Short and long term evaluations of Public Employment Services in Italy

P. Naticchioni and Silvia Loriga

Discussion Paper 2008-30

Département des Sciences Économiques
de l'Université catholique de Louvain

UCL

Short and long term evaluations of Public Employment Services in Italy

Paolo Naticchioni
Univ. Catholique de Louvain, Univ. di Roma La Sapienza
paolo.naticchioni@uniroma1.it

Silvia Loriga
Istat, silvia.loriga@istat.it

October 2008

Abstract

In this paper we evaluate the efficacy of Public Employment Services in Italy (PESs) in increasing the unemployment to employment transition probabilities, through matching techniques. Exploiting the longitudinal dimension of the LFS data we design an evaluation structure that allows observing outcomes in both the short (at most 3 months) and the long run (at most 12 and 15 months). In this framework, PES users show a lower probability of finding a job in the short term, because of a lock-in effect, while in the long term this probability turns out to be positive. Similar results are derived when considering as outcome a proxy for the quality of the job, i.e. having found a permanent job. Robustness checks and sensitivity analysis confirm this evidence, suggesting also that our estimates are to be considered as lower bound in the eventuality of unobserved negative selection into treatment.

Keywords: Public Employment Services, Active Labour Market Policies, Propensity Score Matching, European Employment Strategy, Policy Evaluation, Italy;

JEL codes: J64, J68.

1. Introduction

The last decade has seen lively debate arising over the role and the effectiveness of active labour market policies (ALMP), especially at the European level. Also the European Employment Strategy places great emphasis on the role of active labour market policies. For instance, the European Commission "calls for a strengthened emphasis on activation and prevention policies in order to limit the unemployment spell, and prevent inflow into long term unemployment, detachment from the labour market and inactivity".[1] In line with this institutional framework, most of the European countries have reformed their active labour market policies in order to accomplish the guidelines of the European Employment strategy. These reforms have hence generated a spread of evaluation exercises for most of European countries.[2]

The main aim of this paper is to fill the gap in the literature concerning the Italian case, assessing the efficacy of Public Employment Services (PESs) after the reforms introduced in 1997, 2000 and 2003. Basically, these reforms pursued three main goals: a) improving, at the local level, the PES governance of the labour market; b) enhancing the employability of the unemployed that face greater difficulties in finding a job (unskilled, long term unemployed, women, etc.); c) providing services in order to increase the efficacy of matching processes between labour demand and supply. Subsequent to these reforms, PESs are required to provide job search assistance, counselling, training schemes and job proposals (intermediation) to their clients. It is also worth noting that in Italy passive labour market policies are not widely developed as in other European countries,[3] and they are not actually interacted with active labour market policies. For this reason, our analysis disregards issues related to unemployment benefits, since according to the labour force data (LFS) data they are negligible in the Italian labour market (less than 1% of the unemployed in our analysis).

We use the LFS data for the period 2004-2006. Households selected in the LFS sample have to be interviewed four times during a 15 months period, according to a rotation scheme; merging data collected in the four interviews –at t_1, t_2, t_3 and t_4- we can observe transitions in the labour market (t_2-t_1 = 3 months; t_3-t_1 = 12 months; t_4-t_1 = 15 months).

* We thank Bart Cockx, Muriel Dejemeppe, Stefano Gagliarducci, Claudio Lucifora, Tommaso Nannicini, Enrico Rettore, Henri Sneessens, Barbara Sianesi, Henri Sneessens, and Bruno Van der Linden for their suggestions. We also thank Isfol for having funded this research project. The usual disclaimers apply. Moreover, the opinions expressed in this paper are ours and do not necessarily reflect those of our Institutions.

[1] European Commission (2004), p.26-27.

[2] Among others, see for instance Blundell et al. (2004) for UK, Crepon et al. (2005) for France, Gerfin and Lechner (2002) and Lalive et al. (2008) for Switzerland, Sianesi (2004) for Sweden, Weber and Hofer (2004) for Austria, Van den Berg et al. (2004) for the Netherlands, Lechner and Wunsch (2008) for Germany.

[3] According to OECD (2007), in 2004 the share of GDP referred to passive labour market policies are equal to 0.65% for Italy, 1,71% for France, 2.67% for Denmark, 2.32% in Germany, 1.49% for Spain. Note also that in Italy there is not a welfare policy to better off the economic situation of individuals who do not have any source of income. This is another important difference with other OECD countries.

As far as the econometric technique is concerned, we use propensity score matching, as in other papers in the literature such as Blundell et al. (2004), Gerfin and Lechner (2002) and Sianesi (2004). As treatment variable we consider enrolment in the PESs. However, it would not be appropriate to define as treated those who declare to be enrolled in the PESs in the first of the LFS interviews, because in Italy enrolment in the PESs could have taken place a long time before the LFS interview: an unemployed might be enrolled in a PES for years and at the same time have no contacts with the PES in recent job searching.[4] We claim that it is possible to define a more appropriate treatment variable exploiting the longitudinal dimension of the LFS. More specifically, we select all the unemployed that are not enrolled in a PES at t_1. We then follow these individuals over time defining as treated those unemployed that get enrolled in a PES between t_1 and t_2. Further, we observe the outcome variables, i.e. being employed, both in a short term evaluation at t_2 (after a period of time between one and nineties days from the actual treatment), and in a long term evaluation at t_3 (after a period of time between 9 and 12 months from the actual treatment) and t_4 (after a period of time between 12 and 15 months from the actual treatment). We hence evaluate whether the treated individuals display higher employment probabilities than the untreated, using propensity score matching.

As in other papers (Sianesi, 2004), we show that it is crucial to compute evaluations in both the short and the long run. In particular, we point out that average treatment effects on the treated (ATT) are negative in the short run, while they become positive in the long run. We argue that in the short run the treated might be involved in a sort of lock-in effect, because they spend time in activities such us orientation periods, preparing CVs, training courses, etc. In the long run, when these activities are over, the treated display higher probabilities of finding jobs: after 12 (15) months around 8.3 (7.1) percentage points higher than the baseline probability for untreated of 29.7% (29.6%). Various robustness checks confirm these results, and a regional analysis underlines that our findings are more pronounced in the Centre-North region of the country.

Another important dimension of the analysis concerns the quality of the match, as shown in recent related literature (Blundell et al., 2004, Crepon et al., 2005). In Italy, as in other countries characterized by segmented labour markets, permanent contracts can be considered as a reliable proxy for the willing of the employer and the employee to invest in that match over time. In an additional evaluation exercise we investigate whether a treated in a PES displays a higher probability of finding a permanent job with respect to the untreated. We use the same evaluation structure and treatment variable as in the previous analysis, while the new outcome variable is equal to one if an individual finds a

[4] Note that this definition of the treatment variable was used in previous papers for Italy (Barbieri et al., 2002, 2003), which investigated the efficacy of PESs in an intermediate step of the reform process, using data for the year 2000.

permanent job and to zero in all other cases. The results show that in the short term ATT are still negative (-3.4%), while they become not statistically different from zero at t_3 and positive (3-4%) at t_4. This means that the PES effects are less pronounced (lower in magnitude and not always significant) when considering permanent contracts as outcome variable, both in the short and in the long term.

It is also worth underlying that since LFS data provide some evidence that PESs mainly supply services related to counselling and intermediation activities, our results are in line with other European evaluations that stress the efficacy of these kind of policies (Blundell et al., 2004, for the UK, Crepon et al., 2005, for France, Weber and Hofer, 2004, for Austria).

The main potential critic to our evaluation exercise might be that since propensity score matching is based on selection on observables it is not possible to deal properly with selection on unobservables, which might entail a bias in the computation of our ATT estimates. However, we claim that propensity score matching is the best methodology we can use, for several reasons. First, because the LFS provides a particularly rich set of control variables – a necessary requirement in order to carry out the propensity score matching analysis, based on selection on observables. Second, because convincing instrumental variables are not available to deal with endogeneity issues of the treatment variable, as well as convincing thresholds for regression discontinuity designs. Nevertheless, we make use of the sensitivity procedure developed by Ichino et al. (2008) to assess the robustness of our ATT estimates to possible deviations from the original setting of the Conditional Independence Assumption (CIA), the main untestable assumption of matching procedures.

Furthermore, the sensitivity analysis helps us to investigate further the issue of unobserved heterogeneity. Actually, it is usually argued (Barbieri et al. 2002, 2003) that in Italy a negative selection is at work in enrolment in PESs, i.e. high-ability and high-educated workers resort to PESs less frequently than the low-skilled. The sensitivity analysis allows us to simulate confounders characterized by a negative selection effect and a positive outcome effect. When introducing such a confounder, ATT estimates are higher than the baseline figure. This means that if our set of covariates were not fully able to capture the unobserved heterogeneity, the ATT point estimates would range in a given interval of variation (8.3%, 10.8%), suggesting that our ATT estimates should be considered as lower bound of the 'true' PES impact.

