



Estudios de Economía Aplicada

ISSN: 1133-3197

secretaria.tecnica@revista-eea.net

Asociación Internacional de Economía
Aplicada
España

CANSINO MUÑOZ-REPISO, JOSÉ MANUEL; SÁNCHEZ BRAZA, ANTONIO
Effectiveness of Public Training Programs Reducing the Time Needed to Find a Job
Estudios de Economía Aplicada, vol. 29, núm. 1, 2011, pp. 1-26
Asociación Internacional de Economía Aplicada
Valladolid, España

Available in: <http://www.redalyc.org/articulo.oa?id=30120835015>

- How to cite
- Complete issue
- More information about this article
- Journal's homepage in redalyc.org

redalyc.org

Scientific Information System
Network of Scientific Journals from Latin America, the Caribbean, Spain and Portugal
Non-profit academic project, developed under the open access initiative

Effectiveness of Public Training Programs Reducing the Time Needed to Find a Job

JOSÉ MANUEL CANSINO MUÑOZ-REPISO

*Departamento de Análisis Económico y Economía Política. UNIVERSIDAD DE SEVILLA.
SPAIN. E-mail: jmcansino@us.es*

ANTONIO SÁNCHEZ BRAZA

*Departamento de Análisis Económico y Economía Política. UNIVERSIDAD DE SEVILLA.
SPAIN. E-mail: asb@u.es*

ABSTRACT

This paper estimates the average effect of a binary treatment on the time needed to find a job. Such a treatment is the training public schools program implemented in Seville. The research compares one training group and a control group. Two methods are developed with an interesting database. The first uses an estimator which weights observations by the inverse of the propensity score. This estimator let us conclude that, for participants, the time needed to find a job is reduced in 471 days. The second one is the differences estimator, it let us conclude that the time needed to find a job is reduced in 448 days. The evaluated program works as an active labour market policy with favourable effects on unemployed young people. Compared with other research in Spain, the obtained results show evidence in the same way as most of evaluation but stronger. A similar conclusion is derived from a comparison with international evidence disponible.

Keywords: Training programs evaluation, public policies, propensity score, differences estimator

Eficacia de los programas públicos de formación en la reducción del tiempo necesario para encontrar un empleo

RESUMEN

El artículo estima el efecto promedio de un tratamiento binario sobre el tiempo de búsqueda de empleo. Dicho tratamiento consiste en el Programa de Escuelas taller y Casas de Oficio desarrollado en la Provincia de Sevilla. Los autores comparan un grupo de tratamiento y otro de control. Para la estimación se utilizan dos métodos con los datos ofrecidos por una importante base de datos. El primero utiliza un estimador que pondera las observaciones por la inversa de la probabilidad de participar en el programa de formación ("propensity score"). Este estimador permite concluir que, para los individuos participantes en el programa, el tiempo de búsqueda de empleo se reduce en 471 días. El segundo método de estimación consiste en el estimador de diferencias. Con este segundo método el tiempo de búsqueda de empleo se reduciría en 448 días. El programa evaluado funciona como una política activa en el mercado de trabajo con efectos positivos sobre los jóvenes desempleados. En comparación con investigaciones similares realizadas en España, los resultados obtenidos arrojan evidencia en el mismo sentido aunque muestran un efecto mayor. Una conclusión similar se obtiene de comparaciones con la evidencia internacional disponible.

Palabras clave: Evaluación de programas de formación, políticas públicas, probabilidad de participar, estimador de diferencias

Clasificación JEL: H30, C10

Artículo recibido en diciembre de 2009 y aceptado en enero de 2011

Artículo disponible en versión electrónica en la página www.revista-eea.net, ref. e-29103

ISSN 1697-5731 (online) – ISSN 1133-3197 (print)

1. INTRODUCCIÓN ¹

From Heckman, Clements and Smith (1997), it is generally accepted that social programs impact differently on individuals when they differ in characteristics like the previous educational level, pre-training labour market status, sex, age, earnings, family background, etc. The individual's addequate assignement to the set of programs disposable begins a crucial issue in political decisions.

After controlling by covariates and by knowing the average effect of a program on an appropiatte outcome, the public decisor can decide which program would provoke the best impact on individuals by taking into account the program's average effect on subpopulations previously estimated. Social welfare can improve if public decisors follow an assignment rule which let them determine which individuals must receive which treatment. Manski (2001) theorized this assuming the case of a finite set of rival treatments. Cansino and Román (2007) explored this for the Spanish Accounting Court.

The aim of this paper is to estime the average effect of a binary treatment on an appropriate outcome. Such a treatment is the Training Schools Program (TSP) implemented in Seville between 1997 and 1999. Speciffically, the paper estimates the average effect of this training program on the time needed to find a job. To choice this outcome we followed Zweimüller and Winter-Ebmer (1996) who consider that similar evaluations in USA usually focused on impact on earnings as outcome while in Europe focused on outcomes like the one used in this paper².

We select the province of Seville as this is the zone in Spain with the most widely developed number of training policies until now. The evaluation is carried out by estimating the program's average effect over the individuals' ability of the sample to find a job, individuals being unemployed between 16 and 25 years old.

Following Hirano *et al.* (2003), the paper estimates the average effect by using an estimator which weights observations by the inverse of the estimation of the propensity score. The differences estimator is also implemented to compare results.

Evaluations of training are very scarce in Spain but largely developed in others countries like France, Germany, United Kingdom and USA. This may be due to the difficulty in collecting quality microdata. This paper contributes to

¹ Authors acknowledge comments and suggestions made by two anonymous reviewers although take full responsibility for the content. Authors are also grateful for the financial support received from the project SEJ132.

² A recent exemption is debt to Caparrós *et al.* (2010).

the literature in the sense that no evaluation based of non experimental methods has been applied to this training program in Spain before, apart from Cansino and Sánchez (2008 a y b), Mato and Cueto (2008) and Cueto and Mato (2009).

Paper structures as follow. The model and program characteristics are described in section 2. Section 3 is destined to explain the database used. In section 4, we define the estimator which weights observations by the inverse of the propensity -Hirano *et al.* (2003)- and results are obtained. The differences estimator is defined in section 5 and results are also obtained. Section 6 concludes.

2. CHARACTERISTICS OF THE PROGRAM EVALUATED

The TSP is one of the active labour market policies used in Spain. The program under evaluation is the TSP implemented in Seville between 1997 and 1999. The TSP was designed as a nationwide experimental one implemented by the Spanish Department of Labor (more specifically, by the National Institute of Employment). Considering the first results of the program, the Spanish Department of Labor decided to convert it into a permanent program regulated by the Department Labor's order of march 29st - 1988. Finally, the Department Labor's order of august 3st - 1994 added this program into the set of the national policies of employment until now.

TSP is a vocational training program which is managed in a decentralized way. It is offered free of charge and it involves voluntary access to training. The objective of the TSP is to act as an iniciative of young unemployed minors (less than 25 years old) in finding a job. The TSP organizes its activities into two steps; the first one gives unemployed a theoretical education and the second one offers a professional stage.

In order to judge the interest of this public training program, three parameters have been considered. The first is the number of participants. After the experimental period, the average number of participants had a range of between 45000 and 50000 young unemployed for every year. Compared to the whole of Spain, Seville is the zone with the largest number of implemented projects since 1985. This justifies the geographical focus of the paper and that implies a minor size of the sample.

Secondly, the size of the public funds absorbed as *Table 1* shows.

