IFAU – INSTITUTET FÖR
ARBETSMARKNADSPOLITISK
UTVÄRDERING

Policy programmes only for a few? Participation in labour market programmes among Swedish disabled workers

Maria Melkersson

Working Paper 1999:1

Policy programmes only for a few? Participation in labour market programmes among Swedish disabled workers[1]

Maria Melkersson[2]

1st July 1999

[1]The financial support from the Expert Group on Labour Market Policy Evaluation Studies (the predecessor of IFAU), Ministry of Labour, Stockholm and from *The Jan Wallander and Tom Hedelius Foundation* is gratefully acknowledged. Seminar participants at IFAU in Uppsala, at SOFI and at the Department of Economics, University of Umeå, are thanked for many helpful comments on earlier versions of the paper. The usual disclaimer applies.

[2]email: Maria.Melkersson@sofi.su.se; address: SOFI, University of Stockholm, SE 106 91 Stockholm; phone: +46-8-16 38 62; fax: +46-8-15 46 70.

Abstract

In a large sample of Swedish unemployed disabled workers, the workers participate in between zero and ten policy programmes during their unemployment spell. Clustering of programmes to about half of the sample is prominent.

The number of programmes is modelled as a standard count data model, as a zero-inflated model and as a hurdle model. The most important question is whether disadvantaged workers are more or less probable to participate in programmes. The empirical analysis shows that participants in programmes have a stronger labour market attachment, than do non-participants.

The number of programmes differ across subgroups of the disabled, and workers with impaired hearing or vision and psychical disabilities are expected to participate in more programmes than other groups. The supply of policy programmes is also of importance. In particular, for individuals living in municipalities with many disabled unemployed, the expected number of programmes is lower.

Descriptive inspection of how the unemployment spells end, shows that non-participants, i.e. workers with zero programmes, withdraw from the labour force to a larger extent than do programme participants.

Key Words: Policy programmes, Unemployment, Hurdle models

JEL: C25, J13

1 Introduction

One important part of Swedish labour market policy is the offering of active policy programmes to unemployed workers. There are specific programmes directed to specific groups of unemployed, e.g. youth, immigrants and disabled workers. At the same time, groups with a weak position on the regular labour market may be given priority in programmes directed to all unemployed. Among the priority groups for policy programmes are the disabled workers, a group considered to have a particularly weak position on the labour market. For a large sample of disabled workers entering open unemployment in 1992, the *number* of policy programmes that the workers participate in, show a large variation. The unemployed[1] disabled workers participate in between zero and ten policy programmes during their unemployment spell. As many as about half of the sample do not participate in any programmes at all.

Training and other programmes offered to unemployed workers in Sweden, are paid by the state and administered by the local employment offices. In previous studies on selection into these policy programmes and on the post-programme effects on e.g. earnings, the "treatment" has been considered as a binary variable. This means that it is assumed that the unemployed chose between "training" and "not training", or between "relief work" and "not relief work" (see e.g. Ackum Agell, 1995, Brännäs and Eriksson, 1996). We know that in practice the selection process into policy programmes is much more complicated. There are about ten different policy programmes offered to unemployed in Sweden. Some of the programmes may serve as substitutes to each other while others may be complements. For example, when an individual starts in a policy programme it may already be planned (by the participant together with his official at the employment office) that he/she is then to participate in a second programme, as a continuation of the first.

The facts mentioned above make it extremely difficult to model selection into all the programmes available, and even more so to estimate the post-programme effects on unemployment durations or on earnings. As a simple and realistic example, consider three individuals who spend some time in unemployment. Before leaving

[1] In this paper "unemployed" means "registered at a local employment office". An unemployed worker may be openly unemployed, or participate in a policy programme administered by the employment offices.

unemployment person A participates in training, person B participates in relief work while person C participates in training, then in relief work and then leaves unemployment. How should these individuals be compared to each other? Do all three individuals belong to the same "treatment group"? How should the control group(s) be formed? Already when considering only two programmes (which to some unemployed workers may serve as substitutes and to others as complements) the evaluation problem is extremely complicated.

To my knowledge there are no studies on selection into policy programmes and the subsequent programme effects, that take into account the complex reality of Swedish active labour market policy.[2] This paper is an attempt to take a step towards broadening the understanding of the selection process into policy programmes. The main question to be discussed and modelled in this paper is; how can we explain that about half of the unemployed do not participate in any policy programmes at all? Is there some kind of "barrier" that an unemployed must cross, in order to participate in a first programme, and then possibly in subsequent programmes? A related question is if policy programmes to a large extent are given in order to avoid the expiration of unemployment benefits. If this is the case, workers who are not entitled to benefits are less likely to be offered programmes. Workers entitled to benefits have previous labour market attachment in the form of experience from the regular labour market. This means that eligible workers are probably in many respects better off than workers without this attachment. If workers entitled to benefits are favoured for programmes, workers without previous attachment will have a smaller probability of getting into the programmes, although they may be the most in need.

