Educational, Labor Market, and Welfare Impacts of Scholarships for Private Secondary School: Evidence from Colombia

Preliminary Draft: Please do not circulate

By ERIC BETTINGER, MICHAEL KREMER, MAURICE KUGLER, CARLOS MEDINA, CHRISTIAN POSSO AND JUAN ESTEBAN SAAVEDRA*

November 12, 2014

A large-scale government program in Colombia used a lottery to distribute scholarships for private secondary school to socially disadvantaged students. Based on administrative data up to seventeen years after the scholarship lottery, we document that lottery winners are less likely to repeat grades, more likely to graduate from secondary school on time or ever, and more likely to start and complete tertiary education. Scholarships reduce teen fertility, although there is no significant effect on overall fertility at age 30. Among males, there is some evidence (significant at the 10% level) that winners are less likely to qualify for Colombia's conditional cash transfer program. Point estimates suggest that total formal sector earnings at age 30 are 8 percent greater for lottery winners, a difference that is significant at the 7% level. Impacts on estimated future earnings, including imputed values for those currently in tertiary education suggest a 9.3 percent difference that is significant at the 5% level. Preliminary analyses suggest the expected net present value of increased net tax receipts due to the program exceed the program's fiscal cost, and the program is welfare improving as long as externalities on non-recipients are positive, zero, or negative but less than $1,100 per scholarship recipient.

* Bettinger: Stanford Graduate School of Education, 522 CERAS, 520 Galvez Mall, Stanford, CA 94305 (e-mail: ebettinger@stanford.edu); Kremer: Harvard University Department of Economics, Littauer Center, Cambridge MA 02138 (email: mkremer@fas.harvard.edu); Kugler: United Nations Development Program, 304 E 45th Street, FF-12th Floor, New York, NY 10017 (email: maurice.kugler@gmail.com); Medina: Banco de la República, Calle 50 # 50-21, Medellín, Colombia (email: cmedindu@banrep.gov.co); Posso: University of North Carolina Department of Economics CB 3305, Chapel Hill, NC 27599 (email: cpossosu@email.unc.edu); Saavedra: University of Southern California, Dornsife Center for Economic and Social Research 12015 Waterfront Drive, Playa Vista CA 90094 (email: juansaav@usc.edu). We are grateful to Colombia's Ministry of Education, Ministry of Social Protection, Banco de la República and Instituto Colombiano de Evaluación de la Educación (ICFES) for generously providing access to data. We thank Luis Omar Herrera and Arlén Guarín for excellent research assistance. We thank seminar participants at Harvard, MIT, Michigan, Aarhus, UCLA, LSE, Universidad de los Andes in Colombia, RAND Corporation, Society for Research on Educational Effectiveness, University of Arkansas, University of Texas, Austin, University of Michigan, and The World Bank for helpful comments and suggestions. Saavedra acknowledges financial support from the National Institutes of Health RCMAR Grant P30AG043073), RAND Corporation, Universidad de los Andes and the World Bank. The views expressed in this document are solely those of the authors and do not reflect the views of Banco de la República de Colombia.

1. Introduction

During the 1990s, Colombia's PACES program provided more than 125,000 scholarships to disadvantaged secondary school applicants from elementary public schools who wished to attend private schools. Scholarships covered about two thirds of the cost of the typical private participating school. They were renewable through the end of secondary school, conditional on passing each grade.

We take advantage of a lottery used to allocate scholarships, as well as a unique data compiled from five different sources of Colombian administrative data, to document the impact of receiving these scholarships on educational, family formation, and labor market outcomes up to 17 years after initial scholarship award, when applicants were near age 30. We then use these data to estimate the fiscal impact of the program.

Lottery winners are about 7 percentage points more likely to have graduated from secondary school on schedule (on a 44 percent base) and 5 percentage points more likely to have ever graduated from secondary school (on a 55 percent base). Lottery winners were also 3 percentage points more likely to have attended college at some point, a 16 percent relative increase in attendance. They are 58 percent more likely to be enrolled in tertiary education. (Base rate is 4 percent.) Within tertiary education, the secondary school scholarships increased women's enrollment in five-year universities and men's enrollment in two-year vocational institutions.

Point estimates of the scholarship impact on total formal sector earnings at around age 30 suggest that lottery winners earn 8 percent more in formal employment than scholarship lottery losers (p-value = .07). Imputing values for those currently in tertiary education

implies a 9.3 percent difference in expected future formal sector earnings that is significant at the 5 percent level. Effects seem strongest at the top of the distribution.

Winning a scholarship also reduces teen fertility. Among female applicants, lottery winners are 6 percentage points less likely to have had a child as a teen, a 16 percent relative decrease in teen fertility. Among male applicants, lottery winners are 5 percentage points less likely than lottery losers to be the spouse or partner of a female who gave birth as a teen (of a 16 percent base). We find no effect on total fertility as of age 30, consistent with the notion that the teen fertility effect we observe is mostly an "incarceration" effect rather than an opportunity-cost effect.

The program may have operated through several potential channels. These include allowing more children to attend private secondary schools, allowing those who might have attended private school in any case to attend a wider range of schools, providing income subsidies to families who might have been predisposed to attend private school, changing students' peers, and/or incentivizing students and schools to avoid grade repetition. It seems unlikely that the entire scholarship effect is due to schools' lowering standards since we find effects on tertiary education. Additionally, one might not expect the effects at the top of the distribution if the main channel were increased effort by students concerned that they might lose their scholarships by failing a grade. Scholarship impacts on the sub-population of vocational school applicants are comparable to those in the full sample. The fact that in this subpopulation winners attend schools with peers who are less desirable on observables casts doubt on the notion that peer quality was the only mechanism driving observed scholarship effects (Eric Bettinger, Michael Kremer, and Juan E. Saavedra 2010).

Net fiscal costs depend on direct government expenditures on the scholarship as well as on how the program affects the net present value of subsequent public education expenditure, other public subsidies, and tax revenue. We argue that the net fiscal cost of the program is negative due to the indirect effect of scholarship receipt on the net present value of government expenditure and revenue. If one accepts this calculation, then taxpayers are better off, and the program more than paid for itself by increasing future tax revenue and reducing future public expenditures. Although winners may have spent more on schooling or increased effort, it is possible to construct a lower bound on net benefits to lottery winners based on revealed preference, since payments to intra-marginal applicants who would have gone to private school in any case constitute a pure transfer.

For the program to have been negative from a welfare point of view, externalities on non-participants would have to be negative and greater than the sum of the fiscal benefits to taxpayers and the benefits to participants. We lack data to estimate externalities on those who did not receive scholarships, but many plausible channels are positive. For example, it seems likely that the opening up of an avenue for social mobility may also have positive civic effects.[1]

The design of the program and the Colombian context included several features that had the combined effect of reducing its net fiscal cost. First, scholarships cost less than per

[1] To the extent that there are positive externalities from growth of human capital that are not fully internalized within firms, there may be positive externalities on non-participants. To the extent however that gains for program winners reflect assignment to more favorable peers or signaling benefits in the labor market, there might also be negative effects. In earlier work, a subset of the authors of this paper argues against the view that effects are entirely due to changes in peer assignment (Bettinger, Kremer and Juan E. Saavedra 2010). The effects on secondary school test scores seem consistent with the idea that human capital plays a role in the labor markets gains, although we cannot exclude the possibility of some signaling, which would create negative externalities for non-participants.

pupil expenditure in public schools. Therefore, to the extent that scholarship lottery winners switched from public to private education, government outflows were reduced.

Second, scholarships could be augmented with household funds. To the extent that the program "crowded in" household funds for education, increasing winners' human capital and future taxable earnings, the government budget constraint improves. Allowing “top-ups” also avoids creating incentives for some families to trade down from higher cost private schools to private schools with fees at or below the value of the scholarship, which may have reduced human capital accumulation and future taxable income.

Third, as is fairly standard in many scholarship programs, program rules made retention in the program conditional on satisfactory grade completion. Indeed we find that on-time graduation increased and grade retention fell as a result of the scholarship offers. To the extent that reduced repetition led to fewer years of schooling taking place in public schools, public expenditure in education fell. Better scores on national secondary school exams and improved tertiary educational outcomes among scholarship winners suggest that effects are not simply due to schools “moving the goalposts” for student promotion.

Fourth, scholarships were targeted to the poor, reducing the extent to which the simply subsidized students who would have gone to private school anyway. (This poverty targeting will also of course make a distributional-weighted public finance calculation more favorable.)

Our results should be interpreted as suggesting that, at least in the context of 1990s Bogotá, a suitably designed system of offering students from poor neighborhoods partial scholarships for private schools had a negative net fiscal cost. We take no stand on whether changes at the margin to make the program more generous and expansive —increasing the value of the scholarship, removing the conditionality on grade progression, or expanding

eligibility further up the socio-economic ladder— would at the margin have passed a cost-benefit test. Our data and identification strategy do not allow us to answer such questions. What is clear is that Colombia was able to implement a scholarship program to allow poor children to attend private secondary schools that made both program participants and taxpayers better off than they would have been in the absence of the program. Identifying the impact of the PACES scholarship program "as implemented" is also of considerable interest given the program's scale, the fact that its features are not uncommon among private school scholarship programs.

This paper contributes to several strands of the literature including research on the impacts of private-school voucher programs,[2] the long-term consequences of educational interventions,[3] the economic returns to interventions that target socially disadvantaged children,[4] the effects of education on fertility[5] and recent developments in public finance

[2] For example, Helen Ladd 2002; Angrist, Bettinger and Kremer 2006; Hsieh and Urquiola 2006; Patrick Wolf, Babette Gutmann, Michael Puma, Brian Kisida, Lou Rizzo, Nada Eissa and Matthew Carr 2010; Lisa Barrow and Cecilia Elena Rouse, 2008; Karthik Muralidharan and Venkatesh Sundararaman 2013.

[3] For example, James J. Kemple 2004; Raj Chetty, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Schanzenbach and Danny Yagan 2011; David Deming, Justine Hastings, Thomas Kane and Doug Staiger 2011; Susan Dynarski, Joshua Hyman and Schanzenbach, 2011; Joshua Cowen, David Fleming, John Witte, Wolf and Kisida 2012.

[4] Some argue that interventions that target socially disadvantaged children have highest returns early in the life cycle, during key child-development windows (Flavio Cunha, James J. Heckman, Lance Lochner and Dimitri Masterov 2006; Heckman and Masterov 2007; Heckman 2008). A stronger claim is that interventions in the teen years are doomed to have small impacts. Our findings demonstrate that secondary schooling interventions that target disadvantaged children have the potential to increase earnings and promote social mobility. However, the concentration of effects at the top of the distribution is consistent with the hypothesis of complementarity with initial human capital.

[5] In developing countries, in particular several, studies find a strong causal relationship between educational subsidies and teen fertility (e.g. Lucia Breireova and Esther Duflo 2004; Darwin Cortés, Juan Miguel Gallego and Darío Maldonado 2010; Duflo, Pascaline Dupas and Kremer 2012). In our setting, such effect is the result of people staying in school longer rather than an increase in the opportunity cost of time as a result of additional human capital.

employing reduced form causal estimates of labor-market behavioral responses to policy to measure welfare changes.[6]

The remainder of the paper is organized as follows. Section 2 provides relevant background information. Section 3 we describes the sources of administrative data and the empirical strategy for analyzing long-run outcomes. Sections 4 and 5 discuss long-run educational and labor market outcomes. Section 6 assesses impacts on eligibility for government subsidies. Section 7 discusses fertility. Section 8 presents fiscal impacts. Section 10 concludes.

2. Background

In this section, we provide background on Colombia's educational context, the structure of the PACES scholarship program, and prior evidence on short-term scholarship impacts and mediating channels.

2.1 Colombia's Educational Context

The Colombian education system comprises three levels: elementary school (grades 1-5), secondary (grades 6-11) and tertiary education. Children typically begin elementary school at age 6 or 7 and transition into secondary at around age 12 or 13. Students are legally required to attend school through grade 9 even though the entire secondary school cycle includes two additional grades. Students who complete secondary school on time typically do so by age 18.

Students attending grade 11 of secondary school are compelled by law to take the ICFES exam, and ICFES scores are the primary admission criteria in Colombia's tertiary

[6] See, for example, Nathaniel Hendren 2013 and Sarah Baird, Joan Hamory Hicks, Kremer and Edward Miguel 2013.

education institutions. Since over 95 percent of students take the test (ICFES 2013), we interpret taking the ICFES exam as a proxy for secondary school completion.

Students who wish to continue on to tertiary education can enroll in either vocational colleges or universities. In 2012, the gross tertiary enrollment rate for Colombia was 45 percent (World Development Indicators 2014). Vocational programs typically last two or three years depending on whether they have a technical or technological focus. University programs —more prestigious, expensive and selective— last typically four or five years. Seventy-five percent of tertiary education students in Colombia attend a university and 25 percent attend a vocational college. Public universities are more prestigious and selective. With the exception of a few elite private universities, private tertiary education institutions serve those who do not obtain admission into public universities.

Although we study a large-scale national private secondary school scholarship program, we use data from Bogotá —Colombia's capital city. Colombian municipalities —Bogotá included—are responsible for the administration of public education, funded by transfers from the national government levied through income and value-added taxes (Barrera-Osorio et al. 2011). Nationally, 37 percent of the student population attended a private school when the PACES scholarship program was in place; in Bogotá, however, 58 percent of students did so (Angrist et al. 2002).

