# Evaluation of an Active Labour Market Programme in a Context of High Unemployment

# *Evaluación de un programa de políticas activas de mercado de trabajo en un contexto de elevado desempleo*

Cristina Borra
Luis Palma
M. Carmen González
Luis F. Aguado*

## Abstract

We evaluate the effectiveness of a programme aimed at a group of unemployed in the capital of the South of Spain, within the framework of Active Labour Market Policies (ALMPs). We use high quality administrative data which justifies

---

* Contact information: Cristina Borra is Associate Professor of Microeconomics and Labour Economics at the Department of Economics and Economic History, University of Seville (Spain); her research focuses on Labour and Demographic Economics, including Immigration, Job Satisfaction, Work-Life Balance, and Policy Evaluation; email: cborra@us.es. Luis Palma is Professor of Political Economy and History of Economic Thought at the Department of Economics and Economic History and Chair of the Centre for Competition Economics, University of Seville (Spain); his research interests include Welfare Economics, Social Policy, Cultural Economics, and Industrial Economics; email: lpalma@us.es. M. Carmen González is Associate Professor of Mathematics at the Department of Applied Economics of the University of Seville; her research interests include Health Economics, Welfare Economics, and Old Age Dependency; email: carmengc@us.es. Luis F. Aguado is Associate Professor of Economics at the Department of Economics, Pontificia Universidad Javeriana (Colombia); his research interests include Labour Economics, Policy Evaluation, and Cultural Economics; email: lfaguado@javerianacali.edu.co.

We are grateful to Dr. Almudena Sevilla-Sanz, University of Oxford, and Dr. Begoña Cueto, University of Oviedo, for helpful comments on previous versions of this paper. The usual disclaimer applies.

Este artículo fue recibido el 2 de febrero de 2012; modificado el 21 de agosto de 2012 y, finalmente, aceptado el 26 de octubre de 2012.

the application of propensity score matching methods. The estimated effects are positive with regard to employment, job security, and working hours in the short-term (6 months). However, this is not true in the long-run (36 months). No significant effects have been found on earnings, in neither the short nor long-term. Overall these results are quite robust with respect to the matching algorithm choice and the potential influence of unobserved heterogeneity. Although, the short duration of the programme seems appropriate, the disappointing long-term results highlight the difficulties of putting participants back into stable work in a context of high unemployment.

*Key words*: Unemployment, propensity score matching (PSM), programme evaluation.

*JEL classification*: J08, J60, C14, C52.

## Resumen

Este trabajo evalúa el impacto de un programa dirigido a un grupo de desempleados en la capital del sur de España, en el marco de las políticas activas del mercado de trabajo (PAMT). Se han utilizado para ello datos administrativos de alta calidad, lo que justifica la aplicación de métodos de propensity score matching (PSM). El efecto estimado es positivo en materia de empleo, seguridad en el empleo y horas de trabajo, en el corto plazo (6 meses), pero no en el largo plazo (36 meses). No se ha encontrado un efecto significativo en los ingresos, ni en el corto ni en el largo plazo. En general, estos resultados son bastante consistentes con respecto al algoritmo *matching* elegido y la influencia potencial de la heterogeneidad inobservada. Aunque la corta duración del programa parece apropiada, los pobres resultados a largo plazo reflejan las dificultades de los participantes para conseguir un trabajo estable en un contexto de elevado desempleo.

*Palabras clave*: desempleo, propensity score matching (PSM), evaluación de programas.

*Clasificación JEL*: J08, J60, C14, C52.

## Introduction

This paper evaluates the effectiveness of a local active labour market programme using a very rich administrative data set. Unemployment is a clear example of a trigger event that may have important inequality-enhancing impacts (Gangl, 2006). The persistence of unemployment in European countries, especially in comparison to the United States, has thus drawn attention of academics and policymakers over the last few decades (Kluve and Schmidt, 2002). The case of Spain is especially relevant in this respect. In particular, the South of Spain has shown a persistent differential in unemployment rates with respect to the rest of Europe, of at least 7 percentage points (Eurostat, 2009). This situation has in part determined the need to complement the unemployment subsidies of passive policies with active measures such as job search assistance, classroom or on-the-job training, subsidized employment, or self-employment promotion. [1]

The effectiveness of the Active Labour Market Policies (ALMPs) has been the object of intense debate in recent academic literature. From the theoretical point of view, they have emphasized that the positive effects on worker productivity or improved job matching can be reduced by a "deadweight effect" arising from the workers who would have been employed in any case and by a "substitution effect" arising from the fact that the policy may lead to the substitution of some workers by others, without really generating any employment (Calmfors, 1994). From an empirical point of view, it has been noted that measured ALMP effectiveness depends on the specific country involved, the length of the policy, the characteristics of the participants, the programme type, and the evaluation methodology used (Dar and Tzannatos, 1999; Card, Kluve and Weber, 2010; Kluve, 2010). This uncertainty with regard to the effects of the measures, together with increased budget constraints, suggests the need to regularly evaluate labour market policies.

We evaluate a short-duration combination programme (including training courses, labour orientation, and work placements) targeted at people who

1 In fact, in the European Union (EU-15), spending on Active Labour Market Policies (ALMPs) has increased significantly in most countries over the last few decades, reaching the point in 2005, when it represents between 0.44% and 1.58% of Gross Domestic Product (GDP) (in the United Kingdom and Denmark respectively) (OECD, 2009). Spain is no exception: its spending on active policies has increased from 0.33% of GDP in 1985 to 0.78% in 2005 (OECD, 2009).

enter from registered unemployment that was administered locally from the capital of Andalusia (Spain), with funding from the European Social Fund. We base our analysis on high quality administrative data, enriched with information from two follow-up surveys. This informative data set justifies assuming conditional independence in the application of propensity score matching methods. We claim that this is the best methodology we can use given the particularly rich set of control variables available.[2] We are also able to consider different outcome measures (earnings, probability of employment, job security, and working hours) and for different time periods (at 6 and 36 months since completion of the programme).

