University of the Basque Country
Department of Foundations of Economic Analysis II

Master in Economics: Empirical Applications and Policies (EAP)

MASTER THESIS
by Dimitris Sgourinakis

Novia Salcedo Foundation courses and the probability of employment

Supervisor: Javier Gardeazabal,
University of The Basque Country

Bilbao, 7/2013

Acknowledgments

I do want to acknowledge all the professors of the *Master in Economics: Empirical Applications and Policies (UPV/EHU)*, who offered me this great and productive year full of experiences. Especially I would like to thank my supervisor, *Javier Gardeazabal*, and my professor in Quantitative Methods, *Petr Mariel*. Moreover, I would like to express my gratitude to my classmates who were very supportive and offered me the opportunity to learn their different cultures. I owe a debt to professor *Ainara González de San Román* who provided me with all the necessary data, transformed in a convenient way to use. At this point I would like to thank *Novia Salcedo Foundation* and the people who are working there (they have been remarkably patient), because without them it would not be possible to conduct this survey. It would be remiss of me not to mention and also thank this beautiful city, *Bilbao*, that was not only my co-writer during the past three months but also the source of my concentration. Finally, I dedicate this Master Thesis to my brother, *Antonis Sgourinakis*. It was him who supported me all this year every time I was feeling to give up on myself. Last but not least, I would like to express my deepest feelings of love and gratitude to my parents, *Nikos Sgourinakis* and *Panagiota Stamoulia*.

"Γιατί οι άνθρωποι, σύντροφε, ζουν από τη στιγμή
που βρίσκουν μια θέση στη ζωή των άλλων..."

Τάσος Λειβαδίτης

"Because people, my friend, start to live from the moment
they find a place in the lives of the others..."
Tasos Livaditis

Table of contents

Novia Salcedo Foundation courses and the probability of employment

Abstract

In our days, the economic recession is pushing young people to find more efficient ways in order to come up with a job. This article investigates the contribution of *Novia Salcedo Foundation*, on the employment for young people. The survey mostly focuses on individuals from Spain and more specifically from the Basque country. When analyzing the database, we come up with results regarding the effectiveness of the training offered by the Foundation. Specifically, we focus on some particular non-mandatory courses. Moreover, we observe statistical significant estimations on the effectiveness of those courses, mostly for the period 2007-2009, when the Foundation did not have any mandatory training.

1. Introduction

Firstly, we will refer to our motivation that prompted us to choose and investigate this dataset. Also we will mention some characteristics of *Novia Salcedo Foundation*, so to have a spherical idea on how the Foundation is dealing with the individuals and with which ways is helping them in order to exit unemployment after a scholarship training program in some firm.

1.1 Motivation

The database provided by the Foundation gives us the opportunity to put into practice methods we have learned during this master-year. As a result, the Foundation has been transformed to an ideal ground to apply empirical methods. Another aspect that powered our motivation is the large number of observations we are going to deal with. Almost 17,000 observations were gathered, with the first individual registered in 2007, one year before the break out of the financial crisis in Spain. Statistics related on the unemployment of Spain show us that the trend of unemployed young people in the Basque Country is lower than the country's overall. Despite this lower trend, a significant difference, opposite to what is observed for the whole population, can be observed specifically for young individuals with high level of education. For instance, in the first quarter of 2013 the percentage of the these unemployed individuals in the Basque Country reaches 40.45%, as in the whole Spain is only 24.43%[1]. We will examine how *NSF* treats this difference in the Basque Country.

[1] Source: Encuesta de Población Activa (EPA).

1.2 The Novia Salcedo Foundation (NSF)

The *Novia Salcedo Foundation* is a private, non-profit cultural organization, headquartered in Bilbao with over 30 years of experience. The *NSF*, aims to help young people who are unemployed and are currently searching for a job and also offers various types of courses to individuals. To achieve this, from a vocation of anticipation and collaboration, the Foundation invents, creates and provides social value, developing lines of thought and creative processes of transformation, towards new modes of value exchange between individuals, companies organizations and public institutions.

To cover the constantly evolving changes in the labor market, *NSF* has developed the *NSF Human School*. The *NSF Human School* offers to the participants unique opportunities to acquire skills, talents and values demanded by companies and acts as an active and informal learning platform. Training programs, courses for career orientation and practice in firms, can be some of the features someone can come up with in this guide-school. More specifically: *Training* is provided in these emerging sectors through a diverse range of modules delivered with support and cooperation of leading organizations in the field. *Coaching* is also offered to young people. *NSF Human School* provides individual guidance or group coaching in areas such as interview skills, composing a curriculum, resume and cover letter elaboration and career orientation. *Legal Advice*, which has as main subject services to young people regarding a wide range of legal challenges, including employment issues, labor rights, problems relating to housing rental and consumer's protection. *Local Work Placements*, the skilled employees of *NSF* act as agents between young individuals and local companies, seeking young talent. This task is carried out by promoting work placement opportunities through its website. The filling of these local vacancies is often accompanied by contracts that last from four to nine months. *International Mobility Program*, to further enhance the employability of young people, *NSF Human School*, in partnership with various donors including the EU *Leonardo da Vinci*[2] program, gives unique opportunities of travelling abroad to integrate in the international labor force. Starting from 2004, the Foundation is sending almost one hundred experienced young professionals each year in leading organizations, allowing these individuals to acquire new skill sets and then contribute into the local economy. *Social Entrepreneurship*, is a program which through the collaboration between experts, companies and entrepreneurs, gives the chance to discover opportunities that are transformed into economic activities. In short, serves to highlight the value of collaborative entrepreneurship.

[2] Part of the European Commission's Lifelong Learning Program, Leonardo da Vinci program funds many different types of activities. These include ‘mobility’ initiatives enabling people to train in another country.

2. Data

The data used in this particular Master Thesis, are gathered by *NSF* over the past six and a half years, from 2007 to the first months of 2013 (February 2013), and consist of detailed information of all registered individuals. A wealth of basic respondent characteristics can be found, such as: gender, level of education, knowledge level of foreign languages, a number of skills. Moreover, other variables on which we will base our research have been reported. For instance, we observe if the individual has participated to some of the training courses of *NSF*, if the individual has signed a job contract and the type of the contract.

The initial number of observations reported by *NSF* is almost equal to 17,300. The problem we are facing is that we do not observe, for the whole sample, whether the individuals have signed a contract after their participation to the *NSF* training programs. So we cannot answer to our initial question, that is whether the Foundation is effective in finding a job to individuals, when using the total amount of observations. The sample lacks information on the non-participants in the training-scholarship programs. That is why we are reducing our sample by keeping only 3,186 observations and later on, for our analysis, the sample is reduced even more at those who take the non-mandatory courses in the Foundation. So, our final question will be: Do the training courses of *NSF* help those 3,186 individuals, to exit unemployment? The descriptive statistics for these observations can be shown in the next chapter, in Table 1.

2.1 Descriptive Statistics

Below, we display the means and the standard deviations for the variables we are going to include in our research. Most of the individuals that have registered in *NSF*, during the years we are interested in, are from Spain. The exact number reaches 3,186 respondents with the majority to come from the Basque country, Bizkaia, Gipuzkoa and Araba. More specifically, 2,196 have recorded as citizens of Bizkaia, 605 of Gipuzkoa and 287 of Araba.

Dealing with observational (non-randomized) data, leads us to some substantial differences in the shape of the covariates distributions. In randomized studies, or simply experiments, the covariate distributions for the treated and the control group are expected to be perfectly balanced. For instance, in our case, almost 46% of the respondents are males. So to say we cannot consider the sample to be equally distributed respect to gender. We will deal with this difficulty more explicitly in the third chapter.

Table 1:Descriptive Statistics[a]

Variables	*Mean (Std.Deviation)*	*Variables*	*Mean (Std.Deviation)*
Age	*28.248 (2.756)*	**Knowledge of Programming (percentage)**	*35.687 (0.479)*
Gender		**Knowledge of Databases (percentage)**	*56.057 (0.496)*
Females (percentage)	*54.237 (0.498)*	**Knowledge of Design (percentage)**	*56.340 (0.496)*
Education		***Programs***	
Very low level of Education (percentage)	*1.600 (0.125)*	**Training (percentage)**	*45.072 (0.497)*
Low level of Education (percentage)	*18.989 (0.392)*	**International Scholarships (percentage)**	*1.349 (0.115)*
Medium level of Education (percentage)	*25.768 (0.437)*	**Individuals that have attended previously courses of quality (percentage)**	*43.628 (0.496)*
High level of Education (percentage)	*53.640 (0.498)*	**Individuals that have attended previously courses of ISO (percentage)**	*16.038 (0.367)*
At least one Master (percentage)	*24.427 (0.427)*	***Other***	
Skills		**With recorded Experience (percentage)**	*37.288 (0.483)*
Excellent knowledge of Basque language (percentage)	*50.753 (0.500)*	**Number of contracts signed in the past by the respondent**	*0.774 (1.318)*
Excellent knowledge of English language (percentage)	*36.472 (0.481)*	**Year of completion of the practice**	*2010.274 (1.858)*
Excellent knowledge of French language (percentage)	*3.986 (0.195)*	**Individuals who own a car (percentage)**	*70.495 (0.456)*
Excellent knowledge of German language (percentage)	*0.847 (0.091)*	***Number of observations***	*3,186*

[a] Variables age, year of completion of the practice, number of contracts signed in the past and percentages are computed only for those who got a training-scholarship, either national or international, over the sample from 2007 to 2013.

