A Reappraisal of the Virtues of Private Sector Employment Programmes

Brian Krogh Graversen
Peter Jensen

July 2006

DISCUSSION PAPER SERIES

A Reappraisal of the Virtues of Private Sector Employment Programmes

Brian Krogh Graversen
Danish National Institute of Social Research

Peter Jensen
*Aarhus School of Business
and IZA Bonn*

ABSTRACT

A Reappraisal of the Virtues of
Private Sector Employment Programmes[*]

In this paper, we evaluate the employment effects of Danish active labour market programmes aimed at welfare benefit recipients. We estimate an econometric model with treatment effects and discrete outcomes and we allow the responses to treatment to vary among observationally identical persons. The empirical analysis is based on a register-based dataset that gives information on participation in labour market programmes and subsequent employment. Using a latent variable model, we estimate commonly defined treatment effects, and in particular, the distribution of treatment effects. We do not find any significant mean effects of participation in private sector employment programmes compared to participation in other programmes, but we find evidence of heterogeneity in the treatment effects.

JEL Classification: I38, J64, J68

Keywords: heterogeneous treatment effects, active labour market programmes, welfare
              dependence, employment

Corresponding author:

Peter Jensen
Department of Economics
Aarhus School of Business
Prismet, Silkeborgvej2
DK-8000 Aarhus C
Denmark
E-mail: pje@asb.dk

1 Introduction

Evaluations of active labour market programmes (ALMPs) often point to private sector employment programmes as the most successful type of programme to get unemployed persons back to work. Although the specific features of the programmes tend to vary between countries, this seems to be one of the most persistent lessons that can be learned from the abundance of evaluations that have emerged in recent years. However, most of the evaluation studies estimate mean treatment effects implicitly assuming that all persons respond in the same way to a given programme. Also, in some cases, selection into programmes may be based on observable as well as unobservable characteristics of the potential participants. If such selection is not properly accounted for, the estimation results may be biased. The recent literature on evaluating social programmes points to the fact that different persons respond differently to the same programme. Therefore it is important to evaluate programmes using models that allow for heterogeneous treatment effects[1]. This will provide additional information on distributional issues and allow us to gain further insight into the functioning of such programmes. In this paper, we will therefore take another look at the employment effects of private sector employment (PSE) programmes taking into account selection into programmes and heterogeneous treatment effects.

In this paper, we evaluate the employment effects of Danish ALMPs aimed at welfare benefit recipients. We estimate an econometric model with treatment effects and discrete outcomes and we allow the responses to treatment to vary among observationally identical persons. The empirical analysis is based on a register-based dataset that gives information on participation in labour market programmes and subsequent employment. All previous evaluations of Danish labour market and social policies have estimated mean effects, but have not allowed for heterogeneous treatment effects[2]. In addition, there has not been much focus on the (social) labour market programmes aimed at recipients of welfare benefits, whereas the corresponding labour market programmes aimed at recipients of unemployment insurance benefits have been extensively investigated (this is the case for Denmark as well as for other countries). Hence, our paper contributes to a very sparse empirical literature on heterogeneous treatment effects and our empirical results do also contribute to the knowledge about the effectiveness of social programmes.

---

[1] See Heckman (2001).

[2] See Bolvig et al. (2003) and Graversen (2004a) for some recent examples of this.

Traditionally, evaluations of programme effects consider a treatment group and a control group, where individuals in the latter are non-participants (i.e. they do not participate in any programmes). However, in a Danish context an unemployed person will end up in a programme sooner or later and hence it is not possible to create a 'standard' control group of persons who have not participated in any programmes. Therefore, we are only investigating the relative efficiency of the various ALMPs, i.e. we compare the effects of different programmes, but we do not compare participation to non-participation. Given the fact that the welfare benefit recipients *have to* participate in ALMPs, it is essential to know which programmes fulfil the purposes in the optimal way. A similar situation also arises in other countries with a comprehensive welfare state, see for instance Carling and Richardson (2004) and Sianesi (2002) for their analyses of labour market policy in Sweden, another example of a Scandinavian welfare state. Such a comparison of different programmes also has important implications for the allocation done by caseworkers, see Lechner and Smith (forthcoming) for another analysis of this aspect.

In our analysis, we focus on the employment effects, primarily since that is the main purpose of the programmes, in contrast to many of the US training programmes where the main focus is on increasing earnings. In general, many European countries have put a lot of emphasis on labour market and social policies to reduce their high levels of unemployment. This has lead to an extensive use of various types of programmes and consequently to a strong interest in evaluating the employment effects of these programmes. Some recent examples of such evaluations are Gerfin and Lechner (2002) who evaluate the active labour market policy in Switzerland, and Andrén and Andrén (2002) who evaluate labour market training programmes in Sweden.

The empirical model is a latent variable model with employment as the discrete outcomes and with selection into programmes determined by both observable and unobservable characteristics. The unobservables are specified to follow a one-factor structure model and they enter into the selection equation as well as the outcome equations. The model has originally been formulated by Aakvik et al. (2005), and its structure is particularly convenient for deriving mean treatment effects and distributional treatment effects. This type of model has recently also been applied by Andrén and Andrén (2002) and Aakvik et al. (2003) using Swedish and Norwegian data, respectively.

The estimation of the parameters of the empirical model provides us with a number of important results. First, we get evidence on the mean employment effects of the PSE programmes relative to other ALMPs. Second, our estimation results allow us to derive the distribution of treatment effects across participants and thereby determine to which degree the PSE programmes have different effects on different persons. Third, the estimation results also provide some evidence on the importance of controlling for unobservables in the selection and employment processes.

The remainder of the paper is organized as follows. In Section 2, we describe the institutional settings of the Danish social policy, while we describe the data for our analysis in Section 3. Section 4 presents the econometric model and the various treatment effects that we estimate. The results from the empirical analysis are presented and discussed in Section 5, while Section 6 concludes the paper.

2   Institutional settings

In Denmark there are two different administrative systems for individuals experiencing unemployment, the labour market system (unemployment insurance) and the social system (welfare). The former is responsible for the activities related to unemployed individuals insured against unemployment and eligible for unemployment insurance benefits (UI-benefits). A number of unemployment funds administrate the disbursement of UI-benefits, whereas the public employment service organizes various activities (e.g. job provision and ALMPs) with the intention of alleviating the unemployment problem of the UI-benefit recipients.

Unemployed individuals who are not insured against unemployment or who are not eligible for UI-benefits can receive welfare benefits (when certain conditions related to the overall income and wealth of the household are met). In contrast to UI-benefits that can only be received for a limited period (up to four years), welfare benefits can in principle be received for an unlimited period. The administration of the disbursement of welfare benefits is placed in the hands of the municipalities. The municipalities are also responsible for the organization and administration of the different measures available to assist the welfare benefit recipients in becoming self-supporting.

The measures that can be used to remedy the unemployment problem of UI-benefit recipients and welfare benefit recipients are very similar. For both groups there is a 'right-and-duty' principle associated with the ALMPs. The unemployed have a right to receive assistance in

the form of an ALMP-offer. At the same time, they have a duty to participate in ALMPs and other activities when offered to retain eligibility to the UI-benefits and welfare benefits. At present, approximately 3 per cent of the labour force (measured in full-time equivalents) participates in some kind of ALMP. More than half of the participants are welfare benefit recipients. In international comparisons, Denmark is among the countries that use the largest part of GDP on ALMPs (Martin and Grubb (2001)).

The municipalities have had the possibility to assign welfare benefit recipients into ALMPs since 1977 (Brogaard and Weise (1997)). Before 1994, when a new legislation was implemented, the ALMPs were primarily used to alleviate the unemployment problem of young welfare benefit recipients. But today, all welfare benefit recipients have to participate in an ALMP if they do not become self-supporting after a certain period. In the available dataset there is no information on programme periods that ended before 1994.

At present, welfare benefit recipients below 30 years have to receive an ALMP-offer not later than 3 months after the first day on welfare. If they do not succeed to become independent of welfare benefits after the end of an ALMP, they have to receive a new ALMP-offer 3 months after the end of the previous programme period at the latest. Hence, welfare benefit recipients below 30 years are subject to a more or less continuous treatment in programmes while on welfare. Welfare benefit recipients above 30 years should participate in an ALMP after 12 months at the latest. They only have a right to receive one ALMP-offer during a welfare spell. But most municipalities choose to give a new ALMP-offer if the first programme was not successful in bringing the welfare benefit recipients from welfare to a situation where they are self-supporting.[3]

The decision to assign a welfare benefit recipient into a given programme is made by case-workers in the municipality. For all welfare benefit recipients the type of the ALMP-offer should depend on the background and desires of the individual welfare benefit recipient. If possible, the welfare benefit recipients should receive a selection of programmes to choose from. However, apart from the desires and needs of the welfare benefit recipients, the case-

---

[3] Before mid-1998, the current rules regarding the ALMPs for welfare benefit recipients below 30 years only applied to welfare benefit recipients below 25 years. The rules regarding the ALMPs for welfare benefit recipients above 30 years that are effective today applied to welfare benefit recipient above 25 years. Another difference compared with the present rules is that the municipalities were not obliged to give ALMP-offers to welfare benefit recipient below 25 years with problems in addition to unemployment. However, a large part of the municipalities chose to offer ALMPs to this group as well, even if they were not required to do so.

workers have to take account of the availability of different programme types, the cost of the programmes and the current state of the regional labour market.

The main purposes of the ALMPs are to improve the labour market prospects of the welfare benefit recipients by upgrading their skills and to increase the intensity of job search by testing job-readiness and reducing 'leisure time'. However, for welfare benefit recipients with problems in addition to unemployment (e.g. problems related to health, childcare or lack of self-confidence) participation in a programme should primarily improve everyday life (e.g. through support for the handling of everyday activities and the creation of a basic network). The hope is that the improvement in everyday life in the long run will bring this group of welfare benefit recipients closer to the labour market and a to situation where they are self-supporting. Ideally, the programme chosen for each individual should be the programme that meets these purposes in the best way.

There is a broad range of different types of programmes that the caseworkers can choose from when giving ALMP-offers to the welfare benefit recipients. We have chosen to group the programmes in four main categories: *private sector employment (PSE) programmes*, *public sector employment programmes*, *classroom training* and *other programmes*. However, through most of the paper the three last programme categories are lumped together under the heading *non-PSE programmes*. The programme category 'private sector employment programmes' consists of programmes where the participants work in a private firm. During the programme period the participants receive a wage subsidy. In public sector employment programmes the participants either work in a public institution or in a special employment project created by the municipalities. The welfare benefit recipients in these programmes often carry out work that would otherwise not be done. This can be e.g. snow clearing for senior citizens, nature preservation and assistance of the permanent staff in municipal institutions (schools, youth hostels, theatres, sports centres, museums etc.). Participants in classroom training attend classes to upgrade their qualifications in different fields. The residual group of programmes termed 'other programmes' consists of the self-employment grant programme and programmes that could not be placed in the three other programme categories (mainly because of missing data). A more detailed description of the programmes can be found in Danish Ministry of Labour (2001) and Graversen (2004a).[4]

---

[4] In contrast to Graversen (2004a), the category 'other programmes' does not include rehabilitation programmes. These programmes are available for individuals whose working capacity is reduced (because of e.g. physical or

6

Since all welfare benefit recipients *have to* participate in an ALMP if they do not succeed to become independent of the welfare benefit before an ALMP-offer is given, it is not possible to create a 'standard' control group for the purpose of estimating the effect of the programmes. One possibility to circumvent this problem is to use the variation in the timing of the programmes to estimate the programme effects (Abbring and van den Berg (2003)). However, in this study we choose to estimate the effect of the PSE programmes relative to the effect of other programme types (non-PSE programmes).[5] The advantage of this approach is that we only have to include participants in the empirical analyses. The disadvantage is of course that we do not identify the effect of programme participation relative to the effect of non-participation. Anyway, given the fact that the welfare benefit recipients *have to* participate in ALMPs, it is essential to know which programmes fulfil the purposes of the programmes in the optimal way.

3 Data

In this study we use a data set drawn from *The Register for Analyses relating to the Social Responsibility of Enterprises*. This is a large longitudinal database that has been constructed by The Danish National Institute of Social Research in collaboration with Statistics Denmark. The database is based on administrative data merged from several administrative registers by use of civil registration numbers. It contains a 10 per cent random sample of the Danish Population in the age group 17-66. For each individual the database provides detailed information on various individual characteristics such as e.g. age, gender, nationality, municipality of residence, marital status, children, educational attainment, work experience, employment and unemployment. The database is updated at an annual basis and at present the individuals in the sample can be followed during the period 1984-2000.

A number of conditions have to be met for an individual to be part of the sample used in this study. First, we select a sample of individuals who ended an ALMP during the period 1994-1998.[6] The individuals may have participated in more than one programme during this period

---

psychological problems) to such degree that there is a very limited chance that standard ALMPs will help bringing the individuals from public assistance to employment.

[5] Carling and Richardson (2004) use a similar approach.

[6] The database does not provide information on ALMPs that ended before 1994.

but we only include information on the first programme episode.[7,8] Second, we restrict the sample in various ways to ensure that we have the necessary information for our analysis. Finally, we restrict our attention to men (excluding immigrants and refugees from countries outside EU and North-America) aged 18-59. Below we detail the various sample selection rules together with the ensuing loss of observations:

Individuals who ended an ALMP during 1994-1998	20,105 observations
Restrict to ALMP starting in 1993 or later	20,060 observations
Restrict to ALMP with length between 2 weeks and 5 years	18,454 observations
Exclude individuals with missing basis information	18,339 observations
Restrict to men	9,193 observations
Exclude immigrants and refugees	7,841 observations
Restrict to age 18-59 when starting ALMP	7,181 observations
Restrict to individuals with information for all years incl. one year before and one year after the programme period	6,987 observations
Exclude individuals in public sector employment programmes with missing data	6,822 observations
Exclude individuals from municipalities with less than 10 ALMP participants	6,613 observations

Thus, we end up with a sample of 6,613 men. Of these, 1,391 (21 per cent) participated in a PSE programme and 5,222 (79 per cent) participated in another type of programme.

