for older enrollees, such effect fades away soon thereafter. Finally, although the longer-term effects of QOP on high-school completion and employment are negative, they are not statistically significant, implying no detrimental effects of QOP among this group.

Impacts by Sex

Table 8 considers the effects of heterogeneity with respect to sex. As is frequently found in the literature, QOP had a significantly differential effect by sex. Explanations for this differential impact include the fact that young women may have more self-discipline (Duckworth and Seligman, 2006), may be more likely to delay gratification (Silverman, 2003), or may have lower discount rates than young men (Warner and Pleeter, 2001). Such differential pattern was statistically significant in the short-, medium-, and longer-term. The following discussion first summarizes the results for females, and subsequently discusses the findings for males.

While QOP had beneficial short-term effects on educational outcomes for female enrollees, these effects were short-lived. For instance, QOP significantly increased by 9 percentage points the likelihood that a female enrollee graduated from high school during the last year of the program. QOP's other significant short-term impact on females was a 9 percentage-point increase in the likelihood of engaging in post-secondary education or training, attending high school or a GED class, or working. However, by the time of the 2nd survey, when students were in their early twenties, QOP no longer had any effect on females' high-school graduation or completion, indicating that female students from the control group had caught up to treatment group. Similarly, the medium-term impacts of QOP on females' post-secondary education are small and not significant, except for a not negligible statistically significant 10 percentage points increase in the likelihood of ever having attended post-secondary education or training through its effects on the likelihood of attending vocational or technical school, an apprenticeship, or enlisting in the armed forces.

---

[43] This occurs because individuals who delay employment to obtain further education or training may initially have lower earnings upon entering the labor force than similarly aged but less educated individuals who have accrued more work experience (Mincer, 1974).

The long-term impacts reinforce this notion that QOP got its female enrollees out of high school, into post-secondary training (not necessarily college), and eventually into employment quicker than females from the control group. For instance, five years after the end of the program, QOP female enrollees were 10 percentage points more likely than females from the control group to have a job with health insurance. Although promising, these results call for some caution as they come hand in hand with QOP enrollees being 11 (or 8) percentage points *less* likely to attend college (or post-secondary training) than females from the control group.

While QOP seemed to have made its female enrollees work through the system faster than non-enrollees, the evidence seems to suggest that QOP worked differently for male enrollees. First, QOP had no beneficial short-term impacts on male students and one significant detrimental effect on high-school achievement: it decreased males' GPA average by 0.13. This detrimental impact of QOP on male enrollees' achievement provides evidence consistent with extrinsic financial rewards crowding out intrinsic motivation. Supporting this interpretation is the evidence that QOP increased males' likelihood of risky behaviors in the short-term. As shown in Appendix table A.7, QOP increased by 7 percentage points the likelihood that males in their late teens binge drank over the month preceding the survey.[44]

Just as the short-term beneficial educational effects of QOP on females quickly faded away, QOP's detrimental short-term effects on males also seemed to be short-lived. By the time of the 2nd telephone survey, when students were in their early twenties, QOP had increased by 7 percentage points males' likelihood of getting a GED diploma, and it had increased by 9 percentage points the likelihood of ever going to college. Again, caution is needed when observing the *negative* 10 percentage points statistically significant effect of QOP on the likelihood of being employed for male enrollees as they are also 6 percentage points (albeit not statistically significantly so) more likely to be attending college.

Longer-term impacts measured by the time of the 3rd telephone survey suggest that these medium-term negative effects on employment were long lasting effects, reinforcing the hypothesis of extrinsic motivation crowding out intrinsic motivation. For instance, QOP male enrollees were 15 percentage points less likely to have a job with health insurance by the time of the 3rd telephone survey. More concerning are the findings on detrimental criminal activity, which are also supportive of extrinsic rewards crowding out intrinsic motivation. Results from

---

[44] "Binge" drinking means drinking five or more drinks in a row.

Appendix Table A.7 show that QOP increased by 10 and 12 percentage points the likelihood of committing a crime and being arrested or charged, respectively, when youth were in their mid-twenties. These results are also consistent with earlier evidence on the low effectiveness of most high-school dropout prevention programs in the United States (Dinarsky and Gleason, 2002). Moreover, they suggest that even if the adverse impacts of financial rewards may seem (a priori) short-lived, recovering from them and catching up is not obvious, as getting a GED diploma does not seem to be a good substitute to a high-school diploma (see Cameron and Heckman, 1993; Boesel *et al.*, 1998, and Heckman and Rubinstein, 2001, for evidence on the low rate of return to a GED certification), and given the longer-term consequences on criminal activity that seem to emerge.

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 9. This analysis led to compare Philadelphia and Yakima to the remaining five sites. Overall, we found that DOL-funded sites performed better than Ford-funded sites in the longer-run.

While Ford-funded sites had short-term beneficial high-school and post-secondary education outcomes, these were short lived as they had faded away by the time of the 3rd telephone survey. In the short-term, QOP increased by nearly 2 percentile points the average mathematics achievement test score, and by (a non-statistically significant) 7 percentage points the likelihood of graduating from high-school in Ford-funded sites. QOP also increased by 14 percentage points the likelihood of engaging in post-secondary education or training and by 17 percentage points the likelihood of engaging in postsecondary education or training or working at a good job, that is, a job offering employer-sponsored health insurance. However, by the time of the 2nd telephone survey, the high-school graduation impact had faded away (becoming a negative non-statistically significant 6 percentage points estimate), and the only beneficial outcomes (albeit not significant) were a 15 (or 17) percentage points increase in the likelihood of ever attending college (or post-secondary education). By the time of the 3rd telephone survey, such large effect on post-secondary education or training impacts had practically faded away.

For DOL-funded sites, QOP increased by 7 percentage points the likelihood of high-school graduation in the short-run, and although this impact decreases and is no longer significant, it remains positive. Moreover, QOP had lasting impacts on post-secondary

education. Five years after the end of the intervention, the DOL-funded sites increased by 8 percentage points the likelihood of enrollees' ever engaging in any postsecondary education and by 7 percentage points the proportion of enrollees who completed at least two years of post-secondary training (not shown).