Thirdly, the EU authorities supported this training program allowing the use of the European Social Fund to finance it.

Table 1 shows the total budget for joint employment actions between the Spanish Department of Labor and the European Social Fund (ESF) for the period 2000-2006. Figures in *Table 1* show the importance that the ESF has as a co-financer of the training programs.

Table 1

Joint budget actions between the Spanish Department of Labour (SDL) and the European Social Fund (ESF) (2000-2006)	
TOTAL 9.540.537.406 €	
SDL 3.772.328.490 € 39'54 %	ESF 5.768.208.916 € 60'46 %

Source: INEM (National Institute of Employment)

3. THE BASEVAFOR DATABASE

The BASEVAFOR database has been constructed from individuals who have participated in the TSP carried out in Seville. Individual data came from the official employment agency.

We have selected those individuals who have finished the TSP in 1999, the last year in which information was available when we started the evaluation³. Depending on the length of the program (1 or 2 years), programs finishing in 1999 started in either 1997 or 1998. Only individuals who finished the TSP have been considered, rejecting those who left the program before the end. The total of individuals was 1528 and from that figure we have selected a sample of 150 (around 10%)⁴. The selected individuals make up the group of participants in our investigation.

To complete the database, the official employment agency provided us a control group⁵. To this end, 75 individuals⁶ with similar characteristics to the participants have been selected.

³ The requirements of the individuals were to be between the ages of 16 and 25 years, to remain unemployed and have signed up to the unemployment office.

⁴ The dimension of the selected sample is comparable to similar evaluations included in Dehejia and Wahba (1999). We assume some data are missing at random (Rubin, 1976; Little and Rubin, 1987).

⁵ For the purpose of this paper, the control group is an external one (Friedlander *et al.*, 1997). This type of group represents a sample of individuals who, during the period of time considered, show the same characteristics of participants and have the requirements to participate in the training program, but they have not participate in it. Ashenfelter (1978), Heckman *et al.* (1994) and Dehejia and Wahba (1999) justified the use of external control groups. Friedlander *et al.* (1997) criticize the use of internal groups in evaluations because it was quickly recognized that nonparticipants are likely to be quite different from participants by virtue of the fact that they have been excluded by program staff. In a similar way Bell *et al.*, (1995), have pointed out that

The database includes two types of data related not only to participants but also with the controls.

Firstly, this database gives us information about the periods of employment and unemployment of individuals, including data related to the number of times the individual has applied for a job. We will use this information to construct Y variable defined in next section. Secondly, this database contains information related to the covariates considered: sex, age and residence zone.

4. MODEL CHARACTERISTICS

4.1. The potencial outcome model

The development of public policies evaluation has benefited from the use of causal inference⁷. One of the results is the Potencial Outcome Model –POM-, which allows us to compare participants and non participants in public programs⁸. A prolific development of the POM with regard to training programs evaluation comes thanks to Roy (1951) and Rubin⁹ (1974, 1978). This paper support the Roy-Rubin Causal Model (RRM).

In the implementation of POM and RRM, the individual values of the main variables can be extracted from randomized experiments or from observational data. Both types of data will notably determine the evaluation and will promote different methodological developments.

In social sciences, randomized experiments face important problems related to cost, moral limitations, attrition and problems derived from the Hawthorne effect -Burtless (1995) and Cameron and Trivedi (2005)-. This can be solved by using observational data. In these cases, Rosenbaum (1999) says that the researcher should design a treatment group and a control group from the

‘secreed out’ applicants by definition differ from participants only on factor (both objective and subjective) observable to staff.

⁶ We choice the ratio 2/1 which was used in the evaluation of the National Job Training Partnership Act (JTPA) study. The JTPA Study was commissioned by the U.S. Department of Labour in 1986 to measure the benefits and cost of selected employment and training programs for economically disadvantaged adults and out-of-school youths. See Heckman, Ichimura and Todd (1997, 1998) and Heckman, Ichimura, Smith and Todd (1998).

⁷ Referring to the theoretical approach of causality and its use in randomized experiments, see Cox (1992). Other authors such as Dawid (1979, 2000), Holland (1986), Heckman (1990) and Pearl (2000) also discuss the meaning of causality in such an environment. Finally, in the specific case of training programs, we have referred to the seminal papers of Rubin (1974) and Heckman and Hotz (1989).

⁸ Cameron and Trivedi (2005) expose the POM advantages compared to alternative models.

⁹ The first references that Rubin considered were Neyman (1923, 1935) and Fisher (1928, 1935).

individuals who have or have not been treated. The objective is to reproduce a scenario which is as similar as possible to a randomized experiment¹⁰.

However, the comparison of the two outcomes for the same unit is not possible because they are counterfactual events (like an individual participating and not participating at the same time in a training program). We can at most observe one of these outcomes. Then, models which include counterfactual events are ineffective in individual causal effects estimation. Holland (1986) refers to this as “the fundamental problem of causal identification”. The construction of the counterfactual is one of the major problems in evaluation studies (Imbens and Wooldridge, 2009).

4.2. Definition of D and Y variables.

We define D as the binary variable which indicates the participation of the individuals in the program. So, $D_i=1$ will indicate that individual i has participated in the program and $D_i=0$ will indicate that individual i has not participated.

Otherwise, we consider Y as the response variable from which the program's average effects will be evaluated. We define Y as the ability of the individual i to find a job, and shows how much time he has to spend searching for a job¹¹.

$$Y_i = 1 - \frac{\text{number of consecutive days until the individual "i" find a job}}{\text{total duration of observatio n}} \quad (1)$$

The choice of the response variable is justified because individuals of the sample, both participants and non participants (control group), are initially unemployed and included in the official census of people who are searching for jobs. Most of them have not been working before or have a short labor experience because of their age and lack of experience.

As was mentioned above, the objective of the TSP is to act as an initiative focus on young unemployed minors (less than 25 years old). For this reason, it is relevant for the program evaluation to consider a response variable which allows us to measure the abilities of these people to find jobs.

¹⁰ The seminal papers in this topic were implemented in Medicine. The papers of Cameron and Pauling (1976), Billewicz (1965) and Cochran (1968), must be highlighted. An interesting comment about these seminal papers is contained in Rosenbaum (1995, 1996). Some well-known papers in observational methods are Kiefer (1979), Bassi (1984) and Moertel *et al.* (1985)

¹¹ For a further investigation, we could define two outcomes in an alternative way. The first one let us know the treatment average effect on the individual probability to find a job. The second would be the time needed to find a job conditioned to the unemployments subset (treated and controls) who have found a job.

The period of observation we have considered consists of three years¹² (1095 days). We started to measure this time from the moment participants finished the training program (generally at the end of 1999) and January 1st 2000 for individuals of the control group¹³.

The value of Y varies between 0 and 1. If Y is equal to 0, it means that individual i has not found a job during the period considered. This is the worse scenario for the program's effectiveness. If Y is close to 1, the individual i has found a job in a short period of time and if Y is equal to 1, it implies that individual i has found a job the first day after finishing the training program.

4.3. Identification and selection on observables

The fundamental problem of causal identification makes us look for second best solutions in which researchers leave the estimation of the individual causal effects.

The Average Treatment Effect of the program¹⁴ (ATE) is addressed in a partial equilibrium environment¹⁵ and, by using the potential outcome notation popularized by Rubin (1974), it is obtained as the average expected value from the difference between the potential values of Y_1 (the case of an individual treated) and Y_0 (the case of a non treated individual). Implicit in this notation is the stability assumption or SUTVA (Rubin, 1978) that individuals are not affected by receipt of treatment by others, and there is only one version of the treatment. As a consequence, no general equilibrium effects are considered (Cameron and Trivedi, 2005).

