Workplace Training Programs: Instruments for Human Capital Improvements or Screening Devices?

Abstract

This article analyses the effect of an Italian training program on the re-employment probability of young unemployed workers. The program consists exclusively of workplace training and is coordinated by employment centres, even if it is fully implemented by firms.

We develop a discrete duration analysis. In particular, we compare the re-employment process of individuals that just finished their workplace training program with individuals that just ended their job. We specifically take into account the issue of self-selection adopting the Propensity Score Matching estimation.

Our results suggest that this workplace training program improves only the immediate re-employability of trained workers, failing to bestow them with durable human capital improvements. These results appear to be robust to spurious duration dependence and to self-selection. Our analysis focuses on unobserved heterogeneity and, accounting for it, we show that the training implementation is useful to divide “good” trainees (in terms of unobserved heterogeneity) from “bad” ones. Therefore, we suggest that firms are exploiting training as a screening device and that the implemented program is successful in easing the connection between workers and firms, but it fails to provide a durable improvement in skills and in re-employment prospects.

The evaluation of this program is important because it focus specifically on the workplace component of training, whereas previous analyses focused on generic training, because it evaluates a program targeting youth unemployment which is one the most urgent economic issues and because it helps in understanding the actual processes adopted by firms when implementing workplace training.

Keywords: duration model; policy evaluation; propensity score matching; screening device; workplace training; youth unemployment.

JEL Classification: C41; I38; J64; J68; M53

1. Introduction

Active Labour Market Policies (ALMPs) are measures to improve the situation of the unemployed and of the disadvantaged in terms of employability and wages. This important role of ALMPs justifies the increasing interest in these measures during the last 20 years. The OECD provides a classification of labour market policies, distinguishing active measures from passive ones. The OECD classification considers seven categories of ALMPs: public employment services, labour market training (training for unemployed and employed adults), youth employment and training measures (such as apprenticeship), subsidized employment, employment programs for the disabled, job rotation and job sharing, and direct job creation (see OECD, 2007 and Chapter 14 in Cahuc *et al.*, 2014 for more details on these categories). Each labour market policy has a precise goal and it affects employment in a different way: for example, employment services have the goal of reducing job search costs while training programs and other measures in favour of youth aim to increase the probability to find a job and to rise productivity of this group of individuals. Other policies have the objective of reducing the cost of labour or creating directly public-sector jobs.

Given their important role in the labour market, it is important to assess whether these measures are really effective, in other words, whether they have been able to produce the desired effects. The evaluation of the effectiveness of ALMPs has become a central element in many countries; policy makers and administrations, both in the United States and in Europe, have increasingly become interested in programs effectiveness and efficiency and the evaluation of these aspects has become essential.

In this paper we try to assess the effect of a specific ALMP implemented in Italy, consisting in workplace training for young unemployed. The paper intends to assess the causal effect of this specific training program on the probability of employment after a certain period from the treatment. Broadly speaking, training programs aim to enhance the productivity and employability of the participants but also the human capital by increasing their skills. In general, training programs usually encompass measures like classroom training, workplace training and work experience. These measures can provide general education, specific vocational skills or even firm-specific skills. In addition, although this is not their main objective, training programs can also act as a mean to connect workers with firms and, in some cases, they can also be used by firms as a screening device or, when misused, even as a source of cheap labour (see, e.g., Barnow, 2000 and Barnow and Smith, 2004) .

In the specific case we are observing and evaluating, the program consists solely on workplace training and work experience for young unemployed workers and it is activated and coordinated by Employment Centres but implemented directly by firms. This training program is ruled by Italian laws but some of its specific aspects are overseen and determined by regions, Tuscany in our case.

The evaluation of this kind of program is important for several reasons. First, it is relevant because it focuses specifically on the workplace component of training, something that has not been deeply examined or evaluated in previous analyses on generic training. Second, we evaluate a program targeting youth unemployment which is currently considered one of the most urgent economic issues. Finally, our analysis also helps to understand, from a sociological and economic perspective, the actual processes adopted by firms when implementing workplace training.

The paper is organized as follows: Section 2 discusses the key literature on ALMPs and its evaluation; Section 3 presents the methodology used in general policy evaluation analyses and that adopted in this paper; Section 4 briefly presents the Italian workplace training program; Section 5 describes the data and present the empirical analysis with its results and, Section 6 concludes.

2. A review of the literature on ALMPs and their evaluation

The relevance of the unemployment phenomenon and the economic and social problems it causes have inevitably put the light on ALMPs and on their potential effect to reduce unemployment and to help vulnerable groups to find a job. In the last thirty years there has been a great emphasis on these measures and early studies on labour market policies have stressed the importance of ALMPs in fighting structural unemployment (see Layard *et al.,* 1991 and OECD, 1994). The general conclusion of these early studies and policy analyses has been the need to shift the focus of labour market policies from the passive provision of income support to more active measures that can assist and facilitate re-employment as passive measures have strong negative effects on the length of unemployment and on job search intensity. Active labour measures, on the other hand, can improve the matching of the demand and the supply of labour and reduce the long-term unemployment of disadvantaged workers.

Given the importance of ALMPs it has become crucial to evaluate whether the ALMPs are effective, that is if they produce the expected effects on the target population. The results of such evaluations can be used to monitor and inform the implementation process for example, they can suggest changes in implementation or targeting methods. A substantial number of evaluations of ALMPs effectiveness has been conducted: an overview of some early studies on this subject is contained in Heckman *et al*. (1999), that focuses on Europe and United States evaluations programs before 1994, and in Kluve and Schmidt (2002), that focuses on evaluation studies programs up to 1999 in Europe. Other surveys are in Martin (2000) and Martin and Grubb (2001) and give a descriptive account of OECD countries' experience with active labour market measures.

