

EQUALITAS Working Paper No. 1

Evaluating social assistance reforms under programme
heterogeneity and alternative measures of success

Luis Ayala

Magdalena Rodríguez

2012

Evaluating social assistance reforms under programme heterogeneity and alternative measures of success

Luis Ayala,

Universidad Rey Juan Carlos, EQUALITAS

Magdalena Rodríguez

Instituto de Estudios Fiscales

Key words: welfare, poverty, programme evaluation, work-related activities

Corresponding author:

Luis Ayala

Facultad de Ciencias Jurídicas y Sociales

Universidad Rey Juan Carlos

Paseo de los Artilleros s/n

28032 Madrid, Spain

E-mail: luis.ayala@urjc.es

Abstract

This article presents an assessment of welfare reforms under a framework of programme heterogeneity and alternative measures of success. The study focused on a specific programme – Madrid’s *Ingreso Madrileño de Integración* (Madrid Regional Government’s Welfare Programme) (IMI) – which comprises heterogeneous sub-programmes. We tested whether work-related sub-programmes performed better than general activities aimed at improving life skills by analysing the effects on different types of outcomes. We also tried to identify which work-related sub-programme worked best. Our results show that intensive employment activities yield remarkably better results than general work-related schemes or life skills activities. However, increasing work participation does not automatically lift participants out of material hardship.

Introduction

In recent years, there has been a lively debate on the success of welfare reforms in terms of achieving better results in labour market participation and economic well-being. Most OECD countries have enacted major welfare reform legislation with the aim of increasing work incentives and reducing costs. The heavy emphasis on engaging recipients in work activities has rekindled interest in exploring whether these reforms have resulted in higher levels of well-being for participants. There has been much discussion over US policy changes with considerable empirical evidence on the relevant outcomes (for a synthetic overview, see Blank (2002), Grogger and Karoly (2005) and Moffitt and Ver Ploeg (2001)). Differential effects were found when considering the results of both work-first strategies – trying to push recipients into the labour market as rapidly as possible – and long-term programmes – focused on human capital developments providing intensive training and educational opportunities for recipients (Autor, Houseman & Pekkala Kerr, 2012; Bloom, Hill & Riccio, 2004; Dyke, Heinrich, Mueser & Troske, 2006; Moffitt, 2001; Riddell & Riddell, 2012).

Work-related reforms of welfare policies have also given rise to a considerable European literature. A huge range of experiences has already been assessed including extensive activation programmes in the field of Social Security and labour market policies in Nordic countries (Carling & Richardson, 2004; Forslund & Krueger, 2010; Sianesi, 2004, 2008), specific targeted welfare-to-work initiatives in the Netherlands (Koning, 2009; Van Oorschot, 2002), new policies focusing on low-income families with children combining welfare reforms with earned income tax credits in the United Kingdom (Blundell & Meghir, 2002; Gregg, Harkness & Smith, 2009; Lydon & Walker, 2005), welfare-to-work programmes and job search enforcement in Germany

(Huber, Lechner, Wunsch, & Walter, 2009) or ‘insertion contracts’ embedded in minimum income programmes in France and Spain (Ayala and Rodríguez, 2006; Terracol, 2009; Zoyem, 2001). In general terms, there is voluminous evidence showing that policy changes appear to have mattered.

Even though substantial research has been carried out, caution is important when evaluating these policies. Firstly, we have relatively little insight into the different effects the activities grouped under the notion of work-related welfare schemes have brought about. This is due to the heterogeneous nature of such activities. In many countries, these programmes are not mutually exclusive and welfare recipients can participate in different activities. This fact raises complex methodological issues the standard evaluation literature does not deal with.

Secondly, there is growing evidence that suggests that the effects of these programmes differ between outcomes. It can be difficult to see the whole picture of the programmes’ results if evaluation focuses exclusively on employment. While most of the new programmes were designed to move welfare recipients into the labour market, the ultimate goal of these policies is to improve the economic self-sufficiency of these households. In practice, the assessment of welfare reforms depends crucially on the indicators chosen to measure the programmes’ outcomes. As reviewed by Cancian & Meyer (2004), little systematic analysis deals with the question of whether different measures of success are capturing the same thing. The programmes’ success can vary greatly depending on whether independence from welfare or income poverty are used as outcome indicators instead of employment results (Blank, 2002; Grogger & Karoly, 2005). Studies in welfare evaluation have also paid special attention to the effects on material hardship. While some authors have found that low-wage working mothers experienced higher levels of hardship than welfare recipients (Danziger,

Heflin, Corcoran, Oltmans & Wang 2002; Edin & Lein, 1997), others concluded that material circumstances of single mother families improved modestly after welfare reform in the USA (Meyer & Sullivan, 2006; Schmidt & Danziger, 2012; Winship & Jencks, 2004).

Consequently, one of the outstanding challenges in the evaluation of welfare programmes is to analyse the effects of different and simultaneous treatments in work-related welfare reforms on different types of outcomes. This paper aims at presenting an assessment of welfare reforms under a framework of programme heterogeneity and alternative measures of success. We focus on a specific welfare programme – Madrid’s *Ingreso Madrileño de Integración* (*Madrid Regional Government’s Welfare Programme*; thereafter only IMI)– which comprises heterogeneous sub-programmes. In Southern Europe, active labour market policies usually embedded in new designs of welfare programmes coexist with a variety of initiatives aimed at promoting life skills. Therefore, recipients can simultaneously participate in very different actions and can be considered as treated, as they should take part in some specific activity.

Many reasons make this programme particularly valuable for a comparative evaluation of welfare reforms. Firstly, to the extent that benefit levels fall between the low payments in the USA and the generous ones that are made in Northern Europe, this analysis can provide an interesting comparison for other countries closer to the mid-range of these extremes. Secondly, the IMI approach to self-sufficiency seems less punitive than the US approach based on strict conditions, sanctions and time limits. Thirdly, while the context and programme details are different than those in the USA or the Nordic countries, there are remarkable similarities with the experience of

other Southern European countries. Our results could also apply to some extent to other regions of Spain and Southern Europe or other countries with similar conditions.

To measure the relative effectiveness of the different activities available for IMI recipients, we addressed two different questions. First, we tested whether work-related sub-programmes performed better than general activities aimed at improving life skills. Second, we tried to identify which work-related sub-programme worked best. We performed propensity score matching in a setting of heterogeneous treatments.

The structure of this article is as follows. The opening section summarises the particular design features of the IMI programme and the available data. The second section reviews different approaches to deal with the problem of evaluation in a framework of multiple states and alternative outcomes. The third section tests the extent to which the results are sensitive to alternative definitions of success by comparing the performance of the different treatments. Concluding remarks are presented in the final section.

The IMI programme

Institutional features of the programme

The programme analysed in this study is the Madrid Regional Government's Welfare Programme (IMI), which was set up in 1990. Unemployed individuals who have previously worked have access to two types of benefits: the contributory benefit and the assistance subsidy. In order to gain access to the first level, workers must have previously paid social contributions to cover this risk. The unemployment subsidy is for people whose contributory benefit has expired. Welfare schemes in Spain are designed for those individuals who have exhausted their rights to unemployment

benefits and who are not working. General risk of poverty is covered by regional schemes and welfare in Spain is completely decentralised. The Madrid Programme can be considered an ‘average’ programme within the complex set of regional schemes existing in Spain and Southern Europe. As in other European schemes, all households are entitled to IMI access if they have used up entitlement to other income maintenance programmes.