The average age of the sample is about 28 years old. From our data, it turns out that 24.4% of the individuals with a high level of education[3], obtain at least one master. When revising the data, we come up with 19 observations registered as having a low level of education[4] as well as a master. We treat these observations as outliers and we delete them from our sample. Regarding the rest levels or skills of education we get that, 18.9% of the population seem to have a low level of education and 25.7% a medium level[5]. 1.6 percent of the population has a very low level of education[6]. Moreover, approximately half of the population knows Basque, unlike to only 4% that can speak French. 36.4% of the respondents are reported as excellent users of the English language, unlike to only 0.8% who have excellent knowledge of the German language. We observe that around 37% of the respondents had former job experience before registering in *NSF*. Finally, the treatment (courses) has been received by 45% of the individuals.

Other variables which reveal some interesting information are the ones that show the type of contract (Table 2) that the individual has signed after receiving the training, as well as the variable that depicts which firms have signed a scholarship training contract (Figure 1), which ones have renewed the contract, and so on. We will divide the sample in two groups based on the period of time; from 2007 to 2009 (period 1) and from 2010 untill the first months (February) of 2013 (period 2). The separation is stationed in these exact periods after significant differences that were found when handling the data. An important role to the decision and separation of the data, was also played by the unemployment rate. As we are focusing on the probability of employment, unemployment rate can have a significant influence in our survey. Regarding the two periods, we can observe vital fluctuations in the unemployment rate in the Basque country. For instance, the annual unemployment rate for Bizkaia before 2010, was under 9% and in Gipuzkoa was under 6.5%. After 2010, the annual unemployment rate, ranges from 10% up to 15.8% for Bizkaia and from 7.4% up to 11.1% for Gipuzkoa[7]. Let us cite the tables below with the descriptive statistics of the variables of interest that we mentioned before:

[3] Five years of university.
[4] High school education.
[5] Three years of university.
[6] Primary education.

[7] Source: Euskal Estatistika Erakundea - Instituto Vasco de Estadística (Eustat). Eustat is a public body of the Basque Country that collects, analyses and publishes statistical information concerning aspects of the Basque Country.

From our sample we can observe that 1,664 out of 3,186 individuals (52.23%) have signed a contract, in some firm, after their training period. So, with a first glance we can say that *NSF* is very effective as more than half of the population registered in the records of *NSF* comes up with a job. In this paper we are interested more in particular, non-obligatory courses that are provided by the Foundation.

Table 2: Type of contract signed, after the training period, by period of time

Year	*2007-2009*	*2010-2013*
Indefinite contract	*6.01*	*2.52*
Fixed-term contract	*33.24*	*21.21*
Practice contract	*18.83*	*14.89*
Other contract	*3.27*	*1.43*
Not observed/Not responded	*38.66*	*59.96*
Number of observations	*1,715*	*1,471*

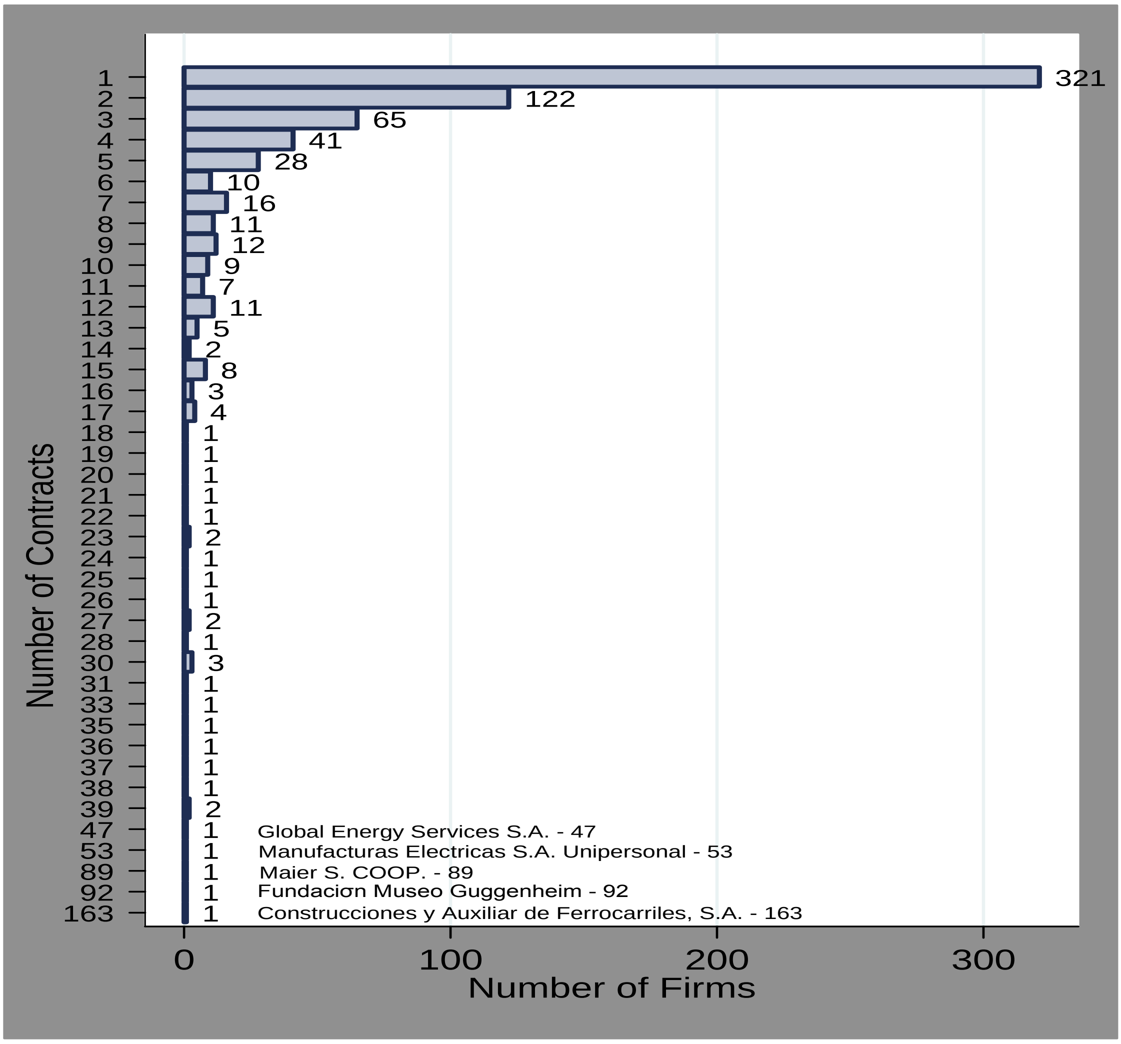

Figure 1: Number of scholarship training contracts signed by amount of firms

From Figure 1 we can observe that 321 firms have signed one scholarship training contract with some individual registered in *NSF* records. 122 firms have signed two contracts, 65 firms have signed three contracts and so on. In the plot we have added the names of the companies that have signed more than 40 contracts during these 6.5 years: first in this ranking, appears to be the firm with name "*Construcciones y Auxiliar de Ferrocarriles S.A.*" with 163 contracts over time. Second, with 92 contracts, is *The Foundation of the Guggenheim Museum*, mostly with short term contracts. The third firm is "*Maier S. COOP*." with 89 contracts. Finally in the two last places, of the firms that have signed more than 40 contracts with individuals through the Foundation, are "*Manufacturas Electricas S.A. Unipersonal*" (53 contracts) and "*Global Energy Services S.A.*" (47 contracts).

3. The Econometric Methods

In this Chapter we will refer to the Econometric Methods which are used in this paper in order to come up with our estimations. The three methods we are going to use are: the Average Treatment Effect, the Matching method and the Sequential Selection of Covariates. The Average Treatment Effect will lead us to results related on the effectiveness of the non-mandatory courses that are provided by the Foundation. The usage of the Matching method will reduce our sample and keep only the registered individuals with similar propensity scores, so to improve the significance of our estimations. And the Sequential Selection of Covariates will help us to choose the covariates that explain better our dependent variables.

3.1 The Average Treatment Effect (ATE)

In order to refer extensively to the econometric method which is going to be used, we should start by introducing the meaning of the Average Treatment Effect and more particularly the case in which the explanatory variable of interest is binary. In our days, estimating the ATEs has become crucial in many fields such as the evaluation of school voucher programs or job-training programs. Also, many of the early applications of the ATE were used in medical interventions, mostly proven to be useful in situations where experiments are clearly impractical. Some of the vocabulary that we are using while describing this method, has been kept because of these medical interventions ("treatment and control group"). The first references appeared in a counterfactual framework pioneered by Rubin (1974). The framework model is known as the Rubin Causal Model (RCM) and since then it has been adopted by many other authors in statistics, econometrics and other fields.