Differences in implementation between the DOL- and Ford-funded sites might help disentangle the mechanisms behind these results. We identified three main differences between DOL- and Ford-funded sites. First, because the Ford Foundation was one of the designers of the QOP model and was involved in the pilot of QOP in the early 1990s (Hahn *et al*., 1994), Ford-funded sites had a greater depth of understanding of the QOP approach to youth development. This implied that fidelity to the program model was highest in the Ford-sites. In contrast, in the DOL-funded sites the lesser expertise on the QOP program (and its complexity, as it was an unusually intensive program) led to a greater flexibility in the implementation. Second, while case managers in DOL-funded sites were employed full-time and had a flat wage, those in Ford-funded sites worked part-time and therefore had also non-QOP job responsibilities. In addition, they 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. Finally, because of Federal regulation restrictions, accrual accounts did not bear interest at DOL-funded sites and enrollees at those sites did not receive regular statements of their accrual accounts, thereby reducing the incentive value of the accounts.[45] In contrast, Ford-funded sites regularly provided account statements to enrollees with the current accumulated balance and a projection of the amount the enrollee could accumulate by the end of the program if he or she participated in QOP for a specified number of hours per month.[46] While it is not possible to identify which element (or combination of elements) explains the differential results by funding source, all in all, the implementation differences between DOL- and Ford-funded sites seem to suggest that, where the financial rewards were better implemented (through incentives on QOP staff and periodic information on the amount of the rewards), the large beneficial short-term outcomes were short-lived suggesting that external incentives reduced intrinsic motivation. This result is supported by greater short- and longer-term significant detrimental effects on several

---

[45] DOL-funded sites relied on case managers to emphasize to enrollees the importance of accrual accounts. However, case managers typically did not have the information to calculate current balances or to project potential balances.

[46] 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.

measures of intrinsic motivation, as shown in Appendix Table A.7—such as, greater substance abuse in the 1st telephone survey, greater criminal activity both in the 1st and 3rd telephone survey, or poorer self-reported health (not shown)—, in Ford-funded sites than in DOL-funded sites (despite the small sample sizes, many of these differences in impacts between sites were statistically significantly so).[47]

V. Conclusion

Recently, a new wave of interventions whose main objective is to improve the educational outcomes of disadvantaged youth has found promising results if not for all participants, for some particular subgroups. While most of these studies analyze the effects of the interventions during or shortly after the students have been exposed to the program, this paper investigates not only the short-term, but also the medium- and longer-term impacts of a five-year intensive after-school program on at-risk youth in the United States using a randomized experimental design. The paper focuses on both educational and employment outcomes: (i) during the last 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 estimating the average effects of the program, the study finds that the beneficial shorter-term effects of QOP on high-school completion and post-secondary education enrollment quickly fade away. The short-lived program benefits are disappointing given the program's cost of $25,000 per enrollee for the full five years of the demonstration. However, the subgroup analysis reveals that heterogeneity matters. First, QOP had lasting beneficial effects on younger youth, giving some hope for interventions in secondary school education. Second, QOP's beneficial effects on female students quickly faded away, as youth from the control group eventually caught up to those of the treatment group. Third, QOP had detrimental short-term impacts on male students' academic achievement. And although, QOP led to higher GED diploma in the longer-term, it also implied detrimental employment outcomes five years after the end of the intervention, suggesting that financial rewards had detrimental effects on intrinsic motivation. Additional

---

[47] As explained in footnote 38 above, the significant benefitial effect of QOP on binge drinking when youth were in their early twenties in the Ford sites is mainly due to a surprisingly and extremely low probability among QOP youth from the Philadelphia site.

evidence by sites' funding source, which led to important implementation differences, supports the hypothesis of short-lived extrinsic motivation.

This paper adds to the growing interest in student incentive programs in primary and secondary education, and is the first real-world context randomized evaluation in the area of youth's educational outcomes to provide evidence that the short-term benefits in response to rewards may well fade away—as feared by cognitive psychologists and behavioral economists. Its results are in line with those found in field and laboratory experiments previously reviewed in the introduction (Deci, 1975; Kruglanski, Friedman and Zeevi, 1971; Lepper, Greene and Nisbett, 1973; and Gneeze and Rustichini, 2000, among others), and suggest that further research on whether and how cash incentives work to motivate students is needed, especially among the younger and more malleable ones.

In addition, this study suggests the following three recommendations for future program design and evaluation.[48] First, the findings in this paper highlight the relevance of measuring impacts at different points in time that include not only educational, but also employment outcomes. 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 employment outcomes in the longer-run. Moreover, observing impacts over time enables us to better identify the mechanisms through which different components of the program may work.[49] Second, this paper highlights that heterogeneity matters and that evaluations ought to be designed such that sample sizes are large enough to identify for whom the program works, as these has important implication for future policy targeting. In addition, heterogeneity analysis combined with economic theory and institutional knowledge may help disentangle the mechanisms through which programs work. However, (and third) to truly identify channels, evaluations ought to be designed to test alternative mechanisms, as is currently and increasingly done in evaluations in developing countries (Duflo, Dupas, and Kremer, 2009; Abhijit Banerjee, Rukmini Banerji, Esther Duflo, and Stuti Khemani, 2008; Duflo, Kremer, and Robinson, 2009, among others).[50]

---

[48] Clearly, the proposed suggestions are costly and, frequently, budgetary restrictions call for middle ground solutions.

[49] While follow-up survey data is extremely expensive, administrative long-term data ought to be cheaper, easier to access and with fewer measurement and response problems (although also more limited in the range of outcome information available).

[50] To the best of my knowledge, Angrist *et al.*, 2009, are the only ones to analyze an experimental evaluation designed to improve academic performance assigning separate types of services to different treatment groups. They find evidence of lasting improvements for college freshmen women in Canada only when education, peer advising

and financial services are combined.