We constructed an algorithm that utilizes the information in the database to determine for every month whether an individual is employed, unemployed, studying or outside the labour market (e.g. individuals receiving early retirement pension). Periods in different programmes

---

[7] Individuals categorized as participants in job training programmes in the public sector (approximately 2 per cent of the ALMP participants) are excluded from the sample. There seems to be a problem with the municipalities' registration of this programme type. Among welfare benefit recipients that the municipalities have registered to be participants in job training programmes in the public sector a large fraction are also registered to have participated in programmes arranged by the public employment service. The reason for this may be that some municipalities erroneously have reported UI-benefit recipients participating in ALMPs in municipal institutions as participants in ALMPs aimed at welfare benefit recipients.

[8] The database contains information on the actual starting date and ending date of ALMP periods. However, there are two major problems with this information, cf. Graversen (2004b). The first problem is that some programme periods are cut up into smaller sub-periods with intervening periods where the programme participants do not participate in the programme. The second problem is that some programme periods are overlapping in time. If two programme periods of the same type overlap or there is less than one week between the end of the first programme period and the start of the second programme period, we choose to collapse the two periods into one programme period. If two programme periods are not of the same type, we simply chose to include information on the programme period that starts first. It should be noted, that to determine whether two programmes are of the same type we use a finer categorization of the programmes than the relatively broad categorization used in the text.

intended to alleviate the unemployment problem of individuals (ALMPs and rehabilitation programmes) are categorized as unemployment. A minor part of the sample (3-4 per cent) could not be categorized into the four different groups mentioned above.

Looking at the raw data, we see that participants in PSE programmes are more likely to be employed after the end of the programme period than participants in the other programme types. One year after the programme period more than half of the participants in PSE programmes are employed, whereas this is only the case for approximately one third of the participants in public sector employment programmes and classroom training, cf. Figure 1. Among the participants in PSE programmes fewer individuals are unemployed and outside the labour market than is the case for non-PSE programme participants, cf. Figures 2 and 3. The proportion of programme participants participating in an ordinary education is almost the same for the different programme types, cf. Figure 4.

The virtues of the PSE programmes could be due to a selective assignment with respect to the characteristics of the participants. Compared to the participants in non-PSE programmes the participants in PSE programmes are a little younger, more likely to have a formal education beyond primary or lower secondary school, have slightly more work experience and spent less time in unemployment prior to the programme period, cf. Table 1. But the differences are quite modest. To obtain a more firmly based assessment of the employment effect of PSE programmes relative to the other ALMPs, we formulate and estimate an econometric model that accounts for the differences in observed and unobserved characteristics of participants in PSE programmes and non-PSE programmes.

In addition to the individual characteristics of the welfare benefit recipients, we also include various characteristics of the municipalities where the welfare benefit recipients live. To account for the fact that the welfare benefit recipients may be confronted with different labour markets (some may live an area with a high unemployment rate and a lack of vacant jobs, whereas other may be living in a area with a booming economy), we include a measure of the regional unemployment rate relative to the countrywide unemployment rate. Our labour market regions are identical to the commuting areas defined by the Ministry of Environment and Energy (2001). The commuting areas are formed in such a way that a relatively large fraction of the residents in a given region work within the region. This also means that the level of

commuting between one commuting area and other commuting areas is limited.[9] The unemployment rate in each region is calculated on an annual basis by using the available information in our large longitudinal database covering 10 per cent of the population of working age.

Another characteristic of the municipalities that we include in the statistical analyses is the number of residents in each municipality.[10] The size of a municipality may have an impact on the type of programmes that can be offered to the welfare benefit recipients. If a municipality is small, there may be less programmes to choose from. There will typically be fewer private firms and public institutions than in larger municipalities. It may also influence the effect of the programmes if the caseworkers in the small municipalities mainly choose programmes on the basis of the availability of different programmes instead of the expected benefits. When the size of the municipalities is used as an explanatory variable, this variable might also act as a proxy for unmeasured characteristics, e.g. because the problems of the welfare benefit recipients in the large municipalities are of another character than the problems of welfare benefit recipients in smaller municipalities.

Finally, we include a variable that measures the use of PSE programmes in each municipality relative to the countrywide use of this programme type. Let W_{it} denote the value of this variable for individual i in year t. More precisely, this variable is defined in the following way:

$$W_{it} \equiv \frac{(N_{it}^{PSE} / N_{it}^{ALMP})}{(N_{t}^{PSE} / N_{t}^{ALMP})} \tag{1}$$

where N_{it}^{PSE} is the number of PSE programme participants in the municipality where individual i lives, N_{it}^{ALMP} is the number of ALMP participants in this municipality, N_{t}^{PSE} is the number of PSE programme participants on national basis, and N_{t}^{ALMP} is the number of ALMP participants on national basis.[11] Obviously, W_{it} has the same value for individuals living in the same municipality.

---

[9] During the sample period there were 275 municipalities in Denmark. These municipalities are grouped into 45 commuting areas by The Ministry of Environment and Energy (2001).

[10] These data are publicly available from StatBank Denmark (www.statistikbanken.dk).

[11] The data used to form W_{it} originate from StatBank Denmark. The number of programme participants in each municipality available from this source is given in full-time equivalents. We would of course prefer to know the actual number of participants in each programme type since this would give the best measure of the likelihood of being assigned into a PSE programme in each municipality. But the available data should do well since the average duration of the different programme types is not that different across municipalities. Unfortunately, the data are only available from 1995 onwards.

The relative use of PSE programmes in the municipalities is used as an instrument in our statistical model that controls for selection into programmes, i.e. this variable is included in the selection equation, but excluded in the employment equations. Obviously, we expect this variable to influence the probability that a welfare benefit recipient participates in a PSE programme. To be a valid instrument, the relative use of PSE programmes in a municipality should not have a direct effect on the employment outcome after training. The only effect on the outcome should be through the possible effect of the actual programme choice (the programme effect). There is a very large degree of variation in the instrument. In some municipalities, 70 per cent of the ALMP participants participate in PSE programmes, whereas in other municipalities no AMLP participants participate in PSE programmes. The average fraction of ALMP participants who participate in PSE programmes is 18 per cent. The value of the instrument may vary between different municipalities for various reasons. One important reason for the variation is that different municipalities make different choices regarding the desired composition of ALMP programmes and exert different effort into obtaining available PSE programme slots. Municipalities that are very much alike in terms of employment possibilities can thus have very different values of the instrument. The value of the instrument may also depend on the occupational structure in the municipality. If there are many private firms in a municipality, it may be easier for the municipality to establish PSE programme arrangements. There also exist potential reasons for variation that may render the instrument invalid, e.g. if the municipalities with a high relative use of PSE programmes are the municipalities that have the most employable ALMP participants. However, in Section 5 we provide further details on the correlation between the instrument and employment probabilities, supporting the validity of this instrument.

The regional treatment intensity has also been used in other studies evaluating labour market policies; see e.g. Frölich and Lechner (2004) who evaluate the Swiss active labour market policies using exogenous variation across Swiss cantons. In their case, the central law obliges the cantons to have a different number of programmes, thereby creating truly exogenous variation. In addition, their methodology differs from ours.

There is, however, an obvious endogeneity problem if W_{it} is included in the statistical model as the variable is defined in (1). If individual i begins an ALMP period during year t this may have an impact on W_{it}. This is problematic, if W_{it} is used to explain the choice of the programme type for an individual beginning a programme period in year t. For this reason we choose to lead W_{it} with two years (i.e. we use $W_{i(t+2)}$ instead of W_{it}).

The reason that we lead the information on the relative use of PSE programmes in each municipality by 2 years instead of 1 year is that a considerable share of the programme periods started in year t will continue into year $t+1$. The average duration of ALMP periods is approximately 6 months. Hence, if an individual starts a programme period in year t this may influence $W_{i(t+1)}$ as well as W_{it}. $W_{i(t+2)}$ is of course only a good instrument if there is some degree of correlation in this variable over time. Municipalities with a high relative use of PSE programmes in one year should also have a relatively high use of this programme type in the surrounding years.

A more obvious approach to overcome the endogeneity problem of W_{it} would have been to lag this variable. But the data used to form W_{it} is only available from 1995 onwards. The sample contains individuals who begin a programme episode in the period 1993-1998. If we chose to lag W_{it} with one period, we would face a problem with missing data for the years 1993-1995. Hence, to avoid this we have chosen to lead the variable.

In some municipalities, very few individuals participate in ALMPs. In these municipalities, the variable W_{it} may not give a good indication of the probability of being assigned into a PSE programme. For this reason, we have chosen to exclude individuals from municipalities where the number of ALMP participants (full-time equivalents) is less than 10.

4   The econometric model

To estimate the employment effect of PSE programmes relative to other ALMPs we use a latent variable model formulated by Aakvik et al. (2005). This type of model has recently also been applied by Andrén and Andrén (2002) and Aakvik et al. (2003). The model explicitly accounts for the differences in observed and unobserved characteristics of the participants in PSE programmes and non-PSE programmes. In contrast to most of the applied research on the effect of ALMPs, the model allows that the employment effect of the programmes vary among observationally identical persons, and that the assignment into programmes may depend on the expected employment effect.[12] Within this framework it is possible to estimate traditionally defined mean treatment effects as well as the distribution of the treatment effect.

---

[12] See Heckman (1997) for a discussion of the problems of conventional IV methods when treatment effects are heterogeneous and person-specific outcomes from treatment influence the assignment into programmes.

4.1 A latent variable model

Let D_i be a dichotomous variable that is equal to 1 if individual i participates in a PSE programme and 0 if the individual participates in another type of programme. The assignment into the two different programme categories is determined by a latent variable D_i^*:

$$D_i^* = Z_i \, \beta_D + U_{Di}$$
$$D_i = 1 \; if \; D_i^* \geq 0, \, D_i = 0 \; otherwise \tag{2}$$

where Z_i is a vector of observed variables and U_{Di} is an unobserved random variable. The latent variable D_i^* can be thought of as the net utility or gain to the caseworker (responsible for the case of individual i) from assigning individual i into a PSE programme rather than another type of programme.

If $D_i=1$, the employment status is Y_{1i}, and if $D_i=0$, the employment status is Y_{0i}. Y_{1i} and Y_{0i} are both dichotomous variables that are equal to 1 if the individual is employed and 0 otherwise. An individual is only observed in one of the two different states ($D_i=1$, $D_i=0$) and hence the observed employment status Y_i is equal to

$$Y_i = D_i Y_{1i} + (1 - D_i) Y_{0i} = Y_{0i} + D_i (Y_{1i} - Y_{0i}) \tag{3}$$

The potential employment status if an individual is assigned into a PSE programme is determined by

$$Y_{1i}^* = X_i \, \beta_1 + U_{1i}$$
$$Y_{1i} = 1 \; if \; Y_{1i}^* \geq 0, \, Y_{1i} = 0 \; otherwise \tag{4}$$

where Y_{1i}^* is a latent index describing the joint effect of the desire and ability to find employment. X_i is a vector of observed variables and U_{1i} is an unobserved random variable.

The potential employment status if individual i is assigned into a non-PSE programme is given by

$$Y_{0i}^* = X_i \, \beta_0 + U_{0i}$$
$$Y_{0i} = 1 \; if \; Y_{0i}^* \geq 0, \, Y_{0i} = 0 \; otherwise \tag{5}$$

where Y_{0i}^* is a latent variable, X_i is the same vector of observed variables that affects Y_{1i}^* and U_{0i} is an unobserved random variable.

Note that in general the model is not identified if the variables in *Z* and *X* are identical.[13] This is the case though if the unobserved variables are normally distributed, as we assume in this paper. However, to improve the empirical identification of the model we impose an exclusion restriction on the model. The variable *W* (the relative use of PSE programmes in the relevant municipality) defined in Section 3 is included in *Z* but not in *X*. More formally, we assume that *Z=(X, W)*.

4.2 Individual-specific treatment effects

Define the individual-specific treatment effect Δ_i in the following way

$$\Delta_i = Y_{1i} - Y_{0i} = \mathbf{1}[X_i \beta_1 + U_{1i} \geq 0] - \mathbf{1}[X_i \beta_0 + U_{0i} \geq 0] \tag{6}$$

where $\mathbf{1}[.]$ is the indicator function. For a given individual i Δ_i measures what the employment status would be in the case when the individual is assigned into a PSE programme compared to the employment status in the case when the individual participates in a non-PSE programme. Δ_i can attain three possible values: 0,1 and –1.

The effect of the unobserved variables U_{1i} and U_{0i} is the same in the two potential states if $U_{1i}=U_{0i}$. In this case Δ can only attain two different values for individuals with a given value of *X*.[14] Let U_i denote the common value of U_{1i} and U_{0i}. Then, if $X_i (\beta_1 - \beta_0) > 0$, Δ can attain the values 0 (if $U_i < -X_i \beta_1$ or $U_i \geq -X_i \beta_0$) and 1 (if $-X_i \beta_1 \leq U_i < -X_i \beta_0$). If $X_i (\beta_1 - \beta_0) < 0$, Δ can attain the values 0 (if $U_i \geq -X_i \beta_1$ or $U_i < -X_i \beta_0$) and –1 (if $-X_i \beta_0 \leq U_i < -X_i \beta_1$). Thus in contrast to a model with continuous outcomes $(Y_{1i}^* = Y_{1i}, Y_{0i}^* = Y_{0i})$ it is possible that the treatment effects vary by unobserved characteristics even if the effect of these characteristics is the same in both states. Some individuals may have unfavourable unobserved characteristics (low U_i –values) such that it is very unlikely that they find employment after the end of a programme no matter what type of programme they are assigned into. Equivalently, some individuals may have very favourable unobserved characteristics making them very likely to be employed after the programme period irrespective of the programme type chosen.