Previous studies have stressed that employability and belonging to an ethnic minority are the main determining factors leading to lengthened spells in the programme (Authors, 2007). These results show that there are different kinds of recipients depending on their ability to enter the labour market. Consequently, they need to be dealt with differently. Previous research has also provided information on the IMI’s recidivism determinants (Authors, 2010). These studies have found that activities to maximise the duration of the off-welfare spells should focus on implementing reforms that improve recipients’ chances of entering into more stable forms of employment and allocate a greater amount of resources to promote the ‘insertion’ of specific groups.¹ There is also evidence that participation in work-related activities reduces recidivism rates among IMI recipients (Authors, 2006).

Among the different institutional features of the programme, the ‘insertion activities’ represent the most prominent trait in a comparative framework. Once benefits are approved by the programme’s managers, recipients must sign an ‘insertion contract’ with the welfare agencies. Participation in these contracts is mandatory while recipients receive benefits. Initially, they are intended to improve the recipients’ self-sufficiency through an individualised design of ‘insertion’ activities

¹ Insertion is the general term used in these countries to summarise the different types of activities aimed at improving life and labour skills of welfare participants. In general terms, it means higher levels of social participation.

adjusted both to individual and households' characteristics. The idea of co-responsibility is at the heart of the programme. Individual assessment is conducted when recipients enter the programme and social services support is provided to help these households to address specific challenges. The contents of the contracts are negotiated by both sides resulting in a final plan of specific public intervention for each household.

Therefore, the programme's outcomes could be conditioned by the efficacy of caseworkers in allocating individuals to government programmes. Recent studies have stressed the key role of caseworkers in the assignment of welfare recipients to different work-related programmes (Behncke, Frölich & Lechner, 2009; Bloom, Michalopoulos, Hill, & Lei, 2003; Huber, Lechner, Wunsch, & Walter, 2009; Lechner & Smith, 2007). Most of the evidence shows that caseworkers matter in allocation decisions. Available data from our survey somewhat confirm that caseworkers also play a relevant role in assigning IMI recipients to the different treatments. Nevertheless, most respondents state that, while the final decision was taken by the caseworker, it was the result of a negotiation process.²

Every recipient must join a specific programme, bringing about a very different scenario than that depicted by the standard theory of programme evaluation. A very relevant issue is that recipients can simultaneously participate in different activities. A first set of activities includes overall actions developed to guarantee the basic preconditions of social participation. They consist of a variety of services

² Households were asked if the final assignment was mainly decided by the caseworker, the recipients, an agreement, an agreement under dominance of caseworkers or an agreement under dominance of recipients. Results did not show large differences in the observed frequencies between the different treatments. Almost one third of the assignments were the result of an agreement under caseworkers' predominance and approximately 30 per cent were made on the basis of a fifty-fifty agreement between caseworkers and recipients.

comprising such different topics as general life skills, family mediation, children's schooling and activities aimed at making it easier for some families to continue their daily routines or helping recipients recognise their strengths.

A second set of activities specifically aim at improving recipients' employment opportunities (labour skills). There are, first, various general services designed to improve the recipients' labour market opportunities; second, there are specific actions trying to push recipients into the labour market as soon as possible, including social enterprises and subsidised employment. The common purpose of these actions is the achievement of basic labour skills and the establishment of a friendly work environment as necessary first steps in the transition to competitive employment. Social enterprises are conducted by government agencies and non-profit organisations. These entities work with a variety of targeted populations, including long-term unemployed.

Data

In this study, we match the programme's administrative records – covering the whole history of the programme – with a specific survey conducted in 2001. This survey covers very different dimensions of the households' economic well-being some years after their participation in the programme. The merging of the two sources may prove successful in creating a comprehensive dataset for evaluation purposes. On the one hand, by examining administrative records, we have very detailed information on the recipients' characteristics at the moment of entering the programme. On the other hand, the IMI survey allows us to assess these households' economic well-being some time after participating in the different treatments. Furthermore, by matching administrative records with outcome variables, such as employment, income, and

living conditions, we can correct some potential biases in evaluation related to omitted information on the characteristics of previous welfare participation.

The examination of the programme's administrative records allows us to study a very diverse set of socioeconomic characteristics of IMI recipients. We have information on over 50,000 spells in the programme corresponding to 39,200 households. Some 8,500 of them had left the programme at some stage and then re-entered it at least once. Recipients' characteristics include some of the variables highlighted as ideal for analysing welfare populations, such as the existence of structural problems (social isolation, alcohol abuse and drug addiction) or the development of behaviour associated with marginal situations (prostitution or begging).

The survey of IMI recipients was conducted by the Madrid Government in 2001 including very detailed information on both participation in sub-programmes during their time in IMI and different dimensions of the current economic situation. The sample size of the survey is about 2,300 households, obtained by stratified random sampling from the programme's administrative records. The population of ex-welfare recipients was divided into four strata and a simple random sample was selected from each stratum.³ The variables used to define the strata were: date of entry, exit type, duration of IMI participation and town size.

The survey contains detailed information on participation in the different 'insertion' activities included in the IMI programme. The different sub-programmes considered in the survey are general information, general counselling, continuous individual support, psychological support, legal support, children intervention, family

³ The survey was conducted using as initial universe those households that had been in the programme at a given moment in the previous decade. Therefore, the sample may include both households who are still in the programmes and welfare leavers.

mediation, group activities, assistance to obtain other benefits, access to specific employment offers, general job search assistance, training, subsidised employment and social enterprises. There is also information on different dimensions of economic well-being, such as employment, subjective economic well-being, material hardship and social difficulties. We will use these dimensions to evaluate the outcomes of the programme.⁴ Some data on socioeconomic characteristics such as age, gender, household type, marital status, educational attainment and labour status are also collected in the dataset.

Methodology

The purpose of this evaluation was to identify which alternative packages of ‘insertion’ activities bring about substantial improvements in the participants’ economic well-being. The IMI survey provides very detailed information on fourteen different treatments, mixing activities aimed at upgrading life skills and work-related initiatives promoting labour-skills. For the purpose of this study, the different treatments were aggregated into four different and mutually exclusive groups taking into account whether recipients took part in work-related initiatives: non-participation in specific work-related sub-programmes (life-skills only) (n=811), participation in general work-related activities (n=594), participation in work-related intensive activities (n=113) and participation both in general work-related general activities and in work-related intensive activities (n=331). The latter group is not a combination of groups 2 and 3 but an independent category comprising those households that simultaneously take part in both activities. The first group includes the households

⁴ We used the survey data to identify treatments and outcomes. Socioeconomic information from administrative records of these households was used to estimate the probability of taking part in a given treatment. Therefore, the sample used in our identification strategy was 2,300 households.

that only took part in activities promoting life-skills. The other three are households that took part in different types of work-related activities (general, intensive or a combination of both).

