I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17254

The Impact of Labor Intermediation and Training in High Informality Contexts. Evidence from Paraguay.

Nicolás Campos
Miguel Chalup
Oscar A. Mitnik
Manuel Urquidi

AUGUST 2024

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17254

The Impact of Labor Intermediation and Training in High Informality Contexts. Evidence from Paraguay.

Nicolás Campos
Inter-American Development Bank

Miguel Chalup
Arizona State University

Oscar A. Mitnik
Inter-American Development Bank and IZA

Manuel Urquidi
Inter-American Development Bank

AUGUST 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Impact of Labor Intermediation and Training in High Informality Contexts. Evidence from Paraguay.*

We provide quasi-experimental estimates of the impact of reforming public training programs offered in Paraguay on formal employment. The *Programa de Apoyo a la Inserción Laboral* (PAIL) revamped training program design in the country by offering courses aligned with the needs of the private sector, enhancing non-cognitive skills, and combining practical work within companies with classroom training. We combine administrative records—which contain detailed information on the employment history and characteristics of all formal workers in Paraguay—with an empirical strategy based on extensions of difference-in-differences models and synthetic difference-in-differences. We find that the probability of obtaining formal employment for women and men increases by 11 percentage points. Even two years after participating, the program has a lasting impact on women, an aspect not observed for men. Additionally, the program's impact is positive only in the metropolitan area of Asunción; the program is less effective in areas far from the urban center, especially for men. The observed results suggest that supply-side interventions are ineffective if no formal jobs are available for the beneficiaries.

JEL Classification: J24, J38, J68

Keywords: public training programs, active labor market policies, formal employment

Corresponding author:
Oscar A. Mitnik
Inter-American Development Bank
1300 New York Avenue, NW
Washington, DC 20577
USA
E-mail: omitnik@iadb.org

* We thank Laura Di Capua, Guillermo Goméz, Laura Guzmán, Monica Pacheco, María Delina Otazú, Solange Sardán, Liliana Serrate, and Nicolás Quiroga for their comments and the information provided for different sections of the paper. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.

I. Introduction

The National System of Labor Training and Vocational Education (SINAFOCAL) and the National Service for Professional Promotion (SNPP), both under the Ministry of Labor, Employment, and Social Security (MTESS), are the institutions responsible for implementing labor training policies in Paraguay.[1] The SNPP promotes and develops professional training at all levels and in all economic sectors, while SINAFOCAL coordinates, controls, and finances labor training by contracting private providers to conduct training.[2] Both institutions are funded through a 1% employer contribution on the wages of private sector workers, a budget allocation that amounted to $36 million (USD) in 2023.[3] Of this total, $7.9 million was allocated to SINAFOCAL and $23 million to the SNPP. The resources assigned to labor training activities are significant. Both allocations represent 70% of the MTESS budget and surpass allocations to other relevant institutions. For instance, the combined budget allocations for these two institutions exceed by 40% the resources allocated to the Office of the Comptroller General of the Republic.[4]

Training programs for employment, within the framework of Active Labor Market Policies, are valuable tools for promoting labor integration and facilitating access to quality jobs in Latin America and the Caribbean (Card et al., 2010; Kluve & Rani, 2016; Mitnik et al., 2016; Card et al., 2018; Escudero, 2018; Escudero et al., 2019; Carranza & McKenzie, 2024). However, for these programs to be effective, they require an appropriate design and effective implementation by the public sector. Indeed, before the modernization process in recent years, public training programs in Paraguay needed to improve their success. For example, the contracted professional training institutions should have offered quality training in line with the demands of the productive sector. They used outdated training programs, lacked an adequate registry of training providers, and had limited capacity for investment in technological infrastructure (Alaimo & Tapia, 2014).

These deficiencies, among other reasons, motivated the creation of the *Programa de Apoyo a la Inserción Laboral* (PAIL, Labor Insertion Support Program in English). This program included improvements to labor intermediation processes, modifications to the training program design by offering courses aligned with the needs of the private sector, enhancements to non-cognitive skills,

[1] On January 1, 2014, Law No. 5.115/2013 came into force, creating the MTESS to replace the Ministry of Justice and Labor. On the other hand, Law No. 1652, enacted in 2000, and Law No. 1253 of 1971 created SINAFOCAL and SNPP, respectively.

[2] Despite differences in their administrative roles, both institutions aim to "provide their beneficiaries with opportunities for training and education in their various modalities" (SINAFOCAL) and "carry out training activities primarily considering the national development process, providing immediate responses to the labor market in terms of training" (SNPP). Both work in coordination with the General Directorate of Employment (DGE), which depends on MTESS.

[3] They also have other funding sources, such as national allocations and subsidies, local government revenues, private sector contributions, and international cooperation. For more details, see: Law No. 1652.

[4] The budget allocation for SINAFOCAL was 57,970,621,392 PYG, while for the SNPP, it was 173,993,114,175 PYG. We used the exchange rate the Central Bank of Paraguay reported in January 2023 (1 USD = 7,305 PYG) for conversion. The General Budget of the Nation is available at Law No. 7050.

and the combination of practical work within companies with classroom training (Urquidi et al., 2023). PAIL focused on SINAFOCAL, considering adjusting courses to be tendered according to private sector demand was a more straightforward first step than adjusting the SNPP's teacher roster and infrastructure. The considerable investment of public resources in intermediation and labor training, along with the recent modernization of the corresponding programs and institutions, motivates us to examine the impact of the new intermediation processes and training course design on the labor market outcomes for its beneficiaries.

This article contributes to the literature by evaluating a developing country with high levels of informality using administrative data and novel developments on difference-in-differences and synthetic control models. We also address potential issues highlighted by recent research, which has identified several weaknesses in the two-way fixed effects (TWFE) estimation strategy when (i) there are more than two time periods, (ii) there is variation in treatment timing, and (iii) there is treatment effect heterogeneity (De Chaisemartin & d'Haultfoeuille, 2020; Freyaldenhoven et al., 2021; Goodman-Bacon, 2021; Sun & Abraham, 2021; Borusyak et al., 2021; Athey & Imbens, 2022; Rambachan & Roth, 2023; Callaway, 2023; Roth et al., 2023). Additionally, we report the ineffectiveness of the program outside metro areas, highlighting that supply-side interventions can be ineffective if no formal jobs are available for the beneficiaries. These results offer valuable insights for designing new employment promotion programs that consider this source of heterogeneity.

In this study, we measure the impact of training programs accompanied by labor intermediation services, implemented by MTESS on formal employment. Formal employment is identified by contributions by employed individuals to social security.[5] We chose a quasi-experimental approach due to the non-random nature of participant assignment to training. To identify the program's causal effect, we combined administrative records—which contain detailed information on the employment history and characteristics of all formal workers in Paraguay—with an empirical strategy based on extensions of difference-in-differences models and synthetic difference-in-differences. The main results are derived from the estimator by Callaway & Sant'Anna (2021). This estimator reports an unbiased treatment effect, is robust to treatment impact heterogeneity across groups or over time, and is suitable for staggered treatment contexts. Additionally, we provide a set of alternative estimators to verify the robustness of our main results. Specifically, we calculate a conventional TWFE difference-in-differences estimator with individual and time-fixed effects and relax the parallel trends assumption by reporting the program impact using the synthetic difference-in-differences estimator by Arkhangelsky et al. (2021).

The results indicate that participation in the training program has positively impacted the formal labor insertion of women and men; participation in the training programs increased formality by 11 percentage points for both groups. These results support the program's effectiveness in reducing

[5] As explained in section 3 we rely on administrative employer-worker data that identifies each month all workers in Paraguay who have contributed to the social security (pension) system in the month.

barriers to formal labor market entry in Paraguay. We highlight three additional findings. First, the program's impact is more significant for women in terms of formal labor insertion over time, and the impact on women does not decrease after two years from training participation. Second, the overall impact is mainly driven by the impact on individuals who live in metropolitan areas; the program's effectiveness outside urban areas is null for men and small for women. Third, the program's impact decreased but was not nullified during the pandemic years (2020 − 2021). The impact is null for men, but it remains positive for women. This finding indicates that the program's effectiveness is higher during favorable economic periods, yet it remains effective even in adverse economic circumstances, at least for women.

The PAIL program, as part of a modernization of the training and labor intermediation system in Paraguay, promoted institutional advances that could have influenced the results. The training courses were designed to consider the needs of the private sector and provide both general skills and specific training outside the classroom, plus labor intermediation support services. Our evidence suggests that, as a result, the beneficiaries faced reduced barriers to labor market entry, leading to greater formality. These findings align with the literature, indicating that well-designed training programs with appropriate incentives can generate lasting improvements in the labor outcomes of beneficiaries, especially in groups facing more significant restrictions in entering the labor market.

Some limitations of this study should be noted. First, our estimates cannot differentiate between the impact of labor intermediation and training as treated individuals received both types of services; therefore, they reflect the comprehensive impact of the PAIL program. Furthermore, this evaluation measures the impact of a newly designed training program, not the impact of all training programs offered by these institutions. Second, our outcome data only captures those individuals who access the formal sector; our estimates do not capture workers operating exclusively in the informal sector. However, our results likely represent a lower bound of the program's true impact, as evidence indicates that active labor market policies are even more effective in labor markets with high levels of informality. In Paraguay, the high level of labor informality is one of the most significant challenges faced by its employment policies (Abuelafia et al., 2020). It should be noted that, from 2010 to 2021, according to data from the Labor Markets and Social Security Information System (SIMS), the average percentage of workers in informal employment in Paraguay reached 77%, one of the highest rates in Latin America and the Caribbean.

The remainder of the paper is organized as follows: Section II provides some background on the program and details the institutional context in which the program was implemented. Section III describes the data sources used in the analysis. Section IV discusses the empirical strategy. Section V presents the main results and the robustness analysis. Finally, Section VI discusses the results and concludes. Two appendices follow these sections, including additional figures and tables.

II. Institutional Context

Institutions. The MTESS is the institution responsible for implementing employment policies and the functioning of the labor market according to the requirements of the Executive Branch. The General Directorate of Employment (DGE) is the body whose mission is to design, execute, and supervise employment policy in the country. Regarding labor training and vocational education, the MTESS has two departments with specific competencies: SINAFOCAL and SNPP. SINAFOCAL was created to ensure training opportunities, while SNPP focuses on offering immediate responses to the labor market regarding training. SINAFOCAL is responsible for regulating and supervising the actions of SNPP and the private institutes that offer training. SNPP provides courses with its teaching staff in its facilities, while SINAFOCAL outsources them. The fact that SINAFOCAL does not use its facilities means it does not require permanent facilities or instructors for course provision. The system enjoys technical and economic autonomy (Casalí, 2016).

Labor Insertion Support Program. The Government of Paraguay implemented the PAIL program with the Inter-American Development Bank (IDB) support. This program emerged from a collaborative analysis between the public and private sectors, identifying an imbalance between institutional offerings in training and labor intermediation and the demands of employers and workers (Alaimo & Tapia, 2014; Casalí, 2016; Urquidi et al., 2023). For example, SINAFOCAL needed better systems to identify the qualified labor needs of the productive sector, as evidenced by its inability to meet employment demands. Additionally, training was predominantly technical, neglecting basic and transversal skills.[6] Most courses focused on classroom instruction without prioritizing practical training in real work contexts.

The PAIL program consists of two elements: labor intermediation and training. Labor intermediation aims to improve access to formal employment by facilitating contact between employers and candidates and providing accurate information on job supply and demand. Training offers technical and basic skills instruction, combining classroom and practical workplace learning. Both components of the PAIL program were designed in close collaboration with the private sector to optimize their effectiveness. As a result, the training courses align with market demands, reflecting the specific needs of the private sector.

The PAIL program was structured around two main components: (i) expansion and consolidation of the DGE and the Employment Office Network, and (ii) support for services to insert young people into the labor market. The first component addressed infrastructure improvements, strengthened DGE's information systems, and reinforced technical training. This aspect is particularly relevant in our evaluation, as it strengthens institutional capacities to offer more effective training programs. DGE and SINAFOCAL personnel updated their supervision and administration skills and SINAFO-

[6] Companies' reluctance to invest in training, particularly in transversal skills, exacerbated this problem. Companies were less willing to finance such training since workers can apply these skills to any job.

CAL's courses through training. They also implemented actions to connect with the productive sector, promoting gender awareness to ensure equal opportunities for women in non-traditional roles.

