XIII ENCUENTRO DE ECONOMIA PÚBLICA
Almería, 2006

Microeconometric Evaluation of Public Employment Programmes

Empirical Analysis for Portugal

Alcina Nunes *

June, 2005

ABSTRACT

In order to better understand the impact of Active Labour Market Policies in the Portuguese economy this paper presents an empirical analysis of Public Employment Programmes. These are programmes of direct job creation for unemployed individuals, out of "regular" labour market, directly aimed at improving participants' employment chances by providing work experience and informal training. To measure the impact of the selected programmes, microeconometric evaluation techniques are applied to a representative Portuguese dataset. The Propensity Score Matching Estimator is the basic non-experimental methodology to address the "evaluation problem" and the results indicate that the programmes under consideration have a strong negative impact on leaving unemployment, both for men and women, thus suggesting that public employment programmes have not been very effective in avoiding the exclusion from the regular labour market.

JEL Classification: C14, J68
Keywords: Active Labour Market Policies, Public Employment Programmes, Administrative Unemployment Data, Social Program Evaluation, Propensity Score Matching.

* PhD Student in Economics, Faculty of Economics – University of Coimbra, Portugal.

School of Technology and Management - Polytechnic Institute of Bragança
Campus de Sta Apolónia
Apartado 134
5301-857 Bragança
Portugal
Phone: +351 96 298 78 17
E-mail: alcina@ipb.pt

1. Introduction

In the European economies, the public policies on labour market have developed increasingly, during the last decades, combining the traditional passive programmes with active labour market policies (ALMPs)[1]. In the Portuguese economy, however, such increased use of labour market programmes has no equivalent in the evaluation of results. The empirical processes of evaluation are scarce and insufficiently publicised, the evaluation being done, essentially, inside public organizations and institutions in reply to institutional demands. In an economy exposed to budget restrictions, the inexistence of an evaluation of results of ALMPs, which absorbs an important share of expenditure of public policies on labour market, is not desirable neither for the credibility of policies nor for the efficiency in the distribution of limited resources.

The scarcity of empirical research is not only a direct consequence of some limitations, like the ones identified by Heckman, Lalonde and Smith (1999)[2], associated with the complex task of the evaluation of social programmes, but it is also the result of a particular institutional design. In Portugal, the ALMPs have universal features resulting from the acceptance of general concepts and goals, but they are also spread on a wide range of distinct programmes and measures across different economies both in their particular targets and goals and in their institutional designs.

In the Portuguese economy, the ALMPs change continuously over time originating a propagation of programmes, sometimes even overlap, making the evaluation process of particular measures difficult and confining a full assessment of results, particularly from a microeconomic angle and in medium and long run. However, the most important limitation on the evaluation of Portuguese ALMPs relies on difficulties of access to a suitable statistical database.

Being aware of the earlier introduced difficulties, this research work aims to adapt evaluation processes, to the reality of the Portuguese labour market, that allow us to understand the impact of ALMPs on the participating individuals. In particular, it will evaluate the Public Employment Programmes (PEPs) that, considering the lack of an

[1] In literature, despite being distinct, the concepts of “policy” and “programme” are sometimes used in an undifferentiated way. While the term “policy” applies to the generic orientation of the public power to reach defined purposes, the word “programme” means the instrument used to put the policy into practice.
[2] The economic literature of evaluation methods is wide and continues expanding. The analysis of the evaluation problem in economics takes shape in several references in specialized literature which is rich in methodological improvements stimulated by new and more advanced challenges in empirical analysis. The last thirty years witnessed a real progress for the understanding how to assess ALMPs like is shown in the synthesis work of Heckman, Lalonde e Smith (1999).

immediate employment or training solution, fill the subsidized unemployed or unemployed with economic difficulties in activities with public interest that do not compete with the normal functioning of a regular labour market.

There is no known work of empirical evaluation of these programmes for the Portuguese economy, so the assessment carried out in this paper will have as references several evaluation studies performed in different European economies. Dealing with databases constructed from administrative records, these different studies perform non-experimental evaluations that rely on matching estimators under the Conditional Independence Assumption (CIA). Their results are ambiguous but, in general, they are little optimistic regarding the impact of PEPs on the employability of participants.

Brodaty, Crépon and Fougére (2001) estimated the causal effect of four programmes on French youth. Among the on-the-job training programmes, vocational training programmes and courses for labour force participation were the public employment programmes. All the four programmes were evaluated in comparison with fixed-term contracts to understand their impact on the stability of the job (a long term labour contract). They estimated the average conditional effect of each treatment using the propensity scores in a kernel matching estimator. The conclusion was that the programmes were designed to increase the probability of leaving unemployment instead of raising the probability of finding a long term labour contract, and, in the whole set of programmes, the PEPs were the less effective.

Eichler and Lechner (2002) evaluated the effect of the participation in PEPs on the probability of unemployment, in the Sachsen-Anhalt German state, applying a nonparametric identification methodology with different comparison groups over time, to correct selection effects. They found that the participation had a significant reduction in the probability of unemployment in the period after the end of the programme. When they analysed different sub-populations, they also found that for men the reduction resulted from a higher employment probability among participants while for women it resulted from a higher probability of dropping out of the labour force.

Gerfin and Lechner (2002) evaluated the ALMPs that took place in Switzerland using a matching methodology, with a multinomial logit model, applied to more than a state including all the Swiss programmes assembled in three major groups: training, subsidized temporary employment and public employment programmes. With a very rich database – combining information from labour market and social security files – containing information about individuals aged between 25 and 55, unemployed for less than a year and for the first time exposed to ALMPs, the authors analysed the effect of participation on the probability to find a non-subsidised job. The conclusion was that PEPs participation had a negative effect on participants.

Using a nonparametric evaluation approach based on the conditional independence assumption (CIA) combined with a matching estimator, Larsson (2003) in Sweden, estimated the average effects of participation of young unemployed in two ALMPs – the on-the-job training and an occupational programme. The results from PEPs on the employability of youth were not encouraging, although these programmes presented results less negative than the on-the-job training programme. The reasons presented were related to an insufficient planning and monitoring process as well as the performance of less qualified tasks, during the participation in the programme, which did not improve the human capital in an attractive way to potential employees. It was also believed that participants were not sufficiently stimulated to find a regular employment.