The criteria used to aggregate participants into four different and mutually exclusive groups were to obtain a compact classification summarising the different strategies available for decision making. A first strategy is to provide only actions developed to guarantee the basic preconditions of social participation (life-skills). This group (treatment) comprises nine activities: general information, general counselling, continuous individual support, psychological support, legal support, children intervention, family mediation, assistance related to other social benefits, and group activities.⁵ The other groups or treatments represent strategies aimed at fostering transitions from welfare to work. As stated before, these actions can be classified as responding to long-term strategies – focused on human capital developments providing intensive training and educational opportunities for recipients – and work-first strategies which try to push recipients into the labour market as rapidly as possible. Therefore, a second group (treatment) includes general labour services for recipients (access to employment offers, general job search assistance and training). A third group (treatment) involves more intensive actions corresponding to work-first strategies (subsidised employment and social enterprises). The last group (treatment) refers to individuals taking part simultaneously in the two types of work-related activities. These households are not included in any of the previous cases. They receive a ‘complete package’ of work-related treatments.

TABLE 1 in about here

⁵ Individuals in Groups 2, 3 and 4 can simultaneously participate in any of the life-skills activities included in the first group. Those that combined work-related with life-skills activities are a small proportion of the total.

A descriptive analysis of the participants allowed us to give a preliminary assessment of possible differences in the characteristics of each group (Table 1). There were marked differences both among households not taking part in work-related activities and the other groups, and among the three types of participants in those activities. The data on age showed a larger presence of middle-aged individuals among households engaged in intensive work-related schemes. Concerning household type, the higher proportion of single-parent households taking part in general work-related activities stood out. Frequencies of recipients' educational attainment suggested a higher percentage of higher levels among households in more intensive and mixed work-related activities. These two groups were also the ones where households were more in touch with the labour market before entering into the programme. There were not, however, very many differences in employability levels, except for the higher proportion of individuals totally unfit for work in participants in life-skills activities. Finally, the second group accumulated a higher number of different social problems that accompany the lack of income. The main reason was the high proportion of households for whom belonging to an ethnic minority limited their possibilities of social integration.

To the extent that every recipient participated in one of the four defined groups, we focused on the relative effectiveness of each treatment. Among the relevant options for policymakers, three specific questions were addressed. First, we evaluated the effects of participation in some work-related scheme as compared with participation in general activities promoting life skills. Second, we assessed the effects of participation in each one of the specific work-related schemes – general, intensive and mixed – as compared with participation in general life-skills activities. Third, we

also examined the relative effectiveness of each specific work-related treatment. Table 2 summarises the pair-wise comparisons.

TABLE 2 in about here

A key question in the analysis of the different effects was the selection problem arising when treatment assignment is not random. A well-known problem of causal inference is how to estimate treatment effects in observational studies in situations where some individuals are exposed to a treatment, but with no methods of experimental design to obtain a control group. In this study we used propensity score matching estimators to build up a sample counterpart by pairing each participant with non-participant recipients with similar characteristics. As is widely known, a necessary assumption is conditional independence between non-treated outcomes and programme participation (Rubin, 1977).

The main limitations of matching were this assumption and that it relied on a sufficiently rich comparison group. As the number of observable covariates increased, it became more difficult to find exact matches for each of the treated units. In a seminal study, Rosenbaum and Rubin (1983) suggested the use of the probability of receiving treatment conditional on covariates (propensity score) to reduce the dimensionality of the matching problem. As stressed by Becker and Ichino (2002), if this balancing hypothesis is satisfied, observations with the same propensity score must have the same distribution of observable characteristics independently of treatment status.

However, most of the evaluation literature of welfare reforms using matching estimators rested on a basic framework in which a programme was administered at a fixed point in time, and individuals were either treated or not treated. For an adequate evaluation of the IMI programme, it was necessary to extend the standard binary

treatment model of only two states to the case of multiple states (Imbens, 1999; Imbens & Wooldridge, 2009; Lechner, 2002ab; Sianesi, 2008). Given a framework of $(M+1)$ mutually exclusive sub-programmes (treatments), every individual will have one observable outcome $\{Y_0, Y_1, \dots, Y_M\}$. Participation in one of the predefined mutually exclusive sub-programmes is indicated by $S \in \{0, 1, \dots, M\}$. We were interested in the effects of participation in one sub-programme (a) compared with participation in other sub-programme (b) for individuals who took part in a :

$$\tau^{a,b}_0 = E(Y^a - Y^b | S=a) = E(Y^a | S=a) - E(Y^b | S=a)$$

where $\tau^{a,b}$ represents the expected effect for a welfare recipient randomly drawn from the group participating in sub-programme a . As in the case of the standard binary treatment, we needed a counterfactual to estimate $E(Y^b | S=a)$.

Under the assumption that conditional independence would hold also in the multiple-states framework, evaluation required observing all the characteristics (X) of the programme's recipients affecting both the probability of participation in the respective sub-programmes and the outcome variables. All participants in sub-programme a needed to have a counterpart in group b for each X (Imbens, 1999). We selected from the participants in b a control group whose distribution of observed variables was as similar as possible to the distribution in the group of participants in sub-programme a .

For an adequate comparison of the outcomes of the different sub-programmes, we needed a balancing score function ($g(X)$) of the recipients' characteristics. Because we were interested in the pair-wise comparisons of the different sub-programmes, we needed to find a balancing score ensuring the balancing of the X 's in the two subpopulations of interest for each comparison:

$$E[Pr(S=a|X, S \in \{a,b\}) | g(X)] = Pr(S=a|X, S \in \{a,b\}) \equiv P^{a|ab}$$

We estimated the respective propensity scores in each one of the comparisons. As discussed by Lechner (2002b), two different approaches can be used to modelling the respective propensity scores for matching with multiple pair-wise comparisons. One approach consists of specifying and estimating a multiple discrete-choice model, such as multinomial logit or probit model (structural approach). A second approach is estimating all conditional probabilities between possible pairs of choices directly (reduced-form approach). This second approach closely mirrors the usual propensity score approach for binary treatments. As we have defined seven pair-wise comparisons the reduced-form approach is not so prohibitive than when using more disaggregated analysis. Additionally, the problem of using the structural approach is that if one choice equation is misspecified all conditional probabilities could be misspecified.

A relevant question was the selection of Xs for balancing the different subsamples in each pair-wise comparison. The resulting quality of the matched samples informed our choice. We used the administrative records to select the characteristics necessary to estimate the propensity score. The covariates were the number of social problems, single-parenthood, educational level, unemployment rate at entry, household size, number of children, single persons and gender – they all were measured at the moment of entering the programme. To compare sub-programme *a* and sub-programme *b*, each participant in the former group was matched to one participant in group *b* based on the balancing score. Different procedures were selected for associating the sets of participants. The results we present below were obtained with nearest neighbour matching estimators without replacement. We carried out different sensitivity analyses with other estimators finding similar results⁶

⁶ Results are available upon request.

Estimates of the effects of ‘insertion’ activities on the recipients’ economic well-being by using propensity score estimators were only reliable if the matching produces credible control groups. There was a common support requirement for all pair-wise comparisons. Figure 1 plots the different density distributions of the propensity score for each comparison. Overlap in compared propensity scores regions seemed to ensure common support across treatment groups. Even though there were slight differences between the seven comparisons – the matching was especially high in evaluations 1, 4 and 6 – the fit was reasonable. The crucial issue of the overlap condition seemed to hold in our estimates.

FIGURE 1

A second general question was the definition of outcomes. As stated above, results can be highly sensitive to the dimensions chosen for the assessment of economic well-being. A large literature has focused on the link between participation in work related activities and employment. Nevertheless, although most of the ‘insertion’ activities aimed to foster transitions from welfare to work, employment activities did not always account for changes in the opportunities for self-sufficiency. Other dimensions garnered a great deal of research attention over the years. Subjective well-being, poverty or material hardship could be as relevant as earnings or working hours’ indicators.

