LONG-TERM DIRECT AND SPILLOVER EFFECTS OF JOB TRAINING:
EXPERIMENTAL EVIDENCE FROM COLOMBIA

Adriana Kugler
Maurice Kugler
Juan Saavedra
Luis Omar Herrera Prada

ABSTRACT

We use administrative data to examine medium and long-term formal education and labor market impacts among participants and family members of a randomized vocational training program for disadvantaged youth in Colombia. In the Colombian program, vocational training and formal education are complementary investments: relative to non-participants, randomly selected participants are more likely to complete secondary school and to attend and persist in tertiary education eight years after random assignment. Complementarity is strongest among applicants with high baseline educational attainment. Training also has educational spillover effects on participants' family members, who are more likely to enroll in tertiary education. Between three and eight years after randomization, participants are more likely to enter and remain in formal employment, and have formal sector earnings that are at least 11 percent higher than those of non-participants.

Adriana Kugler
Georgetown University
McCourt School of Public Policy
37th and O Streets NW, Suite 311
Washington, DC 20057
and NBER
ak659@georgetown.edu

Maurice Kugler
IMPAQ International LLC
1101 Vermont Avenue NW
Suite 1100
Washington, DC 20005
mkugler@impaqint.com

Juan Saavedra
University of Southern California
635 Downey Way
Los Angeles, CA 90089-3331
juansaav@usc.edu

Luis Omar Herrera Prada
Inter American Development Bank
1300 New York Avenue, N.W.
Washington, D.C. 20577
luis.o.herrera@me.com

I. Introduction

Job training programs aim at improving the employment prospects of individuals who would otherwise have difficulty integrating into the labor force. For instance, individuals who drop out of the formal education system before finishing secondary school typically have poorer employment prospects including lower formal sector attachment and lower wages.

Vocational training offers a second chance to those who drop out of the formal education system to obtain and upgrade job-specific skills that can differentiate them from other dropouts in the labor market. For example, Attanasio, Kugler and Meghir (2011)—AKM henceforth—find that a well-designed randomized vocational training program in Colombia, which matched the skills taught to the skills required in the labor market, increases earnings for women by close to 20 percent and increases formal employment for both men and women by up to 7 percentage points. Another paper, circulated contemporaneously with the current one, finds that these early labor effects of the Colombian training program persist in the medium term (Attanasio et al. 2015). AKM and Attanasio et al. (2015) find that the employment benefits of vocational training for women exceed program costs even when assuming that the skills depreciate over time. In other developing countries, however, evidence from the few randomized controlled trials conducted suggests smaller short- and long-term impacts on the employment prospects of participants (Card et al., 2011; Hicks et al., 2013; Hirshleifer et al., 2013; Ripani et al., 2015).

One glaring limitation of most of these studies, however, is that they typically *only* analyze labor market outcomes after participation. Such rationale assumes that among disadvantaged youth, job training is a substitute of formal education. Job training programs in which participants receive vocational training alongside subsidized private sector employment may, however, complement additional educational investments to the extent they enable

participants to learn about their ability, raise expectations or help them relax educational credit constraints. While others have argued earlier that skill begets skill (e.g. Heckman, 2000), no study to date has examined whether skill acquired through vocational training begets formal education skills. Such complementarity would imply, among other things, that welfare calculations based on short-term employment and earnings impacts alone may considerably underestimate the social desirability of these kinds of programs.

In addition, to the extent that families share informational and financial resources, it is plausible that vocational training may also generate spillover benefits on others as well. Omitting potential spillover effects from welfare calculations also underestimates the desirability of vocational training programs. While prior studies have examined the external effects of formal education (e.g. Heckman and Lochner, 2000), no study has examined the extent to which job training generates spillover effects.

In this paper, we aim to address these gaps in the literature by examining experimental evidence on the long-term direct and spillover formal education and labor market effects of a randomized vocational training program for disadvantaged youth in Colombia. We match data from the randomization sample collected by AKM to various sources of administrative data which allow us to follow formal education and labor market trajectories of individuals and family members between three and eight years after participation in the program.

Training lottery winners are 1.4 percentage points (27 percent relative to the lottery losers' mean) more likely than losers to complete secondary school one year after training participation. We find no effects of training on secondary school completion after two years of training participation suggesting that the effect on secondary school completion is concentrated immediately after training participation.

Training lottery winners are 3.5 percentage points (31 percent) more likely than losers to enroll in formal tertiary education between three and eight years after training participation. Tertiary enrollment effects are strongest among applicants with above-average baseline educational attainment.

Training lottery winners are 1.6 percentage points (107 percent) more likely to remain enrolled in college five years after training participation. Among male training applicants, tertiary enrollment effects concentrate on universities whereas among female applicants the effects concentrate on vocational college enrollment. These results demonstrate that vocational training skills complement and beget additional formal education skills of participants.

Training has formal education spillovers on participants' families. Family members of training participants are 1.3 percentage points (35 percent) more likely to enroll in tertiary education and are 1.7 percentage points (38 percent) more likely to remain enrolled in tertiary education five years after training than family members of non-participants. The educational spillovers could be due to informational externalities or due to educational inputs and resources that trainees may share with other family members. The pattern of formal education results on applicants and relatives is most consistent with the training-formal education complementarity arising from a combination of improved skills and changes in expectations.

Training participants are more likely to be employed in formal sector jobs three to eight years after random assignment. Among female applicants, training lottery winners are 5.8 percentage points (10 percent) more likely to be employed in formal sector jobs than losers; among male applicants, winners are 5.1 percentage points (7 percent) more likely to work in formal employment.

Training lottery winners also spend 13 percent more time than losers on formal sector employment three to five years after randomization. Training lottery winners, particularly females, earn at least 11 percent more in total formal sector earnings than losers between three and eight years after randomization.

Compared to the short-term employment results reported by AKM, these results indicate that the impacts of job training on both the extensive and intensive margin of formal employment are long-lasting—particularly among women—which is consistent with prior evidence on the persistent impact of vocational training (Krueger, 2003; Schochet et al., 2008; Hotz, 2006). The impacts of training on formal sector employment outcomes are not necessarily a result of the additional formal education investments of trainees since these are equally strong for training applicants with below-average baseline educational attainment among whom training does not increase the likelihood of enrolling or remaining in tertiary education.

The rest of the paper proceeds as follows. In Section 2, we describe the vocational training program in Colombia, prior evidence on its short term employment effects and related studies on the effects of vocational training for disadvantaged youth. In Section 3, we describe the administrative data sources we use for the analysis and how we linked them to the original applicant data. In Section 4, we present results on the impacts of vocational training on formal educational outcomes of participants and their families in the medium and long-term. In Section 5, we present impacts of vocational training on formal employment in the medium and long-term using social security records. In section 6, we estimate costs and benefits of the program and calculate internal rates of return separately for women and men. We conclude in Section 7.

II.   Program Background and Prior Evidence

A. Program Background

We study the long-term direct and spillover effects of Youth in Action (YIA), a vocational training program introduced in Colombia in the early 2000s. In 2005, the government agreed to use lotteries to assign oversubscribed training courses among interested applicants. This randomized design allowed AKM to evaluate the short-term impacts of training on employment outcomes of trainees one year after training.

YIA consisted of 3 months of classroom training and a 3-month apprenticeship in a job. Classroom training was provided by training institutions which had to be legally registered, show financial solvency and compete in a bidding process in order to be able to participate. The training institutions also had to guarantee apprenticeship in legally registered (formal sector) companies in order to be paid by the government for the provision of training. This pay-for-performance element of the program created incentives for training institutions to offer participants courses and skills that were relevant and in-demand by employers. Participants received a $2.20 daily stipend to cover for transportation and other expenses related to program participation. Female participants with children under age seven received a higher stipend of $3.00 daily to also help them cover for childcare expenses.