Additionally, the PAIL program redefined the identification of productive sectors with growth potential. It established a more efficient design, application, and systematization of quality and relevance measurement instruments for training institutions. These changes significantly enhanced the effectiveness of PAIL training programs. Evidence shows that demand-oriented training, the inclusion of significant guidance and labor intermediation components, and the incorporation of socio-emotional skills development modules are all success factors in implementing labor training and education programs (Heckman & Kautz, 2012; González-Velosa et al., 2012). Moreover, recent evidence suggests that adjusting expectations regarding the jobs and wages workers can realistically access after participating in these programs is essential for improving the long-term effectiveness of these programs (Alfonsi et al., 2022).

The second component improved access to employment opportunities and created new training modalities. These innovations included implementing quality controls, a rigorous review of current training interventions, and establishing collaboration between the DGE and SINAFOCAL. Consequently, SINAFOCAL is committed to training young people in the program and aligning the training with the demands of the productive sector. The labor training program focused on young people aged 18 to 29 with an educational level of 9 to 12 years. The program prioritized vulnerable groups, such as women and heads of households, to ensure that more than 50% of the beneficiaries were women. The program covered transportation costs and provided refreshments to encourage youth participation and reduce dropout rates. For women with children, the program offered financial support for childcare during training sessions. Although the PAIL program established specific age and geographical criteria in its initial design, admitting individuals who did not strictly meet these requirements was possible.

The program introduced multiple innovations, notably in the collaboration plan between the DGE and the private sector. This plan aimed to strengthen dialogue with companies to understand their training and human resources needs. In this framework, meetings were organized with key sector representatives, such as the Paraguayan Industrial Union (UIP) and other sectoral chambers. These meetings allowed the validation and adjustment of curricular programs, ensuring that SINAFOCAL and SNPP training programs aligned with market demands. An example of the effectiveness of this alliance was seen in the management and financing of some training sessions. For instance, courses held in slaughterhouses focused on Kosher meat export. Through a collaborative agreement between the MTESS and the slaughterhouse, the latter provided the animals for training and marketed the resulting meat. This scheme represented significant savings in the traditional costs of such training and eliminated logistical challenges associated with acquiring and managing the animals by the training entity.[7]

III. Data and Descriptive Statistics

Administrative Records. The anonymous data provided by the MTESS allows us to observe the labor history of formal workers in Paraguay, including their monthly earnings, hours worked, participation in the formal labor force, participation in training programs, and demographic characteristics such as occupation type, gender, age, marital status, indigenous status, geographic location, among others. The administrative records come from four institutions: SNPP, SINAFOCAL, the Employer-Worker Registry (REOP), and the Labor Intermediation Service Portal (ParaEmpleo). These institutions supply three interrelated databases—Employment Exchange, Identity, and Eportal—all owned by the MTESS (see Figure A1).

La Bolsa de Empleo, managed by the MTESS, is a portal that facilitates labor intermediation by collecting information on supply and demand in the labor market. It is divided into job seekers, companies, and vacancies. Job seeker data includes personal, educational, and employment information. Company data details the location and economic activity of companies posting vacancies. Vacancy data provides information on job offers, detailing job requirements and specific job characteristics. *Identidad*, which collects information from SINAFOCAL and SNPP, is organized into two main categories: courses and students. The course category details minimum requirements, subject matter, start and end dates, fees, and course duration. The student's category stores information about individuals enrolled in the courses, including gender, number of children, and head of household status. *Eportal* collects information from the REOP and tracks the monthly labor history of all formal workers in Paraguay. It provides details on their current employment status, economic sectors they have worked in, income information, hours worked, and type of occupation. By combining this data with information on training courses, we can trace the labor history of those who have participated in such programs.

A limitation of the labor history data is that it only captures those who occupy or have occupied formal employment. When constructing the formality variable, we assume that workers are only in formal employment if they appear in the records at least once during the quarter. Thus, it implies the limitation of not knowing whether a worker is out of the labor market or has a job in the informal sector.

Database Construction. In the original raw data, we had 44,232,240 observations for the period 2010 - 2022, with 3,402,480 yearly observations. There were 6,121 treated units (954,876 observations) and 277,419 untreated units (43,277,364 observations). There were 3,613 treated women

[7] The training model did not involve additional costs for the slaughterhouse, as the meat processed during the sessions was later sold profitably. Although the meat did not receive Kosher certification due to handling inaccuracies by the students, it was sold in other markets.

and 2,508 treated men. The years between 2010 and 2015 have a low proportion of observations because the system became fully operational in 2016. Based on this fact, we use data only from 2016 to 2021. We lost some treated observations because they graduated at the beginning of 2016, so we could not build a proper control group for them. We are focusing on individuals who are in the age range of 18 to 30 years. Considering this restriction, we lost some observations of treated individuals, which correspond to those who do not have an age record and those who are outside the range. Finally, we did not consider a small group of 40 individuals who participated in a course with different characteristics from the others; therefore, we lost some treated units. After incorporating all these assumptions, we aggregate data by quarter. The database used for the analysis contains 6,625,320 observations, with 4,476 treated units and 271,579 untreated units. Of the treated units, 1,797 are men and 2,679 are women. The panel is fully balanced and represents the period from 2016 to 2021.

Our primary outcome variables correspond to formal employment. We construct this variable from income data. Formality is assigned a value of 1 if the individual has reported an income at any point during the quarter. To maintain data quality and avoid the presence of outliers, we winsorize the income data by replacing incomes above the 99th percentile with this value.

Sample Characteristics. In Table A1, we characterize the main outcome variable and individual characteristics of the units that comprise our database. We note that only 29% are formal employees, consistent with labor informality statistics for Paraguay. Regarding individual characteristics, the average age is 21.7 years, and the median is 21 years, indicating that our sample mainly comprises younger individuals. The geographical distribution shows a high concentration in the Metropolitan Area (46%).[8] Women predominate in our sample, representing 70% of the data. When analyzing the statistical distribution of the outcome variables, we observe that formal employment economic units are characterized by a greater concentration of data near zero, as shown in Figure A2. However, when we analyze PAIL beneficiaries before and after the program (panels (b) and (d) of Figures A2), a change in the employment distributions is detected: an increase in the percentage of formal employed individuals and a decrease in the percentage of unemployed individuals.

Treatment Variable: graduation from a training program. The treatment variable in our study reflects having graduated from one of the PAIL courses. The treated group consists of 4,476 individuals and 107,424 observations. We note that the treatment is staggered, i.e., the units adopt the treatment at different points in time (see Figure A3). 98% of the units do not receive treatment at any period (see Figure A3, panel (a)), while 2% receive it at some point (see Figure A3, panel (b)). The highest concentration of events occurs in the first quarter of 2021 (1,360 units), the fourth quarter of 2018 (1,264 units), the fourth quarter of 2017 (1,020

[8] Metro area includes *Asunción* along with ten districts from the Central Department. These districts are Capiatá, Luque, San Lorenzo, Lambaré, Fernando de la Mora, Limpio, Ñemby, Mariano Roque Alonso, Villa Elisa, and San Antonio.

units), and the fourth quarter of 2019 (832 units) (see Figure A3, panel (c) and Table A2). The cumulative frequency up to the fourth quarter of 2019 is 70% (see Figure A3, panel (d)). Finally, most units treated are concentrated in Central (2,139) and *Asunción* (888). We have a group of 1,274 units without information about their geographic location; we assume these units are outside the metropolitan area (see Table A3). This dataset presents a hybrid structure, with variability in event dates among treated units and a higher percentage of units that never received treatment (98%).

The number of trained individuals is concentrated in five courses: *Debt Collection Skills* (1,273 units), *Customer Service* (1,052 units), *Sales Skills* (683 units), *Oratory - Digital Marketing - Sales - Customer Service* (185 units), and *Community Manager* (146 units). These courses account for 74.5% of the beneficiaries, totaling 3,339 trainees (see Table A4). The gender distribution is not homogeneous across the courses. For example, the *Beverage Preparation*, *User Support*, *Commercial Negotiation*, and *Treasury and Banking Management - Cashier* courses have a reasonably balanced gender distribution. On the other hand, the *Basic Home Electricity*, *Residential Split Air Conditioning*, and *Automobile Air Conditioning Maintenance* courses are exclusively male, while the *Secretariat, Writing, and Spelling* and *Convenience Store Assistant (Minimarket)* courses are exclusively female. On average, each course has 119 students. The course with the highest number of students is *Debt Collection Skills* with 1,273 students, while the *Convenience Store Assistant (Minimarket)* course has the fewest students, with only 6.

Differences Between Groups and Pre-Treatment Balance. Tables A5 and A6 show the descriptive statistics for the treatment and control group considering all the periods. Table A7 compares treatment and control group average characteristics considering only the pre-treatment period for the treated group. When comparing the treated and control groups before treatment, we observe statistically significant differences in almost all observable characteristics and outcome variables (third column, Table A7). Program beneficiaries tend to work fewer hours and show lower levels of labor formality (on average, 16% compared to 29% for untreated units). Regarding their characteristics, they are slightly younger than those who did not participate in the program and have a lower percentage of women. Geographically, treated units are concentrated in metropolitan areas (62% vs. 46%). We also observed that the differences between the individuals who received the treatment and those in the control group are significant overall.

IV. Empirical Strategy

Identification strategy. The main results of this paper come from the estimator of Callaway & Sant'Anna (2021) (DD_{CS}). We supplemented these analyses with event studies, considering some recommendations from Freyaldenhoven et al. (2021) and Miller (2023). We describe the context of our evaluation based on a conventional $TWFE$ model to explain its limitations. Subsequently, we discuss the sensitivity of the conventional $TWFE$ estimator to treatment impact heterogeneity and then justify the choice of the Callaway & Sant'Anna (2021) estimator as our preferred estimator. Finally, we discuss the validity of the parallel trend assumption and the main robustness checks.

The equations (1) and (2) present the models to be estimated within the framework of a conventional $TWFE$ model in their static and dynamic versions, respectively.[9]

$$Y_{it} \;=\; \nu_i \;+\; \rho_t \;+\; \beta \cdot D_{it} \;+\; X_{it} \cdot \gamma \;+\; \varepsilon_{it} \tag{1}$$

$$Y_{i,t} \;=\; \nu_i \;+\; \rho_t \;+\; X_{it} \cdot \gamma \;+\; \sum_{e=-K}^{-2} \delta_e^{\text{anticip}} \cdot D_{i,t}^e \;+\; \sum_{e=0}^{L} \beta_e \cdot D_{i,t}^e \;+\; v_{i,t} \tag{2}$$

Y_{it} denotes the individual's outcome variable i in the period t. The outcome variable is formal employment. The variable D_{it} is dichotomous and takes the value of one when beneficiaries end their training program. v_i represents an individual fixed effect, while ρ_t is a time-fixed effect. Incorporating individual fixed effects allows controlling for time-invariant but variable characteristics across individuals. In contrast, the time-fixed effect ensures that the results are not solely attributed to a time trend. X_{it} represents the set of control variables: age, age squared, marital status, indigenous origin, a dichotomous variable indicating if they have children and economic sector fixed effects. We run the model separately, considering gender and metropolitan areas. ε_{it} is the standard error term. In the equation (1), the coefficient $\hat{\beta}$ is obtained by Ordinary Least Squares (OLS). As usual, standard errors are clustered at the individual level (treatment level), allowing for serial correlation (Bertrand et al., 2004). In the equation (2), $D_{i,t}^e = \mathbb{1}\{t - G_i = e\}$ is an indicator variable that the unit i is at e periods of time away from the initial treatment t. K and L are positive constants. The variable of interest, in this case, corresponds to $\{\beta_e : e \geq 0\}$; these parameters represent the program's causal effect on the outcome variable Y_{it} e periods after exposure to the treatment. Estimating this dynamic specification, often referred to as event studies, helps understand how the impact of participating in the training program changes with the length of exposure to the treatment.

The identification of the causal impact of the program rests on three assumptions. The first is the assumption of parallel trends. Parallel trends assume that, in the absence of the program, the beneficiaries' formal employment trajectories would have followed a course identical to those who did not receive the benefit. The second assumption, the non-anticipation assumption, holds that if a unit does not receive treatment in period t, its outcome is not influenced by the possibility of receiving treatment in future periods. Non-anticipation implies that the treatment has no causal effect before people end the training courses, and it is mandatory to complete the program to effectively increase

[9] This notation is derived from Callaway & Sant'Anna (2021).

the likelihood of accessing formal employment. Finally, it is assumed that the average impact of the treatment is constant across treated units and over time. Under these assumptions, it is possible to state that $\hat{\beta}$ and $\hat{\beta}_e$ acquire a causal interpretation. That is, the observed results are a direct consequence of participation in the training courses and are unrelated to other contemporaneous factors.

