Dipartimento di
Scienze Economiche
Università di Cassino

Working Paper

03/2010

Paolo Naticchioni[(1)] and Silvia Loriga[(2)]

[(1)] Univ. of Cassino, Univ. of Rome La Sapienza, Univ. Catholique de Louvain

[(2)] ISTAT

Short and long term evaluations of Public Employment Services in Italy

Dipartimento di Scienze Economiche
Università degli Studi di Cassino
Via S.Angelo Località Folcara, Cassino (FR)
Tel. +39 0776 2994734 Email dipse@eco.unicas.it

Short and long term evaluations of Public Employment Services in Italy

Paolo Naticchioni*

Univ. of Cassino, Univ. of Rome La Sapienza, Univ. Catholique de Louvain

Silvia Loriga**

ISTAT

Abstract

In the last decade the European Employment Strategy strongly recommended reforms of active labour market policies, reforms that have generated a spread of evaluation exercises for most of European countries. This paper fills the gap in the literature concerning the Italian case, assessing the efficacy of Public Employment Services (PESs) -after the reforms of 1997, 2000, 2003- in increasing the unemployment to employment transition probabilities, through matching techniques. Exploiting the longitudinal dimension of the Labour Force Survey 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. We also show that PES effects in the long term are much less pronounced when considering as outcome variable the probability of finding a permanent job, a proxy for the quality of the job, suggesting that PES impacts are to a large extent driven by the use of temporary contracts. Furthermore, to deal with issues related to selection on unobservables we carry out two different sensitivity analysis, which confirm our baseline findings.

Keywords: Public Employment Services, Active Labour Market Policies, European Employment Strategy, Matching, Policy Evaluation.

JEL codes: J64, J68.

* Corresponding author. *Dipartimento di Scienze Economiche, Via Sant'Angelo 04302 (FR) Italy, tel. +390776302307, fax.+390776302312*, p.naticchioni@gmail.com

** Istat, *via Ravà 150, 00142, Rome,* silvia.loriga@istat.it

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

The main aim of this paper is to enrich the set of the available literature needed to assess the efficacy of the European Employment Strategy, filling the gap concerning the Italian case. We evaluate the efficacy of Public Employment Services (PESs) after the reforms introduced in 1997, 2000 and 2003. Subsequent to these reforms, the Italian PESs are required to provide job search assistance, counselling, training schemes and job proposals (intermediation) to their clients.[2]

We use the Labour Force Survey data (LFS) 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).

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) for the UK, Gerfin and Lechner (2002) for Switzerland, and Sianesi (2004) for Sweden. 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

[1] 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.

[2] It is worth noting that in Italy passive labour market policies are not as widely developed as in other European countries, and are not actually interacted with active labour market policies. For this reason, our analysis disregards issues related to unemployment benefits. Further, according to the labour force survey the share of individuals receiving unemployment benefits is very low in the Italian labour market (less than 1% of the unemployed in our data).

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.[3] We claim that a more appropriate treatment variable can be defined by 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 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 ninety 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.

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 intention 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, using the same evaluation structure and treatment variable as in the previous analysis. The results show that in the short term ATT are still negative (-3.4%),

[3] This treatment variable definition has been used in Barbieri et al. (2002, 2003), which address the efficacy of PES in increasing employment probability before or during the reform process occurred in 1997, 2000, 2003.

while they become not statistically different from zero at t_3 and positive (3-4% and barely statistically different from zero) at t_4. This means that the PES effects are much less pronounced (lower in magnitude and not always significant) when considering permanent contracts as outcome variable, suggesting that the ATT computed after 12 and 15 months derived in the baseline evaluation were to a large extent driven by the use of temporary contracts.

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).

One might argue that our ATT estimates can be biased since propensity score matching cannot control for selection on unobservables. 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 to deal with endogeneity issues of the treatment variable are not available, as well as convincing thresholds for regression discontinuity designs. Nevertheless, we seriously take into account the issue of selection on unobservables carrying out two different sensitivity analysis procedures. First, 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. The sensitivity analysis confirms our findings. Furthermore, we implement another sensitivity analysis proposed by Black and Smith (2004), which also confirms our findings.