In this paper we will study and apply the assumptions which are used in methods of the ATE where the observed outcome is a binary variable. Let us denote by D_i the variable which indicates the binary treatment, in our case this variable shows whether or not the individual has received training courses through *NSF*:

$$D_i = \begin{cases} 0, & \mathit{if\ the\ individual\ did\ not\ participate\ in\ training} \\ 1, & \mathit{if\ the\ individual\ participated\ in\ training} \end{cases}$$

By Y_i we are going to denote the observed outcome, in our case, whether or not the individual has signed a contract with some firm after registering in the Foundation. At this point we remind that we are interested only in those that participate in a scholarship training program:

$$Y_i = \begin{cases} 0, & \mathit{if\ the\ individual\ has\ not\ signed\ a\ contract} \\ 1, & \mathit{if\ the\ individual\ has\ signed\ a\ contract} \end{cases}$$

Let Y_1 denote the outcome with the treatment and Y_0 the outcome without treatment. Because the respondent cannot participate in both outcome states, we are not able to observe both Y_0 and Y_1. The triple vector (Y_0, Y_1, D_i) represents a random vector from the underlying population of interest. If we choose randomly an individual *i* from the whole sample, the triple vector associated with this draw will be (Y_{i0}, Y_{i1}, D_i). In order to measure the effect of the treatment in the population, we are interested in computing the difference in the different outcomes, $Y_1 - Y_0$. As noted before, the difficulty we are facing is that we only observe one of the potential outcomes, Y_1 or Y_0. More precisely, when introducing the treatment D_i, the observed outcome in terms of potential outcomes will be:

$$Y_i = (1 - D_i)Y_0 + D_i Y_1 = Y_0 + D_i(Y_1 - Y_0)$$

Using the triple vector described above, as well as, X_i that denotes the covariates of our sample, we can write the ATE as follows:

$$ATE = E\big(E(Y_{1i}|X_i) - E(Y_{0i}|X_i)\big) = E\big(E(Y_{1i}|X_i, D_i = 1) - E(Y_{0i}|X_i, D_i = 0)\big)$$
$$= E\big(E(Y|X_i, D_i = 1) - E(Y|X_i, D_i = 0)\big) = E\big(\hat{Y}_{1i} - \hat{Y}_{0i}\big)$$

We have managed to express the ATE as an outcome of conditional expectations, as it can be shown above. These conditional expectations depend on observables and they can be consistently estimated. That is, because we are dealing with a binary outcome, the function $\Lambda(X_i'\hat{\beta}_1)$ is obtained by the logistic estimation, using only the respondents who have received the treatment. Respectively, $\Lambda(X_i'\hat{\beta}_0)$ is obtained by using only those that belong to the control group:

$$E(Y_i|X_i, D_i = 1) = \Lambda(X_i'\hat{\beta}_1)$$

Using our sample, a consistent estimator of ATE will be:

$$ATE = \frac{1}{N}\sum_{i=1}^{N}\left(\Lambda(X_i'\hat{\beta}_1) - \Lambda(X_i'\hat{\beta}_0)\right)$$

Another measurement that we are going to derive from our data will be the Average Treatment Effect for the Treated (ATT). This effect depicts the mean effect, as the ATE, but this time only for those who actually participated in the scholarship training program, in our case, for those who participated in the non-mandatory courses of the Foundation. The ATT is equal to: $ATT = E(Y_{1i} - Y_{0i}|\, D_i = 1)$.

Another method we are interested in is the Bootstrap method. Approximations for finite-sample results and related test statistics that help us rely on asymptotic theory, can be provided by this method. We will use the Bootstrap method in order to estimate the standard errors in the logits used in the survey for specific values of the covariates. The bootstrap method introduced firstly by Efron (1979, 1982).

3.2 Matching

As mentioned in chapter 2.1, in this survey we are dealing with unbalanced covariate distributions. For instance let us take the variable gender. As we showed in descriptive statistics, males and females are not equally distributed along the sample. Researchers can come up with this type of difficulties in observational studies in which we have no control over the assignment of the treatment to individuals. In order to avoid problems like this, before a non-randomized survey begins, we can include a "composition" procedure where a sample is constructed in such way that the inferences, in the end, are more robust and credible. In more simple words, by "composition" phase we refer to a well-built sample where the control and treatment subsamples are more balanced. At this point it is crucial to underline that we should not confuse the balance mentioned before, with the covariate balance on the probability of each individual to be exposed on the treatment (propensity score). Balance here refers to the similarity of the marginal covariate distributions in the two treatment arms. The two covariate distributions are expected to be exactly balanced in randomized experiments.

In this chapter we will explain the matching method we are using in order to solve the problem of unbalanced covariate distributions described above. The two most common methods are the Mahalanobis metric matching and the propensity score matching. The two methods differentiate in the way we use them to measure the distances between units. In our case, we will construct the control sample by matching one distinct control to each unit of treatment by using the propensity score matching, where the distance is measured solely in terms of the difference in the propensity score or more precisely, a monotone transformation of the probability of each individual to be exposed on the treatment such as the logarithm of the odds ratio.

If by $p(x) = (D_i = 1|x)$ we denote the observed propensity score, the logarithm of the odds ratio will be:

$$l(x) = ln\{p(x)/[1 - p(x)]\}$$

Let us assume that we observe N_1 treated units in our sample and N_0' control units such as $N_1 < N_0'$. Our goal is to select an amount of control units of size N_0, that is going to be less than the initial one, N_0', in order to construct a sample of size:

$$N = N_0 + N_1$$

In this procedure we can select the number of matches for each unit of treatment by including a number $M \geq 1$, in the matching equation:

$$N_0 = M * N_1$$

Here we are going to focus on the one-to-one case where $M = 1$, not only because almost half of the individuals in our survey were exposed to the treatment but also we consider this choice as the one with the smoothest impact in the variance of the control sample size N_0. As we said previously, the matching will take place based on the propensity score, or more accurately, the highest values of the estimated propensity score will be matched first. The most important part which constitutes a side of the solution to our problem, is that adjusting for differences in the propensity score between control and treated groups, eliminates all biases associated with differences in observed covariates. To include the differences in the logarithm of the odds of ratio rather than the differences in the propensity scores, as the last transformation takes into account the fact that typically the difference in propensity scores of 0.15 and 0.10 is larger in substantive terms than the difference between 0.50 and 0.45, the difference we are interested in can be depicted as follows (Z_i refer to the covariates that exist in different groups of the sample, as X_i). For each X_i, a Z_i is chosen so that it minimizes the following distance:

$$d(x, z) = [l(X_i) - l(Z_i)]^2 = \left[ln\left(\frac{p(X_i)}{(1 - p(X_i)}\right) - ln\left(\frac{p(Z_i)}{(1 - p(Z_i)}\right)\right]^2$$

By using this way of matching we are going to construct a new sub-sample of our original data, where the covariate distributions are going to be well balanced. In the new sub-sample the number of observations that remain is reduced by about 5/6 of the original sample. So to say, we end up with 456 individuals in period 1 (2007-2009) and 494 in period 2 (2010-2013)[8]. All the figures related to the matching method, that depict the distribution of the propensity score before and after the matching for each one of the periods as well as for the whole sample, can be shown in the Appendix (Figure 2, 3, 4, 5, 6, 7).

[8] The procedure was conducted after executing a logistic regression for each period and the logarithm of the odds ratio of the propensity score was matched based on the entire sample (2007-2013).

3.3 Sequential Selection of Covariates

In many studies, the selection of covariates, in order to obtain the best possible coefficient estimations, can create crucial problems. Generally, in every different survey, attention must be paid to a couple of things, in order to conduct the research and export significant results with the most efficient way possible. Firstly, the coherence of our thoughts should be based on a well-built economic model that is going to be surrounded by rational assumptions. For instance, we cannot conduct a survey in which we are willing to estimate the influences of the European art on the standard of living in Europe, in the 70s, and have a sample only from American citizens. We are going to fail even if our sample is the biggest one in the history and contains numerous covariates. Secondly, when we come up with a convenient sample for our survey, we should consider which must be the covariates that are going to be included. The amount of covariates in large-scale surveys sometimes can be very big respectively to the number of observations. For this reason, several methods have been developed in order to choose the most "meaningful" ones which will explain a big part of our dependent variables and give us back accurate estimations.

One of these methods is the Sequential Selection of Covariates and Interactions using a stepwise procedure. This procedure has two simple steps. We will use this method both for selecting the covariates in order to estimate the propensity score, as well as to estimate our dependent variable which is whether the individual has signed a contract or not. The first step consists on setting the basic covariates. So to say, depending on the subject we are dealing with we should select and start our regressions with some fixed covariates. For example, in labor economics, lots of papers find the gender, the experience, the training, the country and a number of other covariates are significant and claim that they should be included as fixed. In this paper we will consider as a fixed variable only the gender which paradoxically, later on it is shown that is significant, in a 5% significance level, only for the treated group, for the period 2 (2010-2013). In the second step we are including in our logit estimation (as our outcome is a binary variable) one covariate at a time and recording the likelihood ratio and *z*-statistics. At every "round" we keep the covariate that is most significant and so on for the rest covariates. The procedure continues until none of the covariates seem to be statistically significant. We are doing the same for the interactions of the covariates. Analytically, in the Appendix in Tables 8, 9, 10, 11, can be show the 540 logistic regressions which were executed for the Sequential Selection of the covariates and their interactions, as well as it can be shown in each round the most statistically significant covariate that was kept.