Our present analysis will use a database created from administrative registers of *Instituto de Emprego e Formação Profissional* (IEFP), the public employment service in Portugal, which holds the monopoly of the application of ALMPs to the recorded population who looks for alternative job solutions – being unemployed or not. These data are the result of individual records in employment services, spread over the entire national territory, providing information on personal and labour characteristics, the status of the labour market and the participation in ALMPs.

Adopting as an interest parameter the average treatment effect on the treated (ATT), the concept of propensity score applied to nonparametric methods of matching regression will be used to construct a counterfactual result under the hypothesis of selection on observables. The aim is to determine the causal effect of the PEPs on the employability of participants using two different measures: (i) the probability of participants to remain unemployed and (ii) the probability of participants to find a regular employment, comparing their results with the results they would have had if they hadn't participated in a programme.

Our findings confirm the results presented in the European literature. We found that, in Portugal, the participation in PEPs has a negative impact on those who have participated in both measures of employability tested, although the negative results decrease over the time period studied. Women have worse results than men.

The paper is organized as follows. In the next section the institutional framework for the ALMPs in Portuguese economy will be presented with a special focus on the PEPs, being followed by section 3, where the evaluation problem is discussed. There is also a possible econometric solution for a nonexperimental evaluation. In section 4, the data will be presented and described and in section 5, the results achieved, by the application of the econometric matching estimator to the available database, will be presented and discussed. The conclusions are presented in section 6.

2. Institutional Design of Portuguese ALMP: The Public Employment Programmes

2.1 – Portuguese Institutional Framework

Over the last decade, Portugal enjoyed a singular labour market situation, combining a low level of unemployment with structural problems such as low levels of educational and professional qualifications. More than quantitative, the structural problem of the Portuguese labour market is a qualitative one, driving the public labour market policies towards activation of policies. This process of activation of the public labour market policies is related to a diversification of programmes, which aim to find suitable answers for distinct population targets, and is strongly characterized by its training nature and its efforts of progressive integration of unemployed individuals in the labour market, trying to avoid situations of labour and social exclusion (DEEP, 2003).

In connection with the enhanced importance given to ALMPs, there is the concern for promoting preventive actions of long term unemployment. Namely, the activation of unemployment benefits was stimulated to conduct unemployed people to a effective search for a regular job, within an Employment Personal Plan. An Employment Personal Plan begins with the identification of individuals with six/twelve months of unemployment spell, if they are under/above the age of 25 years old, in order to determine the unemployed needs.

The intervention of the Portuguese employment public service is divided into effort interventions – defined essentially by the effort of identification of individuals through interviews and monitoring processes – and into answer interventions that lead the identified individuals to one of the several active programmes available in labour market. Among others (not typified) one can find: (i) Professional Guidance Programmes, (ii) Employment (including self-employment) Incentive Programmes, (iii) Public Employment Programmes, (iv) Training-Employment Programmes and (v) Vocational and Rehabilitation Training Programmes.

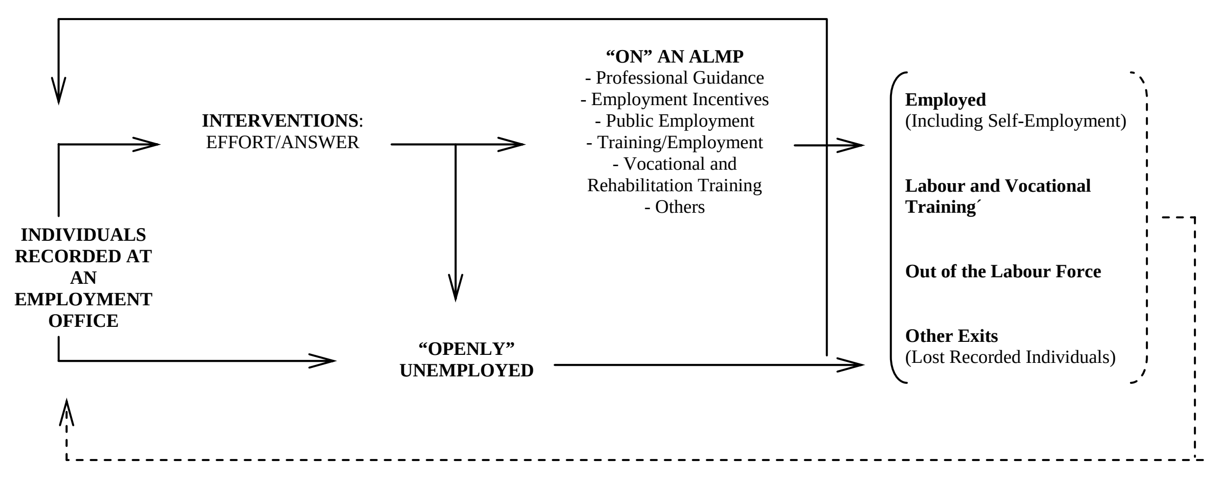
In practice, the set of distinct operational programmes of ALMPs is wide, sometimes running continuously over time, and these are potentially available for all the recorded individuals. On the other hand, the individuals can be recorded repeatedly (and the data show they actually do it) having the right to participate in different periods of time and in different patterns in their observed unemployment spell. It can even happen that the individual takes part in a programme giving continuity to a former distinct one or in result of failure in the specific purposes of that previous programme.

Indeed, although this institutional framework does not fit into a standard evaluation process, where a programme is administered at a fixed point in time and where it is easier to distinguish the individuals by their participation, or no participation at all, in the programme, this is an institutional framework commonly found, in practice, in the European economies (see, for instance, Sianesi, 2004) where one has ongoing programmes and any unemployed individual can potentially become a participant.

After the participation in a programme, there are several destination states for the participant. The main objective of the Portuguese ALMPs is to improve the (re)employability of the unemployed recorded individuals and so employment (including self-employment) represents the main policy outcome. Yet, this is not the only possible outcome. The ALMPs could also have outcomes such as the repeated participation in subsequent programmes, the participation in further labour and vocational training, the leave off the labour force and other non-identified outcomes.

The institutional framework of Portuguese ALMPs is presented in Figure 1. It describes individual flows between different states in the labour market related to a record at a public employment office.

Figure 1: Institutional framework for Portuguese ALMPs – Programmes and Outcomes



2.2 – Public Employment Programmes

The Public Employment Programmes (PEPs)[3] comprise policies of Direct Job Creation. Given the absence of an immediate employment or vocational training solution these programmes directly create occupational activities with public utility, out of the regular labour market. The target population are unemployed individuals collecting unemployment benefits and unemployed in confirmed insolvency, who have to perform occupational activities that do not fill up an existing vacancy, promoted by public and non-profit institutions, by a maximum period of twelve months. The participation acceptance is compulsory for the unemployed who receive unemployment benefits. An occupational benefit, equal to the national minimum wage and partially paid by the public employment service, is given to the insolvent unemployed who agrees to participate in the programme.

