

# Starting on the right track? The effects of first job experience on short and long term labor market outcomes

Lucila Berniell\*

Dolores de la Mata<sup>†‡</sup>

January 15, 2018

## Abstract

For young job seekers barriers to labor market entry are high, especially in developing countries where information frictions are large. Can first job experience impact such barriers and have perdurable effects? This paper exploits a large-scale youth employment program in Argentina that randomly allocated 12-months wage subsidies to investigate what were the effects of (relatively) high quality entry-level jobs on labor market outcomes of the youth. Short and long term outcomes are measured with data gathered from a follow-up survey (short run) and from administrative registries (short and long run). Working in a formal firm caused large short and long run gains in the probability of formal employment, as well as a fall in unemployment. The random assignment design also allows to implement a saturation approach to measure displacement effects, which, if anything, we found to be positive over not beneficiaries. We explore alternative mechanisms that could produce all these impacts of real world first job experience, and we find evidence favoring a reduction in informational barriers over alternative explanations, like on-the-job skills development.

---

\*CAF-Latin American Development Bank, Research Department.

<sup>†</sup>CAF-Latin American Development Bank, Research Department.

<sup>‡</sup>We are grateful to María Laura Alzúa, Guillermo Cruces, Claudia Martínez, Daniel Ortega, Pablo Sanguinetti, Martín Rossi, Santiago Tobón, and Sergio Urzúa for helpful comments and suggestions and to Diego Jorrat and Federico Juncosa for outstanding research assistance. We thank the *Agencia de Promoción del Empleo y Formación Profesional de la Provincia de Córdoba* (Argentina) for granting access to data and for the support provided by Alejandro Pizarro and Marcel Peralta. We also thank José Anchorena, Victoria Castillo and Moira Ohaco from *Ministerio de Trabajo, Empleo y Seguridad Social* (Argentina) for their help to access administrative data about formal employment trajectories. This project would not have been possible without the constant support and financial aid of the Economic and Social Research Department at CAF-Latin American Development Bank.

# 1 Introduction

Unemployment rates are usually two to three times higher for young than for adult workers, a general characteristic in labor markets of both developed and developing economies. In developing countries the presence of the informal sector imposes additional problems to young job seekers. For instance, in Latin America the share of informal salaried workers is about 66% larger for young (16 to 25 years old) compared to older workers.<sup>1</sup> That is, in countries with large informal sectors young job seekers face lower quality first job opportunities. One key question we address in this paper is whether the quality of entry-level jobs can have lasting effects on future labor market trajectories and, if so, what mechanisms can explain those effects.<sup>2</sup>

The reasons behind these worse labor outcomes for the youth can be grouped in two hypotheses: human capital deficiencies (low levels of the basic cognitive, technical or socio-emotional skills required by potential employers) and informational barriers to labor market entry (Palais, 2014). In particular, informational barriers could be critical for the youth: employers may not have sufficient incentives to hire inexperienced workers whose abilities are uncertain. In highly informal economies although some young workers may have some work experience, most of the times this experience took place in the informal sector and for this reason is not easy to certificate, or in case of being certified the information value it conveys is relatively low. Other types of informational barriers specially affecting the youth are the lack of knowledge about how to conduct an efficient job search, low expectations and misperception about labor market conditions, or lack of social contacts that can serve as referrals. In view of these problems, many governments have implemented different active labor market policies targeting the youth population.<sup>3</sup>

To investigate what are the effects of a relatively high quality first job experience on la-

---

<sup>1</sup>According to official monthly or quarterly labor force surveys for ten Latin American countries (Argentina, Brazil, Chile, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Mexico, and Uruguay), the share of informal salaried workers (not covered by social security) is 51% for individuals aged 16 to 25 years old, while it is about 30% for older workers (LABLAC-CEDLAS and The World Bank, 2017).

<sup>2</sup>In this sense our work is connected to the literature showing that analyzes how the way in which young individuals entry the labor market can have scarring effects (Oreopoulos et al., 2012; Kroft et al., 2013; Kawaguchi and Muraio, 2014; Eriksson and Rooth, 2014; Bell et al., 2015; Altonji et al., 2016).

<sup>3</sup>For instance, Card et al. (2017) survey over 200 recent studies of the impacts of active labor market programs, many of which are targeted to young job seekers and typically combine training and short-term internships for disadvantaged youth.

bor market prospects of the youth we study the *Programa Primer Paso* (PPP), a large-scale program in the second largest province in Argentina (Córdoba), aimed at improving employability of individuals between 16 and 25 years old by means of granting them a first formal job opportunity. The program provides a large wage subsidy –equivalent to 90% of the hourly minimum wage– along a period of 12 months during which beneficiaries work part-time in formally registered firms that have previously accepted them as potential employees. That is, an applicant to the program is a match formed by a young job seeker and a firm trying to cover a vacancy, whom have both agreed upon jointly apply to the program. The selection of the matches that finally can receive the benefit is done through a public lottery, because demand largely exceeds the annual number of available subsidies. This feature of the program creates the natural experimental setting we use to estimate the causal effects of the program.

We measure the impacts of PPP on several dimensions related to labor market outcomes, among which there are employment, formal employment, earnings, and skills. We use two main sources of data: administrative records (program application and monitoring registries, as well as social security registries that cover formal employment status and salaries), and a follow-up survey conducted over a representative sample of applicants 12 months after the end of the program.<sup>4</sup> The use of administrative records allows us to follow formal labor employment and wages trajectories of all applicants over time, and the large size of the experiment provides with enough power to detect the effects of interest. The field survey was designed to measure several labor market outcomes that are not observable from administrative registries, but are very relevant to explore potential channels through which the program operates. In this survey we gathered information from around 1,000 individuals for whom we measure employment status, perceptions and expectations regarding the labor market, job search strategies, and skills (cognitive and non-cognitive).

Our results indicate that the PPP causes large gains in the probability of formal employment for the youth, both in the short run –12 months after the program concluded– and in the long

---

<sup>4</sup>We also collected and used qualitative data resulting from focus groups with beneficiaries and in-depth interviews to key stakeholders of the program, as well as census tract data matched to program application administrative records. While qualitative data served as a guide for the design of the survey and for the interpretation of some quantitative results, census tract data was used to enrich the socioeconomic conditions of the neighborhood or city of residence of the youth eligible to the program.

run -4.5 years after the program started, and we find that these effects are larger for women. Additionally, we show that 12 months after finishing the program there are no statistically differences in labor force participation decisions for treated and control individuals, but there is a lower unemployment rate (10% reduction) for beneficiaries as compared to not beneficiaries. We also explore the causal effect of the program on formal salaries and we construct bounds, following Attanasio et al. (2011), from which we cannot rule out a zero wage effect. If we think of wages as proxies of productivity (skills), this evidence, together with the zero impact on psychometric tests of cognitive and non cognitive skills gathered in the follow-up survey, says that it is not very likely that the better employment outcomes observed for treated individuals arise from a positive impact of the program on the human capital accumulation of beneficiaries. On the contrary we explore heterogeneous effects of the program on participants with low and high cognitive skills, and we show that those effects are consistent with a signaling hypothesis, that is, with an strong reduction on informational barriers as a key determinant of the positive impacts driven by the program.

To the best of our knowledge, this paper is the first to isolate the causal impacts of a (relatively) high quality entry-level job, and it is also among the first ones to present causal evidence of real world on-the-job training and job market experience, distinct to the literature that analyzes the effects of in-class training<sup>5</sup> or the effects of in-class training combined with short term internships.<sup>6</sup> One exception is Gelber et al. (2015), which evaluates a summer

---

<sup>5</sup>Previous research has shown mixed results (low to zero) effects of programs only offering in-class training to develop skills for the labor market. This is especially true in developed economies, and although for developing countries the impacts seem to be larger, they are usually short-lived (Kluve et al., 2014).

<sup>6</sup>Due to the limited effectiveness of training programs, most active labor market programs for the youth, both in developed and developing countries, now offer a combination of training with real world practices. Although the evidence for these type of program is scarce for developing countries, there are two programs that have been rigorously evaluated in Latin America: *Juventud y Empleo* (JE, Dominican Republic) and *Jóvenes en Acción* (JA, Colombia) (Card et al., 2011; Attanasio et al., 2011; Ibarrarán et al., 2014; Acevedo et al., 2015). The first group of evaluations of the program JE found no overall effects on employment rates, but, conditional on being employed, the effects are positive on formality –mainly for males– and on wages (Card et al., 2011; Ibarrarán et al., 2014). In the long run the effects on formal employment for males remain positive. However, a group of evaluations of a slightly different version of JE, which includes a treatment arm of training in soft skills (Acevedo et al. (2015) and Acevedo et al. (2017)), found sizable short term employment and wages gains for women but losses for men, although these effects dissipate in the longer run. Results are mixed regarding the effects of JE on non-cognitive skills of participants. While Ibarrarán et al. (2014) found positive effects on these skills, Acevedo et al. (2015) found that the program reduces self-esteem of males in the longer run. Additionally, the program reduces teenage pregnancy and increases youth expectations about the future (Ibarrarán et al., 2014). In the case of the JA Program, Attanasio et al. (2011) find that in the short run the program raises earnings as well as the probability of employment for women, but none of these outcomes are significantly affected for

internship program in the state of New York.<sup>7</sup>

Additionally, this paper addresses several concerns raised by recent reviews of the effectiveness of active labor market policies for the youth (e.g., McKenzie (2017) and Kluve et al. (2017)). In particular, the program we analyze is a large scale program, which assures that the sample sizes are large enough to detect the results of interest. For instance, sample sizes in this natural experiment allow us to test for heterogeneous impacts across sub-population groups (by gender, educational status, and previously accumulated skills). Also, since we use administrative data we can tackle some other concerns, including attrition, the possibility of looking at longer term outcomes, and also doing cost effectiveness analysis. Very importantly, our setup allows us to investigate to what extent a particular type of general equilibrium effect (displacement) is a consequence of this type of active labor market policy. To address this issue, we estimate spillovers of the PPP over control individuals using a saturation approach that, as in the spirit of (Crépon et al., 2013), exploits the as-good-as-random variability in the share of treated individuals across more than 280 cities, which is possible to be done thanks to a specific feature of the random assignment process of the program. We do not find evidence supporting this type of spillover effect.

Both results, the positive impacts in employment outcomes driven by a reduction in informational barriers as well as the absence of displacement effects, are compatible with information frictions, which are common in most developing countries. That is, the evidence in this paper supports that first job opportunities in developing countries can be stepping stones not so much for their potential to accumulate human capital but for their capacity to improve the production and use of information that is useful to improve the efficiency with which labor markets operate.

---

men. Additionally, the authors find that the program has a significant impact on formality, for both men and women. These effects remain positive in the long run only for females (Attanasio et al., 2015). Another similar, but of smaller scale, program is analyzed in Alzua et al. (2016), which using administrative data finds a positive effect on formal employment in the short run that vanishes in the longer run.

<sup>7</sup>In particular, that paper finds that participation in the internship program increases earnings and employment in the year of the program, while it decreases earnings in the three years following participation and has zero effects afterwards.