In order to facilitate targeted subsidies in public services, Colombia divides its population into six strata based on residential location. The poorest two strata —the target population of the PACES scholarship program— represent roughly 55 percent of Colombia's population (49 percent in Bogotá). Close to 12 percent of students in strata 1 and 2

neighborhoods attend private schools in Bogotá (Encuesta Nacional de Calidad de Vida 2010).[7]

2.2 The PACES Scholarship Program

The PACES scholarship program was introduced 1992 as a way of improving secondary school enrollment rates among disadvantaged students. Most poor students in Colombia attend public schools and —especially in large cities— available slots in public secondary schools were limited when the program began in 1992. The program aimed at tapping the excess capacity in private schools by providing scholarships for private secondary schooling among strata 1 and 2 applicants from public elementary schools (King, Laura Rawlings, Marybell Gutierrez, Carlos Pardo, and Carlos Torres 1997).

Participating private schools served lower-income students and charged lower tuition fees than other private schools that chose not to participate. Teacher-pupil ratios were comparable between all public and participating private schools (King et al. 1997). While initially the scholarship covered most tuition fees, the government did not increase its monetary value to keep pace with inflation, and by 1998 the scholarship only covered about 56 percent of the tuition of the average participating school. Families made up for the difference (Angrist et al. 2002).

In order to receive an award, students needed to have applied and been accepted to a participating private school. Scholarships were awarded by lottery if demand exceeded scholarship availability. Students were between 12 and 13 years of age at the time of application. Renewal of the award through the end of students' secondary schooling was

[7] About 90 percent of public schools in Bogotá are secular and 10 percent have a religious affiliation. Of the 10 percent of religiously affiliated public schools, 73 percent are Catholic. Among private schools, 46 percent are secular and 54 percent have a religious affiliation. Of those private schools with a religious affiliation, 54 percent are Catholic.

contingent upon passing grades. The extent to which this conditionality was at all enforced is unclear (Calderón 1996; Ribero and Tenjo 1997).

The program may have operated through other potential channels aside from the conditionality on satisfactory grade progression. These include allowing more children to attend private secondary schools, allowing those who might have attended private school in any case to attend a wider range of schools, providing income subsidies to families who might have been pre-disposed to attend private school, and changing students' peers.

2.3 Previous Evidence on Short-Term Outcomes and Potential Channels of Impact

Prior research on the PACES scholarship program shows that among scholarship recipients, 90 percent of females and 86 percent of males would have attended private school in the first year after the scholarship lottery and that three years after the lottery, winners were 10 times more likely than losers to be using a private-school scholarship —relative to the losers' usage rate of 5 percent— and were 28 percent more likely to be in private school – relative to the loser's rate of 54 percent (Angrist et al. 2002).[8] Winners were 17 percent more likely to have progressed on time through secondary school (eighth grade at the time of the survey), relative to the loser's on-time progression rate of 63 percent. Treatment-on-the-treated estimates of the effect of ever using a private school scholarship, therefore, suggest an impact of 24 percent on on-time progression. Furthermore, girls who won the scholarship were less likely than losers to be married or cohabiting and worked 1.2 fewer hours per week (Angrist et al. 2002).

Relative to the scholarship lottery losers' base rate of 35 percent, winners were 15 percent more likely to graduate on time from secondary school. Accounting for selection bias

[8] The fact that 54 percent of losers were in private school implies that a lower bound on the value of the scholarships to winners is 54 percent of the scholarship value.

into taking the tertiary education entry test, Angrist et al. (2006) also find that scholarship lottery winners scored 0.2 standard deviations higher than the losers.

Bettinger et al. (2010) investigate scholarship effects in the subpopulation of applicants who applied to a vocational school. Among applicants to vocational schools, scholarship lottery winners were more likely to attend schools with higher dropout rates, fewer qualified teachers and lower fees, as compared to the schools attended by scholarship lottery losers. Despite this, scholarship lottery winners still had better educational outcomes including reduced likelihood of repetition and increased likelihood of secondary school graduation, casting doubt on the idea that vouchers improved outcomes for winners solely by matching them to better peers.

3. Data and Empirical Strategy

In this section we describe the various sources of administrative data we employ for our analyses (section 3.1) and the empirical strategy (section 3.2).

3.1 Data

We used students' names, dates of birth and adult identification numbers from the national registrar to complete the matches with the four administrative datasets. In the scholarship applicant list, all applicants reported their full names (typically two first names, two last names) and 89 percent of them reported a valid youth identification number which contains their date of birth embedded in the first six digits. There is no difference between winners and losers in regards to the probability of having a valid youth identifier (Table 1).

Tracking long-run outcomes in some datasets —particularly social security records— relies on having students' adult identification numbers, which citizens obtain when they turn

18 years old. We obtained valid adult identification numbers for 97 percent of applicants, with no difference in the likelihood of having an adult identification number by win-loss status among all applicants or separately by gender (Table 1).[9]

Colombia has comprehensive individual-level administrative data on secondary and tertiary education, female fertility, and labor market experiences. The breadth and depth of the national data provide a unique opportunity to track PACES applicants across a variety of long-run outcomes with little to no attrition in the data.

Like Angrist et al. (2006), we limit our analysis to the Bogotá 1995 lottery (as this lottery and its records are the most complete and accurate) and refer the reader to the reference for additional details. We use the Bogotá 1995 lottery data and four additional administrative data sources.

To track students' educational outcomes, we use:

1. *The ICFES secondary school graduation/tertiary education entry exam database.* We update and improve the prior match conducted by Angrist et al. (2006) in two ways: i) by matching on students' youth identification numbers, adult identification numbers, and names, (Angrist et al. 2006 did not have students' adult identification numbers), and ii) by matching students to the population of test takers through 2007 —7 years after students would have graduated with no grade repetition —whereas Angrist et al. (2006) were only able to match students through 2001.

2. *The tertiary education database.* We use data from Colombia's Education Ministry's *Sistema de Prevención y Análisis de la Deserción en Instituciones de Educación Superior (SPADIES)* to track scholarship applicants through collegiate pathways, including

[9] Youth and adult identification numbers were linked using administrative data from Colombia's national registrar's office and the Department of National Planning.

enrollment and completion. The tertiary education database is an individual-level panel dataset that tracks close to 95 percent of tertiary education students from their first year to their degree receipt beginning in 1998. We obtained data until the first semester of 2012. The tertiary education database is similar to the National Student Clearinghouse in the U.S. It includes information on the timing and institution of students' tertiary attendance. We also observe characteristics of the institution including whether it is a university or a vocational college, whether the institution was public or private, and its selectivity measured by average percentile scores on the tertiary education entrance exam. Using the data, we can construct outcomes characterizing students' trajectories throughout college (enrollment periods, dropout status and graduation). It also contains information on government financial aid receipt.

To track government subsidies' eligibility, informal sector earnings and family formation outcomes we use:

3. *The SISBEN Census.* We use data from the SISBEN household census of 2010. Data from the SISBEN 2010 survey is used to construct and index score to determine eligibility for government subsidies.[10] SISBEN 2010 covers 57 percent of households in all of Colombia and 39 percent of households in Bogotá.

To estimate scholarship impacts on earnings we complement SISBEN data with:

[10] We observe actual receipt of tertiary education loan subsidies in the tertiary education database. For the other subsidies we only observe eligibility as determined by whether households have SISBEN scores below predetermined cutoff points that vary by subsidy. These subsidies include: early childhood care (*primera infancia)*, health care (*régimen subsidiado en salud*), tertiary education loan subsidies (*crédito access*), conditional cash transfers (*familias en acción*) and elderly care (*protección social al adulto mayor*) subsidies. For the healthcare subsidy the only eligibility criterion is SISBEN scores. Eligibility for the remaining subsidies requires additional demographic conditions such as having age-appropriate children (early childhood care and conditional cash transfers), being admitted or attending tertiary education (tertiary education loan subsidies) or living with an elderly relative (elderly care).

4. Colombia's Social Protection Ministry's *Sistema Integral de Información de la Protección Social* (*SISPRO).*[11] SISPRO is an individual-level panel dataset that is updated monthly, and that contains information on contributions to government social programs for health, employment, and retirement. For the purposes of this study we focus on the work module, which contains information on whether individuals have worked in the formal sector, the number of days of formal sector employment, monthly earnings, and social security contributions. We focus on outcomes from 2008 to 2012 —between seven and 11 years after on-time secondary school completion of scholarship applicants in the Bogotá 1995 sample— since SISPRO only began to cover the universe of formal sector workers in 2008. On average, scholarship applicants would have been around 30 years old at the end of this period. We use SISPRO data to examine additional outcomes, including: extensive and intensive margins of formal sector employment, formal sector earnings and payroll taxes.

On the whole, we are able to match close to 95 percent of applicants to any of the four administrative datasets we use to track long-run outcomes.

3.2 Empirical Strategy

Our main empirical strategy is based on an intent-to-treat (ITT) analysis that compares outcomes between scholarship lottery winners and losers, as follows:

$$Y_i = \alpha + \gamma Z_i + \beta X_i + \varepsilon_i$$

where Y_i is an outcome variable for scholarship applicant i, Z_i is an indicator variable for whether applicant i was awarded a private school scholarship through the lottery, X_i is a vector of baseline controls from the scholarship application form that includes age, gender

[11] The SISPRO database only includes people who worked for employers that register their workers or self-employed workers who register themselves. In Colombia and in Bogotá, respectively 50 percent and 55 percent of employment is formally registered (Secretaría de Desarrollo Económico de Bogotá 2012).

and whether the applicant had a phone number at the time of application, and ε_i is an error term. From the scholarship application form we also observe to which school the students applied, and we can infer whether that school has an academic or vocational curriculum. We use this distinction to examine impacts in the subpopulation of applicants to vocational schools.

4. Scholarship Impacts on Long-Run Educational Outcomes

4.1 Secondary Education Completion Outcomes

In this subsection, we examine secondary school outcomes and in the next we examine tertiary education outcomes. We concentrate on on-time secondary school completion and the likelihood of ever completing secondary school. We define on-schedule secondary school completion as having taken the tertiary education entry test no later than six years after applying for the scholarship, that is to say by 2001. Scholarship lottery winners are 17 percent (7.3 percentage points) more likely to complete secondary school on time relative to the loser's on-schedule completion rate of 43.8 percent (Panel A of Table 2).[12] Female lottery losers are more likely than their male counterparts to complete secondary school on schedule; point estimates of scholarship effects are larger (both in percent and percentage point terms) among male applicants, although we cannot reject the hypothesis that the two are the same.

During the six years following on-schedule completion, the difference between the proportion of scholarship lottery winners and losers who have completed secondary school declines with each year. Scholarship lottery winners are 10 percent (5.4 percentage points)

[12] With our updated matching strategy including adult identification numbers, we obtain substantially higher match rates than Angrist et al. (2006). Impact estimates in percentage points are similar.

more likely to complete secondary school within six years after on-schedule completion relative to a base rate of 55 percent.[13]

4.2 Tertiary Education Outcomes

After secondary school, the scholarship effects persist. Scholarship lottery winners are 16 percent (2.9 percentage points) more likely to have ever enrolled in tertiary education relative to the 18.4 percent base rate among lottery losers. Point estimates are slightly larger for females, but the estimated effects are not statistically different by gender (Panel B of Table 2).

Among females, scholarship effects on tertiary enrollment concentrate in 5-year universities. Among males, scholarship effects concentrate in the less academically demanding 2-year vocational programs.

Scholarship lottery winners are 58 percent (2.1 percentage points) more likely to be currently enrolled in tertiary education 11 years after on-schedule secondary school completion, that is to say by 2012, relative to the losers' base rate of 3.6 percent. The effect on being currently enrolled is larger for females, among whom winners are 68 percent (2.5 percentage points) more likely to be currently enrolled in tertiary education relative to the losers' mean of 3.7 percent. Among males, the effect on being currently enrolled is 50 percent (1.7 percentage points) relative to a base of 3.4 percent and is only significant at the 10 percent level.

[13] Three years after the lottery, scholarship lottery winners were less likely than lottery losers to repeat grades in secondary (Angrist et al. 2002). To analyze secondary school graduation, Angrist et al (2006) completed the match focusing on students' probable date of on-time graduation. With the benefit of more data, we observe that many lottery losers eventually complete secondary school, but it is taking them up to six years longer to do so.

Among females, the effect on being currently enrolled is concentrated in private institutions. Among males, current enrollment effects are concentrated in public tertiary education institutions. We find no overall effect on tertiary graduation.

Among females, scholarship lottery winners accumulate 0.12 additional years of tertiary education, an increase of 28 percent relative to a base of 0.44 years among losers. Among males, scholarship lottery winners accumulate 0.035 additional years of education, 9 percent relative to the base rate of 0.37 years among losers. The scholarship impact on additional years of tertiary education is not significant among males.

Close to one percent of scholarship applicants ever received government financial aid. There is no difference in the probability of receiving government aid between winners and losers in the full sample or separately by gender.

5. Scholarship Impacts on Labor Market Outcomes

We examine four labor market outcomes: match rates to earnings data (subsection 5.1); formal sector employment intensity (subsection 5.2); earnings and payroll taxes (subsection 5.3).