Recent literature on $TWFE$ models indicates that it is necessary to be cautious in interpreting the parameter estimates as a causal effect, especially if there are more than two time periods, there is variation in treatment timing, and there is treatment effect heterogeneity (De Chaisemartin & d'Haultfoeuille, 2020; Sun & Abraham, 2021; Borusyak et al., 2021; Callaway & Sant'Anna, 2021; Athey & Imbens, 2022; Rambachan & Roth, 2023; Callaway, 2023; Roth et al., 2023). Moreover, even if we consider that the above problems do not exist, there may be additional issues in staggered treatment contexts that induce bias in the estimates or make it difficult to interpret the estimator (Goodman-Bacon, 2021). Based on these facts, we choose the estimator proposed by Callaway & Sant'Anna (2021) as it is robust to these problems.

TWFE Estimator Sensitivity. To diagnose the presence of heterogeneity in treatment impact, we calculate the "negative weights" based on De Chaisemartin & d'Haultfoeuille (2020). This calculation is based on the fact that the estimator $\hat{\beta}_{TWFE}$ is a weighted average of the average treatment impact for each treated cell, where the weights can be negative. Formally, $E[\widehat{\beta}] = E\left[\sum_{g,t} W_{gt}\Delta_{g,t}\right]$, where W_{gt} are the weights that add up to one and $\Delta_{g,t}$ is the average impact of the treatment on group g in period t. Calculating negative weights is justified since $E[\hat{\beta}]$ may be negative even if all average treatment impacts are positive. Additionally, negative weights are more likely to be assigned to periods with a higher proportion of treated groups or groups treated over many periods. Negative weights pose a problem when the treatment impact differs between periods with many versus few treated groups (as in staggered treatment adoption) or between groups treated for many periods versus few periods. We discuss the results in Section 5.4.

Callaway & Sant'Anna (2021) estimator.[10] This estimator is suitable for our context since it avoids the presence of negative weights and it allows the researcher to specify how the effects are weighted across cohorts (e.g., proportional to the size of the cohort) instead of being determined by OLS, where the weighting is proportional to the variance of the treatment variable. Additionally, it clearly indicates which units are used as the control group, which is especially useful when adopted staggered treatment. The design of this estimator includes a series of aggregations of the average impact of the treatment, among which we can mention the time of exposure to the treatment, the differences between treated groups, and the cumulative impact over time. These aggregations provide information on the temporal evolution of the impact on different treated groups, allowing us to examine the heterogeneity between these groups and understand the program's impact on each

[10]This subsection closely follows the discussion in Callaway & Sant'Anna (2021), Roth et al. (2023), and Callaway (2023).

group during a specific period. Thus, the resulting set of estimators provides a more complete and detailed perspective than the conventional estimator.

The equation (3) represents the average treatment impact for units treated for the first time in period g during period t. $ATT(g,t)$ represents the average treatment impact in period t for those who graduated in g.

$$ATT(g,t) = \mathbb{E}\left[Y_t(g) - Y_t(\infty) \mid G_i = g\right] \qquad \forall g \geq t \tag{3}$$

Under the assumption of parallel trends and no anticipation, it is possible to identify the causal effect of the program by comparing the expected change in the outcome variable for group g between periods $g-1$ and t, with a control group composed of units that have not yet received treatment in period t (Callaway & Sant'Anna, 2021; Roth et al., 2023). This comparison remains valid when averaging over a set of cohorts $g \in \mathcal{G}_{\text{comp}}$ such that $g > t$.[11] Equation (4) shows this comparison.

$$ATT(g,t) = \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i = g\right] - \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i \in \mathcal{G}_{\text{comp}}\right]. \tag{4}$$

By replacing the population means with the sample averages in equation (4), we can obtain the desired estimator for the average treatment impact in each group and period, as indicated by equation (5).

$$\widehat{ATT}(g,t) = \frac{1}{N_g} \sum_{i:G_i=g} \left[Y_{i,t} - Y_{i,g-1}\right] - \frac{1}{N_{\mathcal{G}_{comp}}} \sum_{i:G_i \in \mathcal{G}_{comp}} \left[Y_{i,t} - Y_{i,g-1}\right] \tag{5}$$

The results that we present assume no anticipation and conditional parallel trends. No anticipation assumption means that for all units i and period $t < G_i$ (pre-treatment periods), $Y_{it} = Y_{it}(0)$. Conditional parallel trends mean that we compare the paths of outcomes among treated and untreated units conditional on having the same characteristics.[12] This assumption says that conditional on covariates, paths of untreated potential outcomes are the same for all groups (Callaway, 2023). Equation (6) shows this assumption considering that $\forall t = 2, \ldots, \mathcal{T}$, and $\forall g \in \mathcal{G}$

$$\mathbb{E}\left[\Delta Y_t(0) \mid X, G = g\right] = \mathbb{E}\left[\Delta Y_t(0) \mid X\right] \tag{6}$$

Following this assumption, equation (7) shows an expression $ATT(g,t)$ where $p_g := P(G = g), p_g(X) :=$

[11] It is possible to calculate the estimator using two control group options: (i) only considering untreated units; (ii) considering all units not yet treated. The estimates presented here only consider units that have never been treated.

[12] Note that the software implementations of Callaway & Sant'Anna (2021) only impose parallel trends assumptions from the period right before treatment starts until the last period (Callaway, 2023).

$P(G = g \mid X, \mathbf{1}\{G = g\} + U = 1)$ (which is the probability of being in the group g conditional on covariates and either being in the group g or being in the never-treated group), and $m_{gt}^{nt}(X) := \mathbb{E}\left[Y_t - Y_{g-1} \mid X, U = 1\right]$ as follows:[13]

$$ATT(g,t) = \mathbb{E}\left[\left(\frac{1\{G = g\}}{p_g} - \frac{\frac{p_g(X)U}{p_g(1-p_g(X))}}{\mathbb{E}\left[\frac{p_g(X)U}{p_g(1-p_g(X))}\right]}\right)\left(Y_t - Y_{g-1} - m_{gt}^{nt}(X)\right)\right] \tag{7}$$

In our context, conditioning on covariates in the assumption of parallel trends is beneficial because the trajectories of formal employment, in the absence of treatment participation, are influenced by demographic characteristics and human capital accumulation. If these variables are distributed differently between the treated and untreated groups, the "unconditional" parallel trends assumption is typically violated (Callaway, 2023).

As mentioned above, the estimator also allows us to obtain several measures in addition to the aggregate treatment effect, as indicated by equation (8).

$$\theta = \sum_{g \in \mathcal{G}} \sum_{t=2}^{T} w(g,t) \cdot \mathrm{ATT}(g,t) \tag{8}$$

Where $w(g,t)$ corresponds to specific weights that allow measuring different types of treatment associated with a policy.[14] In our context, they allow us to answer, for example, the following questions: (a) How does the effect of participating in PAIL vary with the duration of treatment exposure? (b) Do groups that graduated earlier have, on average, larger/smaller treatment effects compared to groups that graduated later? Additionally, this estimator is adaptable to an event study that delivers the weighted average treatment impact l periods after adoption across different adoption cohorts, as indicated by equation (9).

$$ATT_l^w = \sum_g w_g ATT(g, g+l) \tag{9}$$

Parallel Trends. One might be concerned about the plausibility of the parallel trends assumption in our context. The assumption of parallel trends requires the selection bias to be constant over time. However, there may be different confounding factors in different periods, or the same confounding factors may affect the outcome across different periods. For instance, the motivation

[13] According to Callaway (2023), this expression for ATT (g,t) is doubly robust in the sense that, given parametric working models $p_g(X;\pi)$ and $m_{gt}^{nt}(X;\beta)$ for the propensity score and outcome regression, respectively, the sample analog of this expression is consistent for ATT (g,t) if either model is correctly specified.

[14] For more details on the different types of weightings and aggregations, see Table 1 in Callaway & Sant'Anna (2021).

to participate in a training program could be more effective during a recession than during a boom. These factors could lead to beneficiaries and non-beneficiaries of the program following different labor market outcome trends pre-treatment. This assumption also requires that exogenous forces affect treated and control groups equally in the post-intervention period (Ryan et al., 2015).

We address the plausibility of parallel trends using three strategies. First, we estimate a fully dynamic version of the equation (2) using the estimators provided by Callaway & Sant'Anna (2021) to check for potential pre-trends. For each event study of Figure 1 we test the joint significance of all the terms simultaneously to be zero based on the hypothesis ($K < 0$): $H_0 : \beta_K = \beta_{-n} = = \beta_{-1} = 0$ vs. $H_1 : H_0$ does not hold. It is important to note that this is only a "partial test" since this assumption requires parallel trajectories without the treatment, which is impossible to observe. It is not a sufficient condition to establish the validity of a difference-in-differences approach (Kahn-Lang & Lang, 2020). Moreover, parallel pre-trends do not necessarily imply parallel post-treatment trends because simultaneous changes in other policies can affect groups differently after the intervention. If this divergence exists, it causes non-parallel post-treatment trends, leading to biased estimates when using difference-in-differences, as the method assumes post-treatment trends would remain parallel without the intervention. Additionally, when we check for parallel trends using pre-treatment differences, this type of test could suffer from low power issues and pre-testing issues, such as selection bias from only analyzing cases with insignificant pre-trends (Roth, 2022; Rambachan & Roth, 2023).

Second, we calculate bounds on relative magnitude following Rambachan & Roth (2023). We calculate bounds on relative magnitude based on the idea that pre-trends are informative about potential violations of parallel trends. We impose that the post-treatment violation of parallel trends will not exceed a constant (M). For instance, ($M = 1$) indicates that the significant result remains robust even if the violations of parallel trends can be as large as the maximum violation observed in the pre-treatment period. We calculate a robust confidence interval for different values of M to understand how robust our findings are to violations of the parallel trends assumption.

Finally, we calculate the Synthetic Difference-in-Differences estimator (SDD) proposed by Arkhangelsky et al. (2021). Systematic differences between the treatment and control groups before the start of treatment may compromise the assumption of parallel trends. The SDD estimator offers an alternative by relaxing this assumption, standardizing the choice of control group units based on observed data. This method selects control units with pre-intervention outcome variables similar to the treated groups. In our context, this allows the selection of control units with work histories similar to those of those who participated in the program. Therefore, these estimators allow us to relax the assumption of parallel trends by generating a synthetic counterfactual that optimally aligns pre-treatment trends based on individuals' employment history.

Synthetic Difference-in-Difference Estimator.[15] According to Arkhangelsky et al. (2021), the goal is to obtain a consistent estimator of the causal effect of a policy (or treatment W_{it}) even when we do not believe that the assumption of parallel trends is met. To obtain an estimator of the average impact, we proceed by estimating the following equation:

$$\left(\tau_a^{\text{sdid}}, \hat{\mu}_a, \hat{\alpha}_a, \hat{\beta}_a\right) = \underset{\tau,\mu,\alpha,\beta}{\arg\min} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} \left(Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau\right)^2 \hat{\omega}_{a,i}^{\text{sdid}} \; \hat{\lambda}_{a,t}^{\text{sdid}} \right\} \tag{10}$$

The average treatment impact estimator for the treated is generated from a regression of individual and time-fixed effects, with weights ω^i_{sdid} and λ^t_{sdid} optimized. According to Clarke et al. (2023), this procedure allows for the presence of time-shared aggregate factors due to the estimation of the time fixed effects β_t, and time-invariant unit-specific factors due to the estimation of the unit fixed effects α_i. The presence of unit fixed effects implies that the synthetic difference-in-differences estimator will seek to match treatment and control units on pre-treatment trends and not necessarily on pre-treatment trends and levels. In this way, the estimator allows for different levels between treatment and control units prior to treatment, thus differing from traditional synthetic control estimators. Unlike the traditional synthetic control estimator, the SDD not only weights the units in the control group but also the pre-intervention periods in order to approximate the counterfactual.