Applicants were randomly assigned to YIA as follows. Since training institutions were paid based on graduation of participants and their ability to place them in apprenticeships, they were inclined to select participants on the basis of who was likely to succeed. AKM suggested to training institutions that they select a list of eligible applicants for every course that included up to 50 percent more applicants than they had capacity to enroll in the course (e.g., a course with capacity for 30 individuals would pre-select a list of 45 individuals). Applicants from these lists

were randomly assigned to receive training or no training in that course. There was close to full compliance. Only 0.18 percent of those offered training turned it down and only 1.29 percent of those who were not offered training ended up receiving training.

AKM show that one year after YIA training, randomly selected female participants are 7 percentage points more likely to be in formal sector employment and earn 20 percent more than non-participants. Among males, YIA increased formal employment by 6 percentage points but had no effect on earnings. Attanasio et al. (2015) show that the earnings increases among participants are persistent. The employment and earnings impacts of YIA are among the largest in the job training literature that have been documented using a credible research design.

B. Prior Evidence on the Effects of Vocational Training for Disadvantaged Youth

Few studies in developing countries examine impacts of training using randomized designs that meet experimental evidence standards to make causal attributions. With one exception, all focus on short-term employment and earnings outcomes. No study to date uses a randomized research design to document long-term direct and spillover effects of training. Ours is the first paper to examine both formal education and employment outcomes of training participants and their families as far as eight years after random assignment.

Card et al. (2011) study a training program in the Dominican Republic (DR) and find small positive impacts on formal employment and earnings. Unlike AKM (2011), however, Card et al. (2011) lacks information on the original random assignment, and the analysis is therefore limited to compare trainees and non-trainees, which compromises the causal interpretation of the analysis. Ripani et al. (2015) find similarly small employment and earnings impacts of the DR program in the longer run.

Hicks et al. (2014) examine the impact of an informational and training voucher intervention in Kenya and find that vouchers increased training participation but did not increase participants' earnings. Hirshleifer et al. (2014) study the impact of a randomized training program in Turkey and find no positive effect on employment and earnings. That study does find a positive effect on employment quality one year after the program, but finds that the effect dissipates three years later.

In developed countries, only a handful of studies examine longer-term employment and earnings effects of training. Coach (1992) studies the impact of the National Supported Work Demonstration (NSWD) eight years after random assignment and finds no earnings impact of training. Krueger (2003) examines longer-term impacts of the Job Training Partnership Act (JTPA) using social security records and finds no impacts of JTPA the first year after participation and positive impacts of about 8 percent on youth earnings up to five years after participation. Schochet et al. (2008) study finds that Job Corps, a residential training program, increases earnings of participants by 12 percent up to 5 years after random assignment. Hotz, Imbens and Klerman (2006) examine short- and long-term impacts of GAIN, a welfare-to-work program in California, and find that the training components of the program were relatively more beneficial in the long-term than the labor force attachment components of the program.

While several studies examine the external effects of formal education and the effects of schooling on other's educational attainment and on voting behavior and criminal activity (Acemoglu and Angrist, 2000; Heckman and Lochner, 2000; Milligan et al., 2004; Moretti, 2004), no study to date examines external effects of vocational training. In this paper, we examine the impact of vocational training on the schooling and employment of family members of the applicants.

Examining long-term effects as well as potential spillovers of programs is important for appropriate welfare calculations. For example, AKM conducted a cost-benefit analysis of the YIA program based on the short-term evaluation. In order to conduct this cost-benefit analysis, AKM had to make assumptions about whether the impacts would persist or not. AKM considered two scenarios: one in which the earnings benefits persisted until the end of the participant's lifetime, and another one in which the earnings gains from training depreciated by 10 percent per year. AKMs cost-benefit analysis was positive under both scenarios for women but not for men. Attanasio et al. (2015) conduct a cost-benefit analysis using longer-term employment outcomes and also conclude that the program has a positive rate of return. In both cases however, the cost-benefit analysis neglects spillover effects, potentially leading to an under-estimate of the YIA program's social desirability.

III.   Data Description

A. Randomization Data

Our analysis sample consists of a random sample of training applicants from the 2005 YIA training cohort originally collected by AKM (2011). Among all 2005 training applicants, AKM (2011) collected baseline information for a sample stratified by initial treatment offer, city and sex. This baseline sample includes 2,066 applicants assigned to treatment and 2,287 assigned to control. AKM (2011) collected baseline data in January 2005, either before the beginning of the training program or during the first week of classes to minimize any influence of participation in the program on interviewees' responses.

As a result of stratification by sex, 54 percent of training applicants in the evaluation sample are female (Table 1). Applicants are, on average 21 years old at the time of application and twenty percent of them are married. Applicants at baseline have completed about 10 years of

schooling and around 52 percent report being employed at the time of application, although less than 10 percent are employed in the formal sector. Average tenure prior to training is between three and four months. Training applicants work, on average, 12 days per month (including zeros for those out of work) and 25 hours per week (including zeros for those out of work).

As a result of randomization, baseline characteristics of training applicants are, on average, balanced across lottery winners and losers. The only statistically significant differences in baseline characteristics among training lottery winners and losers are age, prior educational achievement and job tenure. Training lottery winners are, on average, slightly younger, have higher prior educational attainment and higher prior job tenure. However, overall, the characteristics of winners and losers are not statistically significantly different from each other.

We use the applicants' names and adult national identification numbers to be able to match the original baseline data from YIA with the four administrative datasets. In the training applicant list, all applicants report full their names and over 96 percent report a valid adult identification number.[2] There is no difference between winners and losers in the probability of having a valid adult identification number (Table 2).

We obtain information about applicant's other household members at baseline by matching the sample of training applicants to the 2005 SISBEN Census. Data from SISBEN is used to determine eligibility for government subsidies. The 2005 SISBEN Census covers about sixty percent of households in all of Colombia. We obtained the following identifying information for the relatives of YIA program applicants from the SISBEN Census: names, dates of birth, national identification numbers, and relationship to the applicants. We are able to match close to 88 percent of lottery winners and 88.5 percent of losers to the 2005 SISBEN Census and the difference is not statistically significant (Table 2).

_____

[2] We define a valid identification number as having the correct expected number of digits.

B. Administrative Education Records

To track applicants' and other household members' educational outcomes, we use two national administrative datasets. First, we use Colombia's secondary school graduation exam (known as ICFES for its Spanish acronym − Instituto Colombiano para el Fomento de la Educacion Superior) database. We use these data to determine whether applicants took the ICFES test, which we interpret as a proxy for secondary school graduation since taking the test is a graduation requirement. We match data through the exam-taking cohort of 2012, seven years after training program application. We match 34 percent of lottery winners and 31 percent of losers to the secondary graduation exam database and the difference is statistically significant (Table 2). Note that showing up in the secondary school graduation database is an outcome potentially affected by training treatment status.

The second educational administrative dataset is Colombia's Education Ministry's System for Prevention and Analysis of Dropouts in Institutions of Higher Education (known as SPADIES for its Spanish acronym, Sistema de Prevención y Análisis de la Deserción en Instituciones de Educación Superior) to track applicants and other household members through higher education, including enrollment and completion. The higher education database is an individual-level panel dataset that tracks close to 95 percent of higher education students from their first year to their degree receipt beginning in 1998. We obtained data through 2013, eight years after program application. The higher 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 ICFES entrance exam. Using these data,

11

we can construct outcomes characterizing students' trajectories throughout college (e.g., enrollment periods, retention, and graduation). We match 21 percent of winners and 17 percent of losers to the tertiary education database and the difference is statistically significant (Table 2). As with the secondary education database, showing up in the tertiary education database is an outcome potentially affected by training treatment status.

C. Social Security Records

To track formal employment and earnings outcomes, we match applicants to Colombia's social security records collected by the Ministry of Health and Social Protection, known as SISPRO for its Spanish acronym (i.e., Sistema Integral de Información de la Protección Social).[3] 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 2013 —between three and eight years after randomization. We use data only starting in 2008 since this was the year when SISPRO began to cover the universe of formal sector workers. We use SISPRO data to examine additional outcomes, including: extensive and intensive margins of formal sector employment and formal sector earnings. Seventy eight percent of winners and 71 percent of losers appear at some point in the social security database and the difference is statistically significant (Table 2). As with the educational datasets, showing up in the social security database is an outcome itself potentially affected by training treatment status.