---

[13] See e.g. Heckman (1990) for a discussion of the identification problem.

[14] In the special case where $X_i (\beta_1 - \beta_0) = 0$, Δ is equal to zero for all individuals.

In this paper we allow the effect of the unobservables to be different in the two states (i.e. it may be the case that $U_{1i} \neq U_{0i}$). This creates an additional source of heterogeneity in the treatment effects such that all three different values for Δ (0,1 and -1) may be experienced for different individuals with the same observed characteristics.

Very strong assumptions are required to estimate individual-specific treatment effects (see e.g. Heckman et al. (1999)). For this reason we estimate different population means and distributions of Δ.

4.3 Different treatment parameters and distributions of interest

In most applied research on the effects of ALMPs the focus has been on the estimation of the mean effect of treatment on the treated (TT), i.e. the average treatment effect of those who participate in a given programme. [15] Another parameter that is often referred to is the average treatment effect (ATE). ATE is the mean treatment effect of a randomly selected individual from a given population.

In our setting, the treatment effect is the change in the outcome variable from participating in a PSE programme instead of a non-PSE programme. Hence, TT is the average treatment effect among those individuals who participate in a PSE programme, and ATE is the average treatment effect of a randomly selected individual from the group of all ALMP participants.

For individuals with a given value of X and Z, TT and ATE are defined as

$$\Delta^{TT}(x,z) \equiv E(\Delta \mid X = x, Z = z, D = 1) \tag{7}$$

$$\Delta^{ATE}(x) \equiv E(\Delta \mid X = x)^{\,16} \tag{8}$$

Recently, there has been an increased interest in other parameters characterizing the treatment effect of programmes (e.g. the LATE parameter of Imbens and Angrist (1994)). Heckman and Vytlacil (2000) show that the different available estimators of the treatment effect of programmes can be expressed as a weighted average of a parameter called the marginal treatment effect (MTE). MTE is defined as

$$\Delta^{MTE}(x,u) \equiv E(\Delta \mid X = x, U_D = u) \tag{9}$$

---

[15] See e.g. Heckman et al. (1999) for a comprehensive survey of the methods used in the evaluation of ALMPs.

[16] ATE does not depend on Z since Y_1 and Y_0 (used to calculate Δ) are independent of D by assumption.

For small values of U_D, MTE is the average treatment effect of individuals with unobserved characteristics that make them least likely to participate in a PSE programme. For large values of U_D, MTE is the average treatment effect of individuals with unobserved characteristics that make them most likely to participate in a PSE programme.

If U_D is independent of U_1 and U_0 then all three mean treatment effect parameters defined in (7), (8) and (9) are equal for a given value of X. This will generally not be the case if the PSE programme participants are selected on the basis of unobserved characteristics that affect the outcome in the PSE programme state or the non-PSE state.

In our latent variable model it is actually possible to identify the distribution of the potential outcomes Y_1 and Y_0. For individuals participating in a PSE programme the probabilities characterizing the joint distribution of the two outcome variables is given by

$$P_{Y_1,Y_0}^{TT}(j_1, j_0, x, z) \equiv \Pr(Y_1 = j_1, Y_0 = j_0 \mid X = x, Z = z, D = 1) \tag{10}$$

where $j_1 \in \{0,1\}$ and $j_0 \in \{0,1\}$.

Equivalently the probabilities characterizing the distribution of Y_1 and Y_0 for an individual chosen at random or an individual with a given value of the unobserved variable U_D are given by the following equations:

$$P_{Y_1,Y_0}^{ATE}(j_1, j_0, x) \equiv \Pr(Y_1 = j_1, Y_0 = j_0 \mid X = x) \tag{11}$$

$$P_{Y_1,Y_0}^{MTE}(j_1, j_0, x, u) \equiv \Pr(Y_1 = j_1, Y_0 = j_0 \mid X = x, U_D = u) \tag{12}$$

Another distribution that can be identified is the distribution of the treatment effect Δ. For individuals actually participating in a PSE programme, individuals selected at random and individuals with a given value of the unobserved variable U_D, the probabilities characterizing the distribution of the treatment effect are defined in the following way:

$$P_\Delta^{TT}(j, x, z) \equiv \Pr(\Delta = j \mid X = x, Z = z, D = 1) \tag{13}$$

$$P_\Delta^{ATE}(j, x) \equiv \Pr(\Delta = j \mid X = x) \tag{14}$$

$$P_\Delta^{MTE}(j, x, u) \equiv \Pr(\Delta = j \mid X = x, U_D = u) \tag{15}$$

where $j \in \{-1,0,1\}$.

There is a close connection between the mean treatment effect, the joint distribution of the potential outcomes and the distribution of the treatment effect. Knowledge of the joint distribution of the potential outcomes is more informative than knowledge of the mean treatment effect or the distribution of the treatment effect. If the joint distribution of the potential outcomes is known, then the two other ways to describe the effect of PSE programme participation can easily be derived. For the individuals who actually participate in a PSE programme the mean treatment effect is given by

$$\Delta^{TT}(x,z) = P_{Y_1,Y_0}^{TT}(1,0,x,z) - P_{Y_1,Y_0}^{TT}(0,1,x,z)$$

and the probabilities characterizing the distribution of the treatment effect are given by

$$P_{\Delta}^{TT}(-1,x,z) = P_{Y_1,Y_0}^{TT}(0,1,x,z)$$

$$P_{\Delta}^{TT}(0,x,z) = P_{Y_1,Y_0}^{TT}(0,0,x,z) + P_{Y_1,Y_0}^{TT}(1,1,x,z)$$

$$P_{\Delta}^{TT}(1,x,z) = P_{Y_1,Y_0}^{TT}(1,0,x,z)$$

Similar expressions apply for $\Delta^{ATE}(x)$, $\Delta^{MTE}(x,u)$, $P_{\Delta}^{ATE}(j,x,z)$ and $P_{\Delta}^{MTE}(j,x,u)$ ($j \in \{-1,0,1\}$).

4.4 A one-factor structure assumption

Following Aakvik et al. (2005), we assume that the unobserved variables are normally distributed with a common factor θ, an assumption first introduced by Heckman (1981). The random variables U_D, U_1 and U_0 are generated by the following equations:

$$\begin{aligned} U_{Di} &= \alpha_D \theta_i + \varepsilon_{Di} \\ U_{1i} &= \alpha_1 \theta_i + \varepsilon_{1i} \\ U_{0i} &= \alpha_0 \theta_i + \varepsilon_{0i} \end{aligned} \qquad (16)$$

ε_D, ε_1, ε_0 and θ are independent normally distributed with mean 0 and variance equal to 1. To identify the model we set α_D equal to 1.

From the model (given by (2)-(5) and (16)) the different mean treatment parameters and the probabilities characterizing the distribution of the potential outcomes and the treatment effects defined in Section 4.3 can be expressed as a function of the observed variables and the parameters in the model. The mean treatment parameters can be expressed in the following way:

$$\Delta^{ATE}(x) = \int [\Phi(x\beta_1 + \alpha_1\theta) - \Phi(x\beta_0 + \alpha_0\theta)]\phi(\theta)d\theta \tag{17}$$

$$\Delta^{TT}(x,z) = \frac{1}{\Phi(z\beta_D/\sqrt{2})} \int [\Phi(x\beta_1 + \alpha_1\theta) - \Phi(x\beta_0 + \alpha_0\theta)]\Phi(z\beta_D + \theta)\phi(\theta)d\theta \tag{18}$$

$$\Delta^{MTE}(x,u) = \frac{\sqrt{2}}{\phi(u/\sqrt{2})} \int [\Phi(x\beta_1 + \alpha_1\theta) - \Phi(x\beta_0 + \alpha_0\theta)]\phi(u-\theta)\phi(\theta)d\theta \tag{19}$$

Similar expressions can be derived for the joint distribution of the potential outcomes and the distribution of the treatment effect, cf. appendix A.

4.5 Estimation of the model

Conditioning on the individual specific factor θ, the contribution to the likelihood function of individual i has the following form:

$$\begin{aligned}L_i(\theta_i) &= \Pr(D_i, Y_i \mid X_i, Z_i, \theta_i) \\ &= \Pr(D_i \mid Z_i, \theta_i)\Pr(Y_i \mid D_i, X_i, \theta_i)\end{aligned} \tag{20}$$

Let $P_{Di}(\theta_i)$, $P_{1i}(\theta_i)$ and $P_{0i}(\theta_i)$ be defined as

$$\begin{aligned}P_{Di}(\theta_i) &\equiv \Pr(D_i = 1 \mid Z_i, \theta_i) = \Phi(Z_i\beta_D + \theta_i) \\ P_{1i}(\theta_i) &\equiv \Pr(Y_i = 1 \mid D_i = 1, X_i, \theta_i) = \Pr(Y_{1i} = 1 \mid X_i, \theta_i) = \Phi(X_i\beta_1 + \alpha_1\theta_i) \\ P_{0i}(\theta_i) &\equiv \Pr(Y_i = 1 \mid D_i = 0, X_i, \theta_i) = \Pr(Y_{0i} = 1 \mid X_i, \theta_i) = \Phi(X_i\beta_0 + \alpha_0\theta_i)\end{aligned} \tag{21}$$

where $\Phi(\cdot)$ denotes the standard normal cumulative density function. Then equation (20) can be rewritten as follows

$$\begin{aligned}L_i(\theta_i) &= \mathbf{1}[D_i = 1]P_{Di}(\theta_i)P_{1i}(\theta_i)^{I(Y_{1i}=1)}(1 - P_{1i}(\theta_i))^{I(Y_{1i}=0)} + \\ &\quad \mathbf{1}[D_i = 0](1 - P_{Di}(\theta_i))\; P_{0i}(\theta_i)^{I(Y_{0i}=1)}(1 - P_{0i}(\theta_i))^{I(Y_{0i}=0)}\end{aligned} \tag{22}$$

Let $\phi(\cdot)$ denote the standard normal probability density function. Then the unconditional likelihood contribution of individual i is obtained by integrating out the unobserved factor θ_i from $L_i(\theta_i)$ [17]

---

[17] When integrating out θ_i we assume that $\theta \perp (X, Z)$. This is a standard assumption in random effects models and it can be viewed as a solution to a missing conditioning variables problem in matching; see Heckman and Vytlacil (forthcoming) for further details on this point and discussion of various identifying assumptions.

$$L_i = \int_{-\infty}^{\infty} L_i(\theta_i)\phi(\theta_i)d\theta_i \tag{23}$$

The full likelihood function L is given by

$$L = \prod_{i=1}^{N} L_i \tag{24}$$

where N is the sample size.

The parameters of the model are estimated by the maximum likelihood method and we approximate the integral in (23) by Gauss-Hermite quadrature, cf. appendix B.

To obtain estimates of the different mean treatment parameters and the probabilities characterizing the distribution of the potential outcomes and the treatment effect we insert the estimated parameters from the maximum likelihood procedure into the equations given in (17), (18), (19) and appendix A. The integrals appearing in these expressions are again approximated by Gauss-Hermite quadrature. To compute the associated standard errors we use the delta method (see e.g. Ruud (2000)). From these estimated treatment parameters conditional on X and Z, we estimate the corresponding treatment parameters integrated over the distribution of X and Z by taking sample averages, i.e. integrating with respect to the empirical distribution of X and Z. This means that we, for example, estimate Δ^{ATE} by $\frac{1}{N}\sum_{i=1}^{N}\Delta^{ATE}(x_i)$.

5   Results

We have estimated two versions of the econometric model, one without selection on unobservables ($\alpha_1 = \alpha_0 = 0$) and one with selection on unobservables ($\alpha_1 \neq 0$, $\alpha_0 \neq 0$ and $\alpha_1 \neq \alpha_0$). Below we mainly discuss the results from the latter, since that is the most satisfactory model where we have allowed a more flexible structure of the error terms across equations. We will briefly comment on the differences between the results from the two versions of the model.

Before looking at the estimation results, we notice that the raw data (reported in Table 1) show that participants in PSE programmes have a 18.4 percentage points higher employment rate than participants in other programmes when we compare their employment status 12 months after the end of the programmes. It is even higher, 21.9 percentage points, if we compare their employment status only 6 months after the end of the programmes.

When we control for the observable characteristics of participants, we find that the average treatment effect falls to 13.6 percentage points (see Table 5, last column). When we also control for selection on unobservables the effect increases slightly to 17.2 percentage points, but now it is no longer statistically significant.

5.1 Coefficient estimates

Estimates of the parameters of the econometric model are reported in Tables 2, 3 and 4. The estimates of the parameters of the selection equation are reported in Table 2, while the estimates of the parameters of the employment equations for participants in PSE programmes and participants in non-PSE programmes are reported in Table 3 and Table 4, respectively.

The estimated parameters of the *selection equation* clearly show that there is non-random selection into PSE programmes. The persons that participate in PSE programmes differ significantly from participants in other programmes with respect to observable characteristics. Persons with more work experience, less previous unemployment, and below 30 years of age have a significantly higher probability of being assigned to a PSE programme. The estimated coefficients also show a decreasing trend over the period 1993-1998 in the probability of being assigned to a PSE programme. Persons living in larger municipalities and in municipalities where the relative importance of PSE programmes is high have a higher probability of participating in a PSE programme.

For both *employment equations*, we find that a number of background characteristics have an impact on the probability of getting employed. The estimated coefficients reflect very reasonable effects: younger persons, better-educated persons, persons with more work experience, and persons with less previous unemployment have a significantly higher probability of being employed 12 months after the end of the programme. Also in accordance with our a priori expectations, we find that persons in regions with a high relative unemployment level have the lowest employment probabilities. All of the effects are very similar for the two employment equations.

