ECONOMIC STUDIES

DEPARTMENT OF ECONOMICS
SCHOOL OF ECONOMICS AND COMMERCIAL LAW
GÖTEBORG UNIVERSITY
114

ESSAYS ON TRAINING, WELFARE AND LABOR SUPPLY

Thomas Andrén

Abstract

This thesis contains four essays in applied labor economics. The first 2 papers evaluate labor market training programs in Sweden in terms of earnings gain and reemployment probability. The third paper analyses the exit behavior from social assistance dependency and the last paper analyses the simultaneous relationship between welfare participation, paid childcare utilization and the labor supply of single mothers.

Paper 1 investigates labor market training for three cohorts during the 80s and the beginning of the 90s on its effect on earnings. We separate the analysis between Swedish-born and foreign-born individuals to identify differences in their responses to training. The results indicate that there is positive sorting into training. We find that the proportion of trainees having positive rewards from training was not very different from the proportion having negative rewards. This means that the results do not support the view that from efficiency considerations, too few persons were enrolled in labor market training during this period. Differences in results across cohorts can be interpreted as being caused by rapid changes in the labor market. Further, consistent with results from several previous studies we find that being young often means no positive pay-off from training, and the same is found for persons with only primary education. Rewards from training were higher for foreign-born than for natives and rewards among the former vary by place of birth

Paper 2 uses an econometric framework that allows for heterogeneous training effects on the employment probability. We separate the analysis between Swedish-born and foreign-born individuals. We investigate the importance of the unobservables in the selection to training and how efficient the selection is with respect to the outcome. The results show small positive effects for the Swedish-born. The treatment on the treated is larger than the average treatment effect, indicating that the selection is stronger for the treated, and 40% of those treated gain by participating in training. Foreign-born had a negative training effect the first year, with an average treatment effect larger then the treatment on the treated. From those who participated in training, only 11% experienced positive effect while 38% were hurt by the training. The unobserved factors are important in the selection to training as well as for the outcome.

Paper 3 analyses Swedish-born people who became first-time receivers of social assistance in 1987 and 1992. We find that pattern of social assistance receipt is rather heterogeneous across new recipients. The complex pattern of receipt means that due to choice of perspective, duration of social assistance can appear rather different. On one hand, we find that median duration of social assistance receipt is as low as five months when an eleven-year follow-up period is applied. On the other hand, among people who receive social assistance during one particular year, as many as half had, entered receipt more than four years earlier.

Paper 4 considers the simultaneous relationship of the single mother's decision to choose paid childcare, welfare participation and labor supply, and estimates a structural model that allows for a free error covariance. The results show that there is an association between social assistance, paid childcare and labor supply, but that the relationship is non-symmetric. An increase in the social assistance norms has a relatively small effect on paid childcare utilization, but a relatively larger effect on the mean labor supply. In contrast, a corresponding reduction in the childcare cost has a relatively large effect on the social assistance utilization but a relatively small effect on the mean labor supply. Our estimates suggest that a decrease in childcare cost increases the labor supply of those working rather than encourages non-workers to start work, which implies that childcare cost is foremost a barrier to fulltime work rather then a barrier to work at all.

Keywords: Labor market training, sample selection, heterogeneous treatment effects, social assistance, structural model, simulated maximum likelihood, labor supply.

ISBN 91-88514-74-9

Acknowledgements

Several persons have contributed in various ways to this thesis. First and foremost, I am indebted to my two supervisors, Professor Björn Gustafsson and Professor Lennart Flood for their guidance and support during the work with this thesis. I started to work with Björn as a research assistant in 1996, which in part have resulted in two papers included in this dissertation. This work has been most rewarding and giving me important insights into the art of working with large data sets, and how it is to collaborate in an international project. The need and interest for more advanced econometric methods grew, and two years later, Lennart became my advisor in such matters. Fortunately, Lennart and Professor Thomas MaCurdy gave me the opportunity to spend one year at Stanford University, for which I am very grateful. This year immensely improved my knowledge and understanding in applied econometrics. Additionally, the last paper of this dissertation was built on an idea suggested by Tom.

I would also like to thank Professor Lennart Hjalmarsson for the unconditional help given when most needed.

Financial support from The Office of Labor Market Policy Evaluation, Swedish Council for Work Life Research, HSFR (ICPSR-Stipendium), Stiftelsen Siamon, Knut och Alice Wallenbergs Stiftelse, Kungliga Vetenskaps Akademien, Stiftelsen Lars Hiertas Minne, Resestipendier för forskarstuderande from Göteborg University is gratefully acknowledged.

Nevertheless, I would like to thank Bengt and Ronnie, and my very parents for their support and patience with my absence, and for their effort in trying to understand what I have been working with.

Most of all, I would like thank my wife, Daniela, for her continuous love and support during these years, and for being my favorite co-author.

Göteborg, May 2002

Thomas Andrén

Income Effects from Labor Market Training Programs in Sweden during the 80's and 90's[a]

Thomas Andrén[α] and Björn Gustafsson[β]

Abstract
Swedish labor market programs appear large from an international perspective, yet their consequences are not fully investigated and understood. In this paper, we estimate a switching regression model with training effect modeled as a random coefficient, partitioned in an observed and unobserved component. We investigate labor market training for three cohorts during the 80s and the beginning of the 90s on its effect on earnings. We separate the analysis between Swedish-born and foreign-born individuals to identify differences in their responses to training. The results indicate that there is positive sorting into training. We find that the proportion of trainees having positive rewards from training was not very different from the proportion having negative rewards. This means that the results do not support the view that from efficiency considerations, too few persons were enrolled in labor market training during this period. Differences in results across cohorts can be interpreted as being caused by rapid changes in the labor market. Further, consistent with results from several previous studies we find that being young often means no positive pay-off from training, and the same is found for persons with only primary education. In conflict with what earlier studies have shown, we found that males have a better pay-off from training than females. Rewards from training were higher for foreign-born than for natives and rewards among the former vary by place of birth.

Keywords: labor market training, non-experimental estimator, positive sorting, unobserved heterogeneity to training reward, random coefficient model.

JEL classification: J31, J38.

[a] We thank the Institute of Labor Market Policy Evaluation (IFAU) for financial support, and Christoph M. Schmidt, and participants at IFAU seminar in autumn 2000 for useful comments. The usual disclaimer applies.
[α] University of Göteborg, Department of Economics, Box 640, SE 405 30 Göteborg, E-mail: Thomas.Andren@economics.gu.se
[β] University of Göteborg, Department of Social Work & Institute of Labor (IZA), Bonn, Germany, E-mail: Bjorn.Gustafsson@socwork.gu.se

1 Introduction

The efforts industrialized countries make to train the unemployed and persons at risk of becoming unemployed vary dramatically. Statistics for the 90s for example (OECD 1997, 2000) show one group made up of the Czech Republic, Japan, Luxembourg, Poland and the USA where public expenditures on labor market training programs were less than 0.05% of GDP. The other extreme, with public expenditures on labor market training programs of around 0.5% of GDP or higher is found in the Scandinavian countries (Denmark, Finland and Sweden). Of the three, Sweden has the longest record of allocating massive funds to labor-market training programs. The extensive public involvement in training the unemployed in Sweden started at the beginning of the 1960s although it is possible to find even earlier efforts. During the 60s the number of participants in training, henceforth trainees, increased rapidly, after which followed several examples of contra-cyclical changes.

Who receives training? This is a central issue when setting up as well as evaluating labor market training programs. The selection of trainees can be affected by the preferences of potential trainees as well as by the officials responsible for recruiting trainees, who in turn must follow instructions dictated by politicians. Starting with a potential trainee, one obvious reason for applying is the perception that the training program will improve his or her future position in the labor market when compared to not taking part in the training.

Turning to the role of placement officers, it can be noted that in Sweden public employment offices have a central role of assigning job seekers to training courses. These officers are responsible for providing information on different courses, eligibility rules, training stipends etc. The main motivation for assigning a person to labor market training is that the training should lead to a permanent job. Those eligible are mainly unemployed job seekers and those at risk of becoming unemployed. A person can also be eligible for other reasons. For example, political refugee status makes a foreigner eligible for training within a certain time limit after arrival.

How does training affect a person's subsequent labor market situation? Obviously training can increase the human capital of the trainee by increasing skills. Even if training only serves to preserve the human capital the effect is positive if the alternative

(of continued unemployment) leads to decreased work-capacity. However, there can also be other mechanisms affecting the future labor market situation of a trainee.

Taking part in training might lower the person's reservation wage, making the person more likely to accept a work offer and thereby more likely to be employed. However, being involved in a training program might lead to reduced job search intensity. If this is the case, training can reduce the rate at which job offers arrive, thus reducing employment opportunities. Still another mechanism at work is that a certificate of a completed training course might act as a positive (negative) signal to potential employers. Such persons can be perceived as more (less) ambitious and therefore more (less) productive than other job-searchers.

Given the considerable resources spent on labor market training programs in Sweden, it is not surprising that training programs have been subject to several research efforts. Some authors have studied the enrolment in labor market training programs and the choice set of the unemployed [Brännäs & Eriksson (1996), Eriksson (1997), Melkersson (1999)]. A number of studies have used non-experimental methods to evaluate the subsequent labor market performance of trainees. Some of these studies have focused on particular groups. Examples include Edin (1988) who studied training among workers displaced by a pulp plant closing in 1977, in a small town in the north of Sweden; Ackum (1991) who studied persons aged 16-24 in Stockholm in 1981; and Larsson (2000) who analyzed persons aged 20-24 who became unemployed during 1992 and 1993.

Still other studies have analyzed samples taken from the whole country and without any narrow restrictions on the age of the trainee; this is the approach taken in this paper. There are four previous studies, which in this aspect are similar to ours. First, Björklund & Moffitt (1987) who distinguished between effects for the average and the marginal participant. Using data from the second half of the 70s in which relatively few trainees are found, the average effects were found to be positive while the marginal impacts were found to be negative. Axelsson (1989) compared a sample of 900 persons who completed labor market training programs in 1981 with various control groups. Outcomes, one and two years after the training, were evaluated by several variables. The results show programs to have a significant positive effect on annual earnings amounting to about 20% in 1983.

Another study is Regnér (1993, 1997) where the research strategy is similar but the samples larger and the econometric method of a later vintage. The study investigated training that took place during 1989 and 1990. The results indicate that training did not increase subsequent earnings of the trainee. This study also patterned AMS (1995) in which persons who received training in 1992 and 1994 were investigated along with control groups. It was found that subsequent earnings for the second cohort (as measured half a year later) were positive, but were negative (although not significant), for the first cohort (as measured two years after training).

Our study is inspired by the studies mentioned above but differs in a number of aspects. First, we study three different training cohorts; people who received training during the two-year periods of 1984 and 1985, 1987 and 1988, and 1990 and 1991. The macroeconomic climate varied across these cohorts as the unemployment rate in Sweden fell from a maximum of 3.5% in 1983, reached a minimum of 1.5% in 1989, rose to 3.0% in 1991, and more or less exploded to 8.2% in 1993. Thus we are able to investigate if the outcome of labor market training is affected by the macroeconomic climate, hypothesizing that positive earning effects are easier to find when there is excess demand.

Second, foreign-born persons make up a considerable proportion of all people in labor training in Sweden. We consider this fact at the outset in the sampling process and work with different samples for natives and foreign-born. This research strategy is also motivated by the fact that immigrants often are enrolled in courses other than the courses natives are enrolled in, a fact which provides a strong argument for working with different samples of natives and foreigners.

We follow the three cohorts of trainees and their control groups for three years after completed training. As the primary outcome variable we analyze annual earnings. In the econometric strategy we follow Björklund and Moffitt (1987). We estimate a switching regression model while allowing for unobserved heterogeneity with respect to the reward on training. This allows us to investigate how the reward is distributed over the individuals and their observed characteristics.

The rest of the paper is as follows: In the next section the theoretical framework is laid out while the empirical specification and parameters of interest are discussed in

Section 3. The data is presented in Section 4. Section 5 presents and discusses the results and in Section 6 we sum up the conclusions.

2 The economics of sorting

To discuss the economics of sorting, it is convenient to define two states (treatment and no-treatment) with respect to the outcome variable of interest. We are interested in earnings and the effect on earnings from training. Hence we define two earnings equations representing the potential outcome in the post-training period for the individual:

$$\begin{aligned} Y_1 &= X\beta_1 + U_1 \\ Y_0 &= X\beta_0 + U_0 \end{aligned} \tag{1}$$

A linear decomposition with an additively separable representation is assumed, X being a vector of observables and U_1 and U_0 being mean zero unobservables of the individual. Subscript 1 represents the potential earnings if the individual participate and complete a training program and 0 the potential earnings if the individual choose not to participate in a program. It is assumed that training takes place only once, during a fixed period of time, and no other training has taken place or will take place in the future. Assuming that the individual wishes to maximize the future earnings, the decision to undergo treatment is made on the basis of a net reward function:

$$D^* = Y_1 - Y_0 - C = \alpha - C \tag{2}$$

D^* is a latent variable representing the net reward from training, C the cost associated with training, and α the gross reward in terms of earnings. C can be thought of as some non-earnings related considerations that are relevant to the decision to undertake treatment such as tuition, stigma, distance to training center etc. When $C = 0$ the model coincides with the so-called Roy-Model (Roy, 1951) (Heckman and Honoré, 1990) where an individual's decision to participate in training is a function of potential earnings only.[1] In general, costs are relevant and include variables beside those included in X, capturing differences in cost across individuals.

[1] This model also goes under the name the Neyman-Fisher-Cox-Roy-Quandt-Ruben model, especially at the University of Chicago.

To specify the model further, we formalize the reward and the cost. A general formulation of the model would allow for observed and unobserved heterogeneity in both the rewards and costs associated with training. It is therefore natural to define separate behavioral equations for rewards and costs both of which include observed and unobserved components:

$$\text{Reward: } \alpha = Z\gamma + \varepsilon_\alpha \tag{3}$$

$$\text{Cost: } \quad C = W\delta + \varepsilon_c \tag{4}$$

The unobserved component of the reward equation (3) is defined as the difference between the residuals of the state specific earnings equations ($\varepsilon_\alpha = U_1 - U_0$) and from (2), (3) and (4), the unobserved component of the decision function (D*) is the difference between the unobserved components in (3) and (4) ($\varepsilon = \varepsilon_\alpha - \varepsilon_c$). The full model is now defined and we have access to three behavioral components. The behavioral terms (U_1, U_0, ε) are assumed to be independent of the exogenous variables in the model, with variances (σ^2_j, σ^2_ε) and covariances (σ_{ij}) for i,j = 1,0. The covariances of the pairs (ε, U_1), (ε, U_0) are denoted $\sigma_{1\varepsilon}$ and $\sigma_{0\varepsilon}$. The individual's decision to participate is based on perfect foresight of future net reward. That is if D* is positive the individual will participate in training. In the opposite case, no training will take place for the individual. Relaxing this assumption by assuming that only the expected value of the net reward is known by the individual would not change the reduced form of the decision rule, although ε would not include U_1 and U_0.

To discuss the economics of sorting into the two states we will refer to U_1 and U_0 as state-specific skills [Heckman and Honoré (1990), Vella et al. (1998)]. When $\sigma_{10} < 0$ the state-specific skills are negatively correlated and we have a *comparative advantage structure*. That is, on average those who perform relatively well with the treatment will perform relatively less well without the treatment. When $\sigma_{10} > 0$ the state-specific skills are positively correlated and we have a *hierarchical structure*, where on average those individuals who perform well in one state, will also perform relatively well in the other state. The conditional expectations of the unobserved components of the potential earnings functions are of special interest. If $E[U_i \,|Z,W,D=i,\ i=1,0] > 0$, where D is an indicator that takes the value 1 when training take place and 0 otherwise, we say that the

selection is positive. The conditional expectations of the state specific residuals can be decomposed into two parts:

$$E[U_1 \mid Z, W, D = 1] = \frac{\sigma_{1\varepsilon}}{\sigma_\varepsilon^2} E[\varepsilon \mid \varepsilon > W\delta - Z\gamma] \tag{5}$$

$$E[U_0 \mid Z, W, D = 0] = \frac{\sigma_{0\varepsilon}}{\sigma_\varepsilon^2} E[\varepsilon \mid \varepsilon < W\delta - Z\gamma] \tag{6}$$

with $\sigma_{1\varepsilon} = \sigma^2_1 - \sigma_{10} - \sigma_{1c}$ and $\sigma_{0\varepsilon} = \sigma_{01} - \sigma^2_0 - \sigma_{0c}$. The expectations on the right hand sides of (5) and (6) have fixed signs. In (5) the expectation is always positive and in (6) it is always negative. With that in mind, it is the covariance that determines the sign of the conditional expectations on the left. The signs cannot be determined from the theoretical model and become therefore an empirical question. The sizes and signs of $\sigma_{1\varepsilon}$, $\sigma_{0\varepsilon}$, and σ_{10} discussed above identify three different structures (Willis, 1986).

We consider the case where the unobserved cost component is irrelevant or uncorrelated with the state skills ($\sigma_{1c} = \sigma_{0c} = 0$).[2] The first structure is the *positive hierarchical sorting* and rules when $\sigma^2_1 > \sigma_{10} > \sigma^2_0$, which is equivalent with $\sigma_{1\varepsilon} > 0$ and $\sigma_{0\varepsilon} > 0$. Those who receive training are those who are drawn from the upper portion of the distribution of the potential earnings in state 1, while those who do not enter training are those who are drawn from the lower portion of the distribution of the potential earnings in state 0. In this state we have a positive selection into training and negative selection into non-training.

The second structure is the *negative hierarchical sorting* and rules when $\sigma^2_1 < \sigma_{10} < \sigma^2_0$ which corresponds to $\sigma_{1\varepsilon} < 0$ and $\sigma_{0\varepsilon} < 0$. This is the opposite case of the previous structure, which usually has little empirical importance.

The third structure is the *non-hierarchical sorting* which occurs when $\sigma^2_1 > \sigma_{10}$ and $\sigma^2_0 > \sigma_{10}$ which corresponds to $\sigma_{1\varepsilon} > 0$ and $\sigma_{0\varepsilon} < 0$. In this structure the sign of σ_{10} can be either positive or negative. The signs of the covariances between state and selection imply a positive selection into both training and non-training. In general, this case applies when σ_{10} is sufficiently small or if the scopes of the state-specific skills are about the same. The structure indicates that a person who enters state 1 would have had a higher reward in doing so as opposed to the alternative, while those who enter state 0

[2] In the empirical analysis we impose the same assumption in order to simplify the estimation. Hence the discussion is directly linked to the model that is estimated.

would be better off there as opposed to the alternative. This implies that the selection makes the groups above average in their state specific outcome distribution. In general the non-hierarchical sorting implies that there is more than one distinct ability factor and that the direction of the ability bias is uncertain.

3 The empirical model specification

Most of the empirical literature on evaluating governmental training programs focuses on mean effects and, in particular, on the mean direct effect of treatment on those who receive training [Heckman et al. (1998)]. In this paper we use the standard index sufficient method of the prototypical selection model formulated by Björklund and Moffitt (1987) so that the individual reward from training can be identified, and allowed to be unique for each individual [Heckman et. al. (1985,1986)]. This approach to the selection problem allows for selection on unobservables, which is an important motivation for the choice of estimator since selection into training to a large extent is determined by the ambition of the unemployed.[3] Ambition is usually something that is unobserved and finding a good instrument for it would require unique data that is not at hand.

If we adopt the separability assumption mentioned earlier with a linear restriction in the parameters we may form the observed Y:

$$Y = DY_1 + (1 - D)Y_0 \tag{7}$$

By inserting (1) into (8) we obtain

$$Y = X\beta_0 + D[X(\beta_1 - \beta_0) + (U_1 - U_0)] + U_0 \tag{8}$$

[3] Eriksson (1997) carried out an informal telephone interview with Swedish officials and found that in the contact between the administrator and unemployed individuals, ambition and motivation of the unemployed were important for recruitment to a training program. Åtgärdsundersökning (1998, AMV) interviewed individuals who participated in a program in 1997. This survey showed that 60% of the participants took the initiative to participate in the training program (i.e., by getting informed about different courses and programs from ring binders, billboards, and computer terminals available at the unemployment office). The administrator's role is more important for foreign-born (non-Nordic) people in their decision to participate.

This gives us the two regimes switching regression model (Quandt, 1972). The term multiplying D is the gain from the program where D is the observed binary analogue of the latent continuous variable D*. The gain has two components. The first component, $X(\beta_1 - \beta_0)$, is the gain from the average person with characteristics X in the population. This term is the so-called *experimental treatment average*, and would be the treatment effect in case of a social experiment [(Heckman et. al 1996), (Heckman, 1990)]. Typically this parameter is of limited interest in policy analysis since it constitutes the average gain for a person taken randomly from the population, which is a group that doesn't coincide with the target population of labor market programs. The second component, $U_1 - U_0$, is the idiosyncratic gain from a particular person. This component will be zero if agents do not know their gain or do not act on it. The best forecast would then be zero and no additional effect due to self-selection would be present. This two-component effect is non-standard in conventional econometrics since it combines the "structural" effect, $X(\beta_1 - \beta_0)$, with a stochastic effect, the change in the unobservables, $U_1 - U_0$. With this set-up we can construct three parameters that usually are estimated in the literature. The effect of the *treatment on the treated* (TOT), the effect of the *treatment on the non-treated* (TUT) and the *average treatment effect* (ATE) respectively:

$$E[Y_1 - Y_0 \mid X, D = 1] = X(\beta_1 - \beta_0) + E[U_1 - U_0 \mid X, D = 1] \quad (9)$$

$$E[Y_1 - Y_0 \mid X, D = 0] = X(\beta_1 - \beta_0) + E[U_1 - U_0 \mid X, D = 0] \quad (10)$$

$$E[Y_1 - Y_0 \mid X] = X(\beta_1 - \beta_0) + E[U_1 - U_0 \mid X] \quad (11)$$

All three estimators give the same results when $E[U_1 - U_0 \mid X, D] = E[U_1 - U_0 \mid X] = 0$. This can happen only if $U_1 = U_0$ or if agents either do not know $U_1 - U_0$ or do not act on it. If they are the same it means that a change in one state will result in the exact same change for the individual in the other state. This implies that when we condition the difference on X, everyone with the same X has exactly the same treatment effect. We think this is an unnecessary restrictive assumption and therefore allow for idiosyncratic gain in the model.

3.1 The random coefficient model

In order to account for the unobserved heterogeneity, one needs to make a distributional assumption for the idiosyncratic gain. If no such assumption is made, the individual gain will not be identified and only the mean sum of the two components could be estimated. We separate the reward from training from the cost of training. The selection rule then says that when the reward exceeds the cost, the individual chooses to participate in training. Formally we may express the model in the following way:

$$Y = X\beta + \alpha + U \qquad \text{when } D = 1 \tag{12}$$

$$Y = X\beta + \qquad U \qquad \text{when } D = 0$$

$$\text{Reward: } \alpha = Z\gamma + \varepsilon \tag{13}$$

$$\text{Cost: } C = W\delta \tag{14}$$

$$\text{The selection rule: } D = \begin{cases} 1 & iff \quad D^* = \alpha - C > 0 \\ 0 & iff \quad D^* = \alpha - C < 0 \end{cases} \tag{15}$$

Each regime is allowed to have its own error term with a separate variance, and free correlation between the choice equation and the two regimes. We do not allow for any unobserved heterogeneity with respect to cost, primarily to decrease the complexity of the model, but also since our focus is on the heterogeneity to rewards.[4] In this paper we will estimate a random coefficient model using maximum likelihood technique.[5] We therefore define the following likelihood function:[6]

$$L = \left[P[\varepsilon + u, \varepsilon > W\delta - Z\gamma] \right]^{D} \left[P[u, \varepsilon \leq W\delta - Z\gamma] \right]^{1-D} \tag{16}$$

Few identifying restrictions have been applied. One important restriction is the parameters in the reward equation. In order to identify the variance of the reward we

[4] Björklund and Moffitt (1987) estimate this model allowing for unobserved heterogeneity with respect to cost as well. When they tested if this contributed to the model they received insignificant test results indicating that the unobserved cost components have a minor importance in the model.