Our results indicate that the programme presents positive effects on the participants in the short-term (at 6 months), which are not maintained in the long-term (36 months). These positive program introduction effects, much larger at the beginning than later on, have been previously reported by Blundell, Costa Dias, Meghir and Van Reenen's (2004) analysis of job search programmes in United Kingdom. The opposite result has been found for training programmes by Card et al.'s (2010) meta-analysis. It is difficult to compare our results to those of existing studies, though, because of the comprehensive nature of the programme under evaluation. Other combination programmes evaluated are not as comprehensive. For instance Winter-Ebmer (2006) reports positive short and long run effects on an Austrian combination programme involving training and job-search counselling, whereas Centeno, Centeno and Novo (2009) report small positive to negative effects on a similar Portuguese programme.[3]

More specifically, estimated treatment effects are positive with regard to employment, job security, and working hours in the short-term. No significant effects have been found on earnings, in neither the short- nor long-term. Most previous studies focus on just one or two outcome measures, generally employment probability, unemployment duration, and earnings, so our

---

2 Convincing instrumental variables to deal with endogeneity issues of the treatment variable were not at hand, nor convincing thresholds for regression discontinuity designs. Nevertheless, as explained below, we test the sensitivity of results to the potential influence of unobserved factors.

3 Note that some multi-treatment programmes may have been evaluated as combination programmes. (See for instance Sianesi's (2004, 2008) evaluations of Swedish ALMPs or Heckman et al.'s (1998) and Plesca and Smith's (2007) evaluations of the Job Training Partnership Act in the United States). But ours is a true combination programme which participants join as a whole.

results are also difficult to compare in this regard.[4] Nonetheless, in terms of the meta-analysis conducted by Card et al. (2010), our study belongs to the 39.3% which obtain significantly positive impacts in the short term and the 40.0% which obtain insignificant effects in the long term.

We test the sensitivity of the results to the choice of the matching algorithm and the potential occurrence of unobserved heterogeneity. In fact, even though we are aware of the existence of previous work on these issues (for instance, Amuedo, Malo and Muñoz (2008) use different matching algorithms and Caliendo and Künn (2011) test for unobserved heterogeneity), to our knowledge our work is the first to present sensitivity results for the potential impact of unobserved factors for different matching algorithms.

Our paper adds to the literature in three important ways. First, we contribute to the rather scarce evidence that exists regarding the evaluation of programmes that boost employment in Southern Europe, and specifically in Spain. While international experience in evaluation is relatively wide, especially in The United States, Germany, and Northern Europe (Heckman, Lalonde y Smith, 1999; Card et al., 2010), evaluation in Spain is relatively infrequent and very recent.[5] However the Spanish labour market is characterized by a high and persistent unemployment rate, the highest in Europe (Saint-Paul, 2000)[6]. And additionally, Spain has experienced a significant increase in spending on ALMPs, along with a significant regional decentralization of labour policies over the last decade (Cueto and Mato, 2009). In this regard, our findings are particularly relevant to assess whether the effect of these policies in an economy with high unemployment rates and decentralized spending are similar to those implemented in other European countries, with lower unemployment rates and different degrees of centralization of labour policies. Second, we analyze

4 Few studies evaluate more than one outcome measure. For instance Sianesi (2004) and Winter-Ebmer (2006) consider just two. To our knowledge, less than a handful of studies consider multiple outcomes (i.e., Hardoy, 2005; Cavaco, Fougere and Pouget, 2005; and Mato and Cueto, 2008). Nevertheless none of them consider simultaneously job security or working hours together with the more common probability of employment or earnings measures.

5 To our knowledge, we can only name a handful of studies. We may cite Mato and Cueto (2008), Cueto and Mato (2009), and Arellano (2010), all analyzing training programmes. García-Pérez and Rebollo-Sanz (2009) focus on regional wage subsidies, Malo and Muñoz-Bullón (2006) analyze measures that promote employment for the physically and mentally disadvantaged, and Ramos, Surinach and Artís (2009) cover a diverse group of active employment policies.

6 In fact, in 2009 it has the highest level of unemployment (18%) of all OECD countries (OECD, 2009)

a comprehensive intervention programme, which includes training courses, labour orientation, and work placements. As previously stated, most previous studies refer to training or subsidized employment programmes in isolation and little is known about the likely consequences of combination programs such as the one we consider here. And finally, unlike most previous studies which focus solely on earnings or the probability of employment, given the richness of our administrative data, we are able to evaluate the programme's effects on multiple outcome measures and for different time periods. Thus we are able to asses the programme impact in a more widespread fashion than usually found.

The paper is organized as follows. Section I describes the institutional framework of the Spanish labour market and the data used for the analysis. Section II outlines our evaluation approach and Section III presents the findings. Section IV summarizes our results and concludes the paper.

## I. Institutional Framework and the Data Base Used

Traditionally, the South of Spain has shown a persistent differential in unemployment rates with respect to the rest of Europe, of at least 7 percentage points (Eurostat, 2009). In this context, the programme evaluated intends to favour the employability of those unemployed, by offering them a comprehensive support plan including orientation, training, and professional work placements.