The paper is structured as follows. Section 2 describes the reforms concerning PESs over the last decade, and Section 3 provides a short explanation of the LFS data we use. Section 4 describes the PES evaluation structure, while identification issues are discussed in section 5. Section 6 presents the main results and robustness checks are reported in section 7. Section 8 focuses on the sensitivity analysis and Section 9 investigates further the issues of unobserved heterogeneity. Section 10 concludes.

2. PES reforms in Italy in the last decade

Over the last decade the intermediation role of active labour market policies has been an object of investigation in the economic and political debate, at both the European and the Italian level. The European Employment Strategy (EES), set up in the 1997 and updated several times in the last few years, stressed the importance of reform in public and private employment services in order to enhance employability in the labour market and reduce both the inefficiency associated with the mismatch between labour demand and supply and the social costs due to unemployment.

As far as the Italian PESs are concerned, a number of legislative reforms have been introduced: in 1997 with the so-called 'Pacchetto Treu', then in 2002 with the 297/2002 decree and finally in 2003 with a new decree (30/2003). Before these reforms, the PESs were managed at the national level by the central administration, taking care almost exclusively of the administrative certification of recruiting and of the listing of job offers and job seekers. In particular, the PESs had to record the information concerning the unemployment spells and the transitions towards employment. The efficacy of PESs in matching labour demand and supply was perceived as very low, the core of the PES activities being mostly administrative.

The reforms introduced in 1997, 2002 and 2003, pursued three main goals: improving, at the local level, the PES governance of the labour market; enhancing the employability of the unemployed that face greater difficulties in finding a job (unskilled, long term unemployed, women, etc.); increasing the efficacy of matching between labour demand and supply.[5] In particular, PES activities consist of a complex system of functions designed to reduce unemployment duration and improve the information flow between demand and supply in the labour market. These functions can be summarized with the following general tasks: a) collecting information on labour supply and labour demand in the local labour market; b) identifying priority target groups (long term unemployed, unskilled, women, disabled, immigrant citizens); c) providing individual services and placement programmes; d) supporting job search and participation in professional training courses, easing the access to the labour market; e) providing counselling to companies, information and support on existing specific incentives (collective dismissals policies, tax reductions, assistance on outplacements, etc.); f) promoting self-employment (job creation schemes). Furthermore, the reform process also established that the system of the PESs had to be decentralized to the regions (Regioni – NUTS2), in order to make them more effective in the local labour markets. The regions kept for themselves the

[5]Another goal of these reforms was to introduce the private employment services in the Italian labour market, since labour intermediation was solely public up to 1997. For an evaluation of private agencies in Italy see Ichino et al. (2008). See also section 7 for a robustness check concerning private employment services.

strategic planning of services, while the everyday running of the PESs was in turn decentralized to the provinces (Province– NUTS3).[6]

The new PES setting in Italy establishes that the PESs have to offer their clients one of the following three alternatives:[7]

- a personal counselling interview within three months from the unemployment declaration,[8] in which the staff of the PES illustrate to the unemployed the possibilities to find a job in that provinces (training courses, vacancies opened by firms, etc.);
- a short vocational training course and/or work practice, within 6 months from the unemployment declaration; this time period is reduced to 4 months in the case of young and of women out of the labour market for more than two years;
- a job proposal.

An individual can, on a voluntary basis, decide to enrol in a PES. If enrolled, he has to accept the program established by the PESs.[9]

In this framework, the goal of our analysis is to evaluate whether these PES programmes have an impact on the employment transitions for the unemployed. In particular, we are interested in evaluating the impact of PES treatment on the sample of unemployed aged 15-64 at t_1.[10]

It must also be stressed that, according to OECD (2007), a very high share of ALMP in Italy (about 60%) refers to incentives to create new jobs, such as the incentives provided for apprentiship contracts or for training on the job. As for the policies targeted to the assistance of the job-search process of the pool of unemployed, the PESs represent the most important active labour market policy in Italy. In particular, the LFS data show that of the 2,700,000 individuals looking for a job in 2004 (700,000 involved in on-the-job-search), 28% (about 750,000) had contacts with PESs. A rough and conservative estimate of the total cost of PES is about 600,000 million euros (0.09% of GDP in 2004), partially financed by the European Social Fund.[11]

As last important remark, we underline that in Italy active labour market policies (PESs) are not strictly related to passive labour market policies. In particular, eligibility

[6] For this reason, there may be some geographical differences, due to the different forms of PES organization chosen by local authorities.

[7] As explained in section 6, we do not know exactly in which specific program an individual is enrolled in. In the LFS it is available only the reason of the last contact with a PES.

[8] Note that the unemployment declaration does not correspond to the beginning of the unemployment spell. It is instead related to the time in which the unemployed decide to resort to the PES.

[9] Actually, there are no pecuniary sanctions for unemployed enrolled to PES that refuse proposals from PES. However, in case of refusal they can be excluded from the program.

[10] We do not consider inactive people as well as the assistance provided by PES to the on-the-job search activities carried out by the employed.

[11] See Pirrone and Sestito (2006) for details concerning these estimates, and also for a juridical, economic and political description (and discussion) of the whole reform process.

for unemployment benefits is not affected by the monitoring concerning participations into active labour market policies. The interaction between active and passive policies is mostly formal, in the sense that individuals eligible for unemployment benefits have to get enrolled in a PES before starting receiving the benefits, mainly due to administrative reasons. It is also worth noting that the share of unemployed receiving unemployment benefits is very low, about 3% according to the LFS. The same share in the sample of unemployed that we use in the evaluation exercise in this paper is even lower, less than 1%. We eliminate these observations from the analysis in order to avoid the possibility to detect some enrolled in a PES that are not actually looking for a job using PES, but are enrolled only to receive the unemployment benefit. [12]

3. The Labour Force Survey data

We use the LFS data provided by Istat. This survey was completely overhauled in 2004, from the old RTFL (Rilevazione Trimestrale Forze Lavoro) to the RCFL (Rilevazione Continua Forze Lavoro), according to the Eurostat guidelines. For our analysis two elements need mentioning in particular. First, the data quality has significantly improved, thanks to both the computer assisted technique utilized in collecting data and the new professional interviewers network. Second, the section relating to the PESs has been thoroughly revised, enhancing the set of the collected data (see Istat, 2006).

As for the structure of the Italian LFS, the sample design follows a 2-2-2 household rotation scheme: households participate in the survey for two consecutive quarters, temporarily exit from the sample for the following two quarters, and then re-enter the sample for the last two quarters, after which they finally exit, as shown in figure 1. We use LFS data from the first quarter of 2004 – the official beginning of the new survey – to the third quarter of 2006, bringing together six consecutive household rotation groups. As already mentioned, we observe individuals at four interviews, at t_1, t_2, t_3, t_4 (as in figure 1, the rotation groups from A to F, from the first to the fourth interview). The second interview takes place three months after the first one, the third 12 months after the first one (9 months after the second one), and the fourth 15 months after the first one. We use the four interviews, instead of the two (t_1 and t_2) used by Barbieri et al. (2002, 2003), because in the first one we collect information on the control variables in a pre-treatment status, in the second we derive the treatment variable and the outcome for the short term evaluation, and in the third and the fourth interviews we observe the outcome for the long term evaluation (as explained in the following section). Consistently to the propensity score matching theory, the covariates are observed in pre-treatment status, treatment then takes place, and finally the outcome is observed. This means that treated

[12] Results basically do not change when including in the initial sample those receiving unemployment benefits, suggesting that their elimination from the initial sample is irrelevant to achieve the results.

and untreated are compared using pre-treatment variables, matching individuals that were the same before the treatment took place.[13]

LFS also provides a very rich dataset, containing a wide range of control variables – a necessary requirement in order to carry out the propensity score matching analysis, based on selection on observables. We use the following information at t_1:

- personal information: age, gender, education, education lag,[14] search intensity (number of search actions undertaken to look for a job), previous job experience, potential experience, reasons related to the last job-separation if any (dismissal, retirement, temporary job), occupation in the last job if any, adaptability to accept a fixed term contracts, adaptability to accept part time contracts, adaptability to accept jobs involving geographical mobility, unemployment duration;
- household information: number of household members, number of members of the household enrolled in a PES excluding himself, number of members of the household employed;
- macroeconomic variables estimated using the LFS data at the provincial level (NUTS-3): unemployment rate, employment variation (with respect to the previous interview, in percentage), turnover rate as a measure of labour market dynamics (computed between t_1 and t_2), agriculture employment share, ISFOL index of PES endowments and structures,[15] dummies for the related household rotation group.

4. Structure of the PES evaluation exercise and definition of treatment

In this paper we use matching techniques. One of the main requirements to apply these techniques properly is to make treated and untreated individuals as comparable as possible. Computing the treatment variable appropriately is the major task to address in order to achieve this comparability. The easiest way would be to define as treated all the unemployed that declare to be enrolled in a PES in the first of the LFS interviews. This treatment variable definition, which has already been used in previous papers (Barbieri et al. 2002, 2003), cannot, we hold, be taken as a reliable treatment variable. This is because in Italy the enrolment in a PES is not closely related to the period of time in which the

[13] Another methodological problem related to Barbieri et al. (2002, 2003) is the fact that control variables and treatment variables are observed in the same instant of time (at t_1). We solved this problem exploiting four interviews of the LFS.