When a suitable employment or vocational training offer, supplied by the IEFP, appears it prevails over the participation in PEPs and a refusal ends the participation. The participants need also to carry out, during a given day of the week, an active and confirmed search for a regular job.

The PEPs have as an explicit goal the social useful occupation of the target unemployed individuals and they warrant them receiving one support income. Implicitly, besides allowing unemployed individuals to keep in touch with the labour market while they informally receive basic professional training, the programmes also assure the maintenance of ties with labour market avoiding stigmatization and labour market exclusion.

Given the features presented by PEPs, could the participation make possible a better labour reintegration for those who participate in it, that is, could the participation enhance the employability?

The answer will be based on two different models: a model for the individual probability to leave unemployment records, comparing participation in PEPs with no-participation in any ALMP, during the period of time under analysis, and a model for the individual probability to find a regular job in the labour market when he/she leaves the unemployment records, again comparing participation *versus* non-participation.

[3] Spread by several national legislation, since their creation in 1985, they are now regulated by Portaria 102/96, 30th May.

3. The Evaluation Problem and Nonparametric Resolution

3.1 – The Evaluation Problem

The evaluation problem of social programmes is generally presented as a problem of "lack of information". It can be formalized in a simple way.

At a given moment in time, one individual recorded in the public employment service is in one of two potential situations (D), of which one gave rise to a result (Y): in situation 1, the individual participates in the programme under evaluation, which is called the treatment situation; in situation 0 he does not participate in the programme. The result of this formulation is presented in the Rubin Model, as Heckman and his co-authors call it (1998, 1999):

$$Y = DY_1 + (1-D)Y_0 \qquad (1)$$

So it makes sense to associate both the results and think of their difference as the impact of the programme participation on the individual, that is, the causal effect of the programme participation on the individual is given by the following expression:

$$\Delta = Y_1 - Y_0 \qquad (2)$$

The evaluation problem arises because, for a particular individual in a particular moment in time, it is impossible to observe his participation in the programme (D=1) and, at the same time, his non-participation (D=0) – the individual either participates in the programme under evaluation or not!

This means that only one of the results observed gave rise to the evaluation problem in social policies – it is not possible to know the causal effect of a programme for a given individual because there is no opposite evidence, the counterfactual corresponding to that would have happened in the absence of treatment. The estimation of a treatment effect relies on the artificial construction of the counterfactual result this is, in the inference about a potential result that would have been observed if the individual had not been treated (Rubin, 1974 and Rosenbaum and Rubin, 1983).

In this analysis, to understand the impact of participation in a PEP, that is, on the probability of those who had been treated to find a regular employment, one needs to infer about the same probability if they had not participated.

All the efforts to solve the evaluation problem have been developed to settle the question of construction of the counterfactual. All these efforts are distinguished from each other by the assumption about the relation between the missing data and the data that are available.

Apart from the discussion about the different ways to construct the counterfactual[4] the evaluation literature agrees, in general, with the impossibility to compute Δ for the individual. So, the statistical solution was transferred from the individual level to the global level of population as a way to estimate average interest parameters. In this work, like in several other international research works, the interest parameter is the average treatment effect on the treated (ATT):

$$\Delta^{ATT} \equiv E(\Delta|D=1) = E(Y_1 - Y_0|D=1) = E(Y_1|D=1) - E(Y_0|D=1) \quad (3)$$

This parameter estimates the average impact among those participating in the programme under evaluation.

The available data for programme participants allows us to calculate the mean outcome in the treated state, $E(Y_1|D=1)$. In a social experiment, in which the individuals who would otherwise participate in the programme are randomly attributed to a control group, it is possible to obtain a direct estimate of $E(Y_0|D=1)$. However, in nonexperimental evaluation research no direct estimate of this counterfactual mean is available.

3.2. The Matching Method and the Alternative Matching Estimators

The matching methods have been extensively refined in the most recent evaluation literature. They are now a valuable tool available in empirical methodology, namely in non-experimental evaluation, since it tries to restore the conditions of a social experiment (Blundell and Costa Dias, 2002). These methods are intuitively appealing: an individual who participates in a programme is matched with a non-participant individual who presents the same observational characteristics, so the difference between their results could be attributed to the programme (Deheija and Wahba, 2002).

In a non-experimental evaluation the matching process could be a complex one due to selection problems that could bias the results, so the matching process must rest on strong assumptions. One of these strong assumptions of identification is the Conditional Independence Assumption (CIA) that assumes that treatment assignment (D), conditional on

[4] Blundell and Costa Dias (2002), for instance, divide the evaluation methods into five categories, according to the modality of construction of the counterfactual: randomized social experiments, natural experiments, the matching method, the selection model and the structural simulation model.

observables (X), is independent of the potential results (Y). In formal notation this assumption corresponds to:

$$(Y_1, Y_0) \perp D | X \qquad (4)$$

Under the CIA, the treatment and comparison groups are comparable, on average, when conditioning on X.

$$E(Y_1 | D = 1, X) = E(Y_0 | D = 0, X) = E(Y_0 | X) \qquad (5)$$

A practical implementation problem arises when the vector X is highly dimensional and contains continuous variables. But Rosenbaum and Rubin (1983) showed that matching with a scalar function (X), such as the propensity score, P(X), one uni-dimensional variable defined as the conditional probability of participation given the vector of observed characteristics is sufficient to balance the covariates X, between the treatment and comparison individuals. Therefore, if CIA is conditional on X, it will also be conditional on the propensity score:

$$(Y_1, Y_0) \perp D | P(X), \text{ with } P(X) = \Pr(D = 1 | X) \qquad (6)$$

In order to have empirical content for the ATT, matching also requires that Balancing Hypothesis (Rosenbaum and Rubin, 1983) is satisfied:

$$X \perp D | P(X) \qquad (7)$$

This is, the number of observations with the same propensity score must have the same distribution of observed (and not observed) characteristics, independently from the state of treatment.