The programme is free and participation is voluntary. It is conducted and administered by local public officials. There are two main categories of actions: 1) specialized training, aimed at unemployed with different education levels, and 2) training intended for groups in risk of social exclusion: immigrants, ex-drug addicts, long-term unemployed, and physically or mentally disadvantaged. Interested individuals apply for the specific action and public officials select participants based on the adequacy of their curriculum for the topic involved. High unemployment rates in Southern Spain guarantee a continuous supply of applicants. Selected participants benefit from a comprehensive intervention action, which includes a training course (of approximately 350 hours), a paid internship (200 hours), and labour orientation, everything taking place during a period of three months. Thus, according to the classification used by Card et al. (2010), we are dealing with a combination programme of short duration (less than 4 months) targeted at people who enter from registered unemployment.

The dataset used in this analysis corresponds to the actions offered from October 2004 to May 2005. Our main source of information comes from the administrative details of programme applicants. In order to get additional information regarding the labour market status of participants and non-participants, two follow-up telephone surveys were carried out, one during 2005 and the other during 2008, at 6 months and 36 months approximately since completion of the programme.[7] The total population comprised 990 subjects, 693 participants for the treated group and 297 non-participants for the control group. In the process of telephone interviews it was not possible to locate some of the individuals (as a result of changes in telephone numbers), thus the final sample observed was 520, 363 corresponding to the treated group and 157 to the control group. As shown in Table 1, except for a minor reduction in the proportion of social exclusion actions and a slight increase in the proportion of high educated individuals, dropping those observations had virtually no impact on the characteristics of our sample. It is worth emphasizing that most applicants are highly educated relatively young women. This is no surprise as this group is usually considered as one of the most vulnerable to unemployment (Verick, 2009) and European large cities usually attract a disproportionate share of highly-educated individuals (Carlino and Saiz, 2008).

**Table 1.** Sample Selection

| | Treated Group | | Control Group | | All | |
|---|---|---|---|---|---|---|
| | **Initial sample** | **Final sample** | **Initial sample** | **Final sample** | **Initial sample** | **Final sample** |
| Sample size | 693 | 363 | 297 | 157 | 990 | 520 |
| Female sex (%) | 79% | 86% | 81% | 88% | 81% | 83% |
| Age (years) | 30.39 | 29.60 | 30.96 | 29.37 | 30.15 | 30.45 |
| Married (%) | 28% | 24% | 27% | 16% | 26% | 23% |
| Social exclusion case (%) | 42% | 38% | 37% | 26% | 40% | 34% |
| **Individual's educational level (%)** | | | | | | |
| With no studies | 5% | 5% | 4% | 5% | 5% | 4% |
| Primary studies | 26% | 21% | 18% | 15% | 24% | 17% |
| Secondary studies | 12% | 8% | 11% | 4% | 11% | 9% |
| University studies | 57% | 66% | 68% | 76% | 60% | 71% |

7 A translation of the questionnaire used in the phone surveys is offered in the Appendix.

Our outcome measures are employment probability, job security (measured as the probability of obtaining a permanent, instead of a temporary, contract), working hours (measured as the probability of getting a full-time, instead of a part-time, contract), and earnings.

The profile of those applying for the programme shows that 83% are women, 23% of which are married, with an average age of 30. On average, these women have a university degree and have been unemployed for 8.8 months. Only 14% of applicants have received state help (unemployment subsidy or some type of transfer from the State). 33% of the observations belong to applicants for actions aimed at groups in risk of social exclusion.

## II. Methodology and Estimation Procedure

The aim of our analysis is estimating, in the terminology of Rubin (1974), the causal effect of the programme on the outcomes of a participating individual. Formally, let $Y_1$ denote the outcome if the individual was enrolled in the programme, and $Y_0$, the outcome otherwise. Hence, for a given individual $i$, the impact of agency participation, $\Delta_i$, is defined as:

$$\Delta_i = Y_{1i} - Y_{0i} \tag{1}$$

Suppose $D$ is an indicator variable that equals 1 for individuals who participate in the programme and zero for individuals who do not participate. A variety of labour market impact measures can be estimated (Caliendo, 2006). However, we are mostly interested in the Average Treatment effect on the Treated (ATT), that is:

$$\Delta_{ATT} = E\left(\Delta \middle| D=1\right) = E\left(Y_1 \middle| D=1\right) - E\left(Y_0 \middle| D=1\right) \tag{2}$$

which tells us whether, on average, unemployed participants benefited from joining the programme. The major difficulty in assessing the ATT originates in the complexity of evaluating $E\left(Y_0 \middle| D=1\right)$. This is known as the "Fundamental Problem" in the Evaluation Literature (Holland, 1986, p. 947), as the participants' outcome which would have arisen in the case of their not participating $Y_0$ cannot be observed.

Ideally, social experiments take persons who would otherwise participate in a program and randomly assign them to the participating (treatment) group or the non-participating (control) group. As a result of random assignment, under certain assumptions a simple comparison of the mean outcomes in the experimental treatment and control groups produces a consistent estimate of the impact of the program on its participants (Smith, 2000). Matching methods aim to recreate the conditions of randomness of a laboratory experiment by pairing off treated individuals with "similar" non-treated individuals. In order to do so, they rely on the Conditional Independence Assumption (CIA), which implies that, conditional on a set of observable variables (*X*), assignment between the treatment and control groups is random:

$$Y_0, Y_1 \perp D | X \tag{3}$$

In this way, remaining differences in the outcome variables are exclusively due to the treatment. The CIA is thus crucial for correctly implementing matching methods. The condition implies that all variables that influence treatment assignment and potential outcomes simultaneously have to be observed by the researcher (Caliendo and Kopeinig, 2008). Clearly, this is a strong assumption and has to be justified by the data quality at hand.