All these surveys stress how different programs display different effectiveness and in some cases they even highlight that these programs lead unemployed individuals with skills to not increase their employability (see Heckman *et al.,* 1999) while other contributions suggest that considerable lock-in effect of the programs may even harm the more employable individuals in a group of unemployed people (see Frolich and Lechner, 2010 and Wunsch and Lechner, 2008).

More recent empirical evaluations on ALMPs still find a great variety in terms of effectiveness and efficiency of the programs. Recent meta-analyses on the effectiveness of several forms of ALMPs at the microeconomic level are present in Card *et al.* (2010) and Kluve (2010) and their conclusions seem to suggest that some methods are strictly more efficient than others; subsidized public sector employment programs usually have the least favourable effect and job search assistance programs are quite favourable in the short run while training may have some positive effect in the medium and long run. A recent overall assessment of ALMPs at the macroeconomic level is contained in Martin (2015): its results suggest that activation programs helped the unemployed to get off benefits and into work but that the activation strategies effectiveness varies a lot across countries, further indicating the need to understand which are the effective programs and strategies.

Previous literature and analyses results are so heterogeneous that they can appear contradictory, but this may depend on the great difference among time periods, labour markets and specific program characteristics taken into exam. This heterogeneity of results also stresses the importance to search for specific effective programs and to examine whether and how they are relevant and replicable.

Given these premises, our analysis focuses on training programs and, from this point of view, it must be stressed that these programs are the most widely used active labour market measure: they account for the largest share of expenditures on ALMP across Europe with almost 40% (see Chapter 14 in Cahuc *et al.*, 2014). In particular, several studies investigate the impact of training programs directed at unemployed workers or disadvantaged groups both in the US and Europe, tackling different aspects of training, such as classroom training, on-the-job training or training in job search.

Earlier evaluations of training programs suggest that training hardly has an effect on earnings or employment probabilities of program beneficiaries compared to their counterparts. In most cases, participants fail to do better than the control group both in their post-training employment probability and in their earnings (see Fay, 1996). A more mixed picture arises from the evaluation of the Canadian Job Entry Program: a quasi-experimental evaluation of the program contained in

OECD (1993) shows that young people/teenagers who only undertake classroom training fail to do better than their counterparts but those who undertake enterprise training do significantly better than the control group: however, this positive effect is attributed to the trainees staying-on with the training firms. Bonnal *et al.* (1997) find that workplace training in France increases the transition from unemployment to employment but only for some educational groups, while Brodaty *et al.* (2001), using again evidence from France, find that workplace training is more beneficial for re-employment than some types of classroom training provided by public institutions. Arellano (2005) assesses the effect of four types of training courses on the duration of unemployment in Spain and finds that, apart from the general one, all training programs exhibit significant positive effects on re-employment. More recently, Lechner *et al.* (2011) evaluate the effect of training programs at different time horizon in West Germany during the nineties: using matching methods they find that long training courses (that usually last at least one year and often about two years) seem to have a negative effect on employment in the short run but start to have positive effect in the long run (usually after four years). Similarly, Crepon *et al.* (2012) find that training programs in France have little impact on the exit rate from unemployment in the short run but, after controlling for both observed and unobserved heterogeneity, they have a favourable effect on the duration of the subsequent employment spells. This also confirms the results obtained previously by Winter-Ebmer and Zweimuller (2003) in Austria for men. The role of classroom training programs are also analysed in Osikumino (2013) and in Richardson and Van Den Berg (2013): their studies detect a positive effect on re-employment but also stress that this effect declines over time. In one analysis made for Ireland, McGuinness *et al.* (2014) find that, in general, those who participates in training courses are less likely to be unemployed at the end of the course but this positive effect is present only for skill-intensive training.

As for the Italian case, only few training evaluations exist: a rare example is in Caroleo and Pastore (2001) that focus on long-term unemployed and find that workplace training and participation to ALMPs in Italy does not significantly improve the employability of these workers.

3. Evaluation analysis and its methodology

The aim of our analysis is to assess the effect of a specific form of ALMPs: in particular, we want to evaluate the effect of on-the-job training on the re-employability of young persons who received such training.

From a methodological point of view, the evaluation of the efficacy and success of ALMPs should focus on its causal effect, defined as the difference between the outcome of the units affected by the policy (the actual situation) and the outcome that these *same* units would have experienced if the policy had not been implemented. The fundamental evaluation problem is that it is not possible to observe simultaneously the same unit in the two scenarios, i.e. the scenario in which the policy is not implemented (the *counterfactual*) and the scenario in which the policy is implemented. Therefore it is needed an adequate control group that is similar as possible to the affected one. To find such a *control* group is not easy and there are mainly two different approaches: i) comparing the outcome of interest of the affected units before and after the intervention; and ii) comparing units affected by the intervention with those not affected. Detailed discussions on these strategies to estimate these causal effects are contained in Angrist and Krueger (1999) and in Heckman *et al*. (1999) while an overview of more recent developments in these methods is contained in Fitzenberger *et al.* (2014). In addition, for a discussion on the pro and cons of the two approaches see Loi and Rodrigues (2012).

Our analysis uses the second approach, comparing individuals affected and not affected by the training. In fact, our sample is made of individuals aged between 18 and 30 years that were registered at the Employment Centre of the Province of Pisa and it comprises both individuals that underwent the workplace training program (what we call the "treated") and individuals that did not underwent it (the "non-treated"). In particular, we compare the re-employment process of individuals that just finished their workplace training program with individuals that just ended their job.

In addition, empirical analyses that assess the causal effect of a treatment may suffer from two further problems: *spurious duration dependence* in the survival analysis and *self-selection into treatment*. While survival analysis is usually able to deal with duration dependence, in some cases specific problems may arise and spurious dependence may emerge. In particular, the presence of unobservable heterogeneity and time-varying effect of the treatment could complicate things and we deal with it specifically including a component of unobserved heterogeneity (see Chapter 5 in Cahuc *et al.*, 2015). As for self-selection into treatment, the problem is that it is not easy to distinguish whether a given observed effect is due to the participation to the program or to have been selected to participate to the program. In general, it is possible that the characteristics that induce individuals to be selected into the program also have direct effect on the re-employment probability. There are several methods to overcome the self-selection problem: in our analysis we adopt the propensity score matching (PSM) methodology (for a review of PSM methodology and its applications for policy evaluations see Caliendo and Kopeing, 2008, for a broader overview of the self-selection problems see Chapter 14 of Cahuc *et al.*, 2014).