[5] The distributional assumption made for the likelihood function is that U and ε have a bivariate normal distribution. But the residuals for the two earnings equations are defined differently dependent on D. When D=1, $U_1 = \varepsilon + U$, and when D=0, $U_0=U$. This implies that we implicitly have defined a trivariate normal density with U_1, U_0, and ε. We can therefore indirectly derive the elements of the tri-variate covariance matrix.

[6] The exact definition of the likelihood function is described in Appendix.

have normalized $\gamma = 1$, in the selection equation while allowing it to be unrestricted in the wage equation. This works if we have at least one variable in Z that is not in W. This model does not formally require an exclusion restriction between the selection equation and the earnings equation. That the exclusion restriction can be useful in this case is shown by Monte Carlo studies finding that the estimator performs poorly when exclusion restrictions are not imposed.[7] No other restrictions need to be imposed. The treatment on the treated is therefore defined as

$$E[\alpha \mid D = 1, Z\gamma, W\delta] = Z\gamma + \sigma_{\varepsilon}E[\varepsilon \mid \varepsilon > -Z\gamma + W\delta] \quad (17)$$

The variable specifications pertaining to the different equations are important. The variables explaining the outcome equation are standard, namely age, gender and education. The ambition has been to have the specification as parsimonious as possible yet including what is relevant, and accessible. The reward to training is explained by the same observed factors as in the outcome equation. We find no reason to include anything there that was not in the earnings equation. The cost equation is more complicated. It should include non-earnings related factors such as preferences and foregone income etc, which are not included in our data set. It could be argued that ability to learn decreases by age and therefore induces negative preferences for training. Preference towards training might also differ between genders, in the sense that women and men respond differently to training. Distance to the training center is another factor that might induce a cost. Living in a big city region might therefore create the feeling of being closer to the training center then living elsewhere.

Heckman et al (1999) argue that this model emphasizes changes in the opportunity costs, i.e., foregone earnings, as the major determinant of participation in training programs. They show evidence that suggests that changes in labor force status predict participation in programs. We therefore include number of days of unemployment the year before training as a factor. The variables used as exclusion restrictions are big-city region and previous unemployment. Intuitively we feel that distance to training center is correlated with the selection process while not correlated

[7] It is important, however, that the instrument that constitutes the exclusion restriction is good in the sense that it is correlated with the selection process but uncorrelated with the outcome variable. It can be hard to find good instruments unless one specially designs a survey for this purpose.

with the outcome variable. In the same way we are convinced that earlier unemployment situation is correlated with the selection process, and we have evidence mentioned above that such is the case in the US. In the data section it will be apparent that we have a pre-training dip in earnings in Sweden, which therefore further confirms the relevance of the variable in our case.

The foreign-born group has an extended variable specification in both the earnings and the reward equation. Number of years in the country and country of origin play an important role in the determination of the individual's success in the labor market and therefore also on the reward and earnings of participating in a training program. We have therefore included such variables to control for any observed differences related to those factors.

4 Data

Descriptive statistics presented in Tables 1a and 1b show that the treatment group for the 1984-85 cohorts consists of 495 natives and 982 foreign-born. This corresponds to a population in training programs in Sweden of 59,320 persons for these years. The 1987/88 and 1990/91 samples of trainees are of similar sizes and correspond to populations in training programs of 61,420 and 57,410, respectively.

Tables 1-3 show that the gender composition among people in training programs is relatively even for all cohorts. However, in the reference group, the proportion of females is higher than 50% and a high proportion of females is particularly evident among natives. Although there is a variation in age among the trainees, the majority (or nearly the majority) are in the interval 26 to 45 years with a mean of around 30 years for natives and slightly higher for foreign-born. The foreign-born trainees and their comparison groups are more concentrated to the larger cities than their native counterparts. Few trainees have post-secondary education. Looking at immigrant specific variables, it can be seen that about half of the foreign-born trainees in the two first cohorts have lived in Sweden for more than a decade, while many foreign-born trainees in the third cohort are recent arrivals. Across the cohorts of foreign born there is also a shift in region of birth. People born in other Nordic countries make up a considerable proportion of the trainees in the first cohort while a large portion of people born in countries outside Europe makes up the last cohort of trainees.

Table 1a Descriptive statistics for the 1984/85 Swedish-born cohort

	Trainees		Non-trainees	
	Men	Women	Men	Women
Observations	219	271	852	1156
Age (mean)	30.15	31.87	29.98	30.32
20-25(%)	37	36	44	44
26-45(%)	54	55	46	45
46-55(%)	8	9	10	11
Region (Big city) (%)	32	28	27	23
Married (%)	19	39	19	36
Children 0-6 year (%)	12	25	12	23
Children 7-12 year (%)	8	24	9	16
Education (%)				
Primary	27	40	50	52
Secondary	58	50	46	37
Post secondary	15	10	4	11
Unemployed last year (days)	25	18	30	22

Table 1b Descriptive statistics for the 1984/85 foreign-born cohort

	Trainees		Non-trainees	
	Men	Women	Men	Women
Observations	487	495	907	1007
Age (mean)	33.14	32.73	34.12	33.64
20-25(%)	22	26	20	26
26-45(%)	68	63	65	59
46-55(%)	9	11	15	15
Region (Big city) (%)	54	55	20	39
Married (%)	42	58	38	54
Children 0-6 (%)	22	38	21	34
Children 7-12 (%)	15	27	14	26
Education (%)				
Primary	46	55	61	61
Secondary	49	39	36	31
Post secondary	3	6	6	8
Unemployed last year (days)	19	13	30	27
Number of years in the country (%)				
0-5	27	33	11	8
6-10	20	18	28	22
11-	53	48	61	70
Region of birth (%)				
Nordic	40	41	45	60
Northern Europe	9	8	9	9
Eastern Europe	6	15	7	9
Southern Europe	15	9	16	11
Middle East	16	8	14	6
Rest of the world	14	17	9	6

Table 2a Descriptive statistics for the 1987/88 Swedish-born cohort

	Trainees		Non-trainees	
	Men	Women	Men	Women
Observations	220	298	783	1183
Age (mean)	30.42	32.05	30.97	30.98
20-25 (%)	41	38	38	40
26-45 (%)	49	50	51	49
46-55 (%)	10	12	11	11
Region (Big city) (%)	22	21	21	22
Married (%)	15	35	19	32
Children 0-6 year (%)	9	24	12	26
Children 7-12 year (%)	7	23	8	15
Education (%)				
Primary	28	31	44	50
Secondary	62	46	50	37
Post secondary	10	23	6	13
Unemployed last year (days)	17	11	29	25

Table 2b Descriptive statistics for the1987/88 foreign-born cohort

	Trainees		Non-trainees	
	Men	Women	Men	Women
Observations	514	448	937	1013
Age (mean) (%)	33.23	34.00	34.96	34.14
20-25 (%)	23	20	17	22
26-45 (%)	66	66	69	65
46-55 (%)	11	14	15	14
Region (Big city) (%)	53	46	45	37
Married (%)	39	50	37	51
Children 0-6 (%)	20	34	20	37
Children 7-12 (%)	13	31	13	26
Education (%)				
Primary	36	43	57	54
Secondary	57	45	38	35
Post secondary	8	13	4	10
Unemployed last year (days)	16	15	32	30
Number of years in the country (%)				
0-5	36	26	11	8
6-10	16	21	25	23
11-	48	52	64	69
Region of birth (%)				
Nordic	33	44	41	57
Northern Europe	5	8	8	9
Eastern Europe	9	17	10	11
Southern Europe	10	7	11	9
Middle East	26	8	17	6
Rest of the world	16	16	13	9

Table 3a Descriptive statistics for the 1990/91 Swedish-born cohort

	Trainees		Non-trainees	
	Men	Women	Men	Women
Observations	231	246	683	956
Age (mean)	31.14	33.71	31.58	32.63
20-25 (%)	40	28	35	30
26-45 (%)	48	54	52	55
46-55 (%)	13	17	13	15
Region (city) (%)	20	28	25	26
Married (%)	22	42	23	37
Children 0-6 year (%)	15	26	11	26
Children 7-12 year (%)	6	24	7	15
Education (%)				
Primary	35	39	30	28
Secondary	55	54	56	56
Post secondary	10	7	14	16
Unemployed last year (days)	11	10	17	15

Table 3b Descriptive statistics for the 1990/91 foreign born cohort

	Trainees		Non-trainees	
	Men	Women	Men	Women
Observations	467	504	990	937
Age (mean)	33.68	34.22	34.97	34.97
20-25 (%)	17	19	16	16
26-45 (%)	72	69	68	68
46-55 (%)	10	12	15	15
Region (Big city) (%)	44	43	52	42
Married (%)	50	60	39	55
Children 0-6 (%)	29	35	21	34
Children 7-12 (%)	19	31	14	26
Education (%)				
Primary	74	73	50	47
Secondary	23	23	38	40
Post secondary	3	3	11	13
Unemployed last year (days)	9	8	19	15
Number of years in the country (%)				
0-5	52	45	20	16
6-10	16	15	19	17
11-	32	18	61	67
Region of birth (%)				
Nordic	20	29	35	46
North Europe	4	6	8	9
East Europe	8	14	7	11
South Europe	6	6	11	7
Middle East	38	22	25	12
Rest of the world	23	23	13	14

Earnings is our outcome variable, and it consists of incomes from employment and self-employment and is measured on an annual basis. This means that our outcome variable captures wage effects of training as well as effects on number of hours worked. We follow trainees and their counterfactuals during a period of three years before training until three years after training.

Figure 1 shows for all cohorts, natives and foreign-born, how mean earnings have developed for trainees as well as for the comparison group. With the exception of the period of training, the curves for the treatment group and the comparison group rise until the beginning of the 90s after which they decrease. This reflects the general development of real earnings in the Swedish economy during the period under study.

The curves for trainees and non-trainees among natives start at approximately the same level. The curve for trainees then decreases during the period of training, and after the completion of the training period returns to approximately the same level as for the non-trainees. This makes us suspect that the average training reward for natives cannot be large. Turning to foreign-born the situation is somewhat different. Trainees have considerably lower earnings than non-trainees before training, and reach approximately the same level after training. This makes us suspect that the average reward for foreign-born trainees is positive. Figure 1 also indicates the presence of Ashenfelter's dip in the pre-training earnings (Ashenfelter, 1978), which therefore leads us to believe that employment status before training could work as an indicator for selection into training. The exception is the third cohort of the foreign-born people who do not dip or start their dip much earlier before the observation window. This situation is partly explained by the large number of newly arrived immigrants that apparently had little or no earnings before training.

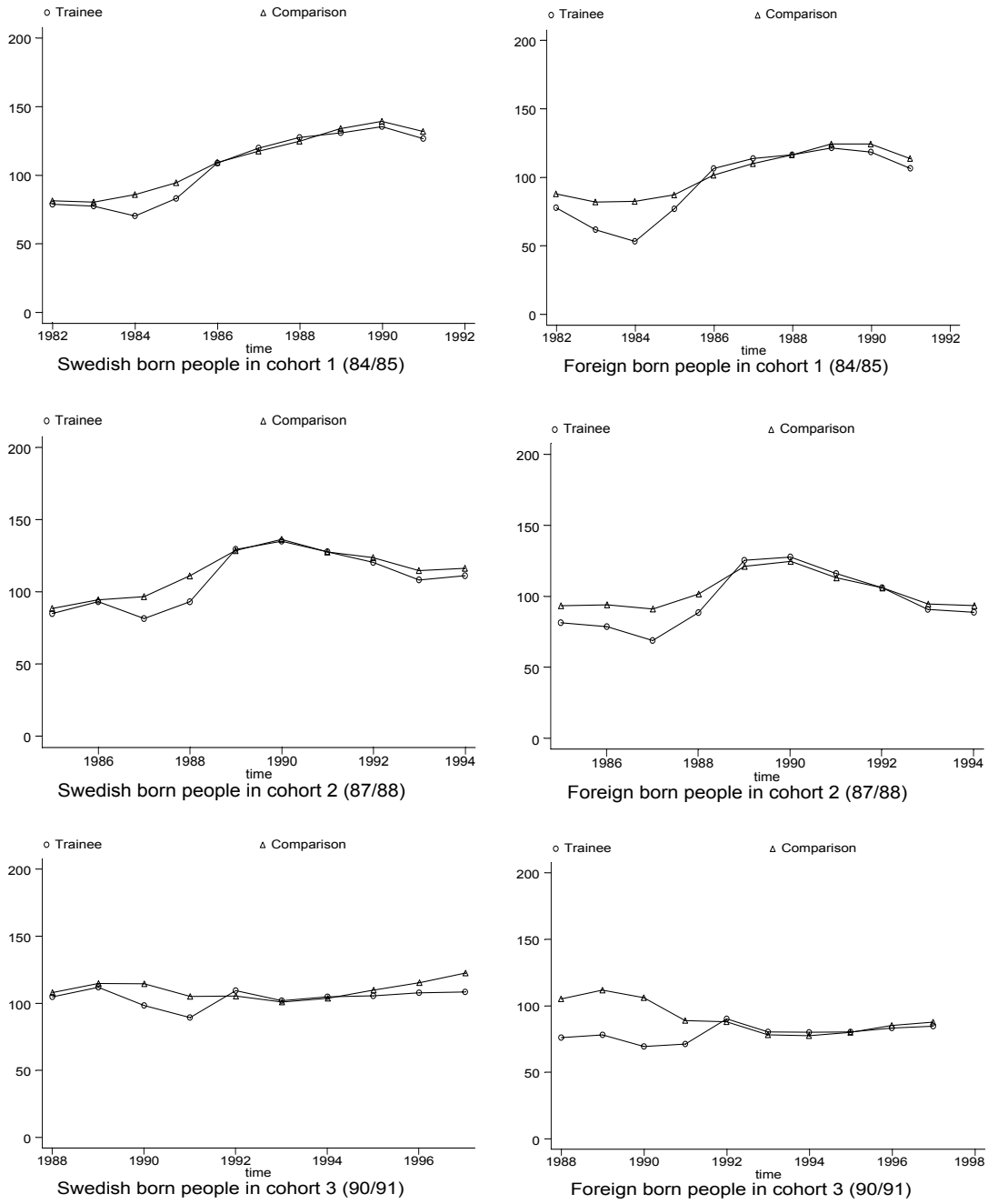


Figure 1 Earnings development over time and cohorts (earnings is deflated using 1999 as base year and is given in thousands of SEK)

5 Results

The estimates of the random coefficient model showing the treatment effect on earnings one year after the training are reported in Tables 4 and 5. Some comments can be made. Starting with the earnings equation, there is hardly a pattern of coefficients being large and estimated with small standard errors in any of the six samples. However, being male not surprisingly is associated with larger earnings among natives in the first two cohorts, and college education appears to yield substantially higher earnings in the third cohort for natives as well as foreign-born. One can also observe among immigrants that certain effects of origin are significant. For example in all cases, coefficients for variables measuring origin from the Middle East and "the rest of the world", respectively, (not in another Nordic country) are negative and estimated with a relatively small standard error.

Looking at the estimates of the reward equation, it can be noted that in many cases belonging to the youngest category (20-25) is associated with lower rewards than belonging to any other age category. In most cases the coefficient for the gender variable is not significant, with the third cohort natives having a positive sign for males as the exception. There is a pattern of education higher than primary school leading to larger rewards in most cases. Looking at variables specific to foreign-born, there is a pattern of a residency period longer than five years leading to larger rewards than a shorter residency. Further, rewards for originating from a country other than a Nordic country is generally positive and estimated with a relatively small standard error.

Table 4 Model estimates for the random coefficient model one year after training[8]

	Swedish-born			Foreign-born		
	Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3
Earnings						
Constant	4.431*	4.488*	4.360*	4.302*	4.507*	4.189*
	(0.028)	(0.030)	(0.057)	(0.074)	(0.069)	(0.083)
Age (26-45)	0.018	0.074*	0.005	-0.018	0.130*	0.025
	(0.029)	(0.030)	(0.060)	(0.037)	(0.041)	(0.058)
Age (46-55)	0.015	0.147*	0.019	0.071	0.128*	0.015
	(0.045)	(0.046)	(0.067)	(0.059)	(0.055)	(0.078)
Male	0.194*	0.185*	-0.050	0.126*	0.079*	-0.056
	(0.027)	(0.029)	(0.041)	(0.038)	(0.033)	(0.045)
High school	0.047	0.059	0.034	-0.020	0.022	-0.019
	(0.030)	(0.031)	(0.048)	(0.040)	(0.035)	(0.051)
College	0.100*	0.088*	0.251*	0.035	0.077	0.428*
	(0.051)	(0.048)	(0.064)	(0.079)	(0.062)	(0.071)
Reward						
Constant	-2.182*	-1.564*	-2.755*	-2.492*	-2.531*	-2.998*
	(0.131)	(0.135)	(0.414)	(0.147)	(0.154)	(0.213)
Age (26-45)	0.350*	0.021	0.459	0.346*	0.138*	0.400*
	(0.068)	(0.088)	(0.489)	(0.078)	(0.071)	(0.093)
Age (46-55)	0.687*	0.112	0.515*	0.267	0.141	0.512*
	(0.148)	(0.130)	(0.218)	(0.141)	(0.134)	(0.201)
Male	0.003	0.101	0.405*	0.103	0.128	-0.131
	(0.094)	(0.086)	(0.147)	(0.093)	(0.086)	(0.126)
High school	0.789*	0.599*	0.763*	0.749*	0.681*	1.073*
	(0.083)	(0.078)	(0.111)	(0.077)	(0.070)	(0.129)
College	0.757*	0.661*	0.769*	0.773*	0.714*	0.190
	(0.128)	(0.107)	(0.199)	(0.161)	(0.120)	(0.257)
Cost						
Constant	0.830	-0.427	-0.293	0.526	0.334	0.989
	(0.640)	(0.534)	(0.550)	(1.599)	(0.629)	(1.154)
Age	0.128*	0.198*	0.243	0.182*	0.205*	0.167*
	(0.040)	(0.032)	(0.335)	(0.097)	(0.038)	(0.069)
Age^2	-0.001*	-0.002*	-0.003	-0.002	-0.002*	-0.001*
	(0.001)	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)
Male	0.995*	0.964*	0.719*	1.059*	0.890*	1.158*
	(0.077)	(0.071)	(0.123)	(0.083)	(0.079)	(0.118)
City	-0.025	0.014	0.032	-0.087	-0.060	0.235*
	(0.070)	(0.069)	(0.1074)	(0.066)	(0.061)	(0.089)
Unemployed	0.002*	0.004*	-0.0005	0.005*	0.005*	0.008*
last year (days)	(0.001)	(0.001)	(0.0014)	(0.0008)	(0.001)	(0.001)
Variance						
σ^2_ε	1.640*	1.400*	3.093*	2.919*	2.594*	6.115*
	(0.142)	(0.126)	(0.432)	(0.172)	(0.155)	(0.392)
σ^2_u	0.352*	0.395*	0.641*	0.656*	0.531*	1.002*
	(0.015)	(0.017)	(0.038)	(0.034)	(0.027)	(0.051)
$\sigma_{\varepsilon u}$	-0.233*	-0.324*	-0.502*	-0.714*	-0.611*	-1.306*
	(0.045)	(0.039)	(0.095)	(0.069)	(0.059)	(0.127)
Log-likelihood	-3713.78	-3774.10	-3899.41	-5421.69	-5274.48	-6118.63
L-L No Cost [9]	-3791.11	-3857.58	-4022.17	-5558.68	-5410.24	-6210.72
Chi-Squared	154.66	166.96	245.52	273.98	271.52	184.18
L-L No Reward	-3771.44	-3779.79	-4027.08	-5522.81	-5382.85	-6184.29
Chi-Squared	115.32	11.38	255.34	202.24	216.74	131.32

Note: * significant at the 10% level. Standard errors are reported within parentheses

[8] The table presents the estimates for the first year after training for each cohort. The estimates for the consecutive years (i.e., the second and the third years) can be found Tables A2 and A3a-b in Appendix.

[9] L-L No Cost and L-L No Reward represent the log likelihood function value when estimating the model excluding observed heterogeneity in the cost and reward equation (except for a constant). The critical value on a 5% confidence level for a Chi-Squared distribution with 5 degrees of freedom is 11.07.

Table 5 Model estimate for random coefficient model one year after the training – extended variable specification for foreign-born[10]

	Cohort 1		Cohort 2		Cohort 3	
Variables	Parameter	Standard err.	Parameter	Standard err.	Parameter	Standard err.
Years in Sweden			Earnings			
6-10	-0.065	0.070	-0.114*	0.064	-0.092	0.074
11-	0.031	0.066	-0.020	0.058	-0.048	0.064
Origin						
Northern E.	-0.029	0.065	-0.168*	0.059	-0.120*	0.082
Eastern E.	-0.099	0.069	-0.267*	0.056	-0.224*	0.078
Southern E.	-0.287*	0.057	-0.306*	0.057	-0.061	0.079
Middle East	-0.130*	0.065	-0.402*	0.057	-0.417*	0.065
Other	-0.260*	0.075	-0.209*	0.058	-0.284*	0.070
Years in Sweden			Reward			
6-10	0.681*	0.117	0.578*	0.108	0.398*	0.148
11-	0.504*	0.104	0.631*	0.095	0.135	0.137
Origin						
Northern E.	0.569*	0.144	0.510*	0.144	0.674*	0.224
Eastern E.	0.714*	0.133	0.770*	0.113	0.915*	0.173
Southern E.	0.724*	0.117	0.855*	0.125	0.500*	0.208
Middle East	0.489*	0.127	0.774*	0.119	0.833*	0.149
Other	0.730*	0.125	0.719*	0.108	0.857*	0.152

Note: 0-5 years represents the reference category for years in Sweden, and Nordic countries represent the reference group for the origin.

In the cost equation the estimated age coefficients generally imply that costs increase with age, but at a decreasing pace. The cost for a male is always positive. With only one exception, the coefficient indicating the number of days in unemployment the year before training is positive and large in relation to its standard error. This is opposite of what we would expect, since it is more reasonable to think that a longer unemployment period would increase the probability of going into a program. What we see here might be a sign of cream skimming in the sense that those most likely to receive an employment after the training are selected into the program, with the believe that longer unemployment periods reduce the employment probability. The two variables that represent the exclusion restriction (city region and days of unemployment last year) have different effects on the selection process. Living in a city region representing the distance to the training center has no effect during good economic

[10] When estimating the models with the foreign-born we extend the variable specification in the earnings and the reward equation, because we believe that time in the country and place of origin are important determinants in the selection effect as well as in the outcome equation. The estimates belong to the results in Table 4 and are separated only to simplify up the presentation.

climates, while it has some effect during a recession for the foreign-born people in the third cohort.

The third cohort is somewhat different in structure compared to the other cohorts since Sweden was faced with a wave of immigrants that to some extent became a target population for the labor market program, since the immigrants' situation on the labor market was difficult. Groups of people among the foreign-born participated in programs that were of preparatory nature such as language courses, and most often the groups were clustered in city regions. That might be one reason for the positive and significant effect received for the third cohort. The second variable measured as number of days of unemployment the year before treatment is significant over the cohorts and groups.[11] The Swedish-born in the third cohort provides an exception.

At the bottom of Table 4 we present the log-likelihood values for specifications where observed heterogeneity with respect to cost and reward are disregarded and set to zero (except for a constant), one at a time. We observe that a likelihood ratio test would reject the null hypothesis (on a 5% significance level) that the observed cost or observed reward heterogeneity had no influence on the model. That is a justification of the statement that heterogeneity in rewards is important to control for. For the foreign-born group it is even more important, since they are more heterogeneous then the Swedish-born group.