Arulampalam and Booth (1997, henceforth A&B) model the number of work-related training courses during a ten-year period, for young British workers. They use a hurdle specification for the number of training courses, which means that the decision to participate in training is assumed to be a process with two stages. The first stage is the decision whether to participate in any training at all (whether to cross the hurdle or not) and in the second stage, the number of training courses is decided (once the hurdle is crossed). The first decision, to cross the hurdle and participate in a first course, may be generated by a different underlying process than

[2] For a study on choice set size of non-disabled Swedish workers, see Melkersson (1997).

the decision on how many courses to participate in.

A&B (1997) state that a majority of young British workers never seem to cross the "training hurdle" and therefore training appears to be mainly for the already trained. Those who do not get into any work related training may be caught in a "low skill, bad job trap". According to the authors, one explanation of this "bunching" of training to a minority of the workers is that the already trained are more easily trained, which is important for work-related training paid by the employer. In addition, A&B test the hypothesis that work-related training is given mainly to workers with past general human capital accumulation, and they find support for this in their data.

My analogue to the work-related training courses analyzed in A&B (1997), is the diversity of policy programmes offered to unemployed workers. For the period in which our sample has been drawn, there were ten different policy programmes offered to unemployed workers, including e.g. labour market training, vocational rehabilitation and relief work. I will use the same kind of model as used by A&B (1997); a count data model. I will, in particular, test the hypothesis that policy programmes to disabled workers are given to the better off workers. That is, test whether programmes are given mainly to workers with education and/or extensive previous labour market experience, i.e. individuals who already have a relatively strong attachment to the labour market.

If there is a tendency to "treat the already treated" among the disabled workers, we may conclude that it is somehow "easier" for placement officers to give policy programmes to former participants and better off workers. One explanation to why this "easy way" may be chosen by the officers, could be that former participants may be easier to administrate than new participants. New participants may also need encouragement, information and support, while former participants are already familiar with policy programmes. Another explanation could be that if programme participation is evaluated, a better result may be achieved if the better off individuals are favoured for participation. The reasons for and possible effects of this so called "cream-skimming", which may be defined as "systematically selecting participants who would have higher employment rates even without the training programme" (Aakvik et al., 1998, p.1), is discussed in e.g. Anderson et al.(1993) and Heckman et al. (1996). One final explanation to the observed bunching of programmes to

about half of the sample, could be that unemployed with past experience of policy programmes may themselves be more active in getting into more programmes.

The paper is organised as follows. The issue of policy programme participation and its relation to unemployment duration is discussed in section 2. Data to be analyzed is presented in section 3, while the relevant econometric issues are discussed in section 4. Empirical results follow in section 5 while conclusions are left for section 6.

2 Unemployment and policy programmes

One important tool available in Swedish active labour market policy is a diversity of policy programmes offered to unemployed workers. The ability of earning income is dependent on the prospect of getting a job. Taking into account the possible stigma caused by unemployment (see Heckman and Borjas, 1980), it may be important to find a job rather quickly, and by this escape unemployment. In order to improve the opportunities of unemployed workers, active labour market programmes are offered by the local employment offices. Participation in training and other policy programmes are supposed to have a positive effect of future labour market prospects; this is actually the main motivation for such programmes (see Edin and Holmlund, 1991).

For one particular unemployed worker himself, it may be difficult to take in the whole situation, to know about all programmes offered to unemployed and foresee their consequences. The role of the placements officers at the local employment offices is guidance and consulting to unemployed workers. Therefore it is reasonable to assume that the worker, with the help of an placement officer, maximises his utility. This role of the officers is their "paternalistic" role discussed in Heckman et al. (1996, p.6); probably a very important role and maybe even more so for disabled unemployed workers.

Participation in all labour market programmes in Sweden, is administered by placement officers at the employment offices. Participation in one particular programme requires self-selection (i.e. the unemployed must be willing to participate) and also officer selection. In officer selection we also include the stipulated

rules which regulate participation in different policy programmes. The officer also has this "paternalistic" role mentioned above, which includes informing the unemployed worker about different programmes, to encourage the worker to apply for programmes, etc. In this role an officer may, of course, be more or less active. In the empirical analysis we will not be able to separate between self-selection and officer selection.

When we model policy programme participation among unemployed workers, we have to take a supply constraint into consideration; unemployed workers may not get more programmes than the "supply stock" available at the local employment office. Restrictions may be financial, since the local employment office must keep within the budget. The restrictions for one particular worker may also be due to regulations. Not everyone has theoretical access to all programmes offered by the employment offices. For example, in order to get a work experience scheme (one of the largest programmes, meaning subsidised employment for a limited period of time), the worker must be entitled to unemployment insurance benefits and have at least one day of benefits left. The local "supply stock" of programmes will surely depend on the local labour market situation. In the empirical estimations a few variables describing the labour market at the municipal level will be included to reflect local conditions and thus the "competition" for programmes among the disabled.