## 2 Description of the program and data

In the 2000's in Argentina, the unemployment rate among the youth was around three times larger than for adults and the share of informal salaried workers was around 75% larger for young workers.<sup>8</sup> In this paper we study the 2012 edition of the *Programa Primer Paso* (PPP 2012). In that year, unemployment rates in Argentina reached 18.2% and 5.4% while the shares of informal salaried workers were 59% and 29% for young and for adult workers, respectively.

The PPP is a program administered by the *Agencia de Promoción del Empleo y Formación Profesional*, a ministry-level agency in the Province of Córdoba (Argentina), and it is aimed at improving the employability of individuals aged 16 to 25 years old by means of providing a first job experience in a formal firm.<sup>9</sup> The PPP basically operates as an internship program, which gives no formal training other than the on-the-job training to young job seekers, and which provides a wage subsidy in the form of a monthly payment to interns.<sup>10</sup> The magnitude of the subsidy is large, as in 2012 it represented around 90% of the hourly legal minimum wage. The number of beneficiaries in 2012 were around 7,300 (see Figure 1).<sup>11</sup>

The PPP was first launched in 1999 and it remained operative on an annual basis until 2007. During those first editions the program selected its beneficiaries in a first-come-first-served basis. The PPP was discontinued between 2008 and 2011 and reappeared in 2012, and since then the selection of beneficiaries is done through a lottery.

The theoretical coverage of the program can be computed comparing the number of benefits granted to the number of young individuals who are eligible to apply to the program, which is obtained from national household survey data (*Encuesta Permanente de Hogares*, EPH). The eligibility criteria only exclude young workers who have been formally registered during at least one month in the 6 months prior to the deadline for application (May 2012 for PPP 2012).

---

<sup>8</sup>From 2003 to 2012, the average unemployment rate for individuals aged 15-24 years old was 22.3% while for older individuals it was 6.9%. In the same period the average share of unregistered (informal) salaried workers was 62% among young and 35% among adults (LABLAC-CEDLAS and The World Bank, 2017).

<sup>9</sup>A formal firm is defined as a firm that have is registered

<sup>10</sup>Firms can voluntarily supplement this pay, and the government pays for an insurance covering job risks.

<sup>11</sup>Indeed, the number of beneficiaries under the PPP 2012 was 10,000, but around 2,700 benefits were given under modalities that are not of the interest of this paper (e.g., special long-term contracts and benefits for disable workers which were not randomly assigned). In November 2014, the Congress of Córdoba passed the "PPP law" (*Ley Provincial* 10236) that enacted the PPP as a permanent labor market policy in that Province. From 2015 onwards the PPP is mandated to cover at least 15,000 beneficiaries a year.

According to this criteria, the theoretical coverage goes from around 2% when computed over all young individuals who were not formally employed, to around 5% when computed over active but not formally employed young people.<sup>12</sup> For the case of firms, eligibility establishes that only formally registered firms –those that pay national taxes– with at least one registered employee –who are paid in accordance to social security regulations– can form a match with a young individual. In the year 2012 there were about 22,800 applications (see Figure 1). To apply to the PPP a young individual need to present his or her application with the endorsement of a firm that is willing to hire him or her in case of resulting beneficiary in the lottery.

The lottery is public and takes place every year around mid of May in the *Lotería de la Provincia de Córdoba*. In the random assignment process there are two different types of quotas. The first one requires that the shares of beneficiaries in each one of the sixteen eligible counties mimic population shares.<sup>13</sup> The second one restricts the number of PPP beneficiaries in each firm according to its size, measured by the number of formally registered employees. These two quotas affect the probability that each application is assigned to the treatment group, and for this reason a flexible form defining these quotas will be included as covariates in all regressions (see Section 3).

## 2.1 The data and the sample

Our data come from several sources. The first source is administrative and includes all the applications submitted to the PPP 2012. For each application we have baseline characteristics provided by applicants in the application form. For individuals, these baseline characteristics are age, gender, educational attainment, marital status, address and contact information, and whether the applicant has a child by the date he or she submitted the application to the program. We matched these baseline characteristics to information coming from the official population

---

<sup>12</sup>There are some other formal requirements that young individuals need to fulfill to be eligible: they cannot be beneficiaries of other national or local employment or social programs (except for the largest national CCT program, *Asignación Universal por Hijo*), and they cannot be related to the owner of the firm together with which they form the match of the application. All these requirements are cross-checked by means of administrative data, prior to the lottery takes place.

<sup>13</sup>We keep out from our analysis ten counties that account for a very small share of population and in which the PPP takes a very different form, since it allows to spend the full year of internship in informal firms.

Census<sup>14</sup>, from which we obtain unemployment, poverty, and labor informality rates at the neighborhood –or city– level for each young job seeker.<sup>15</sup> The characteristics of the firm in the application form are its size (number of formally registered employees), sector of activity and location. For each application we additionally know whether it was submitted in paper format or via internet, whether the match was selected in the lottery, whether the individual finally took the benefit and the number of months it was active in the program. Our sample consists of 22,776 applications submitted by 21,939 individuals who were matched with 10,408 firms.<sup>16</sup> The number of benefits that were randomly assigned were 7305 (see Figure 1).

To analyze formal labor market trajectories, we match the above mentioned individual level data to formal employment administrative records from the SIPA (*Sistema Integrado Previsional Argentino*) by means of the national identification number, which univocally identifies individuals. SIPA is an employer-employee matched dataset setup jointly by the social security administration, ANSES (*Administración Nacional de Seguridad Social*) and the national tax authority (*Administración Federal de Ingresos Públicos*, AFIP), and which records registered worker’s earnings and employment status. This database allows us to construct variables of formal employment status and earnings on a monthly basis during a period of 60 months, from December 2011 to November 2016 (6 pre-treatment, 12 during treatment, and 42 post-treatment months). Additionally, using firms’ tax identification number we are able to determine whether each individual is employed in the same firm where he or she applied to.<sup>17</sup>

Finally, to enrich the analysis of the short run effects of the program we conducted two types of field work, qualitative –focus groups and in-depth interviews– and a follow-up survey. The survey took place by the end of May 2014<sup>18</sup> and covered a random and representative subsample of 1018 beneficiaries and non-beneficiaries residing in the City of Córdoba.<sup>19</sup> In that

---

<sup>14</sup>*Censo Provincial de Población y Vivienda*, which was the pilot for the 2010 National Census and was conducted in 2008.

<sup>15</sup>Information at the neighborhood level is only available for individuals living in the City of Córdoba.

<sup>16</sup>Each individual could match with more than one firm, and this is why the number of applications and of individuals differ.

<sup>17</sup>It is worth noticing that firms were not required to formally register PPP beneficiaries as formal employees during the internship period, hence not all beneficiaries appear as registered workers in the SIPA database during the 12 months of duration of the PPP 2012.

<sup>18</sup>The survey was conducted on the 17th, 24th and 31st of May 2014, at the facilities of the *Universidad Nacional de Córdoba*.

<sup>19</sup>The objective sample size, derived from power calculations, was 1000 individuals.



survey we gathered socioeconomic information of individuals and their families, and a large set of labor market-related outcomes, including measures of cognitive and non-cognitive skills (see Figure 1). The survey was computer-based and all respondents were contacted by phone and invited to participate in the survey. Survey respondents received a stipend equivalent to 12 USD.<sup>20</sup>

## 2.2 Participants characteristics

Table 1 provides descriptive statistics of baseline characteristics of our sample, by assignment to treatment status. About 54% of the sample are females, the average age at the time of application is 21 years old, 94% of applicants are single, 11% have a child, 64% are high school graduates (among those aged 18 years or older) and 9% are college graduates (among those aged 21 years or older).<sup>21</sup> The average poverty rate in the applicants' neighborhoods (or cities) was 7.5%, the unemployment rate was 5.7%, and the informality rate was 42.9%. That table also presents information of the key eligibility criterion: formal employment and wages should be close to zero in the six months prior to the start of the program (December 2011 to May 2012). The numbers in Table 1 confirm that this requirement was met in the vast majority of applications. Randomization guarantees the balance in characteristics between control and treatment groups, as we show in column (2).

## 3 Empirical strategy

Let's define  $Y_{1it}$  and  $Y_{0it}$  as the outcome of interest (having a formal job or formal salaries) with and without the treatment for individual  $i$ ,  $t$  periods after the end of the program. The causal

---

<sup>20</sup>Given that this program was not originally designed to be evaluated, the quality of contact information of applicants was rather poor and only around a third of telephone numbers in the administrative database were correct. Since there is no a priori information indicating that this feature of the data could lead to a selection problem (those individuals with correct contact information being different of those with incorrect contact information), this low fraction is not much of a concern. In addition, as shown in Table 3, there are no significant differences in observed characteristics among the final surveyed population and the overall population of applicants in the City of Córdoba.

<sup>21</sup>Interestingly, the sample of PPP applicants is very much representative of the population of individuals aged 16-25 in the Province of Córdoba.

effect of the program is

$$\tau_t = \mathbb{E}\{Y_{1it} - Y_{0it}\}. \quad (1)$$

Given the random assignment into treatment and that we can follow all applicants in formal employment administrative registries during three and a half years after graduation (starting June 2013 on), then the average outcome under treatment is equal to the average of the observed outcome for those randomly offered the PPP. Similarly, the counterfactual may be estimated using the average of the observed outcome for those not selected to receive the benefit of the program. Hence, the parameter of interest is estimated with the sample analog of

$$\tau_t = \mathbb{E}\{Y_{1it} - Y_{0it}\} = \mathbb{E}\{Y_{it}|D_i = 1\} - \mathbb{E}\{Y_{it}|D_i = 0\}, \quad \text{for } t = 1, \dots, 42, \quad (2)$$

where  $D_i = \{0, 1\}$  is an indicator of whether individual  $i$  was randomly selected through the lottery process and  $Y_{it}$  is the observed outcome  $t$  months after (from June 2013 to November 2016). The parameters  $\tau_t$  in equation 2 can be estimated with OLS regressions of the form

$$Y_{it} = \alpha + \tau_t D_i + X_i \gamma + u_{it}, \quad (3)$$

where  $X$  is a vector of control variables that affect the probability of being assigned to the treatment (size of the firm with which the individual formed the applicant match, the number of applications received by each firm, county level fixed effects, and the number of applications each individual submitted). We need to control for these variables due to the fact that the final assignment of beneficiaries has to fulfill certain quotas at the county and firm level, and because individuals were allowed to submit more than one application.<sup>22</sup>

To estimate the impacts of the program we rely on the original randomization. Although not all beneficiaries were actually treated, compliance was very high: the take up rate of the program was 97% among those offered the treatment and only 3% of individuals in the control

---

<sup>22</sup>For individuals who submitted more than one application and were not resulted as treated we randomly selected the firm's characteristics in only one of the applications.

group received the benefit. Additionally, 83% of treated individuals successfully completed the 12 months of internship. Still, given that there is not full compliance, our estimates should be interpreted as intention-to-treat effects (ITT), but we also produce IV estimates where effective treatment is instrumented with the random assignment (see section 4).<sup>23</sup>

To analyze the short run effects of the program we use information of the follow-up survey which allows to estimate effects on outcomes measures just 12 months after the end of the program (May 2014). Table 2 shows that treatment and control groups in this subsample were also balanced in pre-treatment characteristics.