TABLE 1

QOP's Developmental Activities, Community Services and Educational Services

Activity	*Examples of such types of activities*
Developmental	Life skills activities/ discussion topics (such as, family planning, nutrition, personal hygiene, managing anger, avoiding drug behaviors, among others); pre-employment training; cultural activities; and recreational activities.
Community 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 9th 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)

	QOP group (1)	Control group (2)
Pre-program Characteristics		
Male	52	56
Age when entering 9th grade		
< 14	11	11
14	53	57
> 14	36	31
Hispanic	26	26
Black	68	68
Rank based on 8th 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

QOP demonstration implementation				Post-demonstration	
Fall 1995	*Spring 1999 On time graduation*	*November1999- June 2000*	*September 2000*	*September 2002- April 2003*	*January- September 2005*
Youth entered 9th grade	Paper survey on resiliency factors and Achievement tests in math and reading	1st telephone survey	End of the program	2nd telephone survey	3rd 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

	Cumulative Years 1 through 4	Year 1	Year 4
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)

OUTCOMES	1st telephone survey	2nd telephone survey	3rd telephone survey
High-school performance			
Earned high-school diploma	7* [0.05]	0 [0.91]	0 [0.91]
Earned high-school diploma or GED	5 [0.19]	2 [0.45]	2 [0.45]
Math test scores (percentile)	0.38 [0.43]	--	--
Reading test scores (percentile)	0.50 [0.33]	--	--
GPA	-0.06 [0.18]	--	--
Post-secondary training			
Earned a bachelor's or associate's degree	--	--	-1 [0.68]
Attending college	3 [0.23]	1 [0.32]	-4 [0.28]
Attending or accepted in college	6* [0.08]	--	--
Attending post-secondary education	6* [0.07]	4 [0.56]	-1 [0.76]
Attending post-secondary education or working	3 [0.22]	-1 [0.64]	1 [0.81]
Attending post-secondary education or working in a job with health insurance	5 [0.19]	3	--
Ever in college	--	7* [0.08]	4 [0.30]
Ever in post-secondary education	--	9** [0.03]	6 [0.14]
Employment			
Has a job	--	-7* [0.06]	-1 [0.86]
Has a job with health insurance	--	-2 [0.48]	-3 [0.41]
Has a job that pays more than $10 / hour	--	-2 [0.50]	--
Has a job with paid-off time	--	--	-2 [0.53]
Has a job with pension or retirement benefits	--	--	-1 [0.24]
Ever employed	--	--	0 [0.24]
Total earnings in the past 12 months (dollars)	--	--	-522 [0.66]
Hourly earnings (dollars)	--	--	-1.20 [0.83]
Sample size[a]	893	788	793

Note: All figures were calculated using weights to adjust for the sample and survey designs. P-value in parenthesis.
 *, ** Estimate significantly different from zero at the 90% or 95% confidence level, two-tailed test. [a]Except for high school diploma or GED, in which data has been complemented with high-school transcript information.

TABLE 7
Short-, medium and long-term impacts of QOP, by age
(Percentage points except where noted)

	Age entering 9th grade					
OUTCOMES	Greater than 14 years old			14 years old or less		
	1st telephone survey	*2nd telephone survey*	*3rd telephone survey*	*1st telephone survey*	*2nd telephone survey*	*3rd telephone survey*
High-school performance						
Earned high-school diploma[a]	0	-8†	-8†	12***	7*†	7*†
Earned high-school diploma or GED[a]	6	0	0	7	6*	6*
Math test scores (percentile)	0.55	--	--	0.72	--	--
Reading test scores (percentile)	0.96	--	--	0.67	--	--
GPA	-0.06	--	--	-0.02	--	--
Post-secondary training						
Earned a bachelor's or associate's degree	--	--	-2	--	--	0
Attending college	3	-7††	-3	5	5††	-4
Attending or accepted in college	5	--	--	7*	--	--
Attending post-secondary education	7	-1	0	7*	6	-1
Attending post-secondary education or working	10*	0	-3	1	-3	5
Attending post-secondary education or working in a job with health insurance	8	1	--	6	2	--
Ever in college	--	-4†	0	--	12***†	7
Ever in post-secondary education	--	4	2	--	10**	10**
Employment						
Has a job	--	-5	-4	--	-9**	3
Has a job with health insurance	--	-7	-8	--	-14*	1
Has a job that pays more than $10 / hour	--	0	--	--	-2	--
Has a job with paid-off time	--	--	-11†	--	--	4†
Has a job with pension or retirement benefits	--	--	-7	--	--	2
Ever employed	--	--	-1	--	--	2
Total earnings in the past 12 months (dollars)	--	--	-1524	--	--	465
Hourly earnings (dollars)	--	--	-4.24†	--	--	0.99†

Notes: All figures were calculated using weights to adjust for the sample and survey designs. Achievement test scores are expressed as percentiles in the distribution of scores for tenth graders in the United States. *, **, *** Estimate significantly different from zero at the 90%, 95%, or 99% confidence level, two-tailed test. † Significantly different from the impact for the other subgroups at the 90% confidence level, two-tailed test. [a]High school diploma or GED outcomes have been complemented with high-school transcript information.

TABLE 8

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

OUTCOMES	Males			Females		
	1st telephone survey	*2nd telephone survey*	*3rd telephone survey*	*1st telephone survey*	*2nd telephone survey*	*3rd telephone survey*
High-school performance						
Earned high-school diploma[a]	4	1	1	9*	-1	-1
Earned high-school diploma or GED[a]	0	7*	7*	7	-2	-2
Math test scores (percentile)	0.72	--	--	0.08	--	--
Reading test scores (percentile)	1.10	--	--	0.06	--	--
GPA	-0.13**	--	--	-0.02	--	--
Post-secondary training						
Earned a bachelor's or associate's degree	--	--	-2	--	--	0
Attending college	3	6	$1^{††}$	2	-3	$-11^{***††}$
Attending or accepted in college	4	--	--	5	--	--
Attending post-secondary education	7	5	$4^{†}$	4	3	$-8^{*†}$
Attending post-secondary education or working	$-2^{†}$	-3	-2	$9^{**†}$	2	4
Attending post-secondary education or working in a job with health insurance	4	2	--	6	7	--
Ever in college	--	9*	6	--	3	-1
Ever in post-secondary education	--	6	4	--	10*	6
Employment						
Has a job	--	-10*	-7	--	-2	6
Has a job with health insurance	--	-5	$-15^{**†††}$	--	4	$10^{*†††}$
Has a job that pays more than $10 / hour	--	$-10^{*†††}$	--	--	$7^{†††*}$	--
Has a job with paid-off time	--	--	$-11^{*††}$	--	--	$8^{††}$
Has a job with pension or retirement benefits	--	--	$-11^{*††}$	--	--	$9*^{††}$
Ever employed	--	--	1	--	--	-0
Total earnings in the past 12 months (dollars)	--	--	-1,479	--	--	868
Hourly earnings (dollars)	--	--	1.41	--	--	-0.68

Notes: All figures were calculated using weights to adjust for the sample and survey designs. Achievement test scores are expressed as percentiles in the distribution of scores for tenth graders in the United States. *, **, *** Estimate significantly different from zero at the 90%, 95%, or 99% confidence level, two-tailed test. †, ††, ††† Significantly different from the impact for the other subgroups at the 90% , 95% or 99% confidence level, two-tailed test. [a]High school diploma or GED outcomes have been complemented with high-school transcript information.