Comparing the estimation results for the two versions of the model, with and without selection on unobservables, we find that the coefficient estimates are very similar, in particular for the selection equation where many of them are even identical to two digits significance. Also for the two employment equations, the coefficient estimates are quite similar. This is consis-

tent with the result that for both employment equations the common unobserved factor have a loading coefficient not significantly different from zero.[18]

In the model that controls for selection into programmes, we have used the relative importance of PSE programmes in the municipality as an instrument. Table 2 shows that the effect of this variable is significantly different from zero (with a *z*-value of more than 11). Thus, it satisfies the first requirement for being a valid instrument, namely that it is correlated with the assignment into programme category. To be a valid instrument it also has to satisfy the requirement that it should not affect the employment outcome directly, but only indirectly through the actual programme choice, cf. the discussion in Section 3. As evidence on this, we have calculated the correlation between the instrument and the indicator variable for being employed 12 months after the end of the ALMP period. This correlation is only 0.005 and a statistical test cannot reject that it is different from zero (the *p*-value is 0.69). We have also estimated the model with the instrument included in the two outcome equations. The estimated coefficients of this variable were close to zero (with *z*-values of 0.1 and 1.2, respectively). Thus, there is no indication that the relative importance of PSE programmes in the municipality affects the employment outcome directly and it appears to be a valid instrument. It is not possible to test the exclusion restriction more formally, since the exclusion restriction is almost an identifying assumption in our model. Without the exclusion restriction, the model is only parametrically identified through the normality assumption.

5.2 Mean treatment effects

Based on the maximum likelihood estimates of the parameters, we calculate the different treatment effects. They are reported in Tables 5, 6 and 7. Let us first comment on the mean treatment effects.

The average treatment effect indicates that PSE programmes have a positive employment effect (0.172), although it is not significantly positive when we correct for selection on unobservables. For those who are selected into the PSE programmes, we find that the programmes have a negative employment effect (-0.139), but again the effect is not statistically significant. This result suggests that selection into PSE programmes is negatively related to the expected

---

[18] However, even if the common unobserved factor has loading coefficient estimates that are not individually significantly different from zero, the standard likelihood ratio test rejects the hypothesis that there is no selection on unobservables at a 5 per cent level of significance. Under the hypothesis that $\alpha_1=\alpha_0=0$, the likelihood ratio test statistic is chi-square distributed with 2 degrees of freedom. The value of the likelihood ratio test statistic is 6.08, while the critical value at the 5 per cent level of significance is 5.99.

outcome, a sort of perverse selection since those being assigned to the programmes cannot expect to benefit from them.

Previous analyses of the employment effect have typically found that PSE programmes have a positive effect. Our results indicate that these previous findings may be due to the use of models with too restrictive assumptions, primarily about the role of unobservables. By relaxing the restrictive assumptions we have obtained a more diverse picture, although this to some degree occurs at the expense of increased standard errors. However, our results question the conventional wisdom on the employment effects of PSE programmes.

The large difference between the estimated employment effect (-0.139) and the raw mean effect (0.184) shows that controlling for selection is very important in our setting, despite the fact that the factor loadings are not individually significantly different from zero. Our estimation results show that the factor loadings are large, but very imprecisely estimated. This result is remarkably similar to the results obtained by Aakvik et al. (2005) and Aakvik et al. (2003) using the same empirical model.

The calculation of the marginal treatment effects allows us to investigate further the relationship between the unobservables and the treatment effect. In Table 7, we report the mean MTE for three different values of the unobserved characteristics in the selection equation, U_D. It is seen that the mean MTE is decreasing in U_D. This means that those most likely to participate in PSE programmes (those having large U_D values) benefit the least from these programmes. This is in accordance with the evidence reported above on ATE and TT, since they also showed that a random person would be better off than an actual participant in the PSE programmes.

A possible explanation for this result is that the caseworkers doing the allocation to different programmes select the most (unconditionally) employable persons for PSE programmes, thereby cream-skimming the group of unemployed persons. The reasons behind this may be that the private employers, directly or indirectly, impose some requirements on the potential participants in PSE programmes. This may lead to a selection where persons who would have been employed anyway are chosen for PSE programmes. Below we provide further evidence on the selection on observables and unobservables, see Section 5.4.

5.3 Distributional treatment effects

We have also calculated estimates of the distributional treatment effects (reported in Tables 5, 6 and 7). These results show that if a random ALMP-participant is assigned to a PSE pro-

gramme, then the participant benefits from the programme with probability 0.345, but will be hurt by it with probability 0.174. In other words, roughly twice as many participants will benefit as those being hurt by the programme. But nearly half of the population will not be affected by the type of programme, since they will be either employed or unemployed regardless of whether they participate in a PSE programme or in another programme.

Looking at the effect of treatment on the treated, the picture is a bit different. For those actually participating in the PSE programmes, the probability of benefiting from the programme is only 0.172, whereas the probability of being hurt by it is almost twice as high, 0.312. More than half of those participating in a PSE programme are unaffected by it.

The distributional marginal treatment effects provide further insight into the consequences of selection into the programmes. The distributional impact of the PSE programmes is most negative for those most likely to participate in the programmes (those having large U_D values). When we calculate the distributional MTE for a large value of U_D (in this case equal to 2) the results resemble those for TT. A much better distribution of effects appears for those persons least likely to enter a PSE programme (with $U_D = -2$), where we find that more than half of the potential participants would benefit from participation, i.e. they would become employed because of participation in a PSE programme as opposed to another type of programme. The probability of becoming unemployed would be as low as 0.056 for these persons.

These distributional treatment effects show that there is a considerable amount of heterogeneity in the response to the programmes. They also illustrate that the mean treatment effects hide this heterogeneity in the individual responses.

5.4 Selection on observables and unobservables

As discussed above, there is a considerable amount of heterogeneity in the response to the programmes. The estimated treatment effects vary substantially with observables as well as with unobservables. First, consider the estimated average treatment effect, $\Delta^{ATE}(x)$, which has a standard deviation of 0.087, compared to its mean of 0.172. Similarly, the estimated treatment effect for those participating in PSE programmes, $\Delta^{TT}(x,z)$, has a standard deviation of 0.096, compared to its mean of –0.139. Hence, both these treatment effects show a considerable variation with the observed variables.

Additional information on this heterogeneity can be seen from Table 8 that reports the marginal effects of each observable variable on the mean treatment effects. It is seen that being

married, having no or little work experience, having spent the last 12 months in ordinary education, and living in a region with high unemployment are all associated with smaller treatment effects. This has potentially important implications for the allocation of unemployed persons to programmes, if caseworkers would take these differences into account in the allocation. At present, e.g. married men are over-represented in PSE programmes, but the employment effect of participating in PSE programmes are much smaller for married men than for single men.

We can also gain more insight into the relationship between selection into programmes and employment outcomes by looking at correlations between unobservables and observables, respectively. Given our factor structure model, it is straightforward to calculate the correlations between the unobservables from the estimated parameters (in particular, note that they do not depend on values of X and Z):

$$corr(U_0, U_1) = -0.165$$

$$corr(U_D, U_0) = 0.442$$

$$corr(U_D, U_1) = -0.186$$

The first correlation shows that the unobservables determining the two employment outcomes are negatively correlated, indicating that those individuals with unobservables that make them more likely to be employed after having participated in a PSE programme will be less likely to be employed after having participated in a non-PSE programme. However, the correlation is only –0.165, so it is fairly weak. The other two correlations show that unobservables that are favouring selection into PSE programmes are positively correlated with unobservables that increase the employment chances after a non-PSE programme but negatively correlated with unobservables that increase the employment chances after a PSE programme. Hence, those most likely to participate in a PSE programme (holding constant X and Z) have the lowest treatment effects. This once again illustrates the main point.

We can in the same way examine the dependence between the observables. Since our model captures the effects of the observable variables through the indices $X\beta_1$, $X\beta_0$ and $Z\beta_D$, we can estimate the correlations of these indices:

$$corr(X\beta_0, X\beta_1) = 0.851$$

$$corr(Z\beta_D, X\beta_0) = 0.497$$

$$corr(Z\beta_D, X\beta_1) = 0.399$$

All the correlations are positive and strong. Thus, the observed variables that increase the probability of participating in a PSE programme are also associated with better employment chances after both types of programmes.

Finally, we can also use the correlations to illustrate that individuals most likely to participate in a PSE programme are those who benefit the least from it. This holds for both unobservables and observables, which can be seen from the following two correlations:

$$corr(U_D, U_1 - U_0) = -0.428$$

$$corr(Z\beta_D, X(\beta_1 - \beta_0)) = -0.412$$

We see that the unobservables and the observables reinforce one another, and furthermore it can be noticed that the two correlations are of the same magnitude, such that the impact of the two types of heterogeneity is similar in size.

5.5 Sensitivity analyses

In this subsection we report the results of some estimations that we have performed to test the sensitivity of the parameter estimates to different assumptions. First, let us consider our use of an instrument to improve on the identification of the model. As described in Section 4, we include a variable in Z, but not in X, i.e. we use an exclusion restriction although this is not formally necessary to obtain identification in the presence of a normal factor structure. We have also estimated the model without this variable and we obtain results (not reported in the paper) that are basically the same as those obtained with the instrument included. However, the main difference is that the standard errors of the estimated parameters are much larger without the instrument, which confirms that the instrument improves on the empirical identification of the model.

The results may also depend on the assumption that the common factor θ is normally distributed. Therefore, we have estimated the model with an alternative assumption, namely that θ follows a discrete distribution with a finite number of support points (corresponding to the well-known Heckman-Singer procedure (Heckman and Singer (1984))). We have used three points of support and the estimation results obtained from using the discrete distribution for θ (reported in Tables 9 and 10) are fairly similar to the results obtained by using the normality

assumption.[19,20] The results on the mean treatment effects and the distributional treatment effects (reported in Tables 11, 12 and 13) are qualitatively similar, but they vary considerably in magnitude. Another difference is that the standard errors are smaller when we use the discrete distribution for θ.[21] With the alternative assumption about the distribution of the common factor, we find a less favourable mean ATE (0.082) and an even larger negative mean TT (-0.275). ATE is still insignificant, whereas TT is now strongly significant. In their original analysis of this model, Aakvik et al. (2005) also investigated the sensitivity to the normality assumption and they found that their results were robust with respect to this assumption.

6   Conclusion

In this paper, we have taken a new look at the employment effects of private sector employment programmes taking into account selection into programmes and heterogeneous treatment effects. Our results give a more diverse picture of the employment effects than previous analyses. Even though the raw data and the model without selection on unobservables provide a significantly positive mean effect of participation in PSE programmes relative to other ALMPs, the model that takes into account selection on unobservables gives an insignificant (and even negative) employment effect. We also find evidence of substantial heterogeneity in the employment effects of PSE programmes. Last, but not least, our results show that those most likely to be assigned to PSE programmes are those who benefit the least from it.

Our results thus suggest that there is room for improvement in the way participants are allocated between programmes. If unemployed persons were allocated such that those participat-

---

[19] Let v_1, v_2 and v_3 (where $v_1 < v_2 < v_3$) denote the three points of support, and let p_1, p_2 and p_3 (where $p_1 + p_2 + p_3 = 1$) denote the probabilities associated with each of the three points. We normalize the mean of the common factor θ to 0 and the variance to 1. To simplify the calculations, we assume that v_2 is equal to 0. This means that $v_1 = -\sqrt{p_3/(p_1(1-p_2))}$ and $v_3 = v_1 - (1/(p_1 v_1))$. Hence, compared to the model with a normally distributed common factor, only two additional parameters (p_1 and p_2) have to be estimated. When p_1 and p_2 are given, p_3 can easily be calculated from the restriction that the probabilities p_1, p_2 and p_3 have to sum to 1.

[20] Assuming that there is selection on unobservables (i.e. $\alpha_1 \neq 0$, $\alpha_0 \neq 0$ and $\alpha_1 \neq \alpha_0$), then the model with a normally distributed common factor is not nested within the model where the common factor follows a discrete distribution. But according to the Akaike Information Criterion the latter is the preferred model.

[21] A major reason for the smaller standard errors on the mean and distributional treatment parameters obtained with the alternative model (where θ follows a discrete distribution with three points of support) is that the estimated value of the parameter α_0 is much larger than in the model with a normally distributed common factor. The estimates of α_0 in the two models are 0.801 and 3.905, whereas the standard errors are relatively similar (0.534 and 0.732, respectively).

ing in PSE programmes were those who would benefit most from it, then the overall employment chances of the participants would be considerable improved. The optimal allocation could be based on observables, but it would make an even bigger difference if unobservables were also used in an optimal way. Currently, the caseworkers doing the allocation use this information perversely. This might be changed e.g. by creating the right incentives for the caseworkers. An optimal allocation would require that the caseworkers obtain as much information as possible from the unemployed persons and use it optimally in the allocation. Otherwise, it will not be possible to reap the optimal employment gains from the ALMPs.

The focus that we have taken in our analysis implies some limitations of our study: First, we focus on employment effects – for obvious and good reasons, described in the introduction – but that does not rule out the possibility of other types of effects. For example, the programmes may have effects on occupational choice, on subsequent earnings, or on well-being, but such effects are beyond the scope of our analysis. Second, our analysis does not provide evidence on whether participation in any programme is beneficial or not compared to non-participation. Again, we already gave the reasons for this in the introduction of the paper.

One caveat that should be kept in mind is that the empirical results may depend on the method and the model specification, e.g. the results may be sensitive to the distributional assumptions. Clearly, a model with a more general error structure than the one-factor structure model might yield other results. Also, we have employed a normality assumption, which may be overly restrictive (but obviously it has some very nice properties that make the computations tractable). In line with Aakvik et al. (2005) we have investigated the sensitivity to the normality assumption for the common factor and we find that the results are qualitatively robust with respect to this assumption, even though the magnitudes of the effects vary. Another way to investigate the importance of the various assumptions would be to perform a semi-parametric estimation of the model, but that is beyond the scope of this paper and will be left for future research.