[14] Education lag measures the years of delay in achieving the related educational level for each individual. It is our contention that this variable captures, at least partially, an unobserved heterogeneity within levels of education (we do not have information about marks). In Italy there is a great variability in the number of years taken to complete degrees, especially for graduates and upper secondary school students.

[15] The ISFOL index refers to the year 2004. It is self-declared by each province administration, and consists of various components concerning the quality of PES infrastructure and services (computers, number of employees, range of services, quantity and quality of services, etc.). It can be considered as a proxy of PES endowments and structures, which might be taken into account by the unemployed in their decisions to enrol.

interview takes place: an unemployed can be enrolled in a PES for years and yet have had no contact with the PES for recent job search activities. In other words, an unemployed person enrolled in a PES that does not use PES services is not necessarily cancelled from the list. According to the LFS data, in 1999 in Italy there were 2.6 millions unemployed, while the enrolment lists of PES counted 7 millions individuals, making clear that many unemployed were not cancelled out from PES lists.

For these reasons we design a different structure for our evaluation exercise. More specifically, we consider two pieces of information. First, we use, as in the previous papers, the information related to the enrolment in a PES.[16] Second, we also exploit the information concerning the fact of having had contacts with PES in the last three months (between t_1 and t_2) in order to look for a job. Hence, our treatment variable is structured as follows. To begin with, we select in the first LFS interview all the unemployed not enrolled in a PES at t_1.[17] We then follow these individuals in the period of time between t_1 and t_2, and we define as treated those that in the meantime have enrolled in a PES and have had a contact with a PES to look for a job. In our opinion this is a more appropriate way to focus on those unemployed who effectively used the PES to look for a job in a period of time close to the interview.

As for the binary outcome variable (being employed, according to the Eurostat definition of employment), it is computed for the short term evaluation at t_2 (after a period of time between one and nineties days from the actual treatment), and for the long term evaluation at t_3 (after a period of time between 9 and 12 months from the actual treatment) and t_4 (after a period of time between 12 and 15 months from the actual treatment).

Figure 2 sums up the structure of our evaluation procedure. We select 2759 unemployed "not enrolled" in a PES at t_1. We then go on to observe those individuals who get treated between t_1 and t_2 (348 individuals). We observe the outcome variable at t_2 for a short term evaluation, and at t_3 and t_4 for a long term evaluation.[18] Descriptive statistics and detailed explanations for the variables included in the set of covariates, by treatment status, are reported in Table A1 in appendix: treated and untreated display very similar observed characteristics. Furthermore, it is worth pointing out that the

---

[16] Formally since 2002 PES registration does no longer exist, and it has been replaced by the formal declaration of unemployment status, which is usually still perceived by the unemployed as PES enrolment. These changes were taken into account in the LFS questionnaire in 2005. This means that in 2005 we consider as treated those who declare to have carried out the unemployment declaration.

[17] Note that we consider as untreated at t_1 also the unemployed who declare to be enrolled in a PES and at the same time had the last contact with a PES more that 3 years ago. As already stressed, this situation can take place since in Italy it was possible to be enrolled in a PES without having any contact with a PES recently to look for a job. As robustness check, we considered as untreated at t_1 also the unemployed who declare to be enrolled in a PES and at the same time had the last contact with a PES more that 2 years ago. Results only slightly change and are available on request.

[18] This scheme implicitly assumes that, for those who are treated and are able to find a job between t_1 and t_2, the treatment takes place before the outcome.

difference in elapsed unemployment duration at t_1 is not statistically different from zero between treated and untreated, suggesting that elapsed unemployment durations should not play a relevant role in the selection into treatment, as also confirmed by the probit estimate of the propensity score in section 6.

A last remark concerns the fact that, unfortunately, we cannot exactly identify the program a treated entered in (counselling rather than training or intermediation), since this information is not available in the LFS data. We only have information concerning the reason of the last contact the individual had with a PES. We will deepen this issue in section 6.

5. Propensity score matching and identification

In this paper we use propensity score matching, whose basic assumption is selection on observables (unconfoundedness): selection into treatment is entirely determined by observed variables, and conditional on these variables the assignment into treatment is assumed as random. In comparison with OLS, this technique affords better scope in both dealing with common support issues and using a non-parametric specification in the outcome equation.

The first step of this technique is to compute the propensity score, i.e. the probability of participating in treatment conditional to pre-treatment control variables. Then, by comparing treated and untreated with the same propensity score in the common support region, it is possible to estimate the average treatment effect on the treated (ATT). Since it is often unfeasible to have individuals with exactly the same propensity score, various algorithms are usually applied to match treated and untreated. In this paper we use four different methods in order to test the robustness of results: nearest neighbour, radius, kernel and stratification.[19]

As for the treatment variable, it has been defined in the previous section: T is equal to 1 if between t_1 and t_2 the individual enrols in a PES and have contacts with PES to look for a job, and T=0 if not. In our short (long) term analysis, the outcome variable is computed observing the employment status at t_2 (t_3 and t_4).

If we had data on an experimental design, we would be able to derive the following unbiased ATT:

$$\Delta = E(Y(1) - Y(0) \mid T = 1) = E(Y(1) \mid T = 1) - E(Y(0) \mid T = 1).$$

where *Y(1)* is the outcome in case of treatment and *Y(0)* in case of no treatment. Unfortunately, a controlled experimental design is not available. Hence, while the first term on the right-hand side is identified in the data, the second term –the outcome of the

[19] In this paper we do not go into details of the propensity score matching procedure. See Rosenbaum & Rubin (1973), Dehejia. and Wahba (2002), Caliendo and Sabine. (2005). See also Ichino and Becker (2002) for an explanation of the software we use.

individual in his/her counterfactual treatment situation- is not; this is the usual missing data problem in the program evaluation literature. In order to cope with this matter the conditional independent assumption (CIA) is used, which states that given a set of covariates, X, the counterfactual distribution of the treated is the same as the observed distribution of Y(0) for the untreated, i.e. $[(Y(0), Y(1))\perp T|X]$.

Using the CIA it is then possible to identify the 'missing counterfactual', since:

$$E(Y(0)|T=1) = E_X[E(Y(0)|X,T=1)|T=1] \overset{CIA}{=} E_X[E(Y(0)|X,T=0)|T=1],$$

where in the last term the observed outcomes of the untreated group (*T=0*) are averaged with respect to the distribution of *X* in the *T=1* group. Hence, the ATTs we were looking for become:

$$\Delta = E[Y(1) - Y(0) \mid T=1] = E_X.\{[E(Y(1) \mid X, T=1) - - E(Y(0) \mid X, T=0)] \mid T=1\}$$

In other terms, what matching does is to stratify the data into cells defined by each value of X. Then, within each cell (i.e. conditional on X) it computes the difference between the average outcomes of the treated and the controls, and finally it averages these differences with respect to the distribution of X in the population of treated units.

It is also important to stress that the key choice faced by the unemployed in the period (t_1, t_2) is not simply whether to participate or not, but whether to participate in a program in this time interval or not, continuing the search for a job outside the program with the knowledge that it will always be possible to participate later on. As in Sianesi (2001a, 2002a, 2004), in this paper treatment can be understood in the sense of starting a program in a given period of time (between t_1 and t_2) while as control group we consider those individuals that were untreated at t_1, and that do not get treated in the period (t_1 ,t_2), no matter whether they are to be treated between t_2 and t_3 or between t_2 and t_4.[20]

In this framework, and using this CIA formulation, we compute both a short and long term evaluation. As we will show later on, this distinction is indeed crucial in the interpretation of the results.[21]

---

[20] As Sianesi (2004) claims: "What is required is thus that conditional on having reached the same unemployment duration and conditional on all the relevant information observed, the fact that an unemployed individual goes into a program in a given month while another waits longer is not correlated with the future labour market states the joining individual would have experienced had he instead not entered the program at that time." (p.137). Another way to restate the peculiarity of this CIA assumption is that individuals can be assumed as myopic conditional on observables: given X, outcome-related information about the future (after t_2) plays no role in individual decisions to join a program between t_1 and t_2 or to wait longer. Similar assumptions are made in other papers in the literature, as Lalive, van Ours, Zweimuller (2008) and Fredriksson and Johansson (2003). We carry out a robustness check concerning this assumption in section 7.

[21] Moreover, in this paper we do not have to worry about individuals that do not enter a PES program because they already know that they will soon be starting a new job. In particular, the CIA would be violated if an individual decided not to enroll because he had received an offer for a job that was to start soon. In the Italian LFS data it is possible to identify such individuals, and we drop from the analysis these (very few) cases.

6. PES evaluation: estimates and results

Table 1 shows the propensity score estimates, using a probit.[22] While some variables are not significant (gender, potential experience, number of employed and dimension of the household, adaptability to fixed term, PES performance index, education lag) for the others we derive the expected sign of coefficients. It is also worth pointing out that the pseudo R^2 of the probit is quite low, around 0.10. This confirms that our evaluation structure reduces the observed heterogeneity between treated and untreated.