4. Results and Discussion

In this Chapter we report our results after applying the methods described above. Estimations by treatment and period, when the dependent variables is whether the individual has signed a contract after the training period in the Foundation, are obtained, as well as, estimations on the propensity score. In the section 4.3 we interpret the results of the ATE and ATT. We conclude by summarizing and discussing our results.

4.1 Estimations on the Propensity Score

In the following table (Table 5) we estimate the propensity score of our sample, based on the non-mandatory courses. The estimations are exported using the overall sample (2007-2013).

Table 5: Propensity Score

Variables	*Whole sample (2007-2013)*
Females	***0.008****** (0.004)
ISO courses	***0.080****** (0.013)
Number of courses	***0.012****** (0.001)
Low level of education	***-0.031****** (0.005)
Quality courses	***0.026****** (0.005)
Year of creating the record	***-0.016****** (0.002)
Year of starting the practice	***0.009****** (0.002)
High level of education	***0.016****** (0.005)
Respondent from Bizkaia	***0.016****** (0.005)
Pseudo-R^2	0.296
Number of observations	*3,186*

We can observe that all the estimations of the covariates are significant at the conventional significance level of 1%. This significance is obtained from the Sequential Selection of the Covariates we analyzed in Chapter 3.4.

When interpreting the estimations on the propensity score, we can see that females have 0.8% probability of receiving the treatment. ISO courses give 8% positive probability of treatment. Also, the higher the number of courses attended in the Foundation by the respondent, the higher the probability of receiving the treatment, more specifically the probability is 1.2%. Individuals with low level of education have smaller probability (-3.1%) of receiving the treatment than those who have different levels of education.

On the other hand for the respondents with high level of education the probability of receiving the non-mandatory courses reaches 1.6%. The same percentage is observed also for those who live in Bizkaia.

Young people registered in the Foundation and have attended quality courses, have 2.6% probability of receiving the non-mandatory courses provided. The later someone was registering in the Foundation, the lower the probability (-1.6%) of receiving the treatment. Finally, we can observe that the later an individual was starting the practice, the higher the propensity score (0.9%).

4.2 Estimations by treatment and period

Table 3 reports the marginal effects for the individuals in the control and treated group, taking as the dependent variable a dummy which shows whether the respondent has signed a contract or not:

Table 3: Dependent Variable: If the individual has signed a contract Estimated by Treatment[b]

Variables	*Control Group*	*Treated Group*
Females	*0.002* *(0.020)*	*-0.064* *(0.077)*
Year of creating the record	*-0.019* *(0.012)*	*-0.052* *(0.034)*
Year of completion of the practice	***-0.082****** *(0.012)*	***-0.073***** *(0.028)*
Low level of education	***-0.079 ****** *(0.025)*	*-0.040* *(0.217)*
Very low level of education	***-0.242****** *(0.064)*	--
Knowledge of Programming	***0.072****** *(0.021)*	*-0.041* *(0.076)*
Knowledge of Databases	***-0.046***** *(0.020)*	*-0.047* *(0.070)*
Car	***0.046***** *(0.021)*	*0.056* *(0.089)*
Number of contracts	*0.010* *(0.007)*	*0.025* *(0.028)*
Pseudo-R^2	*0.089*	*0.116*
Number of observations	*2,939*	*247*

[b] All estimations marked by (*) are significant to the conventional significance level of 10%, by (**) are significant to 5% and by (***) are significant to the conventional significance level of 1%.

We can observe that, for both control and treated group, gender is not significant. Nevertheless, a positive sign is reported for the individuals that have not participated to the training courses and, on the other hand, a negative sign for those who have participated. For the year that the record had been created in the foundation, we also obtain not significant estimations but in this case for both groups the probability is negative. The next variable of interest is the one that shows the year of the completion of the practice and both estimations are significant. This variable takes values from 2007 till 2014, which means that in our data exist some pending practices, for instance twenty-four of them are going to finish in 2014. As we can see in Table 3, the later the practice ends, the lower the probability is to sign a contract. This probability is even lower when the person belongs to the control group. From Chapter 2, we should be reminded of the four different categories that describe the variable we have used for education: very low, low, medium, high. After the sequential selection of variables, described in Chapter 3, very low and low level of education are two more variables that are significant, in our case, when estimating the parameters of signing a contract. For the group that is not participating on the training courses, those levels of education are statistically significant and have a negative impact on signing a contract. Moreover, we can see that an individual with very low level of education, that has not participated in the training courses, has even lower probability of signing a contract than an individual with a low level of education. We could claim that the results are as expected. For the treated group, very low level of education is omitted because the dummy variable is zero for all individuals. The variable which depicts the low level of education is not significant, for the group that received the treatment. The knowledge of programming seems to be significant for the control group but not for the treated one. Respondents with knowledge of programming have 7,2% higher probability to sign a contract than those that do not have these skills. Respectively, for the treated group we could say that the skill of programming or the skill of handling databases have a negative impact on signing a contract, but our estimation is not significant. Regarding the knowledge of handling the databases, we get some statistically significant results for the non-formation group. Surprisingly, the knowledge of handling databases, has a negative impact on the employment for the people registered in the Foundation. The estimations of the two last variables of interest can be interpreted in more than one way: we have estimated that an individual who has not taken part to the courses provided by the Foundation and owns a car, has 4.6% higher probability to sign a contract with some firm.

The dummy variable which depicts whether or not the respondent owns a car can have more than one interpretations: usually we consider that a car owner belongs to a gentry family, so to say in this case, individuals who belong to an upper class have 4,6% higher probability to find a job. Another meaning could be that the car owners are preferred to those that do not have a car because of the job requirements. For the treated group the variable "car" also has a positive sign but it is not significant. The last variable, which shows the number of contracts, can be also interpreted in different ways. More often this variable is inseparable with job experience. So to say, for both groups the probability of getting a job is higher when a contract is added into the curriculum. In our case this variable is not significant for none of the two groups.

Table 4 reports the marginal effects for the respondents in the control and treatment group, but this time we have divided the sample in two periods. The variable we are interested in remains the same, whether an individual has signed a contract or not:

Table 4: Dependent Variable: If the individual has signed a contract Estimated by Treatment and by Year

	Treated Group		***Control Group***	
Periods / ***Variables***	***2007-2009***	***2010-2013***	***2007-2009***	***2010-2013***
Females	*0.079* (0.105)	***-0.269***** (0.108)	*0.035* (0.028)	*-0.022* (0.027)
Year of creating the record	*-0.080* (0.080)	***-0.074***** (0.037)	***-0.101****** (0.028)	***-0.028**** (0.015)
Year of completion of the practice	*0.004* (0.069)	***-0.184****** (0.052)	*0.023* (0.023)	***-0.231****** (0.018)
Low level of education	*0.083* (0.271)	*-0.118* (0.231)	***-0.089****** (0.033)	*-0.054* (0.035)
Very low level of education	--	--	***-0.247****** (0.087)	*-0.091* (0.336)
Knowledge of Programming	*-0.080* (0.118)	*-0.076* (0.091)	***0.080****** (0.028)	***0.090****** (0.029)
Knowledge of Databases	*-0.015* (0.099)	*0.030* (0.0880	*-0.012* (0.028)	***-0.078****** (0.027)
Car	*-0.044* (0.110)	*0.156* (0.096)	***0.067***** (0.028)	*0.023* (0.030)
Number of contracts	*0.059* (0.041)	*-0.026* (0.036)	*0.012* (0.010)	*0.009* (0.010)
Pseudo-R^2	*0.043*	*0.178*	*0.032*	*0.150*
Number of observations	*111*	*136*	*1,313*	*1,626*

We can see that gender is significant only for the treated group that belongs to the second period. Females registered in the Foundation after 2010 have 26.9% less probability to have signed a contract. When treating the "year of creating the record" in *NSF* as an explanatory variable, we observe that in every period and group has a negative impact. Except from the first period which this variable is not significant, in all other three periods, the later someone was registering in the records of the Foundation, the smaller the probability to sign a contract. The change of the range of the probability is from -10.1 to -2.8 depending on the period and the group. From the next variable which depicts the year of completion of the practice we can see that for every year added, after 2010, the probability to find a job is negative both for the control and treated group. More specifically, the individuals that have participated in the training courses provided by the Foundation and belong to the second period (2010-2013), have 18.4% less probability to sign a contract as for the individuals that did not participated in the courses the probability is 3.7% more negative, so to say, 23.1%. For the first period, the particular variable is not significant. Very low education and low education are two of the variables that seem to affect only the control group for the period 2007-2009. In the treated group, as mentioned before, none of the respondents has recorded with very low level of education, that is why we do not get any estimations for this variable and is omitted from our regression. Low level of education affected negatively (-8.9%) individuals in the control group who were registered in the Foundation from 2007-2009, to find a job. As well as people that also did not participated in the non-mandatory courses of the Foundation, with a very low level of education for the second period, had 24.7% less probability to sign a contract with some firm after the training program. Knowledge of programming seems to be significant only for the group that did not received the treatment and affect them positively in both periods one and two, 8% and 9% respectively. On the other hand, the knowledge of handling databases seems to have a negative impact on every respondent, except the ones that belong to the treated group of the period two. This variable is significant at 1% significance level only for the individuals that have registered in the Foundation after 2010 and did not participated in the training courses. Specifically, the impact reduces the probability on finding a job by 7.8%. Individuals that have checked the bow of owning a car, have risen up the probability of finding a job by 6.7% (period 1, control group). As in the previous estimations (Table 3), the variable that depicts the number of contracts, is not significant.

