Pontifícia Universidade Católica
do Rio de Janeiro

Luiza Kunz Aires

Vocational Training and Labor Market Outcomes: Evidence from a randomized program in Brazil

Dissertação de Mestrado

Thesis presented to the Programa de Pós–graduação em Economia, do Departamento de Economia da PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Economia.

Advisor : Prof. Claudio Abramovay Ferraz do Amaral
Co-advisor: Prof. Gustavo Maurício Gonzaga

Rio de Janeiro
August 2020

Pontifícia Universidade Católica
do Rio de Janeiro

Luiza Kunz Aires

Vocational Training and Labor Market Outcomes: Evidence from a randomized program in Brazil

Thesis presented to the Programa de Pós–graduação em Economia da PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Economia. Approved by the Examination Committee:

Prof. Claudio Abramovay Ferraz do Amaral
Advisor
Departamento de Economia – PUC-Rio

Prof. Gustavo Maurício Gonzaga
Co-advisor
Departamento de Economia – PUC-Rio

Prof. Miguel Nathan Foguel
IPEA

Prof. André Portela Fernandes de Souza
EESP-FGV

Rio de Janeiro, August the 20th, 2020

Luiza Kunz Aires

B.A., Economics, Universidade de São Paulo (USP), 2016.

Bibliographic data

Kunz Aires, Luiza

Vocational Training and Labor Market Outcomes: Evidence from a randomized program in Brazil / Luiza Kunz Aires; advisor: Claudio Abramovay Ferraz do Amaral; co-advisor: Gustavo Maurício Gonzaga. – 2020.

57 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia, 2020.

Inclui bibliografia

1. Economia – Teses. 2. Educação Vocacional. 3. Resultados no Mercado de Trabalho. 4. Desemprego. 5. Brasil. I. Abramovay Ferraz do Amaral, Claudio. II. Gonzaga, Gustavo Maurício. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Economia. IV. Título.

CDD: 620.11

Acknowledgments

I wish to thank my advisors, Prof. Claudio Ferraz and Prof. Gustavo Gonzaga, for the guidance and patience. I am very grateful for the support through all the difficulties and turns that I took in this work. I would also like to thank Profs. André Portela and Miguel Foguel for participating in the Examination Committee.

I also thank the Brazilian Ministry of Labor, which, through the *Programa de Disseminação de Estatísticas do Trabalho*, has granted PUC-Rio to the information collected by RAIS that made this project possible.

Finally, I thank my family and friends for the unwavering support and for sometimes just letting me know that I would be alright even if everything were to fail.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. The financial support from PUC-Rio is also gratefully acknowledged.

Abstract

Kunz Aires, Luiza; Abramovay Ferraz do Amaral, Claudio (Advisor); Gonzaga, Gustavo Maurício (Co-Advisor). **Vocational Training and Labor Market Outcomes: Evidence from a randomized program in Brazil**. Rio de Janeiro, 2020. 57p. Dissertação de Mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

This study analyzes the effects of an extensive public technical education program in Brazil. Exploring the fact that acceptance to the program was randomized, with students in over-subscribed classes selected through a lottery, we look at the effects of an offer for the program over formal employment, type of occupation and attachment to the labor market. Our findings show no significant effect over formal employment, earnings or the level of attachment after the program. On the other hand, the program significantly improved the match between students and jobs, with an increase of 1 p.p. (100%) in the probability that an individual is employed in an occupation associated with the course chosen. We also find evidence that receiving an offer for the program led to a higher probability of being employed in an occupation with a higher skill demand. When we examine how these results vary by the gender of the applicant, we find that, while gains in associated occupations are similar for both groups, the effects for high skill jobs appear only men and neither experience improvements on employment or wages after the end of the courses.

Keywords

Vocational Education; Labor Market Outcomes; Unemployment; Brazil.

Resumo

Kunz Aires, Luiza; Abramovay Ferraz do Amaral, Claudio; Gonzaga, Gustavo Maurício. **Educação Vocacional e Mercado de Trabalho: Evidências de um programa aleatorizado no Brasil**. Rio de Janeiro, 2020. 57p. Dissertação de Mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Este estudo analisa os efeitos de um programa governamental de educação técnica no Brasil. Explorando o fato de que a seleção para o programa foi aleatorizada, com estudantes em classes com excesso de inscrições selecionados através de uma loteria, nós olhamos para os efeitos de uma oferta para o programa sobre emprego formal, tipo de ocupação e apego ao mercado de trabalho. Nossos resultados não mostram efeitos significantes sobre emprego formal, salários ou apego ao mercado de trabalho após o programa. Por outro lado, o programa melhorou significativamente o match entre estudantes e empregos, com um aumento de 1 p.p. (100%) na probabilidade de um indivíduo estar empregado em uma ocupação relacionada com o curso escolhido. Nós também encontramos evidência de que receber uma oferta para o programa levou a uma maior probabilidade de estar empregado em uma ocupação com alta demanda por habilidade. Quando examinamos como esses resultados variam com o gênero dos aplicantes, encontramos que, enquanto ganhos em ocupações associadas são similares para ambos os grupos, os efeitos sobre empregos de alta habilidade aparecem apenas para homens e que nenhum experimenta melhoras sobre emprego ou salário após o término dos cursos.

Palavras-chave

Educação Vocacional; Resultados no Mercado de Trabalho; Desemprego; Brasil.

Table of contents

List of figures

List of tables

1
Introduction

One of the most pervasive ways that social and economic inequalities manifest is in the form of disparities in educational opportunities. These can have several negative impacts on a person's life, one of the most critical being that those who have less access to education will most likely also fare worse in the labor market. This education disparity will also contribute to the growth of employment and earnings inequality (Autor (2014)). Vocational Education and Training programs are among the most popular active labor market policies that aim to reduce this gap. VET courses could increase workers' productivity, teach demanded skills and help decrease educational and work disparities. This is especially true in Brazil, which experienced a substantial expansion of the offer of VET programs in the last decade. However, literature is divided on the effectiveness of these programs, with little evidence that they lead to better opportunities in the labor market (McKenzie (2017), Card et al. (2010)).

This paper investigates the effects of a large scale governmental technical education program offered by the state of Bahia, in the northeast of Brazil. For this, we look at applicants for courses in the 2011 cohort. The program, which started in 2009, had more than 60 thousand applicants in 2011 trying to obtain one of almost 8 thousand openings offered across the state. With a duration of 2 years, it offered 44 different courses in several different areas ranging from agriculture, nursing, business, construction, industrial mechanics, etc. Aimed at individuals who would otherwise have no access to a professional qualification, it was offered at no cost and demanded that they had completed high school in a public institution before applying. With the intention of not selecting students based on previous education opportunities that would give advantages in more common entrance exams, students were selected through lotteries for each offered class.

We explore this randomized selection method to estimate the intent-to-treat effects of the program[1]. With this over-subscription design, we compare individuals randomly selected to receive an offer through the lottery with individuals who applied but were not selected. This way, we are able to

[1]As we do not have information on enrollment or completion of the program, we are not able to analyze the effects of the program for those that actually completed it.

avoid capturing the confounding effects given by the fact that enrollment in educational courses is generally related to individual characteristics, which is one of the main difficulties in analyzing this type of program, as these characteristics will also affect outcomes.

We make use of the RAIS dataset (*Relação Anual de Informações Sociais*), a large administrative dataset organized by the Brazilian Ministry of Labor that contains information on all formal employment contracts in Brazil. We follow prospective students from 2003 to 2017 (8 years before, the 2 course years, and 5 years after the completion of the course), obtaining annual information on their formal employment history and contract characteristics such as wages, contracted hours, type of contract and occupation. With the list of all 2011 applicants, we are able to uniquely match and follow 42,054 individuals, from which 4,900 were selected in the lottery.

We regress the main outcomes on the result of the lottery for each individual controlling for randomization strata dummies. We do this both aggregating different periods in relation to the program (before, during and after) as well as year by year, which permit that we look both at the dynamic and average aggregate effects of the program. Using this specification, we are also able to show that the sample is balanced in observed characteristics both in our entire sample of applicants as well as the subsamples divided by gender and that the results of the lottery do not predict outcomes before the start of the courses.

We begin by analyzing the impact of an offer to the course over the individuals' formal employment level. We find that the program did not led to higher formal employment, contracted hours or wages after the end of the course. We find, however, a significant negative impact during the two years of the program: individuals selected in the lottery were 1.5 p.p. (3%) less likely to be working formally in those years. They also worked 0.62 hours less on average per week. Although we do not find a related negative effect on wages, this is consistent with the program having an incapacitation effect, with some students not able to conciliate both classes and work. However, we should be careful to note that we do not observe which selected applicants actually enroll and finish the program. This way, these results should only be interpreted as the effects of a selection in the lottery and can not be extrapolated to be the effects for those that actually enrolled and finished the course[2].

Despite selection to the program not bringing benefits over total employ-

[2] A non-significant ITT estimation could mean that this effect is null, but could also come from few selected applicants actually enrolling in the courses or a low level of retention of students until the completion of the program or not selected individuals attending courses in other institutions or cohorts of the program.

ment, there could be impacts over the type of employment that the individuals obtain. For this, we first look at the probability of obtaining an open-ended contract, i.e., a contract without a predetermined end date. Just as there are no effects over employment, there are also no effects over this type of contract, indicating that the program did not lead to a change towards more secure jobs. Similarly, we do not find any effect over the probability that they obtain employment in a white-collar[3] occupation. We also examine the effect of a selection to the program on the chances of employment on an occupation with a high skill requirement. Here, we find that those selected in the lottery experienced a 1 p.p. (7.2%) higher probability of being employed in a high skill content occupation after the course. Although we cannot measure skill gain directly, this is an indication that the program had an impact on the skill level of students that is captured by the labor market.

Additionally, using a list of occupations related to each course given by the Brazilian Ministry of Education, we show that an offer from the lottery is related to a significant increase of 1 p.p. in the probability of having this type of occupation after the course. We also find that the program has a positive impact over tenure in these occupations, with a half-month increase in the fifth year after the completion of the program. As the desire to work in a given profession is one of the primary motivators for doing a course such as the one studied, this indicates that the program not only helped students match to their desired occupations, but also sustain them for longer.

Next, we examine the effects over the level of attachment to the labor market. First, as selected applicants experience less employment during the course, we verify if this impacted their general level of experience, measured as the number of contract-months worked since 2003. If those who were not selected acquired more experience in this period, this could be counterbalancing possible positive effects of the course to generate the null results after the completion of the program. We also look at tenure in the current active contract, as could be the case that those not drawn acquired position specific experience that could act in the same way. We find that the program had no impact on any period in our sample for neither of them. This is consistent with the lack of effects on employment after the course not being a result of long term effects of the lower employment level during the course. Furthermore, we can also see that acceptance to the program did not lead to higher job stability, as tenure maintained at the 13 months level for both groups.

We also investigate job turnover by looking at the rate that individuals

[3]For the classification of occupations in white-collar and high skill, we use the same procedure as Corseuil et al. (2019) and Menezes-Filho et al. (2008) that traces the Brazilian Occupation Code to ISCO-88 and then uses Abowd et al. (2001)'s classification.

separate from their current contract or start working in a new firm. We find that while the program never has a significant effect on the probability of separation from contract, for the probability of employment in a new firm there's a negative effect of 1.8 p.p. in the first year of the course and a positive effect of 1.6 p.p. in the first year after the end of the courses. We interpret this to mean that the negative impacts in employment during the course are not students leaving their jobs more, but starting fewer new positions in the period than those that are not selected. Further, we see that the program has no effect over job turnover after the course.

We also look at how our results vary by the students' gender. The adverse effects during the course are only present for women, that experience 6.3% (2.7 p.p.) less employment and 1.15 less contracted hours per week during the course when selected in the lottery. They are also 5.9% less likely to have an open-ended contract in this period. Additionally, gains on the probability of having a high skill demand occupation appear only for men, for whom we find a 10.8% increase after the course. Regarding the improvement in the probability of obtaining employment in an occupation associated with the course of choice, we find no gender difference in the aggregate period after the end of the course. However, it happens faster for men, that experience full results in the first year after the course, while for women this takes two years. However, these should not be interpreted as causal, as genders differ in several characteristics as well as in the types of course chosen, and we are not able to separate the effects.

Lastly, we address the fact that we exclude applicants that match to more than one person in RAIS from our sample. For this, we select the match with the higher probability of being the correct one (using the program's prerequisites and location) and look again at our main results with the reinsertion of these individuals. We find qualitatively similar results, with no effects on the level of employment after the program. The negative effect during the course becomes slightly more pronounced, with significant negative effects over wages and open-ended contracts. However, while the effect on the probability of obtaining an associated occupation after the course persists, the effect over high skill jobs becomes non-significant.