Table 2 shows the ATT coefficients estimated in the common support region.[23] As for the short term, the first line of table 2 shows that ATTs coefficients are negative and significant, meaning that PES enrolment decreases the probability of going through transition from unemployment to employment in the short run, no matter which the matching procedure used (radius, nearest neighbour, kernel, stratification).[24]

The second line of table 2 shows the corresponding results for the long term evaluation, i.e. employment transitions from t_1 to t_3, between 9 and 12 months as from the treatment. ATT coefficients are positive and mostly significant: the 'PES enrolment' treatment produces an increase in the probability of finding a job by about 8.3 percentage points (using the radius method), with respect to the baseline probability of about 29.7% (defined as the probability of an untreated to be employed at t_3). As robustness check, the third line of table 2 refers to the evaluation of the employment transitions from t_1 to t_4, after a period of time between 12 and 15 months from the treatment. Results are consistent with those observed at t_3: ATT coefficients are positive and slightly lower in magnitude. Note also that the balancing properties are always verified, meaning that the control variables are not significantly different (at 1%) for individuals having similar propensity scores, between treated and untreated (BPNS stands for balancing properties not satisfied, i.e. zero means that all the covariates are balanced in all blocks defined in the propensity score computation).[25]

In order to address for the ATT differences between short and long term, it is plausible to argue, as noted in various other papers (Sianesi 2001a, 2001b, 2004), that individuals enrolled between t_1 and t_2 are involved in a lock-in effect in the short run, probably

[22] It is also worth noting that control variables to be included in the probit have to affect both outcome and treatment equations in order to correct the selection bias into treatment. See for instance Caliendo & Kopeining (2005) or Blundell et al. (2005).

[23] The common support region is actually very wide, from 0.04 to 0.44, and it represents only a slightly reduced interval with respect to the unrestricted variation of the propensity score (0.03 to 0.45). Accordingly, results computed without the common support restrictions are basically the same as the ones computed in table 2.

[24] Previous papers concerning PES evaluation in Italy (Barbieri et al. 2002, 2003) investigated only the short term outcomes, deriving either not significant or negative ATTs.

[25] It is also worth noting that individuals in the control group might display higher probability of being employed at t_2, under the hypothesis that skilled individuals find more quickly a job are also less likely to enrol in a PES. If this hypothesis held, our estimates should be considered as lower bound. We further discuss this issue in sections 8 and 9.

because they spend time on activities such us orientation periods, preparing CVs, training courses, apprenticeships, etc. In the long run, when these activities are over, the treated display higher probabilities of finding jobs. One might argue that the policies provided by PES are often characterized by short durations, in this way casting some doubts about the lock-in explanation. However, it is important to stress that also the short term evaluation takes place after a very short period (from one to 90 days) from the treatment. This might suggest that even if the policies provided by PES were of short durations, still the lock-in effect might plausibly apply.

Another important issue concerning policy evaluation in Italy regards the fact that there are relevant differences between regions, in particular between the South and the Centre-North. Sestito and Pirrone (2006) point out that the number of users of PES in the South is much higher than in the North, also because the unemployment rate is higher. Moreover, they claim that in the Centre-North the reform process has been introduced in a more efficient way. This is confirmed by the ISFOL index, which can be considered as a proxy of the quantity and quality of PES activities at the provincial level: provinces located in the Centre-North display –in average- higher values of this index. This geographical difference might be due either to the fact that the public administration is supposed to be better organized in the Centre-North region, or to the fact that since there are less unemployed per PES employee in this region, services can be supplied more efficiently. Our analysis confirms these conjectures, as shown in table 4. Even if results in the two regions are quite similar to the ones at the national level, in the South ATTs are more often not significant, and also smaller in magnitude, both in the short and the long run. This evidence seems to suggest that PESs are less effective in this region, while in the Centre-North estimates are mostly significant and larger in magnitude, entailing greater negative (positive) effects in the short (long) run.[26]

Recent papers, such as Blundell et al. (2004) and Crepon et al. (2005), have also introduced another interesting dimension to the analysis, investigating the efficacy of active labour market policies in increasing the probability of finding a good job, emphasizing the importance of the quality of a created match. Generally speaking, it could be argued that better matches should result in more productive and, then, longer lasting jobs (Crepon et al., 2005). Unfortunately, we cannot apply duration analysis since we can follow individuals over time only for a fixed period of time (four interviews in 15 months). Nevertheless, in Italy, as in other countries characterized by segmented labour markets, permanent contracts can be considered as a reliable proxy for the willing of the employer and the employee to invest in that match over time. On the contrary, bad matches are usually associated to temporary contracts, mainly because of the lower social

[26] Note that similar regional disparities, i.e. positive ATT in the Centre-North (in Toscany) and not significant ATT in the South (Sicily), have been observed by Ichino et al. (2008), in assessing the efficacy of private employment services in Italy.

security contributions. To address empirically this aspect, we investigate whether a treated in a PES displays a higher probability of finding a permanent job.[27] We use the same evaluation structure and the same treatment variable as in the previous analysis, while the outcome variable is equal to one if an individual finds a permanent job, either in the short or in the long term, and to zero in all other cases. Table 3 shows the results of the analysis. The ATT are still negative (even if lower, i.e. -3.4%), while they become not statistically different from zero after 12 months and slightly positive (3-4%) after 15 months. This means that the PES effects are less pronounced when considering permanent contracts as outcome variable, both in the short and in the long term (at 15 months), suggesting that at least part of the impact of PES observed in table 2 takes place through the use of temporary contracts.

As for the interpretation of our results, it is important to recall that the PESs can offer to the unemployed different kinds of programmes: counselling, training or intermediation, as already stated in section 2. This information in our data is unfortunately not available. From a multi-response question in the LFS questionnaire we only know the reasons for the last contact with PESs, which are reported in table 5. It comes out that 56.9% of the unemployed answers that one of the reasons for the last PES contact was to verify the existence of a job opportunity, and 2.6% because of a call related to a job offer. On the whole, 59.5% of individuals contacted a PES for its intermediation role. On the other hand, only 1.2% of individuals declare that the reason for the last contact with PES concerned training programmes and 19.3% regarded activities related to counselling. However, the shares related to training and counselling could be underestimated if some of the individuals who went to the PES to verify the existence of a job opportunity (56.9%) were waiting for outcomes related to previous training or counselling activities. Even if we cannot disentangle between these two possibilities because this information is not available, we can derive additional information from another question in the LFS questionnaire that investigates whether the individual has attended a training course, not necessarily through PESs, in the last month: only 2.4% of the treated in our sample are involved in training activities (and similarly 2.3% of the untreated). This additional evidence from LFS suggests that PES users in Italy are mostly recipients of counselling and intermediation activities (and less of training courses).

Our findings are then consistent with a number of European studies that have recently stressed the efficacy of intermediation and counselling programmes – an efficacy that had already been underlined by Martin and Grubb (2001). In particular, Blundell et al. (2004)

[27] In the Italian legislation it is straightforward to define a permanent job, since it can be univocally associated to the so called "Contratto a tempo indeterminato", i.e. contract without any limit of time. On the contrary, we define as temporary jobs all those jobs associated to both fixed term contracts included in the Italian legislation ("contratti a tempo determinato", "contratti di apprendistato") and to all forms of self-employment (among which "collaborazione a progetto", "collaborazioni coordinate e continuative").

provide evidence that in the UK the New Deal for Young People program entails an increase in the probability of finding a job of about 5%. Blundell et al. (2004) also claim that at least 1% of this positive impact is related to job search assistance, while the remaining component is related to job subsidies. Crepon et al. (2005) also show that in France the PARE program, which is mainly characterized by intensive job search assistance and counselling, increases the proportion of individuals that has found a job after one year by less than one percentage point, while it decreases the incidence of unemployment recurrence one year after a job is found by 6 percentage points. Weber and Hofer (2004) provide evidence also for Austria, showing that job search assistance programmes significantly decrease unemployment duration while the effect of training is positive.

To sum up, our results have to be considered as additional evidence in favour of intermediation, job search assistance and counselling policies. Moreover, it is worth pointing out that these kinds of policies are usually less expensive than wage subsidy schemes or training courses, implying that the overall social benefits might be even higher.

7. Robustness checks

In this section we apply some robustness checks to previous results. The first point to make here is that in both the short and long term analyses all the algorithms used to match treated and untreated (nearest neighbour, radius, kernel, and stratification) provide very similar results. This represents preliminary evidence of the robustness of the results.

Then, we focus on two robustness check exercises, to answer to two different questions.[28] First, the fact that some unemployed can, at the same time, resort to PESs, private employment services, and other training courses might at least partially drive our ATT coefficient. To address this point, we consider a slightly different evaluation structure, changing the definition of the initial group. So far we have taken into account individuals 'not enrolled' in a PES at t_1. As a check we eliminate from the initial group (both treated and untreated) the individuals benefiting in the previous six months (with respect to t_1) from private employment services or from training courses. The first part of table 6 shows that the results do not change much: ATT have the same signs both in the short and long run, and are slightly smaller in magnitude.