In the same way, the Average Treatment Effect of the program on the Treated (ATET) is defined as the average expected value from the difference between the potential values of Y_1 and Y_0 but only with respect to individuals who have received treatment.

Given the fact that the validity of average effects can be damaged if participants and controls show different characteristics apart from their participation or non participation in the training program, these characteristics must be controlled because of their effect on the values of the response variable.

¹² A relatively broad period of time has been considered, three years (1095 days), due to specific problems of this collective in finding jobs.

¹³ The date fixed to start the test for the control group coincides with the starting date for many of the individuals in the participants group.

¹⁴ Although in this paper only the most well-known average effects are used, Imbens (2004) summarizes all the possible types of average effects of treatment in literature.

¹⁵ The ATE is addressed in a partial equilibrium environment different, for example, to the CGE model mentioned in Cansino, Cardenete y Román. (2007).

If the observed characteristics are the only individual (participants and controls) characteristics that differ, we can therefore control these differences. This is the base of the selection on observable model¹⁶. Selection on observables allows us to isolate the effect of a covariate (or a vector of covariates)¹⁷ maintaining the independence between the treatment indicator variable D and the response variable, Y . This condition can be expressed as

$$(Y_1, Y_0) \perp D | X \quad (2)$$

Selection on observables supports the independence assumption¹⁸ typical in randomized experiments, contributing to the comparison between participants and controls. Following Heckman and Hotz (1989), selection on observables is recommended when the independence between D and Y is because of the covariate X (or vector of covariates), which has influence on the individual selection process, so by controlling X we give a solution to possible biased selection, making the dependency between D and Y disappear.

In the selection on observable context, when the independence assumption is guaranteed, we considered, according to Dehejia and Wahba (1999), that

$$E[Y_1 - Y_0 | X] = E[Y | X, D = 1] - E[Y | X, D = 0] \quad (3)$$

Equation 3 lets us express the ATE as

$$\begin{aligned} ATE &= E[Y_1 - Y_0] = \int (E[Y_1 - Y_0 | X]) dP(X) = \\ &= \int (E[Y | X, D = 1] - E[Y | X, D = 0]) dP(X) \end{aligned} \quad (4)$$

Then, it is possible to determine ATE from the difference between the average observed value of the response variable of the participants and the controls by calculating the difference for every possible value of X .

In a similar way, it is possible to calculate the average effect of the training program only for participants ($ATET$) as the difference between the average

¹⁶ To improve knowledge of selection on observables we recommend Barnow *et al.* (1980). When controlled and treated differ in unobserved characteristics like psychological ones, average effect can be estimated by the differences in differences estimator. For the context of this paper, see Cansino and Sánchez (2008a). Another alternative, when independence condition is not guaranteed, is to use the methods based on instrumental variables if an adequate instrument to determinate D is available. Instrument must be uncorrelated with Y .

¹⁷ As an introduction to the framework of the observational methods we recommended the examples that are used by Rosenbaum (1995) in his exposition about these methods; this also can be said of Cochran (1968) and Cameron and Pauling (1976). We also recommend the papers of Billewicz (1965) and Moertel *et al.* (1985), both of them referring to the two previous examples.

¹⁸ This is also known as the unconfoundedness assumption (Rubin, 1978; Rosenbaum and Rubin, 1983). See also Barnow *et al.* (1980).

observed values in the response variable of participants and the average values of controls for every different value of X when $D=1$. That is what makes (5)

$$ATET = E[Y_1 - Y_0 | D=1] = \int (E[Y|X, D=1] - E[Y|X, D=0])dP(X|D=1) \quad (5)$$

4.4. Definition of the vector X of covariates

With D defined, X will be a covariate¹⁹ with respect to D if, for each of the individuals observed, its values remain the same for each value of D . That is to say $X_{0i}=X_{1i}$, being X_{0i} the X value before the event D ($D=0$) and X_{1i} the X value after happening D ($D=1$).

X covariate is also named contaminant because of the fact that X can contaminate Y by adding its own effects²⁰ to those provoked by D .

The fact that X is predetermined with respect to D does not imply that this independence is bidirectional because it is possible that, as a characteristic of considered population, dependence in an opposite direction can appear making the value of D be affected for X .

From the sample information included in the database, we consider three predetermined characteristics to form the vector of covariates X . The database only allows us to include in the model a complete information about this three characteristics. *Table 2* describes X .

Table 2

Definition of Covariates of Vector X			
Characteristics	Description	Covariates	Potential values
1) Sex	This binary covariate shows if the individual considered is male or female. Base category = female.	X_{11}	1 = male 0 = female
2) Age	This covariate shows the individual's age at the beginning of the observational period. In the case of participants, shows the individual's age when the training program is over. For controls, shows the individual's age as of January 1 st, 2000. Considering that age range for participants in one of the considered training programs is between 16 and 24 years old, and also considering that the program may extend for 1 or 2 years, this covariate will have values of between 17 and 26 years old.	X_2	$17 \leq X_2 \leq 26$

¹⁹ We talk about one covariate but everything we state can be extrapolated for the case that X is a vector of n covariates, as $X^n = (X_1, X_2, \dots, X_n)$.

²⁰ To read more, comments of Rubin (1978) about covariates are very interesting.

Characteristics	Description	Cova riates	Potential values
3) Zone	This characteristic is collected through three binary covariates that show the city where individuals took the training program or, in the case of controls, where individuals lived. The area of Seville has been divided into 4 zones, being the criterion of mapping an operational one: zone 1(Sevilla city), zone 2(east and northeast of Seville), zone 3(south and southwest) and zone 4(west and northwest). Base category = zone 4.	X_{31}	1 = zone 1 0 = otherwise.
		X_{32}	1 = zone 2 0 = otherwise
		X_{33}	1 = zone3 0 = otherwise

Source: Own elaboration

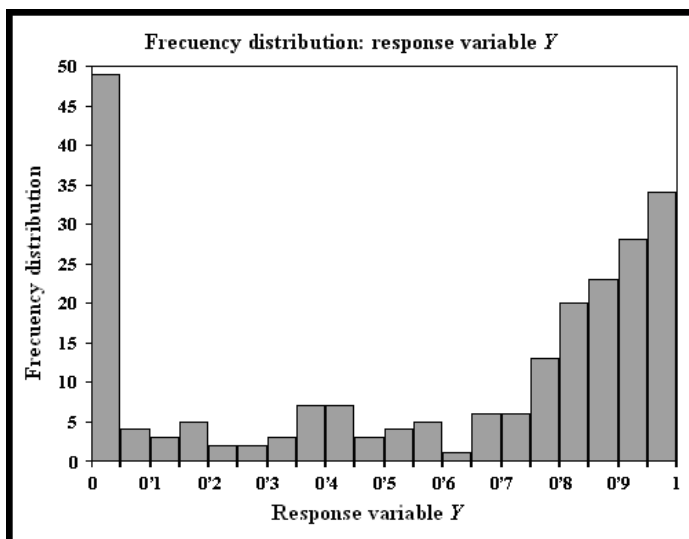
Finally, *Table 3* summarizes the main descriptive data of the response variable Y and X for the total sample and for each of groups (participants and controls). *Figure 1* and *Figure 2* include their frequency distribution.