A large proportion of the sample to be used in the empirical application, have not participated in any programme at all. Pudney (1989, Ch.4) mentions three possible explanations for these so called corner solutions. The first is the length of the observation period, i.e. the unemployment spell is too short to include a programme. This will be taken into consideration in the econometric model, where it will be assumed that the expected number of programmes is a function of unemployment duration. The second explanation to a corner solution is involuntary zero programmes, i.e. the individual lacked the opportunity to participate in any programme, for example due to that demand for e.g. training was larger than the number of places supplied. The third explanation is the outcome of free choice, i.e. the individual chooses not to participate in any programme, e.g. if the worker has a high expectancy of getting a new job relatively quickly. In the empirical analysis I cannot separate between involuntary and voluntary zero programmes, and therefore

the two will be treated alike.

In the study by A&B (1997) it is obvious that zero training programmes is detrimental to the worker, in the sense that he may be trapped in a low skill position for ever. In our example of programmes for disabled unemployed workers, it is not obvious that zero programmes should be considered as prejudicial in the similar sense. It may well be that individuals with zero programmes manage very well, and can find their way out of unemployment without the help offered by the employment offices. If this is the case, the individuals who have not crossed the programme hurdle are the ones with the most favourable position on the labour market. Whether workers with zero programmes are more or less fortunate than others on the labour market, will not be answered in this paper.

3 Data

A large majority (about 90 percent) of unemployed Swedish workers, are registered at a local employment office. These offices provide information and personal counselling by employment officers. In meetings with unemployed workers for subsidised the officers register events such as participation in policy programmes, and the resulting database is hosted by the National Labour Market Board. The database contains individual background variables such as education and professional experience, together with individual labour market histories. This means that we know what policy programmes the unemployed have participated in during their unemployment spell, and we also know the reason for the end of the spell (regular employment, retirement, etc.). The sample consists of 32,609 disabled workers aged 25-55 years, who all registered as unemployed at a local employment office during 1992.[3]

During the unemployment spell, until it ended or was censored after about four years (which is the length our sampling period) the unemployed may have participated in policy programmes. In Table I the sample distribution of the number of programmes is summarised together with a theoretical Poisson distribution with the same mean. The observed mean is 1.1 programmes, with observed standard de-

[3] This is the entire inflow to unemployment of disabled workers in 1992, except for some observations lacking important pieces of information. The excluded observations are assumed to be random, since lack of some information depends more on the routines of the individual official, than on the unemployed individual himself.

viation equal to 1.4, i.e. the unconditional data is overdispersed since the variance exceeds the mean. For individuals with at least one programme the mean is 2.0 programmes with standard deviation 1.3. We see from Table I that nearly half of the sample do not participate in any programmes at all.

Table I: Observed numbers and shares with different numbers of policy programmes, together with a theoretical Poisson distribution with the same mean of 1.1 programmes. Both estimation and holdout samples are added together.

#	Number	Share	Poisson
0	14,918	0.457	0.342
1	8,733	0.268	0.367
2	4,312	0.132	0.197
3	2,405	0.074	0.070
4	1,313	0.040	0.019
5	605	0.019	0.004
6	209	0.006	0.001
7	84	0.003	0.000
8	17	0.001	0.000
9	6	0.000	0.000
10	7	0.000	0.000
n	32,609	1.000	1.000

54 percent of our sample participate in at least one programme and out of these, 30 percent have at least one "repetition", meaning participation in the same kind of programme twice or more. The number of programmes to chose between for the unemployed is ten and within the sample the maximum number of different unique programmes is six.[4] One contributing factor to the "clustering" of the data to a share of the sample is the way participation in programmes is handled at the employment offices. It is not uncommon that the programmes directed to an unemployed worker give rise to a whole "chain" of programmes, for example, first rehabilitation, then training and after that a traineeship. This means that routines used in active labour market policy can explain some of the clustering of the data.

In section 4, I will present a few econometric specifications of the probabilities for participating in $m = 0, 1, 2, ...$ policy programmes. In the application I will take

[4] The phenomenon of "recycling" in labour market programmes is discussed and modelled in Brännäs (1996), in particular with focus on measurement error in exposure times, i.e. in unemployment durations.

into account that there are differences in unemployment duration, i.e. the workers differ in their "exposure to risk" of any programmes. For example, if a worker is between jobs for a very short period of time, the probability that he will participate in a policy programme will be low. In the sample the exposure time varies between one and 1,492 days with a mean of 483 days, which correspond to about 16 months. In the models used, the expected number of programmes is assumed to depend on time spent in unemployment. Duration in unemployment is assumed to end if the worker is deactivated from the register of unemployed, due to regular employment, subsidized/sheltered employment or withdrawal from the labour force (retirement, university studies, etc.).

The individual characteristics which will be used in the empirical analysis are counselling by descriptive statistics in Table II. All variables are measured at the onset of the unemployment spell (i.e. in 1992). In the empirical model I will also include three variables at the municipal level, in order to account for differences in local labour market conditions. The variables are *unemployment rate* (openly unemployed), "*disability rate*" (openly unemployed with a disability) and "*programme rate*" (total number of unemployed who enter policy programmes). The *unemployment rate* is the number of unemployed in relation to the local population aged 16-64, while the other rates are calculated in relation to unemployment. All figures are means across months in 1992 and expressed in percentages.