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 the methodologies and the 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 results of the sensitivity analysis. Section 9 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 process 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 case is concerned, it is worth noting 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, partially financed by the European Social Fund.[4] It is also worth noting that in Italy there are no other public policies and institutions supporting the unemployed in their job search, and this also explain why a high share of unemployed resorts to PESs, even without subsidies and other forms of pecuniary and non-pecuniary benefits.

Another important remark for the Italian case is that passive labour market policies are not as widely developed as in other European countries, and they are not actually interacted with active labour market policies.[5] In particular, eligibility for unemployment benefits is not conditional on the monitoring regarding participations into active labour market policies. The interaction between active and passive policies is basically 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 was very low in the

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

[5] According to OECD (2007), in 2004 the share of GDP devoted to passive labour market policies is 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.

years of analysis, 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 individuals enrolled in a PES that are not actually looking for a job using PES, but are enrolled only to receive the unemployment benefit.[6]

As for the reform process concerning the Italian PESs, a number of legislative acts 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.[7] According to the first goal, the reform process 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 strategic planning of services, while the everyday running of the PESs was in turn decentralized to the provinces (Province – NUTS3). For this reason, there may be some geographical differences, due to the different forms of PES organization chosen by local authorities.[8]

[6] Results basically do not change when including in the initial sample individuals receiving unemployment benefits.

[7] 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).

[8] 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.

Concretely, the new PES setting in Italy establishes that the PESs have to offer their clients one of the following three alternatives:[9]

- a personal counselling interview within three months from the unemployment declaration, in which the staff of the PES illustrates to the unemployed the possibilities to find a job in that province (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. Once an individual is enrolled, he has to accept the program established by the PES.[10] 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. This means we do not take into account in this paper the job search undertaken by inactive people as well as the assistance provided by PES to the on-the-job search activities carried out by the employed.

3. The Labour Force Survey data

We use the LFS data provided by Istat (the Italian National Institute of Statistics). This survey was completely overhauled in 2004, 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 in the questionnaire relating to the PESs has been thoroughly revised, enhancing the set of the collected information (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,

[9] 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.

[10] 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.

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. Since longitudinal data are not yet provided by ISTAT, the first step of the empirical analysis consisted in deriving longitudinal data for all rotation groups, from the first group (from the first quarter of 2004 to the second quarter of 2005) to the last group (from the third quarter of 2006 to the forth quarter 2007), i.e. six consecutive household rotation groups. We then observe individuals at four interviews, at t_1, t_2, t_3, t_4. 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 all the four interviews, instead of the two (t_1 and t_3) used by Barbieri et al. (2002, 2003) that evaluated the efficacy of Italian PESs before or during the reform process: in the first one we collect information on the control variables, 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 matching theory, the covariates are observed in a pre-treatment status, treatment then takes place, and finally the outcome is observed. This means that treated and untreated are compared using pre-treatment variables, matching individuals that were the same before the treatment took place.[11]

LFS also provides a very rich dataset, containing a wide set 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:

(i) a wide set of personal information, which are crucial to capture the individual heterogeneity that matters for selection into PES. The individual variables are: age, gender, education, education lag,[12] 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 fixed term contracts, adaptability

[11] A methodological problem related to Barbieri et al. (2002, 2003) is the fact that in these papers control variables and treatment variables were observed in the same instant of time (at t_1). Using this evaluation structure, Barbieri et al (2002, 2003) derived negative or no impacts on the probability of finding a job for the PES treated with respect to the untreated (after 12 months from the first LFS interview).

[12] 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.

to accept part time contracts, adaptability to accept jobs involving geographical mobility, unemployment duration;

(ii) household information, which we believe are very important in the selection process into PES since 63% of individuals in the sample are under 35 years old, suggesting that most of them might still rely on some help from the family. The household variables are: number of household members, number of members of the household enrolled in a PES excluding himself, number of members of the household employed;

(iii) a set of macroeconomic variables estimated using the LFS data at the provincial level (NUTS3 –103 provinces), which may play a relevant role in the selection process since in Italy there is a well-known dualism between the South and the North of the country, as well as a high NUTS3 (provinces) heterogeneity within the South and the North. The macroeconomic variables are: 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,[13] time (quarterly) dummies.