This is,

$$0 < P(X) = \Pr(D = 1 | X) < 1 \qquad (8)$$

This is, to satisfy this condition there must be both participants and non-participants for each covariate X. Failure to satisfy this assumption restricts the analysis to the region of support common to all treated and comparison individuals, and the estimated treatment effect needs to be redefined as the average treatment effect on the treated within the common region of support.

After the selection of the observations to match, according to the propensity score, several alternative matching estimators can be used.

Under a CIA, the average effect of treatment on the treated can be estimated by contrasting the mean outcome of the treated individuals with a weighted average of observations in comparison groups according to:

$$\frac{1}{n_1} \sum_{i \in \{D=1\} \cap S_P} \left(Y_{1i} - \hat{E}\left(Y_{0i} \middle| D = 1, P(X_i)\right) \right) \tag{9}$$

where,

$$\hat{E}\left(Y_{0i} \middle| D = 1, P(X_i)\right) = \sum_{j \in \{D=0\}} w(i, j) Y_{0j} \tag{10}$$

and, n_1 denotes the number of treated individuals in the region of common support, S_P the region of common support and $w(i, j)$ represents the weight attached to the jth non-participant individual.

The match for each participant, i, is constructed as a weighted average over the outcomes of non-participants, where the weights $w(i, j)$ depend on the distance between $P(X_i)$ and $P(X_j)$ – the more similar are the ith and jth individuals, in terms of their propensity scores, the higher should that weight be.

The alternative matching estimators comprise the definition of a closeness criterion, a neighbourhood, and the selection of a suitable weight function to associate the comparison individual to each treated individual.

In this paper the choice of the comparison group lies on a comparison group in which the matched control observations consist of the most similar non-participants, regarding the characteristics vector – the nearest neighbour matching. The nearest neighbour matching is intuitively clear to understand so it is a traditional and an easy to implement estimator. It consists in a pairwise matching for every treated individual, in which the closest non-treated individual, regarding its propensity score, is chosen. Defining $C(P_i)$ as the neighbourhood for each i in the treatment group, the nearest neighbour matching sets:

$$C(P_i) = \min_j \left\| P(X_i) - P(X_j) \right\|, j \in \{D = 0\} \tag{11}$$

4. Data

Like in other economic contexts, the availability of statistical data will drive the methodological evaluation process.

The database used was built from the statistical information system of IEFP, which contains relevant information relating to all the individuals' records in the public employment service. These records allow us to follow any individual registered, in one of the local employment offices spread over the country, including the participation in ALMPs. The information system, which produces the primary information used, was modified in 1998 to include information on the whole range of interventions administered to record individuals including interventions that confirm the efforts of the public employment service to identify the unemployment individuals through individual and/or collective orientation and evaluation interviews. However, it was only in the second half of 2000 that the information system has covered all the local unemployment offices. During the last six months of 2000, approximately 200 000 different individuals were recorded, including some 85 000 newly registered.

In order to avoid unknown biases, we only consider individuals registered for the first time, in any local employment service over the country. This is a simplifying assumption, also considered by Gerfin and Lechner (2002) and Larsson (2002), to avoid errors in the identification of individuals. The selected individuals are also exposed to a ALMP for the first time, a procedure that prevents the treatment effects from being intensified by former participation in other programmes.

Not all the individuals recorded were exposed to interventions – effort or answer intervention – in the period under analysis. Some of the newly unemployed individuals were never "identified" by the public employment service during this period and were never exposed to any public intervention. So, we can find individuals in three different levels of the admission process to be enrolled on an active programme: individuals "only" registered as unemployed, individuals registered as unemployed and identified by public employment services through an effort intervention and individuals enrolled in an active programme.

For this empirical research we considered two different samples of individuals: the individuals who were exposed to an effort intervention of identification that could lead to the participation in some active programmes and those who were not.

So, from the nearly 85 000 individuals newly registered, we were left with 28 501 individuals who were in the same stage of the admission process to be enrolled on an active programme, that is, the individuals exposed to interventions effort or/and answer [5]. This number represents 34% of all individuals recorded. And from these individuals, exposed to intervention of IEFP, the participation in a Public Employment Programme (PEP) was given as an answer of intervention to 982 individuals.

[5] The number of individuals exposed to intervention no longer contains the observations of individuals that cancelled their records due to compulsory military functions or removal from local employment offices. A total of 695 observations were then dropped from the sample. This option was chosen because the cancelled records (due to the first reason) reported to labour market external causes, while the cancellation of records for removal from the local employment offices registers resulted in the impossibility of following the individual path over his register spell.

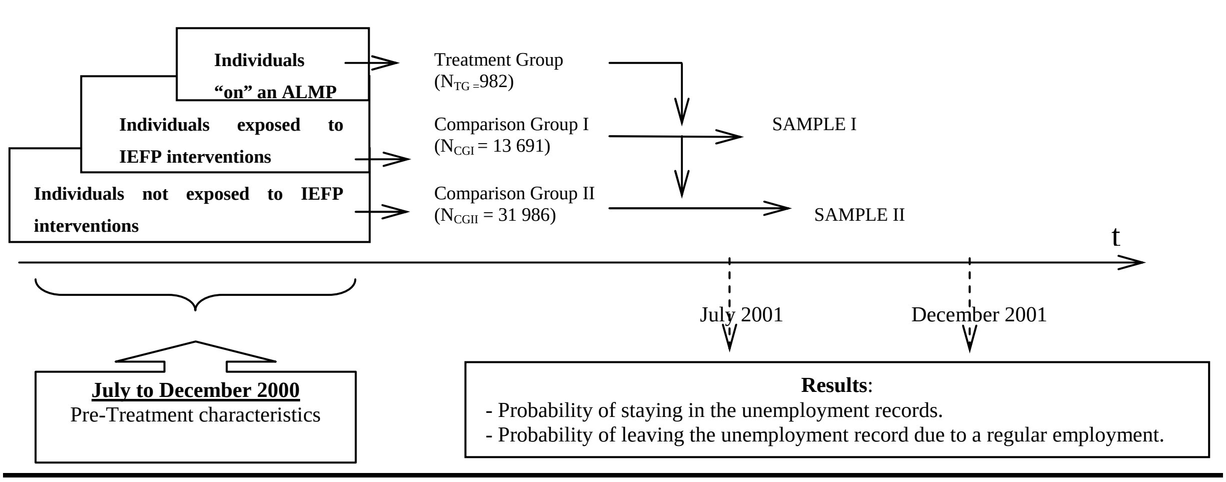
Since the aim of this paper is to assess the effectiveness of PEPs, the set of participants in it will be the group of treatment. The results of participants will be compared with the results of two groups of other individuals, coming from the same local employment services and registered as unemployed, but who had no intervention answer in the sample period. In other words, the comparison group I includes 13 691 individuals not exposed to any kind of active programme, during the period of time under analysis, but identified by the public employment service and the comparison group II includes 31 986 individuals never exposed to any kind of treatment.