Arkhangelsky et al. (2021) indicates that the estimation process can be adapted to the case of staggered treatment, where units are treated at different points in time. In this case, the average treatment impact is calculated by applying the synthetic difference-in-differences estimator to each of the subgroups treated at different points in time. Then, a weighted average is calculated for each of the subgroups based on the number of treated units and periods each subgroup has that adopt a treatment at some point. As indicated by Clarke et al. (2023), the process of estimating the average treatment impact, in the case of phased adoption, is based on the following algorithm:

1. We have data for the outcome variable (Y), the matrix indicating which units are treated per period (W), and the row vector containing the different adoption periods between units (A).
2. Then, for each $a \in A$:
   - A subset of Y and W is selected for units that are pure controls and those who adopted the treatment at $t = a$.
   - A regularizing parameter is computed: ζ.
   - The weights are calculated for each unit: $\hat{\omega}_a^{\text{sdid}}$.
   - Weightings are calculated for each unit: $\hat{\lambda}_a^{\text{sdid}}$

[15] This subsection closely follows the discussion in Arkhangelsky et al. (2021) and Clarke et al. (2023)

- The synthetic difference-in-difference estimator is calculated according to the following equation:

$$\left(\tau_a^{\text{sdid}}, \hat{\mu}_a, \hat{\alpha}_a, \hat{\beta}_a\right) = \underset{\tau,\mu,\alpha,\beta}{\arg\min} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} \left(Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau\right)^2 \hat{\omega}_{a,i}^{\text{sdid}} \, \hat{\lambda}_{a,t}^{\text{sdid}} \right\} \quad (11)$$

3. Once the above process has been completed, the average treatment impact is calculated according to the following equation:

$$\widehat{ATT} = \sum_{\forall a \in \mathbf{A}} \frac{T_{\text{post}}^{a}}{T_{\text{post}}} \times \hat{\tau}_a^{\text{sdid}} \quad (12)$$

Where T_{Post} is the total post-treatment periods observed in the treated units.

V. Results

V.A. Main results

Table 1 shows a summary of the program's results on formal employment. The first row presents the coefficient from a difference-in-differences ($TWFE$) model. Using the Callaway & Sant'Anna (2021) estimator, the second and third rows show aggregated parameter estimates assessing a training program's impact, assuming conditional parallel trends and no anticipation effects. Standard errors are pooled at the individual level. The Simple Weighted Average shows the average of all estimators for each group period, weighted by group size. In contrast, the Event Study shows the average treatment effects relative to the treatment exposure, specifically highlighting the program's impact based on the period since the beneficiaries graduated. Finally, Table 2 shows the group-time average treatment effects, using the Callaway & Sant'Anna (2021) estimator. They can offer valuable insights into the heterogeneity of treatment effects concerning the group. In this case, a group is a set of units treated simultaneously. Therefore, the date indicates the period in which the training was completed.

Table 1. Average Treatment Effect (formal by gender): Two-way fixed-effect (TWFE) and Callaway & Sant'Anna (2021) estimator (Simple Weighted Average and Event Study)

	Women			Men		
TWFE	0.07***			0.01		
	(0.01)			(0.01)		
Simple Weighted Average		0.11***			0.11***	
		(0.01)			(0.01)	
Event Study			0.11***			0.09***
			(0.01)			(0.02)
e = 0			0.08***			0.10***
			(0.01)			(0.01)
e = 1			0.10***			0.11***
			(0.01)			(0.01)
e = 2			0.11***			0.13***
			(0.01)			(0.02)
e = 3			0.12***			0.13***
			(0.01)			(0.02)
e = 4			0.14***			0.17***
			(0.01)			(0.02)
e = 5			0.13***			0.15***
			(0.01)			(0.02)
Observations	4,647,720	4,647,720	4,647,720	1,977,600	1,977,600	1,977,600

Notes: This table presents estimates of various aggregated parameters of the effect of a training program, assuming unconditional parallel trends and standard errors pooled at the individual level. The first row shows the coefficient associated with the estimation of a difference-in-differences model (TWFE). The second row, Simple Weighted Average, displays the weighted average according to group size of all estimators corresponding to each group-period. The third row, Event Study, reports the average treatment effects relative to the duration of exposure to it, specifically showing the impact of the program based on the period that the beneficiaries have graduated from the program. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 2. Average Treatment Effect by group treated (formal by gender): Two-way fixed-effect (TWFE) and Callaway & Sant'Anna (2021) estimator (Group-Specific Effects)

	Women	**Men**
Group-Specific Effects	0.10***	0.09***
	(0.01)	(0.01)
Q4 (2017)	0.12***	0.12***
	(0.02)	(0.02)
Q4 (2018)	0.12***	0.14***
	(0.02)	(0.02)
Q4 (2019)	0.11***	0.11***
	(0.02)	(0.03)
Q1 (2021)	0.04***	0.01
	(0.01)	(0.02)
Observations	4,647,720	1,977,600

Notes: This table presents estimates of Group-Specific Effects, the average treatment effects are summarized based on the timing of training completion for different groups. A group is the set of units treated at the same time, therefore, the date indicates the period in which the training was completed.* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

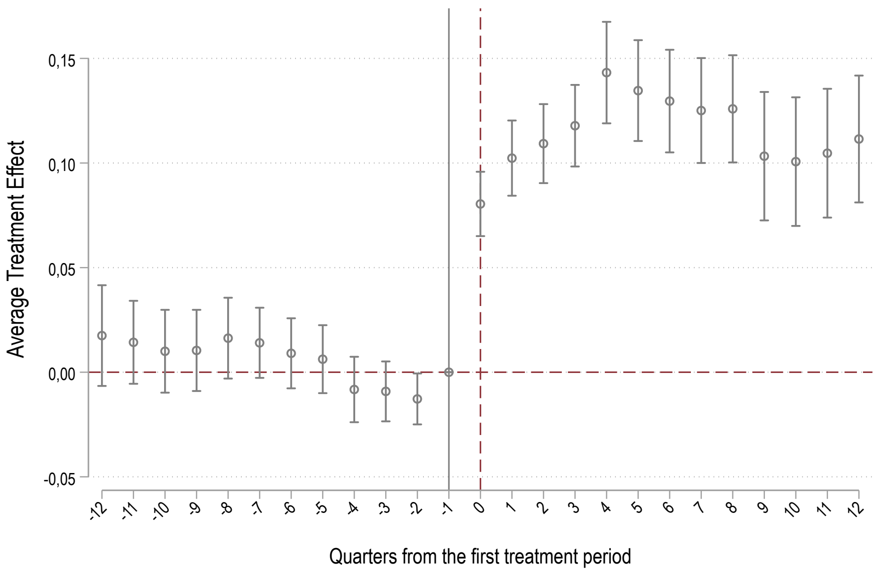
The simple average shows an increase in formal employment by 11 pp. for women and men. The two-way fixed effects model reports an increase in formal employment of 7 pp for women and 1 pp for men. The group-time average treatment effects range from 12 pp to 4 pp for women and 14 pp to 1 pp for men. The program's effectiveness is lower during the pandemic (Q1 - 2021) for women and null for men. These results support that the training program increases formal employment for both men and women, including for women who graduated during the pandemic (Q1 - 2021, Table 2). All the estimators are statistically significant for women, while 3 out of 4 group-time average treatment effects are significant for men. There is also a dynamic impact of graduating from one of the PAIL training programs. The impact is positive and statistically significant for both men and women within one and a half years. Looking at a longer horizon, we observe a long-run impact on women, not men. Overall, our results suggest that the PAIL program increased formal employment relative to what it would have been in its absence.

V.B. Event Studies.

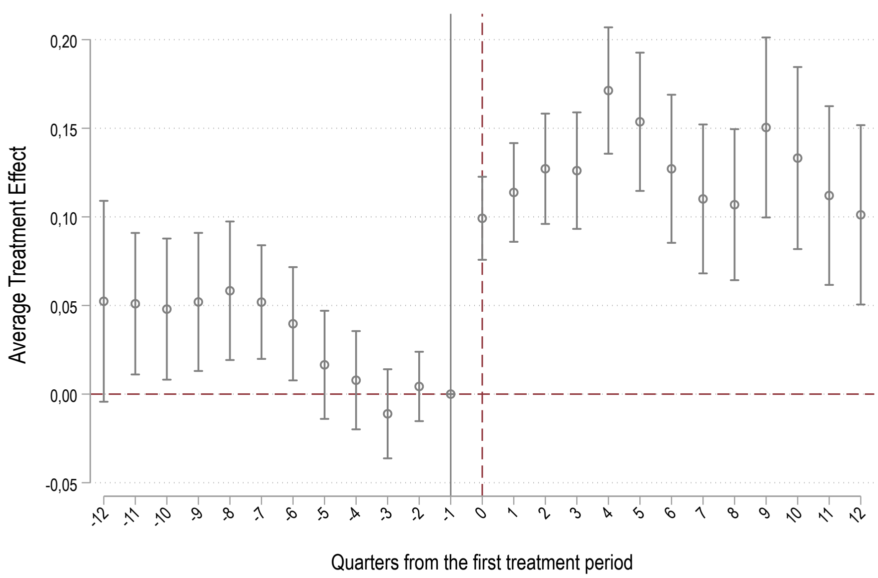
Figure 1 illustrates the impact of the program at different degrees of treatment exposure, based on the estimator by Callaway & Sant'Anna (2021). Figures 2 and 3 illustrate the impact for different treated unit cohorts. Consistent with the results in Table 1, we find a positive impact for both men and women for one and a half years since graduation. After this period, the impact disappears for men but persists for women. The plots include pre-treatment estimates that can be used to "pre-test" the parallel trends assumption and treatment effect estimates for post-treatment periods. We observe differences between the treated and control groups prior to treatment only for men, as shown in each event study in Figure 1. The main source of differences between treatment and control groups comes from the male sample, especially the cohort of Q4 - 2019 (panel (b), Figure 3). Based on this observation, we present several robustness checks to assess the impact of this trend on our results in 5.4.

Figure 1. Dynamic Impact of Treatment - formal employment by gender

(a) Women

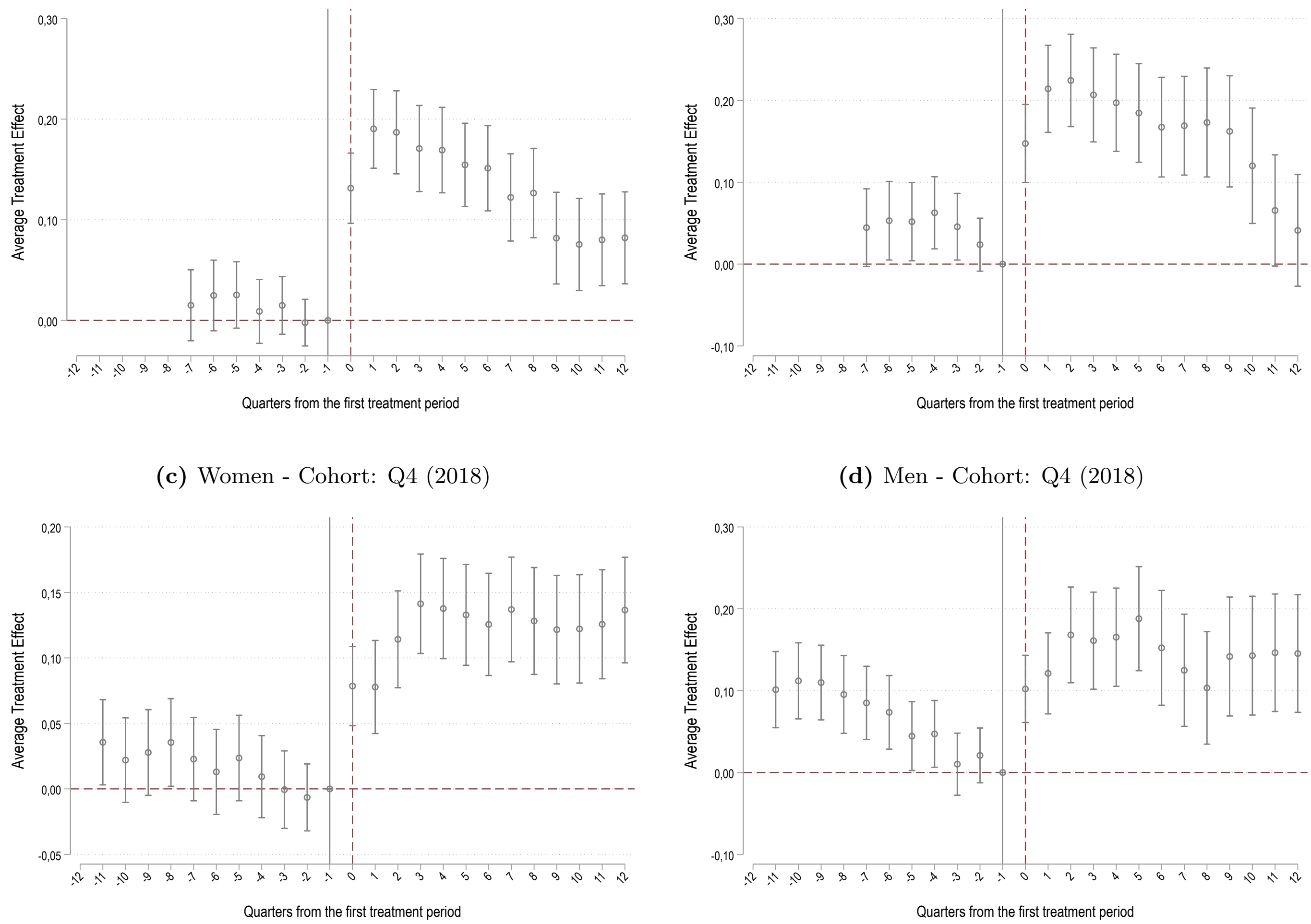


(b) Men



Notes: Panels (a) and (b) present the impact of the treatment over time since adoption for employment variables considering all places but separated by gender. The estimation assumes conditional parallel trends and presents standard errors grouped at the individual level.