---

[3] The SISPRO database only includes people who worked for employers that register their workers or self-employed workers who register themselves.

IV.   Impacts on Formal Education of Participants and their Families

In this section we describe training lottery impacts on secondary school completion of participants (subsection A), tertiary education enrollment and retention impacts on participants (subsection B), and secondary school completion (subsection C) and tertiary education enrollment and retention of family members (subsection D).

A. Impacts on Secondary School Completion of Participants

We define the secondary school completion outcome as having taken the secondary school graduation test after participation in the training program. Twenty seven percent of program applicants had completed secondary school prior to training and there are no differences in secondary school completion rates among training lottery winners and losers prior to training (Table 1).

About 5 percent of training lottery losers goes on to complete secondary school after training (after 2006). Training lottery winners are 1.4 percentage points (27 percent) more likely than losers to complete secondary school one year after training participation, a difference that is statistically significant at the 10 percent level. When we focus on secondary school completion between two and four years after training (after 2007 through 2009), we find no effects of training, suggesting that most of the effect on secondary school completion is concentrated immediately after training participation (Panel A, Table 3). We also do not find statistically significant effects of training participation on secondary school completion when separating by gender (Panels B and C, Table 3).

B. Impacts on Tertiary Education Outcomes of Participants

Training participation increases tertiary education access and retention. Eleven percent of training lottery losers enrolls in tertiary education after training participation. Training lottery

winners are 3.5 percentage points (31 percent) more likely than losers to enroll in tertiary education. Training effects on tertiary enrollment are concentrated in the first year (in 2006) and particularly the second year (in 2007) after program participation. In 2007, for instance, lottery winners are 1.2 percentage points (80 percent) more likely than losers to enroll (Panel A, Table 4).

Tertiary education enrollment effects of training are concentrated among applicants with prior years of completed education above the mean—essentially those who had completed secondary education prior to training participation. Among training applicants with above average prior educational attainment, training increases tertiary education enrollment by 3.9 percentage points, a 27 percent increase relative to the training lottery losers mean of 14 percent. As is the case in the full sample, the effects on tertiary enrollment among applicants with prior above mean educational attainment concentrate in the second year after training (initial enrollment in 2007, Panel B, Table 4). Among applicants with below mean prior educational attainment, the point estimates of training on enrollment in tertiary education are small in magnitude and never statistically significant (Panel C, Table 4).

Disaggregating tertiary education enrollment effects of training by the type of institution applicants register in, we find that two thirds of training lottery losers (11 percent) choose to enroll in universities—the more prestigious academic institutions in Colombia that grant professional undergraduate degrees—and one third (2.4 percent) choose to enroll in vocational colleges, which award technical or technological degrees. Training participation increases university enrollment by 2.5 percentage points (23 percent relative) relative to the loser's rate (Panel A, Table 5). Vocational college enrollment by 1 percentage point for the lottery winners

14

relative to the losers after YIA. The overall training effect on vocational college application is, however, only statistically significant at the 10 percent level.

Among males, all of the tertiary enrollment effect of training concentrates in university enrollment. Male lottery winners are 4.3 percentage points (37 percent) more likely than losers to enroll in universities. The training effect on vocational college enrollment for male applicants is small, negative and statistically insignificant.

Among females, the effect on tertiary enrollment concentrates in vocational colleges. One and one half percent of female training applicants enroll in vocational colleges. Female lottery winners are 2 percentage points (133 percent) more likely than losers to enroll in vocational colleges and this effect is significant at the 1 percent level. Among females, we find no tertiary enrollment effects on universities. For both males and females, winning the training lottery increases tertiary enrollment chiefly into private institutions. We find no effect of winning the lottery on the probability of enrolling in a public tertiary institution (Table 5).

Training also increases tertiary education retention (Table 6). About 1.5 percent of training lottery losers remains enrolled in an institution of tertiary education one, two, three, four and five years after training. Training lottery winners are 3.4 percentage points (200 percent) more likely to remain enrolled one year after training; they are also 2.8 percentage points (187 percent), 2.5 percentage points (192 percent), 2.6 percentage points (186 percent) and 1.6 percentage points (107 percent) more likely than training lottery losers to remain enrolled in tertiary education two, three, four and five years after training. The effects on retention up to four years after are significant at the 1 or 5 percent levels, but the effect five years after training is only statistically significant at the 10 percent level (Panel A, Table 6).

Tertiary education retention effects are similar among male and female training applicants. Due to the reduced sample size, however, retention effects estimated separately for females and males are generally only statistically significant at the 10 percent level up to four years after training participation (Panel B, Table 6).

Tertiary education retention effects of training, as is the case with enrollment effects, concentrate among applicants with above average prior years of schooling completed. Estimates of training on tertiary education retention among this group of applicants are larger and more precisely estimated than those in the full sample of applicants. Among applicants with prior education below the mean there is no impact of training on tertiary education retention (Panel C, Table 6).

Prior evidence uncovered by AKM indicates that one year after the lottery, training had increased the probability of paid employment by 7 percentage points and earnings by 20 percent among female applicants. Training, by contrast, had no labor market effects on male applicants. The earlier labor market results in combination with these formal education impacts of training suggest that the job training-formal education complementarity among males can be due to learning about ability, and/or improved skills due to training. Among females the complementarity may also arise from changes in expectations (for example, increased expected returns to a vocational college degree given that female lottery winners experience an earlier rise in formal sector employment) or from a relaxation of credit constraints due to higher earnings immediately after training.

C. Impacts on Secondary School Completion of Family Members

We also examine whether training participation had external or spillover effects on the educational outcomes of other members of participants' families. If training, for example, helps

ease educational credit constraints due to increased participant incomes and families share educational resources, training participation may also benefit other family members. In addition, there may be informational externalities as other family members may learn of the benefits and returns to education.

We find some evidence that training participation increases secondary school completion of other family members of male applicants (Panel B, Table 7). Family members of male winners are between 2.3 and 2.7 percentage points (20 to 25 percent) more likely to complete secondary schooling one and two years after training. These effects on other family members of male applicants are only statistically significant at the 10 percent level. By contrast, training does not increase secondary school completion rates among family members of female applicants (Panel C, Table 7).

D. Impacts on Tertiary Education Outcomes of Family Members

Training participation increased access to tertiary education and retention in tertiary education among other family members of participants, particularly of male participants (Table 8). Relative to the family members of male lottery losers, family members of male lottery winners are 1.3 percentage points (35 percent relative to the mean of 3.7 percent among members of lottery losers) more likely to enroll in public tertiary institutions. The effect of training on tertiary enrollment of other members of male applicants concentrates in universities. Relatives of male lottery winners are 1.8 percentage points (26 percent) more likely than relatives of lottery losers to enroll in universities (Panel B, Table 8). Training does not improve tertiary education access among relatives of female applicants (Panel C, Table 8).

Training also increases tertiary education retention among relatives of male applicants (Table 9). Relatives of male trainees are 2.2 percentage points (138 percent) more likely to

remain enrolled one year after training; they are also respectively 0.6 percentage points (9 percent), 2 percentage points (32 percent), 1.8 percentage points (33 percent) and 1.7 percentage points (38 percent) more likely than relative of male lottery losers to remain enrolled in tertiary education two, three, four and five years after training. The effect on retention two years after training is only statistically significant at the 10 percent level (Panel B, Table 9), but all other effects are significant at the 1 and 5 percent levels. Training does not improve tertiary education retention among relatives of female trainees (Panel C, Table 9).

Formal education impacts on trainees, as discussed above, could be due to learning about ability, improved skills due to training, changes in expectations and/or relaxation of credit constraints. Differences by gender in training impacts on formal education outcomes among other family members, however, cast doubt on the credit constraints channel. To the extent that family members share educational resources, the earlier rise in employment and earnings among female applicants found by AKM could have helped relax credit constraints, potentially improving formal education outcomes of relatives.