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 has already been used in previous papers (Barbieri et al. 2002, 2003). We doubt about the reliability of this treatment variable definition, since in Italy the enrolment in a PES is not closely related to the period of time in which the 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

[13] 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.

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 exploit the longitudinal dimension of the LFS database using the information concerning the enrolment in a PES between t_1 and t_2. Hence, our evaluation exercise 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.[14] We then follow these individuals in the period of time between t_1 and t_2, and we define as treated those who in the meantime have enrolled in a PES to look for a job.[15] 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.[16]

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 ninety 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 1 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.[17]

Descriptive statistics by treatment status and detailed explanations for the variables included in the set of covariates are reported in Table A1 in appendix: treated and untreated display very similar observed characteristics. The sample is mainly composed by

[14] Note that we consider as not enrolled in a PES 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 consider as not enrolled in a PES 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.

[15] We do not consider as treated those individuals that resorted to PES only to ask for generic information about the PES functioning, which are identifiable in the data.

[16] Formally since 2002 PES enrolment does no longer exist, and it has been replaced by the "formal declaration of unemployment status". The two concepts are basically the same and they are perceived in the same way by the unemployed. 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 'formal' unemployment declaration.

[17] 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.

young individuals (63% has less than 35 years old), and it is balanced between men and women. We also checked out that the difference in elapsed unemployment duration at t_1 between treated and untreated is not statistically different from zero, 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 (table 1). Furthermore, the high value of elapsed unemployment duration for treated and untreated (around 28 months) is not related to the sample selection of our group of analysis (unemployed not enrolled in a PES at t_1), since a similar mean unemployment duration (26 months) is computed for the whole sample of unemployed. It is instead explained by few outliers with very long spells of unemployment durations: the median unemployment duration is around 12 months, in line with other European countries. We will carry out a robustness check on this issue in section 7.

A last remark concerns the fact that, unfortunately, we cannot exactly identify the program a treated entered in (counselling or intermediation rather than training), 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. This issue will be explored in more depth in section 6.

5. Matching, identification, and sensitivity analysis

In this paper we use propensity score matching, whose basic assumption is selection on observables (unconfoundedness): selection into treatment is entirely determined by observed variable, 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 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.[18]

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 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).

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 (2001, 2004, 2008), 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.[19]

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.[20]

One might argue that using propensity score based on selection on observables it is not possible to address the issue of selection on unobservables. To deal with this potential critic, we make use of two sensitivity analysis methodologies. The first one has been proposed by Ichino, Mealli and Nannicini (2008) to assess the robustness of our ATT estimates due to possible deviations from the original setting of the CIA, the main untestable assumption of matching procedures. The central hypothesis of this sensitivity

---

[18] 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.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 Kopeinig (2005). See also Ichino and Becker (2002) for an explanation of the software we use.

[19] 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 (Sianesi, 2004). 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.

[20] 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.

methodology is that the CIA does not hold in the original setting, i.e $P(T=1|Y(0),Y(1),X) \neq P(T=1|X)$, since there is an unobservable variable excluded from the analysis. Ichino et al (2008) introduce an additional variable, a binary confounder U, and suppose that the new version of the CIA -including the confounder- holds: $P(T=1|Y(0),Y(1),X,U) = P(T=1|X,U)$. 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 for the binary values of T and Y by means of four parameters p_{ij} defined in the following way:

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

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.[21] Further, once having fixed the four p_{ij}, U can be assigned in different ways to individuals in order to respect these p_{ij} constraints: the p_{ij} constraints only set the four frequencies of U in the cells defined by the binary values of T and Y, and these four frequencies can be achieved with a very large set of different predictions of U among the N individuals divided in the four cells. 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 as average of all replications for a given set of p_{ij}, using the preferred matching algorithm (radius, nearest neighbour, Kernel).[22]

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 ($p_{1\bullet} - p_{0\bullet} > 0$).[23] The

[21] 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 assumptions do not critically affect the results of the sensitivity analysis.

[22] 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.