The central interest of this study is on treatment-effects using earnings as the outcome variable; these are reported in Table 6 as well as illustrated in Figure 2. There are several findings to comment on. We have positive rewards for a majority (or nearly a majority) of the treated between the first two cohorts, as well as in some cases for the third cohort. Comparing results cross cohorts we find that foreign-born in the third cohort as measured shortly after training, clearly stand out. Only a small proportion of the treated have positive treatment effects one and two years after completed training. However, the proportion was slightly over 50% three years after completed training. The results thus clearly suggest that a deteriorating labor market worsens the prospects for trainees. This comes as no surprise and has been shown in administrative follow-up

[11] Number of days of unemployment the year before training is based on the amount of unemployment compensation (UI or CA) an individual received the year before training. Hence, the variable is an estimate that is based on what an average individual received in compensation per day, which therefore might be exposed to some bias.

studies (See Ds 2000:38, p 195). However, our results indicate that such an effect is limited to foreign-born trainees and to the first two years after training.

Table 6a Treatment on the treated effects for 1984/85-year cohort (standard deviation in parentheses)

Year	Mean effect/Swedish	P(Δ>0)*	Mean effect/foreign	P(Δ>0)
1986	-0.027 (0.188)	38	0.069 (0.313)	57
1987	0.101 (0.160)	74	0.262 (0.273)	83
1988	0.005 (0.172)	51	0.176 (0.316)	66

* Share with positive reward expressed in percentage.

Table 6b Treatment on the treated effects for 1987/88-year cohort (standard deviation in parentheses)

Year	Mean effect/Swedish	P(Δ>0)	Mean effect/foreign	P(Δ>0)
1989	0.084 (0.172)	68	0.062 (0.240)	58
1990	0.035 (0.161)	56	0.194 (0.213)	82
1991	0.129 (0.173)	80	0.323 (0.240)	93

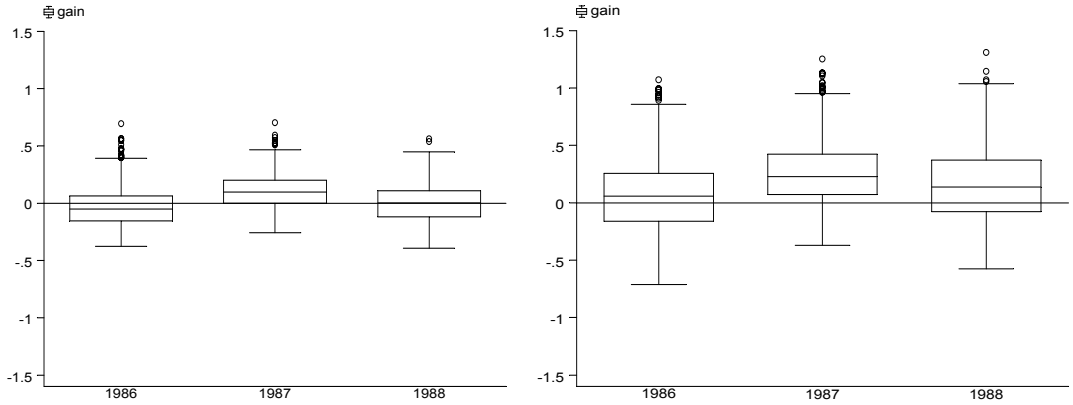
Table 6c Treatment on the treated effects for 1990/91-year cohort (standard deviation in parentheses)

Year	Mean effect/Swedish	P(Δ>0)	Mean effect/foreign	P(Δ>0)
1992	0.016 (0.254)	57	-0.485 (0.479)	18
1993	-0.003 (0.300)	50	-0.472 (0.321)	6
1994	-0.087 (0.292)	39	0.031 (0.350)	59

The structure and composition of the foreign-born group changes dramatically during the beginning of the 90's. Sweden received a large group of immigrants that did not speak Swedish. The labor market programs offer two kinds of courses. The first is of a preparatory nature and the second is of a vocational nature. The relation between the two changes over time due to an increasing number of foreigners taking language

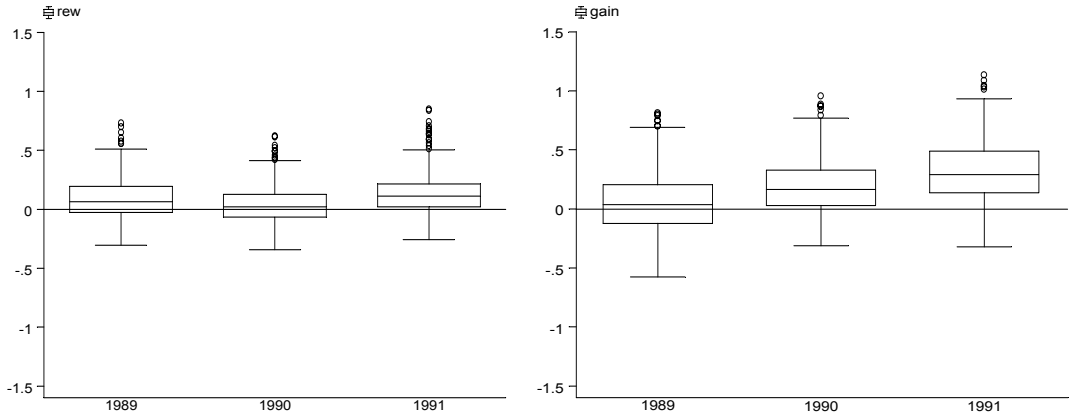
courses. In 1991 around 60% of the foreign-born trainees were in non-vocational courses not designed to increase the probability of employment but. Rather, to prepare for further training (Regnér, 1997). This obviously has some effect on the reward to training since the control group consists of unemployed individuals not taking part in training and, therefore, available to participate in labor market activities. We believe that is the major cause of the discrepancy between trainees and non-trainees for the foreign-born group in the third cohort in Figure 2.

As was conjectured from inspecting Figure 1, the estimated mean effect of training for foreign-born belonging to the first two cohorts is positive and in most cases larger than for natives. However, the standard errors are also large.



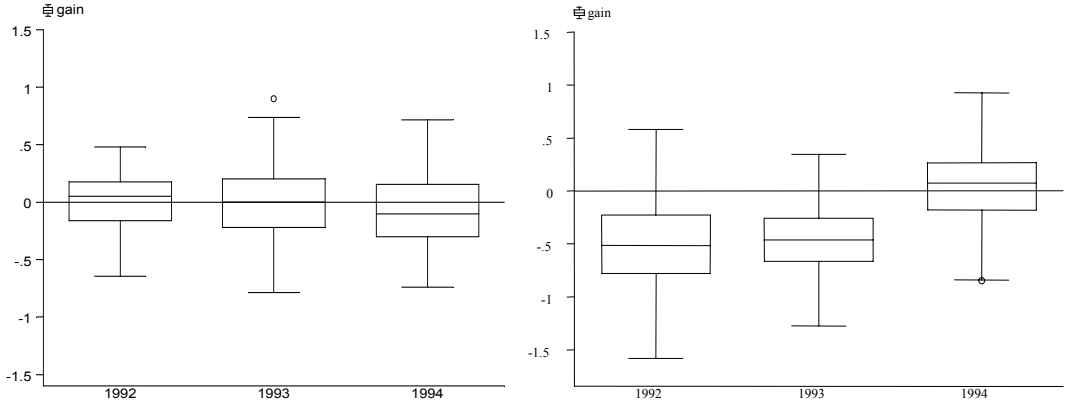
a) Swedish-born trainees b) Foreign-born trainees

Figure 2a Reward dispersion for 1984/1985-year cohort on log earnings[12]



a) Swedish-born trainees b) Foreign-born trainees

Figure 2b Reward dispersion for 1987/1988-year cohort on log earnings



a) Swedish-born trainees b) Foreign-born trainees

Figure 2c Reward dispersion for 1990/1991-year cohort on log earnings

[12] Box-plot explanation: the line in the middle of the box represents the median or 50th percentile of the data. The box extends from the 25th percentile to the 75th percentile, the so-called interquartile range. The lines emerging from the box extend to the upper and lower adjacent values which is defined as plus minus 1.5 times the interquartile range.

Table 7a Characteristics for the lower 25th percentile and upper 75th percentile of the reward distribution, for the 1984/1985-year cohort over the observation period for Swedish-born participants

	1986		1987		1988	
Variables	25th	75th	25th	75th	25th	75th
Age groups (%)						
20-25	50.8	13.0	56.0	15.3	28.7	32.3
26-45	49.2	43.1	43.2	52.4	71.3	37.9
46-55	0.0	43.9	0.8	32.3	0.0	29.8
Education (%)						
Primary	79.2	22.0	64.8	16.1	86.9	6.5
Secondary	16.7	67.5	33.6	62.1	13.1	49.2
Post Secondary	4.2	10.6	1.6	21.8	0.0	44.4
Gender (male) (%)	43.3	42.3	17.6	69.3	36.1	52.4
Married (%)	15.0	44.7	21.6	34.7	28.7	34.6
Number of children age 0-6 (%)	20.8	8.9	26.4	8.1	25.4	14.5
Number of children age 7-12(%)	11.7	11.4	13.6	12.1	18.9	12.1
Unemployed last year (%)	10.1	66.1	9.3	71.3	9.2	65.2
Number of days in training (days)	99.4	101.2	103.9	102.4	108.5	108.1

Table 7b Characteristics for the lower 25th percentile and upper 75th percentile of the reward distribution, for the 1984/1985-year cohort over the observation period for foreign-born participants

	1986		1987		1988	
Variables	25th	75th	25th	75th	25th	75th
Age groups (%)						
20-25	47.8	7.8	48.8	7.8	22.9	19.1
26-45	45.7	77.0	41.1	79.6	68.6	65.0
46-55	6.5	15.2	10.2	12.7	8.6	15.6
Education (%)						
Primary	88.2	15.2	79.2	22.9	90.6	11.0
Secondary	11.4	76.6	14.2	73.1	2.4	84.9
Post Secondary	0.4	8.2	6.5	4.1	6.9	4.1
Gender (male)	28.2	63.9	31.3	62.4	46.9	53.3
Married (%)	46.1	58.6	45.5	60.0	45.7	56.5
Unemployed last year (%)	3.9	60.7	2.88	63.1	1.7	59.5
Number of days in training	124.3	142.9	119.1	146.5	132.5	149.7
0-5 years in the country (%)	45.3	11.4	48.7	8.9	56.3	8.5
6-10years in the country (%)	6.9	33.6	7.3	34.3	8.5	30.8
More then 10 years (%)	47.7	54.9	43.9	56.7	35.1	60.5
Nordic country (%)	54.7	20.9	55.7	18.7	54.3	17.8
Northern Europe (%)	6.9	9.0	6.9	8.5	8.5	10.1
Eastern Europe (%)	2.9	19.7	1.6	21.2	0.0	27.6
Southern Europe (%)	4.1	22.1	2.4	26.1	4.9	19.9
Middle East (%)	21.2	8.2	21.9	7.3	13.4	12.6
Rest of the world (%)	10.2	20.1	11.4	17.9	18.7	11.7

Table 7c Characteristics for the lower 25th percentile and upper 75th percentile of the reward distribution, for the 1987/1988-year cohort over the observation period for Swedish-born participants

	1989		1990		1991	
Variables	25th	75th	25th	75th	25th	75th
Age groups (%)						
20-25	13.3	79.2	24.0	70.9	38.2	48.1
26-45	71.8	17.7	66.7	21.3	53.9	37.9
46-55	14.8	3.1	9.3	7.6	7.8	13.9
Education (%)						
Primary	19.5	34.6	22.5	26.7	44.5	12.4
Secondary	64.1	50.7	68.9	49.6	39.8	75.9
Post Secondary	16.4	14.6	8.5	23.6	15.6	11.6
Gender (male) (%)	32.0	56.9	53.4	39.7	20.3	70.5
Married (%)	27.3	10.7	20.2	15.3	20.3	21.7
Number of children age 0-6 (%)	28.1	17.6	24.8	18.3	21.8	14.7
Number of children age 7-12(%)	17.2	4.6	10.1	8.4	14.8	13.9
Unemployed last year (%)	2.7	39.8	1.5	41.3	1.0	42.5
Number of days in training (days)	146.5	114.2	131.7	136.7	121.9	129.7

Table 7d Characteristics for the lower 25th percentile and upper 75th percentile of the reward distribution, for the 1987/1988-year cohort over the observation period for foreign-born participants

	1989		1990		1991	
Variables	25th	75th	25th	75th	25th	75th
Age groups (%)						
20-25	27.5	13.3	32.5	12.4	30.7	15.4
26-45	65.4	65.8	57.9	70.9	58.9	67.5
46-55	7.1	20.8	9.6	16.5	10.4	17.1
Education (%)						
Primary	72.5	15.8	61.6	13.2	71.3	5.0
Secondary	23.3	72.9	32.1	75.1	18.6	85.8
Post Secondary	4.2	11.3	6.3	11.6	9.9	9.2
Gender (male) (%)	38.3	67.9	40.0	70.1	40.2	70.4
Married (%)	40.0	47.5	35.8	49.7	34.0	48.7
Number of children age 0-6 (%)	32.5	20.8	33.7	25.7	28.2	24.1
Number of children age 7-12 (%)	18.3	20.4	16.6	22.4	19.5	23.7
Percentage unemployed last year (%)	2.2	41.0	2.2	41.2	2.9	37.9
Number of days in training (%)	115.6	158.2	122.6	153.2	115.8	175.8
0 –5 years in the country (%)	47.9	11.6	47.1	12.0	24.8	30.0
6-10years in the country (%)	14.2	27.1	12.9	29.8	20.7	23.7
More then 10 years (%)	37.9	61.2	40.0	58.1	54.3	46.2
Nordic country (%)	55.8	11.6	57.9	11.6	65.5	10.0
Northern Europe (%)	9.2	2.9	2.5	10.7	4.1	6.7
Eastern Europe (%)	6.3	20.8	6.5	19.5	10.4	14.5
Southern Europe (%)	1.3	24.2	2.9	17.8	2.9	17.1
Middle East (%)	13.7	20.8	9.5	25.3	6.2	31.6
Rest of the world (%)	13.7	19.6	20.8	14.9	10.7	20.0

Table 7e Characteristics for the lower 25th percentile and upper 75th percentile of the reward distribution, for the 1990/1991-year cohort over the observation period for foreign-born participants

	1992		1993		1994	
	25th	75th	25th	75th	25th	75th
Age groups (%)						
20-25	57.2	18.5	65.5	9.2	52.9	20.6
26-45	19.6	70.5	21.0	74.7	23.5	70.2
46-55	23.1	10.9	13.4	15.9	23.5	9.1
Education (%)						
Primary	82.1	0	57.1	2.5	70.5	0.8
Secondary	13.6	75.6	9.2	97.4	0.0	99.1
Post Secondary	4.3	24.3	33.6	0.0	29.4	0.0
Gender (male) (%)	15.4	85.7	20.1	64.7	19.3	73.5
Married (%)	30.7	33.6	20.1	35.3	27.7	31.4
Number of children age 0-6 (%)	18.8	21.8	17.6	19.3	18.4	18.1
Number of children age 7-12 (%)	9.4	14.2	6.7	19.3	6.7	15.7
Percentage unemployed last year (%)	9.0	11.6	3.0	18.2	2.3	18.7
Number of days in training (days)	107.7	132.1	115.7	124.3	120.1	126.6

Table 7f Characteristics for the lower 25th percentile and upper 75th percentile of the reward distribution, for the 1990/1991-year cohort over the observation period for foreign-born participants

	1992		1993		1994	
	25th	75th	25th	75th	25th	75th
Age groups (%)						
20-25	40.2	13.6	35.5	5.7	27.2	7.3
26-45	46.5	69.5	53.7	78.5	55.7	84.8
46-55	13.3	16.8	10.7	15.7	16.9	7.7
Education (%)						
Primary	80.1	6.9	61.9	27.6	51.6	34.8
Secondary	2.9	92.5	4.5	72.3	13.6	65.2
Post Secondary	17.0	0.4	33.4	0.0	34.7	0.0
Gender (male) (%)	43.9	57.2	51.6	46.2	37.1	57.3
Married (%)	36.5	61.3	42.1	61.9	42.1	64.3
Number of children age 0-6 (%)	20.7	36.2	26.4	37.6	22.7	37.7
Number of children age 7-12(%)	17.8	25.1	15.2	28.9	19.4	30.3
Unemployed last year (%)	4.0	16.3	2.0	20.4	3.3	17.2
Number of days in training (days)	128.8	157.2	157.1	140.3	136.7	159.7
0-5 years in the country (%)	31.5	50.6	57.4	29.3	32.2	53.2
6-10years in the country (%)	7.8	23.4	7.0	31.4	10.3	22.9
More then 10 years (%)	60.5	25.9	35.5	39.2	57.4	23.7
Nordic country (%)	59.7	5.3	47.9	4.9	69.4	0.0
Northern Europe (%)	4.5	6.2	3.7	4.9	2.5	0.8
Eastern Europe (%)	2.1	20.2	4.1	23.9	3.7	31.1
Southern Europe (%)	12.0	7.8	2.5	10.3	4.1	1.6
Middle East (%)	13.2	33.3	26.4	30.1	13.2	25.8
Rest of the world (%)	8.3	27.2	15.3	25.6	7.0	40.5

In order to summarize the information we have also constructed Table 8. The idea for this table is to count the number of the six samples defined by cohort and country of origin where there are consistent indications of a low respectively high position in the reward distribution. There is also one column for indications of the position in the reward distribution being non-conclusive. When discussing the results we will start from the information in Table 8 and when motivated, also refer to those in Table 7.

Table 8 Summary of results reported in Table 4

Subgroup	Consistent indications of a low position in the reward distribution	Not conclusive	Consistent indication of a high position in the reward distribution
		Number of Indications	
Age of the person			
20-25	5	0	1
26-45	1	2	3
46-55	0	5	1
Education of the person			
Primary	5	1	0
Secondary	0	1	5
Post secondary	2	3	1
Male	0	0	6
Married	0	2	4
Variables specific to native-born			
Number of children aged 0-6	1	2	0
Number of children Aged 7-12	0	3	0
Variables specific to foreign-born			
0-5 years in the country	2	1	0
6-10 years in the country	0	1	2
More than 10 years in the country	1	0	2
Originating from			
A Nordic country	3	0	0
Northern Europe	0	3	0
Eastern Europe	0	0	3
Southern Europe	0	1	2
Middle East	0	1	2
Rest of the world	0	2	1

Note: To be classified as having a consistently high (low) position it is required that the percentage in the 75th percentile differs from the one in the 25th percentile by on average 10 percentage units per year and that a difference of at least 10 percentage units is observed for no less than two years.

Starting with age, there are clear signs of the youngest persons being located in the low position of the reward distribution. The only exception to this is found among the second cohort of Swedish-born; persons who terminated training when the labor market prospects became more favorable. There is much less structure as regards the

location of the other two age groups. This finding can be understood from the background of relatively few trainees being found in the oldest age group.

The conclusion that rewards for young trainees are generally low can also be backed by results from several Swedish studies referred to in the introduction. Ackum (1991) for example, who studied young adults that received training at approximately the same time as our first cohort, drew a very similar conclusion. In addition, results from two studies on persons who received training at approximately the same time as our third cohort are comparable. The studies are Regnér (1997) and Larsson (2000), the latter focusing on young adults. It is interesting to note that Friedlander et al (1997), when summarizing a number of evaluations of labor market training programs in the United States, drew similar conclusions.

Turning to education, the pattern shows that a primary education also leads to a low position in the reward distribution, while the opposite is the case for secondary education. There is not much of a pattern cross the samples when it comes to the position of post-secondary education in the reward distribution. These findings lead to the unanswered question: What can explain why the pay-off from labor market training is higher for those with secondary education, while low for those with only primary education?

According to the findings summarized in Table 8, there is a general pattern of males having a higher position in the reward distribution than females. There is also a pattern, although not equally striking, that married trainees have a higher position in the reward distribution than other trainees. The result mentioned first can be regarded as a controversial finding as it is in conflict with what Regnér (1997) reports for approximately our third cohort.

Finally, we comment on variables specific for foreign-born. First, it seems as there is more of a pattern between country of origin and position in the reward distribution, than between years since immigration and position in the reward distribution. People originating from other Nordic countries are generally found in the lower position of the reward distribution, while people from other parts of the world are under-represented at the bottom of the reward distribution. The results for people originating from Eastern Europe are consistently found in the top of the reward distribution for all three cohorts.

Table 9 presents another way to examine how rewards vary with characteristics one year after completed training. For natives and foreign-born people of a given gender, we present mean and medians the year after completed training, disaggregated by education and age, respectively. Looking at the information in the different cells, the most striking information is that large negative values for foreign-born in the 1990/91 cohort appear in most cells. One can also notice that among natives in the two latter cohorts, the values for males are generally higher than for females.

Table 9a Heterogeneity to reward, treatment on the treated for 1984/85-year cohort (standard deviation in parentheses)

1986		Primary School		Secondary School		Post Secondary School	
	Gender	Mean	Median	Mean	Median	Mean	Median
Swedish	Male	-0.154	-0.204	0.029	-0.003	-0.014	-0.029
		(0.168)		(0.185)		(0.127)	
	Female	-0.124	-0.162	0.036	0.013	0.028	-0.008
		(0.196)		(0.154)		(0.145)	
Foreign	Male	-0.022	-0.021	0.313	0.261	0.335	0.241
		(0.250)		(0.275)		(0.251)	
	Female	-0.167	-0.183	0.153	0.144	0.194	0.147
		(0.232)		(0.246)		(0.244)	
		Age (20-25)		Age (26-45)		Age (46-55)	
	Gender	Mean	Median	Mean	Median	Mean	Median
Swedish	Male	-0.092	-0.114	-0.051	-0.044	0.390	0.393
		(0.115)		(0.147)		(0.137)	
	Female	-0.118	-0.115	-0.052	-0.023	0.329	0.351
		(0.122)		(0.136)		(0.112)	
Foreign	Male	-0.049	-0.029	0.212	0.191	0.277	0.245
		(0.266)		(0.298)		(0.288)	
	Female	-0.207	-0.231	0.043	0.034	0.049	0.060
		(0.256)		(0.279)		(0.218)	

Table 9b Heterogeneity to reward, treatment on the treated for 1987/88-year cohort (standard deviation in parentheses)

1989		Primary School		Secondary School		Post Secondary School	
	Gender	Mean	Median	Mean	Median	Mean	Median
Swedish	Male	0.145	0.136	0.111	0.085	0.152	0.118
		(0.173)		(0.186)		(0.147)	
	Female	0.089	0.056	0.028	-0.006	0.057	0.054
		(0.153)		(0.166)		(0.145)	
Foreign	Male	-0.035	-0.051	0.202	0.166	0.217	0.183
		(0.200)		(0.230)		(0.214)	
	Female	-0.108	-0.127	0.0810	0.059	0.073	0.049
		(0.179)		(0.2084)		(0.201)	
		Age (20-25)		Age(26-45)		Age (45-55)	
	Gender	Mean	Median	Mean	Median	Mean	Median
Swedish	Male	0.233	0.232	0.046	0.057	0.065	0.091
		(0.152)		(0.169)		(0.092)	
	Female	0.161	0.148	-0.024	-0.008	0.027	0.030
		(0.154)		(0.122)		(0.113)	
Foreign	Male	0.081	0.013	0.118	0.115	0.273	0.228
		(0.203)		(0.253)		(0.210)	
	Female	-0.030	-0.037	-0.008	-0.021	0.064	0.044
		(0.209)		(0.216)		(0.211)	

Table 9c Heterogeneity to reward, treatment on the treated for 1990/91-year cohort (standard deviation in parentheses)

1992		Primary School		Secondary School		Post Secondary School	
	Gender	Mean	Median	Mean	Median	Mean	Median
Swedish	Male	-0.036	-0.001	0.268	0.266	0.299	0.327
		(0.159)		(0.150)		(0.139)	
	Female	-0.281	-0.231	0.016	0.077	0.044	0.072
		(0.144)		(0.161)		(0.136)	
Foreign	Male	-0.721	-0.642	-0.002	0.106	-0.861	-0.769
		(0.291)		(0.363)		(0.357)	
	Female	-0.686	-0.612	-0.022	0.005	-0.834	-0.675
		(0.339)		(0.357)		(0.410)	
		Age (20-25)		Age(26-45)		Age (46-55)	
	Gender	Mean	Median	Mean	Median	Mean	Median
Swedish	Male	0.023	0.091	0.279	0.341	0.142	0.094
		(0.181)		(0.169)		(0.199)	
	Female	-0.305	-0.327	-0.007	0.044	-0.143	-0.175
		(0.174)		(0.160)		(0.197)	
Foreign	Male	-0.711	-0.791	-0.398	-0.520	-0.359	-0.370
		(0.418)		(0.468)		(0.539)	
	Female	-0.787	-0.802	-0.466	-0.485	-0.407	-0.301
		(0.502)		(0.451)		(0.389)	

In Table 10 the sorting components are presented for the cohorts and groups over the follow-up period. Two interesting components are $\sigma_{1\varepsilon}$ and $\sigma_{0\varepsilon}$. Those two covariances give you the size and direction of the selection into the two states. Since

$\sigma_{1\varepsilon}>0$ and $\sigma_{0\varepsilon}<0$ we have positive selection into both training and non-training, implying that individuals are rational in the sense that they make the choice on the basis of where they will perform best.