4.3 Estimations on ATE and ATT

We will proceed by dividing, again, the sample in two different time periods like before: the first one will be based on the period 2007-2009 (period 1) and the second one will start from 2010 till the first months of 2013 (period 2). As we said before, after the matching method we end up with different number of observations. To remind the size of the sample, in period 1 the remaining observations are 456 and for the second period, 494. Computing the Average Treatment Effects and the Average Treatment Effects for the Treated we come up with the following numbers:

Table 6: Average Treatment Effect

Periods	*ATE*	*P-value*	*z-statistics*
period 1 (2007-2009)	***0.192***	***0.039***	***2.06***
period 2 (2010-2013)	*0.005*	*0.915*	*0.11*

Table 7: Average Treatment Effect for the Treated

Periods	*ATT*	*P-value*	*z-statistics*
period 1 (2007-2009)	***0.252***	***0.005***	***2.79***
period 2 (2010-2013)	***0.140***	***0.006***	***2.75***

Firstly, let us focus on the significance of the Effects which are shown in the previous tables (Table 6,7): the Average Treatment Effect for the Treated is significant for both periods but this is not the case for the Average Treatment Effect. The ATE estimation is statistically significant at the conventional level of 5% only for the first period. As in the second period (2010-2013), in the Foundation

For the first period, we can observe 19.2% higher probability to find a job, for those that are registered in the Foundation and a ,surprisingly, 25.2% higher probability to exit unemployment for the individuals that

there were introduced some mandatory training courses for those that had a scholarship training contract, the non-mandatory courses did not had a significant contribution on the employment of young people. On the other hand, only for those who participated in the training courses, the probabilities have risen 14% on finding a job after the treatment (significant at 1% significance level).

had taken part in the non-mandatory courses of *NSF*. For the second period, the probability of those who received the treatment, falls at 14% but remains significant and promising.

4.4 Summary and Conclusions

To sum up, the purpose of this paper was to investigate the effectiveness of the *Novia Salcedo Foundation*. More precisely, we were willing to analyze the contribution of the Foundation on the employment of young individuals in the Basque Country. As quoted in the beginning the percentage of young unemployed people in the Basque Country is much bigger than this in the rest of Spain, exceeds 40%. So our motivation was to find out if the particular non-profit organization, helps individuals to exit unemployment by providing some training courses. Unfortunately, our initial sample was reduced significantly after applying all the empirical methods needed. We were led to this reduction also because of the small amount of respondents that had taken this courses. We were able to use as treatment only the non-mandatory courses, otherwise, if all courses were included (mandatory and non-mandatory) our survey would not have sense for estimating the ATE and ATT because we would only have a treatment and not a control group.

In times of crisis, our estimations seem to be consistent as we showed that the later someone is finishing his training period, the more difficult is to find a job. We came up with some discouraging results as well. Such the one about the skill of knowing how to handle databases. Our estimations showed that this skill is working as a disadvantage for someone who had not received the treatment from the Foundation and is willing to exit unemployment. Regarding gender, the only significant estimation we observed, was the one from the treatment group of period 2 (2010-2013). It showed us that females have 26.9% less probability to sign a contract with some firm. The percentage is very high in comparison to other studies. High estimated values also found for the ATT of the second period. Based on our results, those who take the non-obligatory courses in *NSF*, have more than 25% probability to exit unemployment. So to say, one in four can find a job simply by registering in the Foundation and attending these courses. If we do not interpret this result with caution, we will better start writing another paper on how important is the Foundation on a decrease of the unemployment in the Basque Country. Also from the estimation of the ATE we should have big expectations, as almost 19% higher probability on finding a job, have all the young people registered in *NSF*. To close, our results concerning the effectiveness of the Foundation are reasonable and more than good for the community of young Basque unemployed people, nevertheless every single number in this research should be interpreted with caution and used under a frame of awareness.

5. Appendix

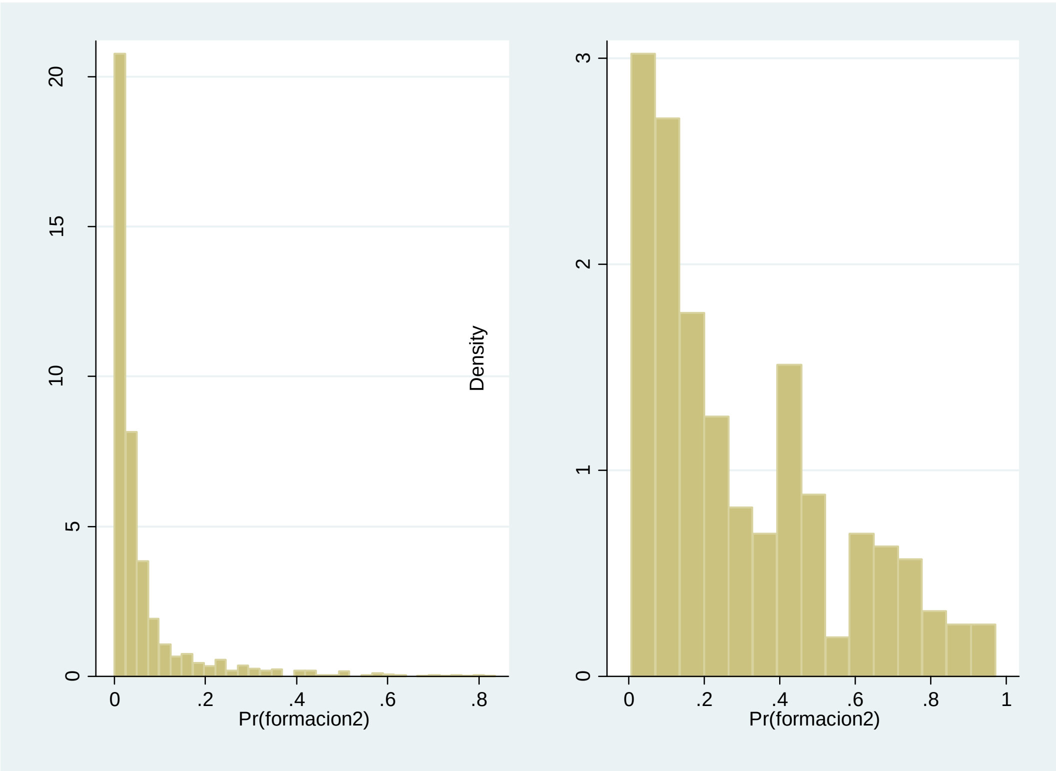

Figure 2: Histogram of the propensity scores for the control and treatment group (full sample)

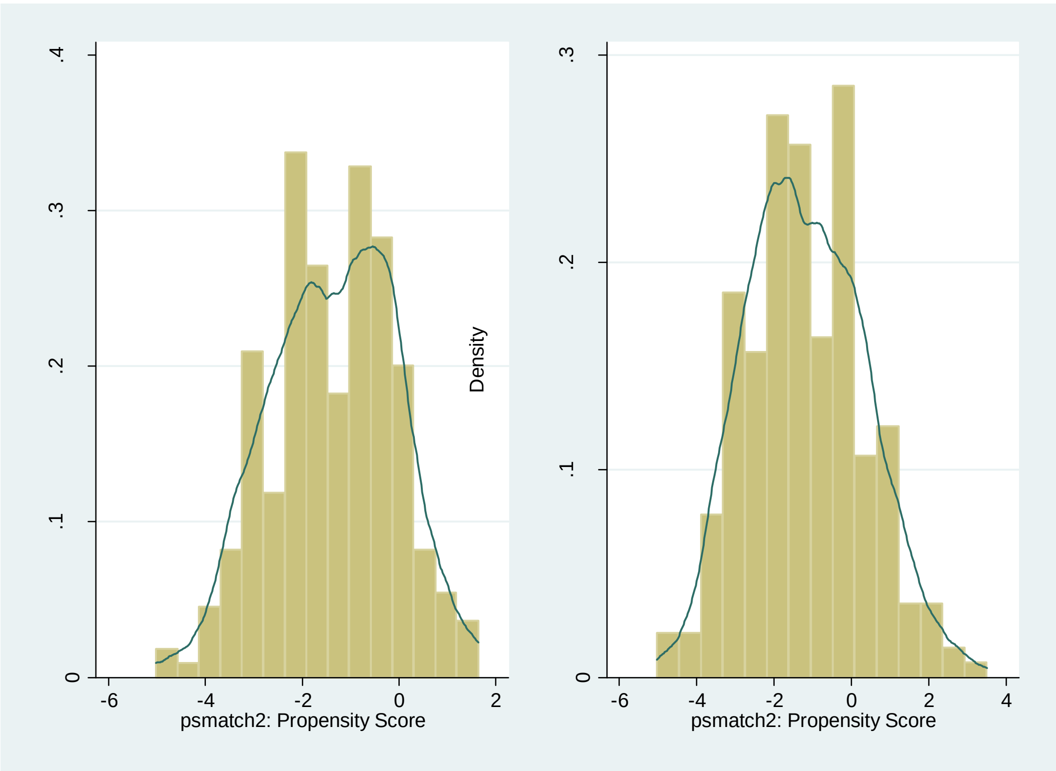

Figure 3: Histogram of the log odds ratio for the control and treatment group (matched sample)