5.1 Match Rates to Earnings Data

We use two earnings data sources: SISBEN, which covers low-SES neighborhoods and includes about 52 percent of the scholarship applicant population fifteen years after initial scholarship award, and SISPRO data. SISBEN 2010 earnings are a cross-section of self-reported earnings for 2010 and include both formal and informal self-reported earnings.[14]

[14] Sixty-four percent of scholarship applicant respondents in the SISBEN 2010 are informal workers as measured by working and earning below minimum wage and not having private health insurance coverage.

SISPRO data contains monthly earnings for all formal sector workers (i.e. those who pay payroll taxes) between 2008 and 2012.

One limitation of this analysis is the fact that lottery winners are about two percent more likely to be still enrolled in tertiary education during the periods we analyze, and this may limit their current earnings while increasing their future earnings. Thus current differences in formal sector earnings between winners and losers may understate future earnings differences. [15]

Among scholarship lottery losers, we match 52 percent to SISBEN data. Lottery winners are 5 percent (2.6 percentage points) less likely to ever appear in SISBEN data, indicating that they are less likely to reside in poor neighborhoods fifteen years after initial scholarship receipt. This difference, however, is not statistically significant (Column 1, Panel A of Table 3). Match rate correlates are not statistically different between scholarship lottery winners and losers (Column 2, Panel A of Table 3).

The implication for bounding earnings effects of the scholarship is that adding in the approximately 5 percent of winners who would have been in SISBEN neighborhoods would likely increase reported earnings in the SISBEN.

Among scholarship lottery losers, we match 69 percent to the Colombian government record of those paying payroll taxes, SISPRO, implying that 69 percent of losers ever show up in formal sector employment between 2008 and 2012. Lottery winners are 3 percent (2.2 percentage points) more likely to ever appear in formal employment records during this

[15] In the Encuesta de Calidad de Vida 2010, a nationally representative household survey, over 80 percent of workers who completed tertiary education worked in the formal sector. When we define an outcome variable as "in formal sector employment or currently enrolled in tertiary education" we find that scholarship winners are 1.9 percentage points (not statistically significant) more likely than losers to be in formal employment or currently enrolled during the period 2008-2012.

period. However, this difference is not statistically significant (Column 3, Panel A of Table 3). Match rate correlates do not systematically differ between winners and losers in the full applicant sample or separately by gender (Column 4, Table 3).

5.2 Formal Sector Employment and Intensity

Point estimates suggest that winners spend more time in formal sector employment. The differences are not statistically significant at conventional levels, but some estimates for males are significant at the 10% level. For the period between 2008 and 2012, we construct an outcome variable that equals one if the applicant spends X months or more in formal sector employment during the period and zero otherwise. We allow X to vary between 5 and 40 months (Table 4).[16] For each cutoff number of months per year in formal sector employment we estimate one regression equation in the full sample and separately by gender with and without controls.

Sixty percent of applicants work at least five months in formal employment. This proportion declines uniformly with the intensity of formal employment such that 23 percent of losers spend at least 40 months in formal sector employment between 2008 and 2012 (Table 4).

Male scholarship lottery winners are 7 percent (4.3 percentage points) more likely than losers to be formally employed for at least 5 months during 2008-2012 and 16 percent (3.6 percentage points) more likely than male losers to be in formal sector employment 40 months or more during 2008-2012. Both of these effects are statistically significant at the 10 percent level. Scholarship effect estimates near the middle of the formal employment intensity

[16] Note that our formal employment data spans from July 2008 through July 2012, or 48 months. Therefore, we view spending 40 or more months in formal sector employment as being formally employed throughout the period covered by our data.

distribution are smaller in magnitude and not statistically significant, although we cannot reject that these effects are different from at the top and at the bottom.

5.3 Earnings and Payroll Taxes

We focus on formal sector earnings, since we have administrative data on this and data on total earnings are only available for a subsample selected on the endogenous variable of location.[17]

We compute annual formal sector earnings by adding inflation-adjusted monthly formal sector earnings during the period covered by our formal employment data (July 2008 to July 2012) including zeroes for months without reported formal sector earnings and dividing by four to get an annual average (Table 5). Since 30 percent of applicants never appear on formal employment records during this period, total formal earnings for them are zero.

Current annual formal earnings for scholarship lottery losers are, on average, $2,201 (including zeros).[18] Scholarship lottery winners earn an additional $173 in formal annual earnings, an 8 percent increase (Panel A of Table 5). The p-value on this difference is 0.07.

Relative to losers, scholarship lottery winners pay higher annual payroll taxes at around age 30. The OLS scholarship impact on payroll taxes is $44.6, which at the losers'

[17] Recall that scholarship winners are 5 percent less likely to appear in the in the SISBEN 2010 survey of those living in poor neighborhoods. This makes it difficult to make strong statements about differences in annual earnings using SISBEN data. Therefore, OLS results on observed SISBEN 2010 earnings (Panel A of Table 5) provide an estimate that is a combination of the causal effect of winning a scholarship on informal earnings plus a selection bias term that is likely negative because winners who are no longer in the SISBEN dataset due to winning the scholarship presumably had higher earnings than those who stayed in SISBEN neighborhoods. In the full sample, the lower bound point estimate on annual earnings is -$164.4. This lower bound has a t statistic of less than 1. Trimming the top 5% of earners among losers that appear in SISBEN provides an upper bound. In the trimmed sample, we obtain an upper bound of the effect of the scholarship on monthly reported informal earnings of $308.4—a 16 percent increase relative to the losers' mean of $1,992.

[18] We report results based on formal earnings reported in health payroll accounts. Results are very similar if we use instead earnings from the pension payroll account.

mean of $597.7 represents an increase of 7.5 percent. The scholarship impact on payroll taxes is statistically significant at the 10 percent level in the full sample of applicants.

Quantile regression results provide some evidence that scholarship impacts are stronger at the top of the formal earnings distribution (Panel B, Table 5). The point estimate of the scholarship effect at the 70th percentile of current formal sector earnings is $92.4, which is not statistically significant. Meanwhile, we find that at the 90th and 95th percentiles, the estimate of the effect of the scholarship on current formal earnings is $368 and $658, 7 percent and 10 percent greater, respectively, than loser earnings at these quantiles (statistically significant at the 10 percent level).

While standard errors on the earnings estimates are too large to make strong inferences, the magnitudes of the point estimates of the earnings increases in Table 5 are too large to be explained fully by additional years of education, suggesting that education quality may also have played a role. Assuming that all of the scholarship effect on earnings came solely through increased educational attainment, instrumental variable estimates (not shown) would imply a return to an additional year of attainment that is 27 percent.[19] Mincer-equation estimates from Colombia suggest that the return to an additional year of educational attainment is about 8 percent for secondary and 18 percent for tertiary education (World Bank 2003). The standard errors on the IV estimate, however, are large enough to include this Mincerian return, indicating that we cannot rule out that additional years of education among scholarship winners entirely drive the formal earnings scholarship effect.

[19] We back out the distribution of years of schooling among scholarship applicants from results in Table 2. In the full sample, scholarship lottery losers accumulate, on average, 10.4 years of schooling. Scholarship lottery winners accumulate 0.2 additional years of schooling. Instrumental variables estimates with application controls in which the dependent variable is monthly formal sector earnings the endogenous regression is years of education and the excluded instrument is winning a scholarship produce an estimate of the return to one additional year of educational attainment of 27 percent.

To summarize, we find scholarship lottery winners have current formal sector earnings that are at least 8 percent greater than those of losers, with the difference significant at the 10% level. The long-run difference in formal earnings, however, is probably greater than any current difference given that those currently in tertiary education will likely earn more later. For instance, scholarship applicants with completed tertiary education have formal sector earnings well above the 90th percentile of the distribution, while those who are still enrolled either have no formal sector experience or have formal earnings at much lower quantiles of the distribution.

To get a sense of how big the effect might be, we assume that in two years current tertiary education enrollees (as of 2012) would earn what 2010 tertiary graduates in our dataset earn, on average, in 2012.[20] The scholarship impact on predicted annual estimated 2014 formal sector earnings is $213.3, and statistically significant at the 5% level in the full sample (Panel A, Table 5). Relative to the losers' mean, this impact represents a 9.3 percent increase. The scholarship impact on projected annual earnings for females is $957 (11 percent increase relative to the losers' mean) and statistically significant at the 10% level.

6. Scholarship Impacts on Government Subsidy Receipt

We estimate scholarship impacts on the probability of receiving *Familias en Acción* conditional cash transfers and on the probability of being eligible to receive benefits from the other three largest government subsidy programs available for urban households: the two tiers

[20] We do not know, however, how many current tertiary enrollees will graduate and in this calculation we assume all do. We expect to relax this assumption in subsequent versions.

of subsidized health care and early childhood care (Table 6).[21] Eligibility for subsidized health care is equivalent to utilization because take up is almost one hundred percent.

Eligibility for these programs is based on SISBEN scores and demographic characteristics such as having age-appropriate children.[22] Applicants without SISBEN scores are not eligible to receive these subsidies so for applicants who do not show up in SISBEN 2010 subsidy receipt is zero.

In general, we find little evidence suggesting that, in the full sample of applicants or among female applicants, winning the scholarship affects government welfare receipt of *Familias en Acción* and subsidized health care programs or eligibility for early childhood care (Table 6). Among male applicants, however, winning a scholarship reduces the likelihood of receiving benefits from Colombia's *Familias en Acción* conditional cash transfer program by 27 percent (1.7 percentage points relative to the losers' mean of 6.2 percent). This point estimate is statistically significant at the 10 percent level (Table 6).

7. Scholarship Impacts on Teen Fertility

Fertility outcomes are only observed for applicants who show up in the SISBEN 2010 data. As noted earlier, scholarship lottery winners have a lower likelihood of appearing in SISBEN 2010 data. We cannot, however, reject the null hypothesis that application correlates that predict showing up in SISBEN 2010 are the same for winners and losers in the full sample and separately by gender (Table 3). Teenage parenthood outcomes conditional on SISBEN appearance are a lower bound as long as lottery winners who do not live in low-SES

[21] In terms of budget and beneficiary population, subsidized health care is the largest government benefit program, followed by *Familias en Acción* and early childhood care.

[22] We define *Familias en Acción* receipt as whether applicants' SISBEN score is at or below the eligibility cutoff and whether they have children between 0 and 17 years of age. Take up of subsidized health care is nearly one hundred percent among eligible families so for subsidized health care eligibility and receipt is almost identical.

neighborhoods covered by SISBEN as a consequence of winning the scholarship have a lower chance of being teenage parents than scholarship lottery winners who remained in SISBEN.[23]

Conditional on showing up in the SISBEN 2010 dataset, winners are 17 percent (4 percentage points) less likely to have a child during their teenage years relative to the lottery losers' mean of 23.4 percent. Among females, winning a scholarship reduces teen motherhood by 16 percent (6.1 percentage points) relative to a base of 37.5 percent. Teen fatherhood is rare in Colombia, because women are usually substantially younger than their partners so rather than examine teen fatherhood, we examine whether males have children with teen partners. Male lottery winners are 31 percent (5 percentage points) less likely to have a spouse or partner who had a child as a teenager relative to a base rate of 16 percent (Table 7).

The reduction in teen fertility could be the result of an "incarceration" effect by which winners stay in school longer and do not want to risk losing the scholarship, or it could be an opportunity cost effect by which additional human capital increases wages, making time valuable. If "incarceration" is the driving mechanism, we might not observe scholarship impacts on total fertility because there may be catch up fertility once schooling is completed. In contrast, if the driving mechanism is opportunity cost, we should observe an effect on total fertility.

[23] Evidence from Angrist et al. (2002) supports the notion that the teenage childbearing effects among female applicants in are likely a lower bound. Among applicants to the Bogotá 1995 lottery cohort, lottery winners were close to 30 percent less likely to have had a child three years after the lottery, although the effect was not statistically significant at conventional levels. Selection, however, was also a concern in their sample. Surveys were conditional on applicants having a phone and the Angrist et al. (2002) sample had a two-percentage point difference in survey response rates between scholarship lottery winners and losers—55 percent response rate for lottery winners and 53 percent for losers—which is similar to what we find in the SISBEN sample (Table 3). The fact that results in samples with different potential sources of selection are consistent yields additional credence to our fertility findings.

We find no evidence that winning a scholarship increased total fertility. At the time of SISBEN 2010, in which applicants are about twenty-eight years old, the average scholarship lottery loser has one child. Estimates on winning a scholarship are small and statistically insignificant although fairly precisely estimated. We also do not find any evidence of scholarship effects on total fertility when we examine the probability of have at least one, two or three children. This pattern of results is consistent with the "incarceration" hypothesis by which winning a scholarship keeps a student in school longer without necessarily affecting her opportunity cost of time. One caveat to this finding is that impacts on total fertility may show up later in the potential childbearing years, so fertility gaps may appear later.

8. Welfare Impacts[24]

In this section we quantify the program's welfare impacts. Welfare impacts are the sum of: i) impacts on taxpayers (subsection 8.1),[25] ii) impacts on scholarship recipients (subsection 8.2), and iii) externality impacts on others. We are not able to identify externality impacts on others; in subsection 8.3, however, we estimate how large net negative externalities would need to be to imply that the program is not welfare improving to society.[26]

[24] This section is still preliminary and undergoing refinements.
[25] For example, see Hendren 2013; Baird et al. 2013.
[26] The discount rate we use is 3.6 percent, which is average interest rate on new external government debt commitments for Colombia between 2002 and 2012 (World Development Indicators database). In all calculations that follow we estimate amounts per scholarship winner, separately for males and females, which assumes that the counterfactual situation is no scholarship program. Throughout the analysis, for each source of cost and revenue, we compute the NPVs converting into United States dollars (if not already) using the year-specific exchange rate (Dec. 31 of that year) between US dollars and Colombian pesos from the Colombian Central Bank, deflating nominal costs back to real value in base year (1995) using the US-CPI change between base year and incurrence of costs (or revenue), taking the present value of the cost and revenue stream. We express the NPV in US dollars for the year of analysis (2013) using US-CPI change between the analysis and the base year.