In the specific case we are assessing, the descriptive data (presented in the next section) on re-employment for treated and non-treated individuals would suggest that the treatment has a positive effect immediately after the end of the training program but this effect quickly dissolves and even reverts to negative during later stages of unemployment. Starting from this simple descriptive evidence we perform an econometric analysis and we account for spurious duration dependence and self-selection into treatment: our results show that the causal effect of training is to all extent initially positive and becomes negligible (but not negative) after few months from the end of the training.

A decline over time of the treatment effect of a generic (classroom) training program is also found in Osikumino (2013) and in Richardson and Van Den Berg (2013). The latter pays special attention to this decline and concludes that, for classroom training, the decline does not concern actually the causal effect of training on the workers' skills but rather is related to: i) the interaction between the treatment effect and the unobservable heterogeneity (so that training is more profitable in terms of skills acquisition for workers with more favourable unobserved characteristics); ii) the boost in the job assistance at the termination of the training. Our conclusions are in part similar to Richardson and Van Den Berg (2013), as we argue that workers with more favourable unobserved characteristics (following Richardson and Van Den Berg, 2013, we refer to these kind of workers as "good" workers, that is workers whose unobserved characteristics make them more likely to be hired) are more likely to increase re-employment as an effect of having participated to training, and in part different, as we suggest that this is due to screening mechanisms inherent in the workplace training program rather than from an increase in the actual workers skills. From a sociological and economic point of view, our results also shed light on how the firms actually use the training programs.

It is worth to note that the presence of a similar screening mechanism was detected in analyses that tried to assess the role of fixed-term contracts in future employability (see Baranowska *et al.*, 2011): from this point of view firms appear to exploit workplace training and fixed-term positions to obtain similar objectives. Instead, there has been little investigation on whether firms use training as screening devices though Barnow (2000) and Barnow and Smith (2004) suggest this behaviour as a possible consequences of the Job Training Partnership Act of 1982 is the United States.

4. The Workplace Training Program

While Italian law contains some guidelines on internships and training (contained in Art. 10, legge 196/97) the actual implementation and regulation of the programs are left to the regional-level governments. We focus on the implementation of this program done in Tuscany: according to this region quality charter on internships and workplace training, “the workplace training program is a measure that aims to create a direct link between a job-seeker and a firm, and to allow the trainees to gain more experience in order to upgrade their curriculum and to facilitate a future possible work relationship with the host firm. It consists in a period of professional training and work counselling that allows young individuals to be close to the production sector”. Therefore, the program aims to accelerate the matching process between demand and supply in the labour market.

The training program duration ranges from 2 months to 6, and it can be extended to 1 year only for specific categories. In our analysis we consider only those training programs with a maximum duration of 6 months.

The workplace training program is associated to work experience and it is reserved to young people, under 30 years of age, who have completed their compulsory schooling. It envisages an agreement between an Employment Centre, that promotes the training, and the host (public or private) firm that provides a training project. The official agreement between the participant and the host firm is not an employment contract and hence no remuneration is compulsory, but the host has to insure the participant against workplace injuries. For each agreement, the Employment Centre appoints an in-firm tutor whose duties consist in following the workplace activities of the trainee.

Several specifications qualify the position of the participants to the program and try to avoid its misuses from the firm. In particular, it is specifically ruled that: 1) the trainee cannot be assigned to activities that do not necessitate actual training, 2) the trainee cannot be assigned to strictly seasonal activities nor can fill in for employees currently on leave, 3) all the activities of the trainee within the firm must be related to the training objectives of the agreement, 4) firms cannot sign more than an agreement with each single trainee. In addition, to avoid that the participant to the workplace training will be considered as an unpaid worker, some further guarantees are entailed within the agreement: pre-determination of promoters, maximum duration of the training and transmission of the agreement to public authority (Regions and Direzioni Provinciale del Lavoro). Finally, workplace training regulation strictly distinguishes it from apprenticeship since the latter is a contract of employment.

5. The Empirical Analysis and its Results

We perform now an empirical analysis on the determinants of re-employment focusing on the effect of the workplace training program we described in previous section.

In our analysis we observe individuals for a total of 8 months after they started searching for a job. We discretize duration of unemployment in blocks of 2 months, obtaining thus a maximum duration of 4 time periods. We choose to use discrete duration because it allows to take into account that the effect of a given variable (workplace training in our case) can differ at different points in time. While continuous models can deal with time-varying effects, their estimation is mostly restricted to those cases in which the effect varies continuously and smoothly with time, failing to capture discontinuity or time threshold. At any rate, we also perform continuous duration regressions as a robustness check for our analysis.

5.1 Data

The dataset used comes from the Employment Centre of the Province of Pisa database. We consider the totality of individuals, aged between 18 and 30 years, enrolled at the Employment Centre: some of them underwent the workplace training program while others did not.

In our evaluation, the treatment group is composed by those individuals that terminated the training program during the last nine months of year 2012, whereas the non-treatment group is composed by individuals that became unemployed during the same period of year 2012. Overall we have a sample of 4087 observations.

The variables used in the analysis are: gender, age, age squared, a dummy for the participation to the workplace training program (TRAINED), education represented by dummies for vocational secondary school (VSS, implying only three years of vocational school), upper secondary school (USS), university degree (UD) and using compulsory education as the reference category. We include also a dummy indicating whether the individual previous training/job ended during the third or fourth quarter of 2012 (Q3 and Q4 respectively and we use the second quarter as the reference category).