Second, one might argue that the positive ATT derived in the long run could be partially related to the composition of the control group, which also includes individuals treated between t_2 and t_3. If these individuals were involved in the above mentioned lock-

[28] We carried out also other robustness checks, which are available on request. For instance, we rule out from the analysis the search intensity variable, which is potentially endogenous. Results do not change.

in effect in the short term, this might affect our ATT estimates in the long run. In order to deal with this problem we exclude from the control group all the individuals treated between t_2 and t_3 that could be potentially affected by the lock-in effect at t_3. Similarly, for the long run evaluation at t_4 we exclude from the control group all the individuals treated between t_2 and t_4. The second part of table 6 shows the ATTs computed using this control group: also in this case results do not change much, and are slightly lower in magnitude with respect to the baseline setting.[29]

8. Sensitivity analysis for possible deviations from the CIA

One might argue that using propensity score based on selection on observables it is not possible to deal with selection on unobservables. To deal with this potential critic, we make use of the sensitivity procedure developed by Ichino, Mealli and Nannicini (2008) to assess the robustness of our ATT estimates due to possible deviations from the original setting of the Conditional Independence Assumption (CIA), the main untestable assumption of matching procedures.

The central hypothesis of this methodology is that the CIA does not hold in the original setting, $P(T=1 \mid Y(0), Y(1), X) \neq P(T=1 \mid X)$, while it is supposed to hold using an additional unobservable variable, a binary confounder U, entailing that $P(T=1 \mid Y(0), Y(1), X, U) = P(T=1 \mid X, U)$. Denoting with $Y = T * Y(1) + (1-T) * Y(0)$ the observed outcome of a given unit, it is possible to fully characterize the distribution of U by T and Y, i.e. by four parameters p_{ij} defined in the following way:

$$p_{ij} = P(U=1 \mid T=i, Y=j) = P(U=1 \mid T=i, Y=j, X), \qquad i, j \in \{0,1\},$$

which gives the probability that $U=1$ in each of the four groups defined by the binary treatment status T and the outcome value Y.[30] Once chosen a set of the four p_{ij}, U might be assigned in different ways to individuals in order to respect these p_{ij} constraints.[31]

To deal with this aspect, for a given set of the p_{ij} we carry out replications (200) computing different predictions of U to the individuals. For each of this prediction, U is introduced in the ATT computation, as any other covariate. Finally, the ATT is computed

[29] The exclusion of those treated between t_2 and t_3 (or between t_2 and t_4) potentially changes the composition of the control group. More specifically, it could be argued that those that get treated in these time spans are less skilled than the average individual, in this way reducing the level of skills in the control group. In this setting, our estimates might be considered as upper bounds.

[30] Note that, in order to make the simulation of the potential confounder feasible, two simplifying assumptions are made: 1) binary U, 2) conditional independence of U with respect to X. Ichino, Mealli and Nannicini (2008) present two Monte Carlo exercises showing that these simulation assumptions do not critically affect the results of the sensitivity analysis.

[31] In other words, the p_{ij} constraints only set the four frequencies of U in the cells defined by the binary values of T and Y. These four frequencies can be obtained with a very large set of predictions of U among the N individuals divided in the four cells.

as average of all replications for a given set of p_{ij}, using the preferred matching algorithm (radius, nearest neighbour, Kernel).[32]

The main question in this procedure is how to choose the p_{ij}, in order to simulate a 'meaningful' confounder. As pointed out by Ichino et al. (2008), the real threat to the baseline estimate comes from a potential confounder that has both a positive effect on the untreated outcome ($p_{01} - p_{00} > 0$) and on the selection into treatment ($p_{1\bullet} - p_{0\bullet} > 0$).[33] The presence of such a confounder, even without a true causal relationship between T and Y, could completely determine a positive ATT estimate. As a consequence, the sensitivity simulations should focus on confounders of this type. Ichino et al. (2008) analytically prove that $d = p_{01} - p_{00} > 0$ entails a positive impact on the untreated outcome, i.e. $\Pr(Y(0)=1 \mid T=0, U=1, X) > \Pr(Y(0)=1 \mid T=0, U=0, X)$, and that $s = p_{1\bullet} - p_{0\bullet} > 0$ produces a positive selection effect, i.e., $\Pr(T=1 \mid U=1, X) > \Pr(T=1 \mid U=0, X)$. In accordance with this framework, we focus our attention on two parameters d and s. However, these parameters cannot be considered as the effective impact of U on outcome and selection since in order to derive the effective impacts it is necessary to take into account the correlation in the data between U and the set of covariates X.[34] For this reason, estimating a logit model of $Pr(Y = 1 \mid T = 0, U, X)$ in every iteration, and hence controlling for the set of covariates X, Ichino et al. (2008) compute the effect of U on the relative probability to have a positive outcome in case of no treatment (the observed "outcome effect" of the simulated U) as the average estimated odds ratio of the variable U:

$$\Gamma \equiv \frac{\dfrac{P(Y=1 \mid T=0, U=1, X)}{P(Y=0 \mid T=0, U=1, X)}}{\dfrac{P(Y=1 \mid T=0, U=0, X)}{P(Y=0 \mid T=0, U=0, X)}},$$

while the selection effect, computed in a similar way, is:

$$\Lambda \equiv \frac{\dfrac{P(T=1 \mid U=1, X)}{P(T=0 \mid U=1, X)}}{\dfrac{P(T=1 \mid U=0, X)}{P(T=0 \mid U=0, X)}}.$$

[32] This method shares some intuitions with other sensitivity methods, such us Rosenbaum and Rubin (1983a) and Imbens (2003), with the main differences of not requiring any parametric assumptions for the outcome equation, and of focusing on point estimates of ATT.

[33] Note that $p_{i\bullet}$, i.e., the fraction of individuals with $U=1$ by treatment status only, is defined as $P_{i\bullet} = \sum_{j=0,1} p_{ij} * P(Y=J \mid T=i)$, where $P(Y=j \mid T=i)$ is the probability observed in the data of a given outcome j for a given treatment status i. Hence, by setting p_{11} and p_{10} appropriately, the assumption $p_{1\bullet} - p_{0\bullet} > 0$ can be imposed.

[34] Note the distribution of U given T and Y is not supposed to vary with X, as stressed in the definition of the p_{ij}. However, there is in the data an empirical association between the simulated U and X, coming indirectly from the association of X with T and Y. For more details see Ichino et al. (2008).

It is worth noting that when d and s are greater than zero the outcome and selection effects must be greater than one, meaning that d and s are positively related to Γ and Λ, respectively.

Ichino et al. (2008) basically propose two exercises to assess the robustness of the ATT estimates from possible deviations of the original CIA. First, it is possible to simulate a confounder having a distribution similar to other covariates in the data, in order to check how the introduction of this confounder would change the ATT (the so-called 'calibration confounder'). Second, the parameters p_{ij} might be properly chosen in order to look for the so-called 'killer confounder', defined as that confounder that -when introduced- would drive the ATT to zero. In this setting it is possible to assess whether the parameters that drive the ATT to zero are characterized by plausible outcome and selection effects.

Using the radius matching algorithm, we implement both exercises to assess whether the ATT computed in the long run might be partially related to some unobserved U.[35] As for the simulation of calibrated confounders, we use the distribution of some binary covariates that were significant in the propensity score estimate: search intensity, primary school, secondary school, adaptability to part time, adaptability to geographical mobility, having a family member enrolled in a PES, having been a low-skilled worker in the previous job. Table 7 summarises the results, reporting the different values of p_{ij} related to the chosen binary covariates, the ATT and the standard errors, and the outcome (Γ) and selection (Λ) effects as previously defined. It comes out that introducing confounders behaving as the chosen binary covariates only slightly alters the ATTs, which are always very close to the baseline value as well as standard errors, remaining always significant at 5%.[36] For instance, introducing a confounder distributed as the search intensity covariate entails an ATT of 0.84, which is basically the same as the baseline of 0.83, with identical standard errors. This represents clear evidence that for various configurations of the confounder U the ATTs do not change.

As for the killer confounder simulation, we let d and s vary from 0.1 to 0.5, in this way entailing increasing outcome and selection effects. As in Ichino et al. (2008), we relate the killer confounder to unobservable skills, which is in our opinion the main variable we cannot fully control for in the original specification.[37] We also claim that values of Γ and

[35] We report in this paper only the sensitivity analysis of the long term results (after 12 months). The sensitivity analyses related to the results of the short term and of the long term evaluation after 15 months are very similar from a qualitative point of view from the sensitivity analysis carried out after 12 months and confirm the baseline ATTs (available on request).

[36] Note that the standard errors are weighted averages of the *within* and *between* standard errors, as in Ichino et al. (2008). This choice leads to conservative inferential conclusions, since the average is always greater than the within and between components. Nevertheless, the ATT estimates we are interested in always prove to be significant. For details see Ichino et al. (2008) and Nannicini (2007).