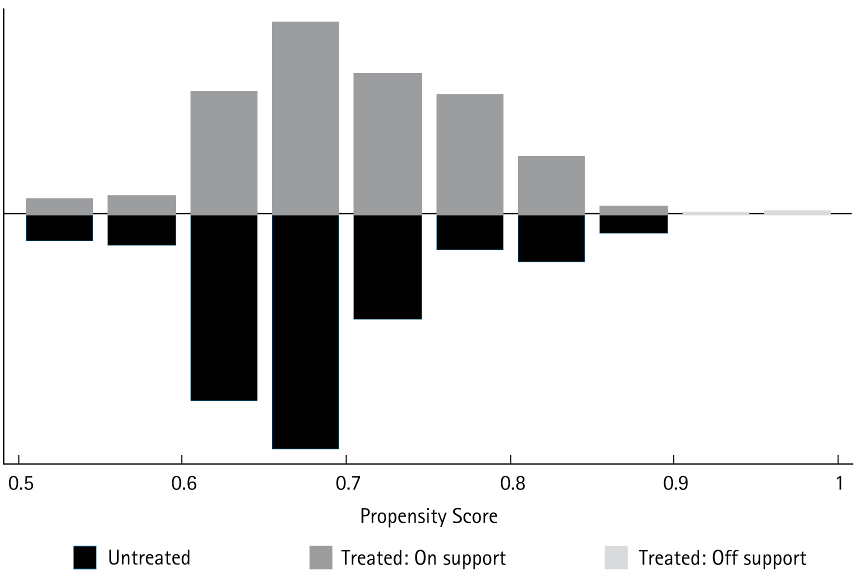
In our view, the dataset used in this analysis contains sufficient information to ensure that the CIA holds. In particular, our information complies with the recommendations for quality of matching by Heckman, Ichimura, Smith and Todd (1998), as: a) the treated groups and the control groups share the same local labour market; b) the information comes from the same source questionnaire in both cases; and c) we have information regarding their work experience along with other socio-demographic data. Actually, our control group is made up of rejected applicants that were offered no other intervention and the fact that they also applied to the programme makes treatment and control groups more similar, also with respect to unobserved characteristics that may affect selection bias (Cueto and Mato, 2009; Raaum and Torp, 2002). Nevertheless we acknowledge that the selection process was not random. It was actually based on a personal interview and on the applicants' observable characteristics. We claim that after conditioning on those observable variables, there should not remain much selection bias. Nonetheless, we additionally perform a sensitivity analysis at the end of Section III to gauge the potential impact of unaccounted selection on unobservables.

For the matching method to provide valid estimates of the impact of programme participation, a further requirement besides independence is the common support or overlap condition. It ensures that persons with the same $X$ values have a positive probability of being both participants and nonparticipants. Formally:

$$0 < P\left(D=1\middle|X\right) < 1 \tag{4}$$

The common support assumption implies that, for each treated individual, there is another non-treated individual who can be used as a matched comparison observation. While there is no formal test for the CIA, the validity of the common support assumption can be tested. Figure 1 shows the propensity score histogram by treatment status.[8] As can be observed, given the high degree of overlap between the two distributions, for the large majority of the treated individuals there is a similar control group individual, in such a way that the common support assumption is satisfied.[9]

**Figure 1.** Distribution of Estimated Propensity Score by Treatment Status




8 The results of the underlying logit model are presented below.

9 In the estimations, only three individuals are discarded.

As noted earlier, we need a large number of exogenous variables to ensure the validity of the CIA. But conditioning on all relevant covariates is limited in the case of a high dimensional vector *X*. To deal with this dimensionality problem, Rosenbaum and Rubin (1983) suggest using the propensity score $P(D = 1 \mid X) = P(X)$, i.e. the probability for an individual to participate in a treatment given his observed covariates *X*. This implies measuring "similarity" between individuals with respect to their estimated probability of participation in the programme. Rosembaum and Rubin (1983) show that if potential outcomes are independent of treatment conditional on covariates *X*, they are also independent of treatment conditional on the propensity score.

Therefore, the first stage in the matching is to model the propensity score. Table 2 displays the results from the probit model of the likelihood of participating in the programme. The results show that men were more likely and college graduates were less likely to participate in the programme. Because of programme design, the longer the unemployment duration, the higher the likelihood to be selected for the programme. Additionally, unemployed with previous placement experience were more likely to join the treated group. As can be observed, following Caliendo and Kopeinig (2008) only variables unaffected by participation (or the anticipation of it) were included in the model.

**Table 2.** Propensity Score Coefficient Estimates

| Variable | Coefficient | | St. Error |
|---|---|---|---|
| Female sex | -0.2864 | ** | 0.136 |
| Married | 0.1025 | | 0.146 |
| University studies | -0.3686 | ** | 0.172 |
| Unemployment duration | 0.0120 | ** | 0.006 |
| Work experience | 0.0215 | | 0.139 |
| Work placements | 0.2827 | * | 0.155 |
| Voluntary work | -0.1799 | | 0.141 |
| Recipient of state benefit | 0.2029 | | 0.179 |
| Special case | -0.1055 | | 0.162 |
| Constant | 0.8175 | *** | 0.211 |

Note: Significance level: *10%; **5%; ***1%.