5.2 ***Descriptive Analysis***

We start our analysis simply presenting the share of individuals that are able to find a job within a certain amount of time, distinguishing the trained from the non-trained. In particular, we focus on time blocks of 2 months so that we report data for 4 blocks of time. The first block reports the share of individuals that found an employment within two months from the end of their previous working experience (either on the job-training or other forms of work), the second block reports the share that found employment within four months conditional to be still employed after two months and so on.

Table 1 reports the share for each time block and it highlights some clear patterns. First, the share of individuals that find employment is decreasing in the duration of unemployment (with 29.29% in the first block and 11.55% in the last block). Second, in the first period, the share of individuals that find an employment is much higher in the trained group (41.15% for the trained versus 28.70% for the non-trained). Third, during later periods, this latter pattern reverses and the non-trained group shows higher re-employment rates than the trained group: this is evident starting from the second period and becomes particularly relevant in the fourth one.

TABLE 1 here

Clearly this evidence is merely descriptive and it does not necessarily imply a causal effect of time or training on re-employment probabilities. However, the reversal of the training effect is particularly interesting and can be further investigated performing a complementary log-log regression where the binary outcome represents the event of finding a job.

The complementary log-log regression is similar to the logit regression but assumes a complementary log-log distribution for the errors and it thus uses a different link function (the function that transforms the actual outcome to its estimated value): in particular, it uses the inverse of the generalized extreme value cumulative distribution function. These assumptions allow to obtain better estimates when one of the possible outcomes (finding a job in our case) is observed to be consistently less likely than the alternative (see Long, 1997).

To emphasize the differences within the different periods we perform the regression for the first period (with all the sample) and for the fourth period (with only individuals that were still unemployed after six months). The results are reported in Table 2 and they represents the effect of the covariates on re-employment probabilities.

TABLE 2 here

As it is apparent from the above regressions, the effect of the training appears to revert from positive in the first period to negative in the fourth period. This result could suggest that the training program initially facilitates finding a job but during later stages it instead reduces the probability of re-employment.

However, extreme caution should be put in formulating such conclusions. In facts, there may be factors that alter and bias the detected relationship between training and re-employment. In particular, we should put great attention in dealing with the issue of i) *spurious duration dependence* that generates an actual selection into sample during later stages of unemployment and ii) *self-selection into treatment*. The first issue could emerge because during later stages we only observe individuals unemployed up to that specific period. If the probability to remain unemployed is only due to observed variables no problems should arise; instead, if there is some unobserved heterogeneity that is individual-specific our result could be biased. Basically, the presence of unobserved heterogeneity could account for the fact that, apart from the observed variables, some individuals are simply "better" or "worse" than others in their employability. In particular, it could exist some mechanisms of selection that relate a certain category of workers (the trained, for example) with the probability that "good" workers are actually awarded with a job. As mentioned before, we use the term "good" to indicate those individuals whose unobservable characteristics are more favourable from an employability point of view. In these instances, the estimated effect of belonging to that category (the trained) could be biased during later stages and there could be an actual selection into (later stages) sample of the "good" or "bad" workers that is asymmetrical across certain categories. Therefore, the difference in different periods in the effect of being trained may be due to an asymmetrical selection into sample during later stages of unemployment. To account and check this aspect we are going to perform a discrete duration regression where we explicitly take into account for unobserved heterogeneity, which is called "frailty" in duration analysis terminology (see Subsection 5.3).

The second issue is related to the possibility that selection into treatment is not random and therefore certain categories of individuals are more likely to participate to it: in these circumstances, the effect of the treatment is confounded with the effect of the characteristics that make the participation to the treatment more likely and the result we obtain from the regression are biased: to explore and account for this issue we perform a propensity score matching analysis (see Subsection 5.4).

5.3 ***Discrete Duration Analysis***

We develop now a survival analysis using discrete duration regressions. As mentioned before, we have partitioned unemployment duration in 4 time periods, each of which is 2 months long. Given this partition, the discrete duration analysis consists in panel binomial regressions, where each individual, in each time period, has a probability to find a job. This probability depends on observed characteristics and on some unobserved characteristics, and, it remains constant in each time period for each individual.

This specification of the model allows to specifically estimate, in each single stage of unemployment, the effect of having participated to the training program. The discrete structure allows this effect to change from period to period and to follow non-linear and non-continuous patterns. From this point of view, the discrete duration estimation is better than the continuous one as the latter would only allow to estimate whether the effect of training is time-varying and what the direction of this variation but would not allow to assess if and when this effect reverts its sign or becomes negligible.

Within this setup, we introduce time dummies (T2, T3 and T4 for second, third and fourth time period, with the first one being the reference period) to capture the duration dependence and the pure effect of time. Given that some individuals find a job during the first period, we lose some observations in the next periods, therefore the panel is unbalanced. However, the robustness of the estimates is not compromised by this self-selection in the later periods. The panel structure in fact allows to capture the unobservable components that produce higher probabilities to remain (or leave) in the sample during each period. In particular, we are de facto assuming a random effect that is individual-specific and that explains the fact that, during the later stages, individuals still unemployed might have "worse" unobservable terms than those that find a job. Within the duration analysis the unobservable term is called "frailty" term and its explicit inclusion and estimation avoid the bias that could arise if the unobservable term becomes asymmetrical across certain categories of workers during the later stages of unemployment.

In our analysis we use again the log-log complementary regression but we do so in a panel model where the random effect component is represented by a gamma frailty: that is, we assume a Gamma distributed individual-specific error term. To check for robustness we also specify and estimate the model assuming a Gaussian distributed frailty. For a detailed discussion of these models specification, their estimation and how to implement it see Jenkins (1995) and Jenkins (2004).

Table 3 reports results for the two specifications of the model (in the table, we represent with Trained1, Trained2, Trained3 and Trained4 the effect that training has during first, second, third and fourth period of unemployment respectively): in the table, the coefficients represent the effect of the covariates on re-employment probability.