Table 3

Descriptive statistics of Y , X_{11} , X_2 , X_{31} , X_{32} , X_{33}						
	Mean	Max	Min	Standard deviation	Kurtosis	Asymmetry coefficient
TOTAL						
Y	0'5781471	1	0	0'3806571	1'6121740	-0'5299899
X_{11}	0'5111111	1	0	0'5009911	1'0019760	-0'0444554
	Mean	Max	Min	Standard deviation	Kurtosis	Asymmetry coefficient
X_2	20'4755600	26	17	2'1087780	2'0620320	0'4149393
X_{31}	0'2088889	1	0	0'4074212	3'0512790	1'4322290
X_{32}	0'2977778	1	0	0'4583009	1'7822600	0'8844544
X_{33}	0'2844444	1	0	0'4521553	1'9131410	0'9555839
GROUP OF PARTICIPANTS						
Y	0'7287001	1	0	0'3005494	3'7215220	-1'4221670
X_{11}	0'5800000	1	0	0'4952120	1'1050900	-0'3241764
X_2	20'7266700	26	17	2'2551850	1'8819180	0'2685756
X_{31}	0'2066667	1	0	0'4062708	3'0992140	1'4488660
X_{32}	0'3000000	1	0	0'4597928	1'7619050	0'8728716
X_{33}	0'2933333	1	0	0'4568152	1'8241850	0'9078465
CONTROL GROUP						
Y	0'2770411	1	0	0'3448190	2'2610600	0'8643034
X_{11}	0'3733333	1	0	0'4869467	1'2743160	0'5237520
X_2	19'9733300	26	17	1'6843740	1'9579370	0'4691556
X_{31}	0'2133333	1	0	0'4124198	2'9586860	1'3995310
X_{32}	0'2933333	1	0	0'4583559	1'8241850	0'9078465
X_{33}	0'2666667	1	0	0'4451946	2'1136360	1'0552900

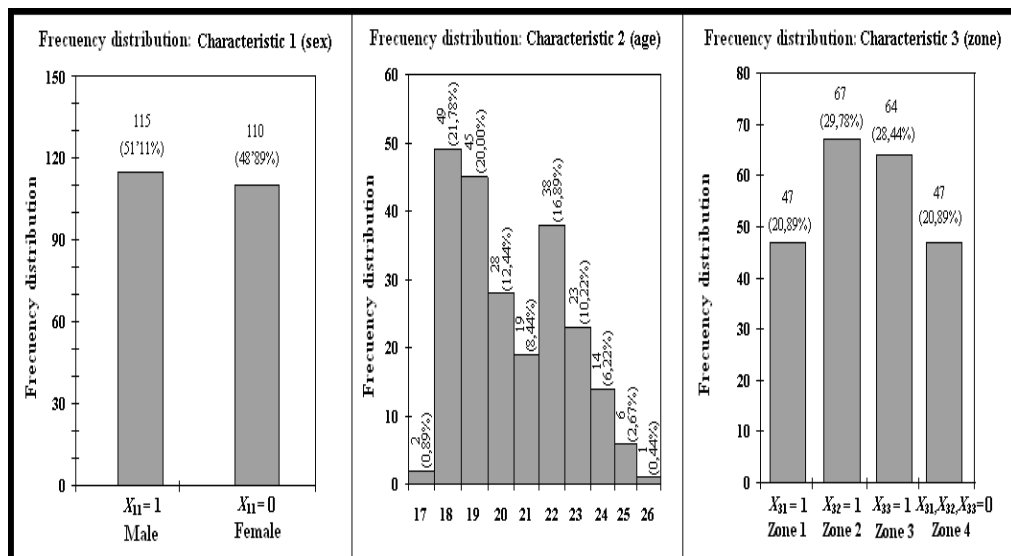
Source: Own elaboration

Figure 1



Source: Own elaboration

Figure 2



Source: Own elaboration

5. ESTIMATION OF THE AVERAGE EFFECT BY USING THE PROPENSITY SCORE

5.1. The propensity score

To avoid the need to match individuals on the values of all covarities, Rosenbaum and Rubin (1983, 1984) developed an approach based on the propensity score, that means the probability of one individual to participate in a program (probability of $D=1$), conditioned to the values of vector X . By making this probability $\varepsilon(X)$, we can express this as:

$$\varepsilon(X) = P(D = 1 | X) \quad (6)$$

which is assumed to be bounded away from zero and one.

This shows that propensity score ε is a function of X , which is usually unknown, and therefore it should be estimated from database.

Rosenbaum and Rubin (1983) proposed to condition on the propensity score. They prove that if $(Y_1, Y_0) \perp D | X$ and $0 < \varepsilon(X) < 1$, then

$$(Y_1, Y_0) \perp D | \varepsilon(X) \quad (7)$$

So the outcome is the same for participants and controls conditioning on the X variables or on the propensity score, $\varepsilon(X)$.

In this way, the independence assumption typical in randomized experiments is guaranteed. This assumption lets us argue that all observations with the same propensity score will have the same distribution of vector X , which means that we can compare the data observed for either participants or controls with the same propensity score.

Following Hahn (1998), the calculation of the conditioned probability of participation in a program, given certain observable characteristics, plays a crucial role in controlling bias in order to obtain an estimator of the program's effects. By using propensity score²¹, we proceed as if it were the case of an unidimensional variable improving evaluation efficiency by avoiding the management of a large number of covariates included in vector X .

To estimate the effect of a training program by using the propensity score we proceed in a two-step way. First, we estimate the propensity score on the vector X . Secondly, we obtain the estimator of the average effect of the training program.

²¹ Really, there is not consensus on the number of covariates which recommended the use of propensity score instead of the covariates vector (Imbens, 2004). In any case, this is recommended when the overlap assumption can not be guaranteed for all the covariates.

5.2. The value of the propensity score

From (6) we can now express the probability of an individual's participation in a program conditioned on the value of vector X , as:

$$\varepsilon(X) = P(D = 1 | X) = F(\beta X) \quad (8)$$

where β is the parameter's vector associated with covariates. The value of this probability will remain conditioned to the value of the distribution function at point βX_j ; X_j being every possible value that can adopt the vector of covariates X , with $j = 1, \dots, k$.

Depending on the specific function F , different selection models of binary response could be specified. We have choiced three of them: the Probit Model, the Logit Model and the Extreme Value Model Type I. There is not a generally accepted selection criterion in choosing one of these three models for the estimation of the propensity score, so the way the choice is made is due only to practical reasons. We estimate the three models and after analysing results, we will choose the best one. The most efficient method will be the one that shows less values of information criterion of Akaike, Schwarz and Hannan-Quinn and a higher value of the log likelihood function. This information is contained in *Table 4*. According to these criteria the Probit Model has been selected.

Table 4

Comparison of the obtained results from the three binary response models applied			
	Probit Model	Logit Model	Extreme Value Model
Log. likelihood function	-134'5663	-134'6875	-134'8964
Criterion Akaike	1'249478	1'250555	1'252412
Criterion Schwarz	1'340574	1'341651	1'343508
Criterion Hannan-Quinn	1'286245	1'287322	1'289179

Source: Own elaboration

Resulting values are shown in *Table 5* and indicate the degree every of the considered covariates contribute to the propensity score. As was explained above, the purpose of the propensity score is to make individuals from the treatment group and the control group as homogeneous as possible as far as the all of the covariates are concerned. In addition, *Table 6* contains the main data of descriptive statistics related to the probability of participation estimated.