In general, the pattern of the components over time is the same for foreign-born and Swedish-born people. The covariance between the two states (σ_{10}) is an exception. For the foreign-born, the sign is negative while it is positive for the Swedish-born. This is the case for all three cohorts and is therefore a difference that is independent of the economic climate. The absolute magnitude changes however, but this is the case for both groups. A positive sign indicates that an individual who performs well in one state also will perform well in the other state, and from the discussion above we know that the state is chosen where the reward is highest. For the foreign-born individuals the situation is different with a negative sign. That implies that if they do relatively well in one state they perform relatively poorly in the other state. This implies that the relative importance of the program for foreign-born people is greater than for Swedes.

A test for the importance of the unobserved component of the reward would be a test of $\sigma_{1\varepsilon} = \sigma_{0\varepsilon} \Leftrightarrow \sigma_{\varepsilon+u,u}= \sigma_{\varepsilon,u}$. In the table, we see that they even have different signs, indicating that controlling for unobserved heterogeneity with respect to the reward is important. Since we know that the individual's ambition to participate in the program is a major factor, we know that we do not have access to all relevant variables for the selection process. This makes it even more important to control for such factors.

Table 10a Behavioral components for Swedish-born people (standard errors in parentheses)

	Estimated variances and covariance								
	Cohort 1			Cohort 2			Cohort 3		
	1986	1987	1988	1989	1990	1991	1992	1993	1994
σ^2_0	0.352	0.403	0.373	0.395	0.332	0.409	0.641	0.835	0.724
	(0.015)	(0.016)	(0.015)	(0.017)	(0.015)	(0.018)	(0.038)	(0.040)	(0.041)
σ^2_ε	1.640	1.543	1.601	1.400	1.359	1.381	3.093	3.336	3.249
	(0.142)	(0.156)	(0.167)	(0.126)	(0.126)	(0.124)	(0.432)	(0.348)	(0.347)
$\sigma_{0\varepsilon}$	-0.233	-0.251	-0.223	-0.324	-0.283	-0.328	-0.502	-0.653	-0.548
	(0.025)	(0.046)	(0.045)	(0.039)	(0.037)	(0.041)	(0.095)	(0.101)	(0.095)
Implied variance and covariances									
σ^2_1	1.526	1.444	1.528	1.147	1.124	1.134	2.729	2.865	2.876
$\sigma_{1\varepsilon}$	1.407	1.292	1.377	1.075	1.075	1.053	2.591	2.683	2.700
σ_{10}	0.119	0.152	0.151	0.071	0.049	0.081	0.138	0.182	0.176

Table 10b Behavioral components for foreign-born people (standard errors in parentheses)

Estimated variances and covariance									
	Cohort 1			Cohort 2			Cohort 3		
	1986	1987	1988	1989	1990	1991	1992	1993	1994
σ^2_0	0.656	0.609	0.531	0.531	0.594	0.759	1.002	1.073	1.378
	(0.034)	(0.034)	(0.030)	(0.027)	(0.033)	(0.040)	(0.051)	(0.055)	(0.073)
σ^2_ε	2.919	2.583	2.555	2.594	2.505	2.672	6.115	6.138	5.838
	(0.172)	(0.069)	(0.137)	(0.155)	(0.150)	(0.168)	(0.392)	(0.382)	(0.362)
$\sigma_{0\varepsilon}$	-0.714	-0.762	-0.687	-0.611	-0.677	-0.840	-1.306	-1.410	-1.779
	(0.069)	(0.063)	(0.058)	(0.059)	(0.061)	(0.072)	(0.127)	(0.131)	(0.142)
Implied variance and covariances									
σ^2_1	2.147	1.667	1.711	1.902	1.745	1.751	4.505	4.391	3.658
$\sigma_{1\varepsilon}$	2.205	1.820	1.868	1.983	1.828	1.832	4.809	4.728	4.059
σ_{10}	-0.058	-0.153	-0.156	-0.080	-0.082	-0.080	-0.303	-0.337	-0.400

6 Summary and Conclusions

In this paper, we have evaluated labor market training programs in Sweden using non-experimental methods. People who received training in 1984/85, 1987/88 and 1990/91 as well as a control groups were followed using register data. The main outcome variable was earnings as evaluated one, two and three years after completed training. Different samples for natives and foreign-born were investigated. We estimate a switching regression model while allowing for unobserved heterogeneity with respect to the reward on training. This allows us to investigate how the reward is distributed across observed characteristics and between individuals.

A number of interesting findings were found and a number of conclusions can be drawn from the study. First, when analyzing treatment effects for trainees and controls, they were found to greatly differ for all cohorts investigated as well as across natives and foreign-born. The difference is found not only when analyzing earnings one year after completed training, but also two and three years after completed training. The differences all mean that there is positive sorting into training.

Second, overall, the proportion having positive rewards from training as evaluated by earnings was not very different from the proportion having negative rewards. The estimates for average rewards from training were in some cases relatively large, but so were the standard errors for the estimate. These results are in line with what was found in earlier studies of training, studies that took place in Sweden during approximately the

same period (Zetterberg, 1997). This means that the results from our study do not support the view, which suggests that from efficiency considerations, too few persons were enrolled in labor market training during this period.

Third, comparing results cross cohorts it was found that rewards stand out for the third foreign-born cohort, as most rewards were negative during the first two years following training. However, this changed during the following year. We interpret these findings as being driven by rapidly deteriorating labor market conditions. Thus it seems as though rapid changes in the labor market can drastically affect rewards, but also that such an influence is concentrated to the foreign-born and vanishes over time.

Fourth, when analyzing how rewards differ by characteristics across samples of the trainee, certain patterns were found. Consistent with several previous studies, we found that being a young adult means a negative or low pay-off from training. We also arrived at the same result for persons possessing only primary education. In conflict with what earlier studies have shown, we found that males have a better pay-off from training than females. Further, the results indicate that among immigrants, the pay-off from labor market training varies by origin. Thus the pay-off for a person from Eastern Europe was found to be better that for someone originating from another Nordic country.

Without additional knowledge it is difficult to offer a well-founded explanation for the finding that rewards were higher for foreign-born than for natives. One plausible explanation stems from the fact that natives and foreigners to some degree attended different training courses. Curriculum's for the courses differ and this might provide a viable explanation for the difference across the two groups. Another reason could be that training reduces the foreigner's reservation wage more than for natives, making them better prepared to accept job offers. A third explanation is that employers use a newly earned certificate for taking part in labor market training as a screening device when selecting foreign workers, but not when selecting native workers.

The estimates imply that we have positive sorting into both training and non-training for both Swedish-born and foreign-born individuals indicating rational behavior with respect to the participation decision. Nevertheless the sorting structure differs between the Swedish-born and the foreign-born. Swedish-born people have a

hierarchical structure while the foreign-born have a comparative advantage sorting structure. This difference is independent of the economic climate.

Although we believe our study has produced new insights into the effects of Swedish labor market training, there are also limitations worth mentioning. For policy purposes, the most important limitation is that we have treated programs as one homogenous category. In reality, programs differ -- by curriculum and length, for example. It is an important task for future research to investigate if and how rewards differ along such dimensions.

References

Ackum, S. (1991), ”Youth Unemployment, Labor Market Programs and SubsequentEarnings”, Scandinavian Journal of Economics 93, 531-43.

AMS (1995), “Den yrkesinriktade arbetsmarknadsutbildningens effekt på inkomst och sysselsättning”, Ubra 1995:1, National Labor Market Board Stockholm, Sweden.

Ashenfelter, O. (1978), “Estimating the Effect of Training Programs on Earnings”, Review of Economics and Statistics 60, 47-57.

Axelsson, R. (1989). ”Svensk Arbetsmarknadsutbildning – En kvantitativ analys av dess effekter”, Umeå Economic Studies No. 197, University of Umeå, Sweden.

Björklund, A. and R. Moffitt, (1987), “The Estimation of Wage Gain and Welfare Gains in Self-Selection Models”, The Review of Economics and Statistics 69, 42- 49.

Brännäs, K. and Eriksson, M (1996), "Enrolment in Labor Market Training Programs", Labor 10, 193-208.

Ds (2000:38), ”En effektivare arbetsmarknadsutbildning”, Ministry for Industry, Employment and Communications, Stockholm, Sweden.

Edin, P.A. (1988), “Individual Consequences of Plant Closure”, PhD thesis, Department of Economics, Uppsala University, Sweden.

Eriksson, M. (1997), “To Choose or Not to Choose: Choice and Choice Set Models”, Umeå Economic Studies No. 443. Umeå University, Sweden.

Friedlander, D., D. Greenberg, and P. Robins (1997), “Evaluating Government training Programs for the Economically Disadvantaged“, Journal of Economic Literature 35, 1809-1855.

Heckman, J. (1990), “Varieties of Selection Bias”. American Economic Review 80, 313-318.

Heckamn, J., and B. Honore (1990). “The Empirical Content of the Roy Model”, Econometrica 58, 1121-1149.

Heckman, J., J. LaLonde and J. Smith. (1998), “The Economics and Econometrics of Training Programs”, *Handbook of Labor Economics*, Volume III, ed. by Orley Ashenfelter and David Card. Amsterdam: North Holland.

Heckman, J. and R. Robb (1985), “Alternative Methods for Evaluating the Impact of Interventions, an Overview”. Journal of Econometrics 30, 239-267.

Heckman, J. and R. Robb (1986), "Alternative Methods for Solving the Problem of Selection Bias in Evaluating the Impact of Treatment on Outcome" in H. Wainer (ed.), *Drawing Inference from Self-selected Samples*, Berlin: Springer, 63-107.

Heckman, J. and J. Smith (1996), "Experimental and Non-experimental Evaluation" in *International Handbook of Labor Market Policy and Evaluation*, Ed. By Günther Schmid et al.

Heckman, J., and J. Smith (1999), "The Pre-program Earnings Dip and the Determinants of Participation in a Social Program: Implications for Simple Program Evaluations Strategies", Economic Journal 109, 313-348.

Larsson, L. (2000), "Evaluation of youth labor market programmes", Working paper 2000:1, Institute for Labor Market Policy Evaluation, Uppsala.

Melkersson, M. (1999), "Explaining Choice Set Size for Unemployed". Applied Economics 31, 1599-1607.

OECD (1997), *Employment Outlook, July 1997*, Paris.

OECD (2000), *Employment Outlook, June 2000*, Paris.

Okeke, S. (2001), "Arbetsmarknadsutbildningen 1999", URA 2001:7, National Labor Market Board, Stockholm.

Olli, Å. (1996), "Arbetsmarknadspolitiskt kalendarium", EFA, Arbetsmarknads departementet.

Quandt, R. E. (1972), "A New Approach to Estimating Switching Regressions", Journal of the American Statistical Association 67, 306-10.

Regnér, H. (1993), "Choosing Among Alternative Non Experimental Methods for Estimating the Impact of Training: New Swedish Evidence", Report no 8, Swedish Institute for Social Research, Stockholm University.

Regnér, H. (1997), "Training at the Job and Training for a New Job: Two Swedish Studies", Swedish Institute for Social Research – Dissertation Series. Stockholm University.

Roy, A. D. (1951), "Some Thoughts on the Distribution of Earnings", Oxford Economic Papers 33, 679-685.

Vella, F., and M. Verbeek (1998), "Estimating and Interpreting Models with Endogenous Treatment Effects", Working paper, Tilburg University.

Willis, J. R (1986), "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions", in *Handbook of Labor Economics* 1, Edited by O. Ashenfelter and R. Layard.

Zetterberg, J. (1997), “Effekter av arbetsmarknadspolitik – en översikt av svensk empirisk forskning”. Bilaga till Arbetsmarknadspolitiska kommittén.

Åtgärdsundersökning (1998), ”En uppföljning av deltagare som avslutat konjunkturberoende åtgärd fjärde kvartalet 1996”, Apra 1998:3, National Labor Market Board, Stockholm.

Appendix

A1 Defining trainees and counterfactuals

We have access to a register database that constitutes a stratified random sample of the population living in Sweden. It is stratified into two parts: the first is a 1% sample of the Swedish-born population and the second is a 10% sample of the foreign-born population. The stratified random sample was drawn by Statistics Sweden using population files from 1978. The individuals drawn at that initial year were followed over time with repeated yearly cross-sections. To each consecutive year a supplement of individuals were added to each cross-sectional unit to adjust for migration and newly born; the intention being to make each and every stratified cross-section representative for the Swedish population with respect to each stratum.

- The Treatment group

We analyse three different cohorts of trainees in this paper. The first cohort was drawn from the 1984 and 1985 cross-sections. Since we have access to data drawn from the total population of Sweden, the sample of trainees using only one cross-sectional year would be very small. Additionally, to be able to include trainees taking courses longer than a year, it was necessary to sample trainees from a two-year cross-sectional window by pooling two cross-sections. The sample of trainees for each cohort can therefore be classified into three groups. The first group consists of people who participated in a program the first year only, the second group of individuals who participated in a program the second year only, and the third group of individuals who participated in a program that started the first year and ended the second year. These individuals were controlled not to have participated in any labor-market training program three years before and three years after the two-year cross-sectional window, which we refer to as the training period. The two following cohorts were drawn in the exact same way but from different cross-sectional years namely 1987/88 and 1990/91.

The critical question when using population files is how to identify the trainees. From the files we have information about how large of a training grant an individual received for a given year. Training grants therefore function as a flag variable, indicating whether or not a person took part in training that particular year. Since this is our only way of identifying trainees, we have no information as to whether the trainee

actually completed the program. Dropouts might therefore be a source of bias in the estimates of the training effects. In order to reduce the training cohort from individuals who dropped out immediately or at the beginning of the program, we decided to truncate the sample with respect to the amount of training grant an individual received. We thought that a training grant corresponding to a **four-week period** would work as a lower truncation point. Since the official rules prescribe that only individuals aged 20 or older may participate in a program, we set the lower age limit to 20 and an upper arbitrary level at 55. The first cohort had no one older then 55 years of age, which therefore made us pick that upper age limit for all three cohorts. There are exceptions to the lower age limit, but we disregard them in this paper.

A problem with our flag-variable is that it contains two sorts of individuals. It includes individuals from labor market training programs but also individuals who have been participating in programs administrated by the Labor Market Institute (AMI). AMI is not a training program but individuals enrolled in AMI received the same kind of allowance, and since we identify the individuals from the grant received, we are not able to separate them. Fortunately, AMI individuals only comprise around 13% of the sample and the proportion stays constant during the observation window. The average schooling time for AMI participants is around 2-3 months, while a training program lasts on average around 5-6 months. This means that we reduced the AMI part proportionally more then the trainee part of the sample with the lower truncation point.

- The comparison group

The main idea is to construct a group that is comparable to the treatment group with respect to the characteristics of the trainees. A natural group to consider is that of individuals who were unemployed during the training period but who did not participate in any training program. When constructing the comparison groups for the three cohorts we use the same observation window as for the trainees. For the first cohort we take unemployed individuals from 1984 and 1985 cross-sectional years and pool them together to one set of non-trainees. 1987/88 and 1990/91 cross-sectional years were used in the same way for the other two cohorts.

Our dataset offers information about how much unemployment insurance (UI and/or CA) a person received in a given year, and this therefore works as an unemployment indicator when we select individuals into the set of non-trainees for the

three cohorts. Individuals participating in training usually have had a period of unemployment before a program may be an option. It is therefore important to have individuals with some length of unemployment in the non-treatment group. From a recent study by Okeke (2001), we know that the average wait before starting a training program is 3 months. This implies that around 50% of all participants have been unemployed a period shorter then 3 months.[13] By imposing a lower truncation point for the non-trainees we will be able to move the average unemployment duration forward and thereby construct a group that looks more like the trainees with respect to this point. We decided that **one month** would be an accurate number for this purpose and deleted those individuals with less then one month of unemployment.

Table A1 Compensation levels in SEK per day [14]

Benefit	1984	1987	1990
Average UI	239.86	307.28	402.07
Maximum UI	(300)	(360)	(495)
Cash Assistance	185	240	297

Source: Olli, 1996. (Nominal figures)

$$Compensated\ \ days\ \ per\ \ month = \frac{(Days\ \ per\ \ year - 2 * number\ \ of\ \ weeks\ \ per\ \ year)}{12} = 21,75$$

To estimate the duration of unemployment for an individual who was entitled to unemployment insurance, we divided the total amount of unemployment insurance a person received for a given year with the daily average unemployment insurance individuals in general received that particular year (see Table A1). That gives us an estimate of how many days that particular individual was unemployed that year. Dividing that number with the average number of days per month an individual could be compensated (21.75), we receive an approximation of the number of months an individual has been unemployed. If the individual was only entitled to cash assistance, we divided by that figure instead. Furthermore, we checked that the individuals did not participate in any training programs three years before and three years after the

[13] If the distribution of waiting times is skewed to the left a smaller proportion of trainees starts earlier then 3 months.

[14] Represents the average compensation amount that particular year. The unemployment insurance covers 80% of the previous income but only up to a max level. The max level is given within parentheses. Cash assistance is paid by a fixed amount. Compensation is paid 5 days per week.

observation window. We applied the same age interval as for the trainees, i.e. only individuals aged 20-55.

A2 The specification of the likelihood function

We define the choice sets in the following way:

$$D = \begin{cases} 1 & if\ \Theta_1 = (\Theta \mid \varepsilon > W\delta - Z\gamma) \\ 0 & if\ \Theta_0 = (\Theta \mid \varepsilon < W\delta - Z\gamma) \end{cases}$$

The likelihood function is:

$$L = \prod_{\Theta} [P_1]^D [P_0]^{1-D}$$

with

$$P_1 = \int_{\Theta_1} f(\varepsilon + u, \varepsilon > W\delta - Z\gamma) d\varepsilon = f(\varepsilon + u) \int_{\Theta_1} f(\varepsilon > W\delta - Z\gamma \mid \varepsilon + u) d\varepsilon$$

$$\left[1 - \Phi \left[\frac{(W\delta - Z\delta)/\sigma_\varepsilon - \left(\frac{\sigma_{\varepsilon\,\varepsilon+u}}{\sigma_\varepsilon \sigma_{\varepsilon+u}} \right) \frac{\varepsilon + u}{\sigma_{\varepsilon+u}}}{\left(1 - \left(\frac{\sigma_{\varepsilon\,\varepsilon+u}}{\sigma_\varepsilon \sigma_{\varepsilon+u}} \right)^2 \right)^{1/2}} \right] \right] \frac{1}{\sigma_{\varepsilon+u}} \phi \left(\frac{\varepsilon + u}{\sigma_{\varepsilon+u}} \right)$$

$$P_0 = \int_{\Theta_0} f(u, \varepsilon < W\delta - Z\gamma) d\varepsilon \quad = f(u) \int_{\Theta_0} f(\varepsilon < W\delta - Z\gamma \mid u) d\varepsilon$$

$$\Phi \left[\frac{(W\delta - Z\delta)/\sigma_\varepsilon - \left(\frac{\sigma_{\varepsilon u}}{\sigma_\varepsilon \sigma_u} \right) \frac{u}{\sigma_u}}{\left(1 - \left(\frac{\sigma_{\varepsilon u}}{\sigma_\varepsilon \sigma_u} \right)^2 \right)^{1/2}} \right] \frac{1}{\sigma_u} \phi \left(\frac{u}{\sigma_u} \right)$$

where Φ is the cumulative normal distribution function and ϕ the standard normal density function.

Table A2 Model estimates for random coefficient model year 2 and 3 after the training period. Swedish-born

	Cohort 1		Cohort 2		Cohort 3	
	1987	1988	1990	1991	1993	1994
Earnings						
Constant	4.5280*	4.5908*	4.5818*	4.5637*	4.2227*	4.3153*
	(0.0298)	(0.0295)	(0.0281)	(0.0307)	(0.0617)	(0.0586)
Age (26 - 45)	0.0017	0.0244	0.0908*	0.0785*	0.0058	0.0306
	(0.0316)	(0.0307)	(0.0291)	(0.0316)	(0.0515)	(0.0495)
Age (46 - 55)	0.0232	0.0104	0.1139*	0.0522	0.0928	0.1155*
	(0.0487)	(0.0465)	(0.0423)	(0.0475)	(0.0724)	(0.0688)
Male	0.1693*	0.2083*	0.2191*	0.1623*	-0.0380	-0.0160
	(0.0299)	(0.0288)	(0.0274)	(0.0304)	(0.0464)	(0.0440)
High school	-0.0004	-0.0185	0.0493*	-0.0232	0.0422	0.0212
	(0.0324)	(0.0314)	(0.0292)	(0.0324)	(0.0538)	(0.0515)
College	0.0260	0.0321	-0.0101	-0.0397	0.3289*	0.2598*
	(0.0552)	(0.0535)	(0.0454)	(0.0492)	(0.0716)	(0.0680)
Reward						
Constant	-2.0821*	-2.1199*	-1.6553*	-1.7427*	-2.7341*	-2.7782*
	(0.1564)	(0.1569)	(0.1367)	(0.1387)	(0.2266)	(0.2271)
Age (26 - 45)	0.3747*	0.2757*	0.0848	0.2037*	0.4890*	0.4101*
	(0.1051)	(0.1079)	(0.0959)	(0.0903)	(0.1476)	(0.1589)
Age (46 - 55)	0.5370*	0.4621*	0.2206*	0.3242*	0.5967*	0.4327*
	(0.1538)	(0.1561)	(0.1262)	(0.1308)	(0.1996)	(0.1969)
Male	0.1051	0.0216	0.0222	0.1309*	0.3959*	0.3583*
	(0.0949)	(0.0959)	(0.0833)	(0.0866)	(0.1362)	(0.1347)
High school	0.7489*	0.8039*	0.6498*	0.7615*	0.8027*	0.8747*
	(0.0865)	(0.0858)	(0.0750)	(0.0803)	(0.1181)	(0.1165)
College	0.7983*	0.9614*	0.7351*	0.7504*	0.3381*	0.4169*
	(0.1327)	(0.1318)	(0.1042)	(0.1091)	(0.2062)	(0.2301)
Cost						
Constant	0.6759	0.8940	-0.2711	-0.2853	-0.4798	-0.2956
	(0.6552)	(0.7426)	(0.7723)	(0.5378)	(0.8483)	(1.2287)
Age	0.1375*	0.1247*	0.1891*	0.1892*	0.2502*	0.2376*
	(0.0397)	(0.0451)	(0.0475)	(0.0328)	(0.0503)	(0.0740)
Age^2	-0.0016*	-0.0014*	-0.0023*	-0.0023*	-0.0032*	-0.0031*
	(0.0005)	(0.0006)	(0.0006)	(0.0004)	(0.0007)	(0.0010)
Male	1.0185*	1.0045*	0.9265*	0.9831*	0.6949*	0.7055*
	(0.0759)	(0.0771)	(0.0701)	(0.0712)	(0.1132)	(0.1128)
City	-0.0249	-0.0576	0.0281	0.0231	0.1237	0.0823
	(0.0703)	(0.0725)	(0.0675)	(0.0702)	(0.1023)	(0.1032)
Unemployed	0.0023*	0.0021*	0.0051*	0.0058*	0.0034*	0.0034*
	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0014)	(0.0013)
Variance						
σ^2_ε	1.5437*	1.6002*	1.3594*	1.3813*	3.3366*	3.2491*
	(0.1565)	(0.1677)	(0.1261)	(0.1241)	(0.3483)	(0.3471)
σ^2_u	0.4038*	0.3738*	0.3328*	0.4096*	0.8350*	0.7249*
	(0.0167)	(0.0157)	(0.0154)	(0.0185)	(0.0401)	(0.0413)
$\sigma_{\varepsilon u}$	-0.2516*	-0.2229*	-0.2837*	-0.3283*	-0.6530*	-0.5488*
	(0.0463)	(0.0455)	(0.0371)	(0.0417)	(0.1012)	(0.0952)
Log-likelihood	-3869.87	-3820.84	-3721.37	-4023.77	-4030.19	-4035.91

Note: * indicates significance at the 10% level. Asymptotic standard errors are within parentheses.