Figure 2. Dynamic Impact of Treatment: formal employment by gender and cohort



(a) Women - Cohort: Q4 (2017)

(b) Men - Cohort: Q4 (2017)

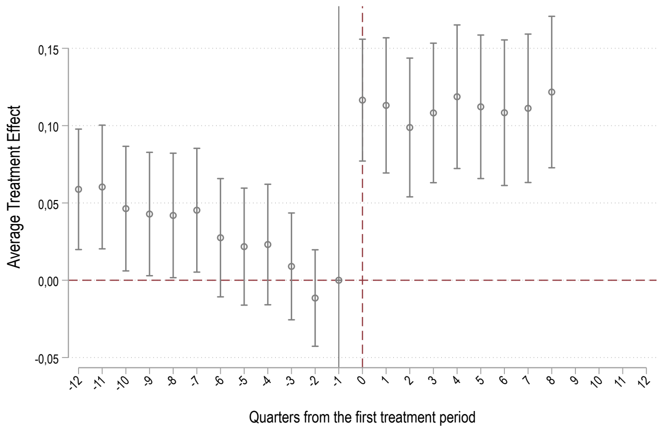
(c) Women - Cohort: Q4 (2018)

(d) Men - Cohort: Q4 (2018)

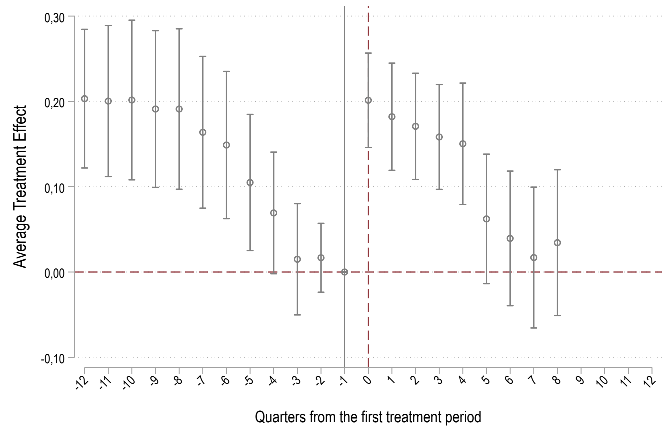
Notes: Panels (a) and (b) present the impact of the treatment over time since adoption for employment variables considering all places but separated by gender and cohorts. The estimation assumes conditional parallel trends and presents standard errors grouped at the individual level.

Figure 3. Dynamic Impact of Treatment: formal employment by gender and cohort

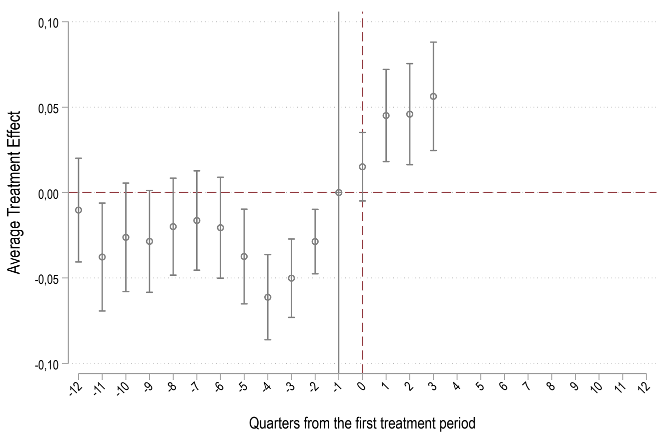
(a) Women - Cohort: Q4 (2019)



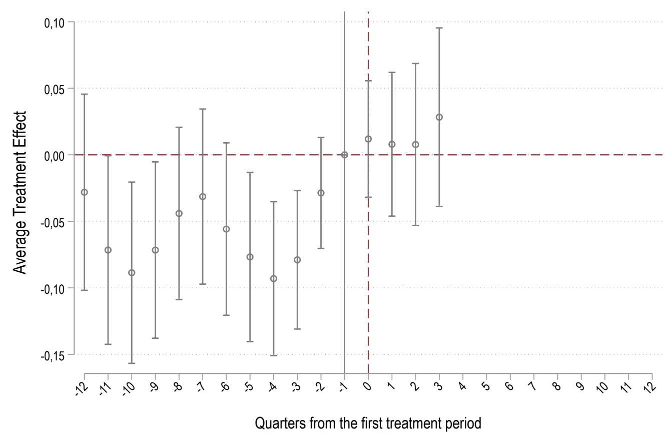
(b) Men - Cohort: Q4 (2019)



(c) Women - Cohort: Q4 (2021)



(d) Men - Cohort: Q4 (2021)



Notes: Panels (a) and (b) present the impact of the treatment over time since adoption for employment variables considering all places but separated by gender and cohorts. The estimation assumes conditional parallel trends and presents standard errors grouped at the individual level.

V.C. Heterogeneity

We examine whether the training program had different impacts depending on the labor market in which the beneficiaries are located. We run our estimations separately for metropolitan and non-metropolitan areas. The metropolitan area includes *Asunción* and ten districts from the Central Department. These districts are Capiatá, Luque, San Lorenzo, Lambaré, Fernando de la Mora, Limpio, Ñemby, Mariano Roque Alonso, Villa Elisa, and San Antonio. Table 3 shows the TWFE estimator and the simple average (weighted by group size) from Callaway & Sant'Anna (2021) by gender and metro-area. Table 4 reports the estimates for units treated at different periods.

Table 3. Average Treatment Effect (formal by gender and place): Two-way fixed-effect (TWFE) and Callaway & Sant'Anna (2021) estimator (Simple Weighted Average - SWA)

	Women				**Men**			
	Metro	**Others**	**Metro**	**Others**	**Metro**	**Others**	**Metro**	**Others**
TWFE	0.099***	-0.033***			0.041***	-0.137***		
	(0.010)	(0.010)			(0.012)	(0.014)		
SWA			0.113***	0.107***			0.157***	0.049**
			(0.013)	(0.014)			(0.018)	(0.023)
Observations	2,113,176	2,534,544	2,113,176	2,534,544	945,216	1,032,384	945,216	1,032,384

Notes: This table presents estimates of various aggregated parameters of the effect of a training program, assuming unconditional parallel trends and standard errors clustered at the individual level. The estimators considers different sub-samples according to gender and place. The first row shows the coefficient associated with the estimation of a difference-in-differences model (TWFE). The second row, Simple Weighted Average (SWA), displays the weighted average according to group size of all estimators corresponding to each group-period. *Metro* includes Asunción along with 10 districts from the Central Department. These districts are Capiatá, Luque, San Lorenzo, Lambaré, Fernando de la Mora, Limpio, Ñemby, Mariano Roque Alonso, Villa Elisa, and San Antonio.* $p < 0.10$; ** $p < 0.05$ *** $p < 0.01$.

We observed a significant heterogeneity in our results based on the labor market in which the program participants are located. The program's overall impact for both men and women is primarily driven by its effect on beneficiaries from the metropolitan area. The impact is even negative for men and women outside urban areas, considering the $TWFE$ estimator. Considering the simple average treatment effect, we observe an increase in formal employment for women of 11.3 pp in metropolitan areas and a 10.7 pp increase in non-metropolitan areas. For men, the results are similar but lower: the treatment leads to an increase in formal employment of 15.7 pp in urban areas and an increase of 4.9 pp in non-metropolitan areas. Figure 2 and 3 report the event studies by gender and cohort. We note that the source of non-parallel trends is mainly driven by men from the cohort of 2018 and 2019 (see Panel (d) in Figure 2 and panel (c) Figure in 3).

Considering the impact of different treated unit cohorts, we note that all the estimators are statistically significant for women (metro and no metro areas). At the same time, only one group-time average treatment effect is significant for men (Q4 - 2019). The group-time average treatment impact for women is 11 pp, ranging from 6 to 13 pp. The impact for men is 13 pp, ranging from 3 pp to 13 pp. The training program's impact for men is very ineffective outside metro areas for almost every cohort of beneficiaries, with the exemption of the cohort of 2019. We note again that the program effectiveness decreased during the pandemic in metro and no metro areas for women and men.

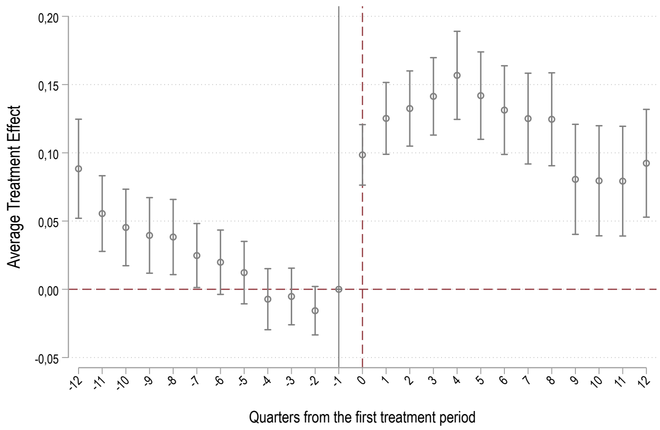
Table 4. Average Treatment Effect on formal employment by group treated (gender and place): Two-way fixed-effect (TWFE) and Callaway & Sant'Anna (2021) estimator (Group-Specific Effects)

	Women		**Men**	
	Metro	**Others**	**Metro**	**Others**
Group-Specific Effects	0.11***	0.08***	0.13***	0.03
	(0.01)	(0.01)	(0.02)	(0.03)
Q4 (2017)	0.11***	0.12***	0.14***	0.07
	(0.02)	(0.03)	(0.03)	(0.04)
Q4 (2018)	0.12***	0.14***	0.20***	0.04
	(0.02)	(0.02)	(0.03)	(0.04)
Q4 (2019)	0.13***	0.08***	0.16***	0.09**
	(0.03)	(0.02)	(0.03)	(0.04)
Q1 (2021)	0.06***	0.03**	0.03	0.00
	(0.02)	(0.01)	(0.03)	(0.04)
Observations	2,113,176	2,534,544	945,216	1,032,384

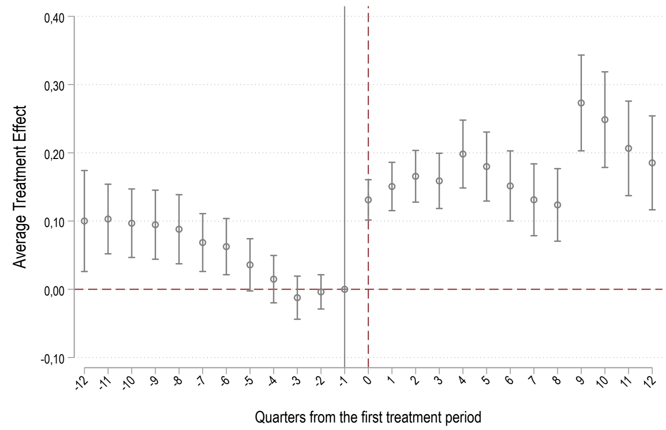
Notes: This table presents estimates of Group-Specific Effects, the average treatment effects are summarized based on the timing of training completion for different groups. A group is the set of units treated at the same time, therefore, the date indicates the period in which the training was completed. *Metro* area includes Asunción along with 10 districts from the Central Department. These districts are Capiatá, Luque, San Lorenzo, Lambaré, Fernando de la Mora, Limpio, Ñemby, Mariano Roque Alonso, Villa Elisa, and San Antonio.* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Figure 4. Dynamic Impact of Treatment: formal employment by gender and metropolitan area