8.1 Costs and benefits to taxpayers

We calculate five cost sources for the government: scholarship costs net of savings from reduced expenditure on public education, cost-savings from reduced secondary school grade repetition, increased tertiary education costs, benefit costs, and foregone tax revenues due to reduced work time among scholarship winners to the extent that they spend more time in school.[27] [28]

8.1.a. Costs associated with secondary school of attendance

There are two competing impacts of the scholarship on public expenditures. First, for students who would have attended private school in the absence of the program, the scholarship increases public expenditure. For example, a substantial proportion (87.7 percent) of lottery losers attended private school in sixth grade. While the government did not have to pay for the fees of the lottery losers, it did have pay for the scholarship value for students who won the lottery. Among these students, the average expenditure increase was $214 ($214 = scholarship value * percentage of lottery losers attending private school = $244 * .877). Note that the proportion of applicants who attended private school among lottery losers quickly deteriorated (53.9 percent by 8th grade).

[27] Based on US evidence, an additional and sizeable source of cost savings to the government is reduced teen fertility (Saul Hoffman 2006). Estimates for Colombia on the monetary costs of teen fertility are only available for society as a whole and already include foregone earnings as a cost (Arturo José Parada-Baños 2005). It is not clear what fraction of these costs accrue to the government. We do not, therefore, account for cost-savings from reduced teen fertility and as such, these fiscal impact estimates err on the side of being conservative.

[28] After the passing of Law 100 of 1993, the pension system created two regimes: average premium (*Regimen de Prima Media)* and individual savings with solidarity (*Regimen de Ahorro Individual con Solidaridad).* In the average premium regime, employee and employer-side contributions go to a common pool of resources and pension benefits are obtained as a function of age, formal sector earnings and time in formal sector employment. There is no minimum pension amount in this regime. The individual savings regime is akin to individual retirement accounts in the US in which accounts belong to the individual and pension benefits do not depend on age or other parameters; they only depend on the principal and interest earned. In neither case, therefore, are there government subsidies to retirees.

We make two other adjustments to these estimates. Among lottery winners, not all private school attendees continued to use the scholarship. Even in the first year, 6 percent of private school attendees who had been offered the scholarship had discontinued using it. By 8th grade, 33 percent of lottery winners who were attending private school were not using the scholarship.[29] The second adjustment involves our assumptions about scholarship usage between 8th and 11th grade. After 8th grade, we have no data on scholarship usage. We know that 32 percent of the overall lottery loser sample finished 11th grade in private school. We assume a constant deterioration from 8th grade to 11th grade in the fraction of losers attending private school. This implies a 40 percent reduction in the fraction attending private school, and we assume that deterioration in scholarship usage among winners follows a similar 40 percent decline from the 8th grade level. As before, we assume constant deterioration over time.

When we combine these figures, we estimate that public expenditure increased by $494 for females and by $451 for females as a result of the awarding of private school scholarships to students who intended to attend private schools regardless of the scholarship program (Row 3 of Table 8).

The scholarship, however, induced some public school attendees to attend private school. The scholarship's value ($244) was considerably lower than the annual cost of public school ($449). We assume that the marginal cost of public education equals the average cost (since this was a period of expanding school enrollment), and thus, for each student who

[29] These students may have repeated grades, transferred schools, or voluntary given up the scholarship.

moved from public to private school, the government saved $205 per year.[30] To figure out the net impact on overall costs, we multiply this cost savings by the proportion of students who attended private schools as a result of the scholarship.[31] We compute that the scholarships reduced public expenditure by $178 for females and by $172 for males as a result of the shift of students from public to private schools (Row 4).

8.1.b. Other government costs

We compute four other costs to the government. These include reduced expenditure as a result of fewer grade repetitions, increased tertiary education costs as a result of increased attendance, changes in welfare program expenditures, and foregone tax revenue for students who remain longer in school.

We defer a detailed explanation of the cost savings from reduced grade repetition among scholarship winners to Appendix B. In short, we assume that i) only public school repetitions cost the government money, ii) among applicants who finish on-schedule, there is no repetition; iii) among applicants who finish with delays, the delays are all a consequence of grade repetition and iv) among those who never finish secondary school they dropped out in 9th grade, so that total repetitions for this group are as reported Angrist et al. (2002). Under these assumptions, the cost savings from reduced grade repetition are $1.75 for females and $6.89 for males (Row 5).[32] Total secondary education costs to the government are six-year

[30] Angrist et al. (2002, p. 1537) reports the annual cost of public school to be $350 and the average scholarship value to be $190, both in 1998 dollars. We calculate that in 2013 prices, these figures correspond to $449 and 244, respectively.

[31] We obtain these impacts on private school attendance for grades 6th through 8th from Table 4, column 2 (for males) and column 4 (for females). After 8th grade, we assume a constant change from the observed 8th grade effect to the eventual effect at graduation.

[32] We derive cost-savings estimates from data on public school costs, which we obtain for 1998 from Angrist et al. (2002, see footnote 29). We discount these to base year by using the US-CPI change between base year and incurrence of costs year (which we assume is 1998 in this case as we do not observe public school costs for any

costs for students who would have attended private schools in the absence of the program (Row 3) net of savings from reduced expenditure on public education for students induced to transfer from public to private schools (Row 4), and cost savings from reduced grade repetition (Row 5), which are $314.85 for females and $271.68 for males (Row 6).

There are two sources of tertiary education costs to the government: additional public tertiary education costs and tertiary education loan subsidies.[33] Additional public tertiary education costs are estimated as annual per-pupil expenditures in public tertiary education times the scholarship impact on years of tertiary education times the fraction of lottery winners attending a public tertiary institution (Row 7). Additional tertiary education loan subsidies are estimated as annual per-pupil tertiary education loan subsidy amounts times the scholarship impact on number of years of subsidy receipt (Row 8). Tertiary education costs (public education plus loan subsidies) are $17.68 for females and $15.29 for males (Row 9).

Costs of benefits receipt are estimated as annual *Familias en Acción* CCT costs times the scholarship impact on CCT receipt.[34] The changes in benefit costs caused by winning the lottery are $2.00 for females and -$3.39 for males (Row 10).

To the extent that winners spent more time in school, the government may have foregone certain tax revenue while they were in school. Based on Table 5 we assume that

---

other year). We then express these costs in US dollars for the year of analysis (2013) using US-CPI change between the analysis and the base year.

[33] Annual tertiary education subsidy data come from ICETEX (2014).

[34] To obtain annual CCT costs we assume that the average scholarship applicant has one child as indicated by the mean number of children at age 28 from Table 7. We assume applicant resides in Bogotá so monthly health subsidy amount is that for Group 1 municipalities (see: http://www.dps.gov.co/Ingreso_Social/FamiliasenAccion.aspx, retrieved October 28, 2014). We assume child is between zero and seven years of age in 2013, so can receive health transfer in the amount of COP 61,200/month. but no education transfer. (Education transfers until the end of primary school are zero in group 1 municipalities and it is unlikely that scholarship applicants have children in secondary school, so the age assumption is conservative from a cost perspective). We obtain annual CCT transfer amount by multiplying the monthly transfer amount by 12. We follow order of operations above to obtain NPV in USD of analysis year. We only assume one year of costs since there is no difference by scholarship status in total fertility, indicating simply a difference in the probability of having age-appropriate children.

there is no difference in informal sector earnings between scholarship winners and losers. Foregone VAT tax revenue, therefore, equals formal annual formal sector earnings of scholarship losers times the scholarship impact on years of education times the average VAT rate of 13.3 percent,[35] which equals $18.09 for females and $22.60 for males (Row 11).[36]

Foregone payroll taxes are annual payroll taxes for scholarship losers times the scholarship impact on additional years of education. Not all foregone payroll taxes, however, represent a net government transfer because a large fraction of these goes back to the worker, for instance, through the pension benefits formula. We estimate that at the margin forty percent of payroll taxes represent a net transfer to the government.[37] Foregone net government transfers from payroll taxes are $18.09 for females and $22.60 for males (Row 12). Total foregone revenue is the sum of foregone VAT taxes and the net transfer from foregone payroll taxes, which totals $41.03 for females and $49.45 for males (Row 13).

[35] Jaramillo and Tovar (2008) Table 3 reports average VAT rates for five consumption groups: Transportation and Communications (15.19%), Food (8.01%), Culture and Entertainment (13.97%), Housing (14.66%) and Other (15.94%) We use data from Colombia's *Encuesta de Ingresos y Gastos* from 2006/2007 to estimate the distribution of consumption across these groups in the two lowest deciles of the consumption distribution, which are 7.4% for Transportation and Communication, 25.6% for Food, 1.5% for Culture and Entertainment, 39.5% for Housing and 25.9% for Other. The average VAT tax rate of 13.3% is a weighted average of the VAT rates across the different consumption groups, with the weights given by the share of consumption among the two lowest deciles in each category.

[36] The current difference in formal sector earnings and payroll taxes between scholarship winners and losers already accounts for foregone earnings due to any additional time in school between 2008 and 2012, which is the period that our formal sector earnings data covers. Evidence from Table 2 indicates that scholarship winners, however, already spent additional time in school prior to 2008, particularly finishing secondary school. Since we do not observe earnings that far back, the assumption that foregone earnings then are similar to those now is fairly conservative.

[37] Ten percent of payroll taxes are earmarked to finance Colombia's national job training agency (SENA) and the national institute for family welfare (ICBF) and therefore represent a net transfer to the government. Thirty percent of total payroll taxes are for health care services and also constitute a net government transfer because the mandatory health plan, known as POS, provides services that do not depend on the amount paid in the system so additional health payroll taxes among winners relax the government budget constraint. We conservatively assume that the pension scheme involves no redistribution.

Total direct scholarship costs are the sum of secondary education costs (6), additional tertiary costs (9), welfare costs (10) and foregone revenue (Row 13). Total expected direct scholarship costs are $375.56 for females and $333.03 for males.

8.1.c Government revenue

There are two sources of future additional government revenue: additional revenue from VAT taxes and additional government revenue from payroll taxes. In this calculation we also use the forty percent figure of additional payroll taxes that represent a net transfer to the government. We use projected annual formal sector earnings (from Panel A, Table 5). We project earnings for losers and winners over a 35-year work horizon allowing for a 3.02 annual growth rate.[38] The NPV of additional earnings is the difference between the NPV of earnings for winners and the NPV of earnings for losers. Multiplying the difference by the 13.3 percent VAT tax rate we get additional VAT tax revenue, which is US$887.18 for females and US$681.51 for males (Row 16).

Annual payroll taxes come from Panel A of Table 5. We project annual payroll taxes for losers and winners over a 35-year work horizon allowing for a 3.02 annual growth rate. The NPV of additional payroll taxes is the difference between the NPV of payroll taxes for winners and the NPV of payroll taxes for losers. The NPV of additional government revenue is this difference multiplied by 40 percent, which is $410.15 for females and $580.68 for males (Row 17). Total expected additional government revenue is $1,297 per female scholarship winner and $1,262 per male scholarship winner (Row 18).

[38] The rate of 3.02 percent is the average annual growth in GDP per capita in Colombia between 2002 and 2012 (World Development Indicators database).

Expected net fiscal costs to taxpayers are negative: -$922 for females and -$929 for males.[39]

The conclusion that the expected cost to taxpayers is negative seems reasonably robust to changes in the assumptions. Expected net fiscal costs to taxpayers are also negative using a higher discount rate of 6 percent: -$265 for females and -$286 for males (Row 15 minus Row 19). Note that net fiscal costs are negative even if one assumes that any increase in formal sector earnings is offset by reduced informal earnings, so there are no gains in VAT revenue, expected net costs for taxpayers are negative due to increased payroll tax receipt (-$61 for females; -$269 for males, Row 15 minus Row 20).[40]

8.2 Benefits to scholarship recipients

Recipients benefited as public school costs exceeded scholarship costs and over 85 percent of recipients would have attended private school anyway (89.7 percent of females and 85.7 percent of males). This implies that gains to infra-marginal recipients were about $262 per female scholarship winner and $235 per male scholarship winner.[41] Hence, even if there were no increase in tax revenue to the government, the program would transfer to

[39] This conclusion is based on the point estimates only. We expect to account for estimation uncertainty in future versions by computing bootstrapped standard errors on the fiscal impact point estimate.

[40] Net fiscal costs are also negative for both females and males if we assume a discount rate of 6 percent instead of the 3.66 percent rate assumed in the text. Net fiscal costs using the 6 percent discount rate are -$235.38 for females and -$335.35 for males.