This paper contributes to the literature on the effects of vocational and technical education. Although vocational training and other active labor market programs are among the most popular education and employment policies, the existing literature is not conclusive about its effectiveness. With a meta-analysis of active labor-market evaluations Card et al. (2010) finds that less than half of estimated program impacts are significantly positive and 25% of one-year estimates are significantly negative. With a focus in developing countries, McKenzie (2017) finds that the effects estimated for vocational training tend to be very modest and most of the time not significantly different from zero.

Various studies estimate the impacts of VET programs in developing countries. In Colombia, Attanasio et al. (2009) and Attanasio et al. (2017) find that being offered training in the program *Jóvenes en Acción* increased the probability of formal employment by 5.3 p.p. in the short-term and by 4 p.p. in the long-run. They also find that, although the initial effects found were concentrated on the women, there are no gender differences long-term. In the Dominican Republic, *Juventud y Empleo* was found to have no significant effect over employment or earnings in the short-term (Card et al. (2007)). For a different cohort, Ibarrarán et al. (2019) finds a positive impact on earnings and formal employment in the long-run, especially for men. Comparing vocational and firm-provided training for disadvantaged youth in Uganda, Alfonsi et al. (2020) finds that both have positive effects on employment and earnings, but that the effects of vocational training are almost twice as large, which is attributed to the certifiability of this type of training that permits that individuals acquire jobs more rapidly when they become unemployed.

For evaluations of programs in Brazil, the evidence is also mixed. The effects found vary with the study and the type of program. There is no consensus on the existence of a positive effect over general employment, with even fewer evidence for formal employment. Demand driven programs, i.e., those created based on the demands of firms, appear to have the most promising results. Effects can also vary with the gender and age of the participants.

Calero et al. (2017) looks at a 6 months long training program for at-risk youth in the city of Rio de Janeiro (*Galpão Aplauso*). The arts-based program counted with basic and work related training and an informal partnership with local firms for post-program placement. They find effects over any type of employment only in the medium-run and none over formal employment. Corseuil et al. (2019) studies the Brazilian apprenticeship programs. They are firm led programs that combine theoretical training in a classroom and hands-on experience in the firm for those younger than 18. Using the same administrative data for all formal employment contracts in Brazil as we do, they find positive effects on the transition to permanent jobs and a negative impact on accumulated experience, which could be related to higher reservation utilities or higher intention for further education.

Reis (2015) uses data from a representative household survey in Brazil to assess the short-term effects of taking a general VET course. Besides a positive effect on earnings, he also finds a small effect over employment for women. However, no significant effect is found on formal employment for neither women nor men. For the more traditional classroom based curses[4], Barría and Klasen (2016) looks at individuals that participated in SENAI's (one of the largest VET providers in the country) courses compared to courses by other institutions. They find that the programs lead to higher earnings for those aged 15-29 but not for older cohorts. They also find positive effects on formality and employability as well as in the probability of working in an occupation related to the training area.

O'Connell et al. (2017) looks at the effects of demand driven VET programs, comparing Pronatec-MDIC (a public-run program that takes inputs from firms to formulate its courses) to courses offered at the same institutions without firm input. They find that the demand driven programs led to a significant 8.6% increase in employment, while non demand driven programs had no significant effects on their main sample. This is in line with the assertions of Souza et al. (2015) that many VET courses in Brazil are created based on the interest of the prospective students and not on the market demand for the skill taught and that this is one of the drivers of the relatively low effectiveness of Brazilian VET courses.

The study that most closely relates to this one is Camargo et al. (2018). Exploiting an over-subscription experimental design, they examine a similar program that offers secondary level VET courses in the south of Brazil. The main difference between the studies is the fact that Camargo et al. (2018) analyzes the effects of a voucher-type scholarship (*Bolsa Formação Estudante*) aimed at low-income students enrolled in secondary school. In this case, the students are currently enrolled in high school and receive a monetary incentive to attend the VET courses. They find no significant intent-to-treat effects over employment probability or earnings, but a significant increase of 14 p.p. in the probability of working the course's area. However, there is a significant gender heterogeneity, with an increase of 32.4 p.p. in the probability of formal employment for women.

While our study has no pretension to be a final answer to the discussion on the effectiveness of VET courses in Brazil, we aim to contribute to the debate with an analysis of a different program that allows us to look at a longer horizon than is commonly studied. Additionally, the randomized method of selection for all students combined with our administrative dataset on formal contracts permit the use of a larger sample than the ones used by most VET

[4]VET courses in Brazil are usually heavily based on classroom instructions, with hands-on experience comprising only a fraction of the curriculum.

evaluations.

The remainder of this paper is organized as follows: Section II provides background information on the program; Section III presents our empirical strategy; Section IV describes the data; Section V shows our main results; In Section VI we explore gender heterogeneities; Section VII presents a robustness exercise; and Section VIII concludes.

2
Background

Basic education in Brazil is comprised of primary, middle and high school and attending is mandatory until an individual turns 18. If someone desires to continue their studies, they can go on to complete higher education in university. Education is offered free of charge at all levels by the government. However, the government provided basic education is openly recognized as of lower quality compared to its private counterpart, and most families with better financial conditions choose to enroll their children in private schools. On the other hand, public universities are generally considered one of the best options for higher education. This creates an impasse for lower income individuals, as openings in public universities are limited and selection is made through a series of tests based mainly on the high school curriculum. This, combined with the costs of the time taken away from work to continue with their studies, causes many people to not continue beyond basic education.

This means that a vast majority of the populating enters the labor market with at most basic education[1], without any job specific or higher level education or training. To compensate for this, starting in the middle 2000s, the Brazilian government put vocational education and training as a policy priority. From 2003 to 2013 the federal spending in vocational education increased from 720.3 million (0.04% of GDP) to 7.6 billion reais (0.2% of GDP) (Souza et al. (2015)), an expansion of approximately 10 times[2]. Although not a perfect substitute for higher education, VET courses are seen as an alternative for those that cannot access it.

Vocational and Technical courses in Brazil are structured in three different levels: Initial or Continued: short term and specific courses that have no age or prior education requirements that do not grant a degree; Technical courses: for students or graduates of secondary schools, they are longer form courses of at least 600 hours that require at least primary education; Techno-

[1]Or even less. In 2014 only 56,5% of 15 to 17 year olds were enrolled in high school (Schwartzman (2016)).

[2]In 2011, an extensive national technical education program (PRONATEC) was created by the national government. This program is not the object of this study, as it only began operating in the courses provided by the state of Bahia at the end of 2012, when this cohort was finishing the course.

logical courses: equivalent to tertiary level courses, they are only available to high school graduates and grant a diploma akin to a university degree.

The courses analyzed in this paper fall on the category of technical courses. This category counted with 1.4 million students in 2013, with an increase of approximately 45% from 2007 (Souza et al. (2015)). This category is also divided into three subcategories: Integrated: for high school students, it is taught together with high school as one single course; Concomitant: also for high school students, it is done separately from regular classes. The student attends a standard high school class in one shift and the VET course in another, generally in a different institution; Sequential: for high school graduates, it requires that the student completes it before enrolling. Despite that, it is still a secondary, and not tertiary, level course, contemplating a High-School Level Technical Certification for those that complete the course[3]. It is this subcategory that encompasses the courses in this paper. It is also the subcategory with more growth in absolute numbers, going from 417 thousand students in 2008 to 789 thousand in 2015 (Schwartzman (2016)).

The sequential category encompassed 47% of the total offer of technical level courses in 2015 (Schwartzman (2016)). While the integrated and concomitant courses tend to concentrate on young students starting high school, the subsequent ones focus on a different type of individual. According to Schwartzman (2016), these students are typically older, around 30 years old, and looking for additional qualifications for the labor market, where they are usually already inserted.

A Brazilian state that has particularly focused on expanding its public provision of VET courses is Bahia. Selecting it as a state priority in 2008, the state implemented the *Plan for Professional Education of Bahia* and created the *Superintendence for Professional Education* (SUPROF) inside its Secretariat of Education. With that, they implemented a broad state network of professional education courses covering more than a hundred municipalities with 48 thousand students in 2011 (Secretaria de Educação do Estado da Bahia (2011)), of which 12 thousand were in the sequential category (DIEESE (2012)). For these students, 60.1% were women, 33.5% had between 20 and 24 years (with 56% older than 25 years) and 86% of those with race information identifying as black (DIEESE (2012)).

The state government provides free VET courses at all levels. Aimed at public school students and the poorer population, the courses are provided

[3]This certification can act as a license for the exercise of some regulated occupations. If an individual attends certain parts of the course but does not conclude it, they can also obtain an intermediary qualification. These qualifications attest to the course's intermediary conclusion but do not act as a license for occupations.

free of charge. In this study, we focus on the technical courses in the sequential category offered by the state of Bahia. More specifically, we study the cohort that started the course in the first semester of 2011. For this cohort, the state offered 7.982 openings in 44 courses distributed through 46 different cities all over the state. The courses available were extremely varied: they covered nine technological axis[4] and ranged from agriculture, commerce, mechatronics, computing, violin, nursing, etc.

The class plans were devised by each individual school for each individual course, so there are heterogeneities across both. However, there were many common guidelines. Each course had a duration of two years and consisted mainly of classroom instructions that could occur in one of three shifts (morning, afternoon or night). All courses need to cover basic instructions that are mandatory for all VET courses in the country and have an average duration of 400 hours[5]. To that are added specific classes tailored to each course, and a recommended minimum of 400 hours of mandatory internship. Corresponding to a total of more than 2000 hours during the 2 years, which is considerably higher than the minimum hours required by law for this type of VET courses (that vary around 800 hours depending on the course).

The mandatory internship needed to be done in the course's area and up to 100 hours could be at the school in observation activities. Students that had a formal job in the area could use it to fulfill the requirement. Although schools are required to have a department focused on the communication with the labor market, obtaining an internship was the student's responsibility.

The selection of students for this program is made through an electronic lottery organized by the state's Secretariat of Education at the end of each year. The courses are broadly publicized in local media and aspiring students then apply online for their desired class, choosing course, school and shift.[6] If a given class has more applicants than openings, a lottery is used to choose which ones will be offered admission.

These lotteries are transmitted online and professionally audited. For each class (the combinations of course, school and shift), individuals are sorted into an ordered list and prescribed a position. Applicants that are given a position equal to or less than the number of spots in their class are selected

[4]Technological axis are a manner that the Brazilian government classifies VET courses. It separates courses in categories such as 'Management and Business', 'Food Production', etc.

[5]This covers disciplines such as organization of work processes, digital inclusion and computers, philosophy and methodology, etc. The institutions are also required to provide students with assistance in case of difficulties with basic education subjects such as math, writing, etc. But this should be done throughout normal classes.

[6]There is a limit of only one entry per person. So, an individual cannot apply to multiple shifts of the same course, or the same course in different schools.

to participate in the course[7]. From there, the selected individuals go through registration at the beginning of the year (the last week of January in 2011). There they have to provide proof of completion of high school at a public institution. If a selected individual fails to enroll, the next person in the draw's list is called. Classes started at the beginning of February.

[7]The exceptions are courses in the Cultural Production and Design area (which includes Furniture Design and a series of music courses), where selected students also have to pass an ability test. As this inserts a non-random component to the selection of students, these courses are removed from our analysis.

3
Empirical Strategy

There are ordinarily many difficulties in evaluating a program like this one. Chiefest of all is that individuals that participate in this type of course are systematically different from individuals that do not. With this, separating what is the effect of the course from what is simply the selection into it can be challenging. However, by selecting students through a lottery, the Secretariat of Education is essentially mimicking an experimental over-subscription design. As the selection is randomized, receiving an offer to participate in the program, conditional on applying for it, ceases to be related to the individual's characteristics. This way, both groups, those who were selected and those who were not, can be compared and the estimates encountered will not reflect the effects of a selection bias.

Thus, we leverage this randomization of students to estimate the intent-to-treat effects of the program, i.e., the effects of offering a spot on the VET program for participants[1]. For this, we utilize the following specification to estimate the effects of being selected and offered a spot through the electronic lottery:

$$Y_{ic} = \alpha + \beta \times \text{Offer}_i + \delta_c + \epsilon_{ic} \tag{3-1}$$

Where Y_{ic} is the outcome of interest for individual i in the lottery strata[2] c, Offer_i is an indicator variable that equals one if person i was drawn in the lottery and offered a spot in the class and δ_c are class/strata fixed effects. These are included as the randomization occurred at the individual level, but within the applicants for each class. As the number of applicants and openings vary across classes, the probability of selection and the proportion of selected / non-selected individuals vary across strata (Figure 4.1). For this, we weight each applicant by $P_{ic}(1 - P_{ic})$, where P_{ic} is the population probability that person i that applied in class c be selected in the lottery[3]. This way, β is our

[1]We do not observe which students selected in the lottery actually attend and conclude the course or if students nor originally selected were offered a spot in place of those that did not enroll. This way, we are not able to estimate the effects of the course on those that completed it (ATE).