However, we only find formal education impacts among relatives of male trainees for whom training does not improve short-term earnings. The pattern of formal education results on applicants and relatives is most consistent with the training-formal education complementarity arising from a combination of improved skills and changes in expectations.

V.   Impacts on Formal Employment of Participants and Relatives

We examine the impact of YIA on participants' likelihood and duration of formal sector employment (subsections A and B, respectively) and formal earnings (subsection C). In subsection D we compare the short and long-term impacts of training on formal employment

outcomes of participants. In subsection E we describe training impacts on formal labor market outcomes of relatives.

A. Formal Sector Employment of Participants

Among training lottery losers, we match 67 percent to the Colombian government records of those paying payroll taxes, SISPRO, implying that 67 percent of losers ever show up in formal sector employment between 2008 and 2013. Training lottery winners are 5.0 percentage points (8 percent) more likely to ever appear in formal employment records during this period (Column 1 Panel A, Table 10).

Separately by gender, seventy-seven percent of male lottery losers ever show up in formal sector employment between 2008 and 2013. Male lottery winners are 5.0 percentage points (7 percent) more likely to ever appear in formal employment during the period (Column 1 Panel B, Table 10). Among female training applicants, 58 percent of lottery losers ever show up in formal sector employment during this period and lottery winners are 5.8 percentage points (10 percent) more likely to every appear in formal employment during this period (Column 1 Panel C, Table 10).

Analyzing participation by prior educational attainment, the effect of training on the extensive margin of formal sector participation is strongest among training applicants with below average education at baseline. Among training applicants with above average education, 72 percent of lottery losers ever show up in formal sector employment between 2008 and 2013. Lottery winners among this group of applicants are 4.0 percentage points more likely to ever show up in formal employment (Column 1 Panel D, Table 10). Among training applicants with below average educational attainment, 55 percent ever show up in formal sector employment

19

during 2008-2003 and lottery winners are 8.2 percentage points (18 percent) more likely to ever appear in formal sector employment during the period (Column 1 Panel E Table 10).

B. Intensity of Formal Sector Employment

Training lottery winners spend more time in formal sector employment than losers. For the period between 2008 and 2013, we construct an outcome variable that equals the total number of days the applicant spends in the formal sector. Training lottery losers spend, on average, 608 days in formal sector employment in between 2008 and 2013—three to eight years after the training lottery. Out of the 1,320 workdays over this six-year period (220 days per year) this implies that lottery losers spend about 46 percent of the time formally employed during the period. Relative to lottery losers, winners spend 78 additional days (13 percent) in formal sector employment in 2008-2013 (Column 2, Panel A of Table 10).

The effects on the duration of formal employment are larger for women than men, just like the effects on being employed in the formal sector. As shown in Column 2 of Panel C of Table 10, female lottery winners are employed 90 additional days compared to female non-winners between 3 and 8 years after random assignment. By contrast, Panel B shows that male lottery winners stay employed 74 more days in the formal sector.

The results also show that those with lower average education before the program benefit substantially more from training in terms of formal sector employment. Panel E of Table 10 shows the effect for those with below average education at baseline. For those with below average education prior to the lottery, training increases formal employment by 129 additional days. By contrast, the benefit in terms of additional days worked in the formal sector is 61 days for those with above average baseline education. Thus, women and those with less education to

begin with are the individuals that most benefit from the training lottery in terms of their formal employment in the medium- to long-run.

C.   Formal Sector Earnings

We focus on two formal sector earnings outcomes. The first is total formal earnings in 2008-2013, the period for which we observe social security records. For this outcome, we add all earnings reported to the system, including zeros in months that the applicant does not report working in the formal sector. The second outcome is daily earnings in 2008-2013, which we obtain by dividing total earnings in 2008-2013 by total formal employment days 2008-2013.[4]

These earnings outcomes are likely lower bounds of the effect of training on formal earnings for two reasons. First, training may have induced infra-marginal applicants into formal sector employment, those who in the absence of training would not be working in formal sector jobs or working at all. Second, training affects formal education and a higher fraction of training lottery winners than losers remains enrolled in tertiary education during the period in which we observe formal earnings data.

Training has a positive impact on formal earnings in the medium- and long-run. Training lottery winners earn $915 more than losers in formal sector earnings 2008-2013. Relative to the lottery losers' mean of $8,225, the estimate represents an 11 percent increase in total formal sector earnings (Column 3, Panel A, Table 10).

Lower bound estimates of the impact of training on formal sector earnings are strongest among female applicants and among applicants with below-average educational attainment. Among males, training participants earn $900 more than non-participants, but this difference is not statistically significant (Column 3, Panel B, Table 10). Among females, training participants earn $1,065 more than non-participants (Column 3, Panel C, Table 10). The strongest effects of

---

[4] All monetary amounts are expressed in 2013 US dollars.

training on earnings among female applicants is consistent with prior evidence on short-term earnings impacts of the program discussed in AKM.

Among applicants with above-average educational attainment, trainees earn in total formal sector earnings $751 more than non-trainees (Column 3, Panel D, Table 10), but this difference is not statistically significant. Moreover, this difference likely understates the long-term impact among applicants with above-average baseline educational attainment since among these, a higher fraction of trainees than non-trainees continues to be enrolled in tertiary education. Among applicants with below-average educational attainment, trainees earn in total formal sector earnings $1,649 more than non-trainees, a 36 percent increase relative to the non-trainees mean (Column 3, Panel E, Table 10).

Trainees also have higher daily earnings than non-trainees between three and eight years after randomization. Training lottery winners earn $0.57 more per day (6 percent) in formal sector employment than lottery losers (Column 4, Panel A, Table 10). As is the case with total formal sector earnings, training effects are strongest among female applicants and applicants with below-average baseline educational attainment. Among female applicants, trainees earn $0.69 (10 percent) more than non-trainees (Column 4, Panel B, Table 10). Among applicants with below-average educational attainment, trainees earn $1.1 (17 percent) more than non-trainees per day of formal sector employment between three and eight years after randomization (Column 4, Panel E, Table 10). Among males and among applicants with above-average baseline educational attainment, estimates are smaller and not statistically significant (Column 4, Panels B and D, Table 10).

D. Comparisons between Short- and Long-term Impacts

One of the benefits of being able to do this analysis having matched our original data to administrative records is that we can verify whether the short-term effects persist, dissipate or even grow over time.

Table 11 shows the short-term results using the data from AKM and the long-term results using the SISPRO data. We re-estimate our short-term effects with the survey data collected by AKM to make the outcome measures comparable with the way we measure formal employment and earnings with the administrative data. First, we estimate formal employment as jobs in which the employer makes pension contributions on behalf of the employee. In the original analysis, individuals where defined as being employed in the formal sector if their employers either made contributions to pensions or health.

In the administrative data it is not possible to distinguish if health contributions were made on someone's behalf through their employer or their spouse's or someone else's employer. For this reason, we use the more restrictive definition based on pension contributions. In addition, for the earnings data, we use daily earnings to make the results from the two datasets comparable.

Table 11 shows that the short-term impacts of training on formal employment persist and do not depreciate. As shown in columns 1 and 2, the likelihood of formal employment is the same in the short- and long-term for women and men. For those with baseline educational attainment below the average, the effect also persists over time (Column 4, Table 11). However, for those with above average education at baseline, the effect becomes smaller over time, showing evidence of depreciation. This may be because this is the group in which a higher fraction of lottery winners goes on attend and remains in tertiary education.

23

The earnings effects are also persistent for women and the less educated. Columns 5 and 8 show the same short- and long-term effects for women and for those with below average education. By contrast, the results show depreciation of earnings gains over the long term for men and for those with above average educational attainment at baseline. This could also well be because it is men and the more educated to begin with who end up going to university and staying in university. This may, thus, be a group who is both working and studying simultaneously and whose earnings capture employment in temporary jobs.