[41] The gain for infra-marginal recipients = (fraction of infra-marginal recipients)*(impact on scholarship amount)*(sum of year-by-year utilization rate). The fraction of lottery losers who attend private school in 6th grade is 0.897 among females and 0.857 among males. The impact on scholarship amount is $93.2 (from Angrist et al. 2002 Table 8, column 3 updated to 2013 dollars). We observe the fraction of winners in private school using the scholarship for grades 6th and 8th only. The 7th grade fraction is the linear combination of the 6th and 8th grade rates. For females the fraction of winners in private school using the scholarship is 0.953 (6th), 0.736 (7th) and 0.519 (8th). For males it is 0.933 (6th), 0.698 (7th) and 0.463 (8th). After 8th grade, we have no data on scholarship usage. We know that 32 percent of the overall lottery loser sample finished 11th grade in private school. We assume a constant deterioration from 8th grade to 11th grade in the fraction of losers attending private school. This implies a 40 percent reduction in the fraction attending private school, and we assume that deterioration in scholarship usage among winners follows a similar 40 percent decline from the 8th grade level. Under these assumptions, scholarship usage rates for females are 0.415 (9th), 0.310 (10th) and 0.206 (11th). For males the scholarship usage rates are 0.374 (9th), 0.284 (10th) and 0.195 (11th).

beneficiaries about 70 percent of what it cost taxpayers, even if one counted only benefits to infra-marginal recipients, thus implicitly treating any benefits of educational and economic gains to beneficiaries as partially offset by effort costs and financial costs to beneficiaries.

8.3 Benefits to society

While we can measure impacts on taxpayers and scholarship recipients, we are not able to identify potential externality impacts on others, and to the extent that such effects exist, they should be part of any welfare calculation. The calculation above implies that as long as any externalities are either positive, or negative and less than about $1,100 per scholarship recipient (Row 22 of Table 8), the program is welfare improving. We think there are likely to be some positive externalities from the program. The reduced teen fertility may be good for the children and the families of winners. The opening up of an avenue for social mobility may also have positive civic effects. However, one could imagine areas of potential negative externalities, for example, if there is labor market signaling or job rationing. The net negative externalities would have to be fairly large relative to the wage gains to winners to change the conclusions.

9. Conclusion

We present evidence on the long-run educational, labor market, welfare, family and fiscal impacts of Colombia's PACES scholarship program, one of the largest private school scholarship programs in the world. As such, it is the first paper that explores the impact of private school scholarships on labor market outcomes. We are able to take advantage of a setting in which: a) there is exogenous variation in private school access due to random assignment of PACES scholarships when demand exceeded availability, b) administrative

data provides credible evidence of impacts, c) we are examining an "as is" implementation of a large-scale government program and d) program rules enable us to examine program effects in a subpopulation of scholarship applicants to vocational schools, amongst whom effects are unlikely to be primarily the result of student re-sorting,

Winning a scholarship for private secondary schooling increases on-time secondary school completion by 17 percent and ever completion by 10 percent (Base rates were 44 percent and 55 percent respectively.) It increases tertiary education access by 16 percent and current enrollment or graduation by 58 percent. (Base rates are 18 and 4 percent respectively.) As a result scholarship lottery winners accumulate 0.12 additional years of tertiary education.

Winning a scholarship increases the estimated amount that applicants pay in future payroll taxes by 8 percent and earnings by 8.5 percent. Higher earnings likely reflect more than just a quantity of schooling effect, but estimates are not precise enough to make definitive statements.

Winning a scholarship reduces teen fertility by 16 percent. (Base rate is 38 percent.) The likely mechanism behind the reduction in teen fertility is an "incarceration" as opposed to an opportunity-cost effect as we find no scholarship impacts on total fertility by age twenty-eight.

There is some evidence of gender differences in program effects. Female scholarship lottery winners are more likely to be currently enrolled in five-year university programs and have accumulated 0.12 additional years of tertiary education. Among males, scholarship lottery winners are more likely to be currently enrolled in two-year vocational colleges, and there is no significant difference between winners and losers in years of tertiary education.

With a single experiment, it's impossible to fully disentangle the channels of program impact. However, gains at the tertiary education level and the labor market suggest that the impact of the program on secondary completion was not simply due to schools gaming of the system by lowering the standards for grade progression.

The scholarship program combined elements of a private school scholarship program with elements of a merit scholarship program insofar as renewal of the scholarship was conditional on grade progression. However, as noted, it's not clear how strongly the later requirement was enforced in practice. If the effect of the program were solely due to its merit scholarship component then one would expect the strongest impacts to occur among those who are near the boundary of failing grades. In fact, it seems that many of the strongest impacts are at the top of the distribution. Effects in tertiary education and formal sector earnings both are relevant at the top of the distribution. We do not observe any effects on the fraction of applicants who are eligible to receive government subsidies. The main place we see an effect that might be at the bottom of the distribution is on teen fertility.

Our results paint a more favorable picture of private school scholarships relative to many of the results coming out of the US school choice literature. One possible explanation is program design, as PACES allowed households to augment scholarship amounts, potentially helping some students who would have attended private schools to trade up to better private schools. Another explanation is contextual differences. Since the program only partially covered costs, it attracted students whose families were willing to pay something for private school and the treatment effect of moving from public to private education in the subpopulation of those willing to pay for private education may be greater than in the population at large.

Our fiscal calculations suggest that the net fiscal cost of the program is negative due to the indirect effect of scholarship receipt on government expenditure and revenue. Some features of the program design minimized the fiscal cost of the program. First, scholarships covered only part of the cost of private school and applicants had to cover the rest of the costs. Indeed, the scholarships crowded-in educational expenses as households invested more total resources in education (Angrist et al. 2002). The conditioning of scholarship renewal created incentives that reduced grade repetition (Angrist et al 2002). The scholarship program was targeted to the poor and the poor typically don't obtain admission to public universities. The increased tertiary education induced by the scholarship was mainly at the expense of households themselves rather than the Colombian treasury. Additional financial (as opposed to time) investments by households in education generate positive fiscal externalities if the additional human capital of scholarship lottery winners increases long-run earnings. Moreover, there is no offsetting reduction on short-run labor supply (and hence short-run tax collection).[42]

The Colombian government has a number of transfer programs designed to support people at the bottom of the income distribution, such as *Familias en Acción*. One natural question is whether it costs more or less to redistribute to strata 1 and 2 households through the PACES private school scholarship program than through alternative means. It seems reasonable to assume that the social cost of transferring one dollar per household through conditional cash transfers is more than one dollar because such programs may distort labor supply or the economic activity among those taxed to pay for the program and among

[42] By contrast, other educational subsidy programs such as state merit aid programs in the US that pay for additional years of school and keep students in school longer (see for example Dynarski 2000; Thomas Kane 2003) will have offsetting effects. They reduce short-run tax revenue by delaying labor market entry and increase long-run revenue by boosting later earnings, with the overall impact on the NPV of tax revenue unclear.

beneficiaries who may seek to remain eligible. The evidence presented here suggests that it likely cost substantially less than a dollar to transfer one dollar in net present value to children born in strata 1 and 2 households through private scholarships.

References

Angrist, Joshua, Eric Bettinger, Eric Bloom, Michael Kremer, and Elizabeth King. 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.

Baird, Sarah, Joan Hamory Hicks, Michael Kremer and Edward Miguel. 2013. "Worms at work: Public finance implications of a child health investment" unpublished manuscript, Harvard University.

Barrera-Osorio, Felipe, Marianne Bertrand, Leigh Linden and Francisco Perez-Calle. 2011. "Improving the design of conditional transfer programs: Evidence from a randomized education experiment in Colombia." *American Economic Journal: Applied Economics* 3(2): 167-195.

Barrow, Lissa and Cecilia Elena Rouse. 2008. "School Vouchers: Recent Findings and Unanswered Questions." *Economic Perspectives*, Federal Reserve Bank of Chicago.

Bettinger, Eric, Michael Kremer, and Juan E. Saavedra. 2010. "Are Educational Vouchers Only Redistributive?" *The Economic Journal* 120(546): F204-F228.

Breierova, Lucia and Esther Duflo. 2004. "The impact of education on fertility and child mortality: Do fathers really matter less than mothers?" National Bureau of Economic Research Working Paper 10513.

Calderón, Alberto. 1996 "Voucher programs for secondary schools: The Colombian experience." World Bank Human Capital Development Working Paper 66.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Schanzenbach, and Danny Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR." *Quarterly Journal of Economics* 126 (4): 1593-1660.

Cortés, Darwin, Juan Miguel Gallego, and Darío Maldonado. 2010. "On the Design of Education Conditional Cash Transfer Programs and Non-Education Outcomes: The Case of Teen Pregnancy." Working Paper 100, Economics Department Universidad del Rosario, Bogotá.

Cowen, Joshua, David Fleming, John Witte, Patrick Wolf, and Brian Kisida. 2012. "Student Attainment and the Milwaukee Parental Choice Program: Final follow-up Analysis." SCDP Milwaukee Evaluation Report # 30.

Cunha, Flavio, James J. Heckman, Lance Lochner and Dimitriy Masterov. 2006. "Interpreting the evidence on life cycle skill formation," in *Handbook of Economics of Education* edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland.

Deming, David, Justine Hastings, Thomas Kane, and Doug Staiger. Forthcoming. "School Choice, School Quality and Postsecondary Attainment." *American Economic Review.*

Dynarski, Susan. 2000. "Hope for whom? Financial aid for the middle class and its impact on college attendance." *National Tax Journal* 53(3): 629-661.

Dynarski, Susan, Joshua Hyman, and Diane Schanzenbach. 2011. "Experimental Evidence on the Effects of Childhood Investments on Postsecondary Attainment and Degree Completion." Mimeo University of Michigan.

Duflo, Esther, Pascaline Dupas and Michael Kremer. 2012. "Education, HIV and early fertility: Experimental evidence from Kenya." Unpublished manuscript, Stanford University.

Encuesta Nacional de Calidad de Vida. 2010. Data source available at: http://190.25.231.249/metadatos/index.php/catalog. Retrieved on March 25, 2014.

Epple, Dennis and Richard Romano. 1998. "Competition between private and public schools, vouchers and peer-group effects." *American Economic Review* 88(1): 33-62.

Gutiérrez, Catalina, Olga Lucía Acosta and Eduardo Alfonso. 2013. "Financiación del sistema de salud: retos y perspectivas." *Notas de Política* 13 (April), Bogotá: Universidad de los Andes.

Heckman, James. 2008. "Schools, Skills, and Synapses." *Economic Inquiry* 46 (3): 289-324.

Heckman, James and Dimitriy Masterov. 2007. "The Productivity Argument for Investing in Young Children." *Review of Agricultural Economics* 29 (3): 446-493.

Hendren, Nathaniel. 2013. "The Policy Elasticity." Unpublished manuscript, Harvard University.

Hoffman, Saul. 2006. "By the numbers: The Public Costs of Teen Childbearing." Washington D.C.: The National Campaign to Prevent Teen Pregnancy.

Hsieh, Chang-Tai and Miguel Urquiola. 2006. "The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile's School Voucher Program." *Journal of Public Economics* 90: 1477-1503.

ICETEX. 2014. *Subsidio y sostenimiento*, http://www.icetex.gov.co/dnnpro5/es-co/cr%C3%A9ditoeducativo/estudiost%C3%A9cnicostecnol%C3%B3gicosyuniversitarios/largoplazoacces.aspx. Retrieved October 17, 2014.

ICFES. 2013. Personal email communication. May 6 of 2013.

Jaramillo, Christian and Jorge Tovar. 2008. "El impacto de impuesto al valor agregado sobre el gasto en Colombia. *Lecturas de Economía* 68 (enero-junio): 67-93.

Kane, Thomas. 2003. "A quasi-experimental estimate of the impact of financial aid on college-going." NBER working paper # 9703. Cambridge: National Bureau of Economic Research.

Kemple, James J. 2004. *Career Academies: Impacts on labor market outcomes and educational attainment.* New York: Manpower Demonstration Research Corporation.

King Elizabeth, Laura Rawlings, Marybell Gutierrez M, Carlos Pardo, and Carlos Torres. 1997. "Colombia's Targeted Education Voucher Program: Features, Coverage, and Participation." Impact Evaluation of Education Reforms Working Paper 3, Development Economics Research Group, World Bank.

Kremer, Michael, Edward Miguel and Rebecca Thornton. 2009. "Incentives to learn." *Review of Economics and Statistics* 91(3): 437-456.

Ladd, Helen. 2002. "School vouchers: A critical review." *Journal of Economic Perspectives,* 16(4): 3-24.

MacLeod, W. Bentley and Miguel Urquiola. 2013. "Competition and educational productivity: Incentives writ large" in *Education Policy in Developing Countries,* edited by Paul Glewwe, Chicago: University of Chicago Press.

Ministerio de Educación Nacional. 2010. "Resumen de estadísticas de las universidades públicas 2003-2010." Sistema Nacional de Información sobre la Educación Superior.

Ministerio de Educación Nacional. 2014a. Estadísticas sectoriales educación básica y media. http://menweb.mineducacion.gov.co/seguimiento/estadisticas/principal.php?seccion=2&consulta_detalle=et&id_categoria=2&nivel=2&et=&consulta=mat_sector, retrieved on February 20, 2014.

Ministerio de Educación Nacional. 2014b. Estadísticas sectoriales educación básica y media. http://menweb.mineducacion.gov.co/seguimiento/estadisticas/principal.php?seccion=9&id_categoria=2&consulta=coberturan_nivel&nivel=9&dpto=&et=&mun=&ins=&sede=&consulta_detalle=et, retrieved on February 20, 2014.

Muralidharan, Karthik and Venkatesh Sundararaman. 2013. "The aggregate effect of school choice: Evidence from a two-stage experiment in India." Unpublished manuscript, University of California, San Diego.