[2]Each lottery strata is a different combination of school, course, and shift. Or a class.

[3]With the differing probabilities of selection in the lottery, we have that the treatment assignment probabilities vary by strata. In this case, the simple comparison between those

coefficient of interest and we have a consistent estimator for the intent-to-treat effect of being offered a place in one of the offered VET courses[4].

We do this analysis with two different time dimensions. To find the weighted average intent-to-treat effects across courses and years, we first present the results of equation (3-1) using panel data in separate regressions for the years before, during and after the realization of the course. For this, we present standard errors clustered at the individual level and include year fixed effects. Additionally, we also present the results of regression (3-1) done year by year, so we can look at the effects over time and differentiate between the short and long run results of the program.

selected or not in the lottery would not provide an unbiased estimation for the effects of the program.

[4]For this, it is also necessary that the offer given to selected individuals does not affect labor market outcomes of non-treated individuals. Although we are not able to test this, we argue that even though this is a large VET program, it is not nearly big enough to affect the general labor market.

4
Data

Our study looks at the 2011 cohort of VET courses in the sequential mode offered by the state of Bahia. For that, we use the originally released list with the results of the electronic lottery. There we have the names, position on the draw and the course, shift and school chosen by each individual that entered the lottery. We match that to the registers of RAIS (*Relação Anual de Informações Sociais*), an administrative dataset that covers all formal employment contracts in Brazil. At the end of any accounting year, every firm in Brazil is required by law to report every worker formally employed at any point in the year. We match prospective students by name[1] in the stock of all workers that appear in the dataset in Bahia from 2003 to 2017 (for a full step-by-step explanation of the process used see Appendix A). Our sample consists of every applicant that has a maximum of 1 exact match in RAIS. Those that match to two or more individuals are removed from the sample. Individuals that have no matches are considered not formally employed in all years.[2]

Our data consists of a yearly panel following 42,054 applicants from 2003 through 2017, eight years before the start of the course, the two course years and five years after. Besides course information (course, shift, school, municipality and if the person was selected or not in the lottery) we also have information on work contracts (wages, type of contract, hours, occupation, reason for separation, etc.) and firm characteristics (size, sector, etc.)[3].

Our main outcomes of interest are: Formal employment, a dummy that equals one when the individual has at least one contract listed for that year in RAIS; Contracted hours per week; Hourly wages, the real value of wages (in 2010 levels) to which we apply the inverse hyperbolic sine transformation[4];

[1]As name is the only personal characteristic present in the original lottery data.

[2]There are concerns that receiving an offer through the lottery could affect employment and the chances that an individual has more than one match in RAIS. However, being selected in the draw has no significant effect on the number of matches in RAIS. We nevertheless present robustness results in Section 7, where we include these individuals in the sample. The results are in line with the ones found with our main sample.

[3]Although RAIS also has individual characteristics, these are available only for those who match in the dataset. The exception for this is gender, that we construct based on the mode of the gender reported for all individuals in RAIS that share a first name with the applicant in the sample.

[4]Defined as $log[x(x^2 + 1)^{12}]$, the inverse hyperbolic transformation allows for similar interpretation of the coefficients as the log transformation. The main difference between both,

Table 4.1: Course Statistics

	No offer (1)	Offer (2)	All (3)
Shift			
Morning	20.7%	24.7%	21.1%
Afternoon	22.6%	29.3%	23.4%
Night	56.8%	46.0%	55.5%
Location			
Capital	49.5%	36.5%	48.0%
Interior	50.5%	63.5%	52.0%
Area			
Environment, Health and Safety	53.2%	42.1%	51.9%
Educational Support	0.2%	0.5%	0.3%
Industrial Processes and Production	20.0%	18.2%	19.8%
Management and Business	11.0%	11.7%	11.1%
Hospitality and Leisure	1.3%	3.3%	1.5%
Information and Communication	0.7%	1.8%	0.8%
Infrastructure	8.4%	10.4%	8.7%
Food Production	1.3%	2.6%	1.5%
Natural Resources	3.9%	9.4%	4.5%
Total	37154	4900	42054

Notes: Numbers are based on the shift, location and area of the courses chosen by the individuals in the main sample. Capital includes both municipalities of Salvador and Feira de Santana.

Open-ended contract, if the individual is employed with a contract without a pre-established end date; High skill job, a dummy that equals one if the individual is in an occupation with a high skill requirement; White collar occupation, an indicator variable for if the individual is employed in a job classified as white collar[5]; Associated job, if the applicant is employed in an occupation that is associated with the chosen course[6]; Experience, measured as the number of months that the individual worked (with a formal contract) since 2003 as of the end of the current year; Tenure, the number of months

however,is that the inverse hyperbolic transformation is also defined when the variable equals zero. As many of the individuals in our sample can have no formal employment contract in a given year, and as a consequence have wage equal to zero, we prefer this transformation.

[5]These measures are based on the one used by Corseuil et al. (2019) and Menezes-Filho et al. (2008) that maps the Brazilian Occupation Code (CBO) into the International Standard Classification of Occupations (ISCO-88) and then separates it into high and low skill groups as well as white and blue collar ones using Abowd et al. (2001)'s classification.

[6]To measure this, we use the National Catalog of Technical Courses made by the Brazilian Ministry of Education, where each course has a list of occupations associated with it. As this list is very small compared with the number of possible occupations, we aggregate it at the family level (4 digits) of the Brazilian Occupation Code. This is the most disaggregated level above each individual occupation, bundling around 2-6 occupations per family.

Table 4.2: Baseline Balance

	Mean			p-value
	No offer	Offer	All	(within)
	(1)	(2)	(3)	(4)
Male	0,323	0,342	0,325	0,548
Worked Before	0,592	0,519	0,583	0,341
Employed	0,395	0,353	0,39	0,363
Hours	16,453	14,66	16,244	0,367
Wage	1,596	1,397	1,573	0,755
# Contracts	0,458	0,405	0,451	0,509
Open-Ended Contract	0,373	0,334	0,369	0,364
Temporary Contract	0,028	0,024	0,027	0,855
Full-time	0,338	0,299	0,333	0,403
Tenure	10,747	9,992	10,659	0,889
Associated Employment	0,003	0,003	0,003	0,908
High Skill	0,113	0,102	0,112	0,847
White Collar	0,241	0,209	0,238	0,666
Experience	19,101	16,906	18,846	0,779
Time Unemployed	4,227	4,628	4,274	0,456
Large Firm	0,174	0,162	0,172	0,853
Total	37154	4900	42054	
	F-stat	p-value		
	0,642	0,852		

Notes: All means are for 2010. Worked Before is an indicator variable that equals one if the individual held at least one contract before 2011. A Large Firm is defined as a firm with at least 250 employees. Column 4 reports the p-value for the difference between the means for each group including strata fixed effects.

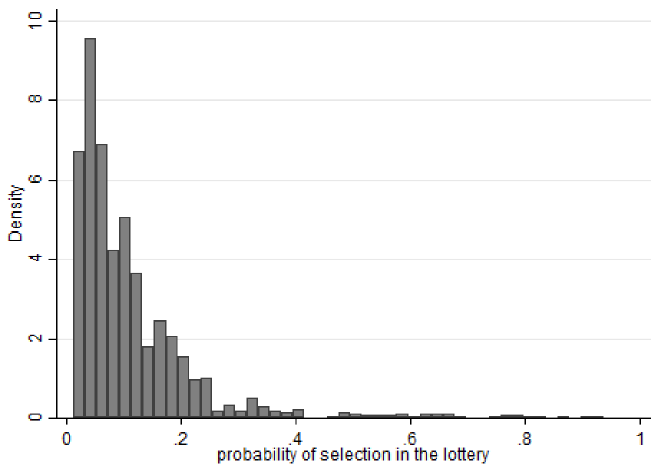
employed in the individual's main active contract; Employment in a new firm, an indicator variable that equals one if the individual starts working in a firm for which they did not work in the previous year; and Separation, a dummy that equals one if the individual terminated at least one contract during the year. For a full list of variables used and how they are coded from RAIS, see Appendix A.2.

4.1 Descriptive Statistics and Balance

Our final sample is comprised of 42,054 applicants, of which 4,900 were selected in the lottery to receive a program offer. Table 4.1 presents the distribution of individuals by the chosen course's shift, location and area of expertise. 56% of prospective students applied for courses that had classes at night, 23% for the ones in the afternoon and 21% in the morning, with course openings following roughly the same distribution. About half of applicants entered lotteries for courses offered in the state capital region[7], although only 36% of sample offers are from courses in this area. Lastly, 52% of individuals applied for courses in the Environment, Health and Safety area, with approximately one third of the sample trying for nursing training. 20% applied for Industrial Processes and Production courses, 11% in Management and Business, 9% in Infrastructure and the rest is divided into five other areas.

Table 4.2 presents averages for characteristics of prospective students in 2010, separating those that received offers (column 2) and those that did not (column 1). Column 3 presents the average for the entire sample and column 4 presents the p-value of the difference between groups calculated using specification 3-1, which includes strata fixed effects. As can be seen, approximately 32% of our sample is male, 59% had at least one contract listed before 2011 with roughly 18 months of experience and 39% were formally employed in 2010 with 10.7 months of tenure on the job. 36.9% had an open-ended contract and 2.7% had a temporary one. The average person worked for 16 hours receiving R$1.57 per hour[8]. 24% had a white collar job, 11% had a job with a high skill requirement and only 0.3% had an occupation associated with the course of choice.

Figure 4.1: Probability of Selection



Notes: Histogram of the probability of being selected in the lottery for all participants in our main sample.

As can be seen in column 4, there is no systematic difference between those that were selected in the lottery and those that were not. Although some of the characteristics seem to differ between groups, not one is statistically

[7]Here we consider both the cities of Salvador, the actual state capital, and Feira de Santana, a large city located about 100 km from Salvador.

[8]As a reference, minimum wage in Brazil was R$2.70 per hour in 2010. As 39% of the sample is employed, the average employed person received R$4.03 per hour and worked for 42 contracted hours per week.

significant, with no p-value below 0.3. This is given by the fact that the difference comes more from disparities across classes with differing chances of selection than from the groups inside each stratum. The mean probability of receiving an offer was 11.6%, but could vary from 1.2% to 93.7%. Figure 4.1 presents the distribution of this probability across the 149 strata in our sample.

5
Results

In this section, we present the estimation results of the effects of a program offer on a series of labor market and contract type outcomes. We look both at year-by-year results and results for the accumulated periods before, during and after the course. We then investigate how an offer for the program affected turnover and attachment to the labor market.

5.1
Employability

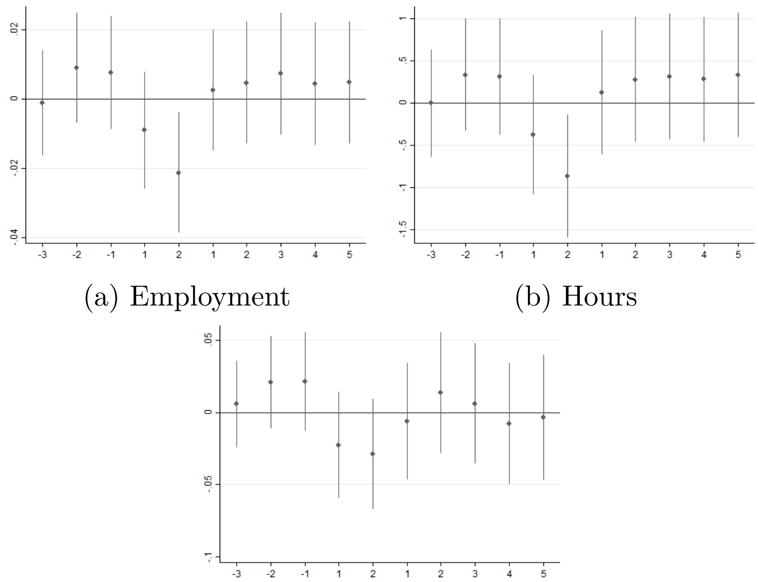
We first ask if the offer for the program resulted in better employability. For this, we look at the effects on the probability of being formally employed in a given year, the number of contracted hours of the individual and hourly wages. In Figure 5.1 we present the results of specification 3-1 done year-by-year and in Table 5.1, columns 1-3, we have results aggregating the periods before, during and after the program in separate regressions[1]. First, as an additional balance test, both Figure 5.1 and Table 5.1, panel A confirm our identifying hypothesis that there were no systematic differences between those that received or not an offer in these outcomes before the lottery[2]. This is in line with what was seen in Table 4.2 and with our expectations that lottery results were random.