Table A3a Model estimates for random coefficient model year 2 and 3 after the training period. Foreign-born

	Cohort 1		Cohort 2		Cohort 3	
	1987	1988	1990	1991	1993	1994
Earnings						
Constant	4.5231*	4.5211*	4.4787*	4.3718*	4.0382*	3.9427*
	(0.0697)	(0.0674)	(0.0725)	(0.0816)	(0.0849)	(0.0967)
Age (26 - 45)	-0.0438*	0.0258	0.1195*	0.0749	-0.0211	-0.0932
	(0.0136)	(0.0348)	(0.0426)	(0.0516)	(0.0558)	(0.0719)
Age (46 - 55)	0.0451	0.0693	0.1187*	0.1395*	0.0892	-0.0371
	(0.0568)	(0.0532)	(0.0586)	(0.0656)	(0.0804)	(0.0891)
Male	0.0674*	0.1407*	0.0363	-0.0423	-0.0159	-0.0029
	(0.0366)	(0.0343)	(0.0360)	(0.0401)	(0.0469)	(0.0523)
High school	-0.0078	-0.0311	0.0341	-0.0289	0.0869	0.0340
	(0.0377)	(0.0365)	(0.0373)	(0.0410)	(0.0521)	(0.0572)
College	-0.0041	0.1109	0.0329	0.0418	0.3834*	0.4252*
	(0.0748)	(0.0706)	(0.0662)	(0.0716)	(0.0742)	(0.0801)
Reward						
Constant	-2.2877*	-2.3707*	-2.3296*	-2.1565*	-2.6978*	-2.1974*
	(0.0856)	(0.1235)	(0.1571)	(0.1881)	(0.1716)	(0.2310)
Age (26 - 45)	0.3916*	0.1748*	0.1481*	0.1819*	0.3109*	0.3084*
	(0.1671)	(0.0601)	(0.0655)	(0.1126)	(0.0980)	(0.1637)
Age (46 - 55)	0.2429*	0.1462	0.0982	0.1743	0.2201	0.1762
	(0.1235)	(0.1271)	(0.1342)	(0.1424)	(0.1998)	(0.1973)
Male	0.0435	-0.0486	0.1193	0.0738	-0.1608	-0.0422
	(0.0840)	(0.0835)	(0.0864)	(0.0909)	(0.1263)	(0.1235)
High school	0.6804*	0.8351*	0.6216*	0.7372*	0.6192*	0.5419*
	(0.0692)	(0.0672)	(0.0723)	(0.0775)	(0.1337)	(0.1287)
College	0.5009*	0.5444*	0.6539*	0.6679*	-0.1955	-0.2591
	(0.1488)	(0.1504)	(0.1222)	(0.1323)	(0.2711)	(0.2599)
Cost						
Constant	1.2051	1.1742	-0.0083	0.1480	0.9480	0.4736
	(2.3768)	(1.2857)	(0.7970)	(0.7858)	(1.9472)	(1.2212)
Age	0.1429	0.1491*	0.2248*	0.2181*	0.1663	0.1951*
	(0.1442)	(0.0783)	(0.0484)	(0.0462)	(0.1156)	(0.0721)
Age^2	-0.0015	-0.0016	-0.0026*	-0.0025*	-0.0017	-0.0021*
	(0.0020)	(0.0011)	(0.0006)	(0.0006)	(0.0016)	(0.0010)
Male	1.0948*	1.0403*	0.8824*	0.8947*	1.1618*	1.1820*
	(0.0784)	(0.0785)	(0.0787)	(0.0812)	(0.1177)	(0.1154)
City	-0.0450	-0.0653	-0.0327	-0.0781	0.3432*	0.1995*
	(0.0609)	(0.0602)	(0.0625)	(0.0664)	(0.0889)	(0.0899)
Unemployed	0.0051*	0.0053*	0.0058*	0.0065*	0.0069*	0.0094*
	(0.0008)	(0.0007)	(0.0007)	(0.0007)	(0.0014)	(0.0015)
Variance						
σ^2_ε	2.5833*	2.5559*	2.5055*	2.6724*	6.1383*	5.8381*
	(0.0692)	(0.1375)	(0.1506)	(0.1687)	(0.3820)	(0.3627)
σ^2_u	0.6093*	0.5313*	0.5948*	0.7596*	1.0732*	1.3787*
	(0.0347)	(0.0302)	(0.0333)	(0.0404)	(0.0557)	(0.0737)
$\sigma_{\varepsilon u}$	-0.7625*	-0.6877*	-0.6775*	-0.8402*	-1.4102*	-1.7790*
	(0.0633)	(0.0585)	(0.0613)	(0.0721)	(0.1313)	(0.1422)
Log-likelihood	-5363.4322	-5166.4741	-5492.8550	-5631.0375	-6169.0407	-6292.2393

Note: * indicates significance at the 10% level. Asymptotic standard errors are within parentheses.

Table A3b Extended variable specification of reward equation for foreign-born people

	Cohort 1 (84/85)		Cohort 2 (87/88)		Cohort 3 (90/91)	
Variables	1986	1987	1990	1991	1993	1994
Earnings						
Years in Sweden						
6 – 10	-0.1549*	-0.1295*	-0.1230*	-0.1035	-0.1020	0.0426
	(0.0664)	(0.0635)	(0.0680)	(0.0740)	(0.0761)	(0.0830)
11 –	-0.0885	-0.0803	0.0319	0.0681	-0.0996	0.0555
	(0.0634)	(0.0600)	(0.0616)	(0.0675)	(0.0667)	(0.0733)
Origin						
Northen E.	-0.0373	-0.0301	-0.0976	-0.0877	-0.0318	-0.0297
	(0.0604)	(0.0593)	(0.0629)	(0.0691)	(0.0848)	(0.0909)
Eastern E.	-0.1488*	-0.1324*	-0.1607*	-0.1919*	-0.2327	-0.4643*
	(0.0628)	(0.0597)	(0.0604)	(0.0654)	(0.8155)	(0.0896)
Southern E.	-0.1813*	-0.1248*	-0.2051*	-0.2853*	-0.0778	-0.2747*
	(0.0535)	(0.0502)	(0.0587)	(0.0657)	(0.8138)	(0.0896)
Middle East	-0.2706*	-0.2551*	-0.3739*	-0.2626*	-0.3261*	-0.4678*
	(0.0617)	(0.0608)	(0.0606)	(0.0656)	(0.0665)	(0.0728)
Other	-0.2020*	-0.1386*	-0.1556*	-0.1587*	-0.2736*	-0.4355*
	(0.0698)	(0.0667)	(0.0638)	(0.0670)	(0.0717)	(0.0791)
Reward						
Years in Sweden						
6 – 10	0.6876*	0.7759*	0.5911*	0.4129*	0.5067*	0.2941*
	(0.1042)	(0.1025)	(0.1116)	(0.1195)	(0.1490)	(0.1513)
11 –	0.5511*	0.6886*	0.5996*	0.4296*	0.4624*	0.2744*
	(0.0931)	(0.0901)	(0.0997)	(0.1073)	(0.1374)	(0.1395)
Origin						
Northen E.	0.5622*	0.5601*	0.6773*	0.6193*	0.5071*	0.6035*
	(0.1198)	(0.1243)	(0.1433)	(0.1508)	(0.2198)	(0.2168)
Eastern E.	0.7584*	0.9151*	0.7346*	0.6527*	0.7411*	0.9433*
	(0.1162)	(0.1126)	(0.1172)	(0.1259)	(0.1731)	(0.1757)
Southern E.	0.7441*	0.6785*	0.7277*	0.8100*	0.6158*	0.5587*
	(0.1071)	(0.1026)	(0.1248)	(0.1400)	(0.2041)	(0.2018)
Middle East	0.4992*	0.6817*	0.7748*	0.8116*	0.6556*	0.7221*
	(0.1127)	(0.1098)	(0.1141)	(0.1198)	(0.1448)	(0.1468)
Other	0.7073*	0.7422*	0.6216*	0.7052*	0.6581*	0.8604*
	(0.1126)	(0.1101)	(0.1140)	(0.1203)	(0.1506)	(0.1523)

Note: * indicates significance at the 10% level. Asymptotic standard errors are within parentheses.

Assessing the Employment Effects of Labor Market Training Programs in Sweden[♣]

Daniela Andrén and Thomas Andrén[α]

Abstract
Several studies have examined the effects of training programs on employment. Most of them assume that the effects of training are constant for all potential trainees. We use an econometric framework that allows studying the heterogeneous training effects on discrete outcomes. The treatment effect is allowed to vary depending on the trainee's observable and unobservable characteristics, and allows selection into training to be determined in part by the trainee's idiosyncratic treatment effect. Furthermore, we investigate the importance of the unobservables in the selection to training and how efficient the selection is with respect to the outcome. The results show small positive effects for the Swedish-born. The treatment on the treated is larger than the average treatment effect, indicating that the selection is stronger for the treated, and 40% of those treated gain by participating in training. Foreign-born have a negative effect from training the first year, with an average treatment effect larger than the treatment on the treated. From those who participated in training, only 11% experienced positive effect, while 38% were hurt by the training. The unobserved factors are important in the selection to training, as well as for the outcome. The effect of the selection is stronger for Swedish-born compared to foreign-born.

Keywords: labor market training, one factor model, selection, heterogeneous response, unobserved heterogeneity, treatment on the treated.

JEL classification: J31, J38.

♣ We would like to thank the Institute of Labor Market Policy Evaluation (IFAU) for the financial support.
α Göteborg University, Department of Economics. E-mail: Daniela.Andren@economics.gu.se, Thomas.Andren@economics.gu.se

1 Introduction

In the beginning of the 1990's Sweden experienced a huge unemployment shock, going from an open unemployment rate of 1.6% in 1990 to 10.3% in 1994. This dramatic change in the labor market placed massive pressure on policy makers, who were led to increase public spending on active labor market activities in decreasing open unemployment. Labor market programs have represented a huge investment for the government: over 3% of the GDP is spent on such measures. In 1994, the participation ratio reached its peak with on average 5.5% of the labor force participating in such programs. Despite the confidence placed in these measures, and their extensive use, there is still a shortage of knowledge about their effectiveness, and voices have been raised criticizing the usefulness of labor market programs in reducing unemployment (Calmfors, 2002).

Even though the number of studies using Swedish data increased rapidly during the 1990s, the Swedish literature on evaluation issues is still small compared to the US literature. US training programs mainly focus on increasing the productivity and earnings of low-income individuals. In contrast, the main purpose of training programs in Sweden is to prevent or reduce unemployment among low-skilled workers by increasing the participants' employment probabilities rather than their earnings. LaLonde (1995) and Heckman et al. (1999) point out that for the US, most of the gains in earnings from training stem from higher employment rates rather than from increased wages. Therefore, this study focuses on estimating the employment effects of training.

There is an increased interest in using matching estimators when determining the treatment effects of social programs [e.g. Larson (2000) and Sianesi (2001)]. The matching estimator solves the problem of creating a comparison group by matching individuals with the same observed characteristics. The drawback is the need of having access to all variables that determine the selection process. This requires that most of the unobserved factors that determine the selection to training are observed. This is a sensitive point since it is believed that unobserved factors such as aptitude and ambition play an important role in the selection to the program, but are not easily observed or approximated. Our choice of model is therefore a latent index sample selection model formulated by Aakvik et al. (2000). This model incorporates the selection process and

allows for unobserved factors to explain the outcome in each state, as well as in the selection process. The structure of the model also makes it easy to derive the mean and distributional treatment parameters, which are expected to shed light on how the treatment effects are distributed for different groups.

By having access to data during the 1993-1997 recession period of Sweden, this study aims to estimate the treatment effect of participating in a training program 1993-1994, on the individuals' employment probability for the next three consecutive years (1995-1997). We choose a model that allows us to study the heterogeneous treatment effect on discrete outcomes, and aim to answer the following questions: (1) What is the overall effect of training on employment probability? (2) How is the treatment response distributed across participants? (3) How important is it to control for unobservables in understanding the selection and outcome process?

The analysis is done separately for the Swedish-born and the foreign-born, since the two groups have different arrangements of characteristics, which determine the selection and treatment process. The group of foreign-born is also more heterogeneous compared to the Swedish-born group, which further emphasizes the importance of analyzing the groups separately. In general, the foreign-born group has a higher frequency of problems during a recession, and is therefore an important target group for labor market training. Moreover, in the beginning of the 1990s, Sweden had a relatively high inflow of immigrants from the South-East Europe that came as refugees. This implies a higher probability of participating in training since the status as political refugee makes a foreigner eligible for training courses during the first three years in Sweden.

The rest of the paper is organized in the following way: Section 2 presents the institutional setting and the main characteristics of the active labor market programs in Sweden for the analyzed period. Section 3 presents the data and main descriptive statistics for both treatment and control groups. The econometric specification is presented in Section 4, and the results in Section 5. Section 6 summarizes the findings of the paper.

2 Institutional setting

The extensive public involvement in training the unemployed in Sweden started in the beginning of the 1960s although it is possible to find earlier public programs of labor market training. Swedish labor market policy has two components: a (passive) benefit system that supports individuals while they are unemployed and various (active) labor market programs offered to improve the opportunities of unemployed workers.

The benefit system has two components: unemployment insurance (UI), and the cash labor market assistance (KAS).[15] UI is the most important form; it is income-related and is available for 60 calendar weeks. The daily compensation is 75% of the previous wages (was 90% before July 1993). A part-time unemployed person registered at a public employment office and actively searching for a job is also eligible for unemployment benefits. The requirements for receiving (full- or part-time) UI are the following: 1) the claimant must have paid the membership fees to the UI fund for at least 12 months prior to the claim; 2) the claimant must have been working for at least 5 months during the 12 months preceding the current spell of unemployment;[16] 3) the claimant must accept an offer of either a 'suitable' job or a labor market program. KAS was designed mainly for new entrants who are not members of any UI fund. This compensation is lower than UI, and in principle is paid for a maximum of 30 calendar weeks.

The public employment offices have a central role in assigning job seekers to training courses. The employment office is responsible for providing information on different courses, eligibility rules, training stipends etc. Those eligible for training are mainly unemployed persons who are job seekers and persons at risk of becoming unemployed. One can also be eligible for other reasons. For example, the status of political refugee makes a foreigner eligible for training courses during the first three years in Sweden. Although there is no formal rule for the offer of labor market training being given to a person who has been unemployed for a long period, there are reasons to

[15] We present the structure and rules of the system valid during 1993-1994, the period analyzed by this study.

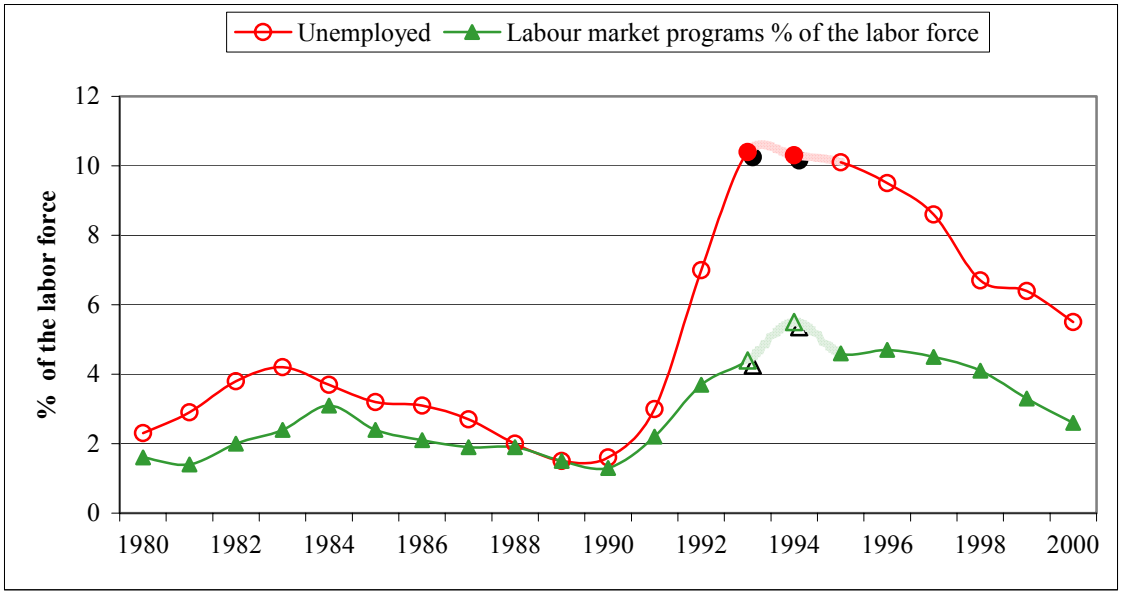
[16] Until 1996, a 5-month participation in practically any labor market program would count as employment in allowing participants to become eligible for the first time.

believe that this is often the case.[17] Since 1986, the time-period a trainee participates in a labor market program is considered equal to time spent on a regular job. Therefore, participation in a labor market program for 5 months counts as an *employment* spell, and thus qualifies for a renewed spell of unemployment compensation.

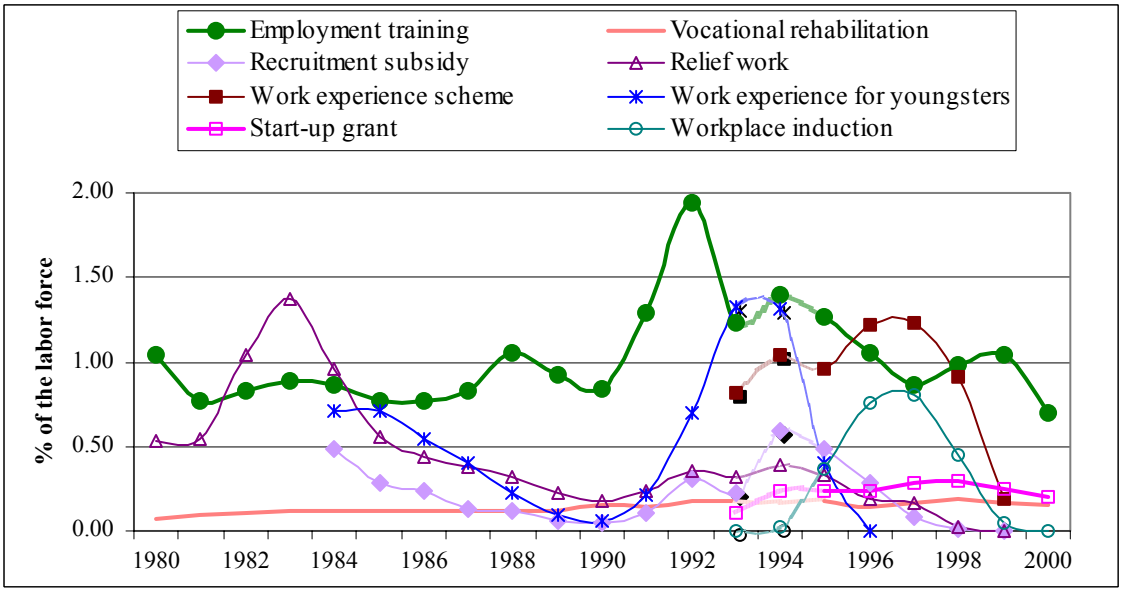
Originally, labor market training mainly consisted of vocational training programs, but over time, schemes comprised of more general educational training have grown more prevalent. During the 1990s, other education programs such as Swedish for immigrants, and computer training have been added to labor market training. There are many other types of publicly funded labor market programs. There are classroom courses as well as courses and activities that stress practical learning. A typical course is full time, five days a week, and last 6 month. Most courses are operated by the state (by AMU-centers), though nowadays the state competes with profit-oriented training organizations. The trainee might also follow courses in the regular school system. Individual firms can also arrange publicly funded training as an alternative to laying off personnel. For their maintenance, trainees receive a training stipend.

Figure 1a shows the unemployed and the participants in labor market programs as percent of the labor force, while Figure 1b shows this percentage by program type (selected categories). During the 1980s the percentage of trainees did not fluctuate very much, but seems to have followed the same trend as unemployment. The percentages coincide during the peak of the business cycle in the end of 1980s, after which the unemployment increased very rapidly. At the beginning of the 1990s, when the Swedish economy was brought to its deepest economic fall in more than 50 years, unemployment quickly reached the highest levels ever. However, the offer of labor market programs continued to expand during these years. The percentage of participants in labor market training decreased during 1993-1997 (i.e., during the recession period), though the offer of programs mainly oriented towards the disadvantaged groups (such as young people without previous experience, immigrants with or without previous work experience, and people in the older age groups) increased.

[17] As many unemployment spells are short a reasonable strategy for officials at labor market offices is to concentrate training offers on people with longer unemployment spells and others who can be assumed to have difficulties being employed without such efforts.



a) The unemployed and participants in labor market programs, % of the labor force



b) Participation in labor market programs, % of the labor force

Figure 1 The unemployed and participants in labor market programs, as percent of the labor force[18]

[18] Data source: National Labor Market Board (Historisk statistik 1980-2000; AMS Statistikenhet; Arbetsmarknadsstyrelsen 2001). The recession period analyzed in this paper is marked by the shaded symbols, which are linked by dashed lines.

3 Data

The data analyzed in this paper come from two longitudinal databases (SWIP and Händel) that have information on personal characteristics, earnings, incomes and unemployment history. SWIP (SWedish Income Panel) has two components: a sample of people that represents 1% of the Swedish-born population, and another sample that represents 10% of the foreign-born. SWIP is a database of *individual incomes*, built on a stratified random sample drawn (by Statistics Sweden) from the 1978 register of total population (RTB). The persons from *this* initial sample (about 77,000 Swedish-born and about 60,000 foreign-born) and the members of their households (the parents, the spouse, and the children) were followed over time using repeated yearly cross-sections. Additionally, each consecutive year (through 1999), a supplementary sample of individuals (varying between 3000 and 7000) and their household's members were added to each cross-sectional unit to adjust for migration in such a way as to make each stratified cross-section representative of the Swedish population with respect to each stratum. Income information is provided by the Swedish tax-register, which also includes information about those who do *not* pay income tax.

Händel is a register-based longitudinal event history database that contains information on *all* persons registered at the *public* employment offices. Its observation period starts in August 1991 and (in this paper) ends in December 1997. Händel has a *multiple spell* structure which provides exact information for the starting and ending dates of registered unemployment spells for each individual (with detailed information about the searching and program episodes that compose each spell). In addition to providing other information related to spells and episodes (e.g., the occupation unemployed people are looking for, the amount of desired labor supply, the location of a possible job, the reason for ending the registration spell, etc.), it provides information about personal characteristics of the job seekers (age, gender, citizenship, education, etc.). The main characteristics of this database are those components that allow us to identify the labor market trainees and counterfactuals.

From SWIP we select only individuals who were randomly selected (i.e., we left out all other members of the "household"), and match-merged this sample with Händel's database. We excluded all dropouts from the labor market training, and then

selected two groups of people: 1) those who participate in labor market training during the recession period, and 2) their unemployed peers (i.e., those who are unemployed during the same period but do not participate in such a program). Given the time horizon of both databases, we choose as *training* window the time-period from January 1993 through December 1994.