TABLE 9

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

OUTCOMES	Ford			DOL		
	1st telephone survey	*2nd telephone survey*	*3rd telephone survey*	*1st telephone survey*	*2nd telephone survey*	*3rd telephone survey*
High-school performance						
Earned high-school diploma[a]	7	-6	-6	7[*]	3	3
Earned high-school diploma or GED[a]	6	-4	-4	4	4	4
Math test scores (percentile)	1.96[*†]	--	--	-0.24[†]	--	--
Reading test scores (percentile)	1.3	--	--	0.21	--	--
GPA	-0.02	--	--	-0.08	--	--
Post-secondary training						
Earned a bachelor's or associate's degree	--	--	-2	--	--	0
Attending college	8	8	-4	1	-2	-4
Attending or accepted in college	11	--	--	3	--	--
Attending post-secondary education	14[*]	13	-7	3	1	1
Attending post-secondary education or working	4	-3	-1	3	0	2
Attending post-secondary education or working in a job with health insurance	17[†**]	6	--	0	2	--
Ever in college	--	15	3	--	4	4
Ever in post-secondary education	--	17	1	--	5	8[*]
Employment						
Has a job	--	-17[*]	0	--	2	-1
Has a job with health insurance	--	-9	-3	--	1	-3
Has a job that pays more than $10 / hour	--	-1	--	--	-3	--
Has a job with paid-off time	--	--	3	--	--	-4
Has a job with pension or retirement benefits	--	--	-4	--	--	-1
Ever employed	--	--	-2	--	--	1
Total earnings in the past 12 months (dollars)	--	--	-3,340	--	--	606
Hourly earnings (dollars)	--	--	-4.34	--	--	0.05

Notes: All figures were calculated using weights to adjust for the sample and survey designs. Achievement test scores are expressed as percentiles in the distribution of scores for tenth graders in the United States. *, **, *** Estimate significantly different from zero at the 90%, 95%, or 99% confidence level, two-tailed test. †, ††, ††† Significantly different from the impact for the other subgroups at the 90% , 95% or 99% confidence level, two-tailed test. [a]High school diploma or GED outcomes have been complemented with high-school transcript information.

REFERENCES

Angrist, Joshua, Eric Bettinger, Erik Bloom, Beth King, and Michael Kremer. 2002. "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment." *American Economic Review*, 92(5): 1535-1558.

Angrist, Joshua, Eric Bettinger, and Michael Kremer. 2006. "Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia." *American Economic Review* 96(3): 847-862.

Angrist, Joshua, Daniel Lang, and Philip Oreopoulos. Forthcoming. "Incentives and Services for College Achievement: Evidence from a Randomized Trial." *American Economic Journal: Applied Economics.*

Angrist, Joshua D., and Victor Lavy. Forthcoming. "The Effect of High School Matriculation Awards: Evidence from Randomized Trials." *American Economic Review.*

Ash, Katie. 2008. "Promises of Money meant to heighten Student Motivation." *Education Week*, February 13.

Ashworth, K., J. Hardman, et al. 2001. "Education Maintenance Allowance: The First Year, A Qualitative Evaluation". Research Report RR257, Department for Education and Employment.

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.

Banerjee, Abhijit, Rukmini Banerji, Esther Duflo, and Stuti Khemani. Forthcoming. "Pitfalls of Participatory Programs: Evidence from a randomized evaluation in education in India." *American Economic Journal: Economic Policy.*

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.

Becker, Gary. 1964. "Human capital; a theoretical and empirical analysis, with special reference to education." New York: Columbia University Press.

Bénabou Roland, and Jean Tirole. 2003. "Intrinsic and Extrinsic Motivation." *Review of Economic Studies* 70, 489–520.

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, 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.

Bettinger, Eric & Slonim, Robert, 2007. "Patience among children," *Journal of Public Economics*, Elsevier, vol. 91(1-2), pages 343-363, February.

Binder, M., P. T. Ganderton, et al. 2002. "Incentive Effects of New Mexico's Merit-Based State Scholarship Program: Who Responds and How?", unpublished manuscript.

Bloom, H.S., Bell, S.H., Doolittle, F., Lin, W., Cave, G., and Orr, L.L. 1996. "Does training for the disadvantaged work? Evidence from the national JTPA study." Washington, DC: The Urban Institute Press.

Bloom, Dan and Colleen Sommo. 2005. "Building Learning Communities: Early Results from the Opening Doors Demonstration at Kingsborough Community College," New York: MDRC.

Bonner, Sarah E. S., Mark Young, and Reid Hastie. 1996. ''Financial Incentives and Performance in Laboratory Tasks: The Effects of Task Type and Incentive Scheme Type,'' Unpublished manuscript, University of Southern California Department of Accounting.

Camerer, Colin, and Robin Hogarth. 1999. "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework." *Journal of Risk and Uncertainty*, 19:1-3; 7-42.

Cameron, J., K. M. Banko, et al. 2001. "Pervasive Negative Effects of Rewards on Intrinsic Motivation: The Myth Continues." *The Behavior Analyst* 24: 1-44.

Cameron, S., & Heckman, J. 1993. The nonequivalence of high school equivalents. *Journal of Labor Economics, 11*(1), 1–47.

Carneiro, Pedro, and James Heckman. 2003. "Human capital policy." Working Paper no. 9495, National Bureau of Economic Research, Cambridge, MA.

Cascio, Elizabeth U., and Ethan G. Lewis. 2006. "Schooling and the Armed Forces Qualifying Test: Evidence from School-Entry Laws." *Journal of Human Resources* 41(2): 294–318.

Cornwell, Christopher, Kyung Hee Lee, and David B. Mustard. 2005. "The effects of merit-based financial aid on course enrollment, withdrawal, and completion in college." *Journal of Human Resources* 40, no. 4:895–917.

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.

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

Currie, Janet, and Duncan Thomas. 2001. "Early test scores, school quality, and SES: Long-run effects on wage and employment outcomes." In *Worker wellbeing in a changing labor market*, Research in Labor Economics, vol. 20, ed. Solomon W. Polachek, 20:103–32. New York: Elsevier Science.