Looking at what happens to the employability of selected candidates during the program, we see in Figure 5.1 that both the level of employment and the number of hours worked fall in this period compared to those that did not receive an offer, with both coefficients for the second year of the course significant at 5%. This is confirmed by panel B of Table 5.1 that finds that individuals that received an offer were 1.5 p.p. (3%) less likely to be working when the courses were occurring. They also worked 0.62 hours less per week

[1]Panel A presents the results for years 2003-2010, i.e., our sample contains the years before the start of the course. Panel B gives the effects during the two course years (2011-2012). And Panel C gives results for all years after the conclusion of the course (2013-2017). Additionally to the inclusion of strata fixed effects and weighting by the probability of receiving an offer from the lottery, these regressions also include year fixed effects and clustering at the individual level.

[2]Although Figure 5.1 presents results only up to three years before the start of the program, no significant coefficient is found for the eight years prior to the program that we have in our sample.

Figure 5.1: Labor Market Outcomes



(a) Employment (b) Hours

(c) Wage

Notes: These figures plot the coefficients for receiving an Offer in specification 3-1 done year by year. Bars represent 95% confidence intervals. Employed is a dummy that indicates if the person has at least one contract in the year. Hours is the number of contracted hours per week in the main contract. Wage is the inverse hyperbolic sine transformation of the hourly wages.

than those that did not receive an offer on average. Despite both being only marginally significant, this is consistent with a possible incapacitation effect of participating in the course[3]. This difference, however, does not translate in smaller wages: although coefficients are negative for this period, they are never significant.

For the period after the course, according to Figure 5.1, there is no apparent effect for any year. Looking at the 5 years aggregate estimate in Table 5.1, panel C, being selected in the lottery is related to a non-significant 0.5 p.p. more likely chance to be formally employed in this period, an increase of only 1% over the group that was not drafted in the lottery. Similarly, those selected worked only 0.3 more hours on average each week. This way, point estimates are in the direction expected if the course had beneficial impacts on labor market performance. However, no effect is statistically significant.

So, besides short term negative effects during the course, it appears that the vocational and technical training had no actual effect on students' ability to find formal employment or in their earnings[4]. Although this is in line with

[3] Although students are in class for only one shift, with half of them studying at night, and are not prohibited from being employed at the same time.

[4] This result is also maintained if we restrict the analysis only to formally employed individuals. When we remove individuals from our sample in the years that they do not have a contract listed, we find that selection for the program had no effect on the number

Table 5.1: Program Offer and Labor Market Outcomes

	Employed	Hours	Wage	Open-Ended Contract	High Skill Occupation	White Collar Occupation	Associated Occupation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Panel A: Before			
Offer	0.001	0.044	0.005	0.001	-0.002	-0.003	0.001
	(0.005)	(0.229)	(0.011)	(0.005)	(0.003)	(0.004)	(0.001)
Observations	336,432	336,432	336,432	336,432	336,432	336,432	336,432
R-squared	0.074	0.075	0.063	0.069	0.036	0.047	0.020
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	42054	42054	42054	42054	42054	42054	42054
Mean Dep Var	0.253	10.58	0.410	0.242	0.074	0.149	0.002
				Panel B: During			
Offer	-0.015*	-0.619*	-0.026	-0.013	-0.003	-0.003	0.001
	(0.008)	(0.336)	(0.017)	(0.008)	(0.005)	(0.007)	(0.001)
Observations	84,108	84,108	84,108	84,108	84,108	84,108	84,108
R-squared	0.063	0.066	0.051	0.064	0.048	0.039	0.041
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	42054	42054	42054	42054	42054	42054	42054
Mean Dep Var	0.497	20.67	0.859	0.470	0.139	0.309	0.004
				Panel C: After			
Offer	0.005	0.268	0.000	0.005	0.011**	0.002	0.010***
	(0.008)	(0.317)	(0.018)	(0.007)	(0.005)	(0.007)	(0.002)
Observations	210,270	210,270	210,270	210,270	210,270	210,270	210,270
R-squared	0.035	0.036	0.035	0.037	0.038	0.020	0.024
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	42054	42054	42054	42054	42054	42054	42054
Mean Dep Var	0.504	21.02	0.953	0.487	0.153	0.317	0.010

Notes: Panel A presents results for the panel of years before the start of the course (i.e., the sample for these regressions are panel years 2003-2010), Panel B for the years during the course (2011-2012) and Panel C for the years after the course (2013-2017). *Offer* is an indicator variable that equals 1 if the individual was selected in the lottery. All regressions control for randomization strata and year fixed effects. Standard errors clusterized at the individual level are presented in parenthesis.
* = p < 0.1, ** = p < 0.05, *** = p < 0.01

what is found in some studies for other programs[5] it is still a surprising result

of hours worked or wages received in any period. Which is consistent with the program having no effect after the course and indicates that the negative effect on hours during the course arises from the smaller level of employment, with those employed not decreasing hours worked. However, this result should be seen with caution, as restricting the sample could insert a selection problem in our analysis. The program could have affected which individuals were employed in a given period, even if has no effect over the quantity.

[5]Card et al. (2010) finds that only 45% of the studies surveyed show positive significant results in the medium term, with 10% yielding significantly negative impacts in the 2-3 year window.

given the longer duration of the program (2 years versus the 3-6 months more typical of programs in other studies). It is also in contrast with some estimates found in studies of other programs in Brazil. There are many possible reasons for this. First, the program requires that students have already completed all basic education prior to enrolling in the course. This means that the effect over completing basic education that is sometimes found from this type of program[6] cannot be present here and do not compose part of the benefits in the labor market. Moreover, although the courses can only be taken after high-school, they are not considered post-secondary education, but a high-school level course, so their effects should not be compared to the ones from a college diploma[7].

However, one must take into consideration the fact that we do not have information on take up and the completion rate of the program. This way, there are two lights under which we can see this result. If the program had a high take up and individuals that are selected in the lottery attend and finish the course, than the program itself had no effect over the level of employment of the students. In this case, even attending two years of classes and concluding the course did not lead to higher probabilities of being formally employed. On the other hand, if few selected individuals participate and conclude the program, then our non-significant intent-to-treat estimates could be reflecting this low take up[8]. Here, the program could have an effect over the employment of those that participate in it, leading to the conclusion that this type of training is effective in increasing employability. But, as few individuals selected in the lottery would have completed it, this would not be reflected in our estimates. This way, we are not able to make any affirmations over the effects of the program for those that completed it, only for those that were selected for it in the lottery, including both those that did or did not accept the offer.

This, however, does not mean that we find no impact of the program on selected individual's job opportunities. Although there are no effects over the quantity, there could be an effect over the type or quality of the jobs obtained.

[6]This is found for Brazil by both Corseuil et al. (2019) and Camargo et al. (2018), that find an increased probability of completing high-school of 43.8% and 20%, respectively.

[7]We are also not able to test the possibility that participating in the program has an effect on procuring post-secondary education, so we cannot answer the question of if the courses are seen as a substitute or stepping-stone to further education. Camargo et al. (2018) tests this in a similar scenario without conclusive results.

[8]This could also be caused by individuals that are not selected in the lottery but attend similar courses either on other institutions or future cohorts of the program. And would also be exacerbated by the fact that when a selected individual failed to enroll, the next person in the lottery list would be called.

5.2 Job Characteristics

Figure 5.2 presents the results for the job characteristics of each individual: if the contract was for a fixed period or open ended, if the current occupation has a high skill demand, if the person was employed in a white collar job and if the occupation is associated with the chosen course. Aggregate results over each different period are presented in columns 4-7 of Table 5.1. Following previous results, there are no significant differences among those who did or did not receive an offer for the program through the lottery before the start of the course.

Examining the offer effects on job characteristics during course years, we see that there are no significant impacts over having a high skill or white collar occupation[9] or over being employed in an associated occupation. The dynamic results in Figure 5.2 present the suggestion of negative effects on having an open-ended contract. Although not significant when both years are put together in table 5.1, the negative effect is significant at 5% in 2012, with a decrease of 1.9 p.p. (3.9%) in the probability of having a contract without a termination date. This follows the effect on formal employment, as most of those that are employed have an open-ended contract[10].

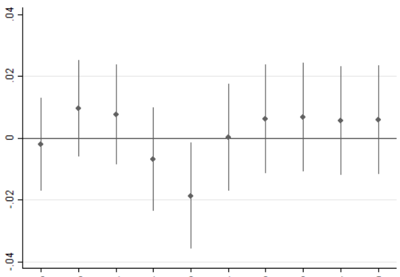
The impacts over job types are concentrated after the end of the course. Although there are no significant effects over open-ended contracts or white collar occupations, we can see in the year by year analysis in Figure 5.2 positive effects over the probability of both high skill and associated occupations. Being drawn in the lottery leads to an increase in the probability of employment in an occupation with a high skill requirement of 1.1 p.p. (7.2%) in the first year after course completion, significant at the 10% level. Point estimates continue relatively stable in the years after that, although not significant in the last two years of our sample. From column 5 of Table 5.1, being drawn in the lottery leads to an increase of 1 p.p. (7.2%) in this probability when we stack all years after the course[11]. This suggests that the skill development given by the

[9]Here, we have white collar occupations as office or administrative assistant, sales assistant, receptionist, etc. In contrast, non-white collar occupations would include production line workers, janitors, construction workers, etc. Examples for high skill occupations would be nursing technicians, high school professors, administrative supervisors, production line workers, etc.

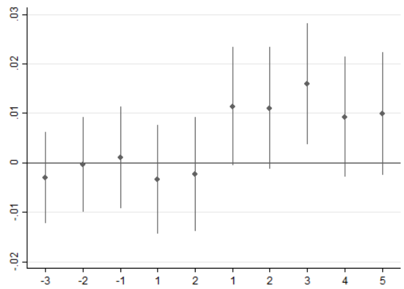
[10]Most of those that have a short-term contract have it in addition to an open ended one. Although we do not show it here, the program has no effect on short-term contracts or the number of jobs that each applicant has.

[11]The point estimate is slightly larger if we restrict our sample only to individuals that are formally employed in a given year. For them, the effect of being selected in the lottery is an increase of 2.3 p.p. in the probability of having a high skill job. Despite that, as approximately 30% of those employed have this type of occupation, this corresponds to a similar 7.6% increase. However, this result should be taken with a grain of salt, as selecting

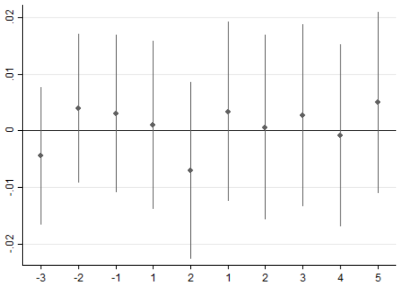
Figure 5.2: Contract Type Outcomes



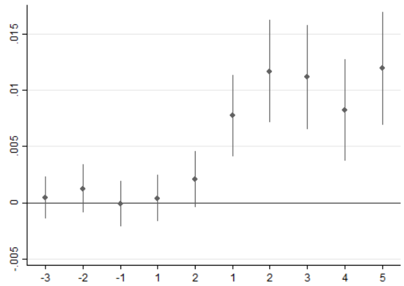
(a) Open-Ended Contract



(b) High Skill Occupation



(c) White Collar Occupation



(d) Employed in an Associated Job

Notes: These figures plot the coefficients for receiving an Offer in specification 3-1 done year by year. Bars represent 95% confidence intervals. Open-Ended Contract is a dummy that equals one if the individual has at least one open-ended contract. High skill and white collar are dummies that equal 1 if the individual's main contract is in a high skill or white collar occupation. Employed in an Associated Occupation is an indicator variable that equals one if the individual has at least one contract in an occupation related to their chosen course.

program is reflected in a change in the type of jobs towards more skill intensive jobs. Even though we cannot directly measure the skill gain from participating in the program, this is an indication that they exist and are reflected by the type of employment that the individual is able to obtain.

Aside from leading students to higher skilled jobs, the program also permitted students to be employed in the area that they chose to major in. Following the occupations listed in the Brazilian National Catalog of Technical Courses as associated with each course[12], from panel 4 of Figure 5.2 we can see that the effect starts already in the first year after the end of the course with an increase of 0.8 p.p. in the probability of having an associated job in 2013 and increases over time reaching 1.2 p.p. in 2017. As Confederação Nacional da Indústria (2014) shows, one of the most cited reasons to participate in a VET course is the desire to obtain a qualification in the chosen profession[13].

on employment could lead to selection bias.