TABLE 3 here

The estimates contained in the table highlight how participating to a training program can have a strong significant and positive effect on re-employment during the very first stages of the unemployment spell, but this effect becomes smaller and non-significant starting from the second time-period (that is, after two months of unemployment).

The table also shows the results on the test on the frailty component and its distribution: while the direct estimation of this component for each individual cannot be obtained, the variance of its distribution can be estimated and used to test the very presence of this component (see Jenkins, 1995 and Jenkins, 2004 for a discussion of this procedure). The test on this component is a Likelihood Ratio test (LR) and is computed as the ratio of the standardized heterogeneity variance to one plus the heterogeneity variance. Therefore, if the hypothesis that this ratio is zero is rejected, then the frailty component is relevant and cannot be excluded. In both our cases (the Gamma and Gaussian distribution), the tests reject the hypothesis of a non-relevant frailty and thus, its inclusion is correct and relevant.

The results shown in Table 3 are particularly interesting when compared to those in Table 2 (in which we are not controlling for unobservable heterogeneity): in particular, we observe that the result for the first period is confirmed, while things change in the fourth one, with training being actually non-significant. These differences between the results in Table 2 and Table 3 suggest that, during the later stages of unemployment, individuals with training are associated with "worse" unobservable characteristics and thus, during these time periods, training may be erroneously considered detrimental to re-employment. In fact, once we explicitly account for individual-specific heterogeneity (which we have tested to be relevant), the negative effect of training disappears.

This line of reasoning suggests that individuals that participate to this training program and that remain unemployed after two months are somehow "worse" than their non-trained counterparts, suggesting that individuals with training are better screened or they are better in signalling their true skills. This result, together with the finding that the training program appears ineffective after the first two months, suggests that the training program fails in bestowing the individuals with better skills or competences but they instead acts as a good screening device. Table 3 also shows other relevant, even if expected, results: men have better chances to find a job than women and older individuals are better-off than younger ones, though the returns from age are decreasing (note that we are referring only to the 18-30 age group). Education is positive but not significant and this result may be conditioned by the lack of information on fields of study: individuals with higher degree may have more re-employment chances but this result may revert for some fields of study.

Time effect is not significant while the dummies representing the quarter when the search process began have significant negative coefficients: this could signal an overall worsening of the employment opportunities, something which is perfectly in line with unemployment rate at the aggregate level (within the Province of Pisa, unemployment rate raised from 6.8% in 2012 to 8.6% in 2013). All the results remain stable in both specifications of the model and are fully compatible with what we obtain estimating the model with continuous duration regressions (see Appendix A).

5.5 Propensity Score Matching Estimation

We deal now with the problem of self-selection into treatment. To understand the extent of the problem when measuring the effectiveness of the training program, suppose that $(Y_1, Y_0)_i$ are the two potential outcomes on the$i$-th population unit when treated or not treated, respectively. If a specific member of the target population receives the treatment then $Y_1$is observable (≡ factual), while Y_0 is irreversibly non-observable and corresponds to what we would have observed if this unit had not received the intervention (≡ counterfactual).

The equation that relates Y_1, the real outcome that is observed on unit i of the target population, to the potential outcomes of the same unit is the following:

$$Yi = Y1iDi + Y0i(1 - Di) = \begin{cases} Y1iifDi = 1 \\ Y0iifDi = 0 \end{cases}$$

where D_i is the treatment status of the i-th population unit: $D_i = 1$ if the unit received the intervention, and 0 otherwise. Given that to calculate individual-unit causal effects is typically impossible (and it is also less interesting from the policy point of view), the literature focuses its attention on the estimation of aggregated causal effects. In this article we estimate the average effect on units in the target population that were assigned to treatment (ATT):

$$ATT = E(Y1i - Y0i \vee Di = 1) = E(Y1i \vee Di = 1) - E(Y0i \vee Di = 1)$$

ATT measures the average treatment effect for the units actually exposed to the intervention and is the parameter of major interest for policy evaluation. Given that $E(Y0i \vee Di = 1)$, the average effect on the treated in the case they had not been treated, is not observed, one could take the outcome of non-participants as an approximation to the outcome that participants would have had without treatment. This would be a correct approach if (and only if) participants and non-participants have similar characteristics, i.e. if they were comparable *a priori* in the absence of the treatment. In general, however, participants and non-participants differ in crucial characteristics that are related both with the participation status and the outcome. This problem is known as "*self-selection* bias". The bias arises when treated and non-treated are systematically different even before the participation to the intervention. This could happen because assignment to the program is non-random: in this case, the estimated effect on re-employment of the participation to the program could be due not to the actual attendance of it, but to the unobserved characteristics that make more likely to attend the program.

A possible way to account for the presence of self-selection and to mitigate its bias is the use of the propensity score matching (PSM) methodology. The basic idea behind PSM is to compare treated individuals that are "similar" to non-treated individuals and to produce then an estimate of the treatment impact given by the difference between the outcome of the treated with the outcome of the matched comparison cases. The key issue is how to define similar individuals and the PSM determines the similarities computing what is called "propensity score": this measure is defined as the probability that an individual receives the treatment given a set of observed variables. Two individuals whose propensity scores have similar values are thus considered "similar". The very computation of the propensity score is obtained with the estimation of a selection model, a probit model in our case, in which the participation to the treatment is regressed on the characteristics of the individual and the probability to participate is thus computed: for a more detailed description of this estimation technique see Bryson *et al.* (2002) and Caliendo and Koeping (2008), for a more practical discussion on how to implement the estimation see Becker and Ichino (2002).

In our analysis we assume that being part of the training program defines the treatment to which individuals are exposed. Once the propensity score is computed, matched individuals can be compared and the effect of the treatment estimated averaging the differences in the outcome of all matched cases. There are several methods to perform the match: to check for the robustness of our results, we use the four most common techniques: stratification, radius, nearest neighbour and kernel. We apply this methodology to assess the effect of training to the probability of finding a job within two months: the results[1] are reported in Table 4, and they measure the average effect on re-employment probabilities of the treatment on the treated.