In regular job search theory it is assumed that unemployment rates both affect the probability of a job offer to an unemployed, and the probability that such an offer is accepted (e.g. Edin, 1991, Edin and Holmlund, 1991). For the issue of this paper, programme participation among disabled unemployed, the effect of the local *unemployment rate* may be either positive or negative; positive if policy programmes are used as substitutes for regular employment, but negative if high unemployment rate means a more stretched economy of the local employment office, and thereby less programmes per unemployed offered. The *disability rate* is expected to have a negative effect on programme participation, since more disabled unemployed means stronger competition for a limited local budget. Finally, the *programme rate* will be expected to have a positive impact on programme participation among disabled. If the local employment office tend to offer many policy programmes to their unemployed, the programme availability will probably "spill over" to the disabled workers

Table II: Descriptive statistics; means and standard deviations in parenthesis, for estimation sample. All characteristics refer to onset of unemployment spells. Sample size is 16,284.

Variable	Mean
Number of programmes	1.072
	(1.359)
Age	38.282
	(8.389)
Male (=1)	0.588
Foreign citizen (=1)	0.063
High school (=1)	0.360
University (=1)	0.044
Stockholm county (=1)	0.133
Vocational education (=1)	0.298
Some experience (=1)	0.223
Long experience (=1)	0.340
UI benefits (=1)	0.502
Cash assistance (=1)	0.051
Impaired hearing/vision (=1)	0.041
Disability (=1)	0.420
Other physical disability (=1) (ref.)	0.207
Psychical/intellectual disab. (=1)	0.118
Social medical disability (=1)	0.214
Unemployment rate, percent	4.813
	(1.071)
"Disability rate", percent	21.092
	(7.476)
"programme rate", percent	10.680
	(3.811)
Unemployment duration, days	482.836
	(426.001)

as well.

4 Econometric models

The number of policy programmes that unemployed worker i participates in, Y_i, is a count variable, i.e. it takes only non-negative integers. The expected number of programmes is assumed to depend on a vector with individual characteristics, $\mathbf{x}_i$, and on the time spent as unemployed, t_i;

$$E(Y_i|\mathbf{x}_i, t_i) = \lambda_i = \exp(\mathbf{x}_i\boldsymbol{\beta} + \delta \ln t_i), \tag{1}$$

where $\mathbf{x}_i$ and t_i are assumed to be fixed and $\boldsymbol{\beta}$ is an unknown parameter vector. Including unemployment as a regressor may be referred to as a "logarithmic offset".[5] As a point of reference the Poisson model will be estimated, for which the density is

$$f(y_i|\mathbf{x}_i, t_i) = \exp(-\lambda_i)\lambda_i^{y_i}/y_i!$$

and for which $E(Y_i|\mathbf{x}_i, t_i) = V(Y_i|\mathbf{x}_i, t_i) = \lambda_i$, the so called "equidispersion property", holds. The likelihood function to be maximised is

$$\ell = \prod_{i\in\Omega} f(y_i|\mathbf{x}_i, t_i).$$

where Ω is the sample to be analyzed.

In section 2 above a so called hurdle specification was discussed. If we assume that programme participation is generated by this kind of two stage decision process, we get the likelihood function

$$\begin{aligned} \ell &= \prod_{i\in\Omega_0} f_1(0|\mathbf{x}_i, t_i) \times \prod_{i\in\Omega_1} (1 - f_1(0|\mathbf{x}_i, t_i)) \frac{f_2(y_i|\mathbf{x}_i, t_i)}{(1 - f_2(0|\mathbf{x}_i, t_i))} = \\ &= \prod_{i\in\Omega} f_1(0|\mathbf{x}_i, t_i)^{1-d_i} (1 - f_1(0|\mathbf{x}_i, t_i))^{d_i} \times \prod_{i\in\Omega_1} \frac{f_2(y_i|\mathbf{x}_i, t_i)}{(1 - f_2(0|\mathbf{x}_i, t_i))} \end{aligned} \tag{2}$$

where $d_i = 0$ if $y_i = 0$ and $d_i = 1$ if $y_i > 0$. Above, Ω refers to the entire sample, Ω_0 refers to the subsample with $y_i = 0$ and Ω_1 to the subsample with $y_i > 0$. The first

[5] Assuming proportionality to unemployment duration by defining the mean function as $E(Y_i|\mathbf{x}_i, t_i) = t_i \exp(\mathbf{x}_i\boldsymbol{\beta})$, is a possible option for the standard count data model. However, for the hurdle model to be presented later, there is no easy way to achieve proportionality by a parameterisation similar to above. Therefore, duration is treated as a regressor in the analysis, with a parameter which is free to vary.

part of the likelihood refers to the decision of crossing the hurdle or not, while the second part is a truncated-at-zero count data model for the positive counts.