In this type of evaluations it is equally convenient to analyze the quality of the matching between treated and non-treated individuals. Rosenbaum and Rubin (1983) suggest that we check whether significant differences between the average values of the variables for both groups exist after matching. Before matching we expect differences, yet after matching the variables should be balanced in both groups and significant differences should not persist. Table 3 presents the mean values of the variables considered for both treated and controls, before and after matching. For the majority of the variables, matching reduces the bias that exists between the distributions. The t-test rejects the null hypothesis of significant differences. However, the variable Secondary studies is worse after matching, with mean values for treated and controls turning significantly different. For this reason, the additional test of stratifi-

**Table 3.** Matching Quality

| Variable | Unmatched | | | | | Matched | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Treated | Control | %bias | t-test | | Treated | Control | %bias | t-test | |
| Female sex | 0.79 | 0.86 | -19.1 | -1.94 | * | 0.79 | 0.79 | -0.1 | -0.01 | |
| Age | 30.37 | 29.60 | 9.0 | 0.94 | | 30.32 | 30.19 | 1.5 | 0.19 | |
| Married | 0.28 | 0.24 | 9.3 | 0.96 | | 0.28 | 0.27 | 1.6 | 0.21 | |
| With no studies | 0.05 | 0.05 | -1.8 | -0.19 | | 0.04 | 0.06 | -8.9 | -1.14 | |
| Primary studies | 0.26 | 0.21 | 11.7 | 1.20 | | 0.26 | 0.28 | -4.7 | -0.61 | |
| Secondary studies | 0.12 | 0.08 | 15.1 | 1.52 | | 0.12 | 0.08 | 12.8 | 1.68 | * |
| University studies | 0.57 | 0.66 | -18.7 | -1.94 | * | 0.57 | 0.57 | 0.3 | 0.03 | |
| Time unemployed | 9.61 | 7.01 | 20.7 | 2.05 | ** | 9.17 | 8.07 | 8.8 | 1.31 | |
| Work experience | 0.75 | 0.73 | 5.7 | 0.60 | | 0.75 | 0.74 | 3.5 | 0.47 | |
| Time work experience | 43.02 | 31.30 | 25.4 | 2.15 | ** | 42.50 | 36.03 | 14.0 | 1.53 | |
| Work placements | 0.54 | 0.54 | -0.5 | -0.06 | | 0.54 | 0.52 | 3.3 | 0.44 | |
| Time work placements | 292.68 | 341.53 | -19.9 | -1.15 | | 292.68 | 312.96 | -8.3 | -0.64 | |
| Voluntary work | 0.32 | 0.40 | -15.5 | -1.64 | | 0.32 | 0.29 | 7.1 | 0.99 | |
| Recipient of state benefit | 0.15 | 0.12 | 9.4 | 0.97 | | 0.14 | 0.14 | 0.8 | 0.11 | |
| Social exclusion | 0.41 | 0.38 | 7.9 | 0.82 | | 0.41 | 0.43 | -4.0 | -0.54 | |

*Note*: Significance level: *10%; **5%; ***1%.

cation suggested by Dehejia and Wahba (2002) was carried out. We divided the observations into sectors based on the estimated propensity score and we later checked whether within each sector significant differences in the distribution of each of the explanatory variables persisted. In the application of this test to our sample we obtained six sectors within which the distribution of the variables was balanced.[10]

## III. Results

As already stated, we use different outcome measures for the calculation of causal effects: probability of employment, job security, working hours, and earnings. For these variables, we estimate ATT using different matching estimators as sensitivity tests: Epanechinikov kernel matching, Gaussian kernel matching, and radius matching. Table 4 shows the estimation results, noticeably robust to different specifications. In the short-term (6 months), the effect of the programme is positive according to all the methods employed and for all the variables considered, although the result is not significant in the case of earnings. The individuals participating in the training programme present a greater employment probability (around 26 percentage points), job security (around 28 percentage points) and probability of obtaining a full-time contract (around 23 percentage points). Calmfors (1994) points out at the potential decrease in job search intensity by programme participants whilst on the programme, what the literature designates as “lock-in effects”. We find no evidence of these adverse lock-in effects in our data. Given the relative short duration of the evaluated programme, it is very unlikely that adverse effects on search effectiveness and effort, such as those highlighted by Sianesi (2008), Lechner, Miquel and Wunsch (2007), or Cueto and Mato (2009), arise in our case.

10 We used the procedure developed by Becker and Ichino (2002) for Stata. The results are available to any researchers that request them.

**Table 4.** Average Treatment Effect on the Treated at 6 and 36 Months since Programme Participation

| Variables | Epanechnikov kernel matching | | Gaussian kernel matching | | Radius matching | |
|---|---|---|---|---|---|---|
| | ATT | Standard error | ATT | Standard error | ATT | Standard error |
| **Effect after 6 months** | | | | | | |
| Employment | 0.257*** | 0.058 | 0.252*** | 0.062 | 0.284*** | 0.069 |
| Permanent contract | 0.276*** | 0.083 | 0.253*** | 0.075 | 0.310*** | 0.090 |
| Full-time contract | 0.226*** | 0.103 | 0.228*** | 0.087 | 0.234** | 0.116 |
| Earnings | 42.223 | 97.860 | 51.013 | 92.516 | 119.243 | 124.529 |
| **Effect after 36 months** | | | | | | |
| Employment | 0.514 | 0.070 | 0.042 | 0.060 | 0.046 | 0.074 |
| Permanent contract | –0.124 | 0.080 | –0.140* | 0.071 | –0.143 | 0.099 |
| Full-time contract | 0.014 | 0.078 | 0.028 | 0.068 | 0.013 | 0.087 |
| Earnings | 2.790 | 65.020 | -23.831 | 61.484 | -30.221 | 79.639 |

*Note*: Significance level: *10%; **5%; ***1%.
Standard errors are computed by bootstrapping with 200 repetitions.

In the long-term (36 months), however, the effect of the programme is by and large insignificant according to the different methods for the majority of the variables. Mato and Cueto (2008) also point out a reduction in the importance of the effects over time, although theirs is less dramatic[11]. Nevertheless, in our case, we can appreciate a negative and significant effect on job security -measured by the probability of having a permanent contract-, for one of the methods. Caliendo (2006) also points out the existence of disappointing results for many groups of unemployed, explained by the existence of "stigma effects". As this author argues, if the programme aims to favour people with disadvantages, there is always a risk that a possible employer takes participation in such schemes as a negative signal. According to our results, however, these stigma effects are not very robust and appear only after the initial short-term benefits of the program have vanished.