References

Aakvik, A., J. J. Heckman and E. J. Vytlacil (2005) Estimating treatment effects for discrete outcomes when responses to treatment vary: an application to Norwegian vocational rehabilitation programs, *Journal of Econometrics*, Vol. 125, 15-51.

Aakvik, A., T. H. Holmås, and E. Kjerstad (2003) A low-key social insurance reform – effects of multidisciplinary outpatient treatment for back pain patients in Norway, *Journal of Health Economics*, Vol. 22, 747-762.

Abbring, J. H. and G. J. van den Berg (2003) The nonparametric identification of treatment effects in duration models, *Econometrica*, Vol. 71, 1491-1517.

Andrén, D. and T. Andrén (2002) Assessing the employment effects of labor market training programs in Sweden, Working Papers in Economics no. 70, Göteborg University.

Bolvig, I., P. Jensen, and M. Rosholm (2003) The employment effects of active social policy, IZA Discussion Paper No. 736, Institute for the Study of Labor, Bonn.

Brogaard, S. og Weise, H. (1997) Evaluering af Lov om kommunal aktivering – Kommuneundersøgelsen. Socialforskningsinstituttet 97:7. København.

Butler, J. S. and R. Moffitt (1982) A computationally efficient quadrature procedure for the one-factor multinomial probit model, *Econometrica*, Vol. 50, 761-764.

Carling, K. and K. Richardson (2004) The relative efficiency of labor market programs: Swedish experience from the 1990s, *Labour Economics*, Vol. 11, 335-354.

Danish Ministy of Labour (2001) Offers to the unemployed, Copenhagen.

Frölich, M. and M. Lechner (2004) Regional treatment intensity as an instrument for the evaluation of labour market policies, IZA Discussion Paper No. 1095, Bonn.

Gerfin, M. and M. Lechner (2002) A microeconometric evaluation of the active labour market policy in Switzerland, *The Economic Journal*, Vol. 112, 854-893.

Graversen, B. K. (2004a) The impact of active labour market programmes on welfare dependence in Denmark. Chapter 1 of PhD thesis 2004-2, Department of Economics, University of Aarhus.

Graversen, B. K. (2004b) Treatment effects of a sequence of active labour market programmes. Chapter 2 of PhD thesis 2004-2, Department of Economics, University of Aarhus.

Heckman, J. J. (1981) Statistical models for discrete panel data, in: C. Manski and D. McFadden (eds.) *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press.

Heckman, J. J. (1990) Varieties of selection bias, *American Economic Review*, Vol. 80, 313-318.

Heckman, J. J. (1997) Instrumental variables: a study of implicit behavioural assumptions used in making program evaluations, *Journal of Human Resources*, Vol. 32, 441-462.

Heckman, J. J. (2001) Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture, *Journal of Political Economy*, Vol. 109, 673-748.

Heckman, J. J., R. J. Lalonde, and J. A. Smith (1999) The economics and econometrics of active labor market programs, in: O. Ashenfelter and D. Card (eds.), Handbook of Labor Economics, Vol. 3, North-Holland.

Heckman, J. J. and B. Singer (1984) A method for minimizing the impact of distributional assumptions in econometric models for duration data, *Econometrica*, Vol. 52, 271-320.

Heckman, J. J. and E. J. Vytlacil (2000) The relationship between treatment parameters within a latent variable framework, *Economics Letters*, Vol. 66, 33-39.

Heckman, J. J. and E. J. Vytlacil (forthcoming) Econometric evaluation of social programs, in: J. J. Heckman and E. Leamer (eds.), Handbook of Econometrics, Vol. 6, North-Holland.

Imbens, G. and J. Angrist (1994) Identification and estimation of local average treatment effects, *Econometrica*, Vol.62, 467-476.

Judd, K. L. (1998) Numerical Methods in Economics, The MIT Press.

Lechner, M. and J. A. Smith (forthcoming) What is the value added by caseworkers? *Labour Economics*, in press (2005).

Martin, J. P. and D. Grubb (2001) What works and for whom: A review of OECD countries' experiences with active labour market policies, *Swedish Economic Policy Review*, Vol. 8, 9-56.

29

Ministry of Environment and Energy (2001) Pendlingen i Danmark år 2000 og udviklingen i 1990'erne, Arbejdsnotat udarbejdet af Landsplanafdelingen, 4. kontor, Miljø- og Energiministeriet.

Ruud, P.A. (2000) An introduction to classical econometric theory. Oxford University Press.

Sianesi, B. (2002) Differential effects of Swedish active labour market programmes for unemployed adults during the 1990s, Working Paper 2002:5, IFAU – Office of Labour Market Policy Evaluation, Uppsala.

Appendix A

The assumptions of our model specified in Section 4 imply that the probabilities characterizing the joint distribution of the potential outcomes Y_1 and Y_0 and the distribution of the treatment effect can be expressed by relatively simple formulas.

For individuals with observed characteristics x and z the following expressions apply if the individuals are randomly selected into a PSE programme:

$$P_{Y_1,Y_0}^{ATE}(1,0,x) = \int [\Phi(x\beta_1 + \alpha_1\theta)(1 - \Phi(x\beta_0 + \alpha_0\theta))]\phi(\theta)d\theta$$

$$P_{Y_1,Y_0}^{ATE}(0,1,x) = \int [(1 - \Phi(x\beta_1 + \alpha_1\theta))\Phi(x\beta_0 + \alpha_0\theta)]\phi(\theta)d\theta$$

$$P_{Y_1,Y_0}^{ATE}(1,1,x) = \int [\Phi(x\beta_1 + \alpha_1\theta)\Phi(x\beta_0 + \alpha_0\theta)]\phi(\theta)d\theta$$

$$P_{Y_1,Y_0}^{ATE}(0,0,x) = \int [(1 - \Phi(x\beta_1 + \alpha_1\theta))(1 - \Phi(x\beta_0 + \alpha_0\theta))]\phi(\theta)d\theta \qquad (A1)$$

$$P_{\Delta}^{ATE}(1,x,z) = P_{Y_1,Y_0}^{ATE}(1,0,x,z)$$

$$P_{\Delta}^{ATE}(-1,x,z) = P_{Y_1,Y_0}^{ATE}(0,1,x,z)$$

$$P_{\Delta}^{ATE}(0,x,z) = P_{Y_1,Y_0}^{ATE}(0,0,x,z) + P_{Y_1,Y_0}^{ATE}(1,1,x,z)$$

If the individuals are selected into a PSE programme using the selection mechanism implied by our model the following results apply:

$$P_{Y_1,Y_0}^{TT}(1,0,x,z) = \frac{1}{\Phi(z\beta_D/\sqrt{2})} \int [\Phi(x\beta_1 + \alpha_1\theta)(1 - \Phi(x\beta_0 + \alpha_0\theta))]\Phi(z\beta_D + \theta)\phi(\theta)d\theta$$

$$P_{Y_1,Y_0}^{TT}(0,1,x,z) = \frac{1}{\Phi(z\beta_D/\sqrt{2})} \int [(1 - \Phi(x\beta_1 + \alpha_1\theta))\Phi(x\beta_0 + \alpha_0\theta)]\Phi(z\beta_D + \theta)\phi(\theta)d\theta$$

$$P_{Y_1,Y_0}^{TT}(1,1,x,z) = \frac{1}{\Phi(z\beta_D/\sqrt{2})} \int [\Phi(x\beta_1 + \alpha_1\theta)\Phi(x\beta_0 + \alpha_0\theta)]\Phi(z\beta_D + \theta)\phi(\theta)d\theta$$

$$P_{Y_1,Y_0}^{TT}(0,0,x,z) = \frac{1}{\Phi(z\beta_D/\sqrt{2})} \int [(1 - \Phi(x\beta_1 + \alpha_1\theta))(1 - \Phi(x\beta_0 + \alpha_0\theta))]\Phi(z\beta_D + \theta)\phi(\theta)d\theta \quad (A2)$$

$$P_{\Delta}^{TT}(1,x,z) = P_{Y_1,Y_0}^{TT}(1,0,x,z)$$

$$P_{\Delta}^{TT}(-1,x,z) = P_{Y_1,Y_0}^{TT}(0,1,x,z)$$

$$P_{\Delta}^{TT}(0,x,z) = P_{Y_1,Y_0}^{TT}(0,0,x,z) + P_{Y_1,Y_0}^{TT}(1,1,x,z)$$

If the individuals are selected into a PSE programme using the selection mechanism implied by our model and the unobserved variable U_D of the individuals is equal to u then the following results apply:

$$P_{Y_1,Y_0}^{MTE}(1,0,x,u) = \frac{\sqrt{2}}{\phi(u/\sqrt{2})}\int[\Phi(x\beta_1+\alpha_1\theta)(1-\Phi(x\beta_0+\alpha_0\theta))]\phi(u-\theta)\phi(\theta)d\theta$$

$$P_{Y_1,Y_0}^{MTE}(1,0,x,u) = \frac{\sqrt{2}}{\phi(u/\sqrt{2})}\int[(1-\Phi(x\beta_1+\alpha_1\theta))\Phi(x\beta_0+\alpha_0\theta)]\phi(u-\theta)\phi(\theta)d\theta$$

$$P_{Y_1,Y_0}^{MTE}(1,1,x,u) = \frac{\sqrt{2}}{\phi(u/\sqrt{2})}\int[\Phi(x\beta_1+\alpha_1\theta)\Phi(x\beta_0+\alpha_0\theta)]\phi(u-\theta)\phi(\theta)d\theta$$

$$P_{Y_1,Y_0}^{MTE}(0,0,x,u) = \frac{\sqrt{2}}{\phi(u/\sqrt{2})}\int[(1-\Phi(x\beta_1+\alpha_1\theta))(1-\Phi(x\beta_0+\alpha_0\theta))]\phi(u-\theta)\phi(\theta)d\theta \quad (A3)$$

$$P_\Delta^{MTE}(1,x,z) = P_{Y_1,Y_0}^{MTE}(1,0,x,z)$$

$$P_\Delta^{MTE}(-1,x,z) = P_{Y_1,Y_0}^{MTE}(0,1,x,z)$$

$$P_\Delta^{MTE}(0,x,z) = P_{Y_1,Y_0}^{MTE}(0,0,x,z) + P_{Y_1,Y_0}^{MTE}(1,1,x,z)$$

Appendix B

The estimation of the model specified in Section 4 involves numerical integration. The following integral have to be calculated during the iterations of the maximum likelihood estimation procedure:

$$L_i = \int_{-\infty}^{\infty} L_i(\theta_i)\phi(\theta_i)d\theta_i = \int_{-\infty}^{\infty} L_i(\theta)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\theta^2}d\theta \qquad (B1)$$

where $L_i(\theta)$ is a known function except for a finite number of parameters ($\beta_D, \beta_1, \beta_0, \alpha_1, \alpha_0$).

Using the linear change of variables, $q = (1/\sqrt{2})\theta$, equation (B1) can be written as

$$\begin{aligned} L_i &= \int_{-\infty}^{\infty} L_i(\sqrt{2}q)\frac{1}{\sqrt{2\pi}}e^{-q^2}\sqrt{2}dq \\ &= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} L_i(\sqrt{2}q)e^{-q^2}dq \end{aligned} \qquad (B2)$$

With the use of the Gauss-Hermite quadrature rule (see e.g. Judd (2000)) this integral can be approximated by

$$L_i \approx \frac{1}{\sqrt{\pi}}\sum_{k=1}^{K} w_k L_i(\sqrt{2}q_k) \qquad (B3)$$

where K is the number of evaluation points used for the approximation, and w_k is the weight associated with the evaluation point q_k ($k=1,...,K$). When K is increased, the accuracy of the approximation in (B3) is improved.

For the empirical analyses in this paper we use 5 evaluation points.[22] The weights and the location of the evaluation points are taken from the online calculator at www.efunda.com.

---

[22] The same choice is made by Aakvik et al. (2005) and Andrén and Andrén (2002), who estimate a similar model. In a somewhat different setting (a panel probit model), Butler and Moffit (1982) show that the estimated coefficients change very little when the number of evaluation points is increased.