## 4 Main results: Short and long run effects of the quality of first job experience

### 4.1 Employment outcomes in the short run

To see whether the PPP affected the employability of beneficiaries we first look at short run effects, that is, we examine causal impacts of the program on outcomes observed 12 months after the program finished (i.e., 24 months after it started). Table 4 summarizes the estimated effects of PPP on labor force participation, the probability of employment (of any kind) and the probability of formal employment for both individuals residing in the city of Córdoba - to compare these results with those of individuals in the survey- and for all applicants in the province. Information for the first two outcomes comes from the follow up survey, while information for formal employment status comes from administrative data.<sup>24</sup> The results are presented for all individuals (panel A), and for female (panel B) and male applicants (panel C), and they include both ITT and IV estimates.

Results in Table 4 indicate that the decision to participate in the labor market is not affected

---

<sup>23</sup>We consider that an individual is effectively treated if he or she remains 6 or more months in the program. This measure should be considered with caution since compliance could be even higher than that we are able to calculate from the administrative data. This is because PPP beneficiaries can change the firm where they spend the year of the internship and if this happens the administrative record about the number of completed months in PPP corresponds to the time completed in the last firm. Hence, ITT are our preferred estimates.

<sup>24</sup>Interestingly, the impacts on formality obtained from self-report in the survey –not shown here– and administrative data are perfectly compatible.

by the PPP, and that there are large and positive effects on the probability of employment (ITT of +7pp, slightly higher in the IV) and of formal employment (+6.1pp in the city of Córdoba and +5.2pp in the entire province, slightly higher in the IVs). Given that the size of the impact on the probability of employment is similar to that one in the probability of formal employment it seems safe to state that most of the new employment induced by the PPP is formal.

Table 4 also shows that the impacts on employment probability seem higher for men than for females, but it is a difference that we cannot rule out it is equal to zero. On the contrary, the effects on the probability of formal employment are larger for females (the null hypothesis that the difference is equal to zero can be statistically rejected in the IV case, as shown in Panel D of Table 5).

## 4.2 Quality of employment in the long run

In the longer run some of the above discussed shorter-run effects could in principle disappear, as shown in previous literature analyzing different employment programs for the youth. However, we find that impacts are large and persistent in the long run, namely 4.5 years after the program started (42 months after it finalized). Table 5 shows these results for 6, 12, 18, 24, 30, 26 and 42 months after the end of the PPP 2012, both as ITT and as IV estimates of the causal impacts of the program.

The first block of results in Table 5 corresponds to the probability of formal employment, while the second block refers to formal labor earnings (unconditional on formal employment). These results indicate that the PPP produces large gains in formal employment and also in formal salaries, a result that in fact is induced by the program augmenting the share of formal workers (see section 6.1). The impacts on formal employment are higher, both in terms of the mean in the control group and in absolute terms, for females. The formal test of differences between impacts for females and males are also shown in the bottom panel of Table 5, and for the case of the IV estimates we can say that the impacts are statistically different across gender groups.<sup>25</sup>

---

<sup>25</sup>Table 6 shows the results for each one of the months for which we have available administrative data. This table also adds a number of robustness checks with alternative specifications –including covariates and

## 5 Displacement

The empirical strategy proposed in equation (2) assumes that the potential outcomes of an individual  $i$  only depend on his own treatment status, regardless of how the treatment is distributed among the eligible population. However, one important criticism against training and employment policies is that individuals receiving treatment may improve their employment prospects at the expense of crowding out other individuals. This potential externality, or displacement effect, violates the stable unit treatment value assumption, or SUTVA (Rubin 1980, 1990), meaning that the employment rate of individuals in the control group is lower than it would have been absent the program. The presence of this phenomenon would imply that our results overestimate the effect of the program, and that the bias would be greater the larger the share of treated individuals assigned to treatment.

Displacement can occur also among treated individuals. As the share of treated individuals rises, more individuals with some job experience will be competing with one another for jobs at the end of the internship period, which implies that –once the displacement effect over the control group is controlled for– the impact of the program would be lower than in the case where this effect is absent.

An ideal setup to identify displacement effects consists of a two-step randomization procedure, where in the first step one randomly assigns to different cities or labor markets the share of eligible individuals to be treated, and afterwards one randomly assigns individuals to the treatment or the control group within the city and at the chosen rate.<sup>26</sup> Under this ideal setup, one can identify displacement effects operating over the control group by comparing the labor market performance of control individuals in cities with positive shares of treated individuals to the performance of individuals in a “super control” group, that is, a group of eligible individuals in 0% assignment areas. Additionally, one can study whether treatment effects are heterogeneous across cities and examine whether they are declining as the share of treatment individuals increases in the a given labor market or city.

---

estimates obtained by an inverse probability weight procedure–, which are all consistent with our baseline preferred specification presented in Table 5.

<sup>26</sup>This is the approach taken by Crépon et al. (2013).

Although the PPP program was not designed to randomly assign the share of treated individuals across cities, we take advantage of the substantial as-good-as-random variation that results from the quotas in the lottery and of the large number of cities that participated in the program (289 cities) to approximate the ideal setup described above. Figure 3 shows that the share of treated individuals at the city level takes values in the whole interval  $[0, 1]$ , and Table 7 shows that this variability is present even within counties. Additionally, Table 8 indicates that these shares are uncorrelated with several observable city characteristics in a baseline period.<sup>27</sup> We approximate the “super control” group with cities with a “low” share of treated individuals.<sup>28</sup> We consider that the share is low whenever less than 20% of applicants received treatment. We compare the labor market performance of individuals in this group of cities with individuals in cities with a moderate share of treatment individuals (20% to 40%) and with individuals in cities with large share of individuals in the treatment group (more than 40%).<sup>29</sup> We then explore the existence of displacement effects running the following regression in the spirit of Crépon et al. (2013):

$$\begin{aligned}
Y_{itk} &= \tau_t^{low} D_i C^{low} + \tau_t^{med} D_i C^{med} + \tau_t^{high} D_i C^{high} \\
&+ \delta^m C^m + \delta^h C^h \\
&+ X_i \lambda_1 + X_k \lambda_2 + u_{itj},
\end{aligned} \tag{4}$$

where  $Y_{itj}$  is formal employment status of individual  $i$  in period  $t$  living in city  $k$ ;  $C^j$ , for  $j = \{low, med, high\}$ , equals 1 if  $i$  lives in a city with low, moderate or high share of treated individuals, respectively;  $X_i$  is the vector of variables that affect the probability of being assigned to the treatment; and  $X_k$  is a vector of city’s  $k$  characteristics in a baseline period.

Coefficients  $\tau^j$  measure the effect of being assigned to treatment in cities with different

---

<sup>27</sup>Since the PPP imposes quotas at the county level and this may affect the proportion of applicants that can receive treatment within the county, in all regressions of Table 8 we control for county fixed effects.

<sup>28</sup>Although there are cities with 0% of treated individuals, the proportion of our sample in those cities is extremely small.

<sup>29</sup>Our results are robust to alternative cutoffs, provided that we have a sufficiently large number of observations in each one of the defined groups.

shares of individuals assigned to treatment, comparing to individuals in the control group in the same type of cities. Coefficients  $\delta^m$  and  $\delta^h$  capture the effect of being assigned to the control group in a city with medium and high share of treated individuals, respectively, compared to being in the control group in areas with a low share of treated individuals.

If the PPP program induces displacement in the control group, then coefficients  $\delta^{med}$  and  $\delta^{high}$  should be negative, and  $\delta^{high}$  should be larger in absolute terms. At the same time, the  $\tau^j$  coefficients should decrease as the share of treated individuals increases. This exercise has its own limitations because if there were displacement over the control group in cities with a low share of treated individuals, we would not be able to identify it. In that case, the  $\delta$  estimates would be a downward bias estimates of the true displacement in areas with moderate and high shares.

We estimate equation 4 for the post-PPP period (June 2014 to November 2016), and for the PPP period (June 2012 to May 2013) – during which individuals in the treatment group are all doing the internship– we run the following simplified version, restricting the analysis only to individuals in the control group:

$$\begin{aligned}
 Y_{itk} &= \alpha + \delta^m C^m + \delta^h C^h \\
 &+ X_i \lambda_1 + X_k \lambda_2 + u_{itj}.
 \end{aligned}
 \tag{5}$$

The results in Table 9 are not consistent with the presence of displacement effects. First, Panel A presents the results of equation 5, which corresponds to the period in which the PPP was in place. Coefficients in columns (7) and (9) indicate that during this period individuals in the control group in cities with moderate (20-40%) and high (+40%) shares of treated individuals do not perform worse in the formal labor market than control individuals in low share cities. Panel B presents the results of equation 4, where heterogeneous effects of the program across types of cities are also allowed. Again, coefficients in columns (7) and (9) indicate that displacements effects over the control group are not operating, and, if anything, the effects go in the opposite direction, as some coefficients for cities +40% are positive and

statistically significant. Regarding the effects of the program in the different groups of cities (columns 1, 3 and 5), there is no clear pattern indicating that the effects are lower as the share of treated individuals rises. We read these results as in favor of no displacement.

## 6 Mechanisms: Information or human capital?

### 6.1 On-the-job learning?

This section explores the mechanisms driving the positive impacts in labor markets outcomes caused by the PPP. First, we analyze whether the job experience granted by this program help youth to develop new skills by a learning-by-doing process (since there are no formal training involved by the program). At this point, it is important to recognize that the possibility to take advantage of an on-the-job learning opportunity is closely related to the type of tasks the worker is supposed to do. Although job opportunities in the PPP are in formal firms, most of the tasks that youth do during the internship are quite basic and basically do not provide interesting learning spaces.<sup>30</sup>

A first way to analyze whether the PPP have caused skills development, or productivity gains, is to look at what happened with salaries. As shown previously, unconditional on working for the formal sector, formal salaries has increased for PPP beneficiaries. However, and as shown in the dashed line in Figure 2, these impacts are close to zero if we condition on being a formally registered worker. The bounds shown in Figure 2 are constructed following Attanasio et al. (2011) and, although they are rather conservative, from them we cannot rule out a zero wage (skill) effect.

A more direct test of skills development after the PPP experience can be obtained from the measures of “cognitive” and “non-cognitive” skills in the survey data. Even though the measures of cognitive skills are, by its definition and construction, not likely to be changed

---

<sup>30</sup>For instance, from information gathered in the survey we know that the qualification of job tasks for individuals who were employed by the date of the survey is more or less the same for control and treatment individuals, except for the fact that for men we observe that the treated and employed individuals are in jobs that require less often that they operate mechanical machinery and more often to use computers, compared to control individuals that were also employed. We also observe that treated individuals are more likely to work in services and retail and less in manufacturing sectors, when compared to control individuals.



by having been part of the PPP<sup>31</sup>, we still report in Panel A of Table 10 the estimates of the PPP on two different measures of cognitive abilities, and find no significant effects. Regarding socio-emotional skills, the estimates (Panel B in Table 10) also show no significant impacts.