Deci, E. 1975. *Intrinsic Motivation* (New York: Plenum Press).

Deci, E., Koestner, R. and Ryan, R. 1999. "A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation", *Psychological Bulletin*, 125 (6), 627–668.

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.

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.

Duflo, Esther, Pascaline Dupas and Michael Kremer, 2009. "Additional Resources versus Organizational Changes in Education: Experimental Evidence from Kenya." Mimeo, MIT, Department of Economics.

Duflo, Esther, Michael Kremer, and Jonathan Robinson, 2009. "Nudging Farmers to Utilize Fertilizer: Theory and Experimental Evidence from Kenya." Mimeo, MIT, Department of Economics.

Dynarski, Mark, and Gleason, Philip. 2002. "How Can We Help? What We Have Learned From Recent Federal Dropout Prevention Evaluations." *Journal of Education for Students Placed at Risk*, 7 (1): 43-69.

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

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

El País, 2009. "Chequera Contra el Fracaso Escolar." Reportaje de J.A. Aunión, 14 October, 2009.

Eckstein, Zvi and K.I. Wolpin. 1999. "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities." *Econometrica,* 67(6): 1295-1340.

Finn, J. (1989). Withdrawing From School. *Review of educational research, 59*, 117–142.

Frangakis, Constantine E., and Donald B. Rubin. 2002. "Principal stratification in causal inference." *Biometrics*, 58: 21-29.

Frey, B. S. 1994. "How Intrinsic Motivation Is Crowded Out and In." *Rationality & Society*, 6(3): 334-352.

Frey, B. S. and Oberholzer-Gee, F. 1997. "The costs of price incentives: An empirical analysis of motivation crowding-out." *American Economic Review*, 87(4):746–755.

Fryer, Roland. 2006. "Incentivising: an intuitive approach to raising achievement." Ongoing research.

Garibaldi, Pietro, Francesco Giavazzi, Andrea Ichino, and Enrico Rettore. 2007. "College Cost and Time to Obtain a Degree: Evidence from Tuition Discontinuities." NBER Working Paper 12863.

Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. "Longer term effects of Head Start." *American Economic Review* 92, no. 4:999–1012.

Gerfin M., and M. Lechner, and H. Steiger. 2005. "Does Subsidized Temporary Employment Get the Unemployed Back to Work? An Econometric Analysis of Two Different Schemes." *Labour Economics,* Vol. 12, pp. 807-835.

Gneezy, U. and Rustichini, A. 2000. "Pay Enough or Don't Pay at All", *Quarterly Journal of Economics*, 115 (3), 791–810.

Gibbons, R. 1997. "Incentives and Careers in Organizations", in D. Kreps and K.Wallis (eds.) *Advances in Economic Theory and Econometrics*, Vol. II (Cambridge, U.K.: Cambridge University Press).

Granger, Robert, and Rachel Cytron. 1998. "Teenage Parent Programs." MDRC report.

Green, L., Fry, A. F., and Myerson, J. 1994. Discounting of delayed rewards: A life span comparison. *Psychological Science, 5*, 33-36.

Greene, A. 1986. "Future Time Perspective in Adolescence: The present of things future revisited", *Journal of Youth and Adolescence*, 15: 99-113.

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.

Gruber, Jonathan. 2001. *Risky Behavior Among Youth: An Economic Analysis.* Chicago: University of Chicago Press.

Hall, J. 2003. "Mentoring and Young People: A literature review." *SCRE Research Report 114*. University of Glasgow.

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, Waltham, Massachusetts.

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

Heckman, James J & Lochner, Lance & Taber, Christopher, 1998. "Tax Policy and Human-Capital Formation," *American Economic Review*, vol. 88(2), pages 293-97, May.

Heckman, James J., and Yona Rubinstein. 2001. "The Importance of Non-Cognitive Skills: Lessons from the GED testing program. *American Economic Review* 91, no. 2:145–49.

Hertwig, Ralph and Andreas Ortmann. 1998. ''Experimental Practices in Economics: A Methodological Challenge for Psychologists,'' *Behavioral and Brain Sciences*.

Holmlund Helena, and Olmo Silva. 2009. "Targeting Non-Cognitive Skills to Improve Cognitive Outcomes: Evidence from a Remedial Education Intervention." IZA DP No. 4476.

James-Burdumy Susanne, Mark Dynarski and John Deke. 2008. "After-School Program Effects on Behavior: Results From the 21st Century Community Learning Centers Program National Evaluation." *Economic Inquiry* , Vol. 46, No. 1, January 2008, 13–18

Jacob, B. and Lefgren L. 2004. "Remedial Education and Student Achievement: A regression Discontinuity Analysis", *Review of Economics and Statistics*, vol. 86, pp. 226-44.

Jackson, C. K. 2007. "A Little Now for a Lot Later: A Look at a Texas Advanced Placement Incentive Program." http://works.bepress.com/c_kirabo_jackson/1.

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.

Jenkins, G. Douglas, Jr., Atul Mitra, Nina Gupta, and Jason D. Shaw. 1998. ''Are Financial Incentives Related to Performance? A Meta-Analytic Review of Empirical Research,'' *Journal of Applied* P*sychology* 83, 777-787.

Kane, Thomas J. 1998. "Savings Incentives for Higher Education." *National Tax Journal*, Vol. 51, No. 3, pp. 609-620.

Kane, Thomas J. 2007. "Evaluating the Impact of the DC Tuition Assistance Grant Program." *Journal of Human Resources*.

Kemple, James J., and Cynthia J. Willner. 2008. "Career Academies. Long-Term Impacts on Labor Market Outcomes, Educational Attainment, and Transitions

to Adulthood" MDRC Report.

Kirby, K. N. 1997. "Bidding on the future: Evidence against normative discounting of delayed rewards." *Journal of Experimental Psychology: General* 126, 54-

Kohn, A. 1993. *Punished by Rewards* (New York: Plenum Press).

Kohn, Alfie. 1999. *Punished by Rewards the Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes*. Bridgewater, NJ: Replica Books.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. Forthcoming. "Incentives to Learn." *The Review of Economics and Statistics*.

Kruglanski, A., I. Friedman, and Zeevi. 1971. "The Effect of Extrinsic Incentives on Some Qualitative Aspects of Task Performance." *Journal of Personality and Social Psychology* 39: 608-617.