TABLE 4 here

According to the PSM analysis, the effect of training in the initial stage of unemployment is strictly positive, confirming the significance and the magnitude of the effect. Therefore, even when we control for the possible self-selection into treatment, the positive effect of training is confirmed. This also means that the self-selection, if present, is not particularly strong and, at any rate, it is not strong enough to bias the findings of the discrete duration regression. This is true at least for the initial effect of training: a similar PSM analysis cannot be performed for later stages of unemployment given that, for those stages, a selection into the surviving sample is also present. However, if self-selection is not present at the beginning (as the PSM analysis suggests) there is no apparent reason why it should arise and begin during the later stage and thus, the results we obtained in the discrete duration analysis should be robust for all the time horizons.

1

5.6 Summary of Results

Our analysis has highlighted two key aspects. First, attending a workplace training program increases the initial probability of finding a job and this result remains true even when we account for self-selection into treatment. Second, the effect of training becomes negligible after about two months and, more in details, it appears that, among the individuals that were unemployed for at least two months, those with training are associated with "worse" unobserved heterogeneity.

These results seem to suggest that workplace training programs only offer more chance to be hired immediately after their end. This can happen either because the training is strictly firm-specific and the firm providing the training directly hires the worker either because the worker may take advantage of the permanence in the firm to improve networks and contacts so that this can be exploited to obtain more easily a job. However, this improvement in re-employment probability is not permanent, suggesting that training is not actually good in permanently providing or improving skills.

Moreover, the asymmetric selection into sample that we observed during later stages of the unemployment suggests that workers with training are better screened than the rest so that, among the trained, the effectively "worse" workers are more likely to remain unemployed than their counterparts. This is also informative on how firms actually use and take advantage from the program and our results strengthen the idea that firms are using the training programs merely as screening devices.

Summing up, the effect of the workplace training programs is partly disappointing as they fail to provide an enduring increase in the skills and re-employment prospects: on the contrary, the program is successful in easing the connection between workers and firms (as trained have a higher initial probability of finding a job) and acts as a screening device for firms. Given that matching and screening issues are known to prevent full employment, workplace training programs are not completely useless but they still appear to fail at least in one of their main aims.

We have to stress that our findings are related to a specific training program and this is thus grounded on a specific mechanism as well as on the legislation and the context of Italy. Even with this limitation, our results are still valuable in assessing what are the characteristics of the training programs that could be useful and what are the aspects that are ineffective or detrimental. As we mentioned before, in Europe and US exist a large variety of training programs and they display different effects and effectiveness making really necessary to understand which mechanisms of the training programs are relevant in determining their overall effect. As we stressed several time, the on the job training we have analysed appears only to act as a mean to connect workers and firms and as a screening device. From this point of view, we offer a further piece of evidence on what the mechanics of training are, on what the training effect are and also on how firms can react to them, with this latter aspect being particularly valuable as this aspect has hardly been analysed up to now.

6. Conclusions

In this article we evaluate the effect of a workplace training program for young unemployed. The core of our analysis was to assess if and how this training is beneficial to the employability of the trained individuals. Descriptive evidence shows that the participation to the training program has a positive effect immediately after the end of it, but this beneficial effect rapidly dissolves and reverts to negative during later stages of unemployment. However, performing a more accurate econometrical analysis, we confirm the initial positive effect on re-employment but we find that the beneficial effect, while not persisting in time, do not revert to negative.

Our analysis indicates that what it is actually happening is that during later stages of unemployment we observe an asymmetric selection into sample among trained and non-trained individuals: this suggests that workers with training are better screened so that, among the trained, the effectively “worse” workers are more likely to stay unemployed than their counterparts.

Our evaluation of this workplace training suggests that the implemented program is successful in easing the connection between workers and firms, and acts as screening devices for firms, but it fails to provide a durable improvement in the skills and in the re-employment prospects.

Given that our analysis was based on an Italian program, it should be interesting to further develop the research to compare other countries and other training programs with the aim to obtain further evidence on firms behaviour and to assess whether firms effectively use systematically training programs as screening mechanisms also in other contexts and other countries.

Notes

1 Propensity score is obtained with a probit regression where participation is regressed on the same variables used in the duration analysis with the exception of the time dummies. The number of blocks for the propensity score is 8 and the balancing property is respected: this ensures that the mean propensity score is not different for treated and controls in each block. Standard errors are computed via bootstrap for the Kernel matching and analytically for the rest of the matching methods.

References

Angrist, J. D. and Krueger, A. B. (1999), “Empirical strategies in labor economics”, in: Ashenfelter O. and Card, D. (Eds.), *Handbook of Labor Economics*, Vol. 3A, Elsevier, Amsterdam pp. 1277-1366.

Arellano, F.A. (2010), “Do Training Programmes Get The Unemployed Back To Work? A Look At The Spanish Experience”, *Revista de Economia Aplicada*, Vol. 18 No. 2, pp. 39-65.

Baranowska, A., Gebel, M. and Kotowska, I. E. (2011), “The role of fixed-term contracts at labour market entry in Poland: stepping stones, screening devices, traps or search subsidies?”, *Work, Employment & Society*, Vol. 25 No. 4, pp. 777–793.

Barnow, B. S. (2000), “Exploring the relationship between performance management and program impact: A case study of the job training partnership act”, *Journal of Policy Analysis and Management*, Vol. 19 No. 1, pp. 118–141.

Barnow, B. S. and J. A. Smith. (2004), "Performance Management of U.S. Job Training Programs", in O'Leary, C. J., Straits, R. A., and Wandner, S. A. (Eds.), *Job Training Policy in the United States*, W.E. Upjohn Institute for Employment Research, Kalamazoo, pp. 21-55

Becker, S. O. and Ichino, A. (2002), “Estimation of average treatment effects based on propensity scores”, *Stata Journal*, Vol. 2 No. 4, pp. 358-377.