[12]This is by no means an exhaustive list of all jobs that could be done with each course, but they are the ones deemed more closely related by the Ministry of Education and should be the ones aimed at by the programs. We expand the Ministry's list by adding all occupations in the same family as the ones listed according to the Brazilian Classification of Occupations.

In this sense, the course is highly effective. Table 5.1, column 7, shows that the effect of being drawn in the lottery was a 1 p.p. increase in the probability of employment in an associated job. Although this may seem a small effect at first, due to the very small proportion of people employed in these occupations (in part because of the restricted nature of the list of occupations), this is equivalent to doubling the probability compared with the group that was not drawn[14]. This effect is also highly significant, with a p-value of less than $5e^{-8}$. The program is therefore successful in its objective of expanding related job opportunities for its students.

5.3 Attachment to the Labor Market

Next, we look at how an offer for the program affected the dynamics of an individual in the labor market. For this, we investigate the effects over total labor market experience, tenure in the current employment, the probability that the individual starts working in a new firm, and of terminating their current contract.

As we find that the program had a negative short term effect over employment but no effect after the end of the courses, a natural question would be if these short term effects could be offsetting possible positive effects of the course. This could be the case if the difference in employment during the course translated in differences in tenure in the current job or general labor market experience. This could lead to those that were not selected accumulating general or occupation-specific skills on the job or being promoted and receiving higher salaries. For this, we present the results of an offer over total labor market experience and tenure on the current main contract year by year in Figure 5.3, panels A and B, and accumulated by period in Table 5.2, columns 1-2.

We find that there is no significant difference between those that were or were not drawn in the lottery over experience or tenure before, during or after

[14]Similarly to results for high skill occupation, if we look at the effects conditional on employment, point estimates are larger (2.1 p.p.) but percentage changes are similar. As before, this result should be treat with caution, as we could have selection bias when conditioning on employment.

[14]When asked the reasons for participating in a VET course, 47% of a representative sample of Brazilians that attended this type of course had this response, second only to obtaining faster access to the labor market (53%) and followed by 'to expand employment opportunities' (28%) (Confederação Nacional da Indústria (2014)). This is also seen for a center in Bahia in de Oliveira (2011): when students were asked the reason for choosing the course 40% responded that it was because they identified with the course's area against 21% for professional qualification and 34% to obtain a better job.

Table 5.2: Program Offer and Attachment to the Labor Market

	Experience (1)	Tenure (2)	New Firm (3)	Separation (4)	Tenure in Assoc. (5)
			Panel A: Before		
Offer	-0.040	0.060	-0.001	-0.002	0.084
	(0.262)	(0.371)	(0.002)	(0.002)	(0.088)
Observations	336,432	336,432	336,432	336,432	336,432
R-squared	0.118	0.020	0.020	0.018	0.007
Strata FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Clusters	42054	42054	42054	42054	42054
Mean Dep Var	8893	7005	0.096	0.059	0.115
			Panel B: During		
Offer	0.082	0.110	-0.011***	-0.003	0.062
	(0.584)	(0.536)	(0.004)	(0.004)	(0.086)
Observations	84,108	84,108	84,108	84,108	84,108
R-squared	0.064	0.020	0.022	0.014	0.014
Strata FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Clusters	42054	42054	42054	42054	42054
Mean Dep Var	26.01	13.11	0.184	0.123	0.209
			Panel C: After		
Offer	0.114	0.222	0.006**	0.002	0.278***
	(0.807)	(0.648)	(0.003)	(0.003)	(0.094)
Observations	210,270	210,270	210,270	210,270	210,270
R-squared	0.084	0.020	0.008	0.008	0.016
Strata FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Clusters	42054	42054	42054	42054	42054
Mean Dep Var	44.11	18.03	0.106	0.114	0.336

Notes: Panel A presents results for the panel of years before the start of the course (i.e., the sample for these regressions are panel years 2003-2010), Panel B for the years during the course (2011-2012) and Panel C for the years after the course (2013-2017). *Offer* is an indicator variable that equals 1 if the individual was selected in the lottery. All regressions control for randomization strata and year fixed effects. Standard errors clusterized at the individual level are presented in parenthesis.
* = p < 0.1, ** = p < 0.05, *** = p < 0.01

the course[15]. In fact, although not significant, Table 5.2 panel C shows that

[15] Although it seems counterintuitive that there is a difference in employment but not in accumulated experience during the course, this stems from the fact that, although fewer people that receive an offer work in those years, those that are employed work for a longer period on average. In fact, there is no significant difference in the number of contract-months worked between groups in any period, but, if we restrict the sample only to those employed

an offer is related to a 0.08 month higher experience and 0.22 month higher tenure in the period after the end of the course. This way, we do not have evidence that possible positive effects of the program are negated by lingering effects of the lower employment level during the course.

Another way that the program could be affecting attachment to the labor market is by influencing the rate at which the individuals terminate a contract or start working in a new firm[16]. Figure 5.3, panel C, shows that those that received an offer had a 1.8 p.p. smaller chance of working in a new firm in the first year of the course, which was compensated with a 1.6 p.p. higher chance in the first year after the course, significant at 1 and 5% respectively, while there were no significant differences in other years. In panel D, we have the effects over termination of contract[17] showing that there is no significant difference in this probability for any year in the sample. Aggregated results in Table 5.2 show the same pattern, with a negative impact over the probability of being employed in a new firm during (-1.1 p.p., or -6%) and a positive one after the course (0.6 p.p., or 5.7%), which comes from the first year significant result. In the same way, there are no significant results over separation in any period. This way, we see that selected applicants were not leaving their jobs at a higher rate because of the courses, but that some failed to start new contracts in this period.

With this, we conclude that the program had no impact over employability for those selected in the lottery after the end of the course. And the only impact over this was a short term incapacitation effect that led to lower employment levels during the course. This, however, did not lead to any negative effects in the long-term. There were also no effects on the level of attachment to the labor market, with both groups changing employment at the same rate and maintaining the same level of tenure and experience after the program. Although we are not able to examine the effects of concluding the course, our ITT estimations suggest that a selection for the program through the lottery was not enough to impact employability in this way.

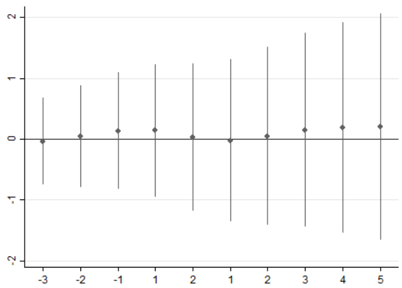
However, as we showed before, selection to the program had a positive impact over the type of employment of selected applicants, leading to a higher probability of having an occupation associated with the chosen course. This

in a given year, selected individuals work on average 0.2 more months during the course.

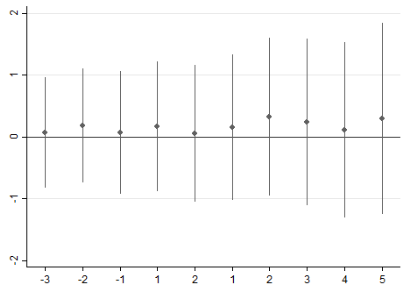
[16]We avoid using the start of any new contracts here because we interpret the case were a person starts a new contract in a firm that they were already working as a renewal of contract or a change in position inside the firm, and not as new employment. Results, however, are qualitatively similar in both cases.

[17]Here we consider separation by any motive (dismissal, resignation, retirement, death, etc.). Results, however, are similar if we consider only cases where individuals were fired or quit their jobs.

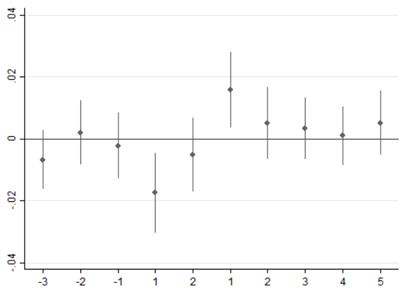
Figure 5.3: Attachment Outcomes



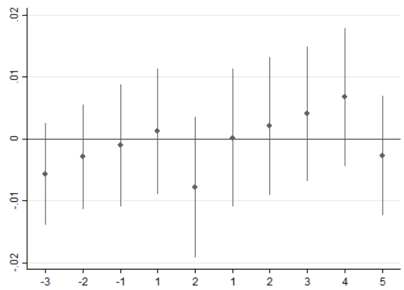
(a) Experience



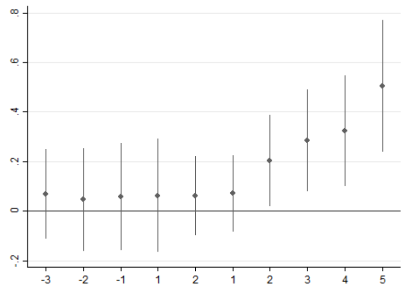
(b) Tenure



(c) New Firm



(d) Separation



(e) Tenure in an Associated Job

Notes: These figures plot the coefficients for receiving an Offer in specification 3-1 done year by year. It covers three years before the course, the two course years and five years after. Bars represent 95% confidence intervals. Experience is the accumulated number of contract x months worked since the beginning of 2003 through to the end of the given year.

way, we also look at tenure in the associated job. Contrary to general tenure, where we do not see an effect, there is an increase in associated jobs due to receiving an offer in the lottery. Table 5.2 column 5 shows a raise of approximately 0.28 months in the five years after the end of the course, an increase of 83% over the mean for those not selected. According Figure 5.3 panel E, this increase in tenure is not a static leap, but an impact that increases with time, starting with an increase of 0.2 months (a 65% increase) in the second year after the course and reaching half a month (a 123% increase) in 2017. I.e., those that received an offer not only have an increased chance to match to their desired occupations, but also to maintain them for longer periods of time.

6
Heterogeneous Effects by Gender

Next, we look at how these effects vary with the student's gender. First, in Table 6.2, we have descriptive statistics and balance for the sample of men and women separately. Our sample is comprised 67% of women, and this distribution is maintained between those selected and not selected in the lottery. Also, as seen in Table 4.2, gender is not a predictor of receiving an offer through the draft. Our sample of women is larger, with 28,374 individuals (3,222 of whom where drawn), but our male sample is still of considerable size with 13,680 applicants, of which 1,678 were selected.

Table 6.1: Course Statistics by Gender

	Women			Men		
	No offer	Offer	All	No offer	Offer	All
	(1)	(2)	(3)	(4)	(5)	(6)
Shift						
Morning	19.8%	22.2%	20.1%	22.4%	29.5%	23.3%
Afternoon	24.5%	29.8%	25.1%	18.6%	28.4%	19.8%
Night	55.7%	48.0%	54.8%	59.0%	42.1%	56.9%
Location						
Capital	48.6%	34.4%	47.0%	51.3%	40.6%	50.0%
Interior	51.4%	65.6%	53.0%	48.7%	59.4%	50.0%
Area						
Environment, Health and Safety	64.8%	52.0%	63.4%	28.7%	23.0%	28.0%
Educational Support	0.3%	0.7%	0.3%	0.1%	0.2%	0.1%
Industrial Processes and Production	10.5%	10.1%	10.5%	39.7%	33.8%	39.0%
Management and Business	10.8%	12.4%	11.0%	11.4%	10.3%	11.3%
Hospitality and Leisure	1.6%	4.2%	1.9%	0.6%	1.7%	0.8%
Information and Communication	0.6%	1.5%	0.7%	1.0%	2.4%	1.2%
Infrastructure	6.5%	8.5%	6.7%	12.4%	14.0%	12.6%
Food Production	1.7%	2.5%	1.8%	0.7%	2.9%	1.0%
Natural Resources	3.2%	8.2%	3.8%	5.3%	11.6%	6.1%
Total	25152	3222	28374	12002	1678	13680

Notes: Numbers are based on the shift, location and area of the courses chosen by the individuals in the main sample. Capital includes both municipalities of Salvador and Feira de Santana. Gender is identified as the mode of the gender of all RAIS observations in Bahia between 2003 and 2017 that share the same first name as the individual.