Figure 1
Proportion of participants in different types of ALMPs who are employed after the end of the programme period
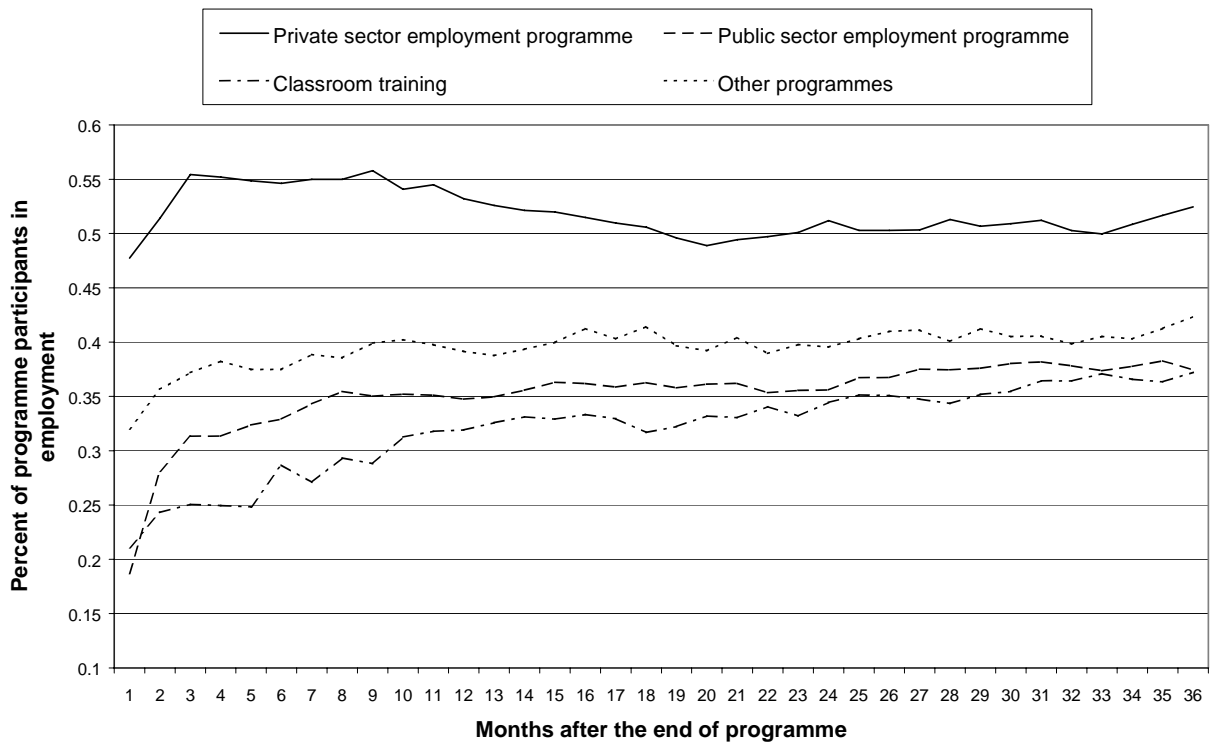
Figure 2
Proportion of participants in different types of ALMPs who are unemployed after the end of the programme period
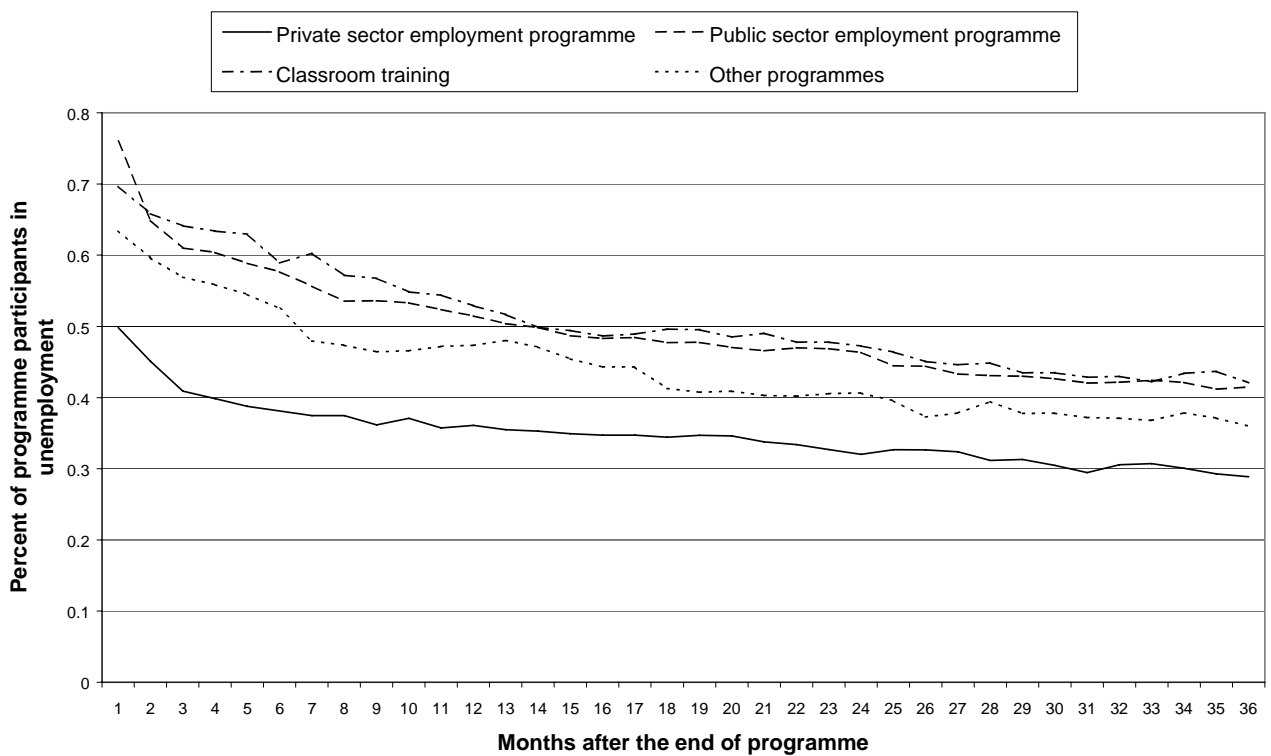
Figure 3
Proportion of participants in different types of ALMPs who are outside the labour market after the end of the programme period
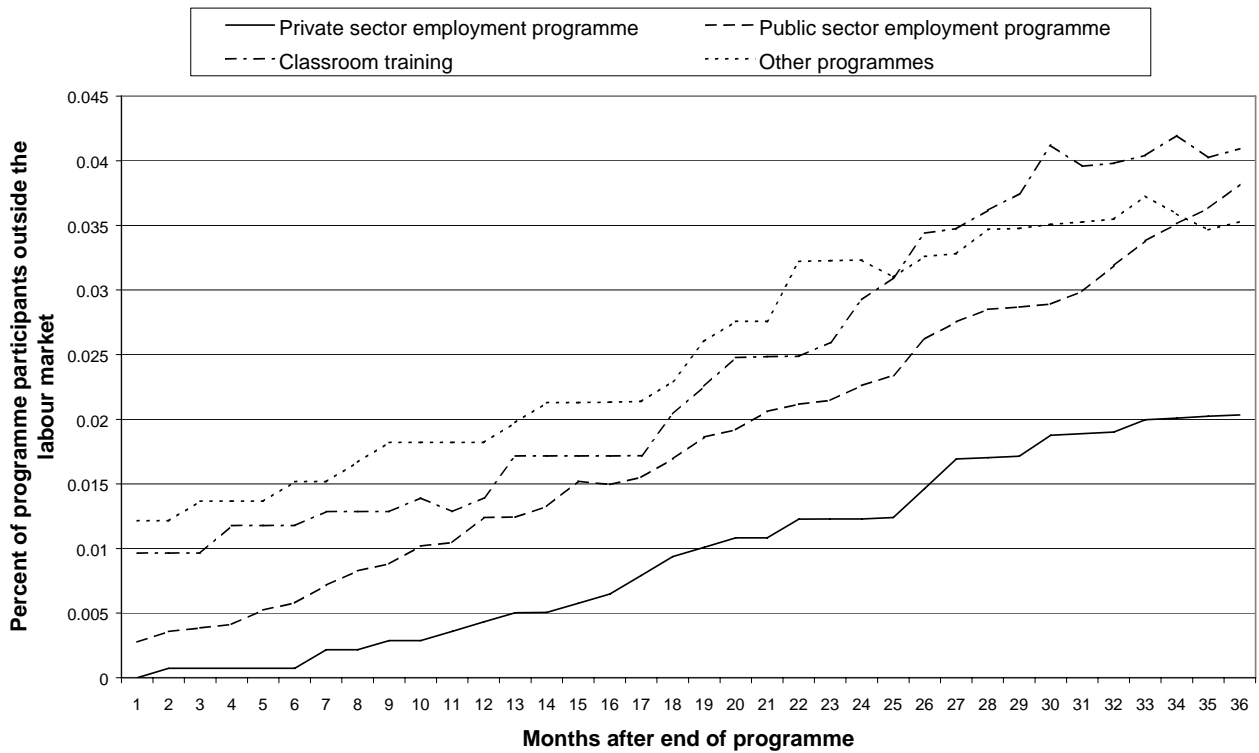
Figure 4
Proportion of participants in different types of ALMPs who are in ordinary education after the end of the programme period
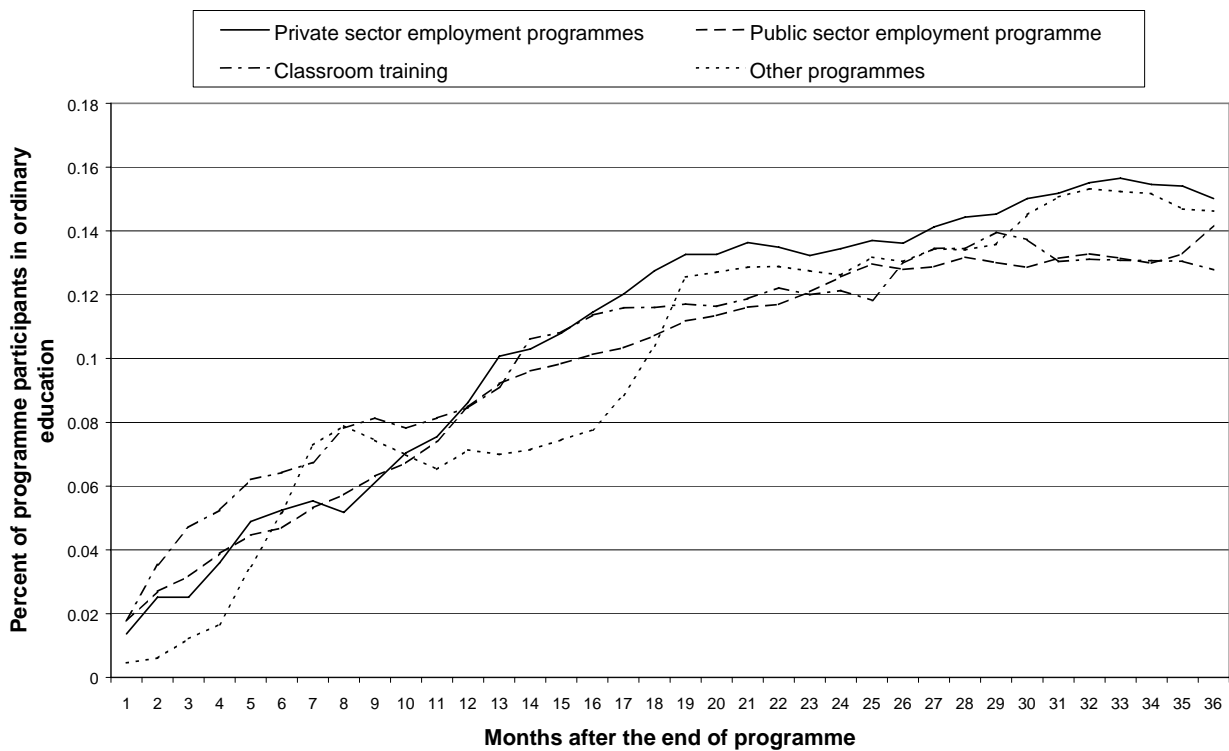
Figure 5
Proportion of participants in different types of ALMPs who cannot be categorized after the end of the programme period
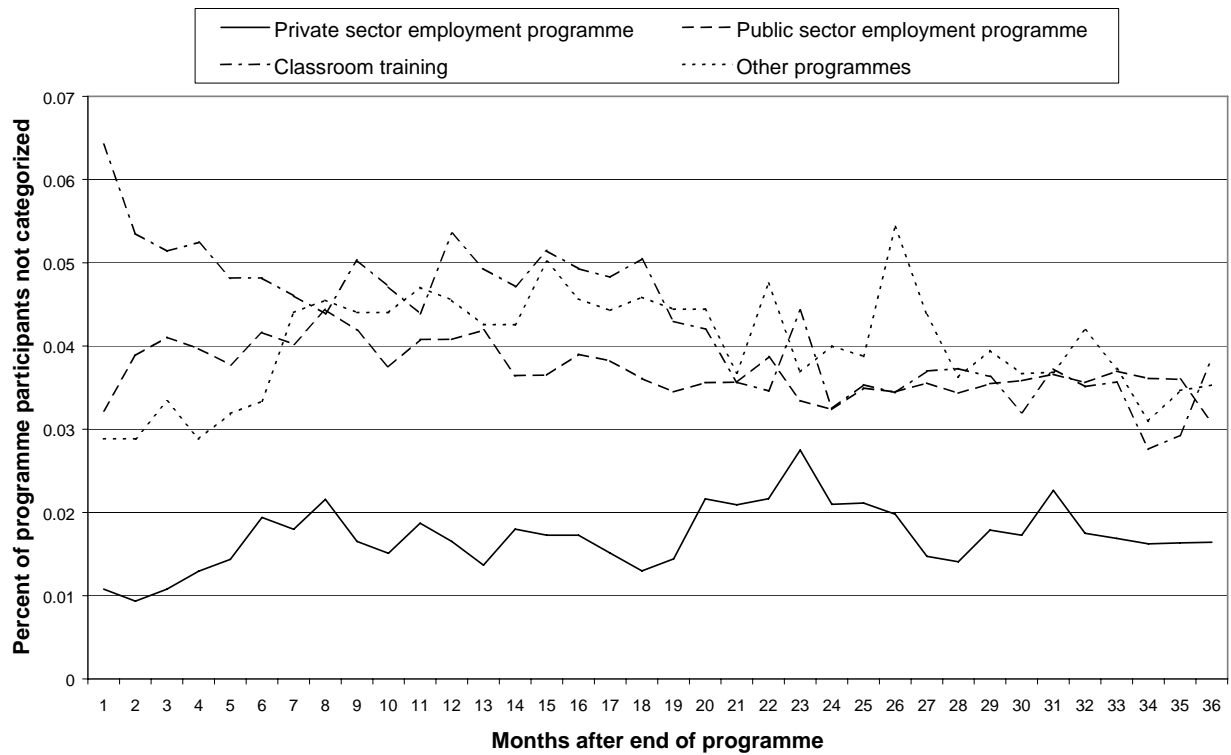
Table 1

Descriptive statistics (mean values) for participants in private sector employment (PSE) programmes and participants in non-PSE programmes