3.1 The construction of the treatment group

Given the available information and the previously mentioned design, we selected individuals who fulfill the following criteria: 1) they completed *one* training program (AMU) during 1993-1994; 2) they did not participate in any AMU program during 1991-1992 and 1995-1997; 3) they were 20-60 years old at the time the program started.[19] Applying these selection filters to Händel, a sample of 4,377 participants was obtained. After match-merging this sample with the SWIP database, the size decreases to 1,915 persons: 735 Swedish-born, and 1,180 foreign-born. Given the different representative selection with respect to their initial populations, and the different behavior of these two groups in the labor market, we will analyze them separately.

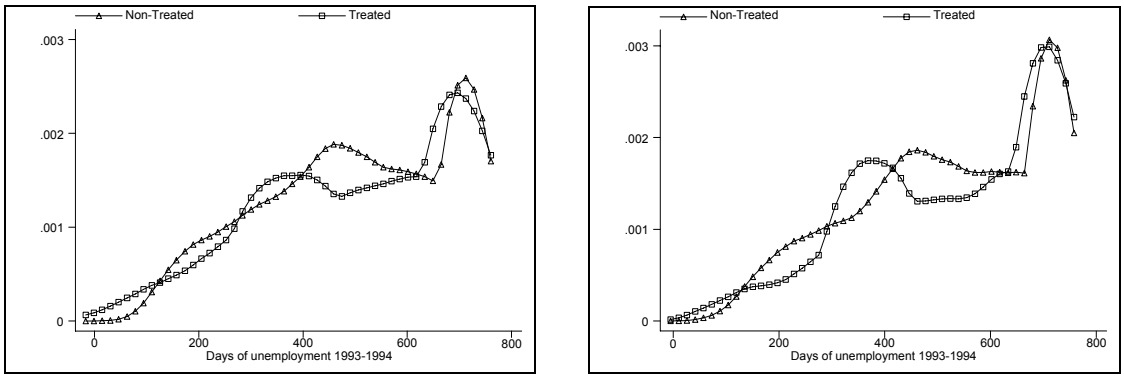
3.2 The construction of the comparison group

Given the available information and the selection criteria for the treatment group, we construct a comparison group, using the following filters: 1) they were unemployed *at least* 30 days in 1993 and *at least* 30 days in 1994;[20] 2) they did not participate in any AMU program during 1991-1997; 3) they were 20-60 years old at the time when the program started. After merging the sample of non-participants from Händel with the SWIP database, a sample of 8,771 persons was obtained: 3,681 Swedish-born, and 5,090 foreign-born. The first filter was imposed in order to harmonize the unemployment behavior between the treated and the untreated. The objective was to form two groups with comparable unemployment characteristics. Figure 2 shows the

[19] The age selection was done considering the following two aspects: 1) in general people are allowed to participate in a vocational training program if they are at least 20 years old: 2) we would like all individuals to be under the mandatory retirement age (65 years) in the last year (1997) of the analyzed period.

[20] This filter was designed in such way that there is a *minimum* unemployment period in both years, when people could qualify for starting a labor market program.

distribution of days of unemployment during the training period for trainees and non-trainees, and suggests that the groups have an acceptable correspondence in the distribution for both Swedish- and foreign-born. The number of days of unemployment considers the sum of the days for the two-year period.



a) Swedish-born b) Foreign-born

Figure 2 Estimated kernel densities for days of unemployment for treated and untreated

3.3 Comparing the treatment and comparison groups

Tables 1-3 present descriptive statistics of the treatment and comparison groups, stratified by country of birth into Swedish-born and foreign-born. Table 1 presents the demographic characteristics for these groups. A first conclusion is that there are slight differences between the groups of Swedish and foreign-born trainees (i.e., those who participated and completed training during 1993-1994), and between each of these groups and their unemployed ("non-trained") peers. Both trainees and non-trainees who were foreign-born were generally older than their native peers. This might reflect the great heterogeneity of the age at which people immigrated to Sweden and/or the age at which they entered the Swedish labor market. This difference might be partially sustained by the group of naturalized Swedes, which represents about half of both treatment and comparison groups of foreign-born. For all four groups, the proportion of men is slightly greater than that of women.

Table 1 Demographic characteristics (mean values),[21] year 1993

	Treatment group		Comparison group	
	Swedish-born n = 735	**Foreign-born** n = 1180	**Swedish-born** n = 3681	**Foreign-born** n = 5090
Women	0.45	0.44	0.47	0.46
Age	34.14	35.33	32.59	35.15
	(10.2)	(9.63)	(11.67)	(10.64)
Age groups				
19-25 years	0.24	0.18	0.39	0.22
26-45 years	0.59	0.65	0.44	0.60
46-60 years	0.17	0.17	0.17	0.18
Married	0.36	0.52	0.27	0.43
Region (counties-groups)				
Mid Sweden	0.38	0.46	0.45	0.53
South Sweden	0.25	0.24	0.23	0.20
West Sweden	0.22	0.23	0.20	0.22
North Sweden	0.14	0.07	0.11	0.05
Municipality groups				
Stockholm	0.11	0.21	0.16	0.32
Göteborg	0.08	0.10	0.09	0.13
Malmö	0.06	0.08	0.06	0.08
Other	0.75	0.60	0.69	0.47
Country of origin				
Nordic (excl. Sweden)		0.28		0.32
Western countries		0.07		0.08
East Europe		0.12		0.08
South Europe		0.09		0.08
Arab countries		0.16		0.17
Africa		0.15		0.12
Latin America		0.07		0.08
Asia and Oceania		0.06		0.07
Naturalized Swedes		0.45		0.56
Years in Sweden		8.21		10.03
		(7.79)		(7.4)
Years in Sweden-groups				
0- 5 years		0.59		0.46
6-10 years		0.13		0.18
> 11 years		0.28		0.36

Table 2 presents the mean figures for education, desired labor supply, the flexibility of accepting commuting, and unemployment duration by year for both treatment and control groups, for natives and foreign-born.

[21] Standard deviations are reported between parentheses only for quantitative variables. The rest of the variables are all dummies (taking value 1 for the mentioned category, and 0 otherwise). This holds true for all tables in this section.

Table 2 Education, desired labor supply and unemployment characteristics

	Treatment group		Comparison group	
	Swedish-born n = 735	**Foreign-born** n = 1180	**Swedish-born** n = 3681	**Foreign-born** n = 5090
Years from last degree	9.41	13.10	7.27	10.75
	(10.78)	(18.02)	(10.01)	(15.68)
Education groups				
Low	0.24	0.36	0.23	0.40
Medium	0.62	0.47	0.58	0.43
High	0.14	0.17	0.19	0.18
Labor supply (wanted job)				
Full-time, only	0.77	0.72	0.60	0.53
Part-time, only	0.05	0.04	0.06	0.05
Full-time or part-time	0.18	0.25	0.34	0.42
Inter-local applicant 1993-1994				
No	0.83	0.87	0.81	0.88
Yes	0.17	0.13	0.19	0.12
Days of unemployment by year			102.57	122.00
1990	1.11	1.81	2.38	5.15
	(11.15)	(17.01)	(20.56)	(31.13)
1991	39.05	39.82	43.52	47.58
	(75.76)	(78.58)	(77.54)	(82.39)
1992	127.20	126.71	121.88	121.92
	(137.53)	(140.58)	(133.77)	(137.43)
1993	247.21	248.37	244.21	249.88
	(131.89)	(135.47)	(111.83)	(111.08)
1994	261.49	283.61	257.70	268.04
	(130.46)	(118.20)	(107.13)	(105.36)
Share with employment				
1995	0.71	0.44	0.66	0.48
1996	0.69	0.46	0.62	0.44
1997	0.69	0.46	0.65	0.45

There are relatively big differences among the four groups regarding the educational background. While the groups of lower-educated native trainees and their unemployed peers are about the same (24%, and 23% respectively), their foreign-born peers are more highly represented in this education group; they are also slightly different from each other (36%, and 40% respectively). If the training were to cover some of the needs related to persons with lower education, we would expect that the foreign-born would have relatively higher rewards from training than natives. On the other hand, the proportion of highly educated trainees is lower than that of their unemployed peers, which does not suggest a straight expectation. It might be more difficult to find suitable training for the highly educated unemployed. If a suitable program were to exist, the reward would be expected to be greater for these people than their less-educated peers. The proportions of medium-educated native trainees and native non-trainees (62%, and 58% respectively) are much higher than their foreign-

born peers (47%, and 43% respectively), which is expected to show up in a higher reward from training if there is a demand for less qualified labor. These differences in human capital are expected to have effect on both selection into and return from training.

There are also differences between the treatment and comparison groups regarding the desired amount of labor supply. Trainees are more often looking for a full-time job compared with their unemployed peers, the figures being slightly higher for the natives than foreign-born for both treatment and comparison groups. On the other hand, the trainees are looking to a lesser extent for both part- and full-time jobs than their unemployed peers, which might imply that the latter group has higher labor supply flexibility than the former. Nevertheless, for all four groups, there is a small proportion (about 5%) of those who are able to accept only part-time jobs.

Looking for a job in another "local" labor market other than in the market of one's residence (i.e., a job which implies either daily, weekly, or monthly commuting) is another consideration for the four analyzed groups. Natives are on average more open to this alternative than the foreign-born, which might imply a higher probability for natives getting the job they are looking for.

We used the yearly days of unemployment as one of the selection filters when constructing the comparison. The filter characteristics were decided using the informational setting for selection into a training program ("some" days of unemployment before the training). Even though we did not use any matching approach regarding the annual days of unemployment, this indicator turned out to show almost the same figures for all (four) groups during 1992 and 1993. Nevertheless, in 1994, the trainees experienced on average more days of unemployment than their unemployed peers, but fewer days in 1990 and 1991.

Important variables in our analysis are the discrete dependent indicators for employment. We construct these variables using information from both Händel and SWIP databases. Händel provides information about both the date and employment status at the beginning and the end of the unemployment spell. Since a person might experience other states such as sickness absenteeism, parental leave, incarceration, etc, these information are not enough to compute the employment duration for a particular year. Therefore, we also use the variables on annual earnings from SWIP. Controlling

for both unemployment dates and employment status, persons were considered to be employed if their annual earnings were at least 40,000 SEK. [22] This was decided after analyzing the percentage of the employed by various ceiling levels, and corresponds to an average of around 3.5 months of full time work, which functions as a threshold level for being considered to be employed.

Table 3 presents the characteristics of training spells for natives and the foreign-born. The training spells for the foreign-born were about three weeks longer than natives' spells. More than half of the training spells took place during the "first" unemployment spell (since August 1991), but less than 1% started after the first visit to the unemployment office. It seems that natives visit the unemployment office more often before starting the training spell than the foreign-born, which is not unexpected if one assumes that the foreign-born might need skills specific to the Swedish labor market. Therefore, it seems obvious to offer them such training (language classes, in many cases) instead of letting them accumulate unemployment experience. About 37% of both natives and the foreign-born participated in training programs organized by AMU-centers. About 10% of the natives and 16% of the foreign-born participate in a municipal adult education program (KomVux).[23] The foreign-born participate more often in primary school education (7.18%), and in training programs organized by adult educational associations (6.5%) than natives do (2.04%, and 5.85%). The proportion of those who participate in vocational training is much higher for Swedish-born (about 73%) than for foreign-born (about 50%). The most frequent vocational training was oriented towards administrative work: 31% of Swedish-born, and 20% of foreign-born, participating in such training. Regarding the non-vocational programs, about 7% of foreign-born have "Swedish for immigrants" as training, and 11% participate in training that contains general or specialized courses.

[22] Assume that an individual has a wage rate of 50 SEK per hour. With an annual income of 40,000 SEK he or she would be working 800 hours per year, which roughly corresponds to 5 months of full-time work. If instead the wage rate were 100 SEK per hour, the corresponding figure would be 2.5 months of full-time work. We believe that the true number of full-time equivalence lies somewhere in-between these two numbers.

[23] Municipal adult education provides education corresponding to the last three forms of primary school and all secondary schooling. In addition, there are special vocational training courses.

Nevertheless, about 12% of the foreign-born trainees have reported work-related disabilities, compared with 8.98% of the natives, which might decrease the probability of securing a suitable job.

Table 3 Training spell's characteristics (mean values)

	Treatment group	
	Swedish-born n = 735	**Foreign-born** n = 1180
Training duration (days)	102.57	122.00
	(81.60)	(83.03)
The program started		
1993	0.44	0.45
1994	0.56	0.55
Unemployment spell *		
1	0.52	0.57
2	0.28	0.27
3	0.13	0.11
>4 (-6, -6)	0.07	0.05
Contacts with the unemployment office*		
1	0.01	0.02
2	0.18	0.23
3	0.20	0.21
4	0.19	0.18
5	0.13	0.14
6	0.10	0.09
7	0.08	0.06
>8 (-15, -13)	0.12	0.07
Course arranger		
AMU	0.37	0.37
Primary school (KomVux)	0.02	0.07
High school (KomVux)	0.08	0.09
High school	0.03	0.02
Adult educational association	0.06	0.07
Other	0.45	0.38
Vocational training		
Health care	0.04	0.04
Administrative	0.31	0.20
Commercial	0.05	0.03
Agriculture, foresting and fishing	0.02	0.01
Transport and communication	0.05	0.03
Manufacture	0.19	0.12
Services	0.07	0.06
Non-vocational training		
Primary school classes	0.02	0.07
High-school classes	0.02	0.02
General and specific courses	0.06	0.11
Swedish for immigrants (SFI)	0.00	0.07
Reported work handicap	0.09	0.12

* Both the (number of) unemployment spells and (number of) contacts with (or visits to) the unemployment office are reported with respect to August 1991 (when the observation period of Händel database started), and show when the training took place. The number between the parentheses shows the maximum number of spells for the natives, and foreign-born respectively.

4 Econometric specification

The fundamental issue of the evaluation problem is that one person is not able to be in two different labor market states at the same time. In the training context, for each trainee there is a hypothetical state of how he or she would have done without training. For each non-trainee, there is the hypothetical state of being a trainee. Let Y_1 be the potential outcome in the treated state, and Y_0 the potential outcome in the untreated state. The gain from going into the program is measured as the difference between the outcomes of the two states ($Y_1 - Y_0$). However, this difference cannot be formed for anyone since one or the other component of the difference is missing. The statistical approach to this problem replaces the missing data on persons using group means or some other group statistic. This does not solve the problem completely since the optimal difference would be $E[Y_1 \mid X, D=1] - E[Y_0 \mid X, D=1]$, in which the second expectation is unobserved and therefore has to be replaced with an approximation. It is typically replaced by $E[Y_0 \mid X, D=0]$, which in general is observed. The discrepancy between $E[Y_1 \mid X, D=1] - E[Y_0 \mid X, D=0]$ and $E[Y_1 - Y_0 \mid X, D=1]$ is the evaluation bias, $B(X)$, and the goal of any evaluation study is to diminish or eliminate this bias. The way we deal with this problem is to model the selection process and thereby reduce the bias using the index sufficient latent variable model (Heckman, 1979). Econometricians have distinguished structural or behavioral relations from conditional expectations and have used unobservable variables to make this distinction.

We postulate a standard framework of potential outcomes:

$$Y_1 = X\beta_1 - U_1 \tag{1}$$

$$Y_0 = X\beta_0 - U_0 \tag{2}$$

$$D^* = Z\beta_D - U_D \tag{3}$$

where X is a matrix of observed characteristics that explains the outcome of the two potential states. Each state also has an unobserved component represented by U_1 and U_0. In (3) we have the selection equation with D^* being a latent variable for the net gain from participating in training and Z a matrix of observed characteristics explaining the selection decision. When D^* is greater then zero the potential trainee chooses to

participate, while if it is negative he or she chooses to renounce. The observed counterpart of D^* is denoted D and takes the value 1 when D^* is positive and 0 otherwise. The standard assumption made for the functional form is a linear specification in the parameters with additive separation between the observed and unobserved components. This assumption has important implications on the structure of the evaluation bias. When the observed and unobserved components are separated additively the bias function turn out to be equal to $B(X) = E[U_0 \mid X, D=1] - E[U_0 \mid X, D=0]$.

The assumptions made about the state-specific unobservables are essential for the interpretation of the results and define a group of models. If $U_1 = U_0$, we obtain the dummy endogenous variable model of Heckman (1978). This assumption is very restrictive from a behavioral point of view, and needs to be relaxed if one would like to model heterogeneous response to training in terms of unobservables. If U_1 and U_0 are deterministically unrelated, with $E[U_1 \mid X] = 0$ and $E[U_0 \mid X] = 0$, we obtain the switching regression model of Goldfeld and Quandt (1972), which is much more flexible than the previous specification.[24] This version of the model solves the problem with unobserved counterfactuals by defining the dependent variable as $Y = DY_1 + (1-D)Y_0$. By substituting (1) and (2) into the expression we end up with the following relationship

$$Y = X\beta_0 + [X(\beta_1 - \beta_0) - (U_1 - U_0)]D - U_0 = X\beta + \alpha D - U \tag{4}$$

which clearly shows how the treatment parameter α is defined by a fixed and observed part, $X(\beta_1 - \beta_0)$, and an idiosyncratic part, $U_1 - U_0$, defined for each individual. This is also the random coefficient model of Heckman and Robb (1985).

4.1 Model with discrete outcome measure

The outcome measure in this paper is discrete and considers the employment probability after training. An important feature of any evaluation study is that of heterogeneous gain from treatment. It is unreasonable to believe that all individuals have one and the same response from the treatment given the observed characteristics. It is therefore important

[24] This model is also known as the Roy model [Roy, (1951), Heckman and Honoré (1990)].

to use a model that accounts for heterogeneous response to training. A second issue is that of unobserved factors being unaccounted for causing inconsistent estimates. Both problems may be taken into account if formulating a properly defined evaluation model. We therefore specify a discrete-choice, latent index model where the unobservables are generated by a normal one factor structure based on the framework discussed above and earlier formulated by Aakvik et al. (2000). We assume that the error terms in equation (1) – (3) are governed by the following one factor structure

$$\begin{aligned} U_1 &= -\rho_1 \xi + \varepsilon_1 \\ U_0 &= -\rho_0 \xi + \varepsilon_0 \\ U_D &= -\rho_D \xi + \varepsilon_D \end{aligned} \tag{5}$$

where ξ constitute the unobserved "ability" factor and ρ_i, $(i = 1,0,D)$, the factor loadings. By formulating the model in this way, we allow both for unobserved factors important for the selection process, and for heterogeneous response to treatment on unobservables.

The factor structure assumption for discrete choice models was introduced in Heckman (1981) and produces a flexible yet parsimonious specification, while making it possible to estimate the model in a tractable fashion. The following normality assumption is imposed: $(\xi, \varepsilon_1, \varepsilon_0, \varepsilon_D) \sim N(0, I)$, where I is the identity matrix, which implies that $(U_1, U_0, U_D) \sim N(0, \Sigma)$, with Σ having the following contents as a result of the one factor structure:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{10} & \sigma_{1D} \\ & \sigma_0^2 & \sigma_{0D} \\ & & \sigma_D^2 \end{bmatrix} = \begin{bmatrix} 1+\rho_1^2 & \rho_1\rho_0 & \rho_1\rho_D \\ & 1+\rho_0^2 & \rho_0\rho_D \\ & & 1+\rho_D^2 \end{bmatrix} \tag{6}$$

Conditioning on ξ, the likelihood function for the one-factor model has the form

$$L = \prod_{i=1}^{N} \int \Pr(D_i, Y_i \mid X_i, Z_i, \xi_i) dF(\xi_i) = \prod_{i=1}^{N} \int \Pr(D_i \mid Z_i, \xi_i) \Pr(Y_i \mid D_i, X_i, \xi_i) dF(\xi_i)$$

Since ξ is unobserved we need to integrate over its domain to account for its existence, assuming that $\xi \perp (X, Z)$. Since the probabilities in the likelihood function are conditioned on ξ, an unobserved factor essential for the selection to training, we have $(Y_1, Y_0) \perp (D \mid X, Z, \xi)$, which implies that $\Pr(Y_i \mid D_i, X_i, \xi_i) = \Pr(Y_i \mid X_i, \xi_i)$. This means that both the selection probability and the outcome probabilities are unconditional probabilities in the likelihood function.

We estimate the parameters using maximum likelihood technique, with a Gaussian quadrature to approximate the integrated likelihood.[25] Identification of the parameters of the model is insured by the joint normality assumption for the unobserved components of the model. The normalization and the joint normality imply that the joint distribution of (U_1, U_0, U_D) is known and given by (6), and that no exclusion restrictions are required.

4.2 Treatment parameters

Three parameters commonly estimated in the literature are the average treatment effect (ATE), the treatment on the treated (TT), and the marginal treatment effect (MTE). The last two parameters are modified versions of the first parameter. There are two ways of applying the parameters just mentioned: as mean treatment parameters, and as distributional treatment parameters. Both are of interest when evaluating effects from social programs.

The ATE answers the question of how much a randomly chosen individual from the population would gain from participating into training. This is a parameter of less interest since publicly funded training is seldom aimed at the total population but at a selected group with problems finding positions in the labor market. However, since it is commonly estimated in the literature we include it for comparative purposes. When the outcome variables are discrete and measure for employment, the probability of the events has to be formed and ATE is simply the difference in mean probabilities between the two states and across the individuals. In order to incorporate the unobserved factor it has to be integrated out. ATE may therefore be expressed in the following way:

[25] We use Gauss-Hermite quadrature to evaluate the integrals in the model, using 5 evaluation points. Points and nodes are taken from Judd (1998).

$$\text{ATE}(X,Z) = \int \left[\Phi(X\beta_1 + \rho_1\xi) - \Phi(X\beta_0 + \rho_0\xi)\right] dF(\xi)$$

Note that $\text{ATE}(X,Z)$ does not depend on Z, so that ATE(*X, Z*) = ATE(*X*). We choose to include *Z* to emphasise that the estimated values of β_1, β_0, ρ_1, and ρ_0 depend on *Z* since the selection equation is estimated jointly with the two outcome equations.

The TT parameter answers the question how much a person (who in fact participated in training) gained compared to the case where no training took place. TT is a modified version of ATE in the sense that it considers the conditional distribution of ξ. Hence, the employment probability of the two states has to be adjusted by the probability of being treated, incorporating the unobserved factors. The parameter is defined as:[26]

$$\text{TT}(X,Z,D=1) = \Phi\left(\frac{Z\beta_D}{\sigma_D}\right)^{-1} \int \left[\Phi(X\beta_1 + \rho_1\xi) - \Phi(X\beta_0 + \rho_0\xi)\right] \Phi(Z\beta_D + \rho_D\xi) dF(\xi)$$

The MTE parameter measures the treatment effect for individuals with a given value of U_D, i.e., the unobserved component of the selection equation.[27] The way the model is defined here induces that a lower value of U_D is associated with individuals that are more likely to participate and vice versa. The parameter is defined in the following way:

$$\text{MTE}(X, U_D = u) = \frac{\int \left[\Phi(X\beta_1 + \rho_1\xi) - \Phi(X\beta_0 + \rho_0\xi)\right] f_{\varepsilon_D}(u + \rho_D\xi) dF(\xi)}{f_{U_D}\left(\frac{u}{\sigma_D}\right)}$$

[26] $\text{TT}(X,Z) = \int \left[\Phi(Y_1 \mid X,\xi) - \Phi(Y_0 \mid X,\xi)\right] dF(\xi \mid D=1, X, Z)$ where $dF(\xi|D=1,X,Z)$ is the distribution of ξ conditional on D=1, *X*, *Z*. By Bayes' rule we have $dF(\xi \mid D=1,X,Z) = dF(\xi \mid D=1,Z) = \frac{\Phi(Z\beta_D + \rho_D\xi) dF(\xi)}{\Phi(Z\beta_D / \sigma_D)}$, which explains the expression given for TT(*X*,*Z*).

[27] This estimator was first introduced into the evaluation literature by Björklund and Moffitt (1987) in the context of Roy model, and developed by Heckman and Vytlacil (1999, 2000).