[37] In order to carry out the killer confounder exercise we have to fix both the incidence of the killer confounder in the sample (as in Ichino et al., 2008, we choose $P(U=1)=0.4$) and the incidence of the confounder on the treated outcome (p_{11}-p_{10}=0). Since these parameters are not expected to represent a threat for the estimated baseline ATT, they can be held fixed and the simulated confounder U can be fully described

Λ greater than 4 have to be considered quite implausible, i.e. the presence of such confounder would increase the outcome and/or the selection probabilities by more than four times. This is also empirically confirmed by table 7, which shows that the outcome and selection effects are always lower than 2 for the chosen calibrated confounders. In table 8 we report the ATT computed for all the possible combinations of d and s, both ranging from 0.1 to 0.5. Moreover, for each combination of d and s we also display the related Γ and Λ. Table 8 shows that for values of d and s lower than 0.3, the outcome and selection effects (Γ and Λ) are lower than 4, the chosen threshold, and the associated ATTs are positive, significant and very close to the baseline estimate (0.083). This confirms the reliability of our ATT estimates due to possible deviations from the original setting of the CIA. Another point to bear in mind is that to drive the ATT to zero, the selection and the outcome effects have to be simultaneously close to 4 – a situation even more improbable.

9. Unobserved heterogeneity and sensitivity analysis

Sensitivity analysis allowed us to assess the robustness of the CIA using simulated confounders as possible deviations from the original CIA setting. Nevertheless, one might argue that even if our control variables group is very large, including much information at the individual, family and macroeconomic levels, there could be some unobserved variables still playing some role. In this last section we make use of the sensitivity analysis to understand what direction a possible unobserved heterogeneity bias might take and to simulate the related ATT using appropriate confounders.

The starting point of this section is the perception in the Italian public opinion and the previous literature (Barbieri et al. 2002, 2003) that PES clients are negatively selected, in the sense that those individuals with relatively worse observed and unobserved characteristics are more likely to be enrolled in a PES, while those with better characteristics carry out their job search activities not through PES, for instance sending curricula directly to employers, using informal networks, etc.

This means that if treated individuals were actually negatively selected, and if we were not able to control thoroughly for this unobserved heterogeneity neither using our wide range of covariates nor the sensitivity analysis, our estimates would represent lower bounds of the true ATTs, in both the short and the long run. This would entail that in the short run the effect might be non negative (either not significant or even positive), while in the long run the ATTs might be even greater in magnitude, i.e. more than the 8.3% using the radius algorithm.

---

by d and s. In this setting, the four parameters p_{ij} can be univocally determined. For further details see Ichino et al. (2008) and Nannicini (2007).

A preliminary test of this hypothesis can be achieved by computing the ATT in a 'thick support', as proposed by Black and Smith (2004). Under plausible assumptions, they argue that if unobserved heterogeneity is still playing some role in the selection into treatment, this bias is minimized when the analysis is restricted to the centre of the distribution of the propensity score, i.e., the thick support region, for instance from the 20th to the 80th percentiles. The underlying intuition is that if some unobserved selection is at work it will more markedly affect the tails of the distribution of the propensity score. Applying this procedure to our data for the long term evaluation (at t_3), and using radius ATT computation as baseline (0.083), we derive a significant ATT of 0.094 in the long run, suggesting that probably there is still some unobserved negative self-selection of individuals into the treatment.

We can actually go a bit further in the analysis. Using the 'killer confounder' procedure we can choose parameters p_{ij} in order to derive a confounder U characterized by the supposed effects due to unobserved heterogeneity, i.e. a negative selection effect and a null or positive outcome effect. Table 9 confirms that when introducing such kinds of confounders the ATT increases. More specifically, when s ranges from -0.1 to -0.3, entailing a negative selection effect, the ATT remains basically unchanged when no outcome effect is at work ($d = 0$) while when the outcome effect is positive ($d > 0$) the ATT increases. According to this simulation exercise, unobserved heterogeneity plays a role, entailing that our estimate might be interpreted as lower bound of the 'true' PES effect, in both the short and long run.

Interestingly enough, in this framework it is also possible to compute a sort of upper bound of ATT estimates. More specifically, using plausible values of Γ and Λ (no greater than 4 and no lower than 1/4, respectively) we observe that the ATTs range between the baseline estimate 8.3% and 10.8%, corresponding to values of s and d equal to -0.25 and 0.3, respectively. This range is also consistent with the ATT estimates computed in a thick support, as suggested by Black and Smith (2004). In other words, the simulation exercise seems to suggest that if our set of covariates did not fully control for unobserved heterogeneity, the ATT estimates would range in the interval [8.3%, 10.8%], thus confirming that in the long run PES services significantly increase the probability of being employed.

10. Conclusion

In accordance with the European Employment Strategy, most of the European countries have reformed their active labour market policies in the last decade. These reforms have hence generated a spread of evaluation exercises for many European countries, such as Blundell et al. (2004) for UK, Crepon et al. (2005) for France, Gerfin and Lechner (2002) for Switzerland, Sianesi (2004) for Sweden, Weber and Hofer (2004) for Austria. Using propensity score techniques, the aim of this paper is then to fill the gap in the literature concerning the Italian case, assessing the efficacy of Public Employment Services (PESs) after the reforms introduced in 1997, 2000 and 2003.

In line with other papers, such as Sianesi (2004), we show that computing both short and long term evaluations really matters in the interpretation of results. In particular, while in the short term the PES impact is negative in the long term the PES users display a higher probability of finding a job with respect to the untreated. We argue that the difference between short and long term results can be accounted for with a lock-in effect. In the short run the impact of the treatment might be negative because PES users are involved in activities such as preparing CVs, taking part in orientation periods, work-practice experiences, training courses, etc., in this way temporarily reducing the probability of finding a job; however, the probability increases in the long run, when these activities are over.

Our results also show that the PES effects are less pronounced (lower in magnitude and not always significant) when considering as outcome a proxy for the quality of the job, i.e. having found a permanent contracts, both in the short and in the long term. We also point out that geographical differences play a role, since in the Centre-North region ATT estimates are greater in magnitude while in the South ATT estimates are not always significant.

Since LFS data provide some evidence that PES users in Italy are mostly recipients of counselling and intermediation activities, our results can be considered as in favour of these kinds of policies, in line with other European evaluations, such as Blundell et al. (2004) for the UK, Crepon et al. (2005) for France and Weber and Hofer (2004) for Austria, which claim that job search assistance programmes produce positive effects on unemployment related outcomes.

Finally, the sensitivity analysis proposed by Ichino et al. (2008) confirms our ATT estimates, using simulated confounders as possible deviations from the original setting of the CIA. Besides, the sensitivity analysis allows us to compute an interval of variation for ATT point estimates, and to claim that our ATT estimates should be considered as lower bound of the 'true' effect of PES in the eventuality of negative unobserved selection into treatment.

References

Barbieri G., Gennari P., Sestito P. (2002), "Do Public Employment Services help people in finding a job? An evaluation of the Italian Case", *Rivista di Statistica Ufficiale*, (3).

Barbieri G., Gennari P., Linfante G., Rustichelli E., Sestito P. (2003), "Valutare i servizi pubblici per l'impiego: implementazione della riforma, attivismo dei servizi e chances lavorative degli utenti", *Politica Economica,* 3, pp. 343-372.

Black D., Smith J. (2004), "How Robust is the Evidence on the Effects of College Quality? Evidence from Matching. Journal of Econometrics", 121(1), pp. 99-124.

Blundell R., Costa Dias M., Meghir C., Van Reenen J. (2004), "Evaluating the employment impact of a mandatory job search program", *Journal of the European Economic Association*, 2, 569-606.

Blundell R., Dearden L., Sianesi B. (2005), Evaluating the effect of education on earnings: models, methods and results from the National Child Development Survey, *Journal of the Royal Statistical Society*, 168, pp. 473-512.

Caliendo M., Kopeinig, S. (2008), "Some Practical Guidance for the Implementation of Propensity Score Matching", *Journal of Economic Surveys*, 22(1), pp. 31-72.

Crepon B., Dejemeppe M., Gurgand M. (2005), "Counseling the unemployed: does it lower unemployment duration and recurrence?", IZA DP no.1796.

Dehejia R.H., Wahba S. (2002), "Propensity Score Matching Methods for Non-Experimental Causal Studies", *Review of Economics and Statistics*, 84(1), 151-161.

European Commission. Joint Employment Report. Bruxelles. 2004.

Fredriksson P., Johansson P. (2003), „Program evalutation and random program starts", IFAU working paper, n.1.

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

Heckman J.J., Ichimura H., Todd P. (1997), "Matching as an Econometric Evaluation Estimator", *Review of Economic Studies*, vol. 65(2), 261-94.

Ichino A., Becker S.O. (2002), "Estimation of average treatment effects based on propensity scores", *The Stata Journal*, Vol.2, 4,pp.358-377.

Ichino A., Mealli F., Nannicini T. (2008), "From Temporary Help Jobs to Permanent Employment: What Can We Learn from Matching Estimators and their Sensitivity?", *Journal of Applied Econometrics*, 23(3), pp.305-327.

Kok W (2003), "Jobs, Jobs, Jobs. Creating more employment in Europe", European Commission.