Same as we had for the entire sample, there is no previous systematic difference between those that did or did not receive an offer in each subsample. And as before, although some differences may appear big at first, they are not significant when we control for randomization strata fixed effects (Table 6.2,

Table 6.2: Baseline Balance by Gender

	Women				Men			
	Mean			p-value	Mean			p-value
	No offer	Offer	All	(within)	No offer	Offer	All	(within)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Worked Before	0,527	0,451	0,518	0,147	0,728	0,65	0,719	0,64
Employed	0,323	0,285	0,319	0,832	0,545	0,485	0,537	0,251
Hours	13,376	11,723	13,188	0,753	22,901	20,299	22,582	0,313
Wage	1,096	0,956	1,08	0,854	2,643	2,243	2,594	0,682
# Contracts	0,365	0,321	0,36	0,556	0,652	0,566	0,642	0,836
Open-Ended Contract	0,304	0,271	0,301	0,522	0,517	0,455	0,509	0,475
Temporary Contract	0,022	0,017	0,022	0,256	0,039	0,038	0,039	0,604
Full-time	0,269	0,233	0,265	0,905	0,481	0,426	0,474	0,286
Tenure	8,643	8,409	8,616	0,593	15,157	13,031	14,896	0,709
Associated Employment	0,003	0,003	0,003	0,351	0,004	0,005	0,004	0,273
High Skill	0,068	0,061	0,067	0,868	0,208	0,181	0,205	0,8
White Collar	0,228	0,192	0,224	0,996	0,269	0,241	0,266	0,551
Experience	15,15	13,472	14,959	0,628	27,383	23,498	26,906	0,993
Time Unemployed	4,78	5,174	4,825	0,954	3,069	3,579	3,132	0,213
Large Firm	0,135	0,126	0,134	0,749	0,254	0,231	0,251	0,386
Total	25152	3222	28373		12002	1678	13680	
	F-stat	p-value			F-stat	p-value		
	0,927	0,532			0,842	0,631		

Notes: All means are for 2010. Worked Before is an indicator variable that equals one if the individual held at least one contract before 2011. A Large Firm is defined as a firm with at least 250 employees. Gender is identified as the mode of the gender of all RAIS observations in Bahia between 2003 and 2017 that share the same first name as the individual. Columns 4 and 8 report the p-value for the difference between the means for those that received or not an offer for each gender including strata fixed effects.

columns 4 and 8). The subsamples are also different from each other. A higher proportion of men worked before and were working on the year before the start of the course compared to women. They also worked more hours and received almost three times the hourly wages. They have higher tenure and work more on high skill jobs on average.

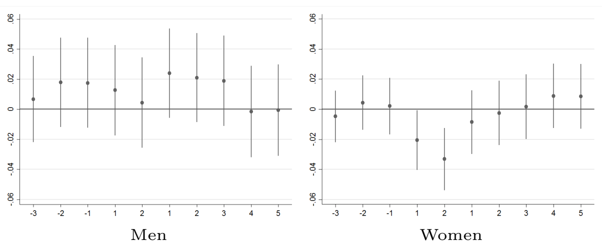
Men and women also differ by the type of course chosen. In Table 6.1, it is possible to see that, although similarly distributed by shift and location, there is a considerable difference in the courses selected between genders. While 63% of the women in the sample applied for a course in the Environment, Health and Safety area, only 28% of men did so. On the other hand, Industrial Processes and Production, the most popular area among men with 39% of applications, was selected by 10% of women. This is also the only category with more men than women applicants in the sample, with 64%. Therefore, as we investigate how the effects of the program vary by gender, we cannot exclude the possibility that this is not a product of the difference in the courses favored by each gender.

First, we look at how both genders differ in the effects over employability. Figure 6.1 presents the effects over formal employment, hours and wages separately for both samples, while columns 1-3 of Tables 6.3, 6.4 and 6.5 present the effects aggregating the three periods in three separate regressions, also presenting the p-value of the difference between estimates for both genders.

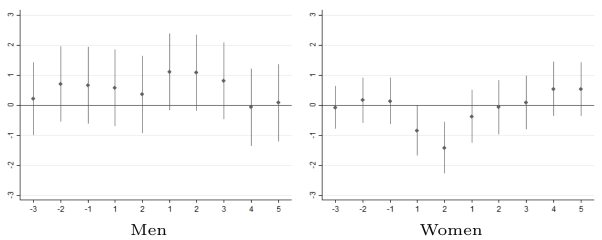
Looking at the result for the probability of being employed and hourly wages, we can see that, just as we find in the main, combined, sample, there is no effect after the completion of the course for any gender. However, when we look at the negative effects during the course, we can see that they are only present in the subsample of women. While these effects are never significant for men, women experience a 2.7 p.p. smaller chance of being employed and work 1.15 hours less per week when drawn in the lottery. With the difference between estimates significant at the 5% level in both cases, the results suggest that this effect is more prevalent for women[1]. Looking at the effects on wages, we can see that the program does not affect either gender.

[1]As mentioned before, a central concern here is the fact that men and women applied for different courses and classes. However, we believe that this is not a major drive in this result as both samples do not differ in the shift of the course applied, which is the other predictor of smaller employment. We do not present this heterogeneity here, but the adverse short-term effects only appear in the subsamples of individuals that applied for morning or afternoon classes, with no significant effects for those that applied for night classes.

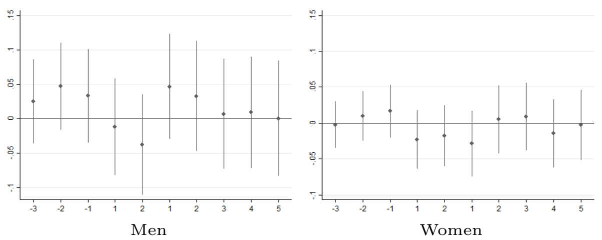
Figure 6.1: Gender Heterogeneity in Labor Market Outcomes



(a) Employed



(b) Hours



(c) Wage

Notes: These figures plot the coefficients for receiving an Offer in specification 3-1 done year by year separately by gender. It covers three years before the course, the two course years and five years after. Bars represent 95% confidence intervals. Employed is a dummy that indicates if the person has at least one contract in the year. Hours is the number of contracted hours per week in the main contract. Wage is the inverse hyperbolic sine transformation of the hourly wages.

Table 6.3: Program Offer and Main Outcomes by Gender - Before

	Employed (1)	Hours (2)	Wage (3)	Open-Ended Contract (4)	High Skill Occupation (5)	White Collar Occupation (6)	Associated Occupation (7)
				Panel A: Women			
Offer	0.001	0.051	0.007	0.002	-0.002	-0.003	0.001
	(0.006)	(0.252)	(0.011)	(0.006)	(0.003)	(0.005)	(0.001)
Observations	226,992	226,992	226,992	226,992	226,992	226,992	226,992
R-squared	0.061	0.061	0.047	0.057	0.017	0.049	0.024
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	28374	28374	28374	28374	28374	28374	28374
Mean Dep Var	0.204	8.452	0.310	0.195	0.047	0.137	0.002
				Panel B: Men			
Offer	0.003	0.103	0.008	0.000	-0.002	-0.002	0.001
	(0.010)	(0.444)	(0.021)	(0.010)	(0.007)	(0.008)	(0.001)
Observations	109,440	109,440	109,440	109,440	109,440	109,440	109,440
R-squared	0.096	0.096	0.086	0.091	0.055	0.048	0.018
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	13680	13680	13680	13680	13680	13680	13680
Mean Dep Var	0.356	15.04	0.619	0.342	0.131	0.174	0.003
				Panel C: p-value of difference			
p-value	0.887	0.918	0.942	0.905	0.967	0.908	0.963

Notes: This table presents results for the period before the beginning of the course (i.e., the sample for these regressions are panel years 2003-2010). Panel A presents results for the subsample of women, Panel B for men and Panel C presents the p-value of the interaction term of the following regression: $Y_{ic} = \alpha + \beta_1 \text{Offer}_i + \beta_2 \text{Male}_i + \beta_3[\text{Offer}_i \times \text{Male}_i] + \delta_c \times \text{Male}_i + \delta_t \times \text{Male}_i + \epsilon_{ic}$. *Offer* is an indicator variable that equals 1 if the individual was selected in the lottery. Gender is identified as the mode of the gender of all RAIS observations in Bahia between 2003 and 2017 that share the same first name as the individual. All regressions control for randomization strata and year fixed effects. Standard errors clusterized at the individual level are presented in parenthesis.
* = p < 0.1, ** = p < 0.05, *** = p < 0.01

For contract characteristics, we have year by year results in Figure 6.2 and aggregate results in columns 4-7 of Tables 6.3, 6.4 and 6.5. Following the results for employment, when we separate genders, the negative effect over open-ended contracts becomes significant for women. They experience 2.4 p.p. smaller chances of having an open-ended contract during the years of the course if they receive an offer through the lottery. This effect does not appear in the male sample, which presents a non-significant increase of 1 p.p.

Looking at the effects over occupation type, we see that, just as in the combined sample, there is no significant effect over the probability of having

Table 6.4: Program Offer and Main Outcomes by Gender - During

	Employed	Hours	Wage	Open-Ended Contract	High Skill Occupation	White Collar Occupation	Associated Occupation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: Women						
Offer	-0.027***	-1.146***	-0.020	-0.024**	-0.005	-0.010	0.001
	(0.010)	(0.395)	(0.019)	(0.009)	(0.005)	(0.009)	(0.001)
Observations	56,748	56,748	56,748	56,748	56,748	56,748	56,748
R-squared	0.057	0.059	0.040	0.059	0.020	0.046	0.051
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	28374	28374	28374	28374	28374	28374	28374
Mean Dep Var	0.432	17.86	0.705	0.407	0.085	0.306	0.004
	Panel B: Men						
Offer	0.009	0.460	-0.025	0.010	0.002	0.009	0.002
	(0.014)	(0.598)	(0.033)	(0.014)	(0.011)	(0.012)	(0.002)
Observations	27,360	27,360	27,360	27,360	27,360	27,360	27,360
R-squared	0.066	0.067	0.059	0.065	0.059	0.041	0.031
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	13680	13680	13680	13680	13680	13680	13680
Mean Dep Var	0.633	26.55	1.181	0.601	0.251	0.316	0.006
	Panel C: p-value of difference						
p-value	0.037	0.025	0.897	0.046	0.589	0.214	0.509

Notes: This table presents results for the period during the course (i.e., the sample for these regressions are panel years 2011-2012). Panel A presents results for the subsample of women, Panel B for men and Panel C presents the p-value of the interaction term of the following regression: $Y_{ic} = \alpha + \beta_1 \text{Offer}_i + \beta_2 \text{Male}_i + \beta_3 [\text{Offer}_i \times \text{Male}_i] + \delta_c \times \text{Male}_i + \delta_t \times \text{Male}_i + \epsilon_{ic}$. *Offer* is an indicator variable that equals 1 if the individual was selected in the lottery. Gender is identified as the mode of the gender of all RAIS observations in Bahia between 2003 and 2017 that share the same first name as the individual. All regressions control for randomization strata and year fixed effects. Standard errors clusterized at the individual level are presented in parenthesis.
$* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$

a white collar job for neither gender. The probability of being employed in an occupation with a high skill requirement, where we encountered a positive effect previously, presents a clear difference between genders. All the effects found come from the sample of men, where an offer for the program is related to an increase of 2.8 p.p. (10.8%) on this probability after the program, while there are no significant effects for women, with a point estimate of 0.4 p.p.

However, as was shown before, the genders distributed themselves differently between courses, with men applying more for the area of Industrial Processes and Production, while women applied more for the service oriented Environment, Health and Safety. So, it could be the case that this result is only reflecting this difference in course type and not an actual difference by

Table 6.5: Program Offer and Main Outcomes by Gender - After

	Employed (1)	Hours (2)	Wage (3)	Open-Ended Contract (4)	High Skill Occupation (5)	White Collar Occupation (6)	Associated Occupation (7)
	Panel A: Women						
Offer	0.002	0.132	-0.006	0.003	0.004	0.006	0.009***
	(0.009)	(0.382)	(0.020)	(0.009)	(0.005)	(0.008)	(0.002)
Observations	141,870	141,870	141,870	141,870	141,870	141,870	141,870
R-squared	0.029	0.030	0.024	0.032	0.015	0.025	0.027
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	28374	28374	28374	28374	28374	28374	28374
Mean Dep Var	0.452	18.74	0.800	0.434	0.101	0.318	0.010
	Panel B: Men						
Offer	0.012	0.596	0.019	0.010	0.028***	-0.007	0.012***
	(0.013)	(0.541)	(0.034)	(0.013)	(0.011)	(0.012)	(0.003)
Observations	68,400	68,400	68,400	68,400	68,400	68,400	68,400
R-squared	0.037	0.037	0.039	0.038	0.044	0.026	0.027
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	13680	13680	13680	13680	13680	13680	13680
Mean Dep Var	0.614	25.81	1.274	0.598	0.260	0.315	0.009
	Panel C: p-value of difference						
p-value	0.494	0.483	0.519	0.655	0.038	0.366	0.418

Notes: This table presents results for the period after the end of the course (i.e., the sample for these regressions are panel years 2013-2017). Panel A presents results for the subsample of women, Panel B for men and Panel C presents the p-value of the interaction term of the following regression: $Y_{ic} = \alpha + \beta_1 \text{Offer}_i + \beta_2 \text{Male}_i + \beta_3 [\text{Offer}_i \times \text{Male}_i] + \delta_c \times \text{Male}_i + \delta_t \times \text{Male}_i + \epsilon_{ic}$. *Offer* is an indicator variable that equals 1 if the individual was selected in the lottery. Gender is identified as the mode of the gender of all RAIS observations in Bahia between 2003 and 2017 that share the same first name as the individual. All regressions control for randomization strata and year fixed effects. Standard errors clusterized at the individual level are presented in parenthesis.
* = p < 0.1, ** = p < 0.05, *** = p < 0.01

gender. Although we cannot test this directly, as the type of course is a choice of the individual and not allocated randomly, we try to address this by identifying what types of courses are the ones generating a gain in high skill jobs. We do not present explicit tables for area heterogeneity, but significant effects are found only in courses in the Infrastructure area, which is chosen similarly by both genders[2]. Additionally, if we go further and look at the heterogeneity inside the sample of Infrastructure applicants, we repeat the findings that re-

[2]This area is comprised of the following courses: Edification, Construction Design and Surveying. It is chosen by 7% of women and 13% of men, which translates into a group comprised 53% of women (as there are more women than men in our sample).

sults are coming only from men[3]. This is consistent with the difference coming from the gender of the student, and not from the different course choices.