For many questions, knowledge of distributional parameters is required. Heckman (1992), Heckman and Smith with Clements (1997) and Heckman and Smith (1998) emphasize that many criteria for the evaluation of social programs require information on the distribution of the treatment effect. Does anyone benefit from the program? Among those treated, what percentage is helped by the program and what percentage is hurt by it? These are interesting questions that only can be answered by the distributional parameter. We will estimate the distributional parameters for TT. Before being able to state the expressions for the distributional parameter we need to define an indicator variable that identifies the parameter. Define $I = Y_1 - Y_0$ keeping in mind that Y_1 and Y_0 are binary. This gives us an indicator variable that takes three values (-1, 0, 1). I=1 is interpreted as a successful treatment in the sense that with training, the individual received employment (Y_1=1) while with no training, no employment would have been received (Y_0=0) (Analogous reasoning for the other values of I). With this in mind, we may define the distributional treatment parameter for TT in the following way:

$$\mathrm{TT}_{\mathrm{dist}}\left[1(I=1) \mid X, Z, D=1\right] = \frac{\int \Phi(Z\beta_D + \xi)\Phi(X\beta_1 + \rho_1\xi)(1 - \Phi(X\beta_0 + \rho_0\xi))dF(\xi)}{\Phi\left(\frac{Z\beta_D}{\sigma_D}\right)}$$

The distributional treatment parameter given above predicts the probability of the event that I=1. In order to receive the probabilities for the remaining values of I the expressions must be elaborated accordingly.

5 Results

This Section reports the results of the one factor model for 1995, i.e., one year after the training period.[28] Table 4 presents the parameter estimates of the normal one factor model, for the Swedish-born people. Although the goodness of fit for discrete choice models are in general fairly low, both Pseudo R^2 and McFadden R^2 indicate that the fit

[28] The model is also estimated for 1996 and 1997 and the estimates are presented in Tables A1-A4 in the Appendix. The estimates in the selection equations over the years do not differ very much, but there are changes in the employment equations in sign as well as in significance; mainly for the foreign-born people.

of the model is quite good, predicting probabilities that are 18-32% better than a model using only constants.[29]

Table 4 Parameter estimates of the one factor model for 1995, Swedish-born

Variables	Employment equation Treated			Employment equation Non-Treated			Selection equation		
	P.E.	S.E.	M.E.	P.E.	S.E.	M.E.	P.E.	S.E.	M.E.
Factor	0.217	0.132	0.081	-0.409	0.193	-0.142	2.039	0.261	0.205
Age	-	-	-	-	-	-	-0.044	0.005	-0.004
Age-groups (CG: 19-25)									
26-45 years	0.246	0.135	0.091	0.116	0.051	0.040	-	-	-
46-60 years	-0.220	0.177	-0.081	0.107	0.058	0.037	-	-	-
Education (CG: primary)									
High School	0.114	0.145	0.042	0.258	0.042	0.089	-0.993	0.160	-0.100
College	0.113	0.189	0.042	0.302	0.059	0.105	-1.243	0.208	-0.125
Has children[30]	0.198	0.111	0.073	0.210	0.059	0.073	0.488	0.121	0.049
Income 1992	-	-	-	-	-	-	0.206	0.085	0.021
City Region[31]	-	-	-	-	-	-	-0.552	0.144	-0.056
L-L model	-4781			σ^2_1	1.047		N (total)		4416
L-L constants	-5861			σ^2_0	1.167		N_1 (trainees)		735
L-L no factor	-4790			σ^2_D	5.157		N_0 (non-trainee)		3681
Chi-squared[32]	17.5			σ_{10}	-0.088				
Pseudo R^2	0.32			σ_{1D}	0.442				
McFadden R^2	0.18			σ_{0D}	-0.833				

Note: **Bolds** are significant at the 10% level, and CG means comparison group; P.E. means parameter estimate; S.E. means standard error; and M.E. means marginal effect.

The constants of the model are replaced by the factor loadings that are designed to capture the effect from unobserved heterogeneity such as aptitude or ambition. The factor loadings are significant in the two employment equations as well as in the selection equation, and a likelihood ratio test of including them in the model confirms their importance. Since the factor loadings define the covariances of the model, the sign of the factor loadings is of importance when determining the stochastic relationship between U_1, U_0 and U_D. The sign of the factor loadings in the two employment

[29] Both R^2 measures are based on a model estimated only with the factors of the models. That is since we do not have ordinary constants included in the model. Pseudo R^2 is a goodness of fit measure defined as $1 - 1/(1+2(\log L_1 - \log L_0)/N)$ with N being the number of observations used in the estimation. McFadden R^2 is defined as $1 - \log L_1/\log L_0$. Several alternative measures for goodness of fit for discrete choice models have been tested, and the conclusion is that different measures give different values, but no one smaller then McFadden R^2, though some even measure 0.4.

[30] This is a dummy variable indicating whether or not the individual has any children under age 18.

[31] City region is a dummy variable indicating if a person is living in one of the municipalities: Stockholm, Göteborg or Malmö.

[32] Chi-squared value generated by a likelihood ratio test statistic using the log-likelihood values from a model with and without the factor component. The critical value is 7.815 at the 5% significance level.

equations differs, indicating different sorting structures. The factor loading of the employment equation for the treated multiplied by the factor loading of the selection equation is defined as the covariance between U_1 and U_D. Since this covariance is positive, the selection to training is positive. That is, the employment probability is greater for the selected group of trainees compared to what it would have been if the selection to training had been random. The factor loading of the employment equation for the non-treated is negative, indicating that the selection to non-treatment is positive.[33] This implies that the employment probability is higher compared to what it would have been if the selection had been random.

The other estimated parameters of the employment equation for the non-treated are all significant, while only two estimates are significant for the treated: the age group of 26-45 years, and the dummy indicating the existence of children younger the 18. For the treated, people aged 26-45 have a better situation in the labor market compared to those aged 19-25. The estimates for the middle age group (26-45 years) are about the same for treated and non-treated, while the estimate for the oldest group (46-60 years) is significant only for the non-treated, suggesting that a person aged 46-60 was better off in the no-treatment state. Having children younger than 18 years has a significant effect, which is almost the same for both treated and non-treated. This might come from an increased responsibility of parents, motivating them to search harder for new jobs.

For the non-treated, high school and college education have a significant positive effect on the employment probability the *first* year after the training period, while for the treated these effects are not significant. This might suggest that the non-treated searched, or even accepted, jobs to a higher extent already when their treated peers still were participating in the programs. Even though training is aimed at people with low education, about 15% of the trainees have some sort of college education, which might indicate that their education did not pay off in the way it was intended. It might also be the case that unemployed with a college degree have a higher reservation wage compared to those with lower earlier education, and therefore reduce their employment opportunities. Another explanation is that being unemployed and participating in a

[33] Non-trainees have higher values of U_D, which corresponds to a lower probability to participate in training. Since σ_{0D} is negative, it follows that those individuals have lower values of U_0, which corresponds to an increased employment probability compared to what the employment probability would have been if the selection were random. This implies a positive selection to non-treatment.

training program might give negative signals for potential employers, thereby reducing the employment probability.

In the selection equation, all parameters are significant. Age has a negative influence on the probability of participating in training, even though the marginal effect is small. High school and college education, and living in a city region decrease the probability of participation in a training program. The pre-training annual earnings have a significant positive effect, suggesting that the higher the earnings a person had, the higher the probability for being selected into the training.[34] The marginal effect is quite low, though according to conventional standard the sign is reversed. The sign of the variable is stable with respect to variables' specification of the selection equation, several alternative specifications leading to the same results. The sign might be a result of the economic recession, in the sense that also highly productive people became unemployed, who therefore to a higher extent were selected into a training program. Given that earnings are related to skills, it might also be the case that the waiting time before the last day of employment and the first day of starting the training program to be shorter for those who had relatively high earnings.

Table 5 reports the parameter estimates of the one factor model for the foreign-born people. The goodness of fit for the model is comparable to the level for the Swedish-born people. The results indicate that the estimated model performs 18-33% better than estimating the model that contains only constants. The likelihood ratio test indicates that the unobserved factor has a significant effect on the performance of the model, suggesting that unobservables are important for foreign-born as well.

[34] The earnings for the year before training is a variable sometimes used as an instrument for the selection to training. Several studies have observed that the earnings of trainees decrease before the training period to a greater extent than for other individuals that are unemployed during the same period. This phenomenon is referred to as the Ashenfelter's dip (Ashenfelter, 1978). The pre-training earnings variable is therefore often used as an exclusion restriction in latent variable sample selection models.

Table 5 Parameter estimates of the one factor model for 1995, foreign-born

Variables	Employment equation Treated			Employment equation Non-Treated			Selection equation		
	P.E.	S.E.	M.E	P.E.	S.E.	M.E.	P.E.	S.E.	M.E.
Factor	0.251	0.095	0.085	0.406	0.096	0.141	2.537	0.021	0.209
Age	-	-	-	-	-	-	-0.038	0.002	-0.003
Age-groups (CG: 19-25)									
26-45 years	-0.075	0.111	-0.025	0.145	0.039	0.051	-	-	-
46-60 years	-0.161	0.151	-0.054	0.032	0.049	0.011	-	-	-
Education (CG: primary)									
High School	-0.044	0.091	-0.015	0.292	0.037	0.102	-0.139	0.069	-0.011
College	0.072	0.129	0.024	0.356	0.051	0.124	-0.157	0.104	-0.012
Has children	0.095	0.087	0.032	0.177	0.041	0.061	-0.176	0.071	-0.014
Country of origin (CG: Nordic countries)									
East Europe	-0.242	0.180	-0.082	0.004	0.077	0.002	-0.126	0.146	-0.010
West Europe	-0.263	0.141	-0.089	-0.157	0.072	-0.055	0.570	0.121	0.047
South Europe	-0.361	0.159	-0.122	-0.067	0.074	-0.023	0.177	0.137	0.014
Arab countries	-0.748	0.149	-0.253	-0.695	0.053	-0.243	-0.374	0.119	-0.031
Africa	-1.002	0.141	-0.339	-0.731	0.061	-0.255	0.061	0.114	0.005
Other nations	-0.531	0.161	-0.179	-0.226	0.055	-0.079	-0.427	0.118	-0.035
Years since immigration (CG: >11 years)									
0- 5 years	-	-	-	-	-	-	0.439	0.079	0.036
6-10 years	-	-	-	-	-	-	-0.515	0.116	-0.042
Income 1992	-	-	-	-	-	-	-0.331	0.045	-0.027
City Region	-	-	-	-	-	-	-1.101	0.066	-0.091
L-L model		-7136.87		σ^2_1	1.063		N (total)		6270
L-L constant		-8689.06		σ^2_0	1.164		N_1 (trainees)		1180
L-L no factor		-7156.04		σ^2_D	7.436		N_0 non-trainees		5090
Chi-squared		38.3		σ_{10}	0.102				
Pseudo R^2		0.33		σ_{1D}	0.636				
McFadden R^2		0.18		σ_{0D}	1.030				

Note: **Bolds** are significant at the 10 % level; CG means comparison group; P.E. means parameter estimate; S.E. means standard error; and M.E. means marginal effect.

The factor loading is positive and significant for all equations, but its magnitude differs between the states and the selection equations. The effect of the unobserved factor for the treated is almost half of the effect for the non-treated, which suggests that the unobservables have a higher effect on employment probability of the non-trainees compared to their treated peers. As discussed earlier, the sign of the factor loadings give important indications of the sorting structure of the unemployed into the two states. Since the factor loadings of both the selection and the employment equations for the non-treated are positive, the covariance between the unobservables of the two equations is positive. This is an indication of a negative sorting into the non-treatment state, which suggest that this group is worse off than the treated. However, the overall effect is a function of both the observed and the unobserved components.

Important variables when analyzing foreigners are the country of origin, and duration in the host country since immigration.[35] However, when it comes to employment probability, the number of years in the country had no effect. Therefore, these variables were excluded from the employment equations. The parameter estimates for country of origin suggest that immigrants born in a country outside Europe are a subgroup with particular problems. Except for the people from East Europe (who had an insignificant parameter), all groups of origin were worse-off than people born in the Nordic countries. Overall, the negative effects on employment were greater for those who participated in training as oppose to the non-trainees. The groups with the bigger negative effect were those from Arab and African countries. The rest of the observed characteristics have no significant effect on the employment probability for the trainees. Hence, for those who participated in training, country of origin was the major factor for the probability of receiving a job one year after the training period. For non-trainees, those aged 26-45, have a higher employment probability than their younger peers. There is also a positive effect of having a high school or college education, or having children younger than 18.

Most of the parameters in the selection equation are significant. The probability of participating in training decreases by age. However, its marginal effect is very low. Both high school education, and having children younger than 18 years, are also associated with a reduced probability of participating in training. The effect of country of origin differs: those born in an Arab country have a lower probability of participating in training, while those born in West Europe, South Europe and Africa have a higher probability (even though the effect for the Africans is not significant). This seems inefficient since the latter groups also lose more from participating in training. One explanation for those born in Africa or former Yugoslavia might be the fact that a high proportion of them came to Sweden as refugees in the 1990s, which implies that many of them take language courses as opposed to vocational training directly aimed at employment. Number of years in the country is also important in selection to training, with both dummies having significant parameters but with different signs. Compared with those who have been residents for more then 10 years, people who have been

[35] Edin and Åslund (2001) describe the labor market situation for foreign-born people in Sweden and find that the immigrants as a group have a weak position in the labor market, especially since large groups came to Sweden as refugees during the 1990's.

residents for less than 6 years, are more likely to enter a program, while those who have been residents 6-10 years are less likely to enter a program.

Both annual earnings before the training, and region of residence have a significant negative effect on the selection to training. In contrast to the Swedish-born people, the income effect suggests that those with lower earnings one year before participating in the program, began training to a higher extent. A gender dummy is not included for any of the groups since its effect was small and insignificant.

Tables 6 presents the mean marginal effects for the treatment on the treated for Swedish-born, while Table 7 presents the corresponding effects for foreign-born. For the Swedish-born, the unobserved factor has the largest effect, and it is positive all three years, suggesting that unobservables increases the employment probability. The effects for older people, the high school educated, and those with children are negative for 1995, and positive for 1996 and 1997. Having college education has a negative effect during 1995-1997, and its magnitude decreases by year, suggesting that the effect of higher education on differences between trainees and non-trainees decreases over the time, compared to their lower educated peers.

For foreign-born, all variables including the unobserved factor have negative effects on the gain from training in 1995, which implies that on average the treatment has a negative effect on the outcome. This situation changed with the time: 2 and 3 years after the training took place, half of the effects were positive. The effect of the unobserved factor is positive the following years, and increases over time. Country of origin was important in the first year after training, but its importance decreased over time, even though the negative effect for the category “other nations” was back on a high level after 3 years.

Table 6 Mean marginal effects on the treatment on the treated for Swedish-born[36]

Variables	1995	1996	1997
Factor	0.233	0.161	0.229
Age-groups (CG: 19-25)			
26-45 years	0.039	0.081	0.081
46-60 years	-0.117	0.138	0.030
Education (CG: primary)			
High School	-0.062	0.011	0.028
College	-0.079	-0.037	-0.035
Has children	-0.014	0.038	0.052

Table 7 Mean marginal effects on the treatment on the treated for foreign-born

Variables	1995	1996	1997
Factor	-0.031	0.063	0.071
Age-groups (CG: 19-25)			
26-45 years	-0.073	-0.006	-0.040
46-60 years	-0.071	0.022	0.036
Education (CG: primary)			
High School	-0.107	-0.051	-0.100
College	-0.083	-0.073	-0.076
Has children	-0.019	-0.011	0.038
Country of origin			
East Europe	-0.092	0.045	0.001
West Europe	-0.050	0.067	-0.041
South Europe	-0.114	-0.043	-0.057
Arab countries	-0.064	-0.014	-0.021
Africa	-0.149	-0.015	-0.026
Other	-0.129	-0.019	-0.101

5.1 Mean and distributional treatment effects

Table 8 reports the mean treatment effects based on the estimated parameters in the model. First year after the training, the ATE parameter is negative for both Swedish- and foreign-born people, the effect being larger for the foreign-born people, suggesting negative effect from training for a randomly chosen individual from the population. This estimate is in accordance with the literature on Swedish data that primarily reports negative or insignificant effects from training. This is not of special concern, ATE being a hypothetical parameter that is of less interest from a policy point of view since publicly funded training is seldom aimed at the total population but at a selected group with problems finding a job. Therefore, the TT parameter is of more interest, since the employment probability of the two states is adjusted by the probability of being treated.

[36] The mean marginal effect of the variables on the treatment of the treatment effect is defined as the partial derivative of TT(X, Z, D = 1) with respect to X, averaged over all individuals in the sample.

TT is positive for Swedish-born people but negative for the foreign-born. This negative effect was already suggested by the marginal effects from the variables explaining the treatment on the treated (Table 7). All marginal effects were negative only the first year, some of them becoming positive afterwards. Thus, we can conclude that training was to some extent beneficial even for foreign-born, but it took longer time. This might be related to the program type, which at least for those who immigrated recently contains mainly language courses.

Table 8 Mean treatment parameter estimates

Parameters	1995		1996		1997	
	Effect	Std-dev	Effect	Std-dev	Effect	Std-dev
Swedish-born						
ATE	-0.038	0.061	0.093	0.053	0.031	0.048
TT	0.181	0.023	0.193	0.024	0.162	0.019
MTE(U_D=4)	-0.391	0.075	-0.124	0,052	-0.227	0.087
MTE(U_D=0)	-0.038	0.076	0.093	0,065	0.031	0.068
MTE(U_D=-4)	0.320	0.064	0,334	0,065	0.375	0.076
Foreign-born						
ATE	-0.073	0.033	-0.038	0.043	-0.053	0.066
TT	-0.091	0.041	0.044	0.045	0.014	0.017
MTE (U_D=4)	0.114	0.111	-0.169	0.074	-0.198	0.106
MTE (U_D=0)	-0.073	0.118	-0.038	0.067	-0.053	0.101
MTE (U_D=-4)	-0.189	0.116	0.131	0.068	0.137	0.107

Since the estimated ATE is smaller than the estimated TT all three years, we conclude that for Swedish-born there is some indication that program administrators select individuals who benefit most from training than a randomly person in the population. For the foreign-born, the selection is negative the first year after the training, and slightly positive afterwards.

The MTE parameter in our case measures the average gain in outcomes for those individuals who are just indifferent to the receipt of treatment when the $Z\beta_D$ is fixed at the value u_D. Evaluating the MTE parameter at low values of u_D averages the outcome gain for those with unobservables making them most likely to participate, while evaluating the MTE parameter at high values of u_D averages the gain for those individuals with unobservables, which make them less likely to participate.[37] The MTE

[37] Recall that high values of u_D imply lower probabilities to participate in training since we have expressed the selection equation with minus in front of the unobservables, i.e., $D^* = Z\beta_D - U_D$.

parameter can be expressed as $\text{MTE}(X,u_D) = X(\beta_1 - \beta_0) - E(U_1 - U_0 \,|\, U_D = u_D)$ an alternative to the expression given in Section 4.2. It is important to treat U_1 and U_0 as deterministically unrelated since the difference between them represents the idiosyncratic gain for the individual. When $u_D = 0$, $\text{MTE} = \text{ATE}$ as a consequence of the symmetry of the normal distribution. We have used the $u_D = -4$ and $u_D = 4$, which are about ± 1.5 times the σ_D. For the Swedish-born, a high value of U_D corresponds to a negative effect of –39%, while a low value of U_D corresponds to a positive effect of 32%. The positive reward for those selected into training remains at the same level over time, while the negative effect for those less likely to participate is reduced over time. The foreign-born have a reversed situation in 1995: those with unobservables making them most likely to participate, gain the least, while the situation is the opposite for those less likely to participate. However, this extreme situation changes in subsequent years into the opposite.

The distributional treatment parameters capture an additional type of treatment effect heterogeneity beyond that of the mean treatment effect. Tables 9 reports the parameter estimates for the distributional version of the treatment on the treated parameter for Swedish- and foreign-born people.

Table 9 Distributional treatment estimates for treatment on the treated

Treatment estimates	1995	1996	1997
Swedish-born			
$E[1(Y_1=1,Y_0=0)\mid X=x, D=1]$	0.405	0.399	0.492
$E[1(Y_1=1,Y_0=1)\mid X=x, D=1]$	0.308	0.294	0.238
$E[1(Y_1=0,Y_0=0)\mid X=x, D=1]$	0.164	0.203	0.194
$E[1(Y_1=0,Y_0=1)\mid X=x, D=1]$	0.123	0.104	0.076
Foreign-born			
$E[1(Y_1=1,Y_0=0)\mid X=x, D=1]$	0.111	0.249	0.233
$E[1(Y_1=1,Y_0=1)\mid X=x, D=1]$	0.342	0.213	0.236
$E[1(Y_1=0,Y_0=0)\mid X=x, D=1]$	0.168	0.335	0.299
$E[1(Y_1=0,Y_0=1)\mid X=x, D=1]$	0.379	0.203	0.232

For the Swedish-born trainees, there is a 41% chance that the participant would benefit from the training, while there is a 12% chance that the trainee would be hurt by participating in the training program. The remaining 47% will neither gain nor lose, but will merely receive the same outcome in either state. The situation is slightly different for the foreign-born trainees, where only 11% of the participants would gain from the

training, while 38% would be harmed by the training. These figures confirm what we have expected from the analysis above. The proportion of individuals that would be hurt by training decreases for both groups over time, the figures being higher and the reduction over time smaller for the foreign-born group.

The measures of the same outcome in the two states are very similar between Swedish and foreign-born people. In 1995, for example, about 34% of the foreign-born would receive a job in any state, which implies that the participation in the labor market program only prolonged the process of receiving employment. This percentage decreases the second and third year after the training for both groups.

5.2 Selection on unobservables

An important question in any evaluation study is whether those most likely to participate in training are those who gain the most. In the previous subsection we reported that, in 1995, TT > ATE for Swedish-born trainees, while the situation was reversed for the foreign-born peers. For the Swedish-born, the selection improves on the gain for the group of those treated, and from the distributional treatment parameters, we also have indications that those who are most likely to go into training also gain most. For the foreign-born, the selection has the opposite effect the first year after the training, while the structure of the selection changed afterwards. Therefore, it is useful to take a closer look at the importance of the unobservables in the selection mechanism, especially at correlation measures, which are very informative. Table 12 reports the correlation among the indices in the model $(Z\beta_D, X\beta_0, X\beta_1)$. The indices are measured without taking into account the unobserved factor that we control for in the full model. For both Swedish and foreign-born, the state specific indices ($X\beta_0$ and $X\beta_1$) are positively correlated all three years after the training period. This implies that a person who does well in one state will also do well in the other state; and those who do poorly in one state, also do poorly in the other state. This situation seems to be much stronger for the foreign-born than for the Swedish-born.

The relationship between observable characteristics that predict participation and observable characteristics that predict employment in the participation state is relatively low, and differs between the groups. For Swedish-born, it is 0.216 in 1995, and decreases to 0.028 in 1997, while for foreign-born it is 0.036 in 1995, 0.097 in 1996,

and -0.015 in 1997. For Swedish-born, the relationship between observable characteristics that predict participation and observable characteristics that predict employment in the non-participation state is relatively low and negative all years, while for foreign-born its magnitude is even lower, but negative only in 1995.

Table 10 Correlations between indices without the unobserved factor

Correlations	1995	1996	1997
Swedish-born			
Corr($X\beta_1$, $X\beta_0$)	0.733	0.899	0.876
Corr($Z\beta_D$, $X\beta_1$)	0.216	0.172	0.028
Corr($Z\beta_D$, $X\beta_0$)	-0.209	-0.205	-0.093
Corr($Z\beta_D$, $X(\beta_1-\beta_0)$)	0.529	0.043	0.191
Foreign-born			
Corr($X\beta_1$, $X\beta_0$)	0.814	0.958	0.878
Corr($Z\beta_D$, $X\beta_1$)	0.036	0.097	-0.015
Corr($Z\beta_D$, $X\beta_0$)	-0.063	0.079	0.017
Corr($Z\beta_D$, $X(\beta_1-\beta_0)$)	0.164	0.039	-0.044

These results show that the Swedish-born people who are most likely to enter the training program are those who gain most from it the first and the second year after the training period. For foreign-born, even though the correlations' sign is the same as for the Swedish-born only in 1995, the effect of observed characteristics is much weaker for them all years. This shows once more that these two groups are different, and it might be the case that the same program works not the same for them.