[23] Note that $p_{i\bullet}$, i.e., the share of individuals with $U=1$ by treatment status only, is defined as $p_{i\bullet} = \sum_{j=0,1} p_{ij} * P(Y=J|T=i)$, where $P(Y=j|T=i)$ 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.

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. $P(Y = 1 \mid T = 0, U = 1, X) > P(Y = 1 \mid T = 0, U = 0, X)$, and that $s = p_{1\cdot} - p_{0\cdot} > 0$ produces a positive selection effect, i.e., $P(T = 1 \mid U = 1, X) > P(T = 1 \mid U = 0, X)$. In accordance with this framework, we focus our attention on the two parameters d and s. However, these parameters cannot be considered as the effective impact of U on outcome and selection, which instead have to take into account the correlation in the data between U and the set of covariates X.[24] The effective impact of U on outcome (Γ) and selection (Λ) are then simply computed using logit models as follows:[25]

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

Ichino et al. (2008) basically propose two exercises to assess the robustness of the ATT estimates from possible deviations of the original CIA. First, the killer confounder, characterized by positive effects on selection and on the untreated outcome, where the parameters p_{ij} are explicitly set to 'kill' the ATT, to drive it to zero. Once having indentified the confounder that drives the ATT to zero, Ichino et al (2008) assess whether this confounder is characterized by plausible outcome (Γ) and selection (Λ) effects. Simply speaking, if the introduction of such a confounder increases the probability to be treated (or to be employed) by –for instance- four or five times, Ichino et al (2008) concludes that the ATT goes to zero only for unrealistic confounders. Second, it is possible to choose randomly a confounder, for instance assigning a distribution in the space (T,Y) similar to other covariates in the data. By doing so it is possible to check how the introduction of a randomly chosen confounder would change the ATT (the so-called 'calibration

[24] Note that 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).

[25] 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.

confounder'), i.e. whether the introduction of such a confounder strongly increase or decrease the ATT.

The second sensitivity analysis methodology to test the presence of unobserved heterogeneity consists in computing the ATT in a 'thick support', as proposed by Black and Smith (2004). Under plausible assumptions, Black and Smith (2004) 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. 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.

6. PES evaluation: estimates and results

Table 1 shows the propensity score estimates, using a probit. 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 that exploits the longitudinal dimension of the LFS data reduces the observed heterogeneity between treated and untreated.

Table 2 shows the ATT coefficients estimated in the common support region.[26] 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 the matching procedure used (radius, nearest neighbour, kernel, stratification).

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

[26] 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.

as the probability of an untreated to be employed at t_3). 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, and the value zero means that all the covariates are balanced in all blocks defined in the propensity score computation).

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 2001, 2004, 2008), that individuals enrolled between t_1 and t_2 are involved in a sort of lock-in effect in the short run, probably 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. To investigate what there is behind the lock-in effect, it is worth noting that even if the policies provided by PESs are often characterized by short durations (such as counselling and intermediation) the short term evaluation takes place after a very short period (from one to ninety days) from the enrolment. Furthermore, PES proposals do not necessarily have to occur just after the enrolment, since PES staff has a period of time of three months (six months) to propose to the unemployed a personal counselling interview (a short training course or work practice). This means that when the outcome for the short term evaluation is observed, the program proposed by PES could be just finished or still ongoing, or sometimes could be bound to start in the next future, and the lock-in effect might plausibly apply. Nonetheless, we claim that for all these reasons the long term evaluations are more reliable and interesting, and deserves more attention than the short term evaluation.

6.1. PES effects and regional differences

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 infrastructures 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 area, services can be supplied more efficiently. Our analysis confirms these conjectures, as shown in table 3. 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.[27]

6.2. PES effects and jobs quality

Recent papers, such as Blundell et al. (2004) and Crepon et al. (2005), have introduced another interesting dimension to the evaluation literature, 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 security contributions. To address empirically this aspect, we investigate whether a treated in a PES displays a higher probability of finding a permanent job.[28] We use the same evaluation

[27] 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.

[28] 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.

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 4 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%) and only barely significant in two out of four cases 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 a not negligible part of the impact of PES observed in table 2 takes place through the use of temporary contracts.