Imbens G.W. (2003), "Sensitivity to Exogeneity Assumptions in Program Evaluation", *AEA Papers and Proceedings*, 93(2), 126-132.

Istat (2006), *La rilevazione sulle forze di lavoro: contenuti, metodologie, organizzazione*, Metodi e Norme, No.32. Roma.

Lalive R., van Ours J., Zweimuller J. (2008), "The Impact of Activle Labour Market Programmes on the Duration of Unemployment in Switzerland", *Economic Journal*, 118, pp. 235-257.

Lechner M., Wunsch C. (2008), "What Did All the Money Do? On the General Ineffectiveness of Recent West German Labour Market Programmes", Kyklos, 61(1), 134-174, 2008.

Nannicini T. (2007), "A Simulation-Based Sensitivity Analysis for Matching Estimators", *Stata Journal*, 7(3), pp.334-350.

OECD (2007), *Employment Outlook*, France.

Pirrone S., Sestito P. (2006), *Disoccupati in Italia: Tra Stato, Regioni e cacciatori di teste*, Il Mulino. Bologna.

Rosenbaum P.R., 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., Rubin D.B. (1983), "Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome", *Journal of the Royal Statistical Society*, Series B, 45, pp.212-218.

Sianesi B. (2001a), "Differential effects of Swedish active labour market programmes for unemployed adults during the 1990s", IFS Discussion Paper, W01/25.

Sianesi B. (2001b), "Swedish Active Labour Market Programmes in the 1990s: Overall Effectiveness and Differential Performance", *Swedish Economic Policy Review*, Vol. 8 (2), pp. 133-169.

Sianesi B. (2004), "An evaluation of the Swedish system of active labour market programmes in the 1990s", *Review of Economics and Statistics*, Vol. 86(1), pp. 133-155.

Van den Berg G.J., Van der Klaauw B., Van Ours J.C. (2004), "Punitive sanctions and the transition rate from welfare to work", *Journal of Labor Economics,* 22, 211-241.

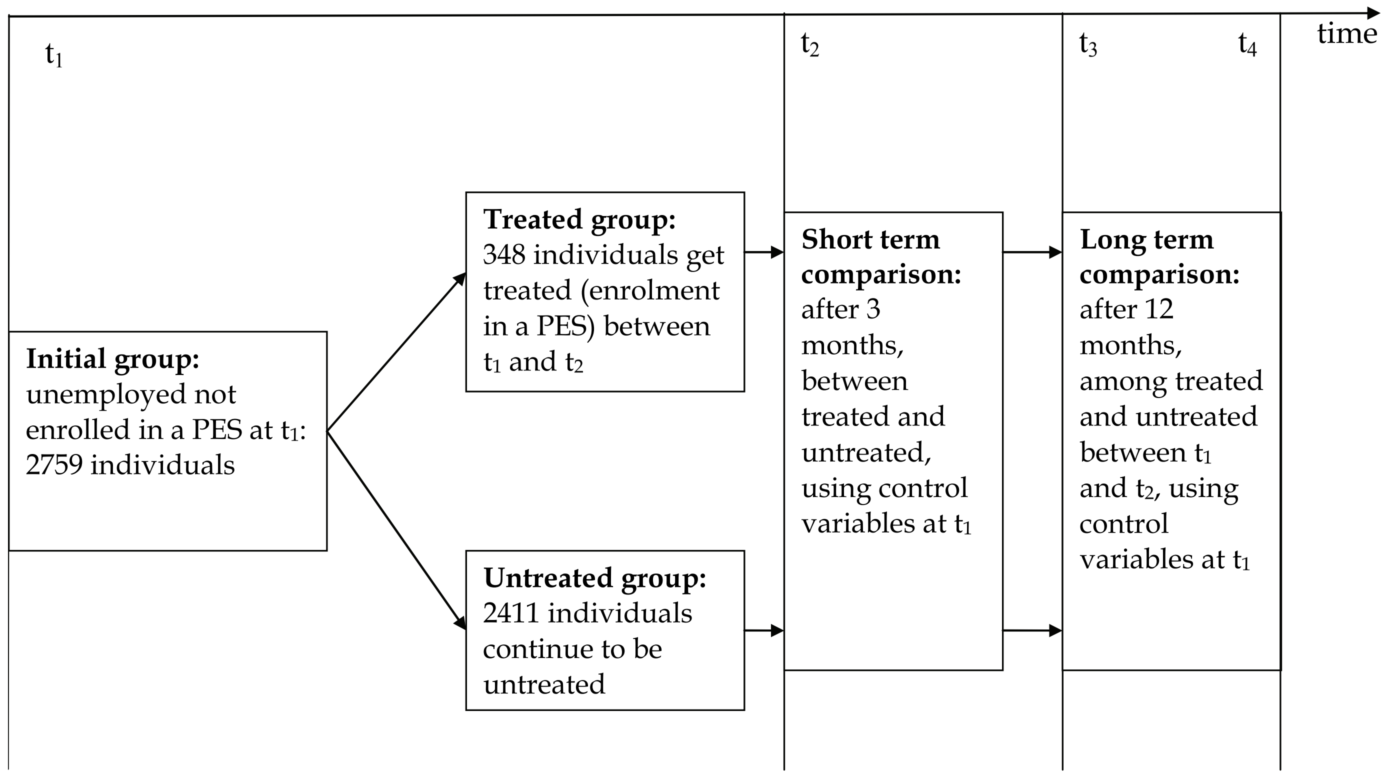
Weber A., Hofer H. (2004), "Are Job Search Programs a Promising Tool? A Microeconometric Evaluation for Austria", IZA DP, no. 1075.

FIGURES

Figure 1. LFS longitudinal dimension and rotation groups

	Quarter		Household rotation group					
			A	B	C	D	E	F
New RCFL	1	2004	A1					
	2	2004	A2	B1				
	3	2004		B2	C1			
	4	2004			C2	D1		
	1	2005	A3			D2	E1	
	2	2005	A4	B3			E2	F1
	3	2005		B4	C3			F2
	4	2005			C4	D3		
	1	2006				D4	E3	
	2	2006					E4	F3
	3	2006						F4

Figure 2. Scheme of the PES evaluation exercise. Treatment: enrolment in a PES.



Tables

Table 1. Probit estimates for the enrolment in a PES		
Covariates	Coeff	p-value
Age	0.036	0.056
Age squared	-0.001	0.040
No school - primary	-	-
Lower secondary	0.193	0.096
Upper secondary (liceo)	0.048	0.786
Upper secondary (no liceo)	0.279	0.025
Humanistic university degree	0.135	0.484
Scientific/giuridic/economic univ. degree	0.124	0.475
No search effort	-	-
Search intensity 1 (1/2 search actions)	0.146	0.089
Search intensity 2 (3/4 search actions)	0.191	0.052
Search intensity 3 (more than four)	0.401	0.002
Adaptability - part time	0.109	0.127
Adaptability - geogr. mobility	0.143	0.075
Unemployment duration (in months)	0.158	0.186
Having being dismissed in the previous job	0.141	0.137
With previous job experience	-	-
Previous experience: low skilled occupation	0.279	0.012
Previous experience: lower blue collar	0.230	0.048
Previous experience: higher blue collar	0.039	0.718
Previous experience: managers & white collar	0.069	0.526
Household members enrolled to PES	0.196	0.000
Turnover rates (NUTS-III level)	-0.016	0.037
Employment variations (NUTS-III level)	0.018	0.050
Agriculture Share (NUTS-III level)	0.014	0.112
Unemployment rate (NUTS-III level)	0.014	0.112
Quarter Dummies	yes	
Constant	-2.324	0.000

Note: All other variables (gender, potential experience, education lag, number of employed and dimension of the household, adaptability to fixed term, performance of PES), are not significant, and have been drop from the analysis.

Table 2: ATT of employment probabilites for unemployed not treated at t1									
Enrolment PES (between t1 and t2)	ATT - Propensity score matching								BPNS
	Radius		Nearest		Kernel*		Stratification		
	coeff	t	coeff	t	coeff	t	coeff	t	
Short Term (3 months)	**-0.074**	-3.80	**-0.072**	-2.46	**-0.077**	-5.41	**-0.080**	-4.39	0
Long Term (12 months)	**0.083**	2.97	0.046	1.61	**0.077**	3.14	**0.069**	2.63	0
Long Term (15 months)	**0.071**	2.59	0.029	0.76	**0.066**	2.40	**0.060**	2.40	0

348 treated. BPNS stands for balancing properties not satisfied. * is for bootstrapped standard errors.

Table 3: ATT of finding a permenent job for unemployed not treated at t1									
Enrolment PES (between t1 and t2)	ATT - Propensity score matching								BPNS
	Radius		Nearest		Kernel*		Stratification		
	coeff	t	coeff	t	coeff	t	coeff	t	
Short Term (3 months)	**-0.034**	-3.41	**-0.034**	-2.16	**-0.034**	-3.28	**-0.035**	-3.48	0
Long Term (12 months)	0.027	1.39	-0.009	-0.32	0.025	1.36	0.022	1.54	0
Long Term (15 months)	**0.038**	1.87	0.017	0.63	**0.037**	1.89	**0.034**	1.87	0

348 treated. BPNS stands for balancing properties not satisfied. * is for bootstrapped standard errors.