We tried to collect a variety of results corresponding to these different dimensions. However, data constraints made it difficult to obtain adequate measures for all possible outcomes. Subjected to these restrictions, we chose three different dimensions to evaluate the relative effectiveness of the different sub-programmes under study. Although there are some limitations for a complete assessment of the

outcomes in each dimension, the survey allowed us to draw a comprehensive picture of results.

The first dimension focused on employment outcomes and two indicators were used: (1) whether the household head was currently employed, and (2) whether there was a legal contract and payment of employer contributions. The IMI survey does not provide information on households' incomes. The survey, however, comprises a set of questions allowing the measurement of subjective well-being. We defined two dummy indicators for this dimension: subjective poverty and changes in living standards as compared with ten years ago (self-assessed). Possible interpretations of the latter measure must be made cautiously as changes in macroeconomic conditions and government support might affect this result.

A third dimension comprised different indicators of material hardship. Most of them were related to housing conditions. The conditions were running water, hot running water, having electricity, having gas, inside toilet with running water, bath or shower, wash basin, kitchen, oven, refrigerator, washing machine, dishwasher, heating, telephone, mobile phone, car and van. Given the different nature of these indicators, it seemed clear that all the items considered carried a different weight in the households' economic well-being. Arithmetic addition implicitly imposes a severe value judgment because it does not differentiate the weighting of each material condition or necessity. The different conditions were summarised into an index of material hardship:

$$Z_i = \sum_j w_j z(d_{ij})$$

or an index of living standards:

$$L_i = \sum_j w_j l(d_{ij})$$

where $z(\cdot)$ and $l(\cdot)$ are non-increasing and non-decreasing functions, respectively, of the amount d_{ij} possessed by the i th household ($i=1, \dots, n$) of the j th attribute ($j=1, \dots, J$), and w_j is the corresponding weight. While some authors give an equal weight to each item, the interpretation of deprivation as a relative situation is at the core of this line of research. One of the most common strategies is to apply weighting systems which give more importance to the goods most widely owned in a society. For the material hardship index we used a normalised weight calculated as:

$$w_j = \frac{v_j}{\sum_j v_j}$$

where v_j is the number of households not lacking item j in the survey. Thus, the weights attached to each item are functions of the spread of the good among the whole population, compared with the spread of the other goods or activities considered. The commodities most people in society enjoy were given more weight. We also defined an index of material well-being using the complementary options. An additional measure provided information on the number of housing problems. Housing difficulties included inadequate housing conditions, overcrowding, excessive spending, and non-payment of dwelling.

Results

To measure the relative effectiveness of the different sub-programmes, we estimated average effects for each pair-wise comparison and for the three types of indicators we defined. We addressed two different questions in each case. First, we tested if work-related sub-programmes performed better than general activities aimed at improving life skills. Second, we tried to identify which work-related sub-programme worked best. In each case, we estimated an unconditional estimator which was 100 times the

treatment B group mean minus the treatment A group mean divided by the treatment A group mean.

One of the main thrusts of the IMI's development has been the provision of skills to welfare recipients so that they are closer to the labour market. Considering the achievement of higher employment rates as a central goal, we might expect a better performance of work-related activities as compared to life-skills activities. The main result shown in Table 3 is that employment effects were substantially higher for those sub-programmes aimed at improving labour opportunities. Although the approaches to foster transitions from welfare to work differed among the different sub-programmes evaluated, they all brought about substantially higher employment rates than general life-skills activities did.

TABLE 3 in about here

While results showed that participation in work-related activities was associated with better employment outcomes, it is interesting that the three kinds of treatments – general activities, direct employment activities and mixed strategies – presented different effectiveness levels. Among the three different options, intensive labour policies intended for welfare recipients stood out as those with the highest capacity to increase employment opportunities. It seems that it was even better to assign most of the available resources to direct activities promoting employment than to combine these actions with very general job assistance activities. It could be the case that placing welfare recipients – especially those hardest to employ – into very different daily labour routines could reduce the programme's efficiency.

If attention is focused on more specific indicators of employment success, the evidence is somewhat mixed. Our second binary indicator showed the possibility of being employed and having a legal contract covered by employer's contributions. The

most relevant result was that very general labour activities were not enough to guarantee a stable position in the labour market. They produced very modest positive effects as compared with general life-skills activities. Furthermore, their results seemed very limited compared with those produced by training or mixed strategies. The second group of activities appeared to be the most suitable strategy in this case.

As stated above, employment results can be considered as intermediate outcomes. While most of the sub-programmes analysed aimed to move welfare recipients into work, the final objective of these policies was to achieve more direct effects on the households' economic well-being. For this reason, it was important to test whether participation in the different sub-programmes helped to improve the economic situation of these households. Among other dimensions, income poverty was at the heart of the mainstream approach to identifying economic well-being in welfare reform.

Subjective poverty indicators are given in Table 4 for each pair-wise comparison. Several points are worth mentioning. First, it is interesting to note that there were substantial differences between the treatments under study. As a rule, work-related activities not only contributed to higher employment rates, but also seemed to reduce income poverty measured on a subjective base. However, results were not significant. Secondly, intensive employment programmes appeared to be the most efficient policy, especially if they were not combined with more general labour-related activities. Results for the latter showed that welfare-to-work sub-programmes not resting on specific forms of subsidised employment or social enterprises did not produce substantially better results than activities aimed to make it easier for some families to continue with their daily routines.

TABLE 4 in about here

The survey also enabled us to carry out intertemporal analysis. More precisely, households were asked to report changes in their economic situation over the last decade. We defined a dummy variable representing upward variation in self-assessed income. It might be expected that households moving from welfare to employment would report improvements in their economic situation. A virtuous circle of getting employed, higher earnings and increasing disposable income could take hold. Furthermore, previous studies found a significant relationship between unemployment and subjective income insecurity in Spain (Ayala, Martínez & Ruiz-Huerta, 1999).

Somewhat mixed evidence came from the estimated effects for the different treatments. As a rule, the range of variation of the estimated average effects was narrow and not significant. Work-related activities did not seem to have substantial effects on the intertemporal changes of economic well-being among welfare participants. In any case, we should be especially cautious about possible outcomes resulting from self-reported income. As stated by Ravallion and Lokshin (2002), concerns about measurement errors and the influence of latent psychological factors on observed respondent characteristics limit the validity of the well-being inferences drawn from answers to subjective survey questions.

A third dimension for the assessment of the different sub-programmes' performance was material well-being. As discussed above, we combined indicators of material hardship to create composite measures of material deprivation and living standards. We assigned weights according to the proportion of households lacking (or not lacking) the respective item. The idea was that the higher the proportion of households with a particular item, the greater the extent to which the item may be deemed to be a necessity. Table 5 presents results for both indicators. A comparison across treatments showed that participation in work-related activities did not sharply

alter the households' levels of material well-being. Average effects were rather small in most cases, suggesting that participation in work-related schemes had not reduced material hardship. A similar result was also found for the measure summarising the number of housing problems. Work-related programmes seemed to produce worse results than life-skills activities did.