6.3. Interpretation of our findings

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 and 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) 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.

7. Robustness checks

In this section we carry out 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 three robustness check exercises, to answer to three different questions.

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 coefficients. 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-in effect in the short term, this might affect our ATT estimates in the long run. Hence, 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.

Third, it might be argued that the composition of our group of analysis in terms of unemployment durations might play a role in the analysis. More specifically, one might expect different PES effects for short term unemployed (unemployed duration lower or equal than 12 months) and long term unemployed (more than 12 months). This robustness check is also crucial since around 45% of individuals in our group of analysis is composed by long term unemployed. We then carried out two separate analyses for short and long term unemployed. Results are basically the same for the two groups and are in line with the baseline results of table 2, both in the short and in the long run, suggesting that the composition in terms of unemployment duration does not play a key role in our analysis.[29]

8. Selection on unobservables: two sensitivity analysis methodologies

In this section we carry out two different sensitivity analysis methodologies. The first one, proposed by Ichino et al. (2008), has been presented in section 5. Using the radius matching algorithm, we implement both killer and calibrated confounder exercises to assess whether the ATT computed in the long run might be partially related to some unobserved U.[30] 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

[29] Since results are very close to the baseline results of table 2, we do not report them in a table. They are available on request.

[30] We report in this paper only the sensitivity analysis of the long term results (after 12 months). The sensitivity analysis 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 ATT_S (available on request).

previous job. Table 7 summarises the results, reporting the 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%.[31] 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 an identical standard error. 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.[32] We also claim that values of Γ and 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 probability by more than four times. 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.

We then make use of the second sensitivity analysis methodology, computing the ATT in a thick support, as proposed by Black and Smith (2004). We define as thick support the interval from the 20th to the 80th percentiles of the propensity score distribution. Applying

[31] 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).

[32] 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 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).

this procedure to our data for the long term evaluation (at t_3), and using radius ATT computation, we derive a significant ATT of 0.094, which is slightly greater than the baseline ATT using radius (8.3%). Similar results are derived at t_4. The fact that the ATT is greater than in the original setting suggests the presence of a slight negative selection into PES in the tails of the distribution. Nonetheless, changes with respect to the baseline ATT are very small confirming that the propensity score matching technique used in this evaluation exercise is robust to the potential critic of selection on unobservables.

9. Conclusion

The aim of this paper is to enrich the set of the available literature concerning the evaluation of the European Employment Strategy, assessing the efficacy of Italian PES after the reforms introduced in 1997, 2000 and 2003, by means of matching techniques and LFS data.

In line with other papers, such as Sianesi (2004), we show that computing both short and long term evaluations 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.

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, suggesting that the baseline ATT are at least partially driven by the use of temporary contracts. 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. Also the sensitivity procedure of Black and Smith (2004) confirms our findings.

Acknowledgments: We thank Bart Cockx, Muriel Dejemeppe, Stefano Gagliarducci, Claudio Lucifora, Tommaso Nannicini, Enrico Rettore, 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.

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, in revision for the *Journal of European Economic Association*.

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, *American Economic Review, 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. (2001), 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.