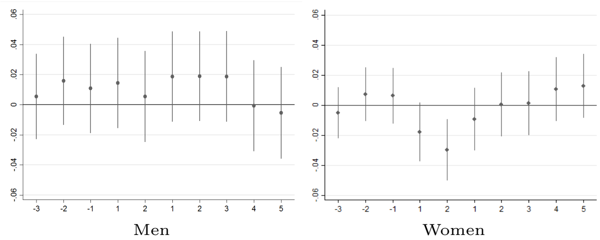
When looking at the probability of obtaining employment in an occupation related to the course of choice, we see in Table 6.5 that the effect is present for both genders. Although point estimates are slightly bigger for men (with a 1.2 p.p. vs. a 0.9 p.p. for women), the difference is not significant, so both genders appear to receive the same effect in the five years after the end of the course overall. However, in Figure 6.2 panel D, we observe that for men, the effect presents itself from the first year out of the course (with a 1.3 p.p. increase), and keeps relatively constant after that. While for women, the effect appears to take more time to present itself fully: there is a 0.5 p.p. increase in the first year that passes to a 1 p.p. increase in the second one before becoming relatively constant at this level. This indicates that results are slower to appear for women, with the 0.7 p.p. difference in the first year significant at the 10% level.

On the whole, the program brings positive results over associated occupations for both men and women. However, they seem to be faster acting for men. Men are also the only ones that experience gains over the probability of obtaining high skill jobs and experience less negative effects over employment during the course. Nevertheless, this result should be taken with a grain of salt, as both samples differ in other characteristics besides gender, and we are not able to separate the gender from its characteristics. Even then, this is an important result, as it demonstrates how men and women are benefiting differently from the program, considering its design and course offering.

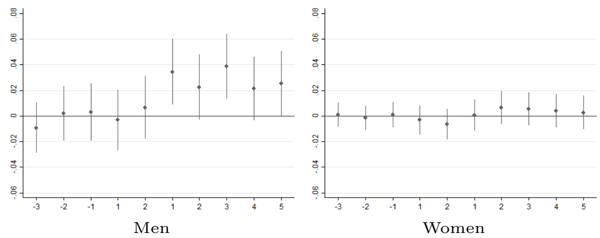
This is somewhat in line with what is seen in the general literature. According to McKenzie (2017) McKenzie (2017), for the VET studies surveyed that test for equality by gender, there is either a similar impact for men and women or a higher impact for men. In the meta-analysis by Card et al. (2010) Card et al. (2010), results point to no general differences in the impact for men and women.

[3]For Infrastructure's prospective students, receiving an offer is related to a 4.3 p.p. higher chance of being employed in a high skill job after the end of the course. For women in this subsample, an offer is related to a non-significant effect in this probability, while for men it leads to a 7.3 p.p. higher chance.

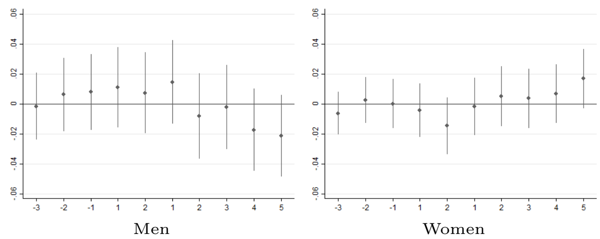
Figure 6.2: Gender Heterogeneity in Job Types



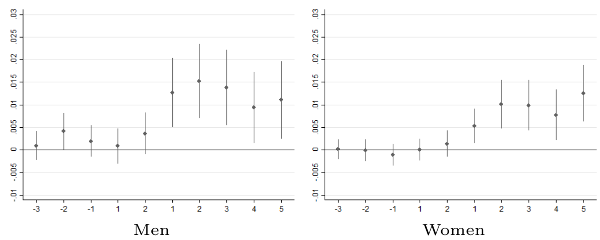
(a) Open-Ended Contract



(b) High Skill Occupation



(c) White Collar Occupation



(d) Employed in an Associated Job

Notes: These figures plot the coefficients for receiving an Offer in specification 3-1 done year by year separately by gender. Bars represent 95% confidence intervals. Open-Ended Contract is a dummy that equals one if the individual has at least one open-ended contract. High skill and white collar are dummies that equal 1 if the individual's main contract is in a high skill or white collar occupation. Employed in an Associated Occupation is an indicator variable that equals one if the individual has at least one contract in an occupation related to their chosen course.

7
Robustness

Robustness checks for sample selection

A point of concern in the analysis made is the fact that we exclude individuals that match more than once to RAIS from the sample. As a higher probability of being employed and appearing in RAIS would also naturally increase the chances that a prospective student matches to more than one person in RAIS, a positive effect of the offer could lead to sample selection and bias in our results[1]. To address this, we re-do the analysis with the inclusion of these individuals. To make this possible, we select the match with the higher probability of being the correct one based on the criteria to participate in the program and the course location[2]. With this, we include 12,938 new individuals, growing our sample by 30%.

The results are presented in Table 7.1, where we can see that the inclusion of these prospective students does not change the previous conclusions. Panel A shows that receiving an offer from the lottery does not predict any outcome before the start of the course. From panel B, we have the negative effects for the duration of the course. Here we see that the effects over employment and hours still appear, and the negative effects over wages, open-ended contracts and high skill employment that before were not significant now are, with similar point estimates from the ones presented before. This indicates that the results found with our main sample are conservative estimates for the negative short term effects of the program.

[1]Although those selected in the lottery have a slightly smaller probability of having more than one match in RAIS, which is consistent with there being only the negative effects during the course. However, the point estimate of a regression of the number of matches over the offer using specification 3-1 (-0.036) is not significant, with a p-value of 0.838.

[2]We do this by consecutively excluding matches that fail to meet certain expected criteria until only one remains. First, we use the city of the course and give preference to individuals that at some point appear with a contract in the same city. Next, we use the fact that only those that have completed high school before the start of the course can participate and exclude those that appear with less schooling after 2011 and those that appear with such schooling that it would not be possible to finish high school in the typical time before 2011. We also give preference to those that did not complete post-secondary schooling before 2011. For the remaining individuals that still had more than one match, we randomized a choice.

Table 7.1: Program Offer and Main Outcomes with Multiple Matches

	Employed	Hours	Wage	Open-Ended Contract	High Skill Occupation	White Collar Occupation	Associated Occupation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Panel A: Before			
Offer	0.001	0.051	0.004	0.002	-0.001	0.000	0.000
	(0.005)	(0.226)	(0.011)	(0.005)	(0.003)	(0.004)	(0.001)
Observations	439,936	439,936	439,936	439,936	439,936	439,936	439,936
R-squared	0.059	0.061	0.051	0.056	0.030	0.037	0.017
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	54992	54992	54992	54992	54992	54992	54992
Mean Dep Var	0.303	12.66	0.511	0.292	0.091	0.176	0.003
				Panel B: During			
Offer	-0.017**	-0.704**	-0.034**	-0.015**	-0.009*	0.001	0.001
	(0.007)	(0.303)	(0.016)	(0.007)	(0.005)	(0.006)	(0.001)
Observations	109,984	109,984	109,984	109,984	109,984	109,984	109,984
R-squared	0.048	0.051	0.037	0.049	0.038	0.031	0.030
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	54992	54992	54992	54992	54992	54992	54992
Mean Dep Var	0.533	22.17	0.953	0.508	0.157	0.328	0.005
				Panel C: After			
Offer	0.006	0.248	-0.005	0.004	0.005	0.006	0.010***
	(0.007)	(0.279)	(0.016)	(0.007)	(0.005)	(0.006)	(0.002)
Observations	274,960	274,960	274,960	274,960	274,960	274,960	274,960
R-squared	0.026	0.029	0.025	0.029	0.031	0.017	0.020
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	54992	54992	54992	54992	54992	54992	54992
Mean Dep Var	0.530	22.06	1.025	0.513	0.166	0.328	0.010

Notes: Panel A presents results for the panel of years before the start of the course (i.e., the sample for these regressions are panel years 2003-2010), Panel B for the years during the course (2011-2012) and Panel C for the years after the course (2013-2017). For the sample in this table, we also include applicants that match to more than one individual in RAIS, choosing the more probable match based on the program's prerequisites and location. *Offer* is an indicator variable that equals 1 if the individual was selected in the lottery. All regressions control for randomization strata and year fixed effects. Standard errors clusterized at the individual level are presented in parenthesis.

$* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$

In panel C, we have the results for the period after the end of the courses. Here we see that there are still no effects over employability and that the effects over formal employment on an associated occupation persist. However, the estimate for the effect on occupations with high skill requirements becomes smaller in size and turns non-significant.

8
Conclusion

In this paper, we estimate the impacts of a government-run large-scale vocational training program in the northeast of Brazil. Exploring the lottery used for the selection of students, we look at the effects over general labor market outcomes as well as over the types of employment contracts obtained by the applicants. For this, we explore a large administrative dataset that covers all formal employment contracts in Brazil (RAIS, *Relação Anual de Informações Sociais*) and allows us to examine outcomes such as formal employment, wages, experience, tenure, long-term or temporary employment and type of occupation year by year up to 5 years after the end of the program for this cohort of students.

We find that selection to the program did not lead to higher levels of formal employment, hours worked or wages. This is in line with what is found in the general literature, although in contrast with some studies from other training programs in Brazil. We also find a short-term negative effect on employment and hours worked during the course. As this could have long term consequences and be compounding the null results of the course, we show that receiving an offer through the lottery had no impact over tenure in the current contract or total labor market experience in any period. This way, as there are no general or job-specific differences in experience, there are no indications that the negative impacts during the course are counterbalancing possible positive effects of the program. We additionally show that the program did not affect the probability of termination of contract or of starting a new job except for a higher hiring rate in the first year after the course that closes the employment gap created during the program. Thus, the program did not lead to more stability in the labor market. However, as we only observe formal employment, we are not able to measure these effects over informal employment and earnings. Further, as we do not have information on take-up, our results only reflect the effects of a selection in the lottery and cannot be extrapolated as the effects of completing the course.

Which does not mean that we find no positive effects for the program. We find that the program was effective in matching its students to jobs with a higher skill demand and to occupations related to their course of choice. Those selected in the lottery had a 7.2% (1.1 p.p.) higher chance of being employed in an occupation that required a higher level of skill. This result is compatible with the program generating gains in the skill level of students. Selected individuals also had a higher probability of obtaining a formal contract in an occupation associated with the course chosen. Being drawn in the lottery was related to a 1 p.p. raise in the probability of being employed in an occupation associated with their course. It is also associated with a higher tenure in associated jobs that increases over time, reaching a half month advantage in the fifth year after the program's completion. So students are not only matching better to their jobs but also keeping them for longer.

We also show how the effects vary by the student's gender. The first difference found is that incapacitation effects appear only in the subsample of women. While women work less on average during the course (with a 2.7 p.p. difference compared to women that were not selected), men appear to keep the same level of employment regardless of being drawn or not in the lottery. Additionally, men are the only ones that experience gains in high skill jobs, with a 10.8% increase. And, while both genders have similar effects over associated employment, for men the effect is already fully realized in the first year after the course, while for women this only happens in the second year. Thus, the program appears to have higher benefits for men, and at the same time, higher costs for women. However, these results should be taken with a grain of salt, as we are not able to isolate gender effects from the difference between the courses chosen by each group.