E. Labor Market Impacts for Relatives

Table 12 explores whether there are labor market impacts on the family members of training lottery winners. The impact on family members could be positive or negative. If those exposed to training refer family members to formal sectors jobs and if the additional education attained by family members of trainees, then the effects would be positive. However, as we saw, lottery winners are more likely to be employed in the formal sectors and receive higher earnings. Thus, the income effect due to higher family income from lottery winners' higher earnings may end up reducing their employment either on the extensive or intensive margin. Table 12 shows no effects on other family members' formal earnings, employment or days worked.

VI.   Cost-Benefit Analysis

In this section we estimate costs and benefits of the program and calculate internal rates of return separately for women and men. We assume a discount rate of 10 percent. Throughout the analysis, for each source of cost and revenue, we compute the Net Present Values (NPVs) converting into United States dollars using the year-specific exchange rate (annual average) 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

accrual 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 (2014) using US-CPI change between the analysis and the base year.

There are two sources of costs. The first is direct program costs. The second is the loss of job tenure associated with program participation. Direct program costs are $770. The loss of tenure was 1.43 months, which evaluated at baseline earnings is $64 (AKM).

There are two sources of program benefits: benefits on participants and benefits on non-participants. Benefits on participants are only those from increased formal sector earnings. This includes realized formal sector earnings (Tables 10 and 11) plus expected additional earnings due to increased tertiary education enrollment and retention, as documented in Sections IV.C and IV.D (see Tables 5 and 6). This is clearly a conservative assumption for a number of reasons. First, as documented in Sections IV.C and IV.D, the program has positive external effects on the educational attainment—particularly tertiary education enrollment—on a significant fraction of family members of lottery winners. Second, we do not account for the non-wage benefits of increased formal-sector employment. Chiefly amongst these are additional payroll taxes that boost the government's revenue. Third, we do not account for other potential benefits such as delayed childbearing, reduced program participation in other government programs (see Bettinger et al. 2014) or reduced criminal engagement.

We assume that women (and men) work continuously for 40 years in formal sector employment at the same intensity as that between 2008 and 2013. This implies that we assume that for 40 years female lottery losers on average spend 94 days per year in formal sector employment; female lottery winners spend 112 days; male lottery losers 153 days and male

25

lottery winners 167 days. We also do not account for potential welfare losses associated with the cost of raising program funds through the tax system.

To account for the additional formal earnings benefits of increased tertiary enrollment and retention among lottery winners, we assume a 10 percent return to an additional year of tertiary education and multiply these gains by the program impact on the 4-year tertiary education retention rate (Table 6).

Under these assumptions, we project earnings based on daily earnings from Tables 10 and 11. Note that the estimates in Table 11 indicate that earnings impacts for females grow over time (in absolute terms) while those for males depreciate over time. To project daily earnings for lottery losers, we run a separate regression by gender of daily earnings 2006-2013 on calendar year and project for the duration of the work life with an annual depreciation rate of 10 percent. To project the earnings of winners we add average 2006-2013 growth/depreciation rates of daily formal sector earnings impacts separately for females and males and the expected earnings benefits of increased tertiary enrollment and retention. Appendix Figure A shows the projected earnings profiles of lottery winners and losers, separately by gender.

The net present value of formal earnings for female lottery losers is $5,015; for female lottery winners is $6,973. Of these, $6,847 are direct earnings benefits and $116 are expected future gains due to increased tertiary retention of female winners. Among males, the net present value of formal earnings for lottery losers is $13,412; for male lottery winners is $16,283. Of these $15,892 are direct earnings benefits and $391 are expected future gains due to increased tertiary retention of male winners. The NPV of the earnings difference is $1,958 for females and $2,870 for males. These net present value calculations differ markedly with the short-term calculations in AKM because in the short-term male lottery winners had no discernible effects on

26

earnings, unlike the long-term effects we document in this paper. Note also that net revenue for males is larger than for females despite the depreciation of male formal earnings impacts. This is so because baseline daily earnings are larger considerable among males than among females (Appendix Figure A).

Benefits to non-participants are only those resulting from additional tertiary education enrollment and retention spillover effects to other family members of participants, which we similarly value using a 10 percent rate of return to an additional year of tertiary education from baseline earnings of relatives in Table 12. These additional educational benefits only accrue to the relatives of male lottery winners and not to the relatives of females (Table 9). We estimate that benefits due to additional education among the relatives of male applicants to be $24.

The program's total revenue is $1,958 for females and $2,894 for males. Revenue net of program costs is $1,124 for females and $2,060 for males. This calculation assumes that there are no negative spillovers to non-participants, for instance, in the form of job displacement effects (e.g. Crépon et al 2013). However, these potential negative spillovers to non-participants would need to be extremely large for the program not to be welfare enhancing. The implied internal rates of return are 19.1 percent for females and 30.1 percent for males.

VII. Conclusion

In this paper we analyze the long-term effects of a randomized vocational training program for disadvantaged youth in Colombia. Beyond the prior evidence of program effects based on short-term employment and earnings effects in AKM and on a recent long-term examination of employment and earnings effects in Attanasio et al. (2015), we demonstrate for the first time using a rigorous experimental research design that: a) vocational training skills complement, rather than substitute additional formal educational investments, particularly in

tertiary education, and b) training participation has spillover effects on participants' family members. The pattern of formal education results on applicants and relatives is most consistent with the training-formal education complementarity arising from a combination of improved skills and changes in expectations. We also demonstrate that the employment and earnings impacts of training participation in the Colombian program amplify rather than diminish over time, at least for the eight years after randomization that we can observe.

The vocational training-formal education complementarity is strongest among trainees with above-average baseline education, which underscores the idea that skill begets skill (Heckman, 2000). However, we also see very strong long-term effects on employment and earnings outcomes among applicants with below-average baseline education, among whom training did not affect tertiary education enrollment or retention after participation. This finding highlights the fact that training by itself—at least in the context of Colombia's YIA program—has large benefits for a particularly disadvantaged population group.

Taken together, the results suggest that a job training program such as Colombia's YIA that combines classroom and apprenticeship time and has pay-for-performance incentives built in is an attractive social investment and a potential avenue for social mobility.

We estimate costs and benefits of the program and calculate internal rates of return separately for women and men. We assume a discount rate of 10 percent. The NPV of the earnings difference is $1,958 for females and $2,870 for males. These net present value calculations differ markedly with the short-term calculations in AKM because in the short-term male lottery winners had no discernible effects on earnings, unlike the long-term effects we document in this paper. The program's total revenue is $1,958 for females and $2,894 for males.

Revenue net of program costs is $1,124 for females and $2,060 for males. The implied internal rates of return are 19.1 percent for females and 30.1 percent for males.

Welfare calculations of this and of similar programs that do not account for complementary educational investments, long-run effects on employment and earnings or potential positive spillover effects likely underestimate the program's social desirability.

References

Attanasio, Orazio, Adriana Kugler Costas Meghir. 2011. "Subsidizing Vocational Training for Disadvantaged Youth in Colombia: Evidence from a Randomized Trial." *American Economic Journal: Applied Economics,* 3 (July): 188-220.

Attanasio, Orazio, Arlen Guarin, Carlos Medina and Costas Meghir. 2015. "Long-term Impacts of Vouchers for Vocational Training: Experimental Evidence from Colombia," NBER working paper no. 21390.

Acemoglu, Daron and Joshua Angrist. 2000. "How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws," *NBER Macroeconomics Annual*, Vol. 15, pp. 9-59.

Arellano Manuel. 2003. *Panel Data Econometrics*. Oxford University Press.

Ashenfelter, Orley. 1978. "Estimating the Effects of Training Programs on Earnings," *Review of Economics and Statistics*, 60: 648-660.

Bettinger, Eric, Michael Kremer, Maurice Kugler, Carlos Medina, Christian Posso and Juan Saavedra. 2014. "Educational, Labor Market and Welfare Impacts of Scholarships for Private Secondary School: Evidence from Colombia." Unpublished manuscript.