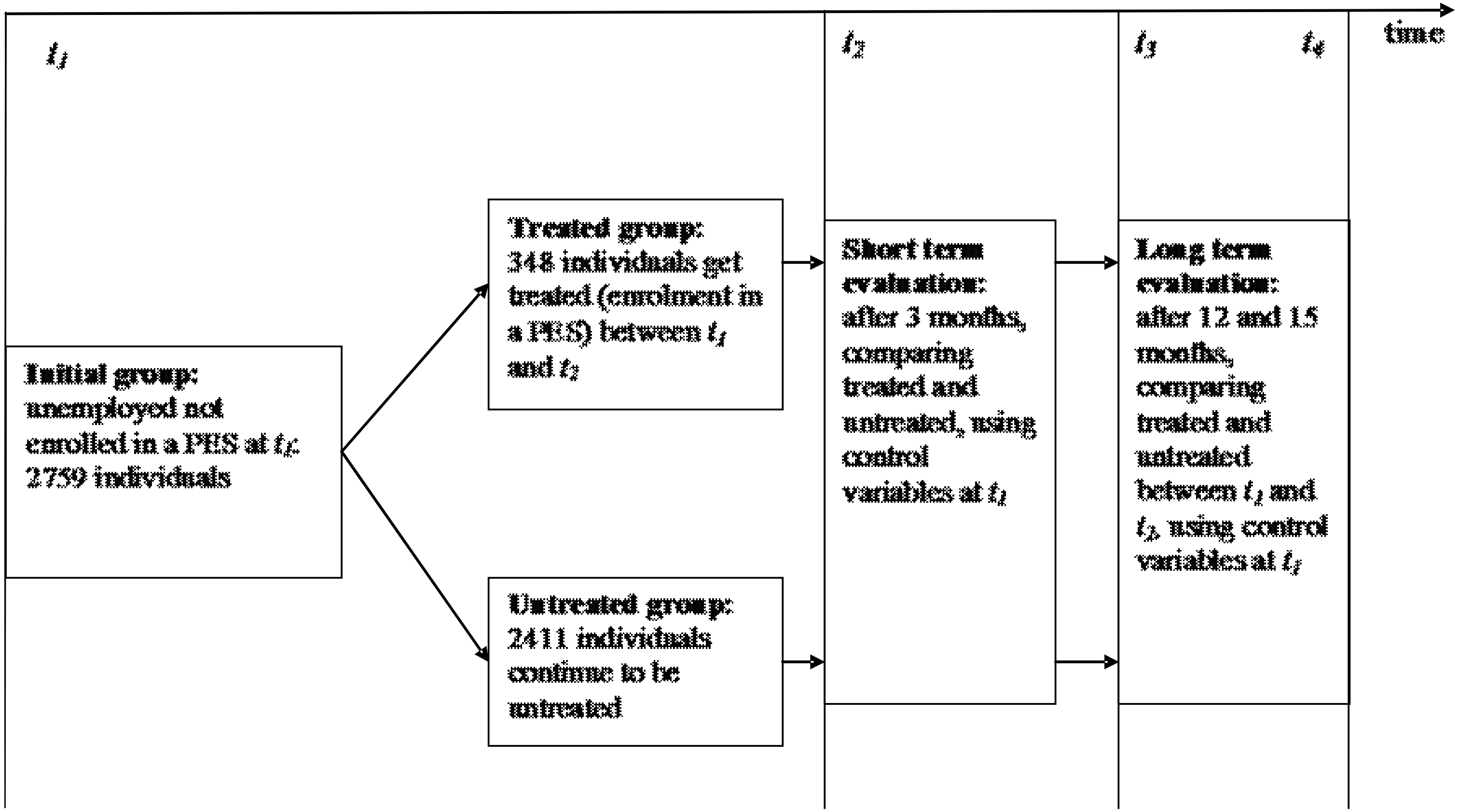
Sianesi B. (2008), Differential effects of active labour market programs for the unemployed, *Labour Economics*, Vol. 15, No. 3, pp 370-499.

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.

Figure

Figure 1. 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 and 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 largely not significant (p-value greater than 0.2) and have been drop from the analysis (for categorical variables we keep all dummies if at least one of them is statistically different from zero).

Table 2: ATT of employment probabilites: short and long term evaluations

Enrolment PES (between t1 and t2)	Num. treated	ATT - Propensity score matching								BPNS
		Radius		Nearest		Kernel*		Stratification		
		coeff	t	coeff	t	coeff	t	coeff	t	
Short Term (3 months)	348	**-0.074**	-3.80	**-0.072**	-2.46	**-0.077**	-5.41	**-0.080**	-4.39	0
Long Term (12 months)	348	**0.083**	2.97	0.046	1.61	**0.077**	3.14	**0.069**	2.63	0
Long Term (15 months)	348	**0.071**	2.59	0.029	0.76	**0.066**	2.40	**0.060**	2.40	0