Kluve, Jochen. 2006. "The Effectiveness of European Active Labor Market Policy." IZA Discussion Paper 2018.

Krueger, Alan B., and Diane M. Whitmore. 2001. "The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from Project STAR." *Economic Journal* 111, no. 468:1–28.

Laibson, David, 1997. "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, vol. 112(2), pages 443-77.

LaLonde, Robert J. 1995. "The Promise of U.S. Employment and Training Programs." *Journal of Economic Perspectives*, 9(2): 149–68.

Lazear, E. (2000), "Performance, Pay and Productivity", *American Economic Review*, 90 (5), 1346–1361.

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

Lepper, M., D. Greene, et al. 1973. "Undermining Children's Interest with Extrinsic Rewards: A Test of the 'Overidentification Hypothesis." *Journal of Personality and Social Psychology* 28: 129-137.

Lepper, M., Greene, D. and Nisbett, R. 1973. "Undermining Children's Interest with Extrinsic Rewards: A Test of the 'Overjustification Hypothesis'", *Journal of Personality and Social Psychology*, 28, 129–137.

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.

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

Machin, Steven, Sandra McNally, and Costas Meghir, 2007. "Resource and Standards in Urban Schools", IZA DP 2653.

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.

Mincer, Jacob. 1974. *Schooling, Experience, and Earnings.* New York: Columbia University Press for the National Bureau of Economic Research.

Nurmi, J. 1991. "How Do Adolescents See Their Future? A review of the development of future orientation and planning", *Developmental Review*, 11:1-59.

Rhodes, Jean. 1994. "Older and Wiser: Mentoring Relationships in Childhood and Adolescence," *The Journal of primary prevention*, v 14 n 3, Spring.

Rhodes, Jean E., Jean B. Grossman, and Nancy L. Resch. 2000. "Agents of Change: Pathways through Which Mentoring RelationshipsInfluence Adolescents' Academic Adjustment." *Child Development*, November/December, Volume 71, Number 6, Pages 1662–1671.

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.

Scrivener, Susan, Michael J. Weiss, and Jedediah J. Teres. 2009. "More Guidance, Better Results? Three-Year Effects of an Enhanced Student Services Program at Two Community Colleges." MDRC report.

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

Titmuss, Richard M. 1970. *The Gift Relationship*. London: Allen and Unwin.

Titmuss, Richard M. 1971. *The Gift Relationship: From Human Blood to Social Policy*. New York: Pantheon.

U.S. Department of Education, National Center for Education Statistics. 2002. "A Descriptive Summary of 1999-2000 Bachelor's Degree Recipients 1 Year Later, With an Analysis of Time to Degree." NCES 2003-165, by Ellen M. Bradburn, Rachael Berger, Xiaojie Li, Katharin Peter, and Kathryn Rooney. Project Officer: James Griffith.

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.

41

Wilson, T., Hull, J. and Johnson, J. 1981. "Awareness and Self-Perception: Verbal Reports on Internal States", *Journal of Personality and Social Psychology*, 40, 53–71.

APPENDIX

(*Not for Publication*)

A.I. Weights

Overall impacts of QOP were estimated using student-, school-, and site-specific weights.[51] 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.[52] 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.[53] 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.

A.II. Sensitivity Analyses to Non-Response Bias

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 be that the observed differences between the shorter- and the longer-term outcomes were due to differential non-response bias between treatment and control group members that would vary across the surveys. 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 lack of positive results in the medium- and long-term outcomes is due to having a higher proportion of "difficult" youth responding in the QOP group relative to the control group.[54] To explore the internal validity of the results presented in the text we have done the following two robustness checks. First, columns 2 and 4 of Table A.2. present unweighted estimates that were derived by making the response rate for the QOP group equal

---

[51] 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.

[52] 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.

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

[54] Notice however that the differential response rate between the treatment and control groups does not always increase across surveys (it is 7 percentage points in the 1st telephone survey, 10 percentage points in the 2nd telephone survey, and 3 percentage points in the 3rd telephone survey.)

43

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 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—as these were those most difficult to find and who had eventually responded because of our intense survey effort. Columns 1 and 3 present the preferred estimates displayed in the paper (in Table 6) for comparison purposes. Overall the estimates in columns 2 and 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.

A second robustness check is to compare the estimate of an outcome that does not change over time across different populations of respondents across surveys. The issue here is to identify an outcome that remains constant across surveys. Fortunately, we can identify such an outcome for the 2nd and 3rd telephone survey that is "high-school graduation". Since by the time of the 2nd telephone survey, more than three years have elapsed between on-time high-school graduation and the survey date, it is very unlikely that individuals who had not yet earned a high-school diploma (and who are about 21 years old) are going to go back and receive a high school diploma. The first row of Table A.3. displays estimates of high-school graduation estimated using *only* the sample members who responded to the 3rd telephone survey, the same sample on which *all* other long-term outcomes reported in the main text are based. The second row of Table A.3 presents estimates of high-school graduation using information on the 3rd telephone survey respondents as well as non-respondents who had reported earning a high-school diploma in a previous survey. In both rows, all sample members' responses to the 3rd telephone survey regarding their high-school completion status are assumed to be accurate. Therefore the only difference between the estimates in rows 1 and 2 is the people who responded, implying that any differences that may appear across the two rows will be due to differential non-response bias between the treatment and the control group changing between the 3rd and the other two surveys. As the differences are minimal, comparing the estimates from rows 1 and 2 suggests that non-response bias did not vary much between the last survey and the two previous ones. The third row of Table A.3 estimates high-school graduation using only information on respondents of the 2nd telephone survey (complemented with 2nd telephone survey non-respondents who had reported earning a high-school diploma in the 1st survey). In this case, comparing the estimates from rows 3 and 2 is equivalent to comparing estimates when only respondents from the first two surveys are observed (row 3) versus when respondents from *all* surveys are observed (row 2). Again

the differences are minimal and if anything they suggest that, had the response rate been the same across surveys, the medium- and longer-term estimates would have been even more detrimental.[55]

---

[55] Notice that the response rate is lower for the comparison group members in the 2nd survey than in the 3rd survey, while the opposite is for the treatment group members.

A.III. Sensitivity to Regression Adjustments

Although the difference-of-means estimates presented in the main paper are unbiased, they may 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.4 presents regression-adjusted impact estimates.   While the regression-adjusted estimates increase the precision of some of the estimates, the main findings and conclusions are unaffected.