In addition, in the survey we also produced a number of survey experiments (endorsement experiments) in order to elicit truthful responses to sensitive questions addressing issues of good behavior, attitudes and beliefs in the workplace.<sup>32</sup> We also find no impacts on these measures.<sup>33</sup>

As a whole, this evidence does not support a channel related to skill acquisition as the driver of the better labor market outcomes caused by the PPP. Therefore, the reduction of informational frictions is the other competing hypotheses, and we discuss it next.

## 6.2 Evidence consistent with impacts on informational barriers

The PPP not only offers beneficiaries the possibility to have a labor market experience and to acquire, eventually, skills through on-the-job-training, but it also provides the opportunity to credibly certify this experience, as long as the internship takes place in a formal firm. This certification is important for employers – who may not have sufficient incentives to hire in-experience workers whose abilities are uncertain– as it provides a signal of the capacities of the young worker. Hence, part of the effects of the PPP on employment outcomes may be explained by a signaling/certification effect, and not by the acquisition of new skills provided by the experience per se.

If the signal of having had a formal work experience is valuable for prospective employers (as a way to infer unobserved productivity of the young worker), then it should be the case that PPP has a greater impact on individuals that lack certification of their abilities. An exercise that allows to explore this hypothesis is to compare the impact of the PPP across individuals with and without certification provided by the educational system (high school completed versus

---

<sup>31</sup>Brenlla (2014) is the technical note explaining the definitions and characteristics of these measures included in the survey. As described in that note, the measure for cognitive skills is not very malleable at the age of applicants to this program.

<sup>32</sup>These experiments were designed to measure the impact of the PPP in the agreement with the following statements: *lack of timekeeping in the workplace should be punished*, *maltreatment at the workplace is unacceptable*, and *teamwork helps in being more efficient at work*.

<sup>33</sup>However, given the sample size these tests are low-powered. These results are not reported here, but are available upon request.

dropouts). Table 11 shows the heterogeneous impacts by educational level for individuals older than 18 years old who completed high school education (HS graduates) and for those who did not (dropouts). A comparison between these two groups indicates that the effects of the PPP are, as consistent with the signaling mechanism, larger (as % of the mean observed for the control group) for individuals who has not been signaled by the educational system.

Another exercise provides further evidence supporting the signaling hypothesis. If this program helps high type workers to demonstrate how large his or her ex ante unobserved productivity is, then heterogeneous impacts by (previously accumulated) skills should be such that employment outcomes would improve more for the high types, as it is the case in the results shown in Table 12.

## 7 Conclusions

Labor market outcomes for the youth are noticeably worse than for adults. Both unemployment and informal employment are more common among younger workers, specially among those from disadvantaged backgrounds. These problems are likely to be associated to barriers in entry-level labor markets. Moreover, a bad start -i.e., a low quality first job experience- is likely to have persistent effects on labor markets prospects, which makes these features of the entry-level jobs more worrying.

In this paper we provide evidence on the importance of entry-level job experience in the formal sector, which proxies for a high quality job for the youth. To identify the causal effects of interest we exploit a lottery that randomly assigns vacancies to an internship program in formal firms.

We measure impacts in several dimensions, using both administrative and survey data. We find that the internship program caused large short-to-medium run gains in the probability of formal employment and a moderate fall in unemployment. The evidence about the impacts on wages is not conclusive and our preferred interpretation of the estimates on this regard says that if there is any positive effect on (formal) salaries, its size is not large. This null to low impacts on wages inform about the importance of mechanisms other than human capital

(skills) accumulation as the main channels through which the program actually boosted the labor market outcomes of beneficiaries.

We also explore in other ways which alternative mechanisms could be driving the observed impacts and we find evidence favoring a signaling effect. Moreover, exploiting a special feature of the random assignment process, we are able to rule out the existence of one type of general equilibrium effects (displacement). Therefore, the overall evidence indicates that the program produces a signal that is valuable in the labor market while it does not crowd out the job opportunities of non beneficiaries.

## References

- Acevedo, P., Cruces, G., Gertler, P., and Martinez, S. (2015). Soft skills and hard skills in youth training programs. long term experimental evidence from the dominican republic. *Mimeo*.
- Acevedo, P., Cruces, G., Gertler, P., and Martinez, S. (2017). Living Up to Expectations: How Job Training Made Women Better Off and Men Worse Off. NBER Working Papers 23264, National Bureau of Economic Research, Inc.
- Altonji, J. G., Kahn, L. B., and Speer, J. D. (2016). Cashier or consultant? entry labor market conditions, field of study, and career success. *Journal of Labor Economics*, 34(S1):S361–S401.
- Alzua, M. L., Cruces, G., and Lopez, C. (2016). Long-Run Effects Of Youth Training Programs: Experimental Evidence From Argentina. *Economic Inquiry*, 54(4):1839–1859.
- Attanasio, O., GuarÃn, A., Medina, C., and Meghir, C. (2015). Long term impacts of vouchers for vocational training: Experimental evidence for colombia. NBER Working Papers 21390, National Bureau of Economic Research, Inc.
- Attanasio, O., Kugler, A., and Meghir, C. (2011). Subsidizing vocational training for disadvantaged youth in colombia: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, pages 188–220.
- Bell, B., Bindler, A., and Machin, S. J. (2015). Crime scars: Recessions and the making of career criminals.
- Card, D., Ibararán, P., Regalia, F., Rosas-Shady, D., and Soares, Y. (2011). The labor market impacts of youth training in the dominican republic. *Journal of Labor Economics*, 29(2):267–300.
- Card, D., Kluve, J., and Weber, A. (2017). What works? a meta analysis of recent active labor market program evaluations. *Journal of the European Economic Association*, (forthcoming).

- Crépon, B., Duflo, E., Gurgand, M., Rathelot, R., and Zamora, P. (2013). Do labor market policies have displacement effects? evidence from a clustered randomized experiment. *The Quarterly Journal of Economics*, 128(2):531–580.
- Eriksson, S. and Rooth, D.-O. (2014). Do employers use unemployment as a sorting criterion when hiring? evidence from a field experiment. *The American Economic Review*, 104(3):1014–1039.
- Gelber, A., Isen, A., and Kessler, J. B. (2015). The effects of youth employment: Evidence from new york city lotteries<sup>1</sup>. *The Quarterly Journal of Economics*, page qjv034.
- Ibarrarán, P., Ripani, L., Taboada, B., Villa, J. M., and García, B. (2014). Life skills, employability and training for disadvantaged youth: Evidence from a randomized evaluation design. *IZA Journal of Labor & Development*, 3(1):1–24.
- Kawaguchi, D. and Murao, T. (2014). Labor-market institutions and long-term effects of youth unemployment. *Journal of Money, Credit and Banking*, 46(S2):95–116.
- Kluve, J., Puerto, S., Robalino, D., Romero, J., Rother, F., Stoeterau, J., Weidenkaff, F., and Witte, M. (2017). Interventions to improve labour market outcomes of youth: A systematic review of training, entrepreneurship promotion, employment services, and subsidized employment interventions. *Campbell Systematic Reviews, The Campbell Collaboration*.
- Kluve, J., Puerto, S., Robalino, D., Rother, F., Weidenkaff, F., Stoeterau, J., Tien, B., and Witte, M. (2014). Interventions to improve labour market outcomes of youth: a systematic review of training, entrepreneurship promotion, employment services, mentoring, and subsidized employment interventions.
- Kroft, K., Lange, F., and Notowidigdo, M. J. (2013). Duration dependence and labor market conditions: Evidence from a field experiment. *The Quarterly Journal of Economics*, 128(3):1123–1167.
- McKenzie, D. (2017). How Effective Are Active Labor Market Policies in Developing Countries?

A Critical Review of Recent Evidence. IZA Discussion Papers 10655, Institute for the Study of Labor (IZA).

Oreopoulos, P., von Wachter, T., and Heisz, A. (2012). The short-and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics*, 4(1):1–29.

Pallais, A. (2014). Inefficient hiring in entry-level labor markets. *American Economic Review*, 104(11):3565–99.

## 8 Figures and tables

Figure 1: Pool of applicants, random assignment and final samples

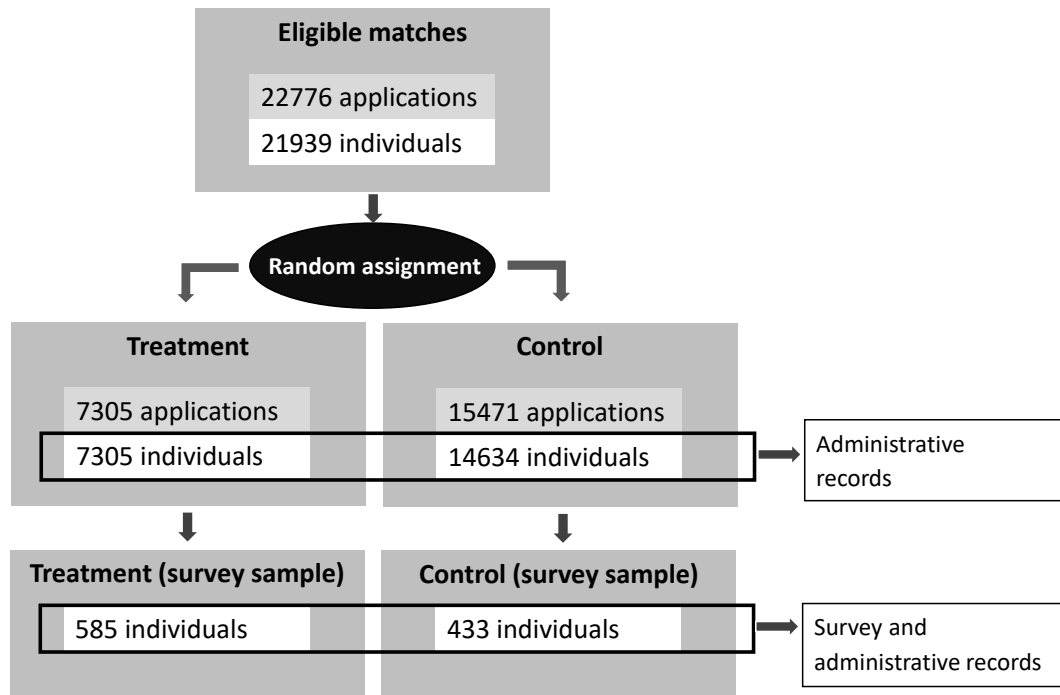
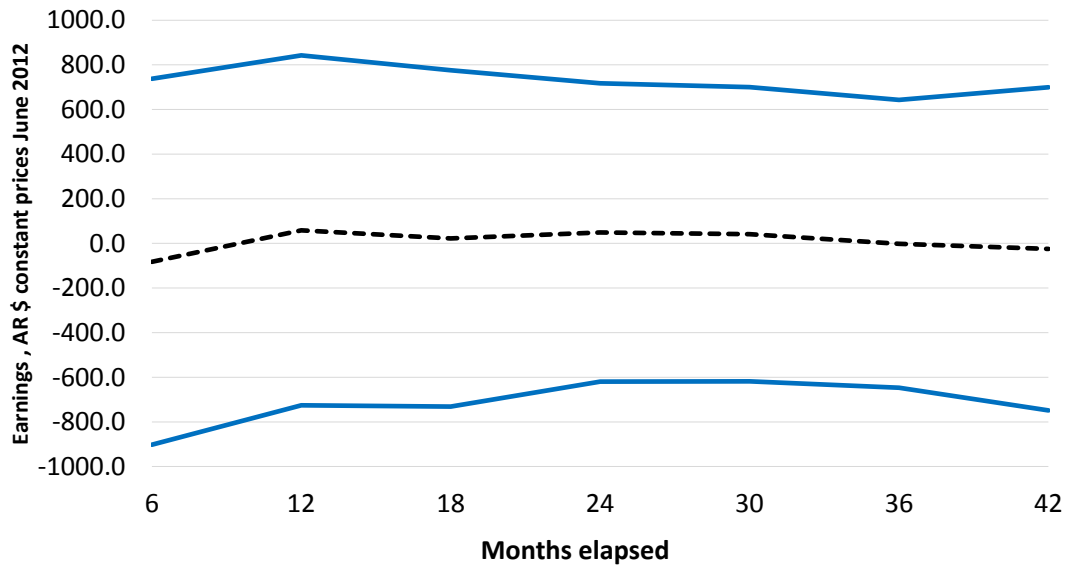