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

Table 3: ATT and regional differences

Enrolment PES (between t1 and t2)	Num. treated	ATT - Propensity score matching								BPNS
		Radius		Nearest		Kernel*		Stratification		
		coeff	t	coeff	t	coeff	t	coeff	t	
Centre-North										
Short Term (3 months)	126	**-0.121**	-3.44	**-0.119**	-2.18	**-0.129**	-3.49	**-0.135**	-3.72	0
Long Term (12 months)	126	**0.111**	2.35	0.110	1.69	**0.108**	2.64	**0.101**	2.29	0
South										
Short Term (3 months)	222	-0.037	-1.59	**-0.090**	-2.41	-0.041	-1.72	-0.043	-1.70	0
Long Term (12 months)	222	**0.076**	2.25	0.072	1.57	**0.067**	2.09	0.060	1.38	0

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

Table 4: ATT of finding a permenent job: short and long term evaluations

Enrolment PES (between t1 and t2)	Num. treated	ATT - Propensity score matching								BPNS
		Radius		Nearest		Kernel*		Stratification		
		coeff	t	coeff	t	coeff	t	coeff	t	
Short Term (3 months)	348	**-0.034**	-3.41	**-0.034**	-2.16	**-0.034**	-3.28	**-0.035**	-3.48	0
Long Term (12 months)	348	0.027	1.39	-0.009	-0.32	0.025	1.36	0.022	1.54	0
Long Term (15 months)	348	**0.038**	1.87	0.017	0.63	**0.037**	1.89	0.034	1.67	0

BPNS stands for balancing properties not satisfied. * 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

Enrolment PES (between t1 and t2)	Num. treated	ATT - Propensity score matching								BPNS
		Radius		Nearest		Kernel*		Stratification		
		coeff	t	coeff	t	coeff	t	coeff	t	
Initial group: unemployed not treated with PES, private empl.services and training at t_1										
Short Term (3 months)	297	**-0.069**	-3.35	**-0.067**	-2.11	**-0.071**	-3.33	**-0.073**	-3.41	0
Long Term (12 months)	297	**0.068**	2.30	0.055	1.50	**0.062**	2.03	**0.057**	2.03	0
Long Term (15 months)	297	**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)	348	**-0.077**	-3.89	**-0.083**	-2.68	**-0.082**	-4.36	**-0.087**	-4.10	0
Long Term (12 months)	348	**0.068**	2.41	**0.109**	2.87	**0.059**	2.07	0.052	1.54	0
Long Term (15 months)	348	**0.066**	2.38	0.066	1.74	**0.059**	2.25	**0.053**	1.87	0

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

Table 7 : Calibrated confounder sensitivity analysis

	p_{11}	p_{10}	p_{01}	p_{00}	ATT	s.e.	Outcome effect (Γ)	Selection effect (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	1.916	1.358
Primary school	0.28	0.45	0.31	0.40	0.082	0.029	0.672	1.101
Secondary school	0.41	0.32	0.36	0.28	0.080	0.028	1.478	1.275
Adapt. part-time	0.41	0.48	0.41	0.39	0.082	0.030	1.127	1.279
Adapt. geogr. mobility	0.21	0.25	0.21	0.17	0.083	0.028	1.260	1.349
If family members enrolled in PES	0.28	0.34	0.19	0.26	0.082	0.029	0.403	1.232
Previously low skilled	0.15	0.15	0.12	0.11	0.083	0.028	1.135	1.414

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**	
	1.54	1.52	1.55	2.25	1.53	3.56	2.37	5.69	1.56	10.42
d=0.2	**0.083**		**0.081**		**0.075**		**0.064**		**0.041**	
	2.32	1.47	2.38	2.26	2.40	3.52	2.60	6.02	2.41	10.64
d=0.3	**0.082**		**0.08**		**0.071**		**0.054***		**0.021***	
	3.54	1.48	3.56	2.23	3.62	3.54	3.69	5.67	3.81	10.31
d=0.4	**0.08**		**0.078**		**0.066***		**0.043 ***		**0.000***	
	5.68	1.52	5.67	2.24	5.71	3.48	5.81	5.66	5.91	10.43
d=0.5	**0.08**		**0.076**		**0.061***		**0.032***		**-0.022***	
	9.29	1.45	9.40	2.24	9.43	3.49	9.59	5.73	9.87	10.36

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

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