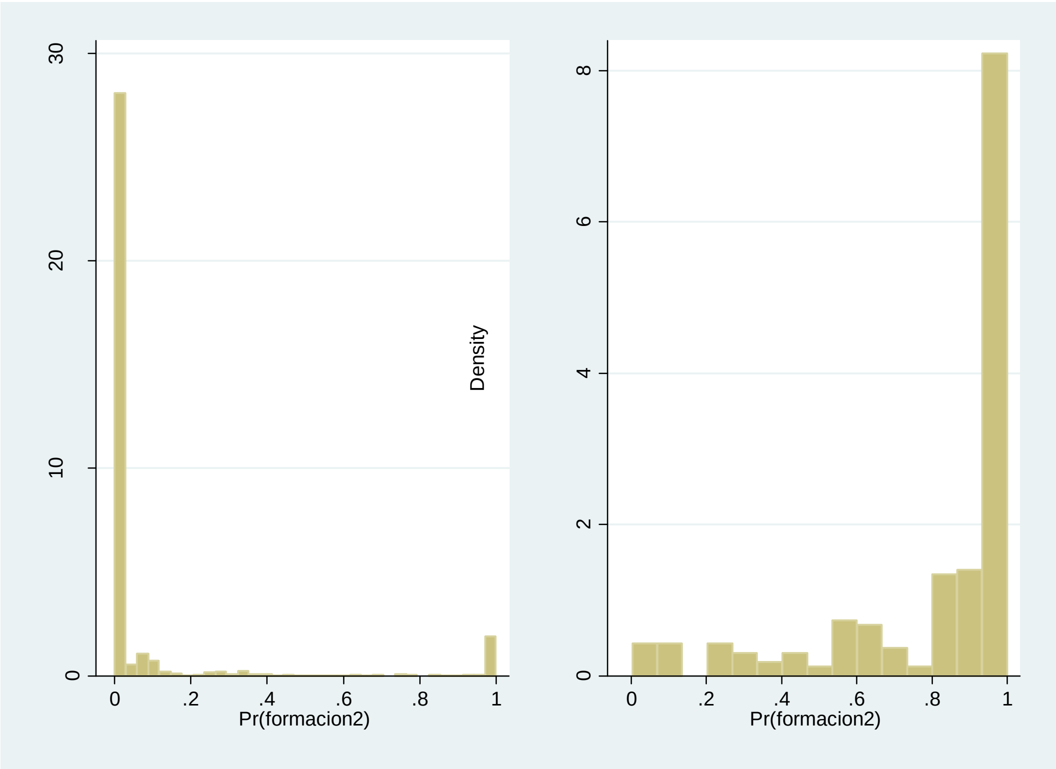

Figure 4: Histogram of the propensity scores for the control and treatment group (2007-2009 sample)

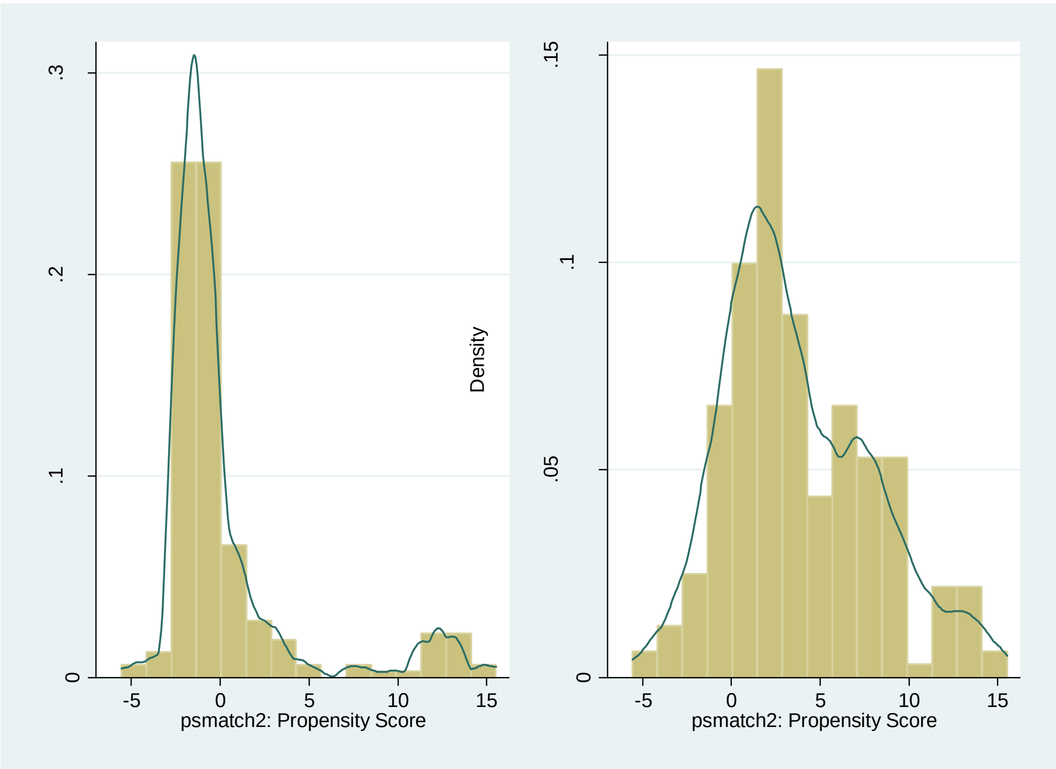

Figure 5: Histogram of the log odds ratio for the control and treatment group (2007-2009 matched sample)

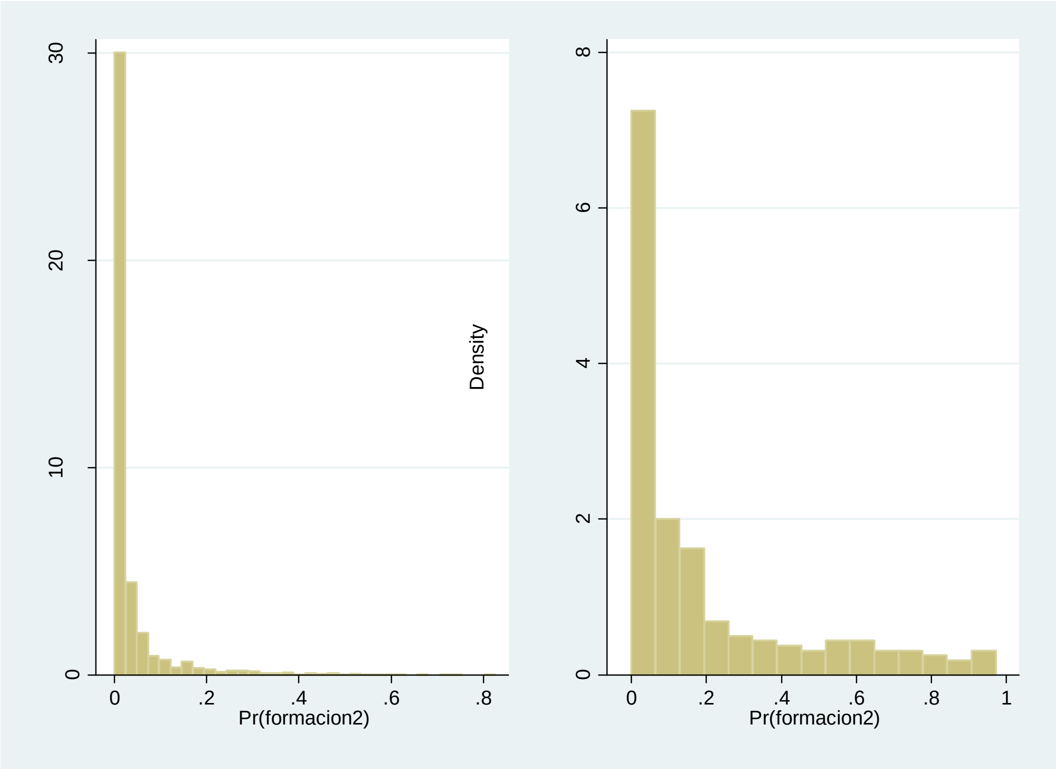

Figure 6: Histogram of the propensity scores for the control and treatment group (2010-2013 sample)