Parada-Baños, Arturo José. 2005. "El embarazo adolescente le cuesta al pais." UN Periódico

http://historico.unperiodico.unal.edu.co/Ediciones/72/05.htm, retrieved March 1, 2014.

Ribero, Rocio and Jorge Tenjo. 1997. "Evaluación del programa de becas PACES." Unpublished manuscript, Universidad de los Andes, Bogotá, Colombia.

Sánchez, Fabio and Jairo Núñez. 1995. "Por qué los niños pobres no van a la escuela? (Determinantes de la asistencia escolar en Colombia)." *Archivos de Macroeconomía* # 39. Departmento Nacional de Planeación, República de Colombia.

Secretaría de Desarrollo Económico de Bogotá. 2012. "Disminuye la informalidad laboral en Bogotá durante el segundo trimestre de 2012." DESR # 73, August.

United Nations. 2000. *World marriage patterns.* http://www.un.org/esa/population/publications/worldmarriage/worldmarriagepatterns2000.pdf, retrieved February 23, 2014.

Urquiola, Miguel. 2005. "Does School Choice Lead to Sorting? Evidence from Tiebout Variation." *American Economic Review* 95 (4): 1310-1326.

Wolf, Patrick, Babette Gutmann, Michael Puma, Brian Kisida, Lou Rizzo, Nada Eissa, and Matthew Carr. 2010. "Evaluation of the D.C. Opportunity Scholarship Program: Final Evaluation Report." U.S. Department of Education, Institute of Education Science, NCEE 2000-4023.

World Bank. 2003. "Tertiary Education in Colombia: Paving the Way for Reform." World Bank Country Study # 26289. Washington D.C.: The World Bank.

Table 1. Descriptive statistics of the Bogotá 1995 PACES scholarship applicant cohort

	All Applicants			Females			Males		
	Loser's Mean	Won a Scholarship	Won a Scholarship, Valid Age	Loser's Mean	Won a Scholarship	Won a Scholarship, Valid Age	Loser's Mean	Won a Scholarship	Won a Scholarship, Valid Age
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Data from PACES Application									
Has Phone	0.874	0.014	0.012	0.862	0.022	0.021	0.887	0.005	0.002
		(0.010)	(0.010)		(0.015)	(0.015)		(0.014)	(0.014)
Age at time of application	12.74	-0.078	-0.078	12.7	-0.080	-0.080	12.8	-0.077	-0.077
	-1.33	(0.043)*	(0.043)*	(1.340)	(0.060)	(0.060)	(1.310)	(0.061)	(0.061)
Male	0.487	0.005	0.007						
		(0.016)	(0.016)						
Applied to Vocational School	0.43	0.013	0.013	0.452	0.002	0.003	0.407	0.025	0.024
		(0.016)	(0.017)		(0.023)	(0.023)		(0.024)	(0.024)
B. National Identification Data									
Valid youth identification number	0.889	-0.002	-0.006	0.884	-0.012	-0.013	0.894	0.009	0.002
		(0.010)	(0.010)		(0.015)	(0.014)		(0.014)	(0.013)
Valid adult identification number	0.974	-0.005	-0.004	0.979	-0.01	-0.007	0.968	0.001	-0.002
		(0.005)	(0.005)		(0.007)	(0.006)		(0.008)	(0.007)
N	1,666	4,044	3,996	854	2,062	2,034	812	1,982	1,962

Notes: Table reports OLS scholarship lottery loser's means and estimated effects of winning a scholarship. Numbers in parentheses are standard deviations in columns of means and standard errors in columns of estimated scholarship effects. Results in Panel A are the same as those in Angrist et al. (2006) for having a phone, age at the time of application and gender, and as those in Bettinger et al. (2010) for having applied to a vocational school.

Table 2. Scholarship impacts on long-run educational outcomes

Outcome	All Applicants		Females		Males	
	Loser's Mean	Won a Scholarship	Loser's Mean	Won a Scholarship	Loser's Mean	Won a Scholarship
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Secondary school completion (ICFES College entry exam database)						
Applicant graduated on schedule	0.438	0.073	0.474	0.069	0.400	0.078
		(0.015)***		(0.021)***		(0.022)***
Applicant graduated with up to a two-year delay	0.513	0.062	0.541	0.065	0.484	0.060
		(0.015)***		(0.02)***		(0.021)***
Applicant graduated with up to a four-year delay	0.538	0.054	0.560	0.063	0.515	0.046
		(0.015)***		(0.02)***		(0.021)**
Applicant graduated with up to a six-year delay	0.550	0.054	0.576	0.060	0.523	0.048
		(0.015)***		(0.02)***		(0.021)**
N	1666	3996	854	2034	812	1962
Panel B. College enrollment and persistence (Higher education database)						
Ever enrolled in tertiary education	0.184	0.029	0.194	0.031	0.172	0.026
		(0.012)**		(0.018)*		(0.017)
Ever enrolled in a vocational college	0.062	0.017	0.069	0.004	0.055	0.031
		(0.008)**		(0.011)		(0.012)**
Ever enrolled in a university	0.129	0.017	0.136	0.028	0.122	0.005
		(0.011)		(0.016)*		(0.015)
Enrolled in tertiary education as of 2012	0.036	0.021	0.037	0.026	0.034	0.016
		(0.007)**		(0.01)**		(0.009)*
Enrolled in vocational college as of 2012	0.008	0.002	0.011	-0.002	0.006	0.006
		(0.003)		(0.004)		(0.004)
Enrolled in university as of 2012	0.028	0.019	0.027	0.027	0.028	0.010

		(0.006)**		(0.009)**		(0.008)
Enrolled in private institution as of 2012	0.030	0.016	0.032	0.024	0.028	0.009
		(0.006)**		(0.009)**		(0.008)
Enrolled in a public institution as of 2012	0.006	0.005	0.006	0.002	0.006	0.009
		(0.003)*		(0.004)		(0.005)*
Graduated from tertiary education as of 2012	0.048	0.009	0.055	0.016	0.041	0.002
		(0.007)		(0.011)		(0.009)
Years of tertiary education	0.405	0.078	0.441	0.120	0.367	0.032
	(1.175)	(0.038)**	(1.229)	(0.057)**	(1.115)	(0.052)
Ever received government financial aid	0.014	0.005	0.015	0.010	0.012	0.000
		(0.004)		(0.006)		(0.005)
Number of years of tertiary education loan subsidies	0.002	0.006	0.003	0.002	0.001	0.011
		(0.005)		(0.005)		(0.008)
N	1666	3996	854	2034	812	1962

Notes: Table reports scholarship lottery loser's means and estimated effects of winning a scholarship with application controls. Controls include age, male and whether the applicant had a phone number at the time of scholarship application. Estimates in columns 2,4 and 6 are from linear probability models. Numbers in parentheses are robust standard errors in columns of estimated scholarship effects. Graduated on schedule is if the applicant took the college entry test in 2001 or before; graduated with up to a two-, four- or six-year delay is if the applicant took the college entry test on or before 2003, 2005 and 2007, respectively. * significant 10%, ** significant 5%, *** significant 1%.

Table 3. Match rates to employment and earnings data

	Matched to SISBEN 2010 data		Matched to SISPRO data	
	(1)	(2)	(3)	(4)
Panel A. All applicants				
Won a scholarship	-0.026	0.002	0.022	-0.092
	(0.016)	(0.163)	(0.015)	(0.145)
Age * won a scholarship		0.003		0.006
		(0.012)		(0.011)
Phone * won a scholarship		-0.083		0.041
		(0.050)		(0.044)
Male * won a scholarship		0.006		0.000
		(0.032)		(0.029)
Loser's mean	0.518		0.689	
p-value on F-stat of joint test of interactions		0.406		0.770
N	3996	3996	3996	3996
Panel B. Females				
Won a scholarship	-0.03	0.052	0.020	-0.204
	(0.023)	(0.228)	(0.021)	(0.204)
Age * won a scholarship		0.002		0.010
		(0.017)		(0.016)
Phone * won a scholarship		-0.117		0.108*
		(0.067)		(0.061)
Male * won a scholarship				
Loser's mean	0.536		0.652	
p-value on F-stat of joint test of interactions		0.221		0.177
N	2034	2034	2034	2034
Panel C. Males				
Won a scholarship	-0.022	-0.053	0.024	0.065
	(0.023)	(0.234)	(0.021)	(0.206)
Age * won a scholarship		0.005		-0.001
		(0.017)		(0.015)
Phone * won a scholarship		-0.044		-0.034
		(0.074)		(0.065)
Male * won a scholarship			0.707	
Loser's mean	0.499			
p-value on F-stat of joint test of interactions		0.797		0.870
N	1962	1962	1962	1962

Notes: Table reports scholarship lottery loser's means and estimated effects of winning a scholarship on the probability of being matched to SISPRO and to SISBEN 2010 data using linear probability models. Additional controls, not shown in the table include, age, male and whether the applicant had a phone number at the time of scholarship application. SISPRO sample is from July 2008 to 2012.

Table 4. Scholarship impacts on formal sector employment intensity

	All Applicants		Female		Male	
Applicant spent at least…months in formal sector employment in 2008-2012	Loser's Mean (1)	Won a Scholarship (2)	Loser's Mean (3)	Won a Scholarship (4)	Loser's Mean (5)	Won a Scholarship (6)
5	0.599	0.025	0.584	0.008	0.615	0.043
	(0.49)	(0.016)	(0.493)	(0.022)	(0.487)	(0.022)*
10	0.541	0.021	0.522	0.007	0.561	0.035
	(0.498)	(0.016)	(0.5)	(0.022)	(0.497)	(0.023)
15	0.489	0.025	0.452	0.031	0.528	0.02
	(0.5)	(0.016)	(0.498)	(0.022)	(0.5)	(0.023)
20	0.444	0.025	0.409	0.027	0.481	0.021
	(0.497)	(0.016)	(0.492)	(0.022)	(0.5)	(0.023)
25	0.402	0.016	0.374	0.007	0.431	0.025
	(0.49)	(0.016)	(0.484)	(0.022)	(0.496)	(0.023)
30	0.356	0.007	0.337	-0.001	0.375	0.016
	(0.479)	(0.016)	(0.473)	(0.021)	(0.485)	(0.023)
35	0.301	0.005	0.293	-0.011	0.309	0.021
	(0.459)	(0.015)	(0.455)	(0.02)	(0.462)	(0.022)
40	0.229	0.015	0.233	-0.004	0.225	0.036
	(0.421)	(0.014)	(0.423)	(0.019)	(0.418)	(0.02)*
N	1622	3903	836	1992	786	1911

Notes: Table reports scholarship lottery loser's means and estimated effects of winning a scholarship with application controls. Controls include age, male and whether the applicant had a phone number at the time of scholarship application. Estimates in columns 2,4 and 6 are from linear probability models. Numbers in parentheses are robust standard errors in columns of estimated scholarship effects. Formal sector employment is based on the health payroll tax form. Sample is restricted to applicants with valid adult identification number (3926) that have complete application controls (3903).

Table 5. Scholarship impacts on annual earnings and payroll taxes

	All Applicants		Female		Male	
	Loser's Mean (s.d) (Panel A) Loser's Earnings at Quantile (Panel B)	Won a Scholarship	Loser's Mean (s.d) (Panel A) Loser's Earnings at Quantile (Panel B)	Won a Scholarship	Loser's Mean (s.d) (Panel A) Loser's Earnings at Quantile (Panel B)	Won a Scholarship
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. OLS regressions						
SISBEN 2010 Self-reported Earnings (observed values)	2,481.6	-164.4	1,926.0	-46.8	3,150.0	-312.0
	(5,888.4)	(201.6)	(2,269.2)	(128.4)	(8,337.6)	(418.8)
N	904	2120	494	1155	410	965
SISBEN 2010 Self-reported Earnings Trimming Top 5% of Loser Earners	1,992.0	308.4	1,676.4	202.8	2,514.0	312.0
	(1,580.4)	(75.6)***	(1,567.2)	(105.6)*	(1,582.8)	(109.2)***
N	842	2059	468	1129	392	948
Mean Formal Sector Earnings 2008-2012	2,201.3	172.7	2,033.2	144.3	2,380.2	200.0
	(2,874.8)	(95.1)*	(2,736.8)	(129.5)	(3,006.1)	(139.4)
N	1622	3903	836	1992	786	1911
Payroll Taxes	597.7	44.6	533.4	36.8	666.2	52.1
	(810.6)	(26.7)*	(751.2)	(35.6)	(864.6)	(40.0)
N	1622	3903	836	1992	786	1911
Projected Formal Sector Earnings 2014	2,285.8	213.3	2,109.7	239.4	2,473.1	183.9
	(2,918.6)	(95.8)**	(2,789.1)	(131.6)*	(3,040.8)	(139.2)
N	1622	3903	836	1992	786	1911
Panel B. Quantile regressions of annual formal earnings 2008-2012						
70th	2,771.3	92.4	2,392.2	174.8	3,012.5	106.4
		(129.6)		(168.3)		(167.4)
75th	3,154.1	134.0	2,905.7	173.7	3,441.3	145.2

		(123.9)		(174.2)		(177.4)
80th	3,647.7	95.3	3,344.9	(62.5)	3,938.4	183.7
		(142.7)		(207.9)		(213.0)
85th	4,235.6	146.3	3,950.1	49.2	4,474.3	197.0
		(173.8)		(253.5)		(260.1)
90th	5,057.5	368.3	4,816.2	312.5	5,452.1	427.2
		(192.2)*		(265.1)		(309.1)
95th	6,631.6	658.0	6,428.8	771.4	6,749.6	650.8
		(377.4)*		(597.2)		(570.3)
N	1622	3903	836	1992	786	1911

Notes: Table reports scholarship lottery loser's means (Panel A), loser's means at quantiles (Panel B) and estimated effects of winning a scholarship on self-reported monthly earnings and total formal sector earnings in USD of 2013. Total payroll taxes include employer and employee contributions. Numbers in parentheses are standard deviations in columns of means and standard errors in columns of estimated scholarship effects. Controls include age, male and whether the applicant had a phone number at the time of scholarship application. Formal sector earnings estimates use the health payroll account. Formal sector earnings sample is restricted to begin in July 2008 and is restricted to applicants with valid adult identification number (3926) that have complete application controls (3903). For projected total sector formal earnings 2014 we assume that current tertiary education enrollees would make two years out what 2010 tertiary graduates in our dataset earn, on average, in 2012..* significant 10%, ** significant 5%.