Table 5

The calculation of the propensity score by using the Probit Model					
Dependent Variable: D (Prob. $D = 1$)					
Variable	Coefficient	Coef. Value	Std. Error	Z - statistic	Prob.
Fixed effect	μ	-2'366701 ***	0'848762	-2'788414	0'0053
X_{11}	β_{11}	0'580077 ***	0'185036	3'134942	0'0017
X_2	β_2	0'125535***	0'040936	3'066652	0'0022
X_{31}	β_{31}	-0'254362	0'285963	-0'889491	0'3737
X_{32}	β_{32}	0'002826	0'254795	0'011091	0'9912
X_{33}	β_{33}	0'058707	0'257340	0'228131	0'8195
Mean dependent var		0'666667	S.D. dependent var		0'472456
S.E. of regression		0'460830	Log likelihood		-134'5663
Sum squared resid		46'50776	Restr. log likelihood		-143'2157
LR statistic (5 df)		17'29884	Avg. log likelihood		-0'598072
Probability (LR stat)		0'003967	McFadden R-squared		0'060394
Note: Std. Errors adjusted by White's method. One, two or three asterisks indicate coefficient significance at the 10-percent, 5-percent and 1-percent levels, respectively.					

Source: Own elaboration

Table 6

Descriptive statistics related to propensity score obtained by using the Probit Model					
Mean	Max	Min	Standard deviation	Kurtosis	Asymmetry coefficient
0'6671581	0'920798	0'358889	0'3806571	-0'795769	-0'145460

Source: Own elaboration

Finally, the estimated value of propensity score must be assigned to every individual (participants and controls). Table 7 shows all results. After doing that, we calculate ATE ($\hat{\alpha}_{ATE}$) and $ATET$ ($\hat{\alpha}_{ATET}$).

Table 7

Propensity score assigned depending on possible values of vector X , according to the Probit Model								
	Male				Female			
	Zone 1	Zone 2	Zone 3	Zone 4	Zone 1	Zone 2	Zone 3	Zone 4
17 years old	0'537092	0'636943	0'657695	0'635882	0'313141	0'409132	0'430973	0'408034
18 years old	0'586537	0'682903	0'702538	0'681896	0'358889	0'458488	0'480714	0'457367
19 years old	0'634645	0'726203	0'74449	0'725261	0'406757	0'508494	0'530757	0'507366
20 years old	0'680721	0'766358	0'783123	0'765491	0'456062	0'558366	0'580319	0'55725
21 years old	0'724162	0'803015	0'818146	0'80223	0'506053	0'607328	0'628636	0'606241
22 years old	0'76448	0'835956	0'849398	0'835257	0'55595	0'654647	0'675006	0'653605
23 years old	0'801313	0'865097	0'876852	0'864483	0'604975	0'699663	0'718811	0'69868
24 years old	0'834439	0'890472	0'900591	0'889942	0'65239	0'74182	0'759548	0'740906
25 years old	0'863765	0'912224	0'920798	0'911773	0'697532	0'780683	0'796839	0'779847
26 years old	0'889321	0'930579	0'93773	0'930201	0'739839	0'81595	0'830443	0'815197

Source: Own elaboration

5.3. Weighting observations by the inverse of the propensity score

By weighting observations of the response variable Y by the inverse of the estimation of the propensity score, we can obtain efficient estimators of the average effect. Estimators of ATE and $ATET$ are expressed as follows²²:

$$\hat{\alpha}_{ATE \rightarrow \text{WEIGHTING PS}} = \frac{1}{n} \sum_{i=1}^n Y_i \frac{D_i - \hat{\varepsilon}(X_i)}{\hat{\varepsilon}(X_i) (1 - \hat{\varepsilon}(X_i))} \quad (9)$$

$$\hat{\alpha}_{ATET \rightarrow \text{WEIGHTING PS}} = \frac{1}{n_1} \sum_{i=1}^n Y_i \frac{D_i - \hat{\varepsilon}(X_i)}{1 - \hat{\varepsilon}(X_i)} \quad (10)$$

²² Hirano *et al.* (2003) developed ATE and $ATET$ estimators obtaining:

$$\hat{ATE} = \frac{1}{n} \sum_{i=1}^n \left(\frac{Y_i D_i}{\hat{\varepsilon}(X_i)} - \frac{Y_i (1-D_i)}{1-\hat{\varepsilon}(X_i)} \right) \quad 0 = \sum_{i=1}^n \hat{\varepsilon}(X_i) \left(\frac{Y_i D_i}{\hat{\varepsilon}(X_i)} - \frac{Y_i (1-D_i)}{1-\hat{\varepsilon}(X_i)} - \hat{ATE} \right)$$

where Y is the outcome, D the binary variable which indicates if individual it's treated or control and the vector X of covariate's vector which let us define $\hat{\varepsilon}(X_i)$ as the probability to participate in the program, conditioned on X .

where $\hat{\varepsilon}(X_i)$ is the estimated value of the propensity score for the i -individual on vector X , n is the sample size and n_1 the number of participants. We can apply bootstrapping techniques to obtain estimates for the standard errors. The obtained results are shown in *Table 8*.

Table 8

The calculation of $\hat{\alpha}_{ATE}$ and $\hat{\alpha}_{ATET}$ estimators by using the "propensity score weighting"				
Estimator (Coefficient)	Coef. Value	Std. Error	t -statistic	Prob.
$\hat{\alpha}_{ATE}$	0'421280 ***	0'060	6'854666	0'0000
$\hat{\alpha}_{ATET}$	0'430409 ***	0'056	7'685875	0'0000
Note: One, two or three asterisks indicate coefficient significance at the 10-percent, 5-percent and 1-percent levels, respectively.				

Source: Own elaboration

The estimated value of the *ATE* is positive. In average, the sample's individuals' ability to find a job increases by 0'421280. In the case of the *ATET*, the estimated value is also positive, meaning that there is a favourable causal effect from the program. This result indicates that participant's ability to find a job has increased, on average, by 0'430409 so the TSP reduces the time needed to find a job and no looking-in effect of trainees is shown.

We can address results in the context of similar evaluations. Mato (2010) and Ramos *et al.* (2010) give an overview of the evaluations of training programs carried out in Spain recently. Compared with other research in Spain, our results show evidence in the same way as most of evaluation but stronger. A similar conclusion is derived from a comparison with international evidence summary in Kluve *et al.* (2007) and Card *et al.* (2009).

6. ESTIMATION OF THE AVERAGE EFFECT BY REGRESSION

It is possible to determine the average effect of a training program on participants (*ATET*) by regression by using Least Squares, given that independence assumption is also guaranteed. According to Stock and Watson (2003), we can obtain an estimator of the average effect of the evaluated training program on participants ($\hat{\alpha}_{ATET}$) using a linear model.

It is also possible to introduce covariates in the model as additional regressors. By doing that, we can measure the effects that covariates have on Y . The inclusion of covariates in the model is shown by the following expression (Wooldridge, 2002):

$$Y = \mu + \alpha D + \beta X + v \quad (11)$$

where Y is the dependent variable and shows the potential results of the individual. D is a binary explanatory variable. X is the covariates' vector and β is the parameter associated of the vector X . Parameter μ collects fixed effects in the model and the parameter v collects the random error of the model, with an average value equal to 0, $E[\varepsilon | D, X] = 0$.

The parameter α will determine the average effect of the program on participants. This parameter is named the "differences estimator with additional regressors". As independence assumption is guaranteed, if necessary assumptions²³ for multiple regression by Least Squares are guaranteed too, this estimator will be unbiased and consistent.