Overall, although we do not find evidence that the program generated gains for employment or wages, we find that it had a significant impact on the types of occupations that the students were able to obtain, leading them to higher skilled and better matched activities. Especially when considered in light of the fact that the program was aimed at individuals that would otherwise not be able to procure other qualifications, this provides evidence that this type of program could bring other benefits besides higher probabilities of procuring a job. By expanding education opportunities and the jobs available to those selected in the lottery, the program is successful in its objective of boosting the job prospects of its students.

Bibliography

Abowd, J. M., Kramarz, F., Margolis, D. N., and Troske, K. R. (2001). The Relative Importance of Employer and Employee Effects on Compensation: A Comparison of France and the United States. *Journal of the Japanese and International Economies*, 15(4):419–436.

Alfonsi, L., Bandiera, O., Bassi, V., Burgess, R., Rasul, I., Sulaiman, M., and Vitali, A. (2020). Tackling Youth Unemployment: Evidence from a Labour Market Experiment in Uganda. *Econometrica, forthcoming*.

Attanasio, O., Guarín, A., Medina, C., and Meghir, C. (2017). Vocational Training for Disadvantaged Youth in Colombia: A Long-Term Follow-Up. *American Economic Journal: Applied Economics*, 9(2):131–143.

Attanasio, O., Kugler, A., and Meghir, C. (2009). Subsidizing Vocational Training for Disadvantaged Youth in Developing Countries: Evidence from a Randomized Trial. IZA Discussion Papers 4251, Institute of Labor Economics (IZA).

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science*, 344(6186):843–851.

Barbosa, A. C. F. (2011). Política pública para a educação profissional na bahia: o plano de educação profissional. Master's thesis, Universidade Federal da Bahia.

Barría, C. V. and Klasen, S. (2016). The impact of SENAI's vocational training program on employment, wages, and mobility in Brazil: Lessons for Sub Saharan Africa? *The Quarterly Review of Economics and Finance*, 62(C):74–96.

Calero, C., Gonzalez Diez, V., Soares, Y. S., Kluve, J., and Corseuil, C. H. (2017). Can arts-based interventions enhance labor market outcomes among youth? Evidence from a randomized trial in Rio de Janeiro. *Labour Economics*, 45(C):131–142.

Camargo, J., Lima, L. S. e., Riva, F. L. R., and Souza, A. P. F. d. (2018). Technical education, noncognitive skills and labor market outcomes: experimental evidence from Brazil. Textos para discussão 480, FGV EESP - Escola de Economia de São Paulo, Fundação Getulio Vargas (Brazil).

Card, D., Ibarraran, P., Regalia, F., Rosas, D., and Soares, Y. (2007). The Labor Market Impacts of Youth Training in the Dominican Republic: Evidence from a Randomized Evaluation. NBER Working Papers 12883, National Bureau of Economic Research, Inc.

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

Confederação Nacional da Indústria (2014). Pesquisa CNI-IBOPE: Retratos da Sociedade Brasileira: Educação Profissional. Technical report, Confederação Nacional da Indústria, Brasília, Brazil.

Corseuil, C. H., Foguel, M. N., and Gonzaga, G. (2019). Apprenticeship as a stepping stone to better jobs: Evidence from Brazilian matched employer-employee data. *Labour Economics*, 57(C):177–194.

da Silva, E. M. (2009). A implementação do currículo integrado em agropecuária: o caso guanambi. Master's thesis, Universidade de Brasília.

de Oliveira, L. W. (2011). O centro estadual de educação profissional da bahia e os desafios da educação profissional: Análise dos cursos subsequentes. Master's thesis, Universidade Federal da Bahia.

DIEESE (2012). Anuário da educação profissional da Bahia, vol II. Technical report, Departamento Intersindical de Estatística e Estudos Socioeconômicos. – Salvador.

Filho, N. M. and Nuñez, D. F. (2011). Estimando os gastos privados com educação no brasil. Technical report, Centro de Políticas Públicas - Insper, São Paulo.

Grosz, M. (2020). The Returns to a Large Community College Program: Evidence from Admissions Lotteries. *American Economic Journal: Economic Policy*, 12(1):226–253.

Ibarrarán, P., Kluve, J., Ripani, L., and Shady, D. R. (2019). Experimental evidence on the long-term effects of a youth training program. *ILR Review*, 72(1):185–222.

McKenzie, D. (2017). How effective are active labor market policies in developing countries? a critical review of recent evidence. Policy Research Working Paper Series 8011, The World Bank.

Menezes-Filho, N. A., Muendler, M.-A., and Ramey, G. (2008). The Structure of Worker Compensation in Brazil, with a Comparison to France and the United States. *The Review of Economics and Statistics*, 90(2):324–346.

O'Connell, S. D., Mation, L. F., Bevilaqua Teixeira Basto, J., and Dutz, M. A. (2017). Can business input improve the effectiveness of worker training? evidence from Brazil's Pronatec-MDIC. Policy Research Working Paper Series 8155, The World Bank.

Reis, M. (2015). Vocational Training and Labor Market Outcomes in Brazil. *The B.E. Journal of Economic Analysis & Policy*, 15(1):1–29.

Schwartzman, S. (2016). *Educação Média Profissional no Brasil: situação e caminhos*. Fundação Santillana, São Paulo.

Secretaria de Educação do Estado da Bahia (2011). Legislação básica 2010-2011.

Souza, A. P., Lima, L., Arabage, A., Camargo, J., de Lucena, T., and Soares, S. (2015). Vocational Education and Training in Brazil. Technical report, Knowledge Sharing Forum on Development Experiences: Comparative Experiences of Korea and Latin America and the Caribbean.

A
Data

A.1
Data Construction

This paper uses two main data sources: the first one is the public results of the electronic lottery of those who applied to participate in the VET courses, made available by the secretariat of education of the state of Bahia's government. From this, we have the names of everyone who signed up to participate in the courses, regardless of whether they were admitted or not. We also have the city, school, course, and shift for which the individual signed up. Finally, it also has the number assigned to the individual in the draw[1] and if they were selected or not. The other dataset used is the RAIS (Relação Anual de Informações Sociais) database. This is a large administrative dataset from the Ministry of Labor containing all formal employment contracts in Brazil. All enterprises in Brazil are required by law to declare all employment relations to the Ministry. The result is a restricted-access panel in the firm-employee level. We use RAIS from 2003-2017, comprising 8 years before the start of the course, the 2 years when classes were held and 5 years after.

The link between both datasets is done using the name of the individuals, as this is the only characteristics of the individuals present in the original lottery list[2]. For that, the following procedure was used: The first step was to restrict RAIS to only contracts in the state of Bahia, where all the courses where held. After, all RAIS years were condensed in one large list with all the combinations of full name and id. This list was used to merge with the list of individuals registered for the draw, using an exact name match. An individual can have zero, one or more matches in RAIS, the ones with zero or

[1] The lottery assigns an order for the students in each school, course and shift combination. A student is selected to participate in the course if their position is below the cut-off of the number of spots in the class.

[2] The list also gives the municipality of the course for which the individual applied. We choose not to use this in the match as we could introduce a bias if individuals that do not live in those cities migrate if selected for the course. If this were the case, selected individuals would have a higher match rate in RAIS. We, however, use this in our robustness analysis, where we select a match for those individuals that have more than one in RAIS.

one comprise our main sample[3]. We then follow the individuals in each RAIS year using id and first name[4].

This way, from the initial 60.487 applicants that entered the draw, 1.968 were removed from being duplicates or having invalid information and 13.429 where removed from the sample for having more than one possible match in RAIS. We also remove from the sample individuals who tried for courses in Cultural Production and Design, where an additional ability test was required and admission was therefore not random. This removes 400 individuals from the sample. We also remove individuals that tried for classes where oversubscription did not occur. In these classes, all individuals were accepted. This removes an additional of 2.389 individuals. Lastly, we remove 247 individual for whom we are not able to obtain information on gender[5]. Our final sample comprises of 42.054 individuals that tried for 37 different courses in 40 cities in a total of 149 classes[6].

As some individuals have more than one contract in a given year, we also choose a main one to get measures such as hours, wages and tenure. For this, we select first the contract with the highest number of hours contracted, then the one with the highest wage. If there are still more than one contract that are equal in these two measures, we randomize a choice.

A.2 Definition of Main Variables

Here are definitions for the main variables used in the analysis:

- Offer: It is a dummy that indicates that an individual was offered a spot in the course pretended based on results of the draw.[7]
- Employment: an indicator variable that equals one if a person was formally employed in a given year. An individual is considered formally employed if they appear with at least one contract for that year in RAIS.

[3] We also do robustness analysis using the sample including the individuals with multiple matches, with tiebreaker criteria using the prerequisites for the course. The results are similar to the ones found for the main sample.

[4] We use both id and first name to avoid some errors that may appear in a dataset of this size, with the id number being mistakenly allocated to a different person in an observation. We use the first instead of the full name to avoid losing observations to spelling mistakes and abbreviations and to still follow individuals that changed last names (due to marriage, etc.) after the course.

[5] As we do not observe gender directly for all applicants, we construct it as the mode of the gender for every observation in RAIS in the sample period that has the same first name as the individual.

[6] A class is defined here as all the students that tried for the same course in the same school in the same shift.

[7] If someone offered a place failed to enroll, another person would be called in their place following the order given in the draw. We do not have information on those additional offers.

- Employment in an associated job: indicator variable that equals one if the individual is formally employed in at least one position related to their course of choice. To determine which positions are associated with each course, we used the 3rd edition of the National Catalog of Technical Courses put forward by the Ministry of Education. There, a small list of CBO codes (the Brazilian Classification of Occupations) is given as the ones targeted by each course. As this is a very restrict list for each course, we use the families of the occupations given[8].
- Employment in an open-ended contract: an indicator for the case that the person is employed with at least one contract without a predetermined expiration date. As individuals can have more than one contract declared at RAIS, it is possible that they also have a short-term contract.
- Employment in a short-term contract: an indicator for the case that the person is employed with at least one temporary contract. As individuals can have more than one contract declared at RAIS, it is possible that they also have an open-ended contract.
- Hours: The number of hours per week worked on the main contract. If the person is not formally employed in a given year, this value is zero.
- Wage: The value of the real wage per hour for the main contract in December[9]. If the person is not employed in a given year, this value is zero. All monetary variables are inflation adjusted for 2010 values. To this, we apply the inverse hyperbolic sine transformation.
- Tenure: Tenure in months in the main contract. If the person is not employed in a given year, this value is zero.
- Tenure in an associated job: Tenure in months for the longest standing contract that is in a position associated with the course pretended. If the person is not employed in an associated job in a given year, this is zero.
- High-skill job: A dummy variable that indicates if the individual's primary contract is in a high-skill position. For this, we use the same procedure utilized by Corseuil et al. (2019) that uses a mapping of CBO to ISCO-88 codes[10] and separates the occupations in high and low skilled groups using Abowd et al. (2001)'s classification. From ISCO-88 codes, high skilled occupations are the ones with codes starting with 1, 2, 3, 6, 7 or 8.

[8] A family contains about five related occupations. This is the more disaggregated classification after the individual occupations, comprised of the 4 first numbers of the code.

[9] wage by hour = monthly wage / (hours $\times$ 4.348)

[10] However, differently from them, we have an additional link, going from ISCO-88 to the CBO in the 2004 version and them to the CBO in the 2012 version, which is the one we have for all years.

- White-collar job: An indicator variable for if the person is employed in a white-collar job. Just as for high skill occupations, we define white-collar jobs following the classification of Abowd et al. (2001) used in Menezes-Filho et al. (2008) and Corseuil et al. (2019), using a mapping from the CBO codes in RAIS to ISCO-88 codes. An occupation is considered a white-collar one if its ISCO-88 code starts with 1, 2, 3, 4 or 5.
- Separation: A dummy variable that indicates if an individual had their main contract ended for any reason. Besides hired or resigned, this could include death, retirement, accumulation of multiple public positions, etc.
- Number of jobs: The number of active contracts for each person in a given year.
- Employed in a new firm: An indicator variable that equals one if the person is employed only on new firms from which they were not employed in the previous year.
- Large Firm: A dummy variable that indicates if the person is employed in a large firm. A firm is considered large if they have at least 250 employees.
- Experience: Number of months that the individual was employed since 2003 as of the end of the year.
- Worked Before: A indicator variable for if the prospective student had any type of formal employment before the year the course started. It equals one if the individual had at least one contract listed between 2003 and 2010.