Table 6. Scholarship impacts government subsidy receipt

Outcome is applicant receives/is eligible for:	All		Females		Males	
	Loser's Mean	Won a Scholarship	Loser's Mean	Won a Scholarship	Loser's Mean	Won a Scholarship
	(1)	(2)	(3)	(4)	(5)	(6)
Familias En Acción CCT program	0.074	-0.003	0.087	0.010	0.062	-0.017
		(0.008)		(0.013)		(0.011)*
Subsidized Health Care Level 1	0.194	-0.009	0.200	0.008	0.188	-0.026
		(0.013)		(0.018)		(0.018)
Subsidized Health Care Level 2	0.245	0.000	0.258	0.006	0.232	-0.005
		(0.014)		(0.020)		(0.019)
Early childhood care (ICBF)	0.271	-0.003	0.286	-0.008	0.255	0.002
		(0.014)		(0.020)		(0.020)
N	1666	3996	854	2034	812	1962

Notes: Table reports scholarship lottery loser's means and estimated effects of winning a scholarship with application controls. Controls include age, male and whether the applicant had a phone number at the time of scholarship application. Estimates in columns 2,4 and 6 are from linear probability models. Numbers in parentheses are robust standard errors in columns of estimated scholarship effects. Receipt of Familias en Acción CCT program is based on having SISBEN 2010 scores below the eligibility cutoff and children under the age of 18. Receipt of subsidized health care levels 1 and 2, and early childhood care is based on having SISBEN 2010 scores below the eligibility cutoff. Applicants who are not in SISBEN 2010 cannot receive these subsidies so for them receipt is zero. * significant 10%, ** significant 5%, *** significant 1%

Table 7. Scholarship effects on teen fertility

	All applicants		Females		Males	
	Loser's Mean	Won a Scholarship	Loser's Mean	Won a Scholarship	Loser's Mean	Won a Scholarship
	(1)	(2)	(3)	(4)	(5)	(6)
			SISBEN 2010 data			
Had any child as a teen	0.234	-0.040	0.375	-0.061	0.074	-0.018
		(0.018)**		(0.03)**		(0.017)
Spouse/partner had a child as a teen	0.103	-0.024	0.054	-0.001	0.160	-0.050
		(0.013)*		(0.014)		(0.023)**
Total number of children	1.064	-0.025	1.355	0.009	0.734	-0.064
	(1.032)	(0.043)	(1.030)	(0.061)	(0.932)	(0.059)
One or more children	0.635	-0.011	0.776	-0.002	0.476	-0.022
		(0.021)		(0.025)		(0.033)
Two or more children	0.322	-0.013	0.435	-0.014	0.195	-0.013
		(0.020)		(0.030)		(0.026)
Three or more children	0.081	-0.003	0.109	0.011	0.050	-0.020
		(0.012)		(0.019)		(0.013)
N	850	2002	451	1053	399	949

Notes: Table reports scholarship lottery loser's means and estimated effects of winning a scholarship on adolescent fertility outcomes based on SISBEN 2010 data. Estimates in columns 2, 4, and 6 are from linear probability models. Numbers in parentheses are robust standard errors with the exception of total number of children in columns 1, 3 and 5, which are standard deviations of the loser's mean. Controls include age, male and whether the applicant had a phone number at the time of scholarship application. * significant 10%, ** significant 5%.

Table 8. Fiscal impacts of scholarships

		FEMALES	MALES	Notes
Row #	*A. Government Costs*			
	Secondary Education Costs			
1	Annual per-pupil cost of public school	$449.08	$449.08	From Angrist et al. (2002), converted to 2013 dollars
2	Annual value of PACES scholarship	$244.00	$244.00	From Angrist et al. (2002), converted to 2013 dollars
3	Expenditure from scholarship costs for students who would have enrolled in private school, aggregated over 6 years	$494.38	$451.28	Row (2)*Proportion of lottery winners attending private school*proportion of winners continuing to use scholarships; computed annually and summed over the six years
4	Expenditure resulting from transfers from public to private schools, aggregated over 6 years	$(177.78)	$(172.71)	(Row (2) – Row (1))*Scholarship Effect on Private School Attendance; computed annually and summed over the six years
5	Cost savings from reduced grade repetition	$(1.75)	$(6.89)	See Appendix B
6	Total secondary education costs to the government	$314.85	$271.68	Row (3) + Row (4) + Row (5)
	Tertiary Education Costs			
7	Additional public tertiary education costs	$16.93	$11.21	Annual per-pupil expenditure in public tertiary education * Scholarship impact on years of tertiary education (Panel B, Table 2, cols. 4 & 6)* Fraction of lottery winners attending a public institution (Panel B, Table 2, col. 1 + col. 3)
8	Additional tertiary education loan subsidies	$0.74	$4.08	Annual per-pupil tertiary education subsidy * Scholarship impact on number of years of subsidy receipt (Panel B of Table 2, cols. 4 & 6)
9	Additional tertiary education costs (public education + loan subsidies)	$17.68	$15.29	Row (7) + Row (8)
	Welfare Receipt Costs			
10	Additional CCT receipt costs	$2.00	$(3.39)	Annual CCT subsidy amount (see notes below) * Scholarship impact on CCT receipt (Table 6, cols. 4 & 6)
	Foregone Revenue			
11	Foregone tax revenue from VAT tax	$22.94	$26.85	Formal annual earnings of losers* Scholarship impact on years of education *VAT tax of 13.3%
12	Foregone net government transfers through payroll taxes	$18.09	$22.60	Annual payroll taxes of losers* Scholarship impact on years of education *0.4
13	Total foregone revenue	$41.03	$49.45	Row (11) + Row (12)
14	**NPV of expected direct scholarship costs to government**	**$375.56**	**$333.03**	Row (6) + Row (9) + Row (10) + Row (13)
15	NPV of expected direct scholarship costs to government, 6% discount rate	$349.43	$312.02	Same calculations as above using 6% discount rate

	B. Government Revenue			
16	Additional VAT tax revenue	$887.18	$681.51	Additional earnings of scholarship winners (see notes below) * VAT tax of 13.3%
17	Additional government transfers through payroll taxes	$410.15	$580.68	Additional payroll taxes of scholarship winners (see notes below) * 0.4
18	**NPV of additional revenue to government**	**$1,297.33**	**$1,262.18**	Row (15) + Row (16)
19	NPV of additional government revenue, 6% discount rate	$614.33	$597.69	Same calculations as above using 6% discount rate
20	NPV of additional government revenue, no VAT revenue	$410.15	$580.68	Assume VAT revenue is zero
	C. Gains to Recipients			
21	**Net gains to scholarship recipients**	**$262.46**	**$231.41**	Fraction of infra-marginal recipients*impact on scholarship amount*sum of year-by-year utilization rate. See notes below
22	**Net Benefits to society**	**$1,184.23**	**$1,160.56**	Additional revenue to government (Row 18)+ Net gains to scholarship recipients (Row 21) - Scholarship costs to government (Row 14)

Notes: We express all figures in 2013 dollars per scholarship winner. For annual per-pupil costs of public school and scholarship impact on scholarship value three years after the lottery: Angrist et al. (2002) report the cost of public schooling in 1998 to be $350 and the scholarship cost to be $190. We follow order of operations described in text to obtain values in analysis year. For Expenditure from scholarship costs for students who would have enrolled in private school, aggregated over 6 years: The fraction of females that attend private school is 0.897 (6^{th}), 0.699 (7^{th}) and 0.535 (8^{th}). For males it is 0.857 (6^{th}), 0.646 (7^{th}) and 0.543 (8^{th}). We do not observe private school attendance for grades 9^{th} or 10^{th}. We observe private school graduation, which is 0.322 for males and 0.314 for males. We interpolate linearly between the 8^{th} grade rate and the graduation rate to obtain the private school attendance rates for grades 9^{th} and 10^{th}, which we estimate to be 0.464 (9^{th}) and 0.393 (10^{th}) for females and 0.469 (9^{th}) and 0.396 (10^{th}) for males. We observe the fraction of winners in private school using the scholarship for grades 6^{th} and 8^{th} only. The 7^{th} grade fraction is the linear combination of the 6^{th} and 8^{th} grade rates. For females the fraction of winners in private school using the scholarship is 0.953 (6^{th}), 0.736 (7^{th}) and 0.519 (8^{th}). For males it is 0.933 (6^{th}), 0.698 (7^{th}) and 0.463 (8^{th}). After 8^{th} grade, we have no data on scholarship usage. We know that 32 percent of the overall lottery loser sample finished 11^{th} grade in private school. We assume a constant deterioration from 8^{th} grade to 11^{th} grade in the fraction of losers attending private school. This implies a 40 percent reduction in the fraction attending private school, and we assume that deterioration in scholarship usage among winners follows a similar 40 percent decline from the 8^{th} grade level. Under these assumptions, scholarship usage rates for females are 0.415 (9^{th}), 0.310 (10^{th}) and 0.206 (11^{th}). For males the scholarship usage rates are 0.374 (9^{th}), 0.284 (10^{th}) and 0.195 (11^{th}). For cost-savings from reduced grade repetition: See Appendix B. For tertiary costs: Average per-pupil government expenditure in tertiary education is COP 3,280,000 in 2010 (Ministry of Education 2010). We use the exchange rate of COP 1913.98/USD (Dec 31, 2010) to convert to nominal USD and follow order of operations described in text to obtain NPV in USD of analysis year. Tertiary education subsidies are COP 682.432 per semester in COP of 2013 (ICETEX 2014). We use the exchange rate of COP 1923.83/USD (Dec 31, 2013) to convert to nominal USD and follow order of operations described in text to obtain NPV in USD of analysis year. Costs of welfare receipt: To obtain annual cost we assume one child, which is the mean number of children of scholarship applicants at age 28 (see Table 7). We assume child is between zero and seven years of age in 2013, so can receive health transfer but no education transfer. We assume applicant resides in Bogotá so monthly health subsidy amount is that for Group 1 municipalities, COP 61,200/month (see: http://www.dps.gov.co/Ingreso_Social/FamiliasenAccion.aspx, retrieved October 28, 2014). We obtain annual CCT transfer amount by multiplying by 12. We follow order of operations above to obtain NPV in USD of analysis year. We only assume one year of costs since there is no difference by scholarship status in total fertility, indicating simply a difference in the probability of having age-appropriate children. Therefore, CCT cost is annual cost * impact on receipt.

Foregone earnings: We estimate annual foregone revenue from average annual formal formal sector earnings of scholarship losers in Panel A of Table 5, columns 3 and 5. We follow order of operations above to obtain NPV in USD of analysis year. Earnings: Annual earnings are projected annual earnings from Panel A, Table 5.. We project earnings for losers and winners over a 35-year horizon allowing for a 3.02% annual growth in earnings per annum, which is the average annual growth in GDP per capita in Colombia between 2002 and 2012, obtained from the World Development Indicators database). US-CPI for years after 2013 is that for 2013. We then follow remaining order of operation to obtain NPV of earnings for winners and losers. Payroll taxes: Annual payroll taxes are from Panel A, Table 5. We follow the same procedure as for earnings to obtain the NPV of payroll taxes in analysis year. Benefits to recipients: The fraction of infra-marginal recipients is the fraction of lottery losers who attend private school in 6^{th} grade, 0.897 among females and 0.857 among males. The impact on scholarship amount is $93.2 (from Angrist et al. 2002 Table 8, column 3 updated to 2013 dollars). For utilization rates see notes above for Expenditure from scholarship costs for students who would have enrolled in private school, aggregated over 6 years.

Appendix A. Impacts on Vocational School Applicants

Among applicants to schools with a vocational curriculum, lottery winners did not attend schools with more desirable peers (Bettinger et al. 2010). We show that scholarship impacts on this sub-population are comparable to those in the full Bogotá 1995 lottery applicant cohort.

Educational Outcomes

Vocational scholarship lottery winners are 18 percent (8.4 percentage points) more likely to finish secondary school on time and 10 percent more likely to ever finish secondary school relative to the losers' means of 46 percent and 59 percent, respectively (Table A1).

Vocational school winners are 42 percent (8.6 percentage points) more likely to enroll in tertiary education relative to the losers' mean of 20.5 percent. Effects are slightly larger among male applicants although statistically indistinguishable from effects on females. Vocational school winners are also 108 percent (3.7 percentage points) more likely to be currently enrolled in tertiary education relative to the losers' mean of 3.4 percent.