11 In this respect, the Spanish case is contrary to the average international evidence, for which classroom and on-the-job training programs are not especially favourable in the short-run (Card et al., 2009).

Taking both short- and long-term results together, we find in fact evidence of what Blundell et al. (2004) refer to as "positive program introduction effects", which are much larger at the beginning than later on. This is in contrast with the evidence reported for training programmes by Card et al.'s (2010) meta-analysis. Nonetheless the programme evaluated is a combination programme which offers not only training, but also job search assistance and work placements. Apparently programme impacts are closer to Blundell et al.'s (2004) job search effects than Card et al.'s (2010) training effects.

One shortcoming of the propensity score matching approach is its reliance on CIA. If participants and non-participants differ in terms of not only observed, but also unobserved characteristics, the CIA is violated and therefore our results are biased. Following Caliendo and Künn (2011), we thus check the robustness of our results with respect to deviations from this assumption. Since testing the CIA directly with non-experimental data is not possible, we address this problem with the bounding approach initially suggested by Rosenbaum (2002). This approach consists of simulating an unobserved component and testing to which degree of unobserved heterogeneity results are robust. The main idea is that in the presence of unobserved factors, identical individuals with respect to observable characteristics ($X_i$) have different probabilities of receiving treatment. Therefore, an artificial factor $\Gamma$ is introduced to simulate an unobserved term. The influence of this unobserved term is gradually increased to assess its effect on the results by comparing the successful number of individuals in the treatment group with the same expected number, given that the treatment effect is zero (Becker and Caliendo, 2007).

Table 5 summarizes sensitivity test statistics for the average treatment effects of Table 4. Clearly, a sensitivity analysis for insignificant treatment effects is not meaningful and hence will be omitted. For the positive estimated treatment effects, we report the test statistic Q+ for the upper bound, under the assumption that we have overestimated the treatment effects and those who participate always have a higher employment probability or likelihood of having a permanent or a full-time contract even in the absence of treatment. Conversely, for the only negative estimated treatment effect we provide the test statistic Q– for the lower bound, under the assumption that we have underestimated the effect and those treated have a lower probability of having a permanent contract anyway. The test statistics are calculated for all three matching algorithms (Epanechinikov kernel, Gaussian kernel matching, and radius matching). Besides the test-statistics and the respective p-values for different values of $\Gamma$, we show the critical values of $\Gamma$ at which the test statistic Q+ turns insignificant with a 95% confidence level, thus implying that the treatment effects are actually due to unobserved factors.

For all outcome measures considered, our departing point is a situation of no unobserved heterogeneity with $\Gamma$ = 1.0. We then gradually increase the value of $\Gamma$, to assess the potential strength of unmeasured influences. For the employment outcome variable, measured 6 months after completion of the programme, results are quite robust to unobserved factors. Critical values of $\Gamma$ are between 1.75 and 1.90 indicating that individuals with the same X-vector would have to differ in their odds of participation by a factor of 1.75 (1.90), or 75% (90%) for treatment effects to turn insignificant at the 5% significance level. As for the estimated effects on job security and working hours, also at 6 months since completion, results are slightly more sensitive to unobserved factors with critical values ranging from 1.15 to 1.30, depending on the matching algorithm used. Finally, the negative estimated treatment effects for job security at 36 months since completion of the programme are very sensitive to potential unobserved heterogeneity. With just a 10 or 15% difference in the odds of participating of individuals with the same observed characteristics, treatment effects turn insignificant. Therefore we feel quite confident on the robustness of our short-term probability-of-employment result, fairly confident on our short-term job-security and working-hours results, and quite unsure on our long-term job-security result.

## IV. Concluding Remarks

In this work we estimate the causal effect of a comprehensive active labour market programme on the probability and the quality of employment of participating individuals, using propensity score matching techniques. Our fundamental result is the existence of positive programme introduction effects, that is, the programme has positive effects in the short-term which are not maintained in the long-term. These findings are quite robust with respect to the matching algorithm choice and the potential influence of unobserved heterogeneity.