	Participants in PSE programmes	Participants in non-PSE programmes
Number of observations	1391	5222
Outcome variables		
Employment status:		
6 months after end of programme period	0.546	0.327
12 months after end of programme period	0.532	0.348
Individual characteristics		
Year when programme started:		
1993	0.201	0.074
1994	0.382	0.326
1995	0.206	0.225
1996	0.101	0.155
1997	0.071	0.134
1998	0.040	0.086
Marital status:		
Single	0.764	0.817
Married	0.069	0.054
Cohabitating	0.167	0.129
Has children	0.134	0.102
Age:		
17-24	0.558	0.517
25-29	0.142	0.121
30-39	0.175	0.199
40-49	0.096	0.129
50-66	0.030	0.034
Completed education:		
Primary or lower secondary school	0.579	0.624
Upper secondary school	0.209	0.183
Vocational education	0.191	0.158
Further or higher education	0.021	0.034
Work experience:		
0-2 years	0.512	0.572
2-5 years	0.207	0.172
5-10 years	0.157	0.142
10+ years	0.124	0.114
Fraction of time spent in different states during the 12 months preceding programme period:		
Employment	0.370	0.346
Unemployment	0.475	0.509
Ordinary education	0.155	0.145

(continued)

Fraction of time spent in different states during 24 months period starting 3 years (and ending 1 year) before the programme period:		
Employment	0.307	0.262
Unemployment	0.320	0.373
Ordinary education	0.093	0.090
No available information regarding the variable above	0.280	0.275
Characteristics of municipality		
Number of residents in municipality in 1996:		
Less than 20,000	0.273	0.320
20,000-40,000	0.183	0.164
40,000-100,000	0.300	0.232
More than 100,000	0.244	0.284
Regional unemployment rate relative to countrywide unemployment rate	1.013	1.028
Percentage of programme participants in PSE programmes relative to countrywide importance of PSE programmes	1.222	0.992

Table 2
Estimates of the parameters determining selection into PSE programmes

	Model *without* selection on unobservables ($\alpha_1=\alpha_0=0$)			Model *with* selection on unobservables ($\alpha_1\neq0$, $\alpha_1\neq0$ and $\alpha_1\neq\alpha_0$)		
	Coeff.	Std.[a]	Marg.[b]	Coeff.	Std.[a]	Marg.[b]
Constant	**-1.480**	0.251	-0.277	**-1.485**	0.244	-0.278
Age (ref. group: 17-24):						
25-29	-0.070	0.099	-0.013	-0.072	0.099	-0.013
30-39	**-0.401**	0.106	-0.075	**-0.394**	0.106	-0.074
40-49	**-0.684**	0.129	-0.128	**-0.672**	0.129	-0.126
50-66	**-0.496**	0.180	-0.093	**-0.487**	0.180	-0.091
Marital status (Ref. group: Single)						
Married	0.193	0.134	0.036	0.176	0.134	0.033
Cohabitating	**0.247**	0.086	0.046	**0.241**	0.086	0.045
Has children	0.164	0.109	0.031	0.172	0.109	0.032
Completed education (Ref. group: Primary or lower sec. school):						
Upper secondary school	0.112	0.082	0.021	0.108	0.082	0.020
Vocational education	0.139	0.078	0.026	0.137	0.078	0.026
Further or higher education	-0.188	0.168	-0.035	-0.182	0.169	-0.034
Work experience (Ref. group: 0-2 years):						
2-5 years	**0.495**	0.085	0.093	**0.488**	0.085	0.091
5-10 years	**0.604**	0.109	0.113	**0.601**	0.109	0.113
10+ years	**0.715**	0.128	0.134	**0.703**	0.128	0.132
Status during the 12 months preceding the programme period (Ref. group: Employment):						
Unemployment	-0.102	0.106	-0.019	-0.114	0.107	-0.021
Ordinary education	-0.028	0.116	-0.005	-0.022	0.117	-0.005
Status during 24 months period starting 3 years (and ending 1 year) before the programme period (Ref. group: Employment):						
Unemployment	**-0.289**	0.116	-0.054	**-0.287**	0.114	-0.054
Ordinary education	-0.137	0.151	-0.026	-0.132	0.151	-0.025
No available information	-0.011	0.099	-0.002	-0.015	0.100	-0.003
Year when programme started (Ref. group: 1994):						
1993	**0.631**	0.084	0.118	**0.635**	0.084	0.119
1995	**-0.196**	0.070	-0.037	**-0.192**	0.071	-0.036
1996	**-0.499**	0.085	-0.093	**-0.497**	0.085	-0.093
1997	**-0.643**	0.095	-0.120	**-0.641**	0.095	-0.120
1998	**-0.765**	0.116	-0.143	**-0.761**	0.116	-0.143
Number of residents in municipality (Ref. group: 0-20,000):						
20,000-40,000	**0.199**	0.081	0.037	**0.197**	0.081	0.037
40,000-100,000	**0.243**	0.073	0.046	**0.239**	0.073	0.045
More than 100,000	**0.186**	0.073	0.035	**0.191**	0.073	0.036
Relative unemployment level	-0.195	0.215	-0.037	-0.191	0.215	-0.036
Relative importance of PSE programmes	**0.473**	0.043	0.089	**0.479**	0.044	0.090
Common unobserved factor	1.000 [c]			1.000 [c]		
Log-likelihood	-7131.7			-7128.7		

Note.　Figures in bold indicate significance at a 5 per cent level.
a.　Asymptotic standard errors.
b.　Mean marginal effect on the probability of participating in a PSE programme.
c.　The parameter is fixed at the given value.

Table 3
Estimates of the parameters determining the employment status 12 months after the end of the programme period (*participants in PSE programmes*)

	Model *without* selection on unobservables ($\alpha_1=\alpha_0=0$)			Model *with* selection on unobservables ($\alpha_1\neq0$, $\alpha_1\neq0$ and $\alpha_1\neq\alpha_0$)		
	Coeff.	Std.[a]	Marg.[b]	Coeff.	Std.[a]	Marg.[b]
Constant	**1.346**	0.369	0.497	**1.578**	0.541	0.553
Age (ref. group: 17-24):						
25-29	0.037	0.135	0.014	0.043	0.139	0.015
30-39	-0.262	0.148	-0.097	-0.228	0.156	-0.080
40-49	**-0.722**	0.186	-0.267	**-0.681**	0.194	-0.239
50-66	**-0.708**	0.251	-0.261	**-0.681**	0.260	-0.239
Marital status (Ref. group: Single)						
Married	0.035	0.182	0.013	0.022	0.188	0.008
Cohabitating	0.026	0.113	0.010	0.006	0.120	0.002
Has children	0.128	0.147	0.047	0.109	0.151	0.038
Completed education (Ref. group: Primary or lower sec. school):						
Upper secondary school	0.028	0.109	0.010	0.017	0.114	0.006
Vocational education	**0.255**	0.105	0.094	**0.250**	0.108	0.088
Further or higher education	0.376	0.268	0.139	0.399	0.281	0.140
Work experience (Ref. group: 0-2 years):						
2-5 years	0.196	0.115	0.072	0.153	0.129	0.054
5-10 years	**0.418**	0.155	0.154	**0.371**	0.167	0.130
10+ years	**0.558**	0.180	0.206	**0.504**	0.193	0.177
Status during the 12 months preceding the programme period (Ref. group: Employment):						
Unemployment	**-0.320**	0.142	-0.118	**-0.311**	0.146	-0.109
Ordinary education	-0.257	0.154	-0.095	-0.263	0.160	-0.092
Status during 24 months period starting 3 years (and ending 1 year) before the programme period (Ref. group: Employment):						
Unemployment	**-0.511**	0.153	-0.189	**-0.497**	0.158	-0.174
Ordinary education	-0.114	0.204	-0.042	-0.109	0.210	-0.038
No available information	0.056	0.134	0.021	0.059	0.137	0.021
Year when programme started (Ref. group: 1994):						
1993	**0.191**	0.098	0.071	0.132	0.120	0.046
1995	0.057	0.098	0.021	0.072	0.103	0.025
1996	0.013	0.125	0.005	0.057	0.141	0.020
1997	-0.142	0.141	-0.053	-0.089	0.160	-0.031
1998	-0.013	0.197	-0.005	0.063	0.225	0.022
Number of residents in municipality (Ref. group: 0-20,000):						
20,000-40,000	0.005	0.107	0.002	-0.006	0.112	-0.002
40,000-100,000	-0.049	0.098	-0.018	-0.074	0.110	-0.026
More than 100,000	0.136	0.103	0.050	0.139	0.105	0.049
Relative unemployment level	**-1.089**	0.332	-0.402	**-1.058**	0.341	-0.371
Relative importance of PSE programmes						
Common unobserved factor				-0.273	0.337	

Note. Figures in bold indicate significance at a 5 per cent level.
a. Asymptotic standard errors.
b. Mean marginal effect on the probability of being employed after a PSE programme.

Table 4
Estimates of the parameters determining the employment status 12 months after the end of the programme period (*participants in non-PSE programmes*)

	Model *without* selection on unobservables ($\alpha_1=\alpha_0=0$)			Model *with* selection on unobservables ($\alpha_1\neq0$, $\alpha_1\neq0$ and $\alpha_1\neq\alpha_0$)		
	Coeff.	Std.[a]	Marg.[b]	Coeff.	Std.[a]	Marg.[b]
Constant	**0.667**	0.156	0.228	**1.034**	0.380	0.286
Age (ref. group: 17-24):						
25-29	-0.020	0.071	-0.007	-0.033	0.089	-0.009
30-39	**-0.357**	0.075	-0.122	**-0.506**	0.171	-0.140
40-49	**-0.504**	0.091	-0.172	**-0.729**	0.242	-0.202
50-66	**-0.660**	0.137	-0.225	**-0.900**	0.294	-0.249
Marital status (Ref. group: Single)						
Married	**0.424**	0.099	0.145	**0.547**	0.179	0.152
Cohabitating	**0.158**	0.063	0.054	**0.238**	0.107	0.066
Has children	0.010	0.081	0.004	0.041	0.103	0.011
Completed education (Ref. group: Primary or lower sec. school):						
Upper secondary school	0.089	0.058	0.030	0.132	0.082	0.037
Vocational education	**0.376**	0.059	0.128	**0.483**	0.137	0.134
Further or higher education	**0.245**	0.109	0.084	0.274	0.145	0.076
Work experience (Ref. group: 0-2 years):						
2-5 years	0.105	0.062	0.036	0.215	0.115	0.060
5-10 years	**0.173**	0.078	0.059	**0.319**	0.153	0.088
10+ years	**0.255**	0.093	0.087	**0.439**	0.194	0.122
Status during the 12 months preceding the programme period (Ref. group: Employment):						
Unemployment	**-0.480**	0.073	-0.164	**-0.599**	0.164	-0.166
Ordinary education	0.045	0.079	0.015	0.060	0.100	0.017
Status during 24 months period starting 3 years (and ending 1 year) before the programme period (Ref. group: Employment):						
Unemployment	**-0.570**	0.083	-0.195	**-0.746**	0.212	-0.207
Ordinary education	-0.115	0.105	-0.039	-0.156	0.138	-0.043
No available information	**-0.177**	0.071	-0.060	**-0.212**	0.099	-0.059
Year when programme started (Ref. group: 1994):						
1993	-0.005	0.075	-0.002	0.127	0.127	0.035
1995	-0.039	0.052	-0.013	-0.082	0.073	-0.023
1996	-0.022	0.058	-0.008	-0.104	0.093	-0.029
1997	-0.042	0.062	-0.014	-0.156	0.113	-0.043
1998	-0.125	0.075	-0.043	-0.273	0.150	-0.076
Number of residents in municipality (Ref. group: 0-20,000):						
20,000-40,000	-0.036	0.057	-0.012	-0.027	0.073	-0.001
40,000-100,000	-0.060	0.051	-0.021	-0.072	0.070	-0.002
More than 100,000	0.006	0.050	0.002	0.032	0.064	0.009
Relative unemployment level	**-0.568**	0.133	-0.194	**-0.732**	0.247	-0.203
Relative importance of PSE programmes						
Common unobserved factor				0.801	0.534	

Note. Figures in bold indicate significance at a 5 per cent level.
a. Asymptotic standard errors.
b. Mean marginal effect on the probability of being employed after a non-PSE programme.

41

Table 5

The average treatment effect (Δ^{ATE}) and associated distributional treatment parameters

	$P_{Y_1,Y_0}^{ATE}(1,0)$	$P_{Y_1,Y_0}^{ATE}(0,1)$	$P_{Y_1,Y_0}^{ATE}(1,1)$	$P_{Y_1,Y_0}^{ATE}(0,0)$	$P_{\Delta}^{ATE}(0)$	Δ^{ATE}
Model *without* selection on unobservables ($\alpha_1=\alpha_0=0$)	**0.298** (0.010)	**0.162** (0.006)	**0.194** (0.006)	**0.346** (0.010)	**0.540** (0.006)	**0.136** (0.016)
Model *with* selection on unobservables ($\alpha_1\neq0$, $\alpha_1\neq0$ and $\alpha_1\neq\alpha_0$)	**0.345** (0.091)	**0.174** (0.027)	**0.242** (0.022)	**0.239** (0.093)	**0.481** (0.073)	0.172 (0.112)

Note.   Figures in bold indicate significance at a 5 per cent level. Standard errors are given in parentheses.

Table 6

The effect of treatment on the treated (Δ^{TT}) and associated distributional treatment parameters

	$P_{Y_1,Y_0}^{TT}(1,0)$	$P_{Y_1,Y_0}^{TT}(0,1)$	$P_{Y_1,Y_0}^{TT}(1,1)$	$P_{Y_1,Y_0}^{TT}(0,0)$	$P_{\Delta}^{TT}(0)$	Δ^{TT}
Model *without* selection on unobservables ($\alpha_1=\alpha_0=0$)	**0.308** (0.009)	**0.164** (0.006)	**0.224** (0.007)	**0.306** (0.009)	**0.529** (0.005)	**0.144** (0.015)
Model *with* selection on unobservables ($\alpha_1\neq0$, $\alpha_1\neq0$ and $\alpha_1\neq\alpha_0$)	**0.172** (0.060)	**0.312** (0.068)	**0.360** (0.060)	**0.157** (0.067)	**0.517** (0.035)	-0.139 (0.123)

Note.   Figures in bold indicate significance at a 5 per cent level. Standard errors are given in parentheses.