In the vocational sample, scholarship lottery winners are 43 percent (2.4 percentage points) more likely to graduate from tertiary education than losers, an effect entirely driven by male applicants. The graduation effect is significant at the 10 percent level. Vocational scholarship lottery winners accumulate 0.19 additional years of tertiary education (40 percent relative to the losers' rate of 0.47 years). Male applicants drive the effect on additional years of tertiary education in the vocational sample.

Labor Market Outcomes

Current annual formal sector earnings are, on average, 12 percent higher for winners who applied to vocational school than for losers among those who applied to vocational

school. This estimate is statistically significant at the 10% level in the full vocational sample, but not separately by gender (Panel A, Table A2).

As is the case in the full applicant sample, the scholarship OLS estimate on total earnings understates the long-term effect due to the higher likelihood of continued tertiary enrollment among scholarship lottery winners. We follow the same procedure as with the full applicant sample described in the main text to estimate the long-term earnings effect. The scholarship impact on annual estimated 2014 formal sector earnings is $326.6, and statistically significant at the 10% level in the full sample (Panel A, Table A2). Relative to the losers' mean, this impact represents a 13.2 percent increase. Future earnings differences will be larger, of course, because the 2008-2014 period includes a period when more winners are still in tertiary education.

Among vocational applicants, scholarship lottery winners also pay higher annual payroll taxes. The OLS scholarship impact on payroll taxes is $78.9, which at the losers' mean represents an increase of 12 percent. This estimate on payroll taxes is statistically significant at the 10 percent (Panel A, Table A2).

Quantile regression results on annual formal earnings suggest that scholarship impacts are stronger higher in the distribution. Estimates, however, are only statistically significant at the 10 percent level for the 85^{th} quantile in the full vocational sample (Panel B, Table A2).

Table A1. Scholarship effects on long-run educational outcomes among vocational school applicants

Outcome	All Applicants		Females		Males	
	Loser's Mean	Won a Scholarship	Loser's Mean	Won a Scholarship	Loser's Mean	Won a Scholarship
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Secondary school completion (ICFES College entry exam database)						
Applicant graduated on schedule	0.461	0.084	0.500	0.062	0.415	0.109
		(0.024)***		(0.032)*		(0.036)***
Applicant graduated with up to a two-year delay	0.557	0.064	0.599	0.050	0.507	0.080
		(0.023)***		(0.031)		(0.035)**
Applicant graduated with up to a four-year delay	0.576	0.064	0.613	0.052	0.533	0.078
		(0.023)***		(0.031)*		(0.035)**
Applicant graduated with up to a six-year delay	0.593	0.060	0.637	0.043	0.539	0.080
		(0.023)***		(0.03)		(0.035)**
N	*559*	*1415*	*306*	*747*	*253*	*668*
Panel B. College enrollment and persistence (Higher education database)						
Ever enrolled in tertiary education	0.205	0.086	0.220	0.067	0.183	0.108
		(0.025)**		(0.033)**		(0.038)**
Ever enrolled in a vocational college	0.058	0.049	0.059	0.025	0.058	0.078
		(0.016)**		(0.020)		(0.026)**
Ever enrolled in a university	0.155	0.051	0.172	0.044	0.131	0.059
		(0.023)**		(0.031)		(0.033)*
Currently enrolled in tertiary education	0.034	0.037	0.044	0.031	0.021	0.046
		(0.013)**		(0.019)		(0.018)**
Currently enrolled in vocational college	0.002	0.012	0.004	0.008	0.000	0.017
		(0.005)**		(0.007)		(0.007)**
Currently enrolled in university	0.032	0.025	0.040	0.023	0.021	0.029
		(0.012)**		(0.018)		(0.016)*
Currently enrolled in private institution	0.032	0.023	0.040	0.026	0.021	0.019

		(0.012)*		(0.018)		(0.015)
Currently enrolled in a public institution	0.002	0.014	0.004	0.004	0.000	0.026
		(0.005)**		(0.006)		(0.009)**
Graduated from tertiary education	0.056	0.024	0.073	0.016	0.031	0.034
		(0.015)*		(0.022)		(0.019)*
Years of tertiary education	0.474	0.191	0.538	0.169	0.382	0.214
	(1.289)	(0.082)**	(1.374)	(0.114)	(1.154)	(0.114)*
Ever received government financial aid	0.015	0.008	0.018	0.007	0.010	0.009
		(0.008)		(0.012)		(0.011)
N	464	1117	273	629	191	488

Notes: Sample is Bogotá 1995 scholarship cohort restricted to applicants to schools with vocational curricula. Table reports scholarship lottery loser's means and estimated effects of winning a scholarship for applicants to vocational schools with application controls. Controls include age and whether the applicant had a phone number at the time of scholarship application. Estimates in columns 2,4 and ,6 are from linear probability models. Numbers in parentheses are standard deviations in columns of means and standard errors in columns of estimated scholarship effects. Graduated on schedule is if the applicant took the college entry test in 2001 or before; graduated with up to a two-, four- or six-year delay is if the applicant took the college entry test on or before 2003, 2005 and 2007, respectively. * significant 10%, ** significant 5%, *** significant 1%.

Table A2. Scholarship impacts on annual earnings and payroll taxes for vocational school applicants

	All Applicants		Female		Male	
	Loser's Mean (s.d) (Panel A) Loser's Earnings at Quantile (Panel B)	Won a Scholarship	Loser's Mean (s.d) (Panel A) Loser's Earnings at Quantile (Panel B)	Won a Scholarship	Loser's Mean (s.d) (Panel A) Loser's Earnings at Quantile (Panel B)	Won a Scholarship
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. OLS regressions						
SISBEN 2010 Self-reported Earnings (observed values)	2,347.2	-27.6	2,068.8	-117.6	2,733.6	128.4
	(2,419.2)	(156.0)	(2,832.0)	(232.8)	(1,618.8)	(176.4)
N	365	868	212	498	153	370
SISBEN 2010 Self-reported Earnings Trimming Top 5% of Loser Earners	2,078.4	249.6	1,752.0	205.2	2,581.2	286.8
	(1,606.8)	(120)**	(1,622.4)	(165.6)	(1,486.8)	(169.2)*
N	344	847	201	487	146	363
Mean Formal Sector Earnings 2008-2012	2,374.5	273.2	2,227.5	184.4	2,550.2	382.6
	(3,029.2)	(166.2)*	(2,844.4)	(221.8)	(3,232.3)	(250.6)
N	652	1605	355	850	297	755
Payroll Taxes	641.8	78.9	585.5	51.6	709.2	112.4
	(841.8)	(46.4)*	(777.6)	(60.9)	(909.3)	(71.0)
N	652	1605	355	850	297	755
Projected Formal Sector Earnings 2014	2,468.0	326.6	2,291.3	293.9	2,679.2	372.6
	(3,071.0)	(167.3)*	(2,869.7)	(223.2)	(3,288.1)	(252.4)
N	652	1605	355	850	297	755
Panel B. Quantile regressions of annual formal earnings 2008-2012						
70th	2,927.3	51.3	2,828.5	95.1	3,125.4	358.4
		(218.4)		(287.8)		(326.1)

75th	3,416.7	124.6	3,243.4	(15.7)	3,643.9	298.0
		(224.6)		(288.9)		(380.5)
80th	3,886.6	129.2	3,665.3	-111.2	4,055.2	294.9
		(268.4)		(357.2)		(420.0)
85th	4,587.9	525.5	4,373.4	224.0	4,818.5	836.3
		(309.9)*		(430.2)		(482.8)*
90th	5,484.5	532.1	5,252.4	452.3	5,615.4	1,036.3
		(432.3)		(523.0)		(759.6)
95th	7,181.3	1,176.2	6,952.2	905.0	7,711.7	1,373.5
		(813.9)		(1,030.1)		(1,038.9)
N	652	1605	355	850	297	755

Notes: Sample is restricted to vocational school applicants with valid adult identification numbers and complete application controls. Table reports scholarship lottery loser's means (Panel A), loser's means at quantiles (Panel B) and estimated effects of winning a scholarship on self-reported monthly earnings and total formal sector earnings in USD of 2013. Total payroll taxes include employer and employee contributions. Numbers in parentheses are standard deviations in columns of means and standard errors in columns of estimated scholarship effects. Formal sector earnings estimates use the health payroll account. Formal sector earnings sample is restricted to begin in July 2008 and is restricted to applicants with valid adult identification number that have complete application controls. For projected total sector formal earnings 2014 we assume that current tertiary education enrollees would make two years out what 2010 tertiary graduates in our dataset earn, on average, in 2012.Controls include age, male and whether the applicant had a phone number at the time of scholarship application. * significant 10%.

Appendix B. Calculation of cost-savings to the government from reduced grade repetition

This appendix explains how we calculate cost savings to the government from reduced grade repetition. We observe public school attendance and repetitions through grade 8 from Table 4 in Angrist et al. (2002) and whether the applicant finished secondary school on time, with delays or did not finish (Table 2 in main text).

To calculate the cost-savings from reduced grade repetition, we assume that: i) only public school repetitions cost the government money, ii) among applicants who finish on-schedule, there is no repetition; ii) among applicants who finish with delays, the delays are all a consequence of grade repetition and iii) among those who never finish secondary school they dropped out in 9th grade, so that total repetitions for this group are the ones reported in Table 4 columns 2 and 4 of Angrist et al. 2002.

There are three types of students: those who finish on time; those who finish with delays; and those who never finish. For those who finish on time, the government receives no cost savings. For those who pass with delays, the government saves from reduced grade repetition in the public sector. We multiply public school costs by the fraction in public and by the overall reduction in grade repetitions to estimate these cost benefits. For those who never graduate, we only record the savings from grade repetition after three years. As before, we multiply public school costs by the fraction in public by the effect of the scholarship on repetitions after three years. We use data from Angrist et al (2002) to compute these effects and the fractions in public.

Based on estimates from Table 2, Table B1 shows the distribution of secondary school completion outcomes for scholarship winners and losers, separately by gender:

Table B1. Distribution of secondary school completion outcomes

	Losers		Winners	
	Female	Male	Female	Male
Completed secondary school on time	0.474	0.4	0.543	0.478
Completed secondary school with delays	0.102	0.123	0.093	0.093
Never completed	0.424	0.477	0.364	0.429

Notes: Completed with delays is completed with up to a six year delay. We assume that the fraction of applicants who never complete secondary school is 1 – (fraction who complete on time + fraction who complete with delays).

For those who complete secondary school with delays, since we assume that the delay is all a consequence of grade repetition, the reduction in grade repetition as a consequence of winning the scholarship is 0.093 – 0.102 = -0.009 for females and 0.093 – 0.123 = -0.030 for males. Annual cost-savings from reduced grade repetition in this group is annual per-pupil public school costs (from Table 8), times the fraction of lottery losers who attend public school, times reduction in the probability of grade repetition. For females this is: \$449.08 * 0.284 * (-0.009) = - \$1.15 and for males it is \$449.08 * 0.300 * (-0.03) = - \$4.04.

We need to multiply these annual amounts by the number of extra years that it takes for winners to graduate from secondary school among those that graduate with delays. Table B2 shows the distribution of delayed graduation, which we obtain from Table 2 in the main text. Using the distribution of delayed graduation in Table B2 and the annual cost-savings from reduced grade repetition, we get the cost savings from grade repetition among scholarship winners who complete secondary school with delays. For females this is: - [$1.15*2 years*0.68 (fraction who completes with up to a two-year delay) + $1.15*4 years*0.18 + $1.15*6 years*0.14] = - $3.35. For males it is: - [$4.04*2 years*0.71 + $4.04*4 years*0.18 + $4.04*6 years*0.11] = - $11.32.

For scholarship winners who never complete secondary school, we assume that they dropped out in 9th grade so the reduction in the total number of repetitions is the one reported by Angrist et al (2002) in Table 4 for the Bogotá sample with controls, which is -0.031 for females and -0.101 for males. Therefore, cost-savings for those who never complete is $449.08*0.284*(-0.031) = -$3.95 for females and $449.088*0.300*(-0.101) = -$13.61 for males.

Table B2. Distribution of delayed secondary school completion for scholarship winners

	Females		Males	
	Percentage points	Percent	Percentage points	Percent
Fraction of winners who complete secondary school with delays	0.093	100%	0.093	100%
Fraction who complete with up to a two-year delay	(0.541+0.065)-0.543 = 0.063	68%	(0.484+0.060)-0.478 = 0.066	71%
Fraction who complete with a 2- to 4-year delay	(0.560+0.063)-0.606 = 0.017	18%	(0.515+0.046)-0.544 = 0.017	18%
Fraction who complete with a 4- to 6-year delay	(0.093-0.063-0.017) = 0.013	14%	(0.093-0.066-0.017) = 0.010	11%

Notes: The first row is from Table B1. The remain rows are from Table 2 in the main text.

Total cost-savings from reduced grade repetition among scholarship winners is the weighted sum of the cost-savings among those who complete secondary school with delays and those who never complete. The weights are given by the fraction of scholarship winners who complete secondary school with delays and who never complete, from Table B1. For females, we have that total cost-savings are [-$3.35* 0.093 - $3.95*0.364] = -$1.75. For males, total cost-savings from reduced grade repetition are [-$11.32*0.093 -$13.61*0.429] = -$6.89