The inclusion of additional regressors in the model lets us improve the estimator's efficiency by reducing the random error variance. On the other hand, the addition allows us to test the randomness in the individual assigning procedure between the participants group and the control group, in the case that the assigning procedure is related to the additional covariates. That is to say, by including the covariates in the model we can control the probability of individuals being assigned between the participants group and the control group by adding characteristics in which participants and controls differ.

The inclusion of predetermined variables in the model will consist of inserting covariables included in the vector X as additional regressors. *Table 9* shows the correlation matrix between explanatory variables included in the model, allowing us to analyze possible multicollineality.

²³ These are the four assumptions: a) the conditioned distribution of the random error, given the explanatory variates $X_{1i}, X_{2i}, \dots, X_{ki}$, is equal to 0 on average (in this case, the explanatory variables are D_i , that indicates program participation, and the covariates). b) all the observations $i = 1, 2, \dots, n$ are distributed both independently and identically random. C) $X_{1i}, X_{2i}, \dots, X_{ki}$ and ε_i have four moments. D) Perfect multicollineality does not exist.

Table 9

Matrix correlation between D and the covariables of vector X						
	D	X_{11}	X_2	X_{31}	X_{32}	X_{33}
D	1'000000	0'194895	0'168779	-0'007731	0'006873	0'027864
X_{11}	0'194895	1'000000	0'005540	0'218229	0'014691	-0'092845
X_2	0'168779	0'005540	1'000000	0'143666	-0'105606	0'030735
X_{31}	-0'007731	0'218229	0'143666	1'000000	-0'334617	-0'323978
X_{32}	0'006873	0'014691	-0'105606	-0'334617	1'000000	-0'410569
X_{33}	0'027864	-0'092845	0'030735	-0'323978	-0'410569	1'000000

Source: Own elaboration

Linear correlation between variables is not obvious because all the coefficients are very low and far from ± 1 . With regard to the relationship between X_{31} , X_{32} and X_{33} , these variables show slightly high correlation because of the fact that they have been constructed to include the zone variable in the model. However, values are never over ± 0.5 . On the other hand, the value of the determinant of the correlation matrix is 0.4464, far from 0. In addition, the condition number of the correlation matrix is $C=2.3240$, far from the limits that determine multicollineality. Everything we have exposed lets us indicate that multicollineality problems are not relevant in the model.

With the previous specifications made, we implement the regression by Least Squares. Results are shown in Table 10.

Table 10

The calculation of the "differences estimator with additional regressors"					
Dependent Variable: Y (Ability to find a job)					
Variable	Coefficient	Coef. Value	Std. Error	t -statistic	Prob.
Fixed effect	μ	-0'206095	0'216058	-0'953891	0'3412
D	α	0'409173 ***	0'050124	8'163276	0'0000
X_{11}	β_{11}	0'128764 ***	0'045467	2'832055	0'0051
X_2	β_2	0'023647***	0'010274	2'301544	0'0223
X_{31}	β_{31}	-0'040937	0'065065	-0'629174	0'5299
X_{32}	β_{32}	-0'028182	0'059367	-0'474715	0'6355
X_{33}	β_{33}	-0'075899	0'058980	-1'286872	0'1995
R-squared		0'360720	Ajusted R-squared		0'343125

Mean dependent var	0'578147	S.D. dependent var	0'380657
S.E. of regression	0'308514	F-statistic	-143'2157
Sum squared resid	20'74949	Prob (F-statistic)	-0'598072
Log likelihood	-51'10855		
<i>Note: Std. Errors adjusted by White's method. One, two or three asterisks indicate coefficient significance at the 10-percent, 5-percent and 1-percent levels, respectively.</i>			

Source: Own elaboration

The α parameter (0'409173) is the differences estimator with additional regressors from the program's effect on participants ($\hat{\alpha}_{ATET}$). The individual significance of the explanatory variables included in the model is demonstrated by the values obtained for the t-statistic and its associated probability. From these values D , X_{11} (sex) and X_2 (age) appear to be significant variables. The X_{31} , X_{32} and X_{33} , defined to include the zone covariate, appear insignificant. The result of the log. likelihood ratio test to assess whether they are jointly significant is also negative (Prob. log. likelihood ratio: 0'617698). In any case, we have decided to maintain them because they help to improve the significance of all the estimated parameters and goodness of fit. The adjustment shows non significance of the fixed effect.

With respect to goodness of fit, the R-squared statistic equals 0'3607 and shows that the explanatory power of the considered variables is equal to 36'07 percent, significantly improving the accuracy of the adjustment over the estimation without additional regressors. The joint significance of all model estimated parameters can also be tested from the value of the probability of the F-Snedecor contrast. In this case the probability is equal to 0'00000 meaning the acceptance of the joint significance of all the parameters of the model. This implies that we can consider all model parameters, jointly taken, significantly different from 0 with a very high probability.

From the estimation we find that the meaning of parameters of the significant variables in the model is important:

- The α coefficient associated with the explanatory variable D , shows that when an individual has participated in the program ($D=1$), the response variable increases by 0'409173. This is the effect on the response variable of participants and means that the ability to find a job increases by 0'409173 over non participants' value.
- The β_{11} coefficient, associated with the explanatory variable X_{11} , shows that in the case of a male participant ($X_{11}=1$), the response variable increases by 0'128794. This means that males have a better position than females in terms of Y , which is higher by 0'128794 than the registered value in the case of females.

- The β_2 coefficient, associated with the explanatory variable X_2 , collects the effect of age on Y . Due to variable X_2 is a quantitative one, this effect will be related to possible values of this variable, adding 0'023647 to the value of the response variable Y for every unitary change registered by X_2 . Therefore, individuals belonging to the sample (individuals between 17 and 26 years old) show a higher ability to find a job as the value of X_2 increases.

To summarize, *Tables 11* (for participants) and *12* (for controls) contain the model estimated values for the response variable Y for every possible value of explanatory variables. As figures show, male participants aged 26 from zone 4 have higher expected values for the response variable Y (though differences between zones are not significant). For any individual, higher age implies a higher expected values of Y . Additionally, is shown that women for all cases always register a lower value of the response variable.

Higher expected values for the response variable Y for male participants must be related with the characteristics of the Spanish Labor Market in which probability to find a job is still higher for male than female (Blanchard and Jimeno, 1995) and (Dolado and Jimeno, 1997). Age seems to be well considered for employers (see the expected values of Y for the case of male participants aged 26).

For political decisions the spread between male and female values of Y could be considered in the sense of Manski (2001) by focusing TSP on groups with higher values of ATET. However this option has legal and moral limits (Cansino and Román, 2007).

If we compare results obtained in this section with papers referred at the end of section 5 a similar discussion can be made.