Table 7

The marginal treatment effect (Δ^{MTE}) and associated distributional treatment parameters

	$P_{Y_1,Y_0}^{MTE}(1,0)$	$P_{Y_1,Y_0}^{MTE}(0,1)$	$P_{Y_1,Y_0}^{MTE}(1,1)$	$P_{Y_1,Y_0}^{MTE}(0,0)$	$P_{\Delta}^{MTE}(0)$	Δ^{MTE}
Model *without* selection on unobservables ($\alpha_1=\alpha_0=0$)	**0.298** (0.010)	**0.162** (0.006)	**0.194** (0.006)	**0.346** (0.010)	**0.540** (0.006)	**0.136** (0.016)
Model *with* selection on unobservables ($\alpha_1\neq0$, $\alpha_1\neq0$ and $\alpha_1\neq\alpha_0$)						
$U_D=-2$	**0.550** (0.194)	0.056 (0.034)	**0.136** (0.060)	0.257 (0.185)	**0.393** (0.161)	**0.494** (0.228)
$U_D=0$	**0.340** (0.083)	**0.157** (0.035)	**0.252** (0.035)	**0.250** (0.084)	**0.503** (0.050)	0.182 (0.117)
$U_D=2$	**0.158** (0.062)	**0.328** (0.079)	**0.332** (0.063)	**0.182** (0.077)	**0.514** (0.037)	-0.170 (0.137)

Note.   Figures in bold indicate significance at a 5 per cent level. Standard errors are given in parentheses.

Table 8
Marginal effect of explanatory variables on the average treatment effect and the effect of treatment on the treated

	$E_X[\dfrac{\delta\Delta^{ATE}(x)}{\delta x}]$		$E_Z[\dfrac{\delta\Delta^{TT}(x,z)}{\delta z}]$	
	$(\alpha_1=\alpha_0=0)$	$(\alpha_1\neq 0, \alpha_1\neq 0$ and $\alpha_1\neq\alpha_0)$	$(\alpha_1=\alpha_0=0)$	$(\alpha_1\neq 0, \alpha_1\neq 0$ and $\alpha_1\neq\alpha_0)$
Constant	0.269	0.267	0.268	0.112
Age (ref. group: 17-24):				
25-29	0.020	0.024	0.021	0.017
30-39	0.025	0.060	0.030	0.015
40-49	-0.095	-0.037	-0.087	-0.118
50-66	0.036	0.011	-0.027	-0.049
Marital status (Ref. group: Single)				
Married	-0.132	-0.144	-0.137	-0.125
Cohabitating	-0.044	-0.064	-0.047	-0.038
Has children	0.044	0.027	0.043	0.047
Completed education (Ref. group: Primary or lower sec. school):				
Upper secondary school	-0.021	-0.031	-0.022	-0.019
Vocational education	-0.034	-0.046	-0.039	-0.029
Further or higher education	0.055	0.064	0.053	0.047
Work experience (Ref. group: 0-2 years):				
2-5 years	0.036	-0.006	0.034	0.050
5-10 years	0.095	0.042	0.091	0.113
10+ years	0.119	0.055	0.114	0.139
Status during the 12 months preceding the programme period (Ref. group: Employment):				
Unemployment	0.046	0.057	0.051	0.042
Ordinary education	-0.110	-0.109	-0.111	-0.115
Status during 24 months period starting 3 years (and ending 1 year) before the programme period (Ref. group: Employment):				
Unemployment	0.006	0.032	0.013	-0.003
Ordinary education	-0.003	0.005	-0.001	-0.011
No available information	0.081	0.079	0.083	0.079
Year when programme started (Ref. group: 1994):				
1993	0.072	0.011	0.070	0.084
1995	0.035	0.048	0.036	0.027
1996	0.012	0.049	0.014	-0.007
1997	-0.038	0.012	-0.036	-0.061
1998	0.038	0.098	0.042	0.013
Number of residents in municipality (Ref. group: 0-20,000):				
20,000-40,000	0.014	-0.001	0.014	0.021
40,000-100,000	0.003	-0.024	0.002	0.002
More than 100,000	0.048	0.040	0.047	0.063
Relative unemployment level	-0.208	-0.168	-0.202	-0.199
Relative importance of PSE programmes				0.054

Table 9

Estimates of the parameters determining selection into PSE programmes when the common factor follows a discrete distribution with 3 points of support and there is selection on unobservables ($\alpha_1 \neq 0$, $\alpha_1 \neq 0$ and $\alpha_1 \neq \alpha_0$)

	Coeff.	Std.[a]	Marg.[b]
Constant	**-1.276**	0.310	-0.200
Age (ref. group: 17-24):			
25-29	-0.102	0.113	-0.016
30-39	**-0.551**	0.129	-0.086
40-49	**-0.908**	0.160	-0.142
50-66	**-0.675**	0.218	-0.106
Marital status (Ref. group: Single)			
Married	0.211	0.154	0.033
Cohabitating	**0.277**	0.096	0.043
Has children	0.213	0.123	0.033
Completed education (Ref. group: Primary or lower sec. school):			
Upper secondary school	0.164	0.096	0.026
Vocational education	0.132	0.089	0.021
Further or higher education	-0.273	0.222	-0.043
Work experience (Ref. group: 0-2 years):			
2-5 years	**0.618**	0.098	0.097
5-10 years	**0.811**	0.131	0.127
10+ years	**0.971**	0.157	0.152
Status during the 12 months preceding the programme period (Ref. group: Employment):			
Unemployment	-0.124	0.124	-0.019
Ordinary education	-0.019	0.136	-0.003
Status during 24 months period starting 3 years (and ending 1 year) before the programme period (Ref. group: Employment):			
Unemployment	**-0.368**	0.135	-0.058
Ordinary education	-0.167	0.174	-0.026
No available information	-0.035	0.118	-0.005
Year when programme started (Ref. group: 1994):			
1993	**0.631**	0.086	0.099
1995	**-0.230**	0.084	-0.036
1996	**-0.654**	0.113	-0.103
1997	**-0.874**	0.132	-0.137
1998	**-1.029**	0.169	-0.161
Number of residents in municipality (Ref. group: 0-20,000):			
20,000-40,000	**0.255**	0.094	0.040
40,000-100,000	**0.248**	0.084	0.039
More than 100,000	**0.221**	0.090	0.035
Relative unemployment level	-0.325	0.268	-0.051
Relative importance of PSE programmes	**0.537**	0.046	0.084
Common unobserved factor	1.000 [c]		
v_1	**-0.729**	0.075	
v_2	0.000 [c]		
v_3	**2.504**	0.209	
$p1 = Pr(\theta = v_1)$	**0.424**	0.044	
$p2 = Pr(\theta = v_2)$	**0.452**	0.049	
$p3 = Pr(\theta = v_3)$	**0.124**	0.017	
Log-likelihood	-7119.1		

Note.　Figures in bold indicate significance at a 5 per cent level.
a.　Asymptotic standard errors.
b.　Mean marginal effect on the probability of participating in a PSE programme.
c.　The parameter is fixed at the given value.

Table 10

Estimates of the parameters determining the employment status 12 months after the end of the programme period when the common factor follows a discrete distribution with 3 points of support and there is selection on unobservables ($\alpha_1 \neq 0$, $\alpha_1 \neq 0$ and $\alpha_1 \neq \alpha_0$)

	Participants in PSE programmes			Participants in non-PSE programmes		
	Coeff.	Std.[a]	Marg.[b]	Coeff.	Std.[a]	Marg.[b]
Constant	**1.382**	0.404	0.511	**2.939**	0.685	0.427
Age (ref. group: 17-24):						
25-29	0.037	0.135	0.014	-0.080	0.142	-0.012
30-39	-0.243	0.156	-0.090	**-0.841**	0.193	-0.122
40-49	**-0.691**	0.198	-0.255	**-1.215**	0.263	-0.176
50-66	**-0.683**	0.258	-0.252	**-1.420**	0.331	-0.206
Marital status (Ref. group: Single)						
Married	0.022	0.184	0.008	**0.936**	0.244	0.136
Cohabitating	0.012	0.117	0.004	**0.447**	0.150	0.065
Has children	0.118	0.149	0.043	0.036	0.160	0.005
Completed education (Ref. group: Primary or lower sec. school):						
Upper secondary school	0.021	0.112	0.008	0.286	0.148	0.042
Vocational education	**0.251**	0.106	0.093	**0.753**	0.165	0.109
Further or higher education	0.394	0.271	0.145	**0.463**	0.233	0.067
Work experience (Ref. group: 0-2 years):						
2-5 years	0.170	0.131	0.063	**0.405**	0.151	0.059
5-10 years	**0.388**	0.171	0.143	**0.528**	0.184	0.077
10+ years	**0.522**	0.201	0.193	**0.780**	0.230	0.113
Status during the 12 months preceding the programme period (Ref. group: Employment):						
Unemployment	**-0.314**	0.143	-0.116	**-1.015**	0.231	-0.147
Ordinary education	-0.263	0.155	-0.097	0.115	0.200	0.017
Status during 24 months period starting 3 years (and ending 1 year) before the programme period (Ref. group: Employment):						
Unemployment	**-0.491**	0.156	-0.181	**-1.261**	0.254	-0.183
Ordinary education	-0.108	0.205	-0.040	-0.315	0.250	-0.046
No available information	0.061	0.134	0.022	**-0.438**	0.172	-0.064
Year when programme started (Ref. group: 1994):						
1993	0.163	0.114	0.060	0.236	0.173	0.034
1995	0.069	0.101	0.025	-0.138	0.110	-0.020
1996	0.048	0.142	0.018	-0.174	0.131	-0.025
1997	-0.093	0.166	-0.034	**-0.314**	0.152	-0.046
1998	-0.037	0.225	-0.014	**-0.551**	0.204	-0.080
Number of residents in municipality (Ref. group: 0-20,000):						
20,000-40,000	-0.004	0.110	-0.001	0.031	0.122	0.004
40,000-100,000	-0.062	0.105	-0.023	0.062	0.112	0.009
More than 100,000	0.134	0.103	0.050	0.082	0.109	0.012
Relative unemployment level	**-1.046**	0.332	-0.387	**-1.353**	0.372	-0.197
Relative importance of PSE programmes						
Common unobserved factor	-0.065	0.129		**3.905**	0.732	

Note. Figures in bold indicate significance at a 5 per cent level.

a. Asymptotic standard errors.

b. Mean marginal effect on the probability of being employed.

Table 11

The average treatment effect (Δ^{ATE}) and associated distributional treatment parameters in a model where the common factor follows a discrete distribution with 3 points of support and there is selection on unobservables ($\alpha_1 \neq 0$, $\alpha_1 \neq 0$ and $\alpha_1 \neq \alpha_0$)

$P_{Y_1,Y_0}^{ATE}(1,0)$	$P_{Y_1,Y_0}^{ATE}(0,1)$	$P_{Y_1,Y_0}^{ATE}(1,1)$	$P_{Y_1,Y_0}^{ATE}(0,0)$	$P_{\Delta}^{ATE}(0)$	Δ^{ATE}
0.281	**0.199**	**0.245**	**0.275**	**0.520**	0.082
(0.054)	(0.016)	(0.016)	(0.054)	(0.039)	(0.069)

Note.   Figures in bold indicate significance at a 5 per cent level. Standard errors are given in parentheses.

Table 12

The effect of treatment on the treated (Δ^{TT}) and associated distributional treatment parameters in a model where the common factor follows a discrete distribution with 3 points of support and there is selection on unobservables ($\alpha_1 \neq 0$, $\alpha_1 \neq 0$ and $\alpha_1 \neq \alpha_0$)

$P_{Y_1,Y_0}^{TT}(1,0)$	$P_{Y_1,Y_0}^{TT}(0,1)$	$P_{Y_1,Y_0}^{TT}(1,1)$	$P_{Y_1,Y_0}^{TT}(0,0)$	$P_{\Delta}^{TT}(0)$	Δ^{TT}
0.107	**0.382**	**0.425**	**0.086**	**0.512**	**-0.275**
(0.019)	(0.021)	(0.021)	(0.018)	(0.029)	(0.027)

Note.   Figures in bold indicate significance at a 5 per cent level. Standard errors are given in parentheses.

Table 13

The marginal treatment effect (Δ^{MTE}) and associated distributional treatment parameters (for selected values of U_D) in a model where the common factor follows a discrete distribution with 3 points of support and there is selection on unobservables ($\alpha_1 \neq 0$, $\alpha_1 \neq 0$ and $\alpha_1 \neq \alpha_0$)

	$P_{Y_1,Y_0}^{MTE}(1,0)$	$P_{Y_1,Y_0}^{MTE}(0,1)$	$P_{Y_1,Y_0}^{MTE}(1,1)$	$P_{Y_1,Y_0}^{MTE}(0,0)$	$P_{\Delta}^{MTE}(0)$	Δ^{MTE}
$U_D=-2$	**0.420**	**0.080**	**0.119**	**0.381**	**0.500**	**0.341**
	(0.082)	(0.015)	(0.016)	(0.081)	(0.068)	(0.096)
$U_D=0$	**0.294**	**0.172**	**0.239**	**0.296**	**0.534**	0.113
	(0.056)	(0.026)	(0.027)	(0.056)	(0.030)	(0.082)
$U_D=2$	**0.077**	**0.418**	**0.413**	**0.091**	**0.505**	**-0.341**
	(0.022)	(0.029)	(0.025)	(0.023)	(0.033)	(0.039)

Note.   Figures in bold indicate significance at a 5 per cent level. Standard errors are given in parentheses.