Since the treatment group was administratively collected from a major group of unemployed individuals and is not possible, from the administrative data, to capture the reasons involved in the selection process, the assumption of selection on observables could be excessively optimistic. Compare the treatment group with comparison groups collected from different levels of the admission process to be enrolled in an active programme could allow us to investigate the reasonability of the conditional independence assumption.

After the selection of the two samples, the results are estimated for two distinct moments in time: six and twelve months after the end of samples construction for a better understanding of the programme impact on the participants. This is, in a moment during which the PEPs are still running (they can run for a maximum of twelve months) and, in a moment, during which all the participants already left the programmes. For the two selected moments in time, two distinct models will be performed: (i) a model to estimate the probability of staying in the unemployment records after participation and (ii) a model to estimate the probability of leaving the unemployment record due to the exist to regular job.

Figure 2 describes the construction of the sample – treatment and comparison groups.

Figure 2: Time plan for samples construction and for the evaluation process



Regarding the Portuguese institutional framework for Portuguese ALMPs, presented in section 2.1, the first model considers as a success all reasons given to leave the unemployment records

This is an over optimistic hypothesis since the cancellation of unemployment records comprises destinations so distinct as the employment, the labour and vocational training, the abandon of labour force or even the "loss" (employed out of labour force or only unofficially unemployed) of recorded individuals. Thus, the "loss" of a recorded individual could bias the estimated results of the probability of being recorded as an unemployed since it is not possible to know the real reasons for leaving the unemployment register, this is, their real labour market status. Sianesi (2004) argues that the direction of the bias cannot be unequivocally established, *a priori*, because the unknown true labour market status over time, once in a lost status, may be systematically different between treated and non-treated individuals.

The second model considers only as a measure of success the cancellation of the unemployment record due to a regular employment.

The observed individual characteristics before participation are presented in Table 1 and Table 2, describing *ex-ante* the treatment and comparison groups. The following analysis is carried out for the total number of individuals in the two samples (treatment and control groups), and for men and women, separately.

Table 1: Sample Pre-Treatment Characteristics – All Individuals

Characteristics	*Treatment Group (%)*	*Comparison Groups (%)*	
		Sample I	Sample II
Number of Observations	982	13691	31986
Sex (Men)	24,34	37,77	43,46
Age (Years Average)	34,16	29,15	29,14
Educational Level			
- None	17,01	5,57	3,31
- Primary (4 years)	36,15	20,02	18,07
- Compulsory Secondary (9 years)	23,93	35,37	38,1
- Secondary (12 years)	15,68	21,86	25,49
- Superior (> 12 years)	7,23	17,17	15,04
Geographical Location			
- Norte	39,51	23,94	38,65
- Centro	21,08	44,63	7,32
- LVT	20,06	17,63	45,54
- Alentejo	12,12	3,83	5,57
- Algarve	7,23	9,97	2,92
Unemployment Category			
- Looking for the first employment	37,17	39,23	35,7
- Looking for new employment	62,83	60,77	64,3

Some differences, between the treatment and comparison groups under analysis, can be seen in both samples of unemployed recorded individuals (Table 1), mainly composed by relatively young women looking for a new employment and living mostly in the North and Centre of the country. The treatment group has more women and the individuals are slightly older than the ones of the non-treatment groups. On the other hand the treatment group is less educated and we could observe a higher rate of individuals without any level of education.

Analysing men and women, separately (Table 2), it is possible to observe that women represent about 75% of the observations in the treatment group and almost 63% and 57% of the observations in the comparison group I and II, respectively. It is also possible to observe slight differences in the distribution of pre-treatment characteristics between the two groups. Geographically, the two samples are not identically distributed across the country. The women are slightly younger than men, they have educational levels either lower or higher than men, who have essentially average levels of education. Women are looking for a first employment rather than for a new job, while men are essentially looking for a new job.

Table 2: Sample Pre-Treatment Characteristics - Men and Women

Characteristics	*MEN*			*WOMEN*		
	Treatment Group (%)	Comparison Groups (%)		Treatment Group (%)	Comparison Groups (%)	
		Sample I	Sample II		Sample I	Sample II
Number of Observations	239	5171	13901	743	8520	18085
Age (Years Average)	36,77	29,64	29,51	33,33	28,85	28,86
Educational Level						
- None	24,27	4,97	3,0	14,67	5,94	3,55
- Primary (4 years)	33,05	21,95	18,57	37,15	18,85	17,68
- Compulsory Secondary (9 years)	24,69	40,75	44,85	23,69	32,11	32,91
- Secondary (12 years)	14,23	21,27	23,87	16,15	22,22	26,74
- Superior (> 12 years)	3,77	11,06	9,71	8,35	20,88	19,13
Geographical Location						
- Norte	40,17	22,53	38,15	39,3	24,8	39,05
- Centro	16,32	45,0	6,71	22,61	44,40	7,79
- LVT	19,67	16,28	46,07	20,19	18,45	45,13
- Alentejo	17,16	4,22	6,1	10,5	3,59	5,17
- Algarve	6,7	11,97	2,98	7,4	8,76	2,86
Unemployment Category						
- Looking for the first employment	26,36	34,11	33,57	40,65	42,34	37,34
- Looking for new employment	73,64	65,88	66,43	59,35	59,49	62,66

5. Results

5.1 – Propensity Score Estimation

To obtain the estimation of the average treatment effect on the treated, this is on the individuals participating in PEPs, the propensity score matching methodology was used. This methodology matches individuals comparable, in terms of their observed characteristics, in both groups in the sample, being the difference in the results the consequence of the participation in the programme.

The propensity score, P(X), was estimated using a logit model. This model estimates the probability of participation in a PEP compared with non-participation in any ALMP. The results, for the all sample and for men and women, are presented in Table 3 and Table 4.