Burghardt, John and Peter Schochet. 2001. "National Job Corps Study: Impact by Center Characteristics," Princeton: Mathematica Policy Research.

Card, David, Pablo Ibarrarán, Ferdinando Regalia, David Rosas, and Yuri Soares. 2011. "The Labor Market Impacts of Youth Training in the Dominican Republic," *Journal of Labor Economics*, 29(2): 267-300.

Card, David and Daniel Sullivan. 1988. "Measuring the Effect of Subsidized Training Programs on Movements In and Out of Employment," *Econometrica*, 56: 497-530.

Couch, K.A. 1992. "New Evidence on the Long-Term Effects of Employment Training Programs," *Journal of Labor Economics*, 10: 380-388.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot and Philippe Zamora. 2013. "Do Labor Market Policies have Displacement Effects? Evidence from a Clustered Randomized Experiment," *Quarterly Journal of Economics*, vol. 128(2), pages 531-580

Gill, Indermit, Claudio Montenegro and Dörte Dömeland. 2002. *Crafting Labor Policy: Techniques and Lessons from Latin America*. Washington, D.C.: The World Bank.

Heckman, James. 2000. "Policies to Foster Human Capital," *Research in Economics*, 54(1): 3-56.

Heckman, James and Alan Krueger. 2003. *Inequality in America: What Role for Human Capital Policies?* Cambridge, Mass.: MIT Press.

Heckman, James, and Lance Lochner. 2000. "Rethinking Education and Training Policy: Understanding the Sources of Skill formation in a Modern Economy." *Securing the Future*, pp. 47-86. *Nueva York: Russell Sage*

Heckman, James, Robert LaLonde and Jeffrey Smith. 1999. "The Economics and Econometrics of Active Labor Market Programs," in Oley Ashenfelter and David Card, eds. *Handbook of Labor Economics*, Vol. 3A, pp. 1865-2097.

Heckman, James, Hidehiko Ichimura, Jeffrey Smith, Petra Todd. 1998. "Characterizing Selection Bias Using Experimental Data," *Econometrica*, 66(5): 1017-1098.

Hicks, Joan Hamory, Michael Kremer, Isaac Mbiti and Edward Miguel. 2013. "Vocational Education in Kenya: Evidence from a Randomized Evaluation about Youth," Mimeo.

Hirshleifer, Sarojini, David McKenzie, Rita Almeida and Cristobal Ridao-Cano. 2013. "The Impact of Vocational Training on the Unemployed: Experimental Evidence from Turkey," IZA Working Paper No. 8059.

Hotz, V.J., G.W. Imbens and J.A. Klerman. 2006. "Evaluating the Differential Effects of Alternative Welfare-to-Work Training Components: A Reanalysis of the California GAIN Program," *Journal of Labor Economics*, 24: 521-566.

Krueger, Alan. 2003. "Inequality, Too Much of a Good Thing," in Ed. Heckman, James and Alan Krueger, *Inequality in America*, Princeton University Press: Princeton.

Kugler, Adriana. 2005. "Wage-Shifting Effects of Severance Payments Savings Accounts in Colombia," *Journal of Public Economics*, 89(2-3): 487-500.

Kugler, Adriana. 1999. "The Impact of Firing Costs on Turnover and Unemployment: Evidence from the Colombian Labor Market Reform," *International Tax and Public Finance Journal*, 6(3): 389-410.

LaLonde, Robert. 1986. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data," *American Economic Review*, 76(4): 604-620.

Manski, Charles. 1994. "The Selection Problem," in Advances in Econometrics, Sixth World Congress, Vol. 1, Ed. by C. Sims. Cambridge, U.K.: Cambridge University Press, 143– 170.

Milligan, Kevin, Enrico Moretti and Philip Oreopoulos. 2004. "Does Education Improve Citizenship? Evidence from the U.S. and the U.K.," *Journal of Public Economics*, 88(9-10), 2004.

Moretti, Enrico. 2004. "Estimating the Social Return to Higher Education: Evidence From Cross-Sectional and Longitudinal Data," *Journal of Econometrics*, 121(1-2).

Orr, Larry, Howard Bloom, Stephen Bell, Winston Lin, George Cave and Fred Doolittle. 1994. The National JTPA Study: Impacts, Benefits, and Costs of Title II-A. Bethesda, Md.: Abt Associates.

Ripani, Laura, Pablo Ibarraran, Jochan Kluve and David Rosas-Schady. 2015. "Experimental Evidence on the Long Term Impacts of a Youth Training Program," mimeo.

Schochet, P.Z., J. Burghardt and S. McConnell. 2008. "Does Job Corps Work? Impact Findings from the National Job Corps Study," *American Economic Review*, 98: 1864-1886.

Smith, Jeffrey and Petra Todd. 2001. "Reconciling Conflicting Evidence on the Performance of Propensity Score Matching Methods," *American Economic Review*, 91(2): 112-118.

Table 1. Randomization Balance (2005 Baseline data)

	Control mean / (sd) (1)	Treated mean / (sd) (2)	Difference mean / (se) (3)
Female	0.54	0.53	0.01
	(0.5)	(0.5)	(0.02)
Age	21.24	21.05	0.19**
	(2.03)	(2.06)	(0.06)
Married	0.2	0.18	0.02
	(0.4)	(0.39)	(0.01)
Years of Education	9.93	10.17	-0.23***
	(1.96)	(1.68)	(0.06)
Secondary School Graduation before 2006	0.26	0.28	-0.02
	(0.44)	(0.45)	(0.01)
Employed	0.51	0.54	-0.03
	(0.5)	(0.5)	(0.02)
Monthly wage and Salary earnings (zero if out of work, USD$)	55.5	57.7	-2.2
	(85.4)	(85.6)	(2.6)
Monthly self-employment earnings (zero if missing, USD$)	12.8	12	0.8
	(44.2)	(42.5)	(1.3)
Formal (zero if out of work)	0.09	0.10	0.01
	(0.29)	(0.30)	(0.01)
Contract (zero if out of work)	0.09	0.08	0.01
	(0.29)	(0.28)	(0.01)
Job Tenure (Months)	3.21	4.06	−0.85**
	(7.69)	(9.97)	(0.28)
Days worked by month (zero if out of work)	12.19	12.80	(0.61)
	(12.67)	(12.73)	(0.39)
Hours worked by week (zero if out of work)	25.33	26.88	(1.44)
	(28.29)	(98.11)	(0.87)
Observations	2,236	2,115	4,351

Notes: Table reports means and standard deviation (error) for control and treated groups, and its difference. Graduates from secondary school before 2006 were used in a placebo test. Wage and earnings expressed in 2011 USD (1 USD = 1848.17 Colombian Pesos). *** p<0.01, ** p<0.05, * p<0.1

Table 2. Match of Baseline Data with Administrative Databases

	Control mean / (sd) (1)	Treated mean / (sd) (2)	Difference mean / (se) (3)
Applicant with valid ID	0.96	0.97	-0.01
	(0.19)	(0.17)	(0.01)
Applicant matched with Sisben data	0.88	0.88	0.01
	(0.32)	(0.33)	(0.01)
Applicant graduated from secondary	0.31	0.34	-0.03*
	(0.46)	(0.47)	(0.01)
Applicant enrolled in Higher Education	0.17	0.21	-0.04***
	(0.37)	(0.41)	(0.01)
Applicant in formal labor market	0.71	0.78	-0.07***
	(0.45)	(0.42)	(0.01)
Observations	2,236	2,115	4,351

Notes: Table reports means and standard deviation (error) for control and treated groups, and its difference. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3. Secondary School Completion

	Completed Secondary School after…			
	2006	2007	2008	2009
	(1)	(2)	(3)	(4)
Panel A: Full Sample				
Treated	0.014*	0.008	0.007	0.006
	(0.007)	(0.006)	(0.006)	(0.005)
Control Mean	0.052	0.041	0.034	0.027
Panel B: Males				
Treated	0.009	0.003	0.005	0.003
	(0.01)	(0.009)	(0.009)	(0.008)
Control Mean	0.045	0.039	0.032	0.026
Panel C: Females				
Treated	0.015	0.009	0.006	0.005
	(0.01)	(0.009)	(0.008)	(0.007)
Control Mean	0.058	0.043	0.036	0.027