A.IV. Impacts by Rank in the 8th-Grade Grade Distribution

Table A.8 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.[56]

According to the short-term impacts shown in Table A.8, QOP was more successful for enrollees in the middle of the distribution than for enrollees at the top or bottom of the distribution. All of the significant short-term impacts on enrollees in the middle third of the baseline grade distribution were beneficial.   The program increased both high-school graduation and completion rates. It increased by 14 percentage points the likelihood of earning a diploma and by 11 percentage points the likelihood of earning a diploma or a GED certificate. QOP also increased by 13 percentage points the likelihood of attending or being accepted into college. All three of these impacts are significantly different from zero. They are larger by at least several percentage points but not significantly different from the impacts for other enrollees.

According to the medium-term impacts, QOP was more successful for enrollees in the bottom and middle thirds of the distribution than for enrollees in the top third of the distribution.   For enrollees in the bottom third of the baseline grade distribution, QOP increased the likelihood of ever attending college by 14 percentage points. The pattern of impacts on post-secondary training is qualitatively similar (although not statistically significant) among enrollees in the middle of the distribution than for those in the bottom of the distribution.   Moreover, QOP increased by 7 percentage points the likelihood of completing at least two years of college for youth in the middle of distribution (not shown).

---

[56] 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.

Finally, although the medium-term impacts of QOP on high-school completion and graduation rates of enrollees in the middle third of the distribution are smaller than in the short-run (and no longer statistically significant), the estimates continue to be a sizeable 6 and 7 percentage points, respectively.[57] In contrast, for enrollees from the top of the baseline grade distribution, QOP had some detrimental impacts but no beneficial educational medium-term impacts. For instance, QOP decreased by 12 percentage points the likelihood of attending college at the time of the 2nd telephone survey.

However, by the time of the 3rd telephone survey, there are very few long-term educational and employment impacts defined by rank in the baseline grade distribution, implying that any beneficial effect of QOP on youth on the lower tail of the baseline grade distribution has faded away. More concerning, QOP decreased by 7 percentage points the likelihood of attending college for enrollees in the bottom of the baseline grade distribution, suggesting that youth in the control group may be catching up to those in the treatment in terms of college enrollment, and providing again evidence that QOP merely speed up the process of getting through the system.

To sum up, the evidence of QOP's impacts by rank in the baseline GPA distribution seems to suggest that QOP affected the timing of college enrollment for youth in the bottom of the baseline grade distribution, but that youth in the control group had caught up with those in the treatment group by the time of the 3rd survey. For youth in the middle of the distribution, QOP seem to have more persistent beneficial (albeit not statistically significant) effects on high-school graduation and completion and college attendance. For youth in the top of the distribution, QOP had no beneficial effects in the short-, medium-, or long-term.

---

[57] When adjusting for random baseline differences using regression methods, I find a statistically significant increase of 9 percentage points in the completion of a high school diploma or GED among QOP enrollees in the middle third of the grade distribution (p-value=0.04). The non-regression-adjusted impact is 7 percentage points and not statistically significant (p-value=0.23).

TABLE A.1

Control group means

	WEIGHTED MEANS		
OUTCOMES	1st telephone survey	2nd telephone survey	3rd telephone survey
High-school performance			
Earned high-school diploma	40	56	60
Earned high-school diploma or GED	49	71	75
Math test scores (percentile)	40.5	--	--
Reading test scores (percentile)	42.7	--	--
GPA	2.19	--	--
Post-secondary training			
Earned a bachelor's or associate's degree	--	--	
Attending college	18	16	17
Attending or accepted in college	25	--	--
Attending post-secondary education	26	26	24
Attending post-secondary education or working	61	78	
Attending post-secondary education or working in a job with health insurance	43	59	--
Ever in college	--	30	34
Ever in post-secondary education	--	53	56
Employment			
Has a job	--	72	68
Has a job with health insurance	--	56	47
Has a job that pays more than $10 / hour	--	45	--
Has a job with paid-off time	--	--	45
Has a job with pension or retirement benefits	--	--	38
Ever employed	--	--	95
Total earnings in the past 12 months (dollars)	--	--	13,198
Hourly earnings (dollars)	--	--	9.14

Note: Estimates were obtained using weights to adjust for differences between respondents and nonrespondents. Achievement test scores are expressed as percentiles in the distribution of scores for tenth graders in the United States. Credits are expressed in Carnegie units that standardize for in-class time. One Carnegie unit corresponds to a class that meets for 45 to 60 minutes every day of the week for an entire academic year.

TABLE A.2

Impacts Using Alternative Approaches to Adjusting for Non-Response
(Percentage points)

	MEANS			
	2nd telephone survey		3rd telephone survey	
OUTCOMES	Preferred Estimates	Same non-response rate for QOP and control group	Preferred Estimates	Same non-response rate for QOP and control group
	(1)	(2)	(3)	(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
Attending post-secondary education or working	-1	-1	1	3
Attending post-secondary education or working in a job with health insurance	3	3	--	--
Ever in college	7*	8**	4	4
Ever in post-secondary education	9**	9**	6	7*
Employment				
Has a job	-7*	-5	-1	-0
Has a job with health insurance	-2	-2	-3	-1
Has a job that pays more than $10 / hour	-2	0	--	--
Has a job with paid-off time	--		-2	1
Has a job with pension or retirement benefits	--		-1	1
Ever employed	--		0	0
Total earnings in the past 12 months (dollars)	--		-522	-349
Hourly earnings (dollars)	--		-1.20	-0.95
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.

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

[a]High school diploma or GED outcomes have been complemented with high-school transcript information.

TABLE A.3

Means and Impacts on High-School Completion
Using Alternative Samples of Respondents

| | WEIGHTED MEANS | | |
OUTCOMES	QOP mean (1)	Control mean (2)	Percentage points (3)
3rd telephone survey respondents only	58	57	1
3rd telephone survey respondents plus 3rd telephone survey non-respondents for whom we have information from earlier surveys or high-school transcripts	60	60	0
2nd telephone survey respondents plus 2nd telephone survey non-respondents for whom we have information from earlier survey or high-school transcripts	58	56	2

Note: Each impact was derived by subtracting the control group mean from the QOP group mean prior to rounding those means; thus, an impact might not equal the difference between the rounded means that are displayed. Estimates were obtained using weights to adjust for differences between respondents and nonrespondents.