The distributions of the first and second part, $f_1(\cdot)$ and $f_2(\cdot)$, may be the same or differ. If they are of the same kind and with identical parameter vectors, the likelihood collapses to that of a standard non-hurdle model. By construction, independence between the two parts is assumed, and the two parts of the hurdle model may be estimated separately. Assuming a Poisson distribution for both parts of the likelihood in eq.(2) leads to the Poisson-hurdle model, with the likelihood function

$$\ell = \prod_{i\in\Omega} \exp(-\lambda_{1i})^{1-d_i}(1-\exp(-\lambda_{1i}))^{d_i} \times \prod_{i\in\Omega_1} \frac{\exp(-\lambda_{2i})\lambda_{2i}^{y_i}}{(1-\exp(-\lambda_{2i}))y_i!}. \tag{3}$$

For the hurdle models the parameterisation of the means are as before;

$$\lambda_{ij} = \exp(\mathbf{x}_i\boldsymbol{\beta}_j + \delta_j \ln t_i), \text{ for } j = 1, 2.$$

That is, the same explanatory variables are allowed to affect both decisions, but the parameter vectors may differ. The null hypothesis of $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$ will be tested for the hurdle model. If the null can not be rejected, the hurdle specification reduces to the standard non-hurdle model. Using a likelihood ratio test we calculate

$$T = -2\left[L_0 - (L_1 + L_2)\right], \tag{4}$$

where L_0 is the value of the restricted loglikelihood function, in which it is assumed that $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$, while L_1 and L_2 are the values of the loglikelihood functions for the first and second steps in the hurdle model, when estimated separately. Under the null, the test statistic is χ^2-distributed with k_0 degrees of freedom, where k_0 is the number of restrictions imposed by the non-hurdle model. In the case the null is rejected, the hurdle specification is considered to be superior to the non-hurdle model. Incorrectly neglecting the hurdle specification leads to an incorrect mean as well as an incorrect variance (see e.g. Winkelmann and Zimmermann, 1995).

In some economic applications the zero outcomes may be of particular interest, e.g. in modelling consumer behaviour where the zero outcomes are suspected to be different in kind from positive observations. Handling zeros as a particular outcome of a first stage in a hurdle model, is one way of treating the zeros. Another approach is to allow for two different "kinds" of zeros. For this latter approach I will use the

zero-inflated specification proposed by e.g. Mullahy (1986). Assume the following density

$$\begin{aligned} f(y_i|\mathbf{x}_i, t_i) &= \psi + (1-\psi)h(0|\boldsymbol{\delta}, \mathbf{x}_i, t_i) = \psi + qh(0|\boldsymbol{\delta}, \mathbf{x}_i, t_i), \quad y_i = 0 \\ f(y_i|\mathbf{x}_i, t_i) &= (1-\psi)h(y_i|\boldsymbol{\delta}, \mathbf{x}_i, t_i) = qh(y_i|\boldsymbol{\delta}, \mathbf{x}_i, t_i), \quad y_i \geq 1 \end{aligned} \tag{5}$$

where $h(y_i|\boldsymbol{\delta}, \mathbf{x}_i, t_i)$ is the "parent" distribution with the unknown parameter vector $\boldsymbol{\delta}$. The extra term ψ is allowed to be both positive and negative, i.e. it is not necessarily a probability which is often assumed for the zero-altered or zero-inflated models (see Greene, 1994, and Lambert, 1992). The sign of ψ may be interpreted directly in terms of departure from the equidispersion property; when $\psi < 0$ there is underdispersion in the data and when $\psi > 0$ overdispersion (cf. Mullahy, 1986, p.355). In eq.(5) the parameter ψ is constant and equal across individuals. The parameter may also be modelled as an individual specific probability, as proposed by e.g. Lambert (1992). The expected value of a zero–inflated model, with distribution as in eq.(5), is

$$E(Y_i|\mathbf{x}_i, t_i) = q\lambda_i. \tag{6}$$

The interpretation of a hurdle model is different from that of a zero-inflated model. The zeros in a setting like eq.(5) may emerge from two sources; from the parent distribution or independently of it, or as put in Greene (1994, p.9) "In one regime, the zero value is automatic, but in the other, it is but one possible outcome". In the hurdle specification, on the other hand, all zeros are lumped together. As an example, unemployed who have *not yet* entered a programme and those who would *never* enter a programme are added together, and treated as the same kind of zeros in a hurdle model. In a zero-inflated model we allow the two kinds of zeros to differ. The zeros according to the parent distribution are, for example workers who might participate in training later on (as t_i increases) or who would participate under other circumstances, e.g. if the local unemployment rate was higher. Zeros outside the parent distribution are workers who would never participate in any programmes, regardless of time spent in unemployment and other factors. Examples of this latter kind of "automatic non-participants" are workers who are waiting passively for retirement, and workers who are quite sure of being recalled to a former employer.[6]

[6] According to a Swedish study by Harkman and Jansson (1995), in a sample of non-disabled

Dividing the sample Ω into those who have participated in zero programmes, Ω_0, and those with one or more programmes, Ω_1, respectively, the likelihood can be written

$$\ell = \prod_{\Omega_0} (\psi + qh(0|\boldsymbol{\delta}, \mathbf{x}_i, t_i)) \prod_{\Omega_1} qh(y_i|\boldsymbol{\delta}, \mathbf{x}_i, t_i) \tag{7}$$

The likelihood is then modelling under some distributional assumption for the parent distribution $h(\cdot)$. In the empirical application the Poisson distribution will be assumed.