Bonnal, L., Fougere, D. and Serandon, A. (1997), “Evaluating the impact of french employment policies on individual labour market histories”, *Review of Economic Studies*, Vol. 64 No.4, pp. 683–713.

Brodaty, T., Crépon, B. and Fougére, D. (2001), “Using matching estimators to evaluate alternative youth employment programs: Evidence from France, 1986-1988”, in Lechner M. and Pfeiffer, F. (Eds.), *Econometric Evaluation of Labour Market Policies*, Physica-Verlag, HD, pp. 85–123.

Bryson, A., Dorsett, R. and Purdon, S. (2002), “The use of propensity score matching in the evaluation of active labour market policies”, *Department for Work and Pensions Working Paper* No.4.

Cahuc, P., Carcillo, P. and Zylberberg, A., (2014), *Labor Economics*, The MIT Press, Cambridge.

Caliendo, M. and Kopeinig, S. (2008), “Some Practical Guidance For The Implementation Of Propensity Score Matching”, *Journal of Economic Surveys*, Vol. 22 No. 1, pp. 31-72.

Card, D., Kluve, J. and Weber, A. (2010). "Active labour market policy evaluations: A meta-analysis", *Economic Journal*, Vol. 120 No. 548, pp. F452–F477.

Caroleo, F. E. and Pastore F. (2001), "How fine targeted is almp to the youth long term unemployed in Italy?", CELPE Discussion Papers 62.

Crépon, B., Ferracci, M. and Fougère, D. (2012) "Training the Unemployed in France: How Does it Affect Unemployment Duration and Recurrence?", *Annals of Economics and Statistics,* Vol. 107/108, pp. 175–199.

Fay, R. (1996). "Enhancing the Effectiveness of Active Labour Market Policies: Evidence From Program Evaluations in OECD Countries", *OECD Labor Market and Social Policy Occasional Paper* 18.

Fitzenberger, B., Lechner, M. and Smith, J. (2013), "Estimation of treatment effects: recent developments and applications", *Empirical Economics,* Vol.44 No. 1, pp. 1-11.

Frolich, M. and Lechner, M. (2010), "Exploiting regional treatment intensity for the evaluation of labour market policies", *Journal of the American Statistical Association,* Vol. 105 No. 491, pp. 1014-1029.

Heckman, J. J., Lalonde, R. J. and Smith, J. A. (1999), "The economics and econometrics of active labor market programs", in *Handbook of Labor Economics,* Vol. 3, Elsevier, Amsterdam, pp. 1865–2097.

Jenkins, S. P. (1995). "Easy Estimation Methods for Discrete-Time Duration Models", *Oxford Bulletin of Economics and Statistics,* Vol. 57 No. 1, pp. 129-138,

Jenkins, S. P. (2004), "PGMHAZ8: Stata module to estimate discrete time (grouped data) proportional hazards models", *Statistical Software Components* S438501, Boston College Department of Economics.

Kluve, J. (2010). "The effectiveness of European active labor market programs", *Labour Economics* Vol. 17 No. 6, pp. 904–918.

Kluve, J. and Schmidt, C. (2002), “Can training and employment subsidies combat European unemployment?”, *Economic Policy*, Vol. 17 No. 35, pp. 409–448.

Layard, R., Jackman R. and Nickell, S. (1991), *Unemployment: Macroeconomic Performance and the Labour Market*, Oxford University Press, Oxford.

Lechner, M., Miquel, R. and Wunsch, C. (2011), “Long-Run Effects Of Public Sector Sponsored Training In West Germany”, *Journal of the European Economic Association, European Economic Association*, Vol. 9 No. 4, pp. 742-784.

Loi, M. and Rodrigues, M. (2012), “A note on the impact evaluation of public policies: the counterfactual canalysis”, *MPRA Paper* (42444).

Long, J. (1997). *Regression Models for Categorical and Limited Dependent Variables*. Advanced Quantitative Techniques in the Social Sciences. SAGE Publications, Thousand Oaks.

Martin, J.P. (2000), “What Works Among Active Labour Market Policies: Evidence from OECD Countries’ Experiences”, *OECD Economic Studies* 3, OECD Publishing.

Martin, J.P., and Grubb, D. (2001), “What Works Among Active Labour Market Policies: A review of OECD Countries’ Experiences”, *Swedish Economic Policy Review*, Vol. 8 No. 2, pp. 9-60.

Martin, J. P. (2015). “Activation and active labour market policies in OECD countries: stylised facts and evidence on their effectiveness”, *IZA Journal of Labor Policy*, Vol. 4 No. 1, pp. 1-29.

McGuinness, S., O'Connell, P. J. and Kelly, E. (2014), “The Impact of Training Programme Type and Duration on the Employment Chances of the Unemployed in Ireland”, *The Economic and Social Review*, Vol. 45 No. 3, pp. 425-450.

OECD (1993), “Active Labour Market Policies: Assessing Macroeconomic and Microeconomic Effects”, in *Employment Outlook*, OECD, Paris, pp. 39-80.

OECD (1994). *The OECD Jobs Study: Evidence and Explanations*, OECD, Paris.

OECD (2007), “Activating the Unemployed: What Countries Do”, in *Employment Outlook*, OECD, Paris, pp. 207-242.

Osikumino, A. (2013). “Quick job entry or long-term human capital development? The dynamic effects of alternative training schemes”, *Review of Economic Studies*, Vol. 80 No. 1, pp. 313–342.

Richardson, K. and Van Den Berg, G. (2013), “Duration dependence versus unobserved heterogeneity in treatment effects: Swedish labor market training and the transition rate to employment”, *Journal of Applied Econometrics*, Vol. 28 No. 2, pp. 325–351.

Wunsch, C. and Lechner, M. (2008), “What Did All the Money Do? On the General Ineffectiveness of Recent West German Labour Market Programmes”, *Kyklos*, Vol. 61 No. 1, pp. 134-174.