Figure 2: Effect of the PPP program on monthly earnings



Note: The dashed line corresponds to the point estimates of an OLS estimation of equation 3 where the outcome variable are monthly earnings in formal employment. Earnings take positive values only for those who have a formal job. The blue lines correspond to the upper and lower bounds computed as explained in the appendix.

Figure 3: Distribution of the share of treated individuals across cities

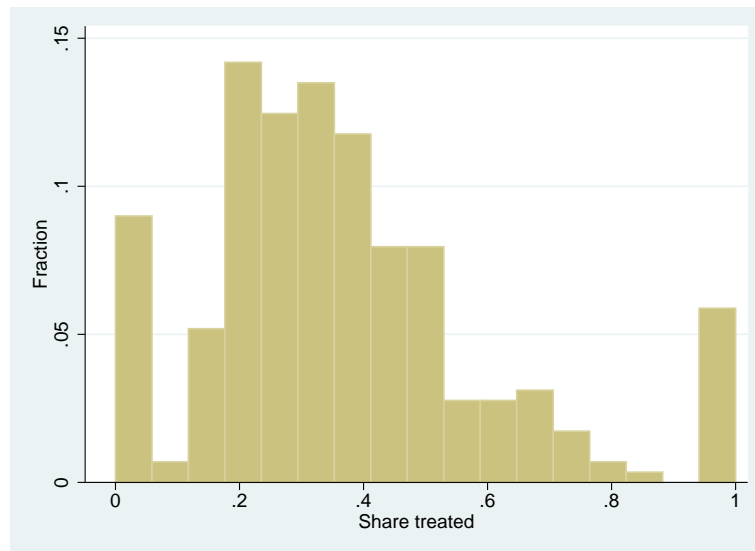




Figure 4: Distribution of the share of treated individuals across cities, by county

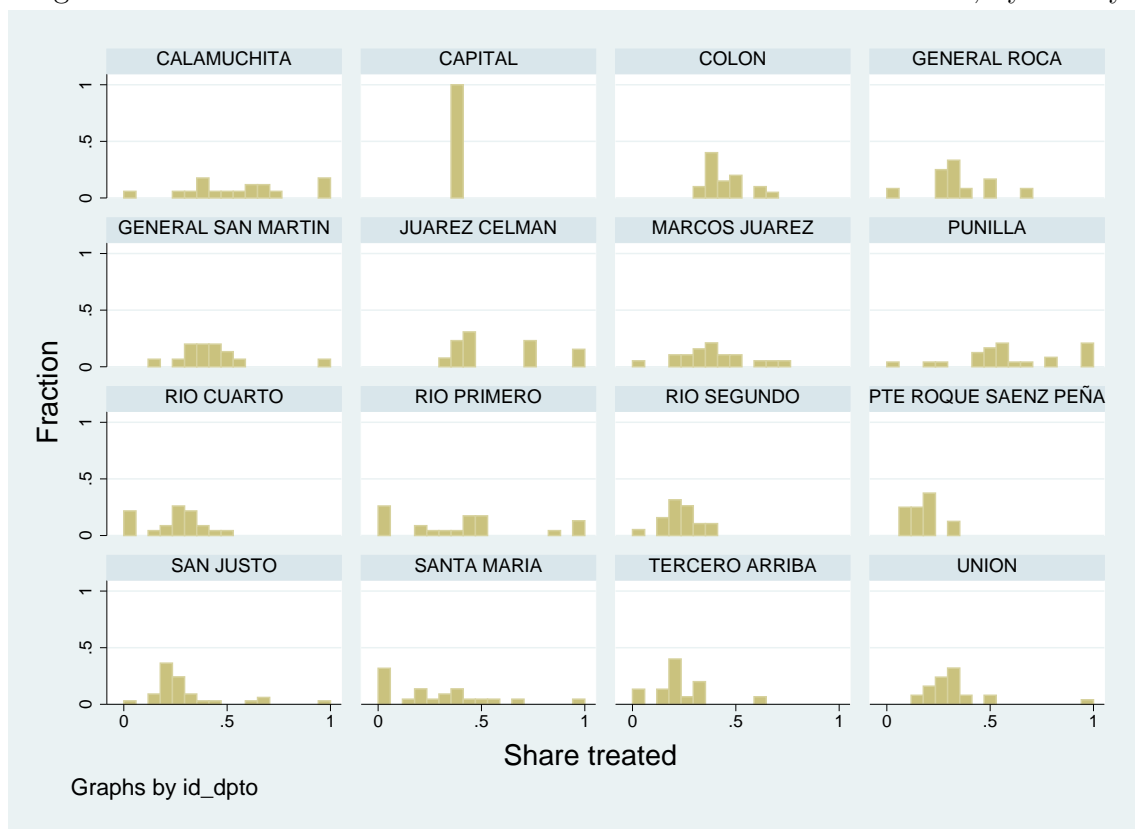


Table 1: Baseline differences between treatment and control group

Baseline Characteristics	(1)	(2)	(3)	(4)
	Mean	Difference	se	N
<b>A. All</b>				
Female	0.544	-0.017**	0.007	21939
Age (in years) <sup>a</sup>	21.017	0.030	0.036	21932
Single <sup>a</sup>	0.937	0.004	0.003	21939
Have children <sup>a</sup>	0.109	-0.011***	0.004	21939
High school graduate (if +18 years old) <sup>a</sup>	0.643	0.011	0.008	16025
College graduate (if +21 years old) <sup>a</sup>	0.090	0.013*	0.008	7271
Poverty rate 2008 (neighborhood or city)	0.075	0.000	0.000	21786
Unemployment rate 2008 (city or neighborhood)	0.057	0.000	0.000	21828
Labor informality rate 2008 (city or neighborhood)	0.429	-0.003**	0.002	21786
Paper application	0.629	-0.001	0.007	21939
Fomal employment Dec 2011	0.017	0.003	0.002	21939
Fomal employment Jan 2012	0.010	0.002	0.002	21939
Fomal employment Feb 2012	0.002	0.001	0.001	21939
Fomal employment Mar 2012	0.005	0.002	0.001	21939
Fomal employment Apr 2012	0.008	0.001	0.001	21939
Fomal employment May 2012	0.017	0.001	0.002	21939
<b>B. Female</b>				
Age (in years) <sup>a</sup>	21.442	0.044	0.049	11811
Single <sup>a</sup>	0.923	0.003	0.005	11814
Have children <sup>a</sup>	0.162	-0.010	0.007	11814
High school graduate (if +18 years old) <sup>a</sup>	0.716	0.010	0.010	9248
College graduate (if +21 years old) <sup>a</sup>	0.111	0.015	0.010	4746
Poverty rate 2008 (neighborhood or city)	0.075	0.000	0.001	11746
Unemployment rate 2008 (city or neighborhood)	0.056	0.000	0.000	11762
Labor informality rate 2008 (city or neighborhood)	0.432	-0.003	0.002	11746
Paper application	0.645	-0.009	0.009	11814
Fomal employment Dec 2011	0.011	0.007***	0.003	11814
Fomal employment Jan 2012	0.007	0.001	0.002	11814
Fomal employment Feb 2012	0.001	0.001	0.001	11814
Fomal employment Mar 2012	0.003	0.004***	0.001	11814
Fomal employment Apr 2012	0.006	0.002	0.002	11814
Fomal employment May 2012	0.013	0.001	0.002	11814
<b>C. Male</b>				
Age (in years) <sup>a</sup>	20.509	0.048	0.050	10121
Single <sup>a</sup>	0.953	0.004	0.004	10125
Have children <sup>a</sup>	0.045	-0.009**	0.004	10125
High school graduate (if +18 years old) <sup>a</sup>	0.541	0.018	0.013	6777
College graduate (if +21 years old) <sup>a</sup>	0.049	0.011	0.010	2525
Poverty rate 2008 (neighborhood or city)	0.075	0.000	0.001	10040
Unemployment rate 2008 (city or neighborhood)	0.057	0.000*	0.000	10066
Labor informality rate 2008 (city or neighborhood)	0.426	-0.004	0.002	10040
Paper application	0.611	0.009	0.010	10125
Fomal employment Dec 2011	0.024	-0.002	0.003	10125
Fomal employment Jan 2012	0.013	0.002	0.003	10125
Fomal employment Feb 2012	0.003	0.000	0.001	10125
Fomal employment Mar 2012	0.006	-0.001	0.002	10125
Fomal employment Apr 2012	0.011	0.000	0.002	10125
Fomal employment May 2012	0.022	0.000	0.003	10125

Note:<sup>a</sup> Measured at the date of application. Column (2) reports the mean difference between the treatment and control group. This difference is the estimated  $\beta$  coefficient of the following regression  $y_i = \alpha + \beta D_i + \gamma X + u_i$ , where  $Y_i$  is the baseline characteristic of individual  $i$ , and  $X$  is a vector of variables that affect the probability of being assigned to the treatment (size of the firm where the individual chose to do the internship, the number of applications received by the firm, county fixed effects, and the number of applications the individual submitted). Column (3) reports robust standard errors. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 2: Survey sample: Baseline differences between treatment and control group