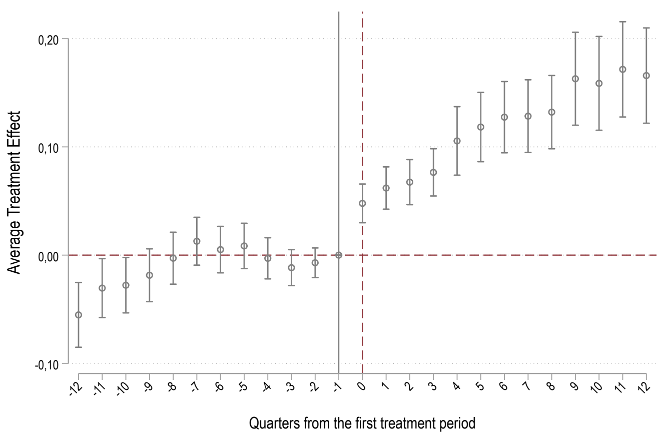
(a) Women - Metro Area



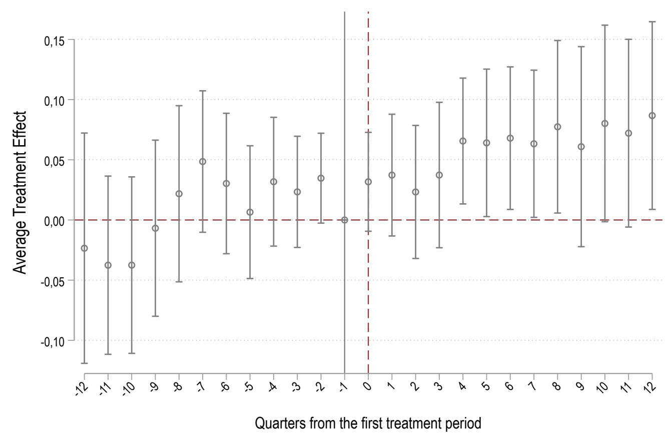
(b) Men - Metro Area



(c) Women - Outside Metro Area



(d) Men - Outside Metro Area



Notes: Panels (a) and (b) present the impact of the treatment over time since adoption for employment variables only considering the metropolitan area, while panels (c) and (d) show the impact of the program at different degrees of exposure to the treatment for non-metropolitan areas. The estimation assumes conditional parallel trends and presents standard errors grouped at the individual level. The Metro area includes *Asunción* and ten districts from the Central Department. These districts are Capiatá, Luque, San Lorenzo, Lambaré, Fernando de la Mora, Limpio, Ñemby, Mariano Roque Alonso, Villa Elisa, and San Antonio.

V.D. Robustness and sensitivity checks

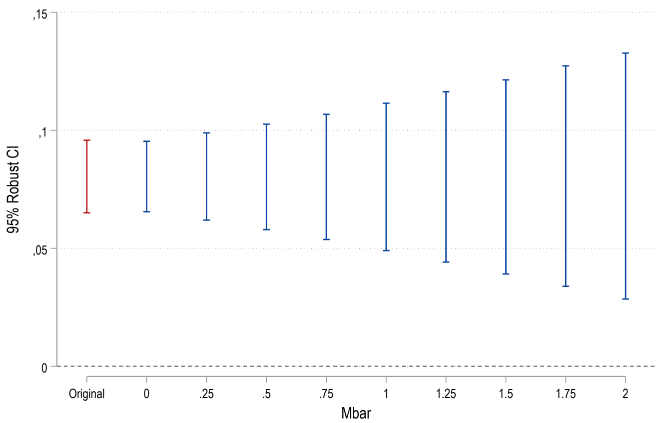
Negative weights. We calculate the negative weights according to the proposal by De Chaisemartin & d'Haultfoeuille (2020) for each $\hat{\beta}_{TWFE}$ for our four main estimators: (i) Women in metro areas (0.000); (ii) Women in non-metro areas (0.000); (iii) Men in metro areas (0.000); (iv) Men in non-metro areas (0.000). Negative weights imply that it is, at least in principle, possible for TWFE regression to deliver very poor estimates of causal effects. For example, in extreme cases, it would be possible for all Average Treatment effects to be positive but, due to negative weighting, to be negative. The results here suggest that the estimator is robust to treatment impact heterogeneity.

Testing violation of Parallel trends. Visual inspection of the event studies is consistent with a violation of the parallel trends assumption, especially for men (see Figures 2, 3, and 4).[16] We calculate, as a sensitivity analysis, the impact on our results if parallel trends do not hold, following Rambachan & Roth (2023). We calculate bounds on relative magnitude based on the idea that pre-trends are informative about potential violations of parallel trends. We impose that the post-treatment violation of parallel trends will not exceed some constant (M). For instance, $M = 1$ means imposing that the post-treatment violation is no greater than the worst pre-treatment violation of parallel trends. Similarly, $M = 2$ indicates that the significant result remains robust even if the violations of parallel trends can be up to twice as large as the maximum violation observed in the pre-treatment period. We calculate a robust confidence interval for different values of M to understand how robust the impact we found is to violations of this assumption. Figure 5 shows the results for both men and women, while Figure 6 shows the results conditional on the location of beneficiaries. We observe that even allowing violations of parallel trends up to twice as large as the maximum violation observed in the pre-treatment period, the result of the training program for women and men remains positive and statistically significant. When we look at the impact by location, we observe that the overall result for women and men is robust for the metropolitan area and weaker outside metropolitan areas (see Panels (c) and (d), Figure 6).

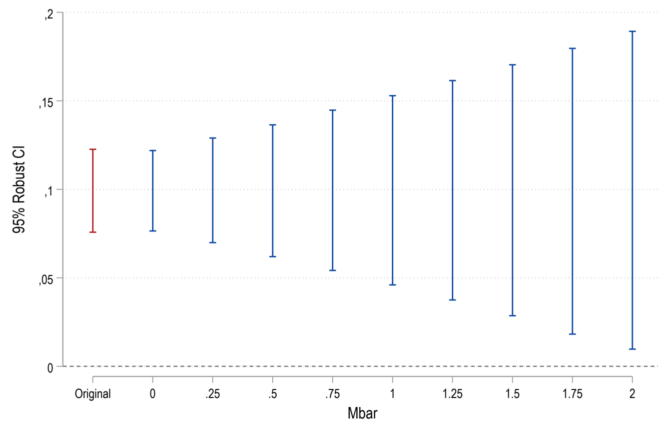
[16] It is essential to note that the assumption of parallel trends may be questionable in our context due to various factors. Although this assumption allows for including variables that affect treatment participation, these variables must have a constant and additive effect on the average outcome over time to be controlled for fixed effects. However, bias will be introduced into the estimates if these factors change non-additively over time. In our case, there may be different reasons to question this assumption. For example, macroeconomic factors may vary non-additively over time, e.g., pandemic impacts. For more details on possible violations of the parallel trends assumption, see: Roth et al. (2023).

FIGURE 5. HONEST DD: FORMAL EMPLOYMENT BY GENDER

(a) Women



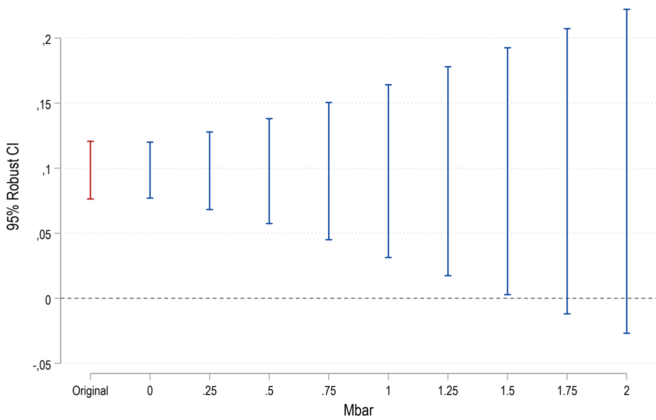
(b) Men



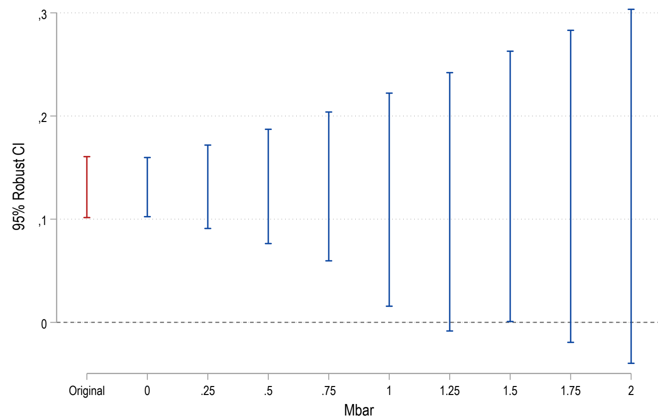
NOTES: The panels show the "breakdown value" (M) for a significant effect for different combinations among gender and place. The blue line is the confidence interval in the original $e = 0$ coefficient.

Figure 6. Honest DD: formal employment by gender and place

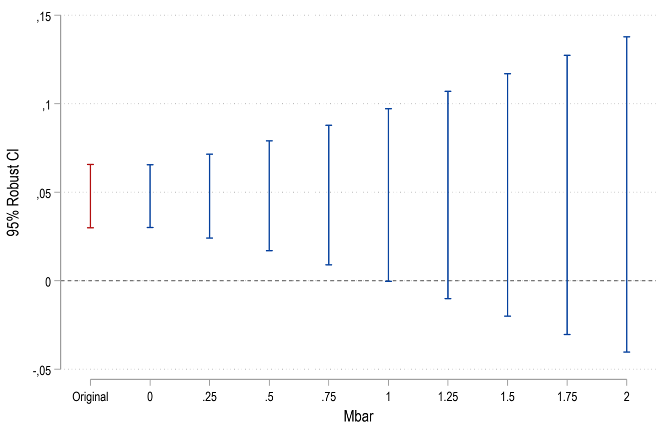
(a) Women - Metro Area



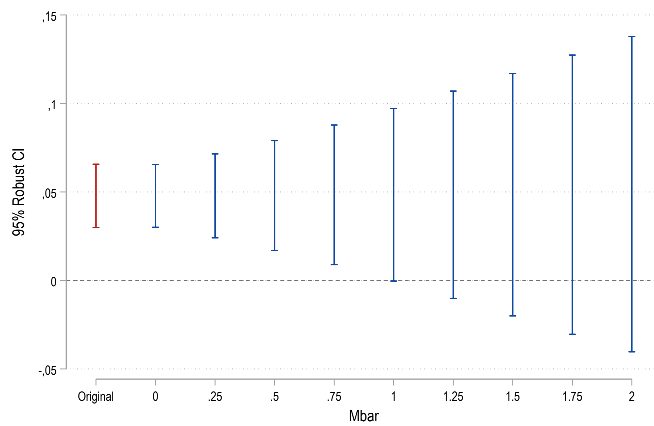
(b) Men - Metro Area



(c) Women - Outside Metro Area



(d) Men - Outside Metro Area



Notes: The panels show the "breakdown value" (M) for a significant effect for different combinations among gender and place. The blue line is the confidence interval in the original $e = 0$ coefficient.

Synthetic difference-in-differences Estimator. The existence of systematic differences between the treatment and control groups before the start of treatment may compromise the assumption of parallel trends. The *SDD* estimator offers an alternative to relaxing this assumption since it standardizes the choice of control group units through a procedure based on observed data. The selection of control units is based on the idea that the pre-intervention outcome variables should be as similar as possible between the control and treated groups. In our context, this method allows the selection of control units with work histories similar to those participating in the program. The results obtained through these estimators highlight a statistically significant and economically important impact of the PAIL program on formal employment, especially for women and men in metropolitan areas.

Table 5. Synthetic Differences in Differences Estimators (formal employment)

	Women		**Men**	
	Metro	**Others**	**Metro**	**Others**
Treatment == 1	0,223***	0,178***	0,061***	-0,150***
	(0,009)	(0,012)	(0,013)	(0,014)
Obs.	2.113.176	2.534.544	945.216	1.032.384

Notes: Each cell displays results based on the estimator proposed by Arkhangelsky et al. (2021). Standard errors are calculated using bootstrapping with 50 repetitions. Significance levels are indicated as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

VI. Conclusion

This paper stresses the relevance of Active Labor Market Policies in increasing access to formal employment in settings where labor informality is high. The results suggest that the PAIL program was especially effective in reducing barriers to formal employment for women and men in metropolitan areas. This finding is consistent with literature suggesting that labor training programs in Latin America and the Caribbean have effectively achieved better labor outcomes, mainly when focused on vulnerable groups.