TABLE 5 in about here

One possible explanation for the little impact of targeted welfare-to-work programmes on material well-being is that increasing work participation does not automatically lift out of material hardship. Family income may not increase significantly and structural forms of deprivation, like housing, remain unresolved. However, this conclusion should be treated cautiously. As discussed by Winship and Jencks (2004), welfare leavers who have experienced serious problems of material hardship are hard to find and the representativeness of hardship variables in the survey could be limited. Additionally, it must be noted that composite or summary measures provide additional information on the concurrence of various hardships, but are at risk of obscuring detail in the individual components.

Results also seem to support the idea that participation in work-related activities would lead to larger reductions in poverty (self-assessed) than in material hardship. This conclusion is in keeping with well-known prior empirical evidence for Spain. Past research has shown only a very moderate association between poverty and hardship measures, both considered in static and dynamic perspective (Pérez-Mayo, 2005).

While the results showed that participation in work-related activities was associated with limited effects on material well-being, it is important to note again that the three labour treatments gave rise to different outcomes. Participation in

intensive work-related activities turned out to be more effective for reducing material hardship than involvement in general work-related activities. The average effects were small and in most cases not significant.

These findings can be very helpful for a better understanding of the programme evaluated. In practice, programme heterogeneity can cause a variety of results depending on the variable chosen as outcome. In this sense, a final matrix of treatments and outcomes can help to assess the overall effects of a huge range of options. It can provide useful feedback on whether the different sub- programmes are generating impacts consistent with long-term expectations. Policy-makers can also choose different input combinations depending on political priorities. If the main goal of the programme is to improve employment for welfare recipients, there is no doubt that some alternatives do indeed work better than others. If the priority is to minimise material hardship, the matrix allows us to identify which combination of sub-programmes yields better results.

TABLE 6 in about here

Table 6 shows how the effects of the different sub-programmes largely depended on the outcome variable. Nonetheless, we found some evidence that could help to clarify the available array of options to policy-makers. One might expect employment and material well-being to be positively correlated, so that engaging welfare participants in work-related activities would improve other dimensions of these households' economic well-being. However, our results called into question the traditional view that transitions from welfare to work bring about improvements in the different dimensions of economic well-being. This implies that there is not a universal solution to the different problems posed by welfare populations. Policymakers must frequently make hard decisions subject to very complex restrictions.

In any case, the multiple states/multiple outcomes matrix can be a useful tool for handling different options. The corresponding analyses of files and columns can be powerful tools for evaluating the various options that often confront programmes' managers in deciding the way ahead. If they have to decide between fostering participation in work-related sub-programmes or general activities promoting life skills, the matrix showed that, for the most part, the former activities had a positive effect on employment and poverty without harming substantially material well-being. Therefore, if employment indicators are used to measure the programmes' success, participation in work-related activities should be encouraged. If the main objective of 'insertion' activities is improving living standards, the options are not so well defined.

Regarding employment activities, there is an additional binary choice. Once the determination to implement work-related activities becomes the policies' guideline, a decision must be made on the best way to promote employment and economic well-being. The overarching finding from the study was that intensive employment activities, such as subsidised employment or engagement in social enterprises, yielded remarkably better results than more general work-related schemes or the combination of both strategies.

Conclusion

Major policy changes have increased interest in outcomes for participants in welfare reforms. In most countries, the main goal of the enacted reforms has been to reduce the dependence of low-income households on government support by improving employment opportunities while continuing to maintain a social safety net for qualified families. This article has assessed the effects of Madrid's *Ingreso Madrileño de Integración* (IMI). This programme is standard among the existing schemes in

Southern Europe. Compared with other welfare models, the development of heterogeneous sub-programmes providing different ‘insertion’ services is the main difference it presents. The merging of two different datasets – administrative records and a survey – together with the use of pair-wise comparisons and a huge set of outcome variables allow us to estimate the average effects of the different sub-programmes.

The picture emerging from the different evaluations is generally consistent. Nevertheless, conclusions need to be drawn cautiously because of the natural limits of these evaluation approaches. For policymakers, work-related activities appear attractive at first sight, as they imply a concentration of resources to reduce welfare participation and improve employment. The results of this study, however, lead us to caution against drawing oversimplified conclusions. Empirical evaluation suggests that the effects of the different ‘insertion’ activities are rather varied depending on the outcome variables. By increasing employment levels, work-related activities seem successful in moving low-income families away from dependence on welfare policies. The long-run impact of these changes on economic well-being could be positive. Work today should raise experience tomorrow and raise future employability. However, higher labour participation does not seem enough to enable low-income families to achieve better results in terms of material well-being. If the overriding goal of social policy is to reduce material deprivation and social hardship, there is no doubt that work-related policies are not entirely suitable. In any case, strict assessments on the validity of these activities could be misleading.

Among the different work-related sub-programme options, intensive employment activities, such as subsidised employment or engagement in social enterprises, yield remarkably better results than general work-related schemes. They

result in higher levels of employment and subjective well-being. In terms of public intervention, however, this finding should be taken with caution. Given the high levels of heterogeneity among welfare populations, possible inferences should be restricted to certain households. In practice, work-first strategies can only be a solution for a segment of the recipients' population. For those unemployable, an upgrading of life skills through specific non-labour related interventions should result more efficient.

Our findings provide new evidence to address some of the central questions the current welfare reform debate has raised. As other countries are discussing similar reforms, our results could contribute to examine the welfare experiments in Southern Europe with greater interest. New evidence on approaches that consider heterogeneous sub-programmes and different types of outcomes might inform and partially shape the future public policy agenda in the welfare reform debate.

Acknowledgements

The authors thank Jeffrey Grogger, Barbara Sianesi, Tim Smeeding and participants at seminars in Ann Arbor, Atlanta, Badajoz, Barcelona, Madrid, Málaga, Oviedo, Murcia, Sevilla, Valencia and Zaragoza for their helpful comments. Luis Ayala acknowledges financial support from the *Ministerio de Ciencia e Innovación* (ECO2010-21668-C03-01).

References

- Ayala, L., Martínez, R., & Ruiz-Huerta, J. (1999). Poverty and Social Policy in Spain in a Period of Political Transition. In: Hutton, S. & Redmond, G. (Eds.) (1999): *Poverty in Transition Economies* (pp. 208–227) London: Routledge.
- Ayala, L. & Rodríguez, M. (2006). The Latin Model of Welfare: Do ‘Insertion Contracts’ Reduce Long-Term Dependence?. *Labour Economics*, 13(6), 799–822.
- Ayala, L. & Rodríguez, M. (2007). Barriers to Employment and Welfare Dynamics: Evidence from Spain. *Journal of Policy Modeling*, 29(2), 237–257.
- Ayala, L. & Rodríguez, M. (2010). Explaining welfare recidivism: what role do unemployment and initial spells have? *Journal of Population Economics*, 23(1), 373–392.
- Autor, D. H., Houseman, S. N., & Pekkala Kerr, S. (2012). *The Effect of Work First Job Placements on the Distribution of Earnings: An Instrumental Variable Quantile Regression Approach*. NBER Working Paper 17972.
- Becker, S. O. & Ichino, A. (2002). Estimation of Average Treatment Effects Based on Propensity Scores. *The Stata Journal*, 2(4), 358–377.
- Behncke, S., Frölich, M., & Lechner, M. (2009). Targeting Labour Market Programmes. Results from a Randomized Experiment. *Swiss Journal of Economics and Statistics* 145(3): 221-268.
- Blank, R. M. (2002). Evaluating Welfare Reform in the United States. *Journal of Economic Literature* 40(4): 1105–1166.
- Bloom, H. S, Hill, C. J., & Riccio, J. A. (2003). Linking Programme Implementation and Effectiveness: Lessons from a Pooled Sample of Welfare-to-Work Experiments. *Journal of Policy Analysis and Management* 22(4): 551-575.
- Bloom, H. S., Michalopoulos, C., Hill, C. J., & Lei, Y. (2004). Can Nonexperimental Comparison Group Methods Match the Findings from a Random Assignment Evaluation of Mandatory Welfare-to-Work Programmes? *The Review of Economics and Statistics* 86(1): 156–179.
- Blundell, R. & Meghir, C. (2002). Active Labour Market Policy vs. Employment Tax Credits: Lessons from Recent UK Reforms. *Swedish Economic Policy Review* 8: 13-37.