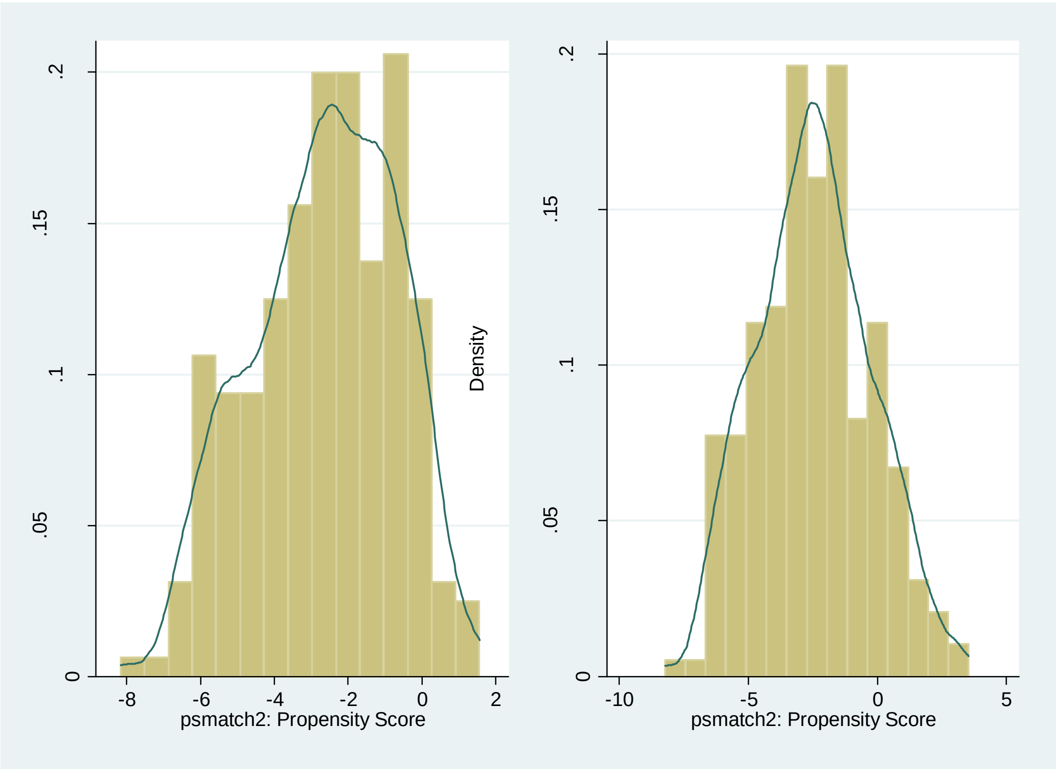

Figure 7: Histogram of the log odds ratio for the control and treatment group (2010-2013 matched sample)

Table 8: Likelihood Ratio and z-statistics for Sequential Selection of Covariates to enter the Propensity Score Estimation for Matching Procedure based on the non-mandatory formation

Variables	z-statistics	LR-statistics	z-statistics	LR-statistics	z-statistics	LR-statistics	z-statistics	LR-statistics	z-statistics	LR-statistics	z-statistics	LR-statistics	z-statistics	LR-statistics	z-statistics	LR-statistics	z-statistics	LR-statistic
mujer	→	→	→	→	→	→	→	→	→	→	→	→	→	→	→	→	→	→
edad	4.84	-847.3345	3.89	-772.1718	5.14	-693.2657	3.35	-668.5323	2.77	-650.3061	-0.76	-637.0047	0.37	-622.6663	0.29	-617.5835	0.39	-612.910
araba	-1.37	-857.8493	-2.06	-777.2989	-1.03	-705.9280	-0.92	-673.6405	-1.18	-653.3665	-1.28	-636.4284	-1.24	-621.9191	-1.18	-616.8894	0.71	-612.738
gipuzk	-4.84	-842.6464	-4.69	-764.8507	-3.00	-701.1264	-2.82	-669.4009	-2.55	-650.3702	-2.31	-634.2815	-2.03	-620.4456	-2.02	-615.3500	-0.13	-612.979
biscai	5.90	-837.0232	6.00	-757.4009	3.77	-698.5184	3.51	-667.2318	3.49	-647.3470	3.29	-631.3325	2.93	-618.0763	2.92	-612.9875	→	→
cursos	12.72	-787.9797	12.94	-706.4912	→	→	→	→	→	→	→	→	→	→	→	→	→	→
año creacion	-5.24	-844.4120	-4.50	-769.1421	-6.19	-685.7165	-6.00	-654.6570	-5.60	-637.2995	→	→	→	→	→	→	→	→
master	5.62	-843.9468	3.50	-773.7856	3.61	-700.1787	1.55	-672.8941	1.16	-653.4454	0.90	-636.8948	0.18	-622.7192	-0.01	-617.6254	0.12	-612.980
carnet	3.11	-853.1665	2.36	-776.5829	2.12	-704.0216	1.60	-672.7119	1.33	-653.1775	1.41	-636.2454	1.12	-622.0824	1.09	-617.0091	1.31	-612.086
coche	3.77	-851.0089	2.45	-776.5269	2.25	-703.8331	1.61	-672.7387	1.21	-653.3574	1.63	-635.9239	1.30	-621.8706	1.37	-616.6661	1.57	-611.712
efqm	7.91	-832.1470	-0.28	-779.6634	-0.17	-706.4760	0.01	-674.0821	-0.20	-654.0888	-0.44	-637.2018	-0.46	-622.6314	-0.58	-617.4577	-0.54	-612.840
medio ambiente	11.34	-802.3581	3.28	-774.3499	2.74	-702.7223	1.99	-672.0982	1.35	-653.1943	0.93	-636.8656	0.83	-622.3934	0.57	-617.4649	0.71	-612.735
experiencia	1.78	-857.3138	1.18	-779.0152	1.67	-705.1003	1.92	-672.2527	0.96	-653.6502	0.53	-637.1597	0.27	-622.6994	0.40	-617.5453	0.52	-612.853
prevencion	7.65	-831.6936	1.28	-778.8958	1.12	-705.8718	1.71	-672.6361	0.89	-653.7182	0.38	-637.2276	0.41	-622.6520	0.78	-617.3250	0.72	-612.732
iso	13.09	-779.704	→	→	→	→	→	→	→	→	→	→	→	→	→	→	→	→
cursos_calidad	9.31	-808.7999	7.24	-750.8957	7.03	-679.2937	6.08	-654.1092	→	→	→	→	→	→	→	→	→	→
bases_datos	-1.47	-857.7982	-1.71	-778.2405	-1.40	-705.5147	-0.56	-673.9238	-0.53	-653.9714	-0.64	-637.0971	-0.51	-622.6071	-0.57	-617.4629	-0.44	-612.891
diseño	1.11	-858.2542	0.29	-779.6607	0.65	-706.2793	0.16	-674.0686	-0.08	-654.1063	0.57	-637.1391	0.31	-622.6889	0.63	-617.4288	0.68	-612.752
programacion	-0.46	-858.7711	-1.04	-779.1541	-1.52	-705.3087	-2.18	-671.6517	-2.12	-651.8109	-1.74	-635.7556	-1.72	-621.2197	-1.39	-616.643	-1.41	-611.979
contratos	0.72	-858.6253	0.40	-779.6259	0.79	-706.1854	1.23	-673.3581	0.19	-654.0914	-0.54	-637.1498	-1.05	-622.1615	-0.90	-617.2102	-0.70	-612.739
beca_internacional	5.90	-845.0939	5.46	-767.3301	4.75	-696.8394	4.18	-666.44031	3.76	-647.7559	3.65	-631.3260	2.33	-620.18071	2.12	-615.5126	2.04	-611.008
añoinicio	1.77	-857.3080	1.64	-778.3480	-0.96	-706.0349	-1.47	-672.9983	-1.55	-652.9033	5.50	-622.7361	→	→	→	→	→	→
añofin	1.47	-857.7877	1.46	-778.6267	-0.87	-706.1112	-1.55	-672.8799	-1.64	-652.7723	4.98	-625.1691	0.22	-622.7122	0.20	-617.6058	0.04	-612.986
n_educ	(omitted)		→	→	→	→	→	→	→	→	→	→	→	→	→	→	→	→
l_educ	-5.52	-826.0856	-5.63	-746.1922	-5.64	-674.0822	→	→	→	→	→	→	→	→	→	→	→	→
m_educ	-1.71	-857.3464	-0.89	-779.2958	-1.43	-705.4339	-3.14	-668.8052	-3.18	-648.6821	-3.03	-632.3817	-2.78	-618.6429	0.03	-616.1924	0.02	-611.436
h_educ	7.04	-830.1810	6.48	-755.9379	6.68	-681.0300	3.54	-667.2905	3.52	-647.4259	3.40	-631.0861	3.10	-617.6254	→	→	→	→
h_engl	3.48	-852.9016	3.92	-772.0976	4.18	-697.8051	2.61	-670.6902	2.53	-650.9162	2.90	-633.1044	2.22	-620.2919	1.81	-615.9843	1.67	-611.590
h_eusk	0.56	-858.7201	0.11	-779.6978	0.75	-706.2106	0.69	-673.8416	0.63	-653.9082	1.10	-636.6985	0.97	-622.2600	1.09	-617.0298	1.46	-611.915
h_fran	-1.06	-858.248	-0.44	-779.6005	-0.34	-706.4320	-0.70	-673.8227	-0.52	-653.9665	-0.22	-637.2753	-0.33	-622.6789	-0.41	-617.5399	-0.52	-615.846
h_germ	0.54	-858.7449	1.17	-779.1286	1.19	-705.8893	0.92	-673.7082	1.18	-653.5124	1.14	-636.7445	1.20	-622.1207	1.26	-616.9490	1.29	-612.282

Table 9: Likelihood Ratio and z-statistics for Sequential Selection of Squared Covariates to enter the Propensity Score Estimation for Matching Procedure based on the non-mandatory formation.