The results of this study suggest that supply-side interventions are ineffective if few jobs are available. We observe that participating in the training program improved labor market outcomes for participants, but only in metropolitan areas. This difference among urban and rural areas suggests that creating training programs in conjunction with the private sector was an important step in the program's design, aligning the training supply with the skills demanded by employers. However, it is necessary to consider the labor market beneficiaries will face more carefully for future interventions. Another aspect to highlight is that part of this program was implemented during a critical period of the pandemic and, despite challenges, demonstrated positive outcomes in terms of employment for the group that graduated during the pandemic.

The results of this paper emphasize the need to implement training, education, integration, and labor reorientation policies with institutions that are effective and efficient in using their resources. An appropriate design of training programs and the institutional capacity to adequately select target groups and technically track beneficiaries is crucial for the programs' effectiveness.

References

Abuelafia, E., Andrian, L. G., Beverinotti, J., Castilleja, L., Díaz Cassou, J., Hirs, J., Manzano, O., and Saboin, J. L. (2020). Alc post covid-19: Retos y oportunidades para países de la región andina.

Alaimo, V. and Tapia, W. (2014). Perspectivas del mercado de trabajo y las pensiones en paraguay. *Desafíos y propuestas.*

Alfonsi, L., Namubiru, M., and Spaziani, S. (2022). Meet your future: Experimental evidence on the labor market effects of mentors. Technical report, Working Paper.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., and Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, 111(12):4088–4118.

Athey, S. and Imbens, G. W. (2022). Design-based analysis in difference-in-differences settings with staggered adoption. *Journal of Econometrics*, 226(1):62–79.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, 119(1):249–275.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint arXiv:2108.12419.*

Callaway, B. (2023). Difference-in-differences for policy evaluation. *Handbook of Labor, Human Resources and Population Economics*, pages 1–61.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, 225(2):200–230.

Card, D., Kluve, J., and Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *The economic journal*, 120(548):F452–F477.

Card, D., Kluve, J., and Weber, A. (2018). What works? a meta analysis of recent active labor market program evaluations. *Journal of the European Economic Association*, 16(3):894–931.

Carranza, E. and McKenzie, D. (2024). Job training and job search assistance policies in developing countries. *Journal of Economic Perspectives*, 38(1):221–244.

Casalí, P. (2016). *Paraguay: panorama de la protección social: diseño, cobertura y financiamiento.* Organización Internacional del Trabajo.

Clarke, D., Pailañir, D., Athey, S., and Imbens, G. (2023). Synthetic difference in differences estimation. *arXiv preprint arXiv:2301.11859.*

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996.

Escudero, V. (2018). Are active labour market policies effective in activating and integrating low-skilled individuals? an international comparison. *IZA Journal of Labor Policy*, 7(1):1–26.

Escudero, V., Kluve, J., López Mourelo, E., and Pignatti, C. (2019). Active labour market programmes in latin america and the caribbean: Evidence from a meta-analysis. *The Journal of Development Studies*, 55(12):2644–2661.

Freyaldenhoven, S., Hansen, C., Pérez, J. P., and Shapiro, J. M. (2021). Visualization, identification, and estimation in the linear panel event-study design. Technical report, National Bureau of Economic Research.

González-Velosa, C., Rippani, L., and Rosas-Shady, D. (2012). ¿cómo mejorar las oportunidades de inserción laboral de los jóvenes en américa latina? Nota Técnica IDB-TN-305, Banco Interamericano de Desarrollo, Unidad de Mercados Laborales y Seguridad Social.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Heckman, J. J. and Kautz, T. (2012). Hard evidence on soft skills. *Labour economics*, 19(4):451–464.

Kahn-Lang, A. and Lang, K. (2020). The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications. *Journal of Business & Economic Statistics*, 38(3):613–620.

Kluve, J. and Rani, U. (2016). *A review of the effectiveness of Active Labour Market Programmes with a focus on Latin America and the Caribbean*. ILO Geneva.

Miller, D. L. (2023). An introductory guide to event study models. *Journal of Economic Perspectives*, 37(2):203–230.

Mitnik, O. A., Ripani, L., and Rosas-Shady, D. (2016). Comparing the results of youth training programs in latin america and the caribbean.

Rambachan, A. and Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90(5):2555–2591.

Roth, J. (2022). Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights*, 4(3):305–322.

Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*.

Ryan, A. M., Burgess Jr, J. F., and Dimick, J. B. (2015). Why we should not be indifferent to specification choices for difference-in-differences. *Health services research*, 50(4):1211–1235.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Urquidi, M., González, R., and Ortega, G. (2023). Trabajo sin papel.

Appendix

The impact of labor training in high informality contexts.
Evidence from Paraguay

List of Figures

List of Tables

I. Appendix: Figures

Figure A1. Administrative data sources

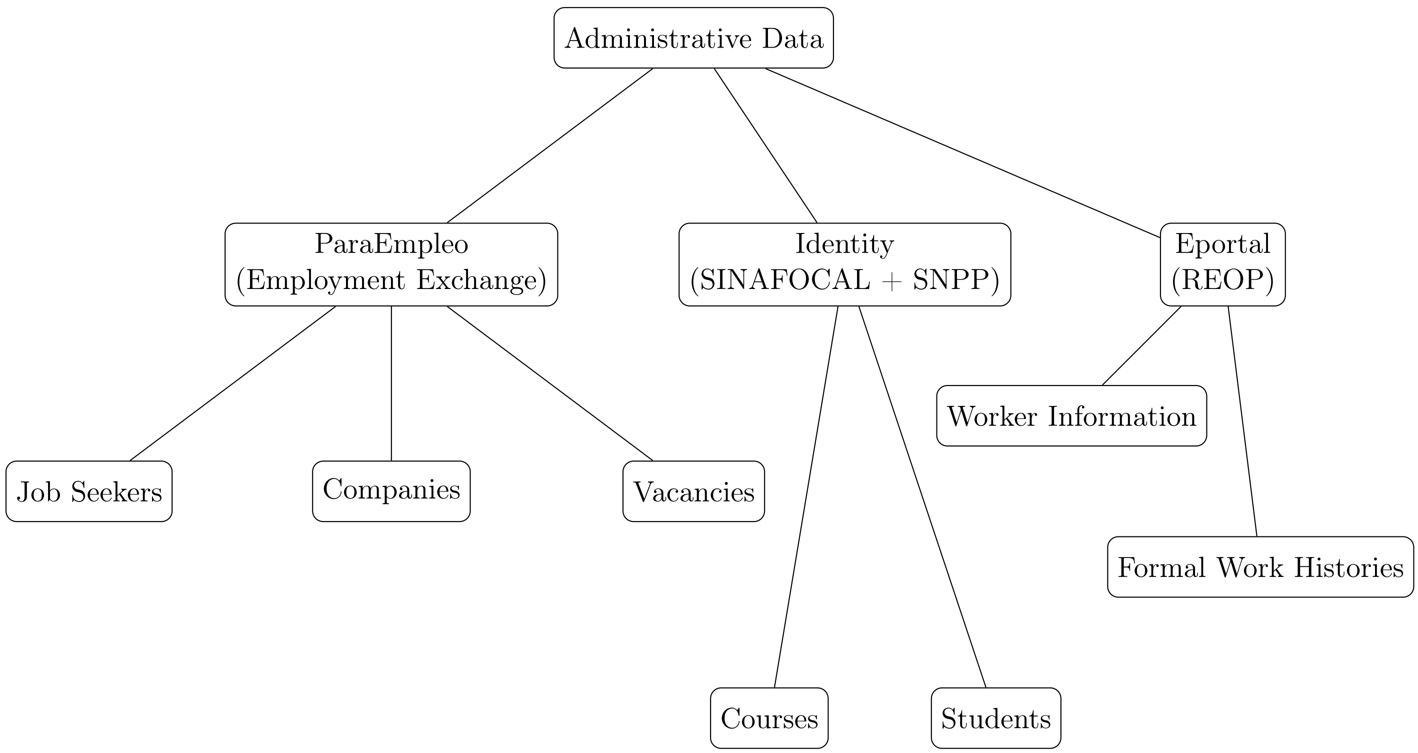
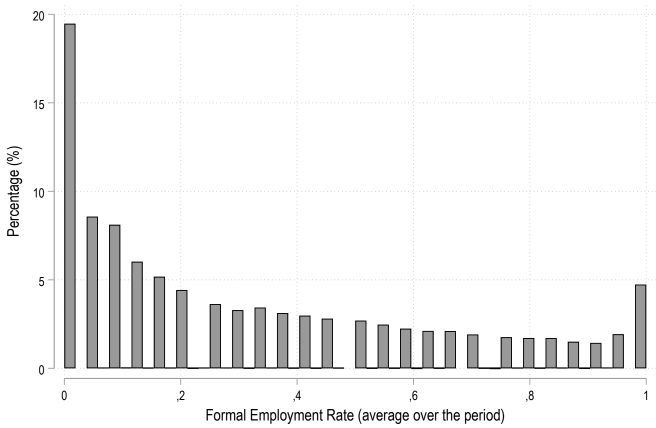
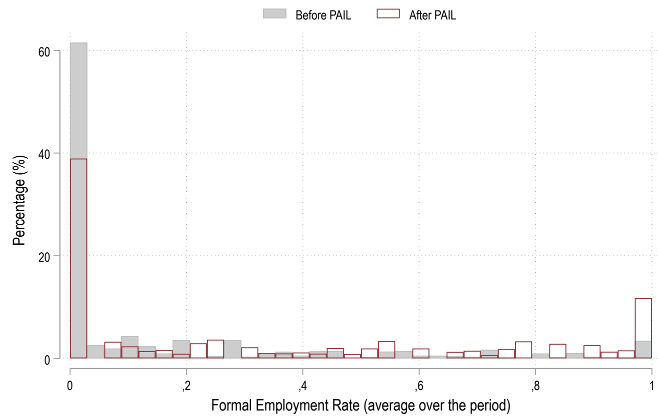


Figure A2. Statistical Distributions of formal employment (women and men)

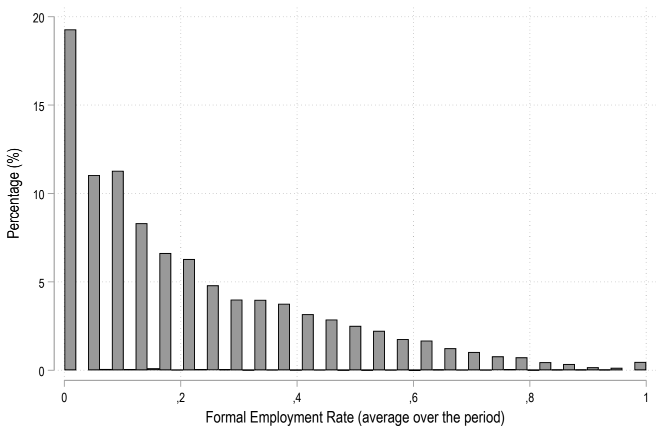
(a) Formal Employment (women)



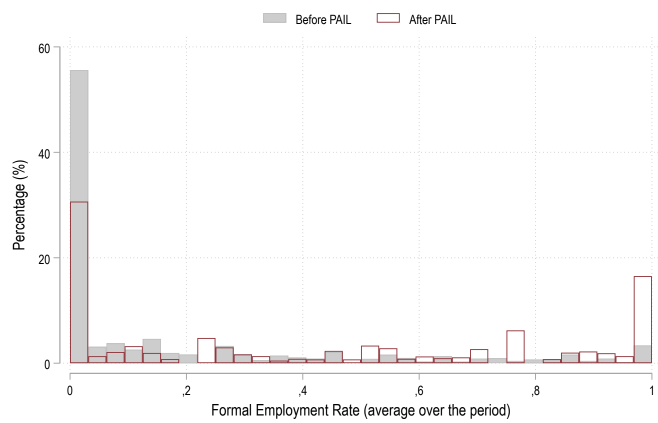
(b) Beneficiaries (women)



(c) Formal Employment (men)