I will model the number of policy programmes by three different specifications; as a standard Poisson count data model, a zero-inflated Poisson model and a Poisson hurdle model. For all models estimated, I will present the value of the log likelihood function, ℓ, and the consistent Akaike information criterion, CAIC, (e.g. Gurmu and Trivedi, 1996)

$$CAIC = -2\ell + k(\ln n + 1)$$

where k denotes the number of estimated parameters, n is the sample size and the minimum CAIC indicates the "best" model.

A number of models will be estimated and their ability to mimic the sample distribution will be compared. In order not to favour more complicated models with more parameters (e.g. Gurmu and Trivedi, 1996) I will split the sample randomly into an estimation sample and a holdout sample. Estimations will be made using the estimation sample (with $16,284$ observations) and then the sample distribution will be predicted for the holdout sample (with $16,325$ observations). For each individual in the holdout sample I will then estimate the probabilities of $m = 0, 1, 2, ...$ policy programmes, conditional on their individual characteristics, $\mathbf{x}_i$, and unemployment duration, t_i. The probabilities of the regular Poisson model are straight forward and the probabilities of the zero-inflated model are given by eq.(5). For the hurdle model the probability of a zero count is $f_1(0|\mathbf{x}_i, t_i)$, while the probabilities of positive counts are given as (see e.g. Winkelmann and Zimmerman, 1995)

$$f(y_i|\mathbf{x}_i, t_i) = (1 - f_1(0|\mathbf{x}_i, t_i)) \frac{f_2(y_i|\mathbf{x}_i, t_i)}{(1 - f_2(0|\mathbf{x}_i, t_i))}, \; y_i = 1, 2, ... \tag{8}$$

unemployed, as many as 45 percent of a who found employment in 1994, were re-employed by a former employer. The authors assume that the recalls is one important explanation to low search intensity among Swedish unemployed, compared to unemployed in other European countries.

which are to be calculated at ML estimates. The estimated probabilities, summed over the sample, are then used as the estimated sample distribution for the holdout sample. I will use a χ^2-test for goodness-of-fit which is based on differences between observed and estimated frequencies with $m = 0, 1, 2, ...$ programmes;

$$X_h^2 = \sum_i \frac{(n_m - \hat{n}_m)^2}{\hat{n}_m} \tag{9}$$

which is χ^2-distributed with $m - 1$ degrees of freedom, where m is the number of cells in the frequency table that the statistic is based on (e.g. Winkelmann and Zimmermann, 1995, p.16).[7]

5 Empirical results

The estimated sample distributions, the value of the log likelihood functions and the CAIC are presented in Table III. According to the test in eq.(4), the null of $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$ is rejected at any conventional risk level. This means that a hurdle specification is superior to a standard Poisson data model. If we use the minimum CAIC as the criterion the hurdle model is also the best, while the ZIP model better predicts the sample distribution of the holdout sample.

In Table IV the estimated parameters of all models are presented. The parameter ψ of the ZIP-model is significantly less than zero, which means that the data are conditionally underdispersed; the conditional variance falls short of the conditional mean. This means that there are *fewer* zeros than expected from a conventional Poisson model, when we condition on unemployment duration and other individual characteristics. The ψ-parameter is significantly different from zero but small in size; about -0.03. The parameters of the standard Poisson model and the ZIP model are apparently very close in size and we leave the ZIP model without further parameterisation of ψ.

If we compare the standard Poisson model and the hurdle model we find that almost all variables with a significant effect in the standard model, are also significant and of the same sign for both stages of the hurdle model. However, there are a couple of exceptions. First, according to the standard Poisson model workers in Stockholm

[7] All models have been estimated using the GAUSS package, with the BHHH estimator of the covariance matrices (e.g. Greene, 1997, Ch.4).

Table III: Observed and estimated sample distributions for the hold-out sample, with n=16,325.

Counts	Obs.	Poisson	ZIP	Hurdle
0	0.456	0.472	0.465	0.538
1	0.268	0.249	0.256	0.216
2	0.132	0.134	0.137	0.112
3	0.072	0.074	0.074	0.064
4	0.042	0.039	0.038	0.036
5	0.020	0.019	0.018	0.019
6	0.007	0.008	0.008	0.009
7-10	0.003	0.005	0.004	0.006
Total	1.000	1.000	1.000	1.000
X^2		55	31	90
$-\ell$		18,013	17,987	17,854
k		19	20	38
$CAIC$		36,229	36,188	36,114

Note: For the hurdle model, the information in the table refer to the sum of both stages.

participate in fewer programmes than others. When the model is divided into the two-stage hurdle model we see that this is mainly a lower participation rate in Stockholm, i.e. fewer "cross the hurdle", but among those who do participate in at least one programme, the difference no longer prevails. Second, the observed higher expected number of programmes for workers with impaired hearing or vision and psychical disabilities as opposed to other disabilities, is explained by a larger number of programmes for individuals who cross the "programme hurdle".

There is no significant difference between men and women, but the number of programmes decreases with age. It is clear that individuals entitled to any kind of benefits at all, have a larger expected number of policy programmes. This may be due to favouring by the officials, to help workers avoid expiration of benefits. It may also be that it is easier to find suitable programmes for individuals with labour market attachment (which is a prerequisite to become eligible for benefits). According to the hurdle model, entitlement to benefits both increases the probability of crossing the "programme hurdle" and also the number of programmes, provided the hurdle is crossed.