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

OUTCOMES	1st telephone survey	2nd telephone survey	3rd telephone survey
High-school performance			
Earned high-school diploma[a]	7**	-1	3
Earned high-school diploma or GED[a]	5	2	4
Math test scores (percentile)	0.52	--	--
Reading test scores (percentile)	0.57	--	--
GPA	-0.04	--	--
Post-secondary training			
Earned a bachelor's or associate's degree	--	--	0
Attending college	3	2	-3
Attending or accepted in college	6*	--	--
Attending post-secondary education	5	5	-0
Attending post-secondary education or working	3	-1	1
Attending post-secondary education or working in a job with health insurance	5	4	--
Ever in college	--	8*	5
Ever in post-secondary education	--	10**	7*
Employment			
Has a job	--	-6	-0
Has a job with health insurance	--	-2	-3
Has a job that pays more than $10 / hour	--	-1	--
Has a job with paid-off time	--	--	-2
Has a job with pension or retirement benefits	--	--	-1
Ever employed	--	--	0
Total earnings in the past 12 months (dollars)	--	--	-211
Hourly earnings (dollars)	--	--	-1.15
Sample size	893	788	793

Note: All figures were calculated using weights to adjust for the sample and survey designs. Achievement test scores are expressed as percentiles in the distribution of scores for tenth graders in the United States. *, ** Estimate significantly different from zero at the 90% or 95% confidence level, two-tailed test. [a]High school diploma or GED outcomes have been complemented with high-school transcript information.

TABLE A.5

Subgroup sample sizes

	Sample sizes
Male	576
Females	493
Age when entering 9th grade	
≤ 14	706
> 14	363
Rank based on 8th grade GPA	
Bottom third	380
Middle third	359
Top third	330

TABLE A.6

Short-, medium and long-term impacts of QOP on Risky Behaviors

(*Percentage points except where noted*)

OUTCOMES	Short-term (1)	Medium-term (2)	Long-term (3)
Drinking in the past 30 days	7**	--	--
Binge drinking in the past 30 days	4	-6	0
Used any illegal drug in the past 30 days	7**	-6**	-0
Committed any crime (1)	3	-2	3*
Arrested or charged (2)	-5	-0	6**
Have first child before age 18	-3	2	2

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

TABLE A.7

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

	Age entering 9th grade					
OUTCOMES	*Short-term*		*Medium-term*		*Long-term*	
	Greater than 14 years old			*14 years old or less*		
Binge drinking in the past 30 days	5	-13*	-7	4	-2	5
Used any illegal drug in the past 30 days	8	-13*	-2	5	-3	-1
Committed any crime (1)	-4	-2	7	5	-2	4
Arrested or charged (2)	-11*	2	9*	-3	-1	5**
Have first child before age 18	6†	4	4	-9**††	-1	-1
	Males			*Females*		
Bing drinking in the past 30 days	7*	-12*	-5	0	-2	4
Used any illegal drug in the past 30 days	7	-8*	-4	7	-3	3
Committed any crime (1)	5	-3	10**†††	2	-1	-4*†††
Arrested or charged (2)	-8	1	12***†††	0	-2	-2†††
Have first child before age 18	-3	3	3	-5	0	0
	Ford			*DOL*		
Binge drinking in the past 30 days	17***†	-13	8	-1†	-2	-3
Used any illegal drug in the past 30 days	14**	-8	-6	4	-5*	2
Committed any crime (1)	16**†	3	7	-2†	-3	4
Arrested or charged (2)	-3	-5	12*	-5	2	4*
Have first child before age 18	-14**†	1	1	1†	4	3

Note: See comments from Table A.6.
 †, ††, and ††† Significantly different from the impact for the other subgroups at the 90%, 95%, 99%, confidence level, two-tailed test, respectively.

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

OUTCOMES	Bottom third			Middle third			Top third		
	1st telephone survey	*2nd telephone survey*	*3rd telephone survey*	*1st telephone survey*	*2nd telephone survey*	*3rd telephone survey*	*1st telephone survey*	*2nd telephone survey*	*3rd telephone survey*
High-school performance									
Earned high-school diploma[a]	3	-1	-1	14**	6	6	4	-3	-3
Earned high-school diploma or GED[a]	1	2	2	11*	7	7	3	-4	-4
Math test scores (percentile)	0.06	--	--	0.28	--	--	0.04	--	--
Reading test scores (percentile)	0.27	--	--	-0.17	--	--	0.91	--	--
GPA	-0.13	--	--	-0.06	--	--	0.03	--	--
Post-secondary training									
Earned a bachelor's or associate's degree	--	--	-1	--	--	-1	--	--	1
Attending college	-2	7	-7*	7	5	-2	4	-12*††	-3
Attending or accepted in college	-3†	--	--	13**	--	--	9	--	--
Attending post-secondary education	9*	5	-3	3	7	-3	4	-2	2
Attending post-secondary education or working	8	-1	-1	1	-2	4	0	3	1
Attending post-secondary education or working in a job with health insurance	7	2	--	9	5	--	4†	4	--
Ever in college	--	14***†	1	--	9	7	--	-5	2
Ever in post-secondary education	--	10	5	--	8	6	--	10	5
Employment									
Has a job	--	-9	1	--	-6	3	--	-1	-5
Has a job with health insurance	--	-8	-5	--	-5	2	--	8†	-5
Has a job that pays more than $10 / hour	--	-7	--	--	-3	--	--	4	--
Has a job with paid-off time	--	--	-2	--	--	-2	--	--	-1
Has a job with pension or retirement benefits	--	--	-3	--	--	-5	--	--	4
Ever employed	--	--	0	--	--	-2	--	--	3
Total earnings in the past 12 months (dollars)	--	--	-1,637	--	--	152	--	--	-571
Hourly earnings (dollars)	--	--	-0.27	--	--	-3.11	--	--	0.34

Notes: All figures were calculated using weights to adjust for the sample and survey designs. Achievement test scores are expressed as percentiles in the distribution of scores for tenth graders in the United States. *, **, *** Estimate significantly different from zero at the 90%, 95%, or 99% confidence level, two-tailed test. †, †† Significantly different from the impact for the other subgroups at the 90% or 95% confidence level, two-tailed test. [a]High school diploma or GED outcomes have been complemented with high-school transcript information.