Variables	z-statistics	Likelihood ratio statistics
edad^2	0.59	-610.8376
añofin^2	0.08	-611.0057
añoinicio^2	(omitted)	

Table 10: Likelihood Ratio and z-statistics for Sequential Selection of Covariates to enter the Bootstrap Estimation based on the non-mandatory formation.

Variables	z-statistics	LR-statistics	z-statistics	LR-statistics	z-statistics	LR-statistics	z-statistics	LR-statistics	z-statistics	LR-statistics	z-statistics	LR-statistics	z-statistics	LR-statistics	z-statistics	LR-statistics
mujer	→	→	→	→	→	→	→	→	→	→	→	→	→	→	→	→
edad	11.44	-2135.0033	2.32	-2051.9176	1.94	-2026.7411	0.95	-2022.6529	0.28	-2016.6874	0.18	-2012.9025	0.12	-2010.327		
araba	1.13	-2204.5056	0.90	-2054.2149	0.76	-2028.3352	0.79	-2022.7861	0.71	-2016.4713	0.62	-2012.7249	0.61	-2010.149		
gipuzk	-1.00	-2204.6517	0.77	-2054.3206	-0.16	-2028.6099	-0.09	-2023.0982	-0.20	-2016.7046	-0.16	-2012.9051	-0.21	-2010.3116		
biscay	0.47	-2205.0409	-1.12	-2053.9908	-0.07	-2028.6202	-0.14	-2023.0923	0.03	-2016.7249	0.04	-2012.9175	0.07	-2010.3314		
cursos	-3.20	-2199.9533	0.12	-2054.613	1.58	-2027.3574	1.59	-2021.8248	1.53	-2015.5502	1.45	-2011.8563	1.46	-2009.2631		
año_creacion	-16.61	-2054.6205	→	→	→	→	→	→	→	→	→	→	→	→	→	→
master	-0.87	-2204.7749	-1.48	-2053.5274	-0.23	-2028.5966	-1.07	-2022.5268	-1.36	-2015.7956	-1.05	-2012.3624	-0.95	-2009.8793		
carnet	0.82	-2204.8106	0.95	-2054.1651	1.30	-2027.7823	1.03	-2022.5714	0.71	-2016.4727	0.63	-2012.7212	0.61	-2010.1493		
coche	1.27	-2204.3425	2.39	-2051.7524	2.71	-2024.9321	2.49	-2020.0067	2.31	-2014.0548	2.22	-2010.4461	2.20	-2007.9164	→	→
efqm	1.38	-2204.1955	0.40	-2054.5408	0.97	-2028.1502	0.97	-2022.6314	0.95	-2016.2719	0.75	-2012.6338	0.81	-2010.004		
medio ambiente	0.50	-2205.0267	-1.15	-2053.9571	-0.34	-2028.566	-0.44	-2023.0056	-0.49	-2016.6065	-0.61	-2012.7319	-0.66	-2010.1159		
experiencia	1.93	-2203.2864	0.32	-2054.5679	0.55	-2028.4687	0.67	-2022.8776	0.65	-2016.5137	0.77	-2012.6207	0.95	-2009.8813		
prevencion	-0.01	-2205.1503	-1.17	-2053.9389	-0.75	-2028.3427	-0.50	-2022.9791	-0.36	-2016.659	-0.44	-2012.8224	-0.44	-2010.2371		
iso	-0.68	-2204.9196	-1.92	-2052.772	-1.04	-2028.0801	-0.99	-2022.6152	-1.00	-2016.2296	-1.11	-2012.3002	-1.07	-2009.7592		
cursos_calidad	0.20	-2205.1311	-1.08	-2054.0404	0.14	-2028.6123	-0.16	-2023.0897	-0.27	-2016.6897	-0.29	-2012.8751	-0.24	-2010.3043		
bases_datos	-0.50	-2205.0246	-1.83	-2052.9474	-2.17	-2026.2659	-1.82	-2021.4474	-1.77	-2015.1591	-2.27	-2010.3342	→	→	→	→
diseño	-0.49	-2205.03	0.42	-2054.5315	0.69	-2028.3858	0.48	-2022.9847	0.25	-2016.6952	-0.81	-2012.5929	-0.82	-2009.9946		
programacion	2.89	-2200.9629	3.18	-2049.5548	3.11	-2023.7556	2.90	-2018.8828	2.76	-2012.9183	→	→	→	→	→	→
contratos	2.68	-2201.5068	1.05	-2054.0643	1.36	-2027.691	1.53	-2021.923	1.54	-2015.5275	1.69	-2011.4822	1.82	-2008.6591	1.68	-2006.5021
beca_internacional	-2.26	-2202.4551	-2.64	-2050.9524	-1.39	-2027.6285	-1.45	-2022.0107	-1.46	-2015.6295	-1.40	-2011.905	-1.40	-2009.316		
añoinicio	-17.54	-2035.1935	-6.47	-2033.2956	-0.90	-2028.2158	-0.96	-2022.6424	-0.87	-2016.3469	-0.88	-2012.5312	-0.79	-2010.0226		
añofin	-17.67	-2030.9791	-7.14	-2028.6226	→	→	→	→	→	→	→	→	→	→	→	→
n_educ	-1.85	-2203.3982	-2.62	-2051.1124	-3.25	-2023.2303	-3.53	-2016.7253	→	→	→	→	→	→	→	→
l_educ	-1.61	-2203.8543	-1.83	-2052.9404	-3.32	-2023.102	→	→	→	→	→	→	→	→	→	→
m_educ	0.71	-2204.8973	1.07	-2054.0452	1.24	-2027.8504	0.28	-2023.0641	-0.10	-2016.7207	-0.37	-2012.8512	-0.39	-2010.2562		
h_educ	1.13	-2204.5109	1.18	-2053.9299	2.34	-2025.8837	0.77	-2022.8046	0.10	-2016.7207	0.37	-2012.8512	0.39	-2010.2562		
h_engl	-0.37	-2205.0834	0.84	-2054.264	1.66	-2027.2497	0.95	-2022.655	0.67	-2016.5018	0.58	-2012.7496	0.65	-2010.1238		
h_eusk	0.17	-2205.1357	1.67	-2053.2229	1.63	-2027.296	1.62	-2021.7908	1.54	-2015.5419	1.48	-2011.8151	1.40	-2009.3572		
h_fran	-0.25	-2205.1191	0.35	-2054.5603	0.35	-2028.5615	0.15	-2023.0911	0.04	-2016.7243	0.16	-2012.9062	0.18	-2010.3182		
h_germ	-0.43	-2205.0263	-0.36	-2054.5561	-0.47	-2028.5103	-0.60	-2022.9222	-0.66	-2016.5106	-0.66	-2012.6992	-0.58	-2010.1633		

Table 11: Likelihood Ratio and z-statistics for Sequential Selection of Squared Covariates to enter the Bootstrap Estimation based on the non-mandatory formation.

Variables	z-statistics	Likelihood ratio statistics
edad^2	-0.51	-2006.3712
añofin^2	(omitted)	
añoinicio^2	-1.08	-2005.9222

6. References

[1] Wooldridge, Jeffrey M. (2002). "Econometric Analysis of Cross Section and Panel Data"

[2] Wooldridge, Jeffrey M. (2010). "Econometric Analysis of Cross Section and Panel Data: Second Edition"

[3] Rubin, Donald B. (1974). Journal of Educational Psychology, Vol. 66, No. 5, 688-701. “Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies”

[4] Imbens, Guido W. & Rubin, Donald B. (2009) (forthcoming). "Causal Inference, Part II"

[5] Cameron, Colin A. & Trivedi, Pravin K. (2005). "Microeconometrics, Methods and Applications"

[6] Euskal Estatistika Erakundea, Instituto Vasco de Estadística, EUSTAT

[7] Keisuke Hirano, Guido W. Imbens, Geert Ridder (2002). "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score"

[8] M. T. Albanese, M. Knott (2000). "Bootstraping latent variable models for binary response"

[9] Paul R. Rosenbaum, Donald B. Rubin (1983). Biometrica, Vol. 70, No. 1, pp41-55. "The Central Role of the Propensity Score in Observational Studies for Causal Effects"

[10] Richard C. Chiburis, Jishnu Das, Michael Lokshin (2011). "A Practical Comparison of the Bivariate Probit and Linear IV Estimators"

[11] Alberto Abadie, Javier Gardeazabal (2003). "The Economic Costs of Conflict: A case Study of the Basque Country"

[12] Alberto Albadie, David Drukker, Jane Leber Herr, Guido W. Imbens (2004). "Implementing matching estimators for average treatment effects in Stata"