Higher municipal *unemployment rate* means more policy programmes, as does

Table IV: Estimation results from a standard Poisson count data model, a zero-inflated Poisson model and a two-stage Poisson hurdle model. Size of estimation sample is 16,284.

	Standard		ZIP		Stage1		Stage 2	
	Par.	t	Par.	t	Par.	t	Par.	t
Constant	-4.463	-27.19	-4.457	-29.33	-3.313	-14.48	-6.730	-26.55
Age, ln	-0.301	-7.46	-0.304	-8.36	-0.464	-8.02	-0.239	-4.54
Male (=1)	0.002	0.10	0.010	0.63	0.003	0.12	0.002	0.10
Foreign citizen (=1)	0.002	0.05	0.003	0.11	-0.046	-0.92	0.019	0.46
High school/univ. (=1)	0.015	0.82	0.017	1.02	0.002	0.05	0.024	1.00
Stockholm (=1)	-0.077	-2.28	-0.073	-2.26	-0.103	-2.09	-0.020	-0.44
Voc. education (=1)	0.012	0.58	0.002	0.11	0.011	0.34	0.017	0.64
Some exp. (=1)	-0.010	-0.42	-0.007	-0.32	-0.012	-0.34	-0.013	-0.41
Long exp. (=1)	-0.047	-2.06	-0.037	-1.81	-0.054	-1.66	-0.049	-1.66
UI benefits (=1)	0.191	8.78	0.194	9.48	0.184	6.48	0.182	5.69
Cash assistance (=1)	0.339	9.12	0.357	9.93	0.279	4.41	0.382	8.00
Hearing/vision (=1)	0.044	1.12	0.030	0.83	-0.058	-0.89	0.098	2.00
Disability (=1)	-0.016	-0.76	-0.020	-1.00	-0.032	-0.95	-0.006	-0.22
Psychical disab. (=1)	0.055	1.77	0.027	1.02	0.002	0.04	0.089	2.20
Social med. disab. (=1)	-0.014	-0.45	-0.017	-0.57	-0.028	-0.66	-0.009	-0.20
Unemployment/10, ln	0.080	1.99	0.082	2.21	0.115	1.92	0.071	1.38
"Disability rate"/10, ln	-0.118	-3.43	-0.119	-3.76	-0.175	-3.46	-0.092	-2.09
"programme rate"/10, ln	0.366	10.48	0.359	11.03	0.567	11.04	0.286	6.28
Duration, ln	0.914	85.30	0.910	87.81	0.850	65.13	1.201	46.87
ψ			-0.029	-8.38				

a higher "*programme rate*", reflecting the availability of policy programmes at the local level. If many unemployed are disabled, reflected by a higher "*disability rate*", competition for programmes increases and the expected number of programmes decreases, as expected. These significant effects suggests that the "supply side" of the problem is of importance. Since the local labour market indicators are used in logarithmic form, the parameters in Table IV should be interpreted as elasticities. This means that if the unemployment rate increases by 1 percent, programme participation among disabled increases by 0.08 percent; a quite low elasticity. On the other hand, if the local programme supply increases by 1 percent, the expected number of programmes for a disabled unemployed increases by 0.37 percent. Since only between five and eight percent of the inflow to unemployment are disabled workers, we see that the disabled get a larger share of policy programmes, than do the non-disabled.

In this paper we have not discussed if and how the workers escape unemployment. In Table V we offer descriptive statistics on unemployment duration by programme participation. It is obvious that the number of policy programmes increases with unemployment duration. Mean unemployment of those who did not participate in any programme at al., is about 7.5 months. When it comes to exit from unemployment, the shares who leave for regular employment, are about the same for both "non-participants" and "participants". However, "non-participants" leave the labour force to a larger extent and go into subsidized/sheltered employment to a lesser extent than "participants". How programme participation affects the probability to leave unemployment, when participation is modelled to account for the complexity of programme structure among participants, is a question that remains to be answered.

6 Conclusions

Selection into active policy programmes is a very complicated question. The selection problem has in studies usually been considered to be between e.g. "training" and "not training". In practice the selection process and programme participation behaviour is much more complicated. Unemployed workers in Sweden may participate in up to ten different policy programmes, of which some serve as substitutes to

Table V: Comparison of unemployed with zero programmes and unemployed with one or more programmes. Mean/proportion and standard deviation in parenthesis. Estimation and holdout sample together, n=32,609.

	Zero	Any	Number of programmes				
	programmes	programme	1	2	3	4	5+
Unemployment	232	694	434	785	1,025	1,145	1,226
duration, days	(238)	(435)	(330)	(373)	(323)	(260)	(208)
Exit from unemployment:							
Regular employment	0.15	0.13	0.15	0.13	0.12	0.08	0.06
Subsidised employment	0.14	0.26	0.25	0.28	0.26	0.24	0.21
Leave labour force	0.52	0.32	0.44	0.29	0.16	0.10	0.07
Censored/lost contact	0.19	0.29	0.16	0.30	0.46	0.58	0.65
n	14,918	17,691	8,733	4,312	2,405	1,313	928

each other while others may be complements. These facts make it extremely difficult to model selection into programmes, not to mention to estimate the post-programme effects on e.g. earnings or unemployment durations.