Baseline Characteristics	(1)	(2)	(3)	(4)
	Mean	Difference	se	N
<b>Control group</b>				
<b>A. All</b>				
Female	0.485	0.019	0.032	1018
Age (in years)	20.993	0.154	0.161	1018
Single	0.961	-0.005	0.013	1018
Have children	0.051	0.010	0.015	1018
High school graduate (if +18 years old)	0.788	-0.050	0.037	559
College graduate (if +21 years old)	0.102	-0.023	0.035	243
Poverty rate 2008 (neighborhood or city)	0.070	0.003**	0.001	1018
Labor informality rate 2008 (city or neighborhood)	0.318	0.002	0.010	1018
Paper application	0.490	0.011	0.031	1018
Fomal employment Dec 2011	0.016	0.012	0.009	1018
Fomal employment Jan 2012	0.007	0.001	0.006	1018
Fomal employment Feb 2012	0.002	-0.002	0.002	1018
Fomal employment Mar 2012	0.012	-0.004	0.006	1018
Fomal employment Apr 2012	0.014	-0.007	0.007	1018
Fomal employment May 2012	0.018	-0.001	0.009	1018
<b>B. Female</b>				
Age (in years)	21.400	0.101	0.238	507
Single	0.948	0.014	0.019	507
Have children	0.086	0.009	0.027	507
High school graduate (if +18 years old)	0.890	-0.069	0.044	289
College graduate (if +21 years old)	0.103	0.023	0.049	139
Poverty rate 2008 (neighborhood or city)	0.071	0.002	0.002	507
Labor informality rate 2008 (city or neighborhood)	0.322	0.003	0.014	507
Paper application	0.500	-0.015	0.045	507
Fomal employment Dec 2011	0.024	0.002	0.014	507
Fomal employment Jan 2012	0.010	-0.004	0.009	507
Fomal employment Feb 2012	0.005	-0.006	0.006	507
Fomal employment Mar 2012	0.019	-0.020	0.010	507
Fomal employment Apr 2012	0.024	-0.026	0.012	507
Fomal employment May 2012	0.014	-0.010	0.010	507
<b>C. Male</b>				
Age (in years)	20.610	0.154	0.216	511
Single	0.973	-0.024	0.017	511
Have children	0.018	0.007	0.013	511
High school graduate (if +18 years old)	0.673	-0.033	0.060	270
College graduate (if +21 years old)	0.100	-0.077	0.048	104
Poverty rate 2008 (neighborhood or city)	0.070	0.004	0.002	511
Labor informality rate 2008 (city or neighborhood)	0.313	0.000	0.013	511
Paper application	0.480	0.033	0.045	511
Fomal employment Dec 2011	0.009	0.022	0.012	511
Fomal employment Jan 2012	0.004	0.006	0.008	511
Fomal employment Feb 2012	0.000	0.000	0.000	511
Fomal employment Mar 2012	0.004	0.008	0.008	511
Fomal employment Apr 2012	0.004	0.010	0.009	511
Fomal employment May 2012	0.022	0.010	0.014	511

Note:<sup>a</sup> Measured at the date of application. Column (2) reports the mean difference between the treatment and control group. This difference is the estimated  $\beta$  coefficient of the following regression  $y_i = \alpha + \beta D_i + \gamma X + u_i$ , where  $Y_i$  is the baseline characteristic of individual  $i$ , and  $X$  is a vector of variables that affect the probability of being assigned to the treatment (size of the firm where the individual chose to do the internship, the number of applications received by the firm, county fixed effects, and the number of applications the individual submitted). Column (3) reports robust standard errors. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 3: Survey sample representativeness (city of Cordoba)

Baseline Characteristics	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	A. All				B. Female				C. Male			
	Eligible in Cordoba  city	Diff.	se	N	Eligible in Cordoba  city	Diff.	se	N	Eligible in Cordoba  city	Diff.	se	N
Female	0.523	-0.019	(0.017)	8,882								
Age (years)	20.459	0.101	(0.083)	8,862	20.841	0.086	(0.120)	4,598	20.040	0.146	(0.113)	4,264
Single	0.947	0.011*	(0.007)	8,882	0.937	0.022**	(0.010)	4,612	0.958	-0.000	(0.009)	4,270
Paper application	0.515	-0.022	(0.017)	8,882	0.527	-0.038	(0.023)	4,612	0.503	-0.005	(0.024)	4,270
Material deprivation rate in households 2008 (neighb.)	0.153	-0.009***	(0.003)	8,842	0.157	-0.004	(0.005)	4,598	0.149	-0.013***	(0.005)	4,244
UBN rate 2008 (neighb.)	0.092	-0.003	(0.003)	8,842	0.092	0.002	(0.004)	4,598	0.092	-0.009***	(0.003)	4,244
Head of hhd dropout high school 2008 (neighb.)	0.491	-0.003	(0.008)	8,842	0.485	0.010	(0.011)	4,598	0.498	-0.018	(0.011)	4,244
Head of hhd with college degree 2008 (neighb.)	0.117	-0.000	(0.004)	8,842	0.119	-0.007	(0.005)	4,598	0.115	0.006	(0.006)	4,244
Illiteracy rate 2008 (neighb.)	0.025	-0.001	(0.000)	8,842	0.024	0.000	(0.001)	4,598	0.025	-0.001*	(0.001)	4,244
Population aged 15-19 y.o. in formal education 2008 (neighb.)	0.698	0.005	(0.005)	8,842	0.701	-0.001	(0.007)	4,598	0.696	0.012*	(0.006)	4,244
Population aged 20-24 y.o. enrolled in univ. 2008 (neighb.)	0.303	-0.003	(0.007)	8,842	0.311	-0.011	(0.010)	4,598	0.294	0.005	(0.010)	4,244
Population aged 20-24 y.o. in post-sec. educ. 2008 (neighb.)	0.387	-0.002	(0.008)	8,842	0.395	-0.012	(0.011)	4,598	0.378	0.009	(0.011)	4,244
Labor informality rate 2008 (neighb.)	0.327	-0.006	(0.005)	8,842	0.325	0.003	(0.007)	4,598	0.329	-0.015**	(0.007)	4,244
Unemployment rate 2008 (city or neighb.)	0.073	-0.001	(0.001)	8,842	0.073	-0.000	(0.001)	4,598	0.073	-0.001	(0.001)	4,244
Fomal employment Jan 2012	0.014	-0.004	(0.004)	8,882	0.011	0.002	(0.005)	4,612	0.017	-0.010**	(0.004)	4,270
Fomal employment Feb 2012	0.004	-0.002	(0.002)	8,882	0.003	0.003	(0.003)	4,612	0.005	-0.006***	(0.001)	4,270
Fomal employment Mar 2012	0.007	0.003	(0.003)	8,882	0.006	0.004	(0.005)	4,612	0.009	0.002	(0.005)	4,270
Fomal employment Apr 2012	0.011	-0.001	(0.004)	8,882	0.011	0.001	(0.005)	4,612	0.012	-0.003	(0.005)	4,270
Fomal employment May 2012	0.023	-0.005	(0.005)	8,882	0.019	-0.010*	(0.005)	4,612	0.026	-0.000	(0.008)	4,270
Wage in formal job Jan 2012	20.437	11.072	(12.599)	8,882	20.692	15.831	(18.758)	4,612	20.159	6.166	(16.802)	4,270
Wage in formal job Feb 2012	4.832	1.536	(5.861)	8,882	4.740	9.187	(11.508)	4,612	4.933	-6.218***	(1.823)	4,270
Wage in formal job Mar 2012	8.304	8.061	(7.137)	8,882	7.340	15.272	(12.687)	4,612	9.360	0.754	(6.444)	4,270
Wage in formal job Apr 2012	14.580	9.400	(10.013)	8,882	14.647	11.957	(13.638)	4,612	14.507	6.801	(14.672)	4,270
Wage in formal job May 2012	46.952	-12.027	(12.098)	8,882	40.884	-11.302	(16.163)	4,612	53.596	-13.412	(18.041)	4,270

Note: Column (2), (6) and (10) report the difference in each variable between eligible individuals in the city of Cordoba and surveyed individuals. Robust standard errors in parenthesis. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 4: Short run effect of the PPP program on labor outcomes (12 month after the end of the program)

Outcomes	ITT			IV			N (5)	Mean control group (6)	
	Coef. (1)		Std. Error (1)	Coef. (3).		Std. Error (4)			
<b>PANEL A. All individuals</b>									
Labor force participation	0.013		0.020	0.015		0.023	1018	0.891	
Employment	0.070	**	0.033	0.086	**	0.038	915	0.334	
Formal employment	Córdoba (City)	0,061	***	0.009	0.073	***	0.011	8,886	0.189
	Córdoba (Province)	0,052	***	0.006	0.063	***	0.007	21938	0.163
<b>PANEL B. Female</b>									
Labor force participation <sup>a</sup>	0.009		0.028	0.009		0.031	508	0.900	
Employment <sup>a</sup>	0.040		0.046	0.051		0.052	463	0.317	
Formal employment	Córdoba (City)	0,062	***	0.012	0.074	***	0.015	4,558	0.156
	Córdoba (Province)	0,059	***	0.007	0.071	***	0.008	11,817	0.121
<b>PANEL C. Male</b>									
Labor force participation <sup>a</sup>	0.006		0.029	0.007		0.033	511	0.883	
Employment <sup>a</sup>	0.089	*	0.047	0.111	**	0.054	452	0.350	
Formal employment	Córdoba (City)	0,058	***	0.014	0.069	***	0.017	4,328	0.224
	Córdoba (Province)	0,042	***	0.009	0.050	***	0.011	10,134	0.213

Note:<sup>a</sup> outcomes measured in the follow survey. <sup>b</sup> outcomes from administrative data. Column (1) reports the intention to treat (ITT) effect of the program 12 months after its end. Each coefficient is the corresponding OLS estimate of the parameter  $\tau_{12}$  in equation 3 (note that only for the outcome formal employment the regression jointly estimates the effects for each month, while for outcomes coming from the survey there is only one estimated ITT). Column (4) reports the IV estimates where effective treatment is instrumented with random assignment. All regressions control for size of the firm where the individual chose to do the internship, the number of applications received by the firm, county fixed effects, and the number of applications the individual submitted). Columns (1) and (4) report either robust standard errors (for labor force participation and employment outcomes) or standard errors clustered at the individual level (formal employment outcomes). Column (5) report the mean of the corresponding outcome for the control group. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 5: Long run effects of the program on the probability of being employed in the formal sector and on labor earnings (number of months after the end of the PPP)