**Table 5.** Sensitivity to Unobserved Heterogeneity

| | | | | | | |
|---|---|---|---|---|---|---|
| **After 6 months** | | | | | | |
| **Gamma** | **Q+** | **p+** | **Q+** | **p+** | **Q+** | **p+** |
| | **Epanechnikov kernel matching** | | **Gaussian kernel matching** | | **Radius matching** | |
| Employment | | | | | | |
| 1.00 | 3.99 | 0.000 | 3.99 | 0.000 | 4.26 | 0.000 |
| 1.25 | 3.03 | 0.001 | 3.03 | 0.001 | 3.31 | 0.000 |
| 1.50 | 2.26 | 0.011 | 2.26 | 0.011 | 2.54 | 0.005 |
| 1.75 | 1.61 | 0.053 | 1.61 | 0.053 | 1.90 | 0.028 |
| 2.00 | 1.05 | 0.145 | 1.05 | 0.145 | 1.35 | 0.088 |
| Critical value 5% | 1.75 | | 1.75 | | 1.90 | |
| Permanent contract | | | | | | |
| 1.00 | 2.12 | 0.017 | 1.96 | 0.024 | 2.06 | 0.019 |
| 1.25 | 1.47 | 0.071 | 1.31 | 0.095 | 1.43 | 0.075 |
| 1.50 | 0.94 | 0.173 | 0.77 | 0.220 | 0.93 | 0.173 |
| 1.75 | 0.50 | 0.309 | 0.32 | 0.373 | 0.50 | 0.308 |
| 2.00 | 0.11 | 0.454 | -0.07 | 0.526 | 0.13 | 0.447 |
| Critical value 5% | 1.20 | | 1.15 | | 1.20 | |
| Full-time contract | | | | | | |
| 1.00 | 2.36 | 0.009 | 2.25 | 0.012 | 1.93 | 0.026 |
| 1.25 | 1.71 | 0.043 | 1.60 | 0.055 | 1.31 | 0.095 |
| 1.50 | 1.18 | 0.119 | 1.07 | 0.143 | 0.80 | 0.213 |
| 1.75 | 0.73 | 0.231 | 0.62 | 0.267 | 0.38 | 0.353 |
| 2.00 | 0.35 | 0.362 | 0.23 | 0.407 | 0.01 | 0.496 |
| Critical value 5% | 1.30 | | 1.25 | | 1.15 | |
| **After 36 months** | | | | | | |
| **Gamma** | **Q-** | **p-** | **Q-** | **p-** | **Q-** | **p-** |
| | **Epanechnikov kernel matching** | | **Gaussian kernel matching** | | **Radius matching** | |
| Permanent contract | | | | | | |
| 1.00 | 1.93 | 0.026 | 1.93 | 0.026 | 1.93 | 0.026 |
| 1.25 | 1.21 | 0.112 | 1.21 | 0.112 | 1.21 | 0.112 |
| 1.50 | 0.63 | 0.263 | 0.63 | 0.268 | 0.63 | 0.268 |
| 1.75 | 0.14 | 0.444 | 0.14 | 0.444 | 0.14 | 0.444 |
| 2.00 | -0.03 | 0.511 | -0.03 | 0.511 | -0.03 | 0.511 |
| Critical value 5% | 1.10 | | 1.10 | | 1.15 | |

*Note*: Results achieved by using *mhbounds.ado* (Becker and Caliendo, 2007).
Critical values refer to the exact values of Gamma at which results turn insignificant at the 5% level.

From a public policy standpoint the short duration of the programme seems appropriate, given the absence of lock-in effects. However, we find rather disappointing long-term results. The more convincing explanation we can give for this fact is related to the difficulties of putting participants back into stable work in a context of high unemployment, as suggested by Sianesi (2008).[12] Probably in Spain, as in East Germany, ALMPs can certainly not solve the deep structural problems in the labour market. They may alleviate the symptoms, but cannot cure the disease, as emphasized by Lechner and Wunsch (2009b). Overall, the Spanish institutional rigidity constitutes a challenging environment for any ALMP, certainly worth continuing to research and evaluate.

## References

1. AMUEDO-DORANTES, C., MALO, M. y MUÑOZ-BULLÓN, F. (2008). "The role of temporary help agency employment on temp-to-perm transitions", *Journal of Labor Research*, 29:138-161.

2. ARELLANO, F. A. (2010). "Do training programmes get the unemployed back to work? A look at the Spanish experience", *Revista de Economía Aplicada*, forthcoming.

3. BECKER, S. O. y CALIENDO, M. (2007). "Sensitivity analysis for average treatment effects", *Stata Journal*, 7(1):71-83.

4. BECKER, S. O. y ICHINO, A. (2002). "Estimation of average treatment effects based on propensity scores", *Stata Journal*, 2:358-377.

5. BLUNDELL, R., COSTA DIAS, M., MEGHIR, C. y VAN REENEN, J. (2004). "Evaluating the employment impact of a mandatory job search program", *Journal of the European Economic Association*, 2(4):569-606.

12 Even though existing evidence on this topic seems to support the hypothesis of a clear positive relation between the effectiveness of the programmes and the unemployment rate over time (Lechner and Wunsch, 2009a; Kluve, 2010), the Spanish context in which the programme takes place is not one increasing unemployment (the financial crisis had not yet appeared), but one of persistently high unemployment rates.

6. CALIENDO, M. (2006). *Microeconometric evaluation of labour market policies.* Berlin, Springer-Verlag.

7. CALIENDO, M. y KOPEINIG, S. (2008). "Some practical guidance for the implementation of propensity score matching", *Journal of Economic Surveys*, 22:31-72.

8. CALIENDO, M. y KÜNN, S. (2011). "Start-up subsidies for the unemployed: Long-term evidence and effect heterogeneity", *Journal of Public Economics*, 95(3-4):311-331.

9. CALMFORS, L. (1994). "Active labour market policy and unemployment – a framework for the analysis of crucial design features", *OECD Economic Studies*, 22(1):7-47.

10. CARD, D., KLUVE, J. y WEBER, A. (2010). "Active labor market policy evaluations: A meta-analysis", *Economic Journal*, 120(548):452-477.

11. CARLINO, G. A. y SAIZ, A. (2008). City beautiful (Discussion Paper 3778). IZA. *http://ftp.iza.org/dp3778.pdf.*

12. CAVACO, S., FOUGERE, D. y POUGET, J. (2005). Estimating the effect of a retraining program for displaced workers on their transition to permanent jobs (Discussion Paper 1513). IZA.

13. CENTENO, L., CENTENO, M. y NOVO, Á. A. (2009). "Evaluating job-search programs for old and young individuals: Heterogeneous impact on unemployment duration", *Labour Economics*, 16(1):12-25.

14. CUETO, B. y MATO, J. (2009). "A nonexperimental evaluation of training programmes: Regional evidence for Spain", *Annals of Regional Science*, 43:415-433.

15. DAR, A. y TZANNATOS, Z. (1999). Active labor market programs: A review of the evidence from evaluations (Social Protection Discussion Paper 9901). World Bank.