For a labour market policy like the Swedish, which relies heavily on the offering of active policy programmes to unemployed workers, we may ask to whom the programmes are given. From a large sample of unemployed disabled workers we know that the clustering of programmes to some individuals is prominent, and about half of the sample do not participate in any programmes at all. According to the estimated hurdle model, the expected number of policy programmes the individuals participate in decreases with age, while there is no significant difference between men and women. Being eligible to unemployment benefits increases the probability of "hurdle crossing" and once the hurdle has been crossed, eligible workers also participate in more programmes.

The most important question is if some disadvantaged workers (with little or no education, without prior labour market attachment, etc.) are more or less excluded from programme participation. The empirical analysis confirms that this seems to be the case. Among the Swedish disabled workers, those who are entitled to unemployment benefits or cash assistance participate in more programmes than others. This means that those who already have an attachment to the labour market, are favoured for policy programmes.

For future research it would be interesting to investigate if some policy programmes are used as substitutes for retirements, leading to an observed "recycling" in policy programmes? (see Riphahn, 1997). The group of unemployed analyzed here are particularly prone to retirements. In reality there is a complicated pattern between open unemployment, participation in policy programmes, sheltered/subsidized employment aimed specifically at disabled unemployed, temporary disability pensions and retirements. This paper is a first attempt to model the large variation in programme participation among disabled unemployed in Sweden.

References

Aakvik, A., Heckman, J.J. and Vytlacil, E.J. (1998). Training Effects on Employment when the Training Effects are Heterogenous: An Application to Norwegian Vocational Rehabilitation Programs. *Working paper.*

Ackum Agell, S. (1995). Swedish Labour Market programmes: Efficiency and Timing. *Swedish Economic Policy Review*, **2**, 65-98.

Anderson, K.H., Burkhauser, R.V. and Raymond, J.E. (1993). The Effect of Creaming on Placement Rates under the Job Training Partnership Act. *Industrial and labour Relations Review*, **46**, 613-624.

Arulampalam, W. and Booth, A.L. (1997). Who Gets over the Training Hurdle? A Study of the Training Experiences of Young Men and Women in Britain. *Journal of Population Economics*, **10**, 197-217.

Brännäs, K. (1996). Count Data Modelling: Measurement Error in Exposure Time. *Umeå economic studies,* **423**.

Brännäs, K. and Eriksson, M. (1996). Enrolment in Labour Market Training Programmes. *Labour*, **10**, 193-208.

Cameron, A.C. and Trivedi, P.K. (1986). Econometric Models Based on Count Data: Comparison and Applications of some Estimators and Tests. *Journal of Applied Econometrics,* **1**, 29-53.

Edin, P-A. (1991). Unemployment Duration and Competing Risks: Evidence from Sweden. *Scandinavian Journal of Economics*, **91**, 639-653.

Edin, P-A. and Holmlund, B. (1991). Unemployment, Vacancies and Labour Market Programmes: Swedish Evidence. In F. Padoa Schioppa (ed.), *Mismatch and Labour Mobility*; Cambridge University Press.

Greene, W.H. (1997). *Econometric Analysis.* 3rd edition. Prentice Hall International, Inc., New Jersey.

Gurmu, S. and Trivedi, P.,K. (1996). Excess zeros in Count Models for Recreation Trips. *Journal of Business & Economic Statistics*, **14**, 469-477.

Harkman, A. and Jansson, F. (1995). *Sökaktivitet,återanställning och chansen att få jobb* (*Search Activities, Re-employment and the Chance of Getting a job*). Ura 1995:4, National Labour Market Board, Stockholm. Available in Swedish only.

Heckman, J.J. and Borjas, G.J. (1980). Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence. *Economica*, **47**, 247-283.

Heckman, J.J., Smith, J.A. and Taber, C. (1996). What do Bureaucrats do? The Effects of Performance Standards and Bureaucratic Preferences on Acceptance into the JTPA Program. NBER Working Paper No. 5535.

Lambert, D. (1992). Zero-Inflated Poisson Regression with an Application to Defects in Manufacturing. *Technometrics*, **34**, 1-14.

Melkersson, M. (1997). Explaining Choice Set Size for Unemployed. *Umeå economic studies,* **438**. Forthcoming in *Applied Economics.*

Melkersson, M. (1999). Unemployment Duration and Heterogenous Search Behavior among Disabled Workers. Unpublished manuscript.

Mullahy, J. (1986). Specification and Testing of Some Modified Count Data Models. *Journal of Econometrics,* **33**, 341-365.

Pudney, S. (1989). *Modelling Individual Choice: The Econometrics of Corners, Kinks and Holes*, Oxford U.K. Blackwell.

Winkelmann, R. and Zimmermann, K.F. (1995). Recent Developments in Count Data Modelling: Theory and Application. *Journal of Economic Surveys*, **9**, 1-24.