	Probability of formal employment							Labor earnings						
	ITT		IV			(5)	ITT		IV			(10)		
	(1)	(2)	(3)	(4)	(6)		(7)	(8)	(9)					
	Coef.	se	Coef.	se	mean	Coef.	se	Coef.	se	mean				
<b>PANEL A. All individuals</b>														
6 months	0.051	***	0.006	0.061	***	0.007	0.151	218.8	***	33.92	286.3	***	38.52	771.6
12 months	0.052	***	0.006	0.063	***	0.007	0.163	330.4	***	41.92	411.7	***	48.37	993.2
18 months	0.050	***	0.006	0.059	***	0.007	0.195	397.5	***	53.8	471.4	***	62.26	1452
24 months	0.042	***	0.006	0.053	***	0.007	0.220	403.6	***	63.3	496.1	***	74.03	1867
30 months	0.043	***	0.007	0.053	***	0.008	0.251	474.8	***	76.05	560.3	***	89.05	2445
36 months	0.040	***	0.007	0.047	***	0.008	0.262	553.3	***	97.39	621.7	***	113.8	3193
42 months	0.046	***	0.007	0.056	***	0.008	0.285	672.7	***	113.6	811.2	***	132.6	4032
<b>PANEL B. Female</b>														
6 months	0.056	***	0.007	0.067	***	0.008	0.110	217	***	38.64	298.5	***	43.08	506.2
12 months	0.059	***	0.007	0.071	***	0.008	0.121	333.9	***	48.48	424.3	***	54.91	672.6
18 months	0.052	***	0.008	0.062	***	0.009	0.149	405.9	***	64.64	471.1	***	73.25	1035
24 months	0.047	***	0.008	0.057	***	0.009	0.165	438.9	***	73.41	514.3	***	84.09	1267
30 months	0.041	***	0.008	0.05	***	0.009	0.193	423.2	***	89.35	474.7	***	102.4	1737
36 months	0.048	***	0.008	0.055	***	0.01	0.200	585.7	***	110.4	601.2	***	126.7	2223
42 months	0.053	***	0.009	0.063	***	0.01	0.222	746.7	***	131.8	826.7	***	150.2	2818
<b>PANEL C. Male</b>														
6 months	0.044	***	0.009	0.050	***	0.011	0.200	219	***	56.15	244.6	***	65.97	1088
12 months	0.042	***	0.009	0.050	***	0.011	0.213	320.6	***	68.9	367	***	82.14	1375
18 months	0.044	***	0.009	0.052	***	0.011	0.251	373.8	***	86.79	437.8	***	103.3	1949
24 months	0.035	***	0.010	0.044	***	0.012	0.285	337.3	***	104.2	427.6	***	124.7	2582
30 months	0.042	***	0.010	0.052	***	0.012	0.321	497.2	***	124.1	605.3	***	148.7	3289
36 months	0.028	***	0.010	0.033	***	0.012	0.335	460.9	***	163	558.8	***	194	4350
42 months	0.034	***	0.010	0.041	***	0.012	0.360	513.8	***	187.7	672.2	***	223.2	5478
<b>PANEL D. Differences Female-Male (p-values)</b>														
6 months	0.238		0.086			0.977		0.269						
12 months	0.126		0.032			0.874		0.302						
18 months	0.478		0.075			0.766		0.344						
24 months	0.422		0.086			0.424		0.332						
30 months	0.980		0.614			0.628		0.846						
36 months	0.139		0.084			0.526		0.667						
42 months	0.131		0.037			0.310		0.555						
No. of observations	1,338,340						1,338,340							

Note: Columns (1) and (6) report the intention to treat (ITT) effect of the program. Each coefficient is the corresponding OLS estimate of the parameter  $\tau_t$  in equation 3 pooling all available periods and including months fixed effects. Columns (3) and (8) report the IV estimates where effective treatment is instrumented with random assignment. Labor earnings of individuals with no formal employment are set equal to 0. All regressions control for size of the firm where the individual chose to do the internship, the number of applications received by the firm, county fixed effects, and the number of applications the individual submitted. Columns (2), (4), (7) and (9) report robust standard errors clustered at the individual level. Columns (5) and (10) report the mean of the corresponding outcome for the control group. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 6: Robustness check: Long run effects of the program on the probability of being employed in the formal sector (number of months after the end of the PPP) for alternative specifications

Months after PPP	ITT (Baseline)			ITT (with covariates)			ITT (IPW logit)		
	Coef.	***	Std. Error	Coef.	***	Std. Error	Coef.	***	Std. Error
Month 1	0.019	***	0.005	0.013	***	0.005	0.017	***	0.005
Month 2	0.030	***	0.005	0.024	***	0.005	0.029	***	0.005
Month 3	0.038	***	0.005	0.034	***	0.006	0.037	***	0.005
Month 4	0.046	***	0.005	0.038	***	0.006	0.045	***	0.005
Month 5	0.053	***	0.006	0.047	***	0.006	0.052	***	0.006
Month 6	0.051	***	0.006	0.046	***	0.006	0.050	***	0.006
Month 7	0.057	***	0.006	0.052	***	0.006	0.054	***	0.006
Month 8	0.058	***	0.006	0.054	***	0.006	0.051	***	0.006
Month 9	0.054	***	0.006	0.050	***	0.006	0.049	***	0.006
Month 10	0.051	***	0.006	0.046	***	0.006	0.049	***	0.006
Month 11	0.053	***	0.006	0.050	***	0.006	0.052	***	0.006
Month 12	0.052	***	0.006	0.049	***	0.006	0.052	***	0.006
Month 13	0.050	***	0.006	0.048	***	0.006	0.050	***	0.006
Month 14	0.050	***	0.006	0.047	***	0.006	0.050	***	0.006
Month 15	0.051	***	0.006	0.048	***	0.006	0.051	***	0.006
Month 16	0.051	***	0.006	0.048	***	0.006	0.051	***	0.006
Month 17	0.050	***	0.006	0.048	***	0.006	0.051	***	0.006
Month 18	0.050	***	0.006	0.050	***	0.006	0.051	***	0.006
Month 19	0.049	***	0.006	0.050	***	0.007	0.049	***	0.006
Month 20	0.046	***	0.006	0.046	***	0.007	0.043	***	0.006
Month 21	0.043	***	0.006	0.042	***	0.007	0.041	***	0.006
Month 22	0.047	***	0.006	0.045	***	0.007	0.046	***	0.006
Month 23	0.044	***	0.006	0.042	***	0.007	0.044	***	0.006
Month 24	0.042	***	0.006	0.040	***	0.007	0.044	***	0.006
Month 25	0.042	***	0.006	0.040	***	0.007	0.043	***	0.006
Month 26	0.039	***	0.006	0.038	***	0.007	0.041	***	0.006
Month 27	0.039	***	0.006	0.039	***	0.007	0.041	***	0.006
Month 28	0.041	***	0.006	0.040	***	0.007	0.043	***	0.006
Month 29	0.039	***	0.006	0.037	***	0.007	0.040	***	0.007
Month 30	0.043	***	0.007	0.040	***	0.007	0.044	***	0.007
Month 31	0.039	***	0.007	0.037	***	0.007	0.040	***	0.007
Month 32	0.045	***	0.007	0.041	***	0.007	0.043	***	0.007
Month 33	0.043	***	0.007	0.039	***	0.007	0.042	***	0.007
Month 34	0.044	***	0.007	0.040	***	0.007	0.044	***	0.007
Month 35	0.038	***	0.007	0.036	***	0.007	0.039	***	0.007
Month 36	0.040	***	0.007	0.038	***	0.007	0.041	***	0.007
Month 37	0.045	***	0.007	0.042	***	0.007	0.045	***	0.007
Month 38	0.040	***	0.007	0.039	***	0.007	0.041	***	0.007
Month 39	0.041	***	0.007	0.038	***	0.007	0.042	***	0.007
Month 40	0.044	***	0.007	0.041	***	0.007	0.044	***	0.007
Month 41	0.045	***	0.007	0.042	***	0.007	0.047	***	0.007
Month 42	0.046	***	0.007	0.044	***	0.007	0.047	***	0.007
No. Of observations	1,338,340			1,338,340			1,338,340		

Table 7: Descriptive statistics of city's share of treated individuals

County	Num. of cities	mean	sd	min	max
CALAMUCHITA	17	0.56	0.27	0.00	1.00
CAPITAL	1	0.35	.	0.35	0.35
COLON	20	0.45	0.09	0.33	0.67
GENERAL ROCA	12	0.34	0.16	0.00	0.67
GENERAL SAN MART	15	0.43	0.19	0.14	1.00
JUAREZ CELMAN	13	0.57	0.24	0.30	1.00
MARCOS JUAREZ	19	0.39	0.17	0.00	0.75
PUNILLA	24	0.60	0.27	0.00	1.00
RIO CUARTO	23	0.24	0.15	0.00	0.50
RIO PRIMERO	23	0.39	0.33	0.00	1.00
RIO SEGUNDO	19	0.24	0.09	0.00	0.40
PTE ROQUE SAENZ	8	0.17	0.08	0.10	0.33
SAN JUSTO	33	0.29	0.19	0.00	1.00
SANTA MARIA	22	0.27	0.26	0.00	1.00
TERCERO ARriba	15	0.22	0.14	0.00	0.60
UNION	25	0.32	0.17	0.14	1.00
Total	289	0.37	0.24	0.00	1.00

Table 8: Correlation between the share of treated individuals at the city level and municipality characteristics

Variables	(1) coeff.	(2) sd	(3) N
Population aged 15-19, 2008	-186.432	(170.624)	278
Population aged 20-24, 2008	-164.651	(157.846)	278
UBN rate 2008	0.027	(0.023)	281
Unemployment rate 2008	-0.002	(0.008)	281
Unemployment rate, population aged 15-24	-0.003	(0.019)	280
Labor informality rate 2008	-0.000	(0.023)	281
Health insurance coverage 2008, population aged 15-24 (%)	-0.049	(0.036)	281
Literacy rate 2008	-0.006	(0.007)	281
Educ. enrollment rate, overall population	0.009	(0.011)	281
Educ. enrollment rate, population aged 15-24	0.036	(0.035)	281
High-school graduates, population aged 20+	-0.006	(0.032)	281
Labor force participation 15-24	-0.054**	(0.026)	280

Note: Column 1 reports the coefficient of an OLS regression of each city characteristics on the share of treated individuals at the city level, controlling for department fixed effects. Robust standard errors in parenthesis. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .



Table 9: Impact of the program on formal employment accounting for displacement effects.