Table 3: Probability of participation in the program for all individuals

	Sample I	Sample II
	N=14673	N=32968
	N_{TG} =982 (6,69%)	N_{TG} =982 (2,98%)
	N_{CGI} =13691 (93,31%)	N_{CGII} =31986 (97,02%)
	$LR\chi^2(12) = 829{,}48$	$LR\chi^2(12) = 1248{,}18$
	$PseudoR^2 = 11{,}51\%$	$PseudoR^2 = 14{,}13\%$
VARIABLES	Log Likelihood=-3189,1371	Log Likelihood=-3793,594
Sex	-0,622 (*) (-0,080)	-0,805 (*) (0,078)
Age	0,168 (*) (0,020)	0,227 (*) (0,020)
Age2	-0,002 (*) (0,000)	-0,003 (*) (0,000)
Unemployment Category		
Looking for a first employment	0,405 (*) (0,085)	0,980 (*) (0,091)
Looking for a new employment	(a)	(a)
Geographic Category		
Norte	0,624 (*) (0,137)	-1,110 (*) (0,140)
Centro	-0,553 (*) (0,143)	-0,190 (0,151)
Lisboa e Vale do Tejo	0,434 (*) (0,146)	-1,819 (*) (0,148)
Alentejo	1,506 (*) (0,163)	-0,310 (**) (0,163)
Algarve	(a)	(a)
Educational Level		
None	(a)	(a)
Primary (4 years)	-0,421 (*) (0,108)	-1,041 (*) (0,107)
Compulsory Secondary (9 years)	-1,033 (*) (0,125)	-1,920 (*) (0,124)
Secondary (12 years)	-1,011 (*) (0,139)	-2,11 (*) (0,139)
Superior (15 or more years)	-1,703 (*) (0,161)	-2,60 (*) (0,161)
Constant	-5,055 (*) (0,408)	-4,780 (*) (0,415)

(a) Reference Variable; (*) Statistical significance at 1%; (**) Statistical significant at 5%; (***) Statistical significant at 10%
Reference standard errors are presented in parentheses

Since there is no algorithm to choose the set of characteristics X, that satisfies the identification condition, the variables considered relevant to estimate the propensity score are: sex, age, level of education, the unemployment category (unemployed with or without a former employment) and geographical location.

Table 4: Probability of participation in a PEP for Men and Women

	MEN		WOMEN	
	Sample I	Sample II	Sample I	Sample II
	N= 5410	N= 14140	N= 9263	N= 18828
	N_{TG} = 239 (4,42%)	N_{TG} = 239 (1,69%)	N_{TG} = 743 (8,02%)	N_{TG} = 743 (3,95%)
	N_{CGII} = 5171 (95,58%)	N_{CGII} = 13901 (98,31%)	N_{CGII} = 8520 (91,98%)	N_{CGII} = 18085 (96,05%)
	$LR\chi^2(11) = 300{,}91$	$LR\chi^2(11) = 328{,}37$	$LR\chi^2(11) = 495{,}39$	$LR\chi^2(11) = 803{,}34$
	$PseudoR^2 = 15{,}36\%$	$PseudoR^2 = 13{,}54\%$	$PseudoR^2 = 9{,}57\%$	$PseudoR^2 = 12{,}83\%$
VARIABLES	Log Likelihood= -828,7564	Log Likelihood= -1047,9732	Log Likelihood= -2339,3318	Log Likelihood= -2728,1475
Age	0,188 (*) (0,040)	0,219 (*) (0,040)	0,177 (*) (0,024)	0,244 (*) (0,024)
Age2	-0,002 (*) (0,000)	-0,002 (*) (0,000)	-0,002 (*) (0,000)	-0,003 (*) (0,000)
Unemployment Category				
Looking for a first employment	0,654 (**) (0,208)	(a)	(a)	(a)
Looking for a new employment	(a)	-0,971 (*) (0,224)	-0,373 (*) (0,093)	-1,016 (*) (0,100)
Geographic Category				
Norte	-0,861 (*) (0,213)	-0,821 (*) (0,198)	-0,85 (*) (0,148)	-0,750 (*) (0,136)
Centro	-2,461 (*) (0,245)	-0,072 (0,238)	-1,898 (*) (0,154)	0,216 (0,149)
Lisboa e Vale do Tejo	-1,114 (*) (0,238)	-1,670 (*) (0,224)	-1,013(*) (0,158)	-1,409 (*) (0,148)
Alentejo	(a)	(a)	(a)	(a)
Algarve	-1,947 (*) (0,314)	0,098 (0,313)	-1,294 (*) (0,194)	0,418 (**) (0,192)
Educational Level				
None	2,090 (*) (0,380)	(a)	(a)	(a)
Primary (4 years)	1,049 (**) (0,370)	-1,533 (*) (0,187)	-0,218 (***) (0,131)	-0,860 (*) (0,131)
Compulsory Secondary (9 years)	0,682 (***) (0,367)	-2,305 (*) (0,218)	-0,896 (*) (0,153)	-1,760 (*) (0,151)
Secondary (12 years)	0,725 (***) (0,384)	-2,321 (*) (0,257)	-0,899 (*) (0,168)	-2,029 (*) (0,167)
Superior (15 or more years)	(a)	-2,937 (*) (0,381)	-1,608 (*) (0,185)	-2,516 (*) (0,185)
Constant	-6,600 (*) (0,838)	-4,728 (*) (0,704)	-3,379 (*) (0,440)	-4,414 (*) (0,439)

(a) Reference Variable; (*) Statistical significance at 1%; (**) Statistical significant at 5%; (***) Statistical significant at 10%

Reference standard errors are presented in parentheses

The results, for the sample with all observations, show a lower probability for men participation. These are young men who have an educational level beyond compulsory level. Being an unemployed looking for a new job (rather than a first one) also contributes negatively to the probability of participation in the programme. Indeed, women less young, less educated (those with primary education or even with none at all) and unemployed looking for a job for the first time are the ones with the bigger probability to engage in occupational activities with public interest. This is not a surprising conclusion taking into account that the institutions that promote this kind of activities render community assistance traditionally made by women with the characteristics here reported. These results are reinforced by the estimated propensity scores for the samples with either men or women.

5.2 – Balancing Property

The variables includes in the models to compute the propensity score guarantee the balancing property – for a given propensity score, the exposure to treatment is random and so the observations in both groups are, on average, observed in an identical way.

Heckman, Ichimura and Todd (1998) and Lechner (2002), for instance, show the importance of the choice of variables to include in the estimation of the propensity score. They consider that, in practice, an unrefined set of variables tends to increase bias, but economic theory – and, in this case, the nature of the institutional database – does not provide much guidance on how to choose the characteristics vector, X. According to Smith and Todd (2005), the set of characteristics X that satisfies the matching conditions is not necessarily the most comprehensive one because increasing the set that satisfies the identification conditions could lead to a violation of those conditions and to exacerbate a common support problem. The balancing property (equation 7) is not very helpful in choosing the variables to include in vector X either, but it helps to determine interaction and higher order terms to include in it.