- Cancian, M. & Meyer, D. R. (2004). Alternative Measures of Economic Success among TANF Participants: Avoiding Poverty, Hardship, and Dependence on Public Assistance. *Journal of Policy Analysis and Management* 23(3): 531–548.
- Carling, K. & Richardson, K. (2004). The Relative Efficiency of Labor Market Programmes: Swedish Experience from the 1990s. *Labour Economics* 11(3): 335–354.
- Danziger, S., Heflin, C., Corcoran, M., Oltmans, E., & Wang, H. (2002). Does It Pay to Move from Welfare to Work?. *Journal of Policy Analysis and Management* 21(4): 671-692.
- Dyke, A., Heinrich, C. J., Mueser, P. R., Troske, K. R. (2006). The Effects of Welfare-to-Work Programme Activities on Labor Market Outcomes. *Journal of Labor Economics* 24(3): 567-608.
- Edin, K. & Lein, L. (1997). *Making ends meet: how single mothers survive welfare and low-wage work*. New York: Russell Sage Foundation.
- Forslund, A. & Krueger, A. (2010). Did Active Labor Market Policies Help Sweden Rebound from the Depression of the Early 1990s. In R. B. Freeman, B. Swedenborg, & R. Topel (Eds.), *Reforming the Welfare State: Recovery and Beyond in Sweden* (pp. 159–187). Massachusetts: National Bureau of Economic Research.
- Gregg, P., Harkness, S., & Smith, S. (2009). Welfare Reform and Lone Parents in the UK. *Economic Journal* 119(535): 38-65.
- Grogger, J., Karoly, L. A. (2005). *Welfare Reform: Effects of a Decade of Change*. Cambridge, MA: Harvard University Press.
- Huber, M., Lechner, M., Wunsch, C., & Walter, T. (2009). Do German Welfare-to-Work Programmes Reduce Welfare and Increase Work?. IZA Discussion Paper n° 4090.
- Imbens, G. W. & Wooldridge, J. M. (2009). Recent Developments in the Econometrics of Programme Evaluation. *Journal of Economic Literature* 47(1): 5-86.
- Koning, J. (2009): Reforms in Dutch active labour policy during the last 20 years: an evaluation. SEOR Working Paper n°2009/2.
- Lechner, M. (2002a). Some Practical Issues in the Evaluation of Heterogeneous Labor Market Programmes by Matching Methods. *Journal of the Royal Statistical Society Series A* 165(1), 59-82.

- Lechner, M. (2002b). Programme Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labor Market Policies. *The Review of Economics and Statistics* 84(2): 205–220.
- Lechner, M. & Smith, J. (2007). What Is the Value Added by Caseworkers?. *Labour Economics* 14(2): 135-151.
- Lydon, R. & Walker, I. (2005). Welfare to work, wages and wage growth. *Fiscal Studies* 26(3): 335-370.
- Meyer, B. D. & Sullivan, J. (2006). Consumption, income, and material well-being after welfare reform. NBER Working Paper Series n°11976.
- Moffitt, R. (2001). From Welfare to Work: What the Evidence Shows? The Brookings Institution, Policy Brief 13/2002.
- Moffitt, R. & Ver Ploeg M (eds.) (2001). *Evaluating Welfare Reform in an Era of Transition*. Washington DC, National Academy Press.
- Pérez-Mayo, J. (2005). Identifying deprivation profiles in Spain: a new approach. *Applied Economics* 37(8): 943-955.
- Ravallion, M. & Lokshin, M. (2002). Identifying Welfare Effects from Subjective Questions. *Economica* 68(271): 1453-1473.
- Riddell C, Riddell WC (2012). The Pitfalls of Work Requirements in Welfare-to-Work Policies: Experimental Evidence on Human Capital Accumulation in the Self Sufficiency Project. IZA DP No. 6378
- Rosenbaum, P. R. & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70(1): 41–55.
- Rubin, D. (1977). Assignment to a Treatment Group on the Basis of a Covariate. *Journal of Educational Statistics* 2(1): 1–26.
- Schmidt, L. & Danziger, S. (2012). Filling Holes in the Safety Net? Material Hardship and Subjective Well-being among Disability Benefit Applicants and Recipients after the 1996 Welfare Reform”. *Social Science Research* (forthcoming).
- Sianesi, B. (2004). An Evaluation of the Swedish System of Active Labor Market Programmes in the 1990s. *The Review of Economics and Statistics* 86(1): 133–155.
- Sianesi, B. (2008). [Differential effects of active labour market programmes for the unemployed](#). *Labour Economics* 15(3), 370-399.
- Terracol, A. (2009): Guaranteed Minimum Income and Unemployment Duration in France. *Labour Economics* 16: 171-182

- Van Oorschot, W. (2002). Miracle or Nightmare? A Critical Review of Dutch Activation Policies and their Outcomes. *Journal of Social Policy* 31(3): 339–420.
- Winship, S. & Jencks, C. (2004). How Did the Social Policy Changes of the 1990s Affect Material Hardship among Single Mothers? Evidence from the CPS Food Security Supplement. John F. Kennedy School of Government n°04-027.
- Zoyem, J. P. (2001). Contrats d'Insertion et Sortie du RMI. *Économie et Statistique* 346-347: 75–102.

Figure 1. Common support.