Zweimuller, J. and Winter-Ebmer, R. (2003), “On-the-job-training, job search and job mobility”, *Swiss Journal of Economics and Statistics*, Vol. 139 No. 4, pp. 563-576.

APPENDIX 1

In this appendix we present results from the estimation of the survival model using continuous survival models. This allows to test the robustness of our estimation and to check whether our results depend too much on the discrete hazard model we used.

In Table 5 we present the estimates for Cox parametric regression assuming that baseline hazard is distributed according to a Weibull distribution: in the regressions we also assume that the effect of training is time varying and that a frailty component is present. In specification (i) the frailty component is distributed as a Gamma distribution while in specification (ii) it is distributed according to an inverse-gaussian distribution.

TABLE 5 here

We are mainly interest in the effect of training and from the above table we see that its effect on re-employment is, in both specifications, initially positive (direct effect is significantly positive) but it declines with time (interaction with time is significantly negative), possibly reverting to negative. This is clearly compatible with the results we obtained in the discrete hazard regression. The continuous regression, however, is not able to exactly pin whether and when the effect of training becomes negative so that the discrete hazard model appears to be more informative from this point of view. The results in the above table also confirm all the signs we found in the discrete hazard regression though significance in some case is slightly different. The only relevant difference between the discrete and continuous version stems from the effect of time (the baseline hazard in the continuous version). In the estimations in the above table we find that the effect of time is negative in specification (i) and positive in (ii). The change in its sign might be a symptom that the Weibull distribution is not the correct choice to describe the baseline hazards and it thus reinforces our choice of using time dummies in the discrete hazard model.

Tables

Table 1: **Share of individuals finding a job within 2 months (conditional of still being unemployed after *x* months).**

	Immediately	After 2 months	After 4 months	After 6 months
OVERALL	0.293	0.165	0.131	0.115
Among TRAINED	0.411	0.150	0.094	0.069
Among NON-TRAINED	0.287	0.166	0.133	0.117
Number of obs.	4,087	2,890	2,412	2,095

Table 2: **Estimate of log-log regression for the first and fourth period**

Vairables	Within 2 months	After 6 months
Female	-0.203***	-0.268**
	(0.0590)	(0.131)
Age	0.412***	-0.298
	(0.147)	(0.301)
Age Squared	-0.008***	0.005
	(0.00290)	(0.00599)
Trained	0.592***	-0.876**
	(0.122)	(0.419)
Education: VSS	0.117	0.260
	(0.126)	(0.286)
Education: USS	0.0225	0.279*
	(0.0653)	(0.143)
Education: UD	0.0285	0.620***
	(0.100)	(0.217)
Q3	-0.375***	0.505***
	(0.0681)	(0.182)
Q4	-0.383***	0.965***
	(0.0754)	(0.182)
Constant	-6.103***	1.445
	(1.839)	(3.723)
Observations	4,087	2,095

Notes: Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table 3: **Estimation of discrete duration model.**

Variables	Gamma frailty	Gaussian frailty
Female	-0.262**	-0.147*
	(0.122)	(0.0850)
Age	0.623**	0.466**
	(0.250)	(0.211)
Age Squared	-0.0118**	-0.00884**
	(0.00490)	(0.00411)
Trained1	1.048***	0.822***
	(0.329)	(0.277)
Trained2	0.656	0.291
	(0.463)	(0.350)
Trained3	0.288	-0.0546
	(0.515)	(0.402)
Trained4	-0.0479	-0.318
	(0.556)	(0.457)
Education: VSS	0.253	0.249
	(0.206)	(0.167)
Education: USS	0.116	0.124
	(0.101)	(0.0836)
Education: UD	0.208	0.213
	(0.159)	(0.131)
T2	0.168	-0.0697
	(0.289)	(0.323)
T3	0.425	-0.0493
	(0.473)	(0.460)
T4	0.700	0.00538
	(0.628)	(0.556)
Q3	-0.689***	-0.531***
	(0.163)	(0.159)
Q4	-0.600***	-0.371***
	(0.187)	(0.139)
Intercept	-8.148***	-7.383***
	(2.981)	(2.803)
LR test on Frailty Component	13.5953***	6.96***
Numb of Observations	4087	4087

Notes: Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table 4: **Estimation of the effect of training via Propensity Score Matching.**

	Stratification Matching	Radius Matching	Nearest Neighbour Matching	Kernel Matching
Effect of Training (ATT)	0.151***	0.135***	0.120**	0.138***
	(0.038)	(0.036)	(0.05)	(0.036)

Notes: Number of Blocks: 8. Balancing properties satisfied for all blocks.
Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 5 **Estimation of continuous duration model.**

Variables	(*i*) Gamma Frailty	(*ii*) Inverse-Gaussian
Female	-0.144**	-0.123
	(0.0640)	(0.0793)
Age	0.452***	0.525***
	(0.151)	(0.194)
Age Squared	-0.00861***	-0.0100***
	(0.00298)	(0.00383)
Trained	1.122***	1.407***
	(0.181)	(0.226)
Effect of Time on Training	-0.0121***	-0.0155***
	(0.00252)	(0.00345)
Education: VSS	0.220	0.292*
	(0.135)	(0.173)
Education: USS	0.112	0.154*
	(0.0683)	(0.0878)
Education: UD	0.215**	0.303**
	(0.106)	(0.133)
Q3	-0.514***	-0.582***
	(0.0796)	(0.0932)
Q4	-0.369***	-0.346***
	(0.0839)	(0.0997)
Constant	-9.530***	-10.40***
	(1.894)	(2.426)
Log of Ancillary parameter	-0.248***	0.0928***
	(0.0393)	(0.0343)
LR test on Frailty Component	38.19***	79.01***
Numb. Of Observations	4,087	4,087

Notes: Standard errors in parenthesis: *** p<0.01, ** p<0.05, * p<0.1.