The tests used in this paper – provided by Leuven and Sianesi (2003) – calculate, for each variable included in the matching the t-tests for equality of means in the treated and comparison groups, both before and after matching, and the standardized bias, also before and after matching, as formulated by Rosenbaum and Rubin (1985) and implemented, among others, by Lechner (2002). Eichler and Lechner (2002) implemented a variant of this measure suggested in Rosenbaum and Rubin (1985) that was based on standardized differences between treatment and comparison groups in terms of means of each variable in the vector X, squares of each variable in X and first-order interaction terms between each pair of variables in X.

The results from the applied tests suggest no misspecification in the models presented in Tables 3 and 4.

5.3 – Matching Estimation

After the estimation of the propensity score we performed the matching to estimate the difference in the average results between participants and non-participants in order to conclude about the effectiveness of PEPs: (i) in the probability of staying in the unemployment records and, (ii) in the probability of leaving the unemployment record due to a regular employment.

The nearest neighbour matching (with number of neighbours equal to one) was applied in this paper. In the nearest neighbour matching, those who are closer to each participant, according to the estimated propensity score and the number of neighbours, are selected for the comparison group.

Our findings are in Tables 5 and 6.

To assess the precision of our matching estimates, the bootstrap method is used to estimate standard errors for ATT in both matching estimators. For each matching estimator the treatment and comparison groups are matched inside a common support region and the standard error is calculated with 100 replications.

In non-experimental databases is possible to find a support problem if there is a failure of the common support condition – if we want to estimate the counterfactual for a given individual in the treatment group we need to have someone similar to that individual in the non-treatment state. So, to deal with the problem of the inexistence of members of the comparison group, who are truly similar to the ones in treatment group in terms of the propensity score, we imposed a common support, this is, we only used values of the propensity score for which both the density of the treatment group and the comparisons groups are positive. This is a procedure similar to Heckman, Ichimura and Todd (1997) but, unlike these authors, our distributions are sufficiently close so only few values of the propensity score were eliminated. The values obtained for the ATT without common support are very similar to those obtained with the common support, therefore only the last ones are presented.

5.3.1 - Probability of staying in the unemployment records

When we look at the results for staying in the unemployment records (Table 5) six months after the beginning of the analysis period, we can observe strong negative results for the participants in PEPs

In sample I, the participants have almost 45% higher probability to maintain the unemployment record than the non-participants. This value is very high, so the conclusion is that the programmes have a strong negative average effect on the participants in PEPs. For men, the results are less negative, around 41%, essentially because the probability of keeping the unemployed record is not so high for the treatment group. For women the results are even more negative, around 46%, being their results more similar to those found in the sample with all individuals, both for treatment and comparison groups.

Table 5: Probability of staying in the unemployment records (%).

		ALL			MEN			WOMEN		
		Sample		(I-II)	Sample		(I-II)	Sample		(I-II)
		I	II		I	II		I (h_S=0,005)	II (h_S=0,002)	
6 Months	Treatment Group	83,3	83,2		76,99	76,89		85,33	85,25	
	ATT	44,8 (0,031)	63,5 (0,023)	-18,7	40,17 (0,051)	55,46 (0,047)	-15,29	46,03 (0,032)	66,17 (0,032)	-20,14
	Comparison Group	38,5	19,7		36,82	21,43		39,30	19,08	
12 Months	Treatment Group	67,0	66,8		61,51	61,76		68,78	68,61	
	ATT	41,3 (0,025)	51,5 (0,021)	-10,2	34,31 (0,058)	47,90 (0,045)	-13,59	42,13 (0,032)	53,18 (0,028)	-11,05
	Comparison Group	25,7	15,3		27,20	13,87		26,65	15,43	

(a) The bootstrapped standard errors are presented in parentheses for the ATT.

Is important to notice that women perform worse in both groups, this is, independently of treatment the probability of remain in the unemployment records is bigger for the women.

For sample II, the results for the average treatment on the treated are even worse. Considering de group of all participants, we find that they have almost 64% higher probability to remain register as unemployed that the non-participants never identified by the public employment services. This values decrease for men and increase for women, like has been observed for sample I.

The difference between the computed values for the ATT in the samples I and II is present in the column (I-II). It is possible to verify that the differences are driven essentially by the sub-group of women, still participants in PEPs, both men and women, present worse employability results than the non-participants never exposed to some kind of intervention from the public employment service.

Some explanations, for the worse results of treatment group comparing with the comparison groups six months after the beginning of the programme, could be provided.

In the first place, a PEP could last until 12 months, maximum. So, this result could suggest that participants remain in the programme until the maximum term of participation. Indeed, if we analyse the maintenance of the unemployment record twelve months after the beginning of the analysis period, we can verify that the probability of maintaining the unemployment record decreases – the ATT values decreases to 41% in sample I and to 52% in sample II for the group with all the individuals. After the end of the programme, although the participants remain presenting worse results than the non-participants, their labour market perspectives improve. This is especially true for men than for women.

This explanation can also be found in Larsson (2003). Both in Sweden and Portugal, the participants are not stimulated to look for a regular employment in the labour market during the participation period. At the same time, the performed activities might not visibly improve the labour qualification of participants, and, therefore, they do not allow a desirable adjustment of the participants' qualifications to a highly demanding labour market.

Another explanation that can help to understand the existence of a higher probability of maintenance of unemployment records in the first months of participation is related to the nature of (the public and non-profit) organizations that promote PEPs. As in Sweden (Larsson, 2003) these institutions tend to give participants in the programmes tasks that, in another way, have as result a regular employment.

5.3.2 - Probability of leaving the unemployment records due to a regular employment

Table 6 shows less negative results, for participants, from the estimation of leaving the unemployment records due to an exit to a regular job.

The participants still compare relatively badly with non-participation, but the difference is much smaller. In the total sample I, the probability of a participant to leave the unemployment records due to a regular employment, is around 12 % lower six months after the beginning of the evaluation period, and around 9% lower twelve months after this same time period, this is at the end of the legal period of participation in PEPs. In the total sample II, the probability of a participant to leave the unemployment records due to a regular employment, is around 27% lower six months after the beginning of the evaluation period, and around 22% lower twelve months after this same time period.