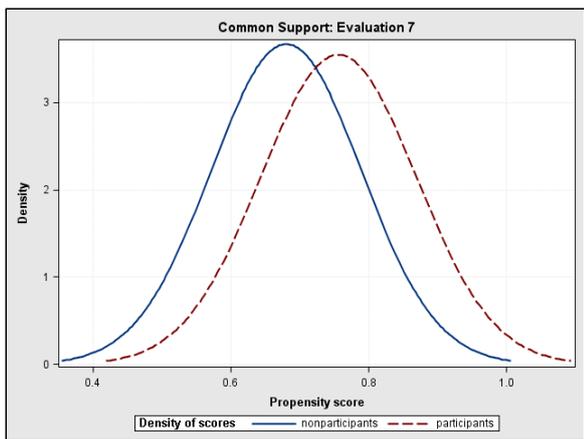
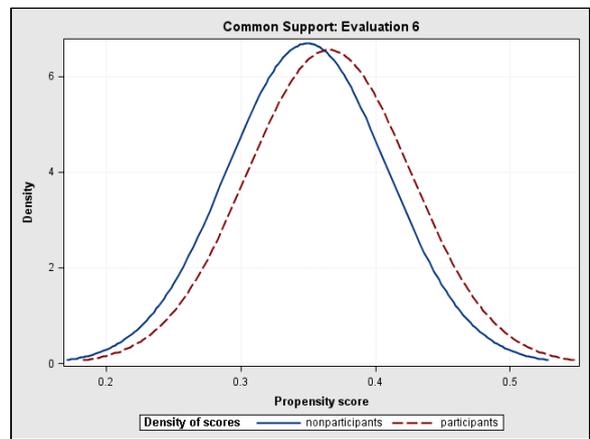
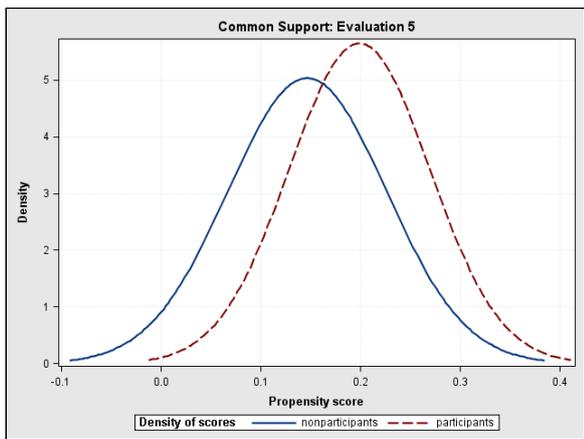
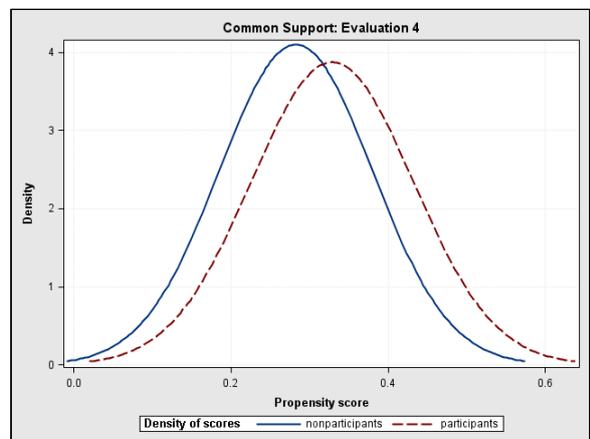
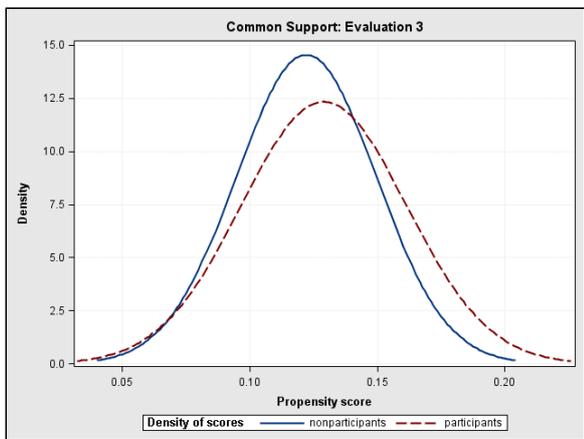
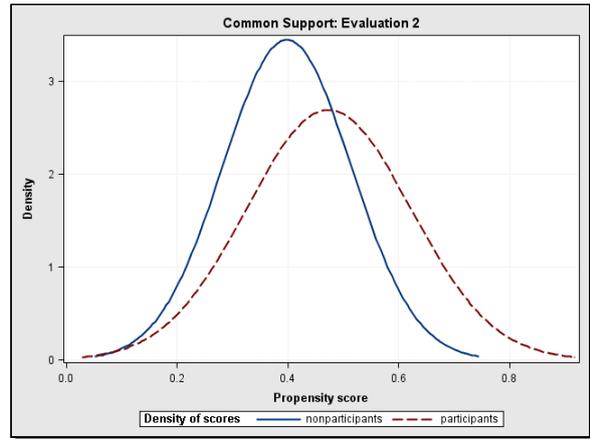
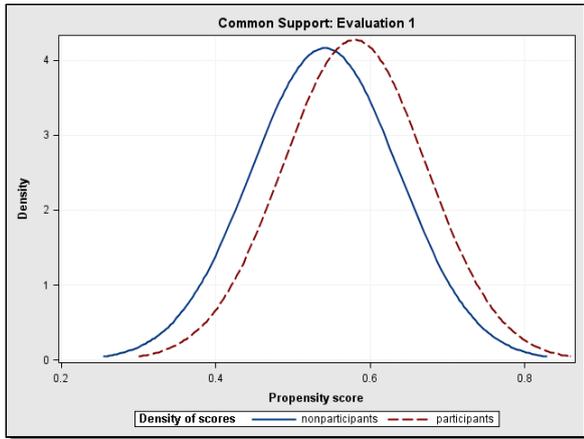


Table 1. Socio-economic characteristics of participants in each treatment (*frequency distribution*).

	<i>Life-skills only</i>	<i>General work-related schemes</i>	<i>Intensive work-related schemes</i>	<i>Mixed work-related schemes</i>
AGE				
<26	5.6	11.5	3.4	7.9
26-35	23.8	28.9	27.6	31.0
36-45	24.2	30.4	34.5	36.2
46-55	21.9	18.5	25.3	18.3
56-65	24.5	10.7	9.2	6.6
GENDER				
Males	34.9	29.9	34.5	35.8
Females	65.1	70.1	65.5	64.2
HOUSEHOLD TYPE				
Single person	24.9	21.1	20.3	26.6
Single-parent household	30.2	37.2	28.3	32.6
Other households with children	18.5	21.7	25.7	19.3
Other households without children	26.4	20.0	25.7	21.5
EDUCATION				
Does not read or write	10.6	12.2	8.1	7.4
No academic qualifications (only reads and writes)	25.4	19.9	18.9	19.3
Low	38.9	41.2	37.8	36.8
Medium	20.6	24.2	31.5	33.1
High	4.5	2.5	3.6	3.4
LABOR FORCE STATUS				
Employed	16.0	15.6	14.3	15.7
Unemployed	56.4	59.5	74.1	69.2
Inactive	27.6	24.9	11.6	15.1
EMPLOYABILITY				
Totally unfit for normal work	12.3	5.3	3.3	3.1
Low	26.8	28.8	27.8	28.2
Medium	14.6	23.7	24.4	23.8
High	46.3	42.2	44.4	44.9
NUMBER OF SOCIAL PROBLEMS				
0	34.6	29.8	42.3	34.3
1	43.7	44.7	39.2	45.0
2	15.9	19.1	13.4	15.1
3	4.5	4.9	3.1	4.1
4 or more	1.3	1.5	2.0	1.5
SOCIAL PROBLEMS ¹				
Drug abuse	5.0	6.6	6.2	6.9
Alcohol abuse	5.4	4.7	1.0	6.6
Non-payment of dwelling, eviction	8.5	8.9	9.3	7.9
Debt accumulation, non-payment	12.1	13.5	12.4	10.7
Beggary	0.3	0.8	0.0	1.6
Prostitution	0.8	1.5	0.0	0.6
Social isolation	14.1	14.6	14.4	17.0
Ethnic minority	15.0	21.4	13.4	14.5
Number of observations	811	594	113	331

¹The categories appearing in social problems are non-excluding dummy variables. A household can therefore suffer from more than one problem. The figures show percentages of recipients affected by each problem.