	ITT (city <20%)		ITT (city 20-40%)		ITT (city +40%)		City 20-40%		City +40%		N				
	Coef.	se	Coef.	se	Coef.	se	Coef.	se	Coef.	se					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)				
<b>Panel A. PPP period</b>															
2 months after PPP start							0.000	0.006	-0.009	0.008	14510				
4 months after PPP start							0.008	0.007	-0.009	0.009	14510				
6 months after PPP start							0.013	0.008	0.007	0.011	14510				
8 months after PPP start							0.012	0.009	0.004	0.013	14510				
10 months after PPP start							0.019	*	0.010	0.010	14510				
12 months after PPP start							0.014	0.010	0.017	0.014	14510				
<b>Panel B. Post PPP period</b>															
6 months after PPP end	0.024	0.020	0.051	***	0.007	0.061	***	0.014	0.016	0.012	0.021	0.017	21776		
12 months after PPP end	0.044	**	0.022	0.053	***	0.007	0.054	***	0.015	0.010	0.013	0.028	0.018	21776	
18 months after PPP end	0.043	*	0.024	0.050	***	0.007	0.052	***	0.016	0.021	0.014	0.032	*	0.019	21776
24 months after PPP end	0.024	0.024	0.048	***	0.007	0.031	*	0.017	0.021	0.015	0.038	*	0.020	21776	
30 months after PPP end	0.030	0.025	0.042	***	0.007	0.052	***	0.018	0.024	0.016	0.053	**	0.021	21776	
36 months after PPP end	0.006	0.026	0.044	***	0.007	0.020		0.018	-0.001	0.016	0.029		0.022	21776	
42 months after PPP end	0.040	0.027	0.048	***	0.008	0.042	**	0.018	-0.002	0.017	0.006		0.022	21776	

Panel A reports coefficients of the corresponding OLS estimates of equation 5, which is run on the subsample of individuals in the control group only. Panel B reports coefficients of the corresponding OLS estimates of equation 4, which is run on the whole sample. All regressions control for size of the firm where the individual chose to do the internship, the number of applications received by the firm, county fixed effects, the number of applications the individual submitted, and city of residence characteristics in a baseline period (2008) (population, poverty rate, unemployment rate and labor informality rate). Robust standard errors are reported. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 10: Mechanisms: Skills development (Survey sample)

	All				Female				Male			
	(1) Mean control	(2) Diff.	(3) se	(4) N	(5) Mean control	(6) Diff.	(7) se	(8) N	(9) Mean control	(10) Diff.	(11) se	(12) N
<b>A. “Cognitive” skills</b>												
Cognitive (conceptual verbalization)	7.395	-0.050	(0.151)	1.017	7.632	-0.197	(0.187)	512	7.167	0.057	(0.237)	505
Seconds to complete survey	2514.191	-137.182	(108.477)	1.019	2498.415	-144.889	(158.916)	512	2529.257	-114.854	(148.321)	507
<b>B. “Non-cognitive” skills</b>												
Negative strategies for stress management	12.638	-0.000	(0.132)	1.019	12.585	0.045	(0.187)	512	12.689	-0.068	(0.186)	507
Resolute strategies for stress management	8.622	0.114	(0.120)	1.019	8.476	0.144	(0.173)	512	8.761	0.085	(0.166)	507
Social strategies for stress management	8.106	0.157	(0.131)	1.019	8.212	0.165	(0.194)	512	8.005	0.134	(0.177)	507
Personal projects	8.376	-0.066	(0.069)	1.019	8.491	-0.046	(0.093)	512	8.266	-0.097	(0.101)	507
Self-control	8.157	0.070	(0.070)	1.019	8.217	0.082	(0.095)	512	8.099	0.047	(0.101)	507
Self-efficacy	8.949	0.064	(0.100)	1.019	8.877	-0.027	(0.143)	512	9.018	0.153	(0.139)	507
Time planning	17.392	0.037	(0.184)	1.019	17.759	0.169	(0.262)	512	17.041	-0.132	(0.257)	507

Note: Columns (2), (6) and (10) report the intention to treat (ITT) effect of the program. Each coefficient is the corresponding OLS estimate of the parameter  $\tau_t$  in equation 3, where the outcome variables are measures of skills (Brenlla, 2014). All regressions control for size of the firm where the individual chose to do the internship, the number of applications received by the firm, county fixed effects, and the number of applications the individual submitted. Columns (1), (5) and (9) report the mean of the corresponding outcome for the control group. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 11: Long run effects of the program on the probability of being employed in the formal sector by level of education (number of months after the end of the PPP)

		Probability of formal employment ITT		
		(1)	(2)	(3)
		Coef.	Std. Error	Mean control group
<b>PANEL A. Edu Low (High School dropout)</b>				
6 months	0.047	***	0.009	0.137
12 months	0.054	***	0.009	0.151
18 months	0.053	***	0.009	0.173
24 months	0.045	***	0.010	0.195
30 months	0.046	***	0.010	0.216
36 months	0.050	***	0.010	0.226
42 months	0.055	***	0.010	0.237
<b>PANEL B. Edu high (High School complete)</b>				
6 months	0.052	***	0.011	0.179
12 months	0.051	***	0.011	0.191
18 months	0.048	***	0.012	0.227
24 months	0.042	***	0.012	0.251
30 months	0.034	***	0.012	0.286
36 months	0.033	***	0.013	0.295
42 months	0.038	***	0.013	0.319
<b>PANEL D. Differences Edu low-Edu high (p-values)</b>				
6 months			0.708	
12 months			0.863	
18 months			0.721	
24 months			0.848	
30 months			0.444	
36 months			0.277	
42 months			0.292	

Note: Column (1) reports the intention to treat (ITT) effect of the program. Each coefficient is the corresponding OLS estimate of the parameter  $\tau_t$  in equation 3. The regression control for size of the firm where the individual chose to do the internship, the number of applications received by the firm, county fixed effects, and the number of applications the individual submitted). Column (2) reports standard errors clustered at the individual level. Columns (3) reports the mean of the corresponding outcome for the control group. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 12: Long run effects of the program on the probability of being employed in the formal sector by cognitive skills level (number of months after the end of the PPP)

		Probability of formal employment ITT		
		(1)	(2)	(3)
		Coef.	Std. Error	Mean control group
<b>PANEL A. Cognitive Low</b>				
6 months		0.024	0.03	0.1569
12 months		0.044	0.027	0.1131
18 months		0.042	0.032	0.1715
24 months		-0.02	0.033	0.2226
30 months		0.002	0.036	0.2628
36 months		0.034	0.036	0.2628
42 months		0.015	0.038	0.3139
<b>PANEL B. Cognitive high</b>				
6 months		0.039	0.041	0.165
12 months	**	0.092	0.039	0.120
18 months	**	0.087	0.044	0.190
24 months		0.075	0.047	0.241
30 months	**	0.110	0.048	0.266
36 months	**	0.106	0.049	0.278
42 months	**	0.100	0.05	0.304
<b>PANEL D. Differences low-Edu (p-values)</b>				
6 months			0.772	
12 months			0.299	
18 months			0.404	
24 months			0.097	
30 months			0.072	
36 months			0.237	
42 months			0.174	

Note: Column (1) reports the intention to treat (ITT) effect of the program. Each coefficient is the corresponding OLS estimate of the parameter  $\tau_t$  in equation 3. The regression control for size of the firm where the individual chose to do the internship, the number of applications received by the firm, county fixed effects, and the number of applications the individual submitted). Column (2) reports standard errors clustered at the individual level. Column (3) reports the mean of the corresponding outcome for the control group. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

## 9 Appendix: Bounds calculation for the causal effect of the program on wages

The effect of the program on wages in Figure 2 are estimates of the following expression.

$$\text{Wage effect} = \mathbb{E}[\text{wage}|L = 1, D = 1] - \mathbb{E}[\text{wage}|L = 1, D = 0] \quad (6)$$

However this is not a causal effect. In this appendix we show (following Attanasio et al. (2011)) how we bound the causal effect of interest (productivity effect).

Let  $L(j)$ , for  $j \in \{0, 1\}$ , be the formal working status given treatment assignment. The population can be divided in four groups:

- Always takers:  $L(1) = 1, L(0) = 1$
- Never takers:  $L(1) = 0, L(0) = 0$
- Compliers:  $L(1) = 1, L(0) = 0$
- Defiers:  $L(1) = 0, L(0) = 1$

Assuming monotonicity, we rule out the existence of defiers. Additionally, given treatment randomization, the size of each group is independent of treatment status. Given these two assumptions, we can write the two terms of the wage effect as

$$\begin{aligned} \mathbb{E}[\text{wage}|L = 1, D = 1] = & \mathbb{E}[\text{wage}|\text{complier}, D = 1] \times \frac{\text{Pr}[\text{complier}]}{\text{Pr}[\text{complier}] + \text{Pr}[\text{Always taker}]} \\ & + \mathbb{E}[\text{wage}|\text{Always taker}, D = 1] \times \frac{\text{Pr}[\text{Always taker}]}{\text{Pr}[\text{complier}] + \text{Pr}[\text{Always taker}]} \end{aligned}$$

$$\mathbb{E}[\text{wage}|L = 1, D = 0] = \mathbb{E}[\text{wage}|\text{Always taker}, D = 0]$$

Lets call

$$K = \frac{Pr[complier]}{Pr[complier] + Pr[Always\ taker]}$$

Then

$$\begin{aligned} \text{Wage effect} &= \mathbb{E}[wage|L = 1, D = 1] - \mathbb{E}[wage|L = 1, D = 0] \\ &= \mathbb{E}[wage|complier, D = 1] \times K + \mathbb{E}[wage|Always\ taker, D = 1](1 - K) \\ &\quad - \mathbb{E}[wage|Always\ taker, D = 0] \end{aligned}$$

Adding and subtracting the term  $\mathbb{E}[wage|complier, D = 0] \times K$  and rearranging we get

$$\begin{aligned} \text{Wage effect} &= \left\{ \underbrace{E[wage|complier, D = 1] - E[wage|complier, D = 0]}_{\text{Causal effect on compliers}} \right\} \times K \\ &+ \left\{ \underbrace{E[wage|Always\ taker, D = 1] - E[wage|Always\ taker, D = 0]}_{\text{Causal effect on Always takers}} \right\} \times (1 - K) \\ &+ \left\{ \underbrace{E[wage|Complier, D = 0] - E[wage|Always\ taker, D = 0]}_{\text{Baseline wage difference between compliers and always takers}} \right\} \times K \end{aligned}$$

Then, we can estimate the causal effect of interest (productivity effect), which is a weighted average of the causal effect of the program on the wages of compliers and always takers, as

$$\text{Causal effect} = \text{wage effect} - \{E[wage|Complier, D = 0] - E[wage|Always\ taker, D = 0]\} \times K$$

Note that the wage effect and  $K$  can be estimated from the data.

$$Pr[Always\ taker] = Pr[L = 1|D = 0] \quad (7)$$

$$Pr[complier] + Pr[Always\ taker] = Pr[L = 1|D = 1]$$

$$\Rightarrow Pr[complier] = Pr[L = 1|D = 1] - Pr[L = 1|D = 0]$$

Hence,

$$\begin{aligned} K &= \frac{Pr[complier]}{Pr[complier] + Pr[Always\ taker]} \\ &= \frac{Pr[L = 1|D = 1] - Pr[L = 1|D = 0]}{Pr[L = 1|D = 1]} \end{aligned}$$

However, the baseline wage difference between compliers and always takers is not observed. This difference is the selection effect partially explaining the wage effect. This term can be bounded by using the mean wage in the tenth and ninetieth percentile for those in the control group who have formal employment, assuming that it could be  $\pm \{\mathbb{E}[wage(p_{0.90,it})] - \mathbb{E}[wage(p_{0.10,it})]\}$ . However, the lower bound estimate with this proxy would be extremely conservative. One can assume that the average wage of the always takers without treatment is at least as large as the average wage of the compliers, resulting in the following bounds

$$\text{Lower bound Causal effect} = \text{wage effect}, \quad (8)$$

$$\text{Upper bound Causal effect} = \text{wage effect} - \{\mathbb{E}[wage(p_{0.90,it})] - \mathbb{E}[wage(p_{0.10,it})]\} \times K. \quad (9)$$