Notes: The table reports the difference in complete secondary school between treatment and control groups controling for baseline controls (gender, age, married, years of education, employed, salary, profit, formal employment, contract, tenure, days worked, hours worked) and training institution fixed effects. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4. Training Impacts on University and College Enrollment, by Year

	Initial Enrollment After YA (2006 and after)	Initial Enrollment in 2006	Initial Enrollment in 2007	Initial Enrollment in 2008	Initial Enrollment in 2009
	(1)	(2)	(3)	(4)	(5)
			Panel A. Full Sample		
Treated	0.035***	0.008*	0.012**	0.003	0.004
	(0.011)	(0.005)	(0.005)	(0.004)	(0.004)
Control Mean	0.110	0.017	0.015	0.013	0.014
		Panel B. Applicants with prior education above mean			
Treated	0.039***	0.010	0.016**	0.000	0.006
	(0.014)	(0.006)	(0.006)	(0.005)	(0.004)
Control mean	0.142	0.021	0.020	0.017	0.020
		Panel C. Applicants with prior education below mean			
Treated	0.021	0.005	0.003	0.009	0.004
	(0.014)	(0.005)	(0.006)	(0.006)	(0.004)
Control mean	0.040	0.006	0.006	0.004	0.001

Notes: The table reports the difference in enrollment in Tertiary Education by year between treatment and control groups controling for baseline controls (gender, age, married, years of education, employed, salary, profit, formal employment, contract, tenure, days worked, hours worked) and training institution fixed effects. Prior education mean is 10 years. *** p<0.01, ** p<0.05, * p<0.1

Table 5. Training Impacts on University and College Enrollment, by Type of Institution

	Enrolled in University (1)	Enrolled in Vocational College (2)	Enrolled in a Public Institution (3)	Enrolled in a Private Institution (4)	Enrolled in University (5)	Enrolled in Vocational College (6)	Enrolled in a Public Institution (7)	Enrolled in a Private Institution (8)	Enrolled in University (9)	Enrolled in Vocational College (10)	Enrolled in a Public Institution (11)	Enrolled in a Private Institution (12)
Panel A.		Full Sample				Male				Female		
Treated	0.025**	0.010*	0.005	0.031***	0.043**	(0.002)	0.000	0.041***	0.011	0.020***	0.010	0.021*
	(0.011)	(0.005)	(0.008)	(0.009)	(0.017)	(0.009)	(0.014)	(0.014)	(0.014)	(0.007)	(0.011)	(0.011)
Control Mean	0.110	0.024	0.070	0.064	0.116	0.035	0.081	0.070	0.105	0.015	0.060	0.059
Panel B.		Full Sample				Applicants with prior education above mean				Applicants with prior education below mean		
Treated	0.025**	0.010*	0.005	0.031***	0.032**	0.012*	0.004	0.040***	0.007	0.008	0.011	0.004
	(0.011)	(0.005)	(0.008)	(0.009)	(0.014)	(0.007)	(0.011)	(0.011)	(0.014)	(0.007)	(0.012)	(0.011)
Control Mean	0.110	0.024	0.070	0.064	0.139	0.032	0.088	0.083	0.046	0.007	0.030	0.023

Notes: The table reports the difference in get enrolled in Tertiary Education by type of Higher Education Institution (HEIs) between treatment and control groups controling for baseline controls (gender, age, married, years of education, employed, salary, profit, formal employment, contract, tenure, days worked, hours worked) and training institution fixed effects. Prior education mean is 10 years. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6. Training Impact on Higher Education Retention

		Continued Enrollment in Tertiary Education X years after training				
		One year	Two years	Three years	Four years	Five years
		(1)	(2)	(3)	(4)	(5)
Panel A: Full Sample	Treated	0.034***	0.028***	0.025**	0.026***	0.016*
		(0.011)	(0.011)	(0.010)	(0.009)	(0.009)
	Control mean	0.017	0.015	0.013	0.014	0.015
Panel B: Male	Treated	0.043**	0.030*	0.031*	0.029*	0.013
		(0.018)	(0.017)	(0.016)	(0.015)	(0.014)
	Control mean	0.021	0.014	0.016	0.019	0.020
Panel C: Female	Treated	0.027*	0.026*	0.022*	0.023*	0.017
		(0.015)	(0.014)	(0.013)	(0.012)	(0.011)
	Control mean	0.012	0.017	0.011	0.011	0.012
Panel D: Applicants with prior education above mean	Treated	0.043***	0.033**	0.032**	0.033***	0.024**
		(0.014)	(0.014)	(0.013)	(0.012)	(0.011)
	Control mean	0.053	0.052	0.046	0.041	0.034
Panel E: Applicants with prior education below mean	Treated	0.009	0.010	0.002	0.003	-0.008
		(0.015)	(0.014)	(0.013)	(0.012)	(0.011)
	Control mean	0.171	0.158	0.146	0.124	0.105

Notes: The table reports the difference in persistance in Tertiary Education between treatment and control groups controling for baseline controls (gender, age, married, years of education, employed, salary, profit, formal employment, contract, tenure, days worked, hours worked) and training institution fixed effects. Prior education mean is 10 years. Persistance is defined as continued enrollment of the student in the same Higher Education Institution and in the same program for X years. *** p<0.01, ** p<0.05, * p<0.1

Table 7. Training Impact on Relatives´ Secondary School Completion

	Relative completed Secondary School after…			
	2006	2007	2008	2009
	(1)	(2)	(3)	(4)
	Panel A: Full Relatives Sample			
Treated	0.012	0.009	0.004	0.000
	(0.010)	(0.009)	(0.007)	(0.006)
Control Mean	0.136	0.094	0.065	0.045
	Panel B: Relatives with Male Applicants			
Treated	0.027*	0.023*	0.008	0.002
	(0.016)	(0.014)	(0.011)	(0.010)
Control Mean	0.142	0.096	0.069	0.05
	Panel C: Relatives with Female Applicants			
Treated	0.007	0.005	0.005	0.002
	(0.015)	(0.012)	(0.010)	(0.009)
Control Mean	0.129	0.089	0.6	0.038

Notes: The table reports the difference in complete secondary school for relatives between applicants´ treatment and control groups controling for   applicant´s baseline controls (gender, age, married, years of education, employed, salary, profit, formal employment, contract, tenure, days worked, hours worked), and training institution fixed effects. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 8. Training Impact on Relatives' Enrollment in Higher Education, by Year and Type of Institution

	Initial Enrollment After YA (2006 and after) (1)	Initial Enrollment in 2006 (2)	Initial Enrollment in 2007 (3)	Enrolled in a Public HEI (4)	Enrolled in a University program (5)	Enrolled in a Public Institution (6)	Enrolled in a Private Institution (7)
			Panel A. Full Sample				
Treated	0.005	0.002*	0.002	0.008**	0.008*	0.005	-0.007
	(0.004)	(0.001)	(0.002)	(0.003)	(0.005)	(0.005)	(0.005)
Control Mean	0.06	0.006	0.009	0.037	0.069	0.037	0.038
			Panel B. Relatives with Male Applicants				
Treated	0.010	0.002	0.004*	0.013**	0.018**	0.022***	0.008
	(0.007)	(0.002)	(0.002)	(0.006)	(0.008)	(0.008)	(0.007)
Control mean	0.059	0.007	0.008	0.037	0.069	0.037	0.036
			Panel C. Relatives with Female Applicants				
Treated	0.000	0.002	-0.000	0.003	0.001	-0.006	-0.017**
	(0.005)	(0.002)	(0.002)	(0.004)	(0.006)	(0.007)	(0.007)
Control mean	0.06	0.006	0.009	0.036	0.07	0.036	0.040