Table 2. Comparison between Treatment A and Treatment B.

<i>Evaluation</i>	<i>Treatment A</i>	<i>Treatment B</i>
1	Participation in a work-related scheme	Non-participation in a work-related scheme
2	Participation in general work-related schemes	Non-participation in a work-related scheme
3	Participation in intensive work-related schemes	Non-participation in a work-related scheme
4	Participation in mixed work-related schemes	Non-participation in a work-related scheme
5	Participation in intensive work-related schemes	Participation in a general work-related scheme
6	Participation in mixed work-related schemes	Participation in a general work-related scheme
7	Participation in mixed work-related schemes	Participation in an intensive work-related schemes

Table 3. Employment effects (PS matching estimates).

TREATMENT	OUTCOME VARIABLES	
	Employment (currently)	Legal contract and payroll taxes
Non-participation in a work-related scheme	0.262 (0.440) ¹	0.382 (0.486)
Participation in a work-related scheme	0.292 (0.455)	0.468 (0.499)
<i>Average effect</i>	11.5*	22.5**
Non-participation in a work-related scheme	0.244 (0.430)	0.382 (0.486)
Participation in general work-related schemes	0.256 (0.437)	0.393 (0.489)
<i>Average effect</i>	4.9	2.9
Non-participation in general work-related schemes	0.269 (0.444)	0.407 (0.492)
Participation in intensive work-related schemes	0.402 (0.493)	0.563 (0.499)
<i>Average effect</i>	49.4**	38.3**
Non-participation in a work-related scheme	0.244 (0.430)	0.382 (0.486)
Participation in mixed work-related schemes	0.296 (0.457)	0.580 (0.495)
<i>Average effect</i>	21.3*	51.8***
Participation in general work-related schemes	0.240 (0.427)	0.400 (0.490)
Participation in intensive work-related schemes	0.385 (0.489)	0.500 (0.503)
<i>Average effect</i>	60.4**	25.0*
Participation in general work-related schemes	0.265 (0.442)	0.412 (0.493)
Participation in mixed work-related schemes	0.313 (0.464)	0.558 (0.497)
<i>Average effect</i>	18.1	35.4***
Participation in intensive work-related schemes	0.376 (0.487)	0.473 (0.502)
Participation in mixed work-related schemes	0.296 (0.457)	0.577 (0.495)
<i>Average effect</i>	-21.3	22.0*

¹ Standard deviation in brackets.

***Significant at 99%, **Significant at 95%, *Significant at 90%.

Table 4. Effects on poverty and subjective well-being (PS matching estimates).

TREATMENT	OUTCOME VARIABLES	
	Subjective poverty	Economic situation compared with 10 years ago
Non-participation in a work-related scheme	0.366 (0.482) ¹	0.237 (0.425)
Participation in a work-related scheme	0.339 (0.474)	0.251 (0.434)
<i>Average effect</i>	-7.4	5.9
Non-participation in a work-related scheme	0.375 (0.484)	0.226 (0.418)
Participation in general work-related schemes	0.393 (0.489)	0.256 (0.437)
<i>Average effect</i>	4.8	13.3
Non-participation in general work-related schemes	0.312 (0.464)	0.243 (0.429)
Participation in intensive work-related schemes	0.282 (0.453)	0.230 (0.423)
<i>Average effect</i>	-9.6	-5.3
Non-participation in a work-related scheme	0.375 (0.485)	0.226 (0.418)
Participation in mixed work-related schemes	0.300 (0.459)	0.225 (0.419)
<i>Average effect</i>	-20.0**	-0.4
Participation in general work-related schemes	0.393 (0.489)	0.251 (0.434)
Participation in intensive work-related schemes	0.277 (0.450)	0.260 (0.441)
<i>Average effect</i>	-29.5**	3.6
Participation in general work-related schemes	0.370 (0.483)	0.268 (0.443)
Participation in mixed work-related schemes	0.299 (0.458)	0.229 (0.421)
<i>Average effect</i>	-19.2**	-14.6
Participation in intensive work-related schemes	0.326 (0.471)	0.226 (0.420)
Participation in mixed work-related schemes	0.301 (0.460)	0.225 (0.418)
<i>Average effect</i>	-7.7	-0.4

¹ Standard deviation in brackets.

**Significant at 95%

Table 5. Effects on material hardship (PS matching estimates).

TREATMENT	OUTCOME VARIABLES		
	Material well-being	Material hardship	Housing problems
Non-participation in a work-related scheme	9.991 (2.737) ¹	1.711 (1.592)	0.585 (0.493)
Participation in a work-related scheme	9.885 (2.778)	1.777 (1.581)	0.639 (0.480)
<i>Average effect</i>	-1.1	3.9	9.2*
Non-participation in a work-related scheme	9.845 (2.886)	1.755 (1.590)	0.596 (0.491)
Participation in general work-related schemes	9.889 (2.810)	1.694 (1.368)	0.649 (0.478)
<i>Average effect</i>	0.4	-3.5	8.9*
Non-participation in general work-related schemes	10.153 (2.545)	1.684 (1.603)	0.552 (0.498)
Participation in intensive work-related schemes	10.186 (2.354)	1.733 (1.697)	0.546 (0.501)
<i>Average effect</i>	0.3	2.9	-1.1
Non-participation in a work-related scheme	9.845 (2.886)	1.755 (1.590)	0.596 (0.491)
Participation in mixed work-related schemes	9.433 (3.123)	2.073 (1.877)	0.645 (0.480)
<i>Average effect</i>	-4.2	18.1	8.2
Participation in general work-related schemes	9.845 (2.882)	1.695 (1.415)	0.665 (0.472)
Participation in intensive work-related schemes	10.234 (2.311)	1.691 (1.552)	0.600 (0.492)
<i>Average effect</i>	4.0	-0.2	-9.8
Participation in general work-related schemes	9.969 (2.756)	1.655 (1.377)	0.640 (0.480)
Participation in mixed work-related schemes	9.594 (2.959)	2.040 (1.884)	0.648 (0.478)
<i>Average effect</i>	-3.8	23.3	1.3
Participation in intensive work-related schemes	10.264 (2.149)	1.747 (1.657)	0.630 (0.485)
Participation in mixed work-related schemes	9.424 (3.136)	2.070 (1.872)	0.639 (0.481)
<i>Average effect</i>	-8.2	18.5	1.4

¹ Standard deviation in brackets.

*Significant at 90%.

Table 6. Treatments-outcomes matrix.

	<i>Comparison 1</i>	<i>Comparison 2</i>	<i>Comparison 3</i>	<i>Comparison 4</i>	<i>Comparison 5</i>	<i>Comparison 6</i>	<i>Comparison 7</i>
Employment (currently)	++*	+	+++**	++*	+++**	++	---
Legal contract and payroll taxes	+++**	≈	+++**	+++***	+++*	+++***	++*
Subjective poverty	-	≈	--	--**	---**	--**	-
Economic situation compared to 10 years ago	+	+	-	≈	≈	--	≈
Material well-being	≈	≈	≈	≈	≈	≈	-
Material hardship	≈	≈	≈	++	≈	+++	++
Housing problems	+*	+*	≈	+	-	≈	≈

(≈): <5%; (-/+): 5-10%; (--/++): 10-20%; (+++/---): >20%

***Significant at 99%, **Significant at 95%, *Significant at 90%.