Table 6: Probability of leaving the unemployment records due to a regular employment (%)

		ALL			MEN			WOMEN		
		Sample		(I-II)	Sample		(I-II)	Sample		(I-II)
		I	II		I	II		I (h_S=0,005)	II (h_S=0,002)	
6 Months	Treatment Group	2,44	2,46		4,18	4,2		1,88	1,89	
	ATT	-12,0 (0,020)	-27,4 (0,023)	15,4	-15,06 (0,037)	-26,47 (0,045)	11,41	-11,98 (0,022)	-28,82 (0,027)	16,84
	Comparison Group	14,5	29,8		19,25	30,67		13,86	30,72	
12 Months	Treatment Group	8,45	8,5		7,95	7,98		8,61	8,66	
	ATT	-8,66 (0,020)	-21,8 (0,024)	13,1	-14,23 (0,038)	-24,79 (0,052)	10,56	-7,27 (0,020)	-22,60 (0,028)	15,33
	Comparison Group	17,1	30,33		22,18	32,77		15,88	31,26	

(b) The bootstrapped standard errors are presented in parentheses for the ATT.

The results for both samples are consistent with the conclusions presented for the probability of staying in the unemployment records – for participants, the results improve with time, for the entire samples and for only men or only women. However, these results show that, in general, the average treatment effect of leaving the unemployment records due to a regular employment is lower for men than for women. The treated women do not have bigger probabilities to leave unemployment due to a regular employment than treated men. The ATT is only better for women because the comparison group of women has worse performance than the comparison men group with respect to a regular employment.

The time improvement for the comparison groups, in both samples, is not so clear. The results presented for the comparison groups remain almost unchangeable from six to twelve months.

The former results renew the explanation that the improvement of participants' results, over time, could be explained because a PEP could last until twelve months, maximum.

Is important to notice that sample II presents much worse results than sample I, sometimes over duplicating the negative results of sample I, in particular for women and for the twelve months period. As for the probability of staying in the unemployment records we find that the comparison groups in sample II perform better than the comparison groups that had been identified by the public employment service and, for that reason, were considered in sample I.

6. Conclusions

The evaluation research presented in this paper, applying a propensity score matching methodology, assuming the conditional independence assumption, to a non-experimental database, is not less pessimistic than the typical European literature on the average treatment effects on the treated. Indeed the employability does not improve with participation and the results are worse for women than for men. Actually, women seem to have more difficulties in leaving the unemployment records, besides the treatment state.

Regarding the sample where the comparison group consists on those identified by the public employment service but had not participated in a PEP, the participants have a probability of maintenance of the unemployment record almost 45% higher than non-participants in a time span in which is possible to maintain the programme participation, and the difference between the two groups decreases about 4% in the period after participation. The estimated results are not so negative for the probability to leave the unemployment records due to a regular employment. The difference between participants and non-participants is, in this case, of about 12% and 9%, for six and twelve months, respectively.

Regarding the sample where the comparison group consists on those never identified by the public employment service but had not participated in a PEP, the results appear much worse for participants. This occurs essentially because the comparison group presents results clearly better.

The difference of results between the two samples, considering the CIA, shows that could exist some kind of administrative selection bias, not capture for any measurable variable. Individuals that have a bigger propensity to participate in a PEP seem to be the individuals with lower probability to succeed in the labour market.

Much in line with the literature, our findings suggest that Public Employment Programmes have not been very effective, particularly for women, in avoiding the exclusion of the regular labour market.

Bibliography

Blundell, R. and Costa Dias, M., (2002), Alternative Approaches to evaluation in Empirical Microeconomics. Portuguese Economic Journal, 1(2), pp. 91-115.

Brodaty, T., Crépon, B. and Fougère, D. (2001), Using Matching Estimators to Evaluate Alternative Youth Employment Programs: Evidence from France, 1986-1988. In Econometric Evaluation of Labour Market Policies, ZEW Economic Studies, 13, Lechner, M. and Pfeiffer, F. (eds), Physica-Verlag Heidelberg, pp 85-124.

Dehejia, R. H., and Wahba, S., (2002), Propensity Score-Matching Methods for Nonexperimental Causal Studies. The Review of Economics and Statistics, 84(1), pp.151-161.

DEEP, (2003), Avaliação do Impacto da Estratégia Europeia para o Emprego em Portugal – Políticas Activas para a Empregabilidade. Colecção Cogitum, nº 2, Ministério da Segurança Social e do Trabalho.

Eichler, M. and Lechner, M. (2002), An Evaluation of Public Employment Programmes in the East German State of Sachsen-Anhalt. Labour Economics, 9, pp. 143-186.

Gerfin, M and Lechner, M., (2002), Microeconometric Evaluation of the Active Labour Market Policy in Switzerland. The Economic Journal, 112 (482), pp. 854-893.

Heckman, J., Ichimura, I. and Todd, P., (1998), Matching as an Econometric Evaluation Estimator. The Review of Economic Studies, 65(2), pp. 261-294.

Heckman, J., Lalonde, R. and Smith, J., (1999), The Economics and Econometrics of Active labour Market Programs. In Handbook of Labor Economics, Volume 3A, Card, O.C. Ashenfelter and D. (eds), Amsterdam: North-Holland, pp.

Larsson, L., (2003), Evaluation of Swedish Youth Labor Market Programs. The Journal of Human Resources, 38(4), pp. 891-927.

Lechner, M., (2002), Some Practical Issues in the Evaluation of Heterogeneous Labour Market Programmes by Matching Methods. Journal of the Royal Statistical Society Series A, 165 (Part I), pp 59-82.

Leuven, E. and Sianesi, B., (2003), "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing". http://ideas.repec.org/c/boc/bocode/s432001.html.Version 1.2.3.

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., (1985), Constructing a Control Group Using Multivariate Matched Sampled Methods that Incorporate the Propensity Score. American Statistician, 39, pp. 33-38.

Rubin, D.B., (1974), Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. Journal of Educational Psychology, 66, pp. 688-701.

Sianesi, B. (2004), An Evaluation of the Swedish System of Active Labour Market Programs in the 1990s. The Review of Economics and Statistics, 86 (1), pp.133-155.

Smith, J.A. and Todd, P.E., (2005), Does Matching Overcome LaLondes's Critique of Non-Experimental Estimators? Journal of Econometrics, 125, pp.305-353.