Notes: The table reports the difference in enrollment (by 1) enrollment year, 2) type of Higher Education Institution (HEI) where relative get enrolled, and 3) the level of the program where relative get enrolled) in Higher Education for relatives between applicants´ treatment and control groups controling applicant´s baseline controls (gender, age, married, years of education, employed, salary, profit, formal employment, contract, tenure, days worked, hours worked) and training institution fixed effects. Prior education mean is 10 years. *** p<0.01, ** p<0.05, * p<0.1

Table 9. Training Impact on Higher Education Retention of Relatives

| | | Relative continued Enrollment in Tertiary Education X years after training | | | | |
| | | One year | Two years | Three years | Four years | Five years |
		(1)	(2)	(3)	(4)	(5)
Panel A: Full Relatives Sample	Treated	0.009**	0.005**	0.008*	0.006	0.006*
		(0.005)	(0.002)	(0.004)	(0.004)	(0.004)
	Control mean	0.017	0.069	0.063	0.055	0.046
Panel B: Relatives with Male Applicants	Treated	0.022***	0.006*	0.020***	0.018**	0.017***
		(0.008)	(0.003)	(0.007)	(0.007)	(0.006)
	Control mean	0.016	0.067	0.062	0.054	0.045
Panel C: Relatives with Female Applicants	Treated	-0.001	0.005	-0.002	-0.004	-0.001
		(0.006)	(0.003)	(0.006)	(0.005)	(0.005)
	Control mean	0.019	0.071	0.064	0.056	0.056
Panel D: Relatives with Applicants with prior education above mean	Treated	0.009*	0.007**	0.008	0.006	0.008*
		(0.006)	(0.003)	(0.005)	(0.005)	(0.005)
	Control mean	0.018	0.073	0.066	0.056	0.045
Panel E: Relatives with Applicants with prior education below mean	Treated	0.007	0.002	0.005	0.003	0.003
		(0.008)	(0.004)	(0.008)	(0.007)	(0.007)
	Control mean	0.016	0.062	0.057	0.052	0.052

Notes: The table reports the difference in persistance in Higher Education for relatives between applicants´ treatment and control groups controling for applicant´s baseline controls (gender, age, married, years of education, employed, salary, profit, formal employment, contract, tenure, days worked, hours worked) and training institution fixed effects. Prior education mean for applicants is 10 years. Persistance is defined as the continue enrollment of the student in the same Higher Education Institution and in the same program for X years. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 10. Training Impact on Labor Market Outcomes

	Ever in Formal Sector Employment 2008-2013	Formal Sector Tenure (Days) 2008-2013	Total Formal Sector Earnings 2008-2013	Daily Earnings 2008-2013 (Total Earnings/Formal Sector Tenure in Days)
	Panel A: All Applicants			
Treated	0.050***	77.360***	914.544***	0.568***
	(0.014)	(21.259)	(329.628)	(0.194)
Control Mean	0.663	599.275	8073.746	8.286
N	4,194	4,194	4,194	4,194
	Panel B: Male Applicants			
Treated	0.050***	73.804**	900.298	0.542*
	(0.019)	(34.683)	(563.240)	(0.311)
Control Mean	0.774	763.063	10632.654	10.069
N	1,874	1,874	1,874	1,874
	Panel C: Female Applicants			
Treated	0.058***	90.208***	1,065.175***	0.690***
	(0.020)	(27.481)	(399.119)	(0.250)
Control Mean	0.577	471.582	6078.773	6.896
N	2,320	2,320	2,320	2,320
	Panel D: Applicants with Above Average Baseline Education			
Treated	0.040**	60.562**	751.286*	0.435*
	(0.016)	(26.493)	(416.864)	(0.236)
Control Mean	0.712	707.321	9,717.211	9.152
N	2,971	2,971	2,971	2,971
	Panel E: Applicants with Below Average Baseline Education			
Treated	0.082***	129.449***	1,648.978***	1.080***
	(0.029)	(36.645)	(537.303)	(0.368)
Control Mean	0.553	360.753	4,456.741	6.354
N	1,207	1,207	1,207	1,207

Notes: The table reports the difference in each variable between treatment and control groups controling for baseline data (gender, age, married, years of education, employed, salary, profit, formal employment, contract, tenure, days worked, hours worked) and by course fixed effects. Earnings expressed in 2013 US Dollars (1 USD= 1,869.1 Colombian pesos). *** p<0.01, ** p<0.05, * p<0.1

Table 11. Difference between Short- and Long-term Impacts

		Formal				Daily Earnings			
		Females	Males	Education above mean	Education below mean	Females	Males	Education above mean	Education below mean
Long Run	Treated	0.046**	0.054**	0.030*	0.090***	0.807***	0.814*	0.643**	1.213***
		(0.020)	(0.023)	(0.018)	(0.030)	(0.308)	(0.428)	(0.311)	(0.464)
	Control Mean	0.385	0.564	0.521	0.336	5.128	8.451	7.561	4.409
Short Run	Treated	0.057***	0.077***	0.061***	0.076***	0.975***	1.665***	1.286***	1.057***
		(0.016)	(0.021)	(0.016)	(0.022)	(0.258)	(0.384)	(0.274)	(0.359)
	Control Mean	0.147	0.249	0.227	0.112	2.208	3.934	3.488	1.805
	Difference	0.011	0.023	0.031**	0.014	0.168	0.851**	0.643***	0.156

Notes: The table reports the absolute difference in each sample between long run and short run impacts among treatment and control groups controling for baseline data (gender, age, married, years of education, employed, salary, profit, formal employment, contract, tenure, days worked, hours worked) and by course (Ecap) fixed effects. Formal expressed as pension fund contribution. Earnings expressed in 2013 US Dollars (1 USD= 1,869.1 Colombian pesos). Short run as baseline data (2004), Long run as 2013 data. *** p<0.01, ** p<0.05, * p<0.1

Table 12. Training Impacts on Labor Market Outcomes of Relatives

	Ever in Formal Sector Employment 2008-2013	Formal Sector Tenure (Days) 2008-2013	Total Formal Sector Earnings 2008-2013	Daily Earnings 2008-2013 (Total Earnings/Formal Sector Tenure in Days)
	Panel A: All Relatives Sample			
Treated	0.005	0.858	-89.656	-0.102
	(0.007)	(9.549)	(214.898)	(0.204)
Control Mean	0.245	268.384	4,778.901	4.466
N	12,717	12,717	12,717	12,717
	Panel B: Relatives from Male Applicants			
Treated	-0.002	0.295	206.179	-0.188
	(0.010)	(13.647)	(317.294)	(0.296)
Control Mean	0.226	240.579	4,186.704	4.155
N	6,118	6,118	6,118	6,118
	Panel C: Relatives from Female Applicants			
Treated	0.009	-1.614	-415.693	-0.083
	(0.010)	(13.797)	(314.797)	(0.295)
Control Mean	0.261	292.819	5,299.357	4.740
N	6,599	6,599	6,599	6,599
	Panel D: Relatives from Applicants with Above Average Baseline Education			
Treated	0.004	2.480	-239.728	-0.170
	(0.009)	(11.721)	(257.022)	(0.237)
Control Mean	0.254	288.285	5,285.111	4.696
N	9,054	9,054	9,054	9,054
	Panel E: Relatives from Applicants with Below Average Baseline Education			
Treated	0.005	-1.757	282.518	-0.001
	(0.014)	(17.231)	(377.498)	(0.370)
Control Mean	0.226	225.135	3,675.901	3.990
N	3,593	3,593	3,593	3,593

Notes: The table reports the difference for relatives in each variable between treatment and control groups from applicants controling for baseline data (gender, age, married, years of education, employed, salary, profit, formal employment, contract, tenure, days worked, hours worked) and by course (Ecap) fixed effects. Earnings expressed in 2013 US Dollars (1 USD= 1,869.1 Colombian pesos). *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix Figure A. Projected earnings

Projected Earnings of Participants
By Gender and Lottery Status