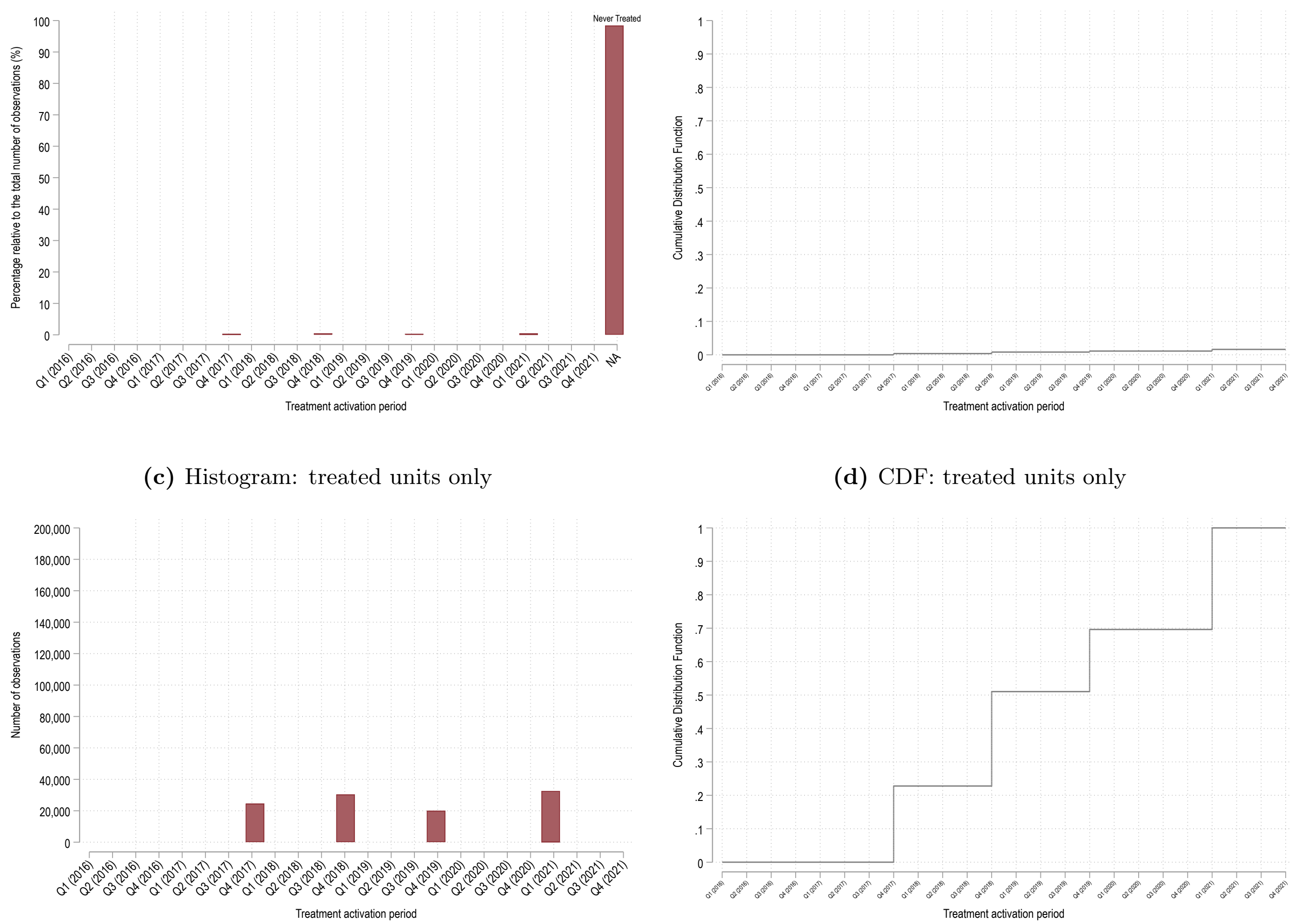


(d) Beneficiaries (men)



Notes: The histograms presented describe the levels of formal employment and hours worked, capturing two distinct perspectives. Figures (a and c) show the percentage of observations found in a certain range for the entirety of the observed sample. On the other hand, Figures (b and d) focus specifically on those individuals who have been beneficiaries of the PAIL program at some point. The figures consider the cumulative averages over the entire period analyzed.

Figure A3. Distribution of treated and untreated units by the period in which treatment is activated



(a) Histogram: treated and untreated units

(b) CDF: treated and untreated units

(c) Histogram: treated units only

(d) CDF: treated units only

Notes: Figures (a) and (c) present histograms of the dates on which treatment was activated, including one observation per unit, for both treated and untreated units. While Figure (a) shows the percentage that the events represent concerning the total number of observations, Figure (c) uses absolute frequencies. On the other hand, Figures (b) and (d) illustrate the Cumulative Distribution Function. In this context, Figure (b) encompasses all treated and untreated units, while Figure (d) only considers those observations belonging to the treated units. This dataset has a hybrid structure characterized by variability in event dates among treated units and a high percentage of units that never received treatment.

II. Appendix: Tables

Table A1. Descriptive Statistics - Entire Sample

	Average	**Std. Dev**	**Median**	**Obs.**
Quarterly income (Guarani in thousands, active)	5.906,09	2.729,81	6.150,00	1.915.300
Quarterly real income (Guarani in thousands, active)	62.943,49	28.805,13	67.100,06	1.915.300
Quarterly Income (logs)	15,41	0,74	15,63	1.915.300
Quarterly real income (logs)	10,87	0,74	11,11	1.915.300
Hourly real income in a quarter (actives)	0,13	0,16	0,11	1.902.052
Formal job == 1	0,29	0,45	0,00	6.625.320
Hours worked in the quarter (actives)	509,62	189,88	572,00	1.910.698
Characteristics				
Age in years	21,70	3,56	21,00	6.625.320
Woman == 1	0,70	0,46	1,00	6.625.320
Metro Area == 1	0,46	0,50	0,00	6.625.320
National == 1	0,99	0,09	1,00	6.625.320
Children == 1	0,06	0,24	0,00	6.625.320
Not single == 1	0,01	0,09	0,00	6.625.320

Table A2. Distribution of Men and Women by Time of Treatment Adoption

Quarter (year)	**Men**	**Women**	**Total**
Q4 (2017)	375	645	1,020
Q4 (2018)	508	756	1,264
Q4 (2019)	318	514	832
Q1 (2021)	596	764	1,360
Total	**1,797**	**2,679**	**4,476**

Table A3. Treated Units by District

District name	**Untreated**	**Treated**	**Total**
Central	99,106	2,139	101,245
Without info	54,665	1,274	55,939
Capital (Asunción)	38,987	888	39,875
Alto Paraná	28,577	24	28,601
Itapúa	9,252	2	9,254
Caaguazú	5,667	67	5,734
Boquerón	4,550	9	4,559
Concepción	4,428	2	4,430
Amambay	3,936	4	3,940
Cordillera	3,903	18	3,921
Presidente-Hayes	3,591	21	3,612
San Pedro	3,353	8	3,361
Canindeyú	3,143	1	3,144
Paraguarí	2,205	9	2,214
Misiones	1,820	1	1,821
Guairá	1,754	3	1,757
Caazapá	1,250	0	1,250
Ñeembucú	809	4	813
Alto Paraguay	583	2	585
Total	**271,579**	**4,476**	**276,055**

Table A4. Distribution of Men and Women by Course

Course	Men	Women	Total	% Men	% Women
Residential Split Air Conditioning	15	0	15	100	0
Automobile Air Conditioning Maintenance	14	0	14	100	0
Basic Home Electricity	17	0	17	100	0
Live Line Electrician Distribution	49	1	50	98	2
Plumbing	14	1	15	93.3	6.7
Senior Technician in Industrial Mechatronics	38	8	46	82.6	17.4
Basic PLC	13	3	16	81.3	18.8
Supermarket Security Guard	13	4	17	76.5	23.5
Bakery Assistant	9	4	13	69.2	30.8
Beverage Preparation	25	22	47	53.2	46.8
User Support	10	9	19	52.6	47.4
Butchery Assistant	10	9	19	52.6	47.4
Commercial Negotiation	40	38	78	51.3	48.7
Customer Service	32	40	72	44.4	55.6
Treasury and Banking Management - Cashier	8	10	18	44.4	55.6
Digital Marketing	21	28	49	42.9	57.1
Oratory - Digital Marketing - Sales - Customer Service	78	107	185	42.2	57.8
HR Management	5	7	12	41.7	58.3
Customer Service, Sales Skills, Oratory, and Digital Marketing	49	71	120	40.8	59.2
Sales Skills	277	406	683	40.6	59.4
Debt Collection Skills	503	770	1273	39.5	60.5
Commercial Cashier	17	26	43	39.5	60.5
Negotiation and Sales Closing	26	40	66	39.4	60.6
Customer Service	384	668	1052	36.5	63.5
Service Station Assistant - Gas Station	4	7	11	36.4	63.6
Supermarket Replenisher	4	7	11	36.4	63.6
Medium and Large Business Administration	5	9	14	35.7	64.3
Basic Cooperative Administration	6	11	17	35.3	64.7
Community Manager	47	99	146	32.2	67.8
Social Media and Community Manager	5	11	16	31.3	68.8
Professional Cashier	16	39	55	29.1	70.9
IT Tools - Office 2013	5	13	18	27.8	72.2
Executive Secretariat	4	13	17	23.5	76.5
Negotiation and Sales	3	11	14	21.4	78.6
Elderly Care	7	30	37	18.9	81.1
Senior Technician in Pharmacy	17	77	94	18.1	81.9
Customer Service and Effective Communication	3	25	28	10.7	89.3
Pharmacy Assistant	4	35	39	10.3	89.7
Convenience Store Assistant (Minimarket)	0	6	6	0	100
Secretariat, Writing, and Spelling	0	14	14	0	100
Total	**1,797**	**2,679**	**4,476**		

Table A5. Descriptive Statistics - Treatment Group

	Average	Std. Dev	Median	Obs.
Quarterly income (Guarani in thousands, active)	5,398.73	2,616.56	5,857.00	28,663
Quarterly real income (Guarani in thousands, active)	57,333.32	27,355.29	63,704.05	28,663
Quarterly Income (logs)	15.30	0.80	15.58	28,663
Quarterly real income (logs)	10.75	0.79	11.06	28,663
Hourly real income in a quarter (actives)	0.15	0.22	0.11	28,419
Formal job == 1	0.27	0.44	0.00	107,424
Hours worked in the quarter (actives)	469.22	204.48	512.00	28,536
Characteristics				
Age in years	21.78	3.28	22.00	107,424
Woman == 1	0.60	0.49	1.00	107,424
Metro Area == 1	0.65	0.48	1.00	107,424
National == 1	0.99	0.08	1.00	107,424
Children == 1	0.24	0.43	0.00	107,424
Not single == 1	0.01	0.09	0.00	107,424

Table A6. Descriptive Statistics - Control Group

	Average	Std. Dev	Median	Obs.
Quarterly income (Guarani in thousands, active)	5,913.79	2,730.77	6,160.00	1,886,637
Quarterly real income (Guarani in thousands, active)	63,028.72	28,818.18	67,100.06	1,886,637
Quarterly Income (logs)	15.42	0.74	15.63	1,886,637
Quarterly real income (logs)	10.88	0.74	11.11	1,886,637
Hourly real income in a quarter (actives)	0.13	0.15	0.11	1,873,633
Formal job == 1	0.29	0.45	0.00	6,517,896
Hours worked in the quarter (actives)	510.23	189.58	572.00	1,882,162
Characteristics				
Age in years	21.70	3.56	21.00	6,517,896
Woman == 1	0.70	0.46	1.00	6,517,896
Metro Area == 1	0.46	0.50	0.00	6,517,896
National == 1	0.99	0.10	1.00	6,517,896
Children == 1	0.06	0.23	0.00	6,517,896
Not single == 1	0.01	0.09	0.00	6,517,896

Table A7. Descriptive Statistics: Treated and Control Units (Prior to Treatment)

	Controls	**Treated**	**Difference**
Quarterly income (Guarani in thousands, active)	5,913.79	4,652.68	1,261.12***
Quarterly real income (Guarani in thousands, active)	63,028.72	51,827.17	11,201.56***
Quarterly Income (logs)	15.42	15.10	0.31***
Quarterly real income (logs)	10.88	10.61	0.27***
Hourly real income in a quarter (actives)	0.13	0.13	0.00**
Formal job == 1	0.29	0.16	0.13***
Hours worked in the quarter (actives)	510.23	438.39	71.84***
Age in years	21.70	20.66	1.04***
Woman == 1	0.70	0.59	0.11***
Metro Area == 1	0.46	0.62	-0.16***
National == 1	0.99	0.99	-0.00***
Children == 1	0.06	0.26	-0.21***
Not single == 1	0.01	0.01	-0.00
Joint Significance Test	0.000		