Table 11

Estimated values for the response variable Y for the participant individuals according to the “differences estimator with additional regressors” model								
	<i>PARTICIPANT INDIVIDUALS</i>							
	Male				Female			
	Zone 1	Zone 2	Zone 3	Zone 4	Zone 1	Zone 2	Zone 3	Zone 4
17 years old	0'692904	0'705659	0'657942	0'733841	0'56414	0'576895	0'529178	0'605077
18 years old	0'716551	0'729306	0'681589	0'757488	0'587787	0'600542	0'552825	0'628724
19 years old	0'740198	0'752953	0'705236	0'781135	0'611434	0'624189	0'576472	0'652371
20 years old	0'763845	0'7766	0'728883	0'804782	0'635081	0'647836	0'600119	0'676018
21 years old	0'787492	0'800247	0'75253	0'828429	0'658728	0'671483	0'623766	0'699665
22 years old	0'811139	0'823894	0'776177	0'852076	0'682375	0'69513	0'647413	0'723312
23 years old	0'834786	0'847541	0'799824	0'875723	0'706022	0'718777	0'67106	0'746959
24 years old	0'858433	0'871188	0'823471	0'89937	0'729669	0'742424	0'694707	0'770606
25 years old	0'88208	0'894835	0'847118	0'923017	0'753316	0'766071	0'718354	0'794253
26 years old	0'905727	0'918482	0'870765	0'946664	0'776963	0'789718	0'742001	0'8179

Source: Own elaboration

Table 12

Estimated values for the response variable Y for the control individuals according to the “differences estimator with additional regressors” model								
	<i>CONTROL INDIVIDUALS</i>							
	Male				Female			
	Zone 1	Zone 2	Zone 3	Zone 4	Zone 1	Zone 2	Zone 3	Zone 4
17 years old	0'283731	0'296486	0'248769	0'324668	0'154967	0'167722	0'120005	0'195904
18 years old	0'307378	0'320133	0'272416	0'348315	0'178614	0'191369	0'143652	0'219551
19 years old	0'331025	0'34378	0'296063	0'371962	0'202261	0'215016	0'167299	0'243198
20 years old	0'354672	0'367427	0'31971	0'395609	0'225908	0'238663	0'190946	0'266845
21 years old	0'378319	0'391074	0'343357	0'419256	0'249555	0'26231	0'214593	0'290492
22 years old	0'401966	0'414721	0'367004	0'442903	0'273202	0'285957	0'23824	0'314139
23 years old	0'425613	0'438368	0'390651	0'46655	0'296849	0'309604	0'261887	0'337786
24 years old	0'44926	0'462015	0'414298	0'490197	0'320496	0'333251	0'285534	0'361433
25 years old	0'472907	0'485662	0'437945	0'513844	0'344143	0'356898	0'309181	0'38508
26 years old	0'496554	0'509309	0'461592	0'537491	0'36779	0'380545	0'332828	0'408727

Source: Own elaboration

7. CONCLUSIONS

The TSP's average effect estimated by weighting observations by the inverse of the estimation of the propensity score let us conclude that, for treated, the time needed to find a job is reduced in 471 days. As the program was designed to improved the employ between youngers unemployed; this result supports the effectiveness of this public policy.

The TSP's average effect estimated by the differences estimator let us conclude that, for treated, the time needed to find a job is reduced in 448 days. This result also supports the effectiveness of this public policy.

Another conclusion can be obtained from using the covariates information contained in the BASEVAFOR. Males have a better position than females in terms the response variable considered. More specifically, on average, the period needed for a treated to find a job, is reduced in 141 days. By considering age, the same period is reduced in 26 days per year from 16 to 25 years old.

According to obtained results, the effectiveness of TSP from the "differences estimator with additional regressor" is positive too, so both evaluations show evidence of that this program contribute to reduce the time needed to find a job.

Summing up results of the analysis, TSP works as an active labour market policy with favourable effects on young unemployment.

Although there is no consensus in literature related with training programs evaluation about its effects our results show evidence in the same way as most of evaluation but stronger. A similar conclusion is derived from a comparison with international evidence. So we have to be prudent with evidence founded.

Further investigations might improve conclusions if public authorities let researchers extend the database information with data related with others individual characteristics.

REFERENCIAS BIBLIOGRÁFICAS

- ASHENFELTER, O. C. (1978): "Estimating the Effect of Training Programs on Earnings", *Review of Economics and Statistics*, 60(1), pp. 47-57.
- BARNOW, B., CAIN, G. and GOLDBERGER, A. (1980): "Selection on Observables", in Stromsdorfer, E. W. and Farkas, G. (eds.), *Evaluation Studies Review Annual*, vol. 5, pp. 43-59,.
- BASSI, L. J. (1984): "Estimating the Effect of Training Programs with Non-Random Selection", *Review of Economics and Statistics*, 66 (1), pp. 36-43.

- BELL, S. H., ORR, L. L. and BLOMQUIST, J. D. (1995): Program applicants as a comparison group in evaluating training programs. W.E. Upiohn Institute for Employment Research. Kalamazoo, MI.
- BILLEWICZ, W. Z. (1965): "The Efficiency of Matched Samples: an Empirical Investigation", *Biometrics*, 21 (3), pp. 623-644.
- BLANCHARD, O. and JIMENO, J. F. (1995): "Structural Unemployment. Spain and Portugal", *American Economic Review (Papers and Proceedings)*, 85 (2), pp. 212-218.
- BURTLESS, G. (1995): "The Case for Randomized Field Trials in Economic and Policy Research", *Journal of Economic Perspectives*, 9 (2), pp. 63-84.
- CAMERON, E. and PAULING, L. (1976): "Supplemental Ascorbate in the Supportive Treatment of Cancer: Prolongation of Survival Times in Terminal Human Cancer", *Proceedings of the National Academy of Sciences of the United States of America*, 73 (10), pp. 3685-3689.
- CAMERON, A. C. and TRIVEDI, P. K. (2005): *Microeconomic. Methods and Applications*, New York, Ed.: Cambridge University Press.
- CANSINO, J. M., CARDENETE, M. A. and ROMÁN, R. (2007): "Regional evaluation of a tax on retailer sales of some fuels through social accounting matrix", *Applied Economics Letters*, 14 (12), pp. 877-880.
- CANSINO, J. M. y ROMÁN, R. (2007): "Evaluación de políticas públicas sobre poblaciones heterogéneas ¿pueden los órganos de control externo contribuir a su avance?" *Revista Española de Control Externo*, 9 (25), pp. 107-129.
- CANSINO, J. M. and SÁNCHEZ, A. (2008a): "Evaluación del programa de Escuelas Taller y Casas de Oficios a partir de su efecto sobre el tiempo de búsqueda del primer empleo. El caso de Sevilla", *Estudios de Economía Aplicada*, vol. 27 (1) ref 27101.
- CANSINO, J. M. and SÁNCHEZ, A. (2008b): "Average effect of training programs on the time needed to find a job. The case of the training programs schools in the south of Spain (Seville, 1997-1999)". *Funcas Workpapers*. Available in www.funcas.es. [Last access: November, 2010].
- CAPARRÓS, A., NAVARRO, M. L. and RUEDA, M.F. (2010): "Rentabilidad salarial de la formación laboral: Un análisis con datos de panel", *Estudios de Economía Aplicada*, 28 (2), pp. 1-20.
- CARD, D., KLUVE, J. and WEBER, A. (2009): "Active labor market policy evaluations: A meta-analysis". *CESIFO Working Paper* nº 2570.
- COCHRAN, W. G. (1968): "The Effectiveness of Adjustment by Subclassification in Removing Bias in Observational Studies", *Biometrics*, 24 (2), pp. 295-313.
- COX, D. R., (1992): "Causality: Some Statistical Aspects," *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, 155 (2), pp. 291-301.