16. DEHEJIA, R. y WAHBA, S. (2002). "Propensity score-matching methods for nonexperimental causal studies", *Review of Economics and Statistics,* 84:151-161.

17. EUROSTAT. (2009). *Regional statistics*. Disponible en *http://epp.eurostat.ec.europa.eu/portal/page/portal/region_cities/regional_statistics/data/main_tables* (accessed 20 January 2010).

18. GANGL, M. (2006). "Scar effects of unemployment: An assessment of institutional complementarities", *American Sociological Review*, 71:986-1013.

19. GARCÍA-PÉREZ, J. I. y REBOLLO SANZ, Y. F. (2009). "The use of permanent contracts across Spanish regions: Do regional wage subsidies work?", *Investigaciones Económicas*, 33:97-130.

20. HARDOY, I. (2005). "Impact of multiple labour market programmes on multiple outcomes: The case of Norwegian youth programmes", *Labour,* 19(3):425-467.

21. HECKMAN, J., ICHIMURA, H., SMITH J. y TODD, P. (1998). "Characterizing selection bias using experimental data", *Econometrica*, 66:1017-1098.

22. HECKMAN, J. J., LALONDE, R. J. y SMITH, J. A. (1999). "The economics and econometrics of active labor market programs", en O. Ashenfelter y D. Card (Eds.), *Handbook of Labor Economics* (vol. 3A, pp. 1865-2097). Amsterdam, North-Holland.

23. HOLLAND, P. W. (1986). "Statistics and causal inference", *Journal of the American Statistical Association*", 81:945-960.

24. KLUVE, J. (2010). "The effectiveness of European active labor market programs", *Labour Economics*, 17:904-918.

25. KLUVE, J. y SCHMIDT, C. (2002). "Can training and employment subsidies combat European unemployment?" *Economic Policy: A European Forum*, 35:409-443.

26. LECHNER, M., MIQUEL, R. y WUNSCH, C. (2007). "The curse and blessing of training the unemployed in a changing economy: The case of East Germany after unification", *German Economic Review*, 8:468-509.

27. LECHNER, M. y WUNSCH, C. (2009[a]). "Are training programs more effective when unemployment is high?", *Journal of Labor Economics*, 27(4):653-692.

28. LECHNER, M. y WUNSCH, C. (2009b). "Active labour market policy in East Germany", *The Economics of Transition*, 17(4):661-702.

29. MALO, M. A. y MUÑOZ-BULLÓN, F. (2006). "Employment promotion measures and the quality of the job match for persons with disabilities", *Hacienda Pública Española*, 179:79-111.

30. MATO, J. y CUETO, B. (2008). "Efectos de las políticas de formación a desempleados", *Revista de Economía Aplicada*, 46:361-383.

31. OECD. (2009). *OECD statistics*, Disponible en *http://stats.oecd.org/Index.aspx? DatasetCode=ALFS_SUMTAB*, (accessed 22 January 2010).

32. PLESCA, M. y SMITH, J. (2007). "Evaluating multi-treatment programs: Theory and evidence from the U. S. Job Training Partnership Act experiment", *Empirical Economics*, 32(2):491-528.

33. RAAUM, O. y TORP, H. (2002). "Labour market training in Norway - effect on earnings", *Labour Economics*, 9:207-247.

34. RAMOS, R., SURINACH, J. y ARTÍS, M. (2009). The effectiveness of regional active labour market policies to fight against unemployment: An analysis for Catalonia (Discussion Papers 4649). Institute for the Study of Labor (IZA).

35. ROSENBAUM, P. R. (2002). *Observational studies*. Nueva York, Springer.

36. ROSENBAUM, P. R. y RUBIN, D. B. (1983). "The central role of the propensity score in observational studies for causal effects", *Biometrika,* 70:41-55.

37. RUBIN, D. B. (1974). "Estimating causal effects of treatments in randomised and non-randomised studies", *Journal of Educational Psychology*, 66:688-701.

38. SAINT-PAUL, G. (2000). Flexibility vs. rigidity: Does Spain have the worst of both Worlds? (Discussion Paper 144). IZA.

39. SIANESI, B. (2004). "An evaluation o f the Swedish system of active labour market programs in the 1990s", *Review of Economics and Statistics*, 86:133-155.

40. SIANESI, B. (2008). "Differential effects of active labour market programs for the unemployed", *Labour Economics,* 15:370-399.

41. SMITH, J. (2000). "A critical survey of empirical methods for evaluating active labour market policies", *Swiss Journal for Economics and Statistics,* 136:1-22.

42. VERICK, S. (2009). Who is hit hardest during a financial crisis? The vulnerability of young men and women to unemployment in an economic downturn (Discussion Paper 4359). IZA. *http://www.politiquessociales.net/IMG/pdf/dp4359-3.pdf.*

43. WINTER-EBMER, R. (2006). "Coping with a structural crisis: Evaluating an innovative redundancy-retraining project", *International Journal of Manpower*, 27(8):700-721.

## Appendix

From the University of Seville, we are conducting a study for the Project Redes, in which you participated in 2004 (or 2005, as appropriate)

*for the treatment group*

From the University of Seville, we are conducting a study on the effectiveness of the Project Redes in promoting employability in Seville. Note that the data provided will be treated with confidentiality as required by law and in no case will be used for commercial purposes.

*for the control group*

1. Are you currently working?
2. (If the person is working) Employed or self-employed?
3. (If employed) Is it a permanent contract?
4. (If working) What is your monthly income?
5. (If not working) Do you study?, Are you actively seeking employment? Are you inactive?
6. *Only for treatment group:* How useful is for you today having completed of the Redes course? Has it helped in finding work? (Current perception of the usefulness of the course) Rate of 1 (none) to 5 (to a very high degree)