Table 4: Differences between the Centre-North and the South regions

Centre-North									
Enrolment PES (between t1 and t2)	ATT - Propensity score matching								BPNS
	Radius		Nearest		Kernel*		Stratification		
	coeff	t	coeff	t	coeff	t	coeff	t	
Short Term (3 months)	**-0.121**	-3.44	**-0.119**	-2.18	**-0.129**	-3.49	**-0.135**	-3.72	0
Long Term (12 months)	**0.111**	2.35	0.110	1.69	**0.108**	2.64	**0.101**	2.29	0
South									
Short Term (3 months)	**-0.037**	-1.59	**-0.090**	-2.41	-0.041	-1.82	-0.043	-1.80	0
Long Term (12 months)	**0.076**	2.25	0.072	1.57	**0.067**	2.09	0.060	1.38	0

126 treated for the Centre-North, 222 for the South. BPNS stands for balancing properties not satisfied. * is for bootstrapped standard errors

Table 5: Reasons of the last contact with the PES*	%
To verify the existence of a job opportunity	56.9
Because of a call related to a job-offer	2.6
To carry out vocational training	1.2
For activities related to counselling	19.3

* Multiresponse question. Only items related to job search activities are reported, and not the ones related to enrolment.

Table 6: Robustness checks

Initial group: unemployed not treated with PES, private empl.services and training at t_1

Enrolment PES (between t1 and t2)	ATT - Propensity score matching								BPNS
	Radius		Nearest		Kernel*		Stratification		
	coeff	t	coeff	t	coeff	t	coeff	t	
Short Term (3 months)	**-0.069**	-3.35	**-0.067**	-2.11	**-0.071**	-3.33	**-0.073**	-3.41	0
Long Term (12 months)	**0.068**	2.30	0.055	1.50	**0.062**	2.03	**0.057**	2.03	0
Long Term (15 months)	**0.084**	2.82	0.057	1.42	**0.078**	2.95	**0.072**	2.42	0
Initial group: unemployed not treated with PES at t_1, and removing from the control group those treated between t_2 and t_3									
Short Term (3 months)	**-0.077**	-3.89	**-0.083**	-2.68	**-0.082**	-4.36	**-0.087**	-4.10	0
Long Term (12 months)	**0.068**	2.41	**0.109**	2.87	**0.059**	2.07	0.052	1.54	0
Long Term (15 months)	**0.066**	2.38	0.066	1.74	**0.059**	2.25	**0.053**	1.87	0

297 treated for the upper panel, 248 for the bottom panel. BPNS stands for balancing properties not satisfied. * is for bootstrapped standard errors

Table 7 : 'Calibrated confounder' sensitivity analysis

	p_{11}	p_{10}	p_{01}	p_{00}	ATT	s.e.	Γ	Λ
Baseline	0.00	0.00	0.00	0.00	0.083	0.028	-	-
Confounder like:								
Search intensity	0.45	0.33	0.41	0.27	0.084	0.028	1.92	1.36
Primary school	0.28	0.45	0.31	0.40	0.082	0.029	0.67	1.10
Secondary school	0.41	0.32	0.36	0.28	0.080	0.028	1.48	1.28
Adapt. part-time	0.41	0.48	0.41	0.39	0.082	0.030	1.13	1.28
Adapt. geogr. mobility	0.21	0.25	0.21	0.17	0.083	0.028	1.26	1.35
If family members enrolled in PES	0.28	0.34	0.19	0.26	0.082	0.029	0.403	1.23
Previously low skilled	0.15	0.15	0.12	0.11	0.083	0.028	1.135	1.41

The matching algorithm is radius. 200 replications. To make the confounder variables binary we used the following classification: search intensity is equal to 1 if the individual has carried out more than 3 job search actions to look for a job; secondary school is specific for those who achieved a secondary degree not in a "liceo"; 'if household members enrolled in a PES' is equal to 1 is at least one person in the household is enrolled in a PES. Note that p_{11} refers to the probability that $U=1$ when Y and T are equal to 1, and similarly for p_{10}, p_{01}, p_{00}.

Table 8 : 'killer confounder' sensitivity analysis

	s=0.1		s=0.2		s=0.3		s=0.4		s=0.5	
	ATT		ATT		ATT		ATT		ATT	
	Γ	Λ	Γ	Λ	Γ	Λ	Γ	Λ	Γ	Λ
d=0.1	**0.083**		**0.083**		**0.08**		**0.075**		**0.064**	
	1.536	1.52	1.545	2.25	1.525	3.564	2.37	5.69	1.56	10.421
d=0.2	**0.083**		**0.081**		**0.075**		**0.064**		**0.041**	
	2.317	1.47	2.375	2.26	2.396	3.524	2.6	6.02	2.41	10.638
d=0.3	**0.082**		**0.08**		**0.071**		**0.054***		**0.021***	
	3.54	1.48	3.564	2.23	3.624	3.543	3.69	5.67	3.81	10.307
d=0.4	**0.08**		**0.078**		**0.066***		**0.043 ***		**0.000***	
	5.679	1.52	5.665	2.24	5.708	3.479	5.81	5.66	5.91	10.433
d=0.5	**0.08**		**0.076**		**0.061***		**0.032***		**-0.022***	
	9.288	1.45	9.395	2.24	9.434	3.492	9.59	5.73	9.87	10.355

* Not significant at 10%. Radius matching algorithm. 200 replications.

Table 9 : Unobserved heterogeneity and ATT range of variation

	s=- 0.1		s=- 0.2		s=- 0.25		s=- 0.3	
	ATT		ATT		ATT		ATT	
	Γ	Λ	Γ	Λ	Γ	Λ	Γ	Λ
d=0.0	**0.084**		**0.084**		**0.083**		**0.085**	
	1.017	0.653	1.018	0.392	1.014	0.292	1.508	0.641
d=0.1	**0.085**		**0.089**		**0.092**		**0.096**	
	1.508	0.641	1.499	0.393	1.524	0.286	1.531	0.224
d=0.2	**0.087**		**0.093**		**0.099**		**0.109**	
	2.327	0.644	2.304	0.386	2.266	0.287	2.27	0.197
d=0.3	**0.088**		**0.099**		**0.108**		**0.12**	
	3.539	0.646	3.488	0.375	3.518	0.277	3.489	0.197

All ATT are significant at 1%. Radius matching algorithm. 200 replications.

Département des Sciences Économiques
de l'Université catholique de Louvain
Institut de Recherches Économiques et Sociales

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique

ISSN 1379-244X D/2008/3082/030

APPENDIX

Table A1. Means of the observed characteristics by treatment status

Variables	Initial group	
	Untreated	Treated
Age	32.93	32.16
Gender *(0 Male, 1 Female)*	0.54	0.50
Educational levels (in dummies):		
No school - primary	0.13	0.10
Lower secondary	0.37	0.39
Upper secondary (liceo)	0.30	0.36
Upper secondary (no liceo)	0.07	0.05
Humanistic university degree	0.05	0.04
Scientific/giuridic/economic univ. degree	0.07	0.07
Education lag	0.91	0.82
Potential Experience	14.85	14.24
Unemployment duration *(in months)*	27.58	28.19
No search effort	0.26	0.20
Search intensity 1 (1/2 search actions)	0.43	0.42
Search intensity 2 (3/4 search actions)	0.23	0.26
Search intensity 3 (more than four)	0.08	0.12
Adaptability to fixed term contracts	0.87	0.89
Adaptability to part time contracts	0.40	0.45
Adaptability to geographical mobility	0.18	0.23
Having being dismissed in the previous job	0.14	0.19
Fixed term contract in the previous job	0.18	0.22
Previous job experiences	0.61	0.64
Previous experience: low skilled occupation	0.11	0.15
Previous experience: lower blue collar	0.11	0.14
Previous experience: higher blue collar	0.13	0.13
Previous experience: managers and white collar	0.13	0.14
Members of the household	3.61	3.70
Members of the household enrolled in a PES	0.31	0.42
Members of the household employed	0.84	0.82
Unemployment rate	11.30	11.88
Employment variation	0.47	0.92
Turnover rate	13.08	12.88
Agriculture rate	5.62	5.97

Classification of categorical variables. Education lag: 1) less than average; 2) in average; 3) more than average. Potential experience: difference between the current age and the age when the individual attained the highest educational level. Job search intensity (number of search actions in the last 4 weeks): 1) 0 search actions; 2) 1-2 search act.; 3) 3-4 search act; 4) more than 4 search actions. Occupation in the previous job: 1) low skilled; 2) blue collar; 3) high skilled blue collar; 4) employees, executives: these dummies are identified along with the dummy for having had job experience since there are few individuals with job experience but without a specification for the occupation (because the related job ended more than 7 years ago, information not collected in the LFS). Adaptability to fixed term or part time contracts, and to geographical mobility are equal to 1 if an individual declares he/she would accept these conditions in the new job.