- CUETO, B. and MATO, F. J. (2009): "A nonexperimental evaluation of training programmes: regional evidence for Spain", *Annals of Regional Science*, 43 (2), pp. 415-433.
- DAWID, A. P. (1979): "Conditional Independence in Statistical Theory", *Journal of the Royal Statistical Society, Series B (Statistics Methodological)*, 41 (1), pp. 1-31.
- DAWID, A. P. (2000): "Causal Inference Without Counterfactuals", *Journal of the American Statistical Association*, 95 (2), pp. 407-448.
- DEHEJIA, R. H. and WAHBA, S. (1999): "Causal Effects in Non-Experimental Studies: Reevaluating the Evaluation of Training Programs", *Journal of the American Statistical Association*, 94 (448), pp. 1053-1062.
- DOLADO, J. J. and JIMENO, J. F. (1997): "The causes of Spanish Unemployment: A Structural VAR Approach", *European Economic Review*, 41 (7), pp. 1281-1307.
- FISHER, R. A. (1928): *Statistical Methods for Research Workers*, 2nd ed., London, Ed.: Oliver and Boyd.
- FISHER, R. A. (1935): *The design of experiments*, Edinburgh, Ed.: Oliver and Boyd.
- FRIEDLANDER, D., GREENBERG, D. H. and ROBINS, P. K. (1997): "Evaluating Government Training Programs for the Economically Disadvantaged", *Journal of Economic Literature*, vol. 35 (december), pp. 1809-1855.
- HAHN, J. (1998): "On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects", *Econometrica*, 66 (2), pp. 315-331.
- HECKMAN, J. J. (1990): "Varieties of Selection Bias", *American Economic Review, Papers and Proceedings of the Hundred and Second Annual Meeting of the American Economic Association*, 80 (2), pp. 313-338.
- HECKMAN, J. J., CLEMENTS, N. and SMITH, J. (1997): "Making the Most out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts", *Review of Economic Studies*, 64 (4), pp. 487-535.
- HECKMAN, J. J. and HOTZ, V. J. (1989): "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training", *Journal of the American Statistical Association*, 84 (408), pp. 862-874.
- HECKMAN, J. J., ICHIMURA, H., SMITH, J. and TODD, P. E. (1998): "Characterizing Selection Bias Using Experimental Data", *Econometrica*, 66 (5), pp. 1017-1098.
- HECKMAN, J. J., ICHIMURA, H. and TODD, P. E. (1997): "Matching As an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme", *Review of Economics Studies*, 64 (4), pp 605-654.

- HECKMAN, J. J., ICHIMURA, H. and TODD, P. E. (1998): "Matching As an Econometric Evaluation Estimator", *Review of Economics Studies*, 65 (2), pp. 261-294.
- HECKMAN, J., SMITH, J. and TABER, C. (1998): "Accounting for Dropouts in Evaluations of Social Programs", *Review of Economics and Statistics*, 80(1), pp. 1-14.
- HIRANO, K., IMBENS, G. W. and RIDDER, G. (2003): "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score", *Econometrica*, 71(4), pp. 1161-1189.
- HOLLAND, P. W. (1986): "Statistics and Causal Inference (with discussion)", *Journal of the American Statistical Association*, 81 (396), pp. 945-970.
- IMBENS, G. W. (2004): "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review", *Review of Economics and Statistics*, 86 (1), pp. 4-29.
- IMBENS, G.W. and WOOLDRIDGE, J. M.(2009): "Recent Developments in the Econometrics of Program Evaluation", *Journal of Economic Literature*, 47 (1), pp. 5-86
- KIEFER, N. (1979): *Economic Benefits from Four Manpower Training Programs*, Garland Series of Outstanding Dissertations in Economics, New York, Ed.: Garland Press.
- KLUGE, J., CARD, D. FERTIG, M., GORA, M., JAVOBI, L., JENSEN, P., LEETMAA, R., NIMA, L., PATACCHINI, E., SCHAFFNER, S., SCHMIDT, C., VAN DER LLAUW, B., and WEBER, A. (2007): *Active Labor Market Policies in Europe: Performance and perspectives*, Berlin, Ed.: Springer.
- LITTLE, R. and RUBIN, D. (1987): *Statistical Analysis with Missing Data*, Nueva York, Ed.: Wiley.
- MANSKI, C. F. (2001): "Designing programs for heterogeneous populations: The value of covariate information", *American Economic Review*, 91 (2), pp. 103-106.
- MATO, F. J. (2010): "La formación continua en España desde una perspectiva comparada: Balance y propuestas de mejora", *Papeles de Economía Española*, 124, pp. 266-280.
- MATO, F.J. and CUETO, B. (2008): "Efectos de las políticas de formación a desempleados", *Revista de Economía Aplicada*, 16 (46), pp. 61-83.
- MOERTEL, C., FLEMING, T., CREAGAN, E., RUBIN, J., O'CONNELL, M. and AMES, M. (1985): "High-dose Vitamin C Versus Placebo in the Treatment of Patients with Advanced Cancer Who Have Had no Prior Chemotherapy: a Randomized Double-blind Comparison", *New England Journal of Medicine*, 312, pp. 137-141.
- NEYMAN, J. (1923): "On the Application of Probability Theory to Agricultural Experiments. Essay on Principles.", re-edited in *Statistical Science* (with discussion), 1990, 5 (4), pp. 465-472.

- NEYMAN, J. (1935): "Statistical Problems in Agricultural Experimentation", Supplement to the Journal of the Royal Statistical Society, 2, pp. 107-180.
- PEARL, J. (2000): *Causality: Models, Reasoning and Inference*, Cambridge, Ed.: Cambridge University Press.
- RAMOS, R., SURINACH, J. and ARTÍS, M. (2010): "¿Es necesario reformar las políticas activas de mercado de trabajo en España?", *Papeles de Economía Española*, 124, pp. 281-300.
- ROSENBAUM, P. R. (1995): *Observational Studies*, Springer Series in Statistics, Nueva York, Ed.: Springer-Verlag.
- ROSENBAUM, P. R. (1996): "Observational Studies and Nonrandomized Experiments", in S. Ghosh and Rao, C. R. (eds.), *Handbook of Statistics*, vol. 13, chapter 6: 1277-1366, New York, Ed.: Elsevier.
- ROSENBAUM, P. R. (1999): "Choice As an Alternative to Control in Observational Studies (with discussion)", *Statistical Science*, 14 (3), pp. 259-304.
- ROSENBAUM, P. R. and RUBIN, D. B. (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika*, 70 (1), pp. 41-55.
- ROSENBAUM, P. R. and RUBIN, D. B. (1984): "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score", *Journal of the American Statistical Association*, 79 (387), pp. 516-524.
- ROY, A. (1951): "Some Thoughts on the Distribution of Earnings", *Oxford Economic Papers*, 3 (2), pp. 135-146.
- RUBIN, D. B. (1974): "Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies", *Journal of Educational Psychology*, 66 (5), pp. 688-701.
- RUBIN, D. B. (1976): "Inference and Missing Data", *Biometrika*, 63 (3), pp. 581-592.
- RUBIN, D. B. (1978): "Bayesian Inference for Causal Effects the Role of Randomization," *Annals of Statistics*, 6 (1), pp. 34-58.
- STOCK, J. M. and WATSON, M. W. (2003): *Introduction to Econometrics*, Boston, MA, Ed.: Addospm Wesley.
- WOOLDRIDGE, J.M. (2002): *Econometric Analysis of Cross-Section and Panel Data*, Cambridge, Mass., Ed.: MIT press.
- ZWEIMÜLLER, J. and WINTER-EBMER, R. (1996): "Manpower Training Programmes and Employment Stability", *Economica*, 63, pp. 113-130.