Table 11 reports the estimated correlations among the unobservables.

Table 11 Correlations between the unobservables

Correlations	1995	1996	1997
Swedish-born			
Corr(U_1, U_0)	-0.079	-0.013	-0.026
Corr(U_D, U_1)	0.190	0.031	0.045
Corr(U_D, U_0)	-0.339	-0.338	-0.483
Corr(U_D, U_1 - U_0)	0.364	0.228	0.394
Foreign-born			
Corr(U_1, U_0)	0.092	-0.005	-0.004
Corr(U_D, U_1)	0.226	0.120	0.147
Corr(U_D, U_0)	0.350	-0.037	-0.025
Corr(U_D, U_1 - U_0)	-0.101	0.111	0.122

Overall, the levels of the correlations are quite small, and the signs are in accordance with what we discussed above. First year after the training, the level of the correlations is about the same for the Swedish-born and the foreign-born, but except for the correlation between selection and participation, the signs differ. For the Swedish-born, the unobservables support the state chosen so that those who enter training are better off there compared to the alternative.

The unobserved factors are important determinants of the outcome. This was shown earlier with the likelihood ratio test performed for each group testing if the factor contributed to the model. A way to further elucidate the importance of the unobserved factor is to determine correlation measures as in Table 10, while controlling for the unobserved factor. The results are presented in Table 12. The correlation between the two states is negative for the Swedish-born. This corresponds to a situation with comparative advantage in the sense that those with a high value in state 1 will have a corresponding low value in state 0. That is, on average those who perform relatively well with the training will perform relatively less well without the training.

Table 12 Correlations between indices with the unobserved factor

Correlations	1995	1996	1997
Swedish-born			
$Corr(X\beta_1-U_1, X\beta_0-U_0)$	-0.053	-0.028	-0.037
$Corr(Z\beta_D-U_D, X\beta_1-U_1)$	0.192	0.011	0.044
$Corr(Z\beta_D-U_D, X\beta_0-U_0)$	-0.332	-0.289	-0.461
$Corr(Z\beta_D-U_D, X(\beta_1-\beta_1)-(U_1-U_0)$	0.363	0.217	0.383
Foreign-born			
$Corr(X\beta_1-U_1, X\beta_0-U_0)$	0.163	0.122	0.132
$Corr(Z\beta_D-U_D, X\beta_1-U_1)$	0.208	0.117	0.131
$Corr(Z\beta_D-U_D, X\beta_0-U_0)$	0.309	-0.023	-0.021
$Corr(Z\beta_D-U_D, X(\beta_1-\beta_1)-(U_1-U_0))$	-0.087	0.106	0.113

For Swedish-born, the correlations' pattern is stable over time, while for foreign-born it is somewhat different. While the correlation between the two states is negative for the Swedish-born, it is positive for the foreign-born. This means that foreign-born who do well in one state also do well in the other state, which implies that if the performance is poor, they will do poorly in both states.

For both Swedish- and foreign-born, the correlation between selection and participation was positive all three years after the training, while the correlation between

selection and non-participation is negative in 1996 and 1997. In 1995, it was negative for Swedish-born, and positive for foreign-born. The positive correlation between selection and participation indicates that those who are most likely to enter training gain in doing so. According to the reported correlations, this is true both for observed and unobserved characteristics. In terms of observed characteristics, except for the 1997's value of the foreign-born, $\text{Corr}(Z\beta_D, X(\beta_1 - \beta_0))$ was positive all years, taking values between 0.039 and 0.529 (the levels being always higher for Swedish-born). In terms of unobserved characteristics, except for the 1995's value of the foreign-born, $\text{Corr}(U_D, U_1 - U_0)$ was positive all years, taking values between 0.111 and 0.394 (the levels being always higher for Swedish-born).

Table 13 reports the sorting gain from unobservables, i.e., $E[U_1 - U_0 \mid D = 1]$, suggesting a larger effect for the Swedish-born than for foreign-born. Except for 1995, when it was negative for foreign-born, the effect was positive all years for both groups. For the Swedish-born, the effect is stronger the first year after the training, while for the Swedish-born is the opposite.

Table 13 The sorting gain from unobservables

	1995	1996	1997
Swedish-born	0.219	0.100	0.131
Foreign-born	-0.018	0.082	0.067

6 Summary and conclusions

Using data that cover the 1993-1997 recession period of Sweden, this study estimated the treatment effect of participating in a training program 1993-1994, on the individuals' employment probability for the next three consecutive years (1995-1997). The analysis was done separately for the Swedish-born and the foreign-born, since the two groups have different characteristics, which determine the selection and treatment process.

Assuming a normal one-factor structure on the unobservables, we estimated a latent variable sample selection model that assesses the effect of training on the employment probability. Additionally, we investigated how the effect is distributed

across the participants, and explored the relationship between selection into training and the outcome.

For Swedish-born, the employment effect of labor market training is driven by being in the age bracket 26-45, having less education, no children and a heavy load of the unobserved factor. The predominant component is the unobserved factor that has a larger effect on the outcome then the other components. The ATE parameter is negative for the first year after training period, suggesting negative effect from training for a randomly chosen individual from the population. The TT parameter is stable and positive during the whole period, suggesting that the participation in training increases employment probability by around 18%. Moreover, TT > ATE the whole period, indicating that the selection to training is positive. The first year after the training, the distributional parameter suggests that about 40% of the trainees gain from treatment, while 12% are harmed by it. The proportion of those being hurt decreases over time.

For foreign-born, the employment effect of labor market training is driven by being in the age bracket 20-25, having less education, no children, and being from a Nordic country. The unobserved factor is not the predominant component, but its magnitude increases over time. However, the effect was negative first year, and positive afterwards. The first year after training, the ATE > TT, but the treatment from training turns positive afterwards. The distributional parameter for the treatment on the treated shows that first year after the training, 11% of trainees gain from treatment, while 38% are harmed by it. For the following years, the proportion of those who gain from training is almost double, while the proportion of those who are hurt is half.

For both Swedish-, and foreign-born, the state specific indices ($X\beta_0$ and $X\beta_1$) are positively correlated all three years after the training period. This implies that a person who does well in one state will also do well in the other state; and more importantly, those who do poorly in one state will also do poorly in the other state. This situation seems to be much stronger for the foreign-born than for the Swedish-born.

The relationship between *observable* characteristics that predict participation and *observable* characteristics that predict employment in the participation state is relatively low, and differs between the groups. In addition, for Swedish-born, the relationship between observable characteristics that predict participation and observable characteristics that predict employment in the non-participation state is relatively low

and negative all years, while for foreign-born its magnitude is even lower, but negative only in 1995. These results indicate that the Swedish-born people who are most likely to enter the training program are those who gain most from it. For foreign-born, even though the correlations' sign is the same as for the Swedish-born only in 1995, the effect of observed characteristics is much weaker for them all years. This shows, once more, that these two groups are different, and it might be the case that the same program works in different directions for them.

Overall, the levels of the correlations among the *unobservables* are quite small. While the correlation between the two states is negative for the Swedish-born, it is positive for the foreign-born. This means that foreign-born that do well in one state also do well in the other state, and if the performance is poor, they will do poorly in both states. For both Swedish- and foreign-born, the correlation between selection and participation was positive all three years after the training, while the correlation between selection and non-participation is negative in 1996 and 1997. In 1995, it was negative for Swedish-born, and positive for foreign-born. The positive correlation between selection and participation indicates that those who are most likely to enter training gain in doing so.

According to the reported positive correlations for both observed and unobserved characteristics, there is a *weak* indication that those most likely to participate in the training program are those who benefit the most from it. This evidence is much lower for Foreign-born than for Swedish-born.

References

Aakvik, A., J. Heckman, and E. Vytlacil (2000). "Treatment Effects for Discrete Outcomes When Responses to Treatment Vary Among Observationally Identical Persons: An Application to Norwegian Vocational Rehabilitation Programs", NBER, Technical Working Paper 262.

Ashenfelter, O. (1978). "Estimating the Effect of Training Programs on Earnings", Review of Economics and Statistics 60, 47-57.

Björklund, A., and R. Moffitt (1987). "The Estimation of Wage Gain and Welfare Gains in Self-Selection Models", Review of Economics and Statistics 69, 42-49.

Edin, P.A., and O. Åslund (2001). "Invandrare på 1990-talets Arbetsmarknad", IFAU, Forskningsrapport, 2001:7.

Goldfeld, S. and R. Quandt (1972). Non-linear Methods in Econometrics. Amsterdam: North Holland.

Heckman, J. (1978). "Dummy Endogenous Variables in A Simultaneous Equations System", Econometrica 46, 1931-60.

Heckman, J. (1981). "Statistical Models For Discrete Panel Data", in: C. Manski and D. McFadden, eds., Structural Analysis of Discrete Data with Econometric Applications. M.I.T. Press.

Heckman, J. (1997). "Instrumental Variables: A Study of Implicit Behavioural Assumptions Used in Making Program Evaluations", Journal of Human Resources, 32, 441-462

Heckman, J. and B. Honoré (1990). "Empirical Content of The Roy Model", Econometrica 58, 1121-49.

Heckman, J. and R. Robb. (1985). "Alternative Methods For Evaluating The Impact of Interventions", In Longitudinal Analysis of Labor Market Data, ed. James Heckman and Burton Singer, 156 – 245. New York: Wiley.

Heckman, J. and J. Smith (1998). "Evaluating the Welfare State", in Strom, S. (ed.), Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial, Econometric Society Monograph Series, Cambridge: Cambridge University Press.

Heckman, J., R. LaLonde., J. Smith (1999). "The Economics and Econometrics of Training Programs", in: O. Ashenfelter and D. Card., Eds., Handbook of Economics, Volume III, (North Holland, Amsterdam)

Heckman, J., J. Tobias, and E. Vytlacil (2000). “Simple Estimators for Treatment Parameters in a Latent Variable Framework with an Application to Estimating the Return to Schooling”, NBER, Working paper 7950.

Judd, L. K. (1998). Numerical Methods in Economics, The M.I.T. Press.

LaLonde, J. R. (1995). “The Promise of Public Sector-Sponsored Training Programs”, Journal of Economic Perspective 9, 149-168.

Roy, A.D. (1951). “Some Thoughts on the Distributions of Earnings”, Oxford Economic papers 3, 135-146.

Quandt, R. (1972). “Methods For Estimating Switching Regressions“, Journal of The American Statistical Association 67, 306-10.

Åtgärdsundersökning (1998). “En Uppföljning av Deltagare som Avslutat Konjunktureberoende Åtgärd Fjärde Kvartalet 1996”, Apra 1998:3, AMS, Enhet för arbetsmarknadsprogram, ISSN 1402-6813.

Appendix

Table A1 Parameter estimates of the one factor model for 1996, Swedish-born

Variables	Employment equation Treated			Employment equation Non-Treated			Selection equation		
	P.E.	S.E.	M.E.	P.E.	S.E.	M.E.	P.E.	S.E.	M.E.
Factor	0.035	0.133	0.011	-0.412	0.181	-0.139	1.931	0.268	0.207
Age	-	-	-	-	-	-	-0.043	0.005	-0.005
Age-groups (CG: 19-25)									
26-45 years	0.305	0.137	0.101	0.041	0.051	0.013	-	-	-
46-60 years	-0.376	0.177	-0.124	-0.708	0.075	-0.240	-	-	-
Education (CG: primary)									
High School	0.389	0.145	0.128	0.309	0.041	0.104	-0.943	0.147	-0.101
College	0.363	0.192	0.120	0.421	0.059	0.142	-1.208	0.199	-0.130
Has children[38]	0.438	0.112	0.144	0.277	0.058	0.094	0.488	0.117	0.052
Income 1992	-	-	-	-	-	-	0.220	0.079	0.023
City Region[39]	-	-	-	-	-		-0.478	0.131	-0.051
L-L model	-4727			σ^2_1	1.047		N (total)		4416
L-L constants	-5957			σ^2_0	1.167		N_1 (trainees)		735
L-L no factor	-4735			σ^2_D	5.157		N_0 (non-trainee)		3681
Chi-squared[40]	14.7			σ_{10}	-0.088				
Pseudo R^2	0.357			σ_{1D}	0.442				
McFadden R^2	0.206			σ_{0D}	-0.833				

Note: CG means comparison group; P.E. means parameter estimate; S.E. means standard error; and M.E. means marginal effect.

[38] This is a dummy variable indicating whether or not the individual has any children under age 18.

[39] City region is a dummy variable indicating if a person is living in one of the municipalities: Stockholm, Göteborg or Malmö.

[40] Chi-squared value generated by a likelihood ratio test statistic using the log-likelihood values from a model with and without the factor component. The critical value is 7.815 at the 5% significance level.

Table A2 Parameter estimates of the one factor model for 1996, foreign-born people

Variables	Employment equation Treated			Employment equation Non-Treated			Selection equation		
	P.E.	S.E.	M.E	P.E.	S.E.	M.E.	P.E.	S.E.	M.E.
Factor	0.129	0.117	0.049	-0.039	0.194	-0.014	2.721	0.263	0.203
Age	-	-	-	-	-	-	-0.044	0.005	-0.003
Age-groups (CG: 19-25)									
26-45 years	0.001	0.119	0.001	0.019	0.055	0.006	-	-	-
46-60 years	-0.487	0.155	-0.175	-0.558	0.075	-0.204	-	-	-
Education (CG: primary)									
High School	0.041	0.089	0.014	0.180	0.053	0.065	-0.160	0.117	-0.012
College	0.145	0.116	0.052	0.348	0.059	0.127	-0.183	0.162	-0.013
Has children	0.178	0.081	0.064	0.211	0.041	0.077	-0.241	0.131	-0.018
Country of origin (CG: Nordic countries)									
East Europe	0.099	0.158	0.035	-0.022	0.076	-0.008	-0.107	0.218	-0.008
West Europe	-0.053	0.129	-0.019	-0.236	0.074	-0.086	0.632	0.217	0.047
South Europe	-0.225	0.141	-0.081	-0.111	0.075	-0.041	0.234	0.251	0.017
Arab countries	-0.686	0.127	-0.247	-0.663	0.057	-0.242	-0.321	0.193	-0.024
Africa	-0.942	0.129	-0.339	-0.921	0.064	-0.336	0.167	0.194	0.012
Other nations	-0.298	0.133	-0.107	-0.252	0.056	-0.092	-0.361	0.189	-0.027
Years since immigration (CG: >11 years)									
0- 5 years	-	-	-	-	-	-	0.475	0.141	0.035
6-10 years	-	-	-	-	-	-	-0.536	0.167	-0.040
Income 1992	-	-	-	-	-	-	-0.167	0.094	-0.012
City Region	-	-	-	-	-	-	-1.247	0.174	-0.093
L-L model		-7039.99		σ^2_1	1.017		N (total)		6270
L-L constant		-8666.95		σ^2_0	1.001		N_1 (trainees)		1180
L-L no factor		-7053.61		σ^2_D	8.403		N_0 non-trainees		5090
Chi-squared		27.2		σ_{10}	-0.005				
Pseudo R^2		0.341		σ_{1D}	0.351				
McFadden R^2		0.187		σ_{0D}	-0.106				

Note: CG means comparison group; P.E. means parameter estimate; S.E. means standard error; and M.E. means marginal effect.

Table A3 Parameter estimates of the one factor model for 1997, Swedish-born

Variables	Employment equation Treated			Employment equation Non-Treated			Selection equation		
	P.E.	S.E.	M.E.	P.E.	S.E.	M.E.	P.E.	S.E.	M.E.
Factor	0.051	0.130	0.016	-0.642	0.221	-0.204	2.014	0.254	0.206
Age	-	-	-	-	-	-	-0.046	0.005	-0.004
Age-groups (CG: 19-25)									
26-45 years	0.295	0.135	0.097	0.033	0.054	0.011	-	-	-
46-60 years	-0.220	0.176	-0.072	-0.299	0.070	-0.095	-	-	-
Education (CG: primary)									
High School	0.426	0.143	0.141	0.316	0.041	0.101	-0.983	0.148	-0.101
College	0.501	0.191	0.165	0.579	0.070	0.184	-1.256	0.203	-0.129
Has children[41]	0.328	0.110	0.108	0.151	0.062	0.048	0.493	0.121	0.051
Income 1992	-	-	-	-	-	-	0.274	0.081	0.028
City Region[42]	-	-	-	-	-		-0.469	0.131	-0.048
L-L model	-4743			σ^2_1	1.002		N (total)		4416
L-L constants	-5874			σ^2_0	1.412		N_1 (trainees)		735
L-L no factor	-4754			σ^2_D	5.056		N_0 (non-trainee)		3681
Chi-squared[43]	21.9			σ_{10}	-0.032				
Pseudo R^2	0.338			σ_{1D}	0.102				
McFadden R^2	0.192			σ_{0D}	-1.292				

Note: CG means comparison group; P.E. means parameter estimate; S.E. means standard error; and M.E. means marginal effect.

[41] This is a dummy variable indicating whether or not the individual has any children under age 18.

[42] City region is a dummy variable indicating if a person is living in one of the municipalities: Stockholm, Göteborg or Malmö.

[43] Chi-squared value generated by a likelihood ratio test statistic using the log-likelihood values from a model with and without the factor component. The critical value is 7.815 at the 5% significance level.

Table A4 Parameter estimates of the one factor model for 1997, foreign-born

Variables	Employment equation Treated			Employment equation Non-Treated			Selection equation		
	P.E.	S.E.	M.E	P.E.	S.E.	M.E.	P.E.	S.E.	M.E.
Factor	0.160	0.114	0.057	-0.027	0.013	-0.009	2.590	0.252	0.208
Age	-	-	-	-	-	-	-0.040	0.005	-0.003
Age-groups (CG: 19-25)									
26-45 years	-0.113	0.118	-0.041	-0.008	0.049	-0.003	-	-	-
46-60 years	-0.180	0.155	-0.065	-0.275	0.058	-0.100	-	-	-
Education (CG: primary)									
High School	-0.003	0.088	-0.001	0.261	0.045	0.094	-0.145	0.117	-0.011
College	0.256	0.116	0.092	0.458	0.055	0.166	-0.167	0.162	-0.013
Has children	0.292	0.081	0.105	0.191	0.041	0.069	-0.200	0.129	-0.016
Country of origin (CG: Nordic countries)									
East Europe	0.001	0.158	0.000	-0.004	0.075	-0.002	-0.127	0.217	-0.010
West Europe	-0.295	0.128	-0.106	-0.186	0.072	-0.067	0.571	0.217	0.045
South Europe	-0.177	0.141	-0.063	-0.025	0.074	-0.009	0.203	0.251	0.016
Arab countries	-0.695	0.123	-0.251	-0.641	0.056	-0.232	-0.355	0.189	-0.028
Africa	-0.867	0.128	-0.312	-0.799	0.063	-0.290	0.092	0.197	0.007
Other nations	-0.454	0.132	-0.163	-0.188	0.055	-0.068	-0.410	0.191	-0.032
Years since immigration (CG: >11 years)									
0- 5 years	-	-	-	-	-	-	0.472	0.136	0.038
6-10 years	-	-	-	-	-	-	-0.512	0.167	-0.041
Income 1992	-	-	-	-	-	-	-0.263	0.082	-0.021
City Region	-	-	-	-	-	-	-1.170	0.174	-0.09
L-L model		-7104.61		σ^2_1	1.025		N (total)		6270
L-L constant		-8673.49		σ^2_0	1.001		N_1 (trainees)		1180
L-L no factor		-7120.65		σ^2_D	7.708		N_0 non-trainees		5090
Chi-squared		32.1		σ_{10}	-0.004				
Pseudo R^2		0.333		σ_{1D}	0.414				
McFadden R^2		0.181		σ_{0D}	-0.069				

Note: CG means comparison group; P.E. means parameter estimate; S.E. means standard error; and M.E. means marginal effect.

Patterns of Social Assistance Receipt - Experiences from Sweden during a Period of Rapidly Deteriorating Labor Market Conditions

Thomas Andrén[α] and Björn Gustafsson[β]

Abstract
This paper analyses Swedish-born people who became first-time receivers of social assistance in 1987 and 1992. The macroeconomic situation at the time of entry was rather different for the two cohorts, and the number of new entrances increased by almost 50 percent. Much of the increase consisted of young adults, and experience of social assistance has been shown to be fairly widespread among those under age of 25. The role of bridging the period for individuals of being supported by their parents to being established as wage earners has become an important function of social assistance in Sweden. We find that pattern of social assistance receipt is rather heterogeneous across new recipients. The complex pattern of receipt means that due to choice of perspective, duration of social assistance can appear rather different. On one hand, we find that median duration of social assistance receipt is as low as five months when an eleven-year follow-up period is applied. On the other hand, among people who receive social assistance during one particular year, as many as half had, entered receipt more than four years earlier.

Keywords: Social assistance, unemployment duration, Sweden, panel data

JEL classification: D31, D6.

[α] Göteborg University, Department of Economics, Box 640, SE 405 30 Göteborg, Sweden;
E-mail: Thomas.Andren@economics.gu.se
[β] Göteborg University, Department of Social Work, Box 720, SE 405 30 Göteborg, Sweden;
E-mail: Bjorn.Gustafsson@socwork.gu.se

1 Introduction

All industrialized countries have safety-net programs for people with financial problems. How many people enter receipt and what characterizes them? How long do people receive assistance? Are such patterns of receipt affected by the macroeconomic situation? The answers to these questions are relevant for understanding how social assistance programs function and for shaping policies aimed at reducing the need of receipt.

It is true that statistical authorities in many countries publish information on the numbers of people dependent on social assistance during a given period. Such information is important for monitoring how the extent of financial problems at the household level changes. However, these numbers are not useful in describing the patterns of receipt; one and the same rate of receipt can stand for very different realities. It is possible that the same persons receive assistance year after year, in which case the duration of receipt is long. This particular case suggests that when the observation period is extended, the number of recipients does not change much. The opposite case is short periods of receipt, as well as proportions of the population receiving social assistance increasing rapidly when the observation period is extended. To distinguish between these alternatives, panel data is needed.

In the literature, an increasing number of studies of social assistance receipt using panel data can be found. In particular, there have been increases in the number of studies where those who already have entered the system are followed over time. Such studies can provide answers to how long people stay on social assistance receipt and how such periods of receipt vary. This study contributes to the literature by providing evidence for Sweden.

We study people who entered social assistance receipt for the first time in 1987 respectively 1992 and follow them until 1997. Within this period, the Swedish labor market changed dramatically. Sweden was transformed from a country with a remarkably low unemployment rate (about 1% in 1990) into a country possessing unemployment rates similar to other European countries (about 12% in 1997). The rapid deterioration in the labor market hit young adults trying to find a job particularly hard. As many young job seekers did not have previous work experience, they were ineligible

for unemployment compensation. However, unlike their counterparts in several other countries, they were eligible for social assistance.

This study is not the first on patterns of social assistance receipt in Sweden. Recent comparative studies on exits from social assistance put the duration of social assistance in Sweden in perspective of its counterparts in other European countries. A first example is Gustafsson and Voges (1997), which compares the periods of receipt for new recipients in one Swedish city (Göteborg) with its counterparts in a German city (Bremen). The results showed periods until the first exit in the Swedish sample to be shorter on average than in the German city. The results also showed that a considerable proportion of those who exited receipt in the Swedish city returned for a second period and in some cases even more periods of receipt.

A second example is Saraceno ed. (2002) which includes Göteborg and Bremen as well as two cities in Italy (Milan and Turin) respectively Spain (Barcelona and Vitoria) one in Portugal (Lisbon) and an additional Swedish city (Helsingborg). Heikkilä and Keskitalo eds (2001) added the same information for one city in Austria (Linz) and two cities in Denmark (Aarhus and Copenhagen) respectively Finland (Tampere and Vaasa). The results from these two studies show social assistance receipt in the cities of Portugal and Spain stands out as being of much longer duration than in other cities investigated. This can be looked upon as a paradox as benefit levels were found to be lower in the south of Europe. However, the results should be seen in the light of social policies being more ambitious in the north of Europe. The studies also show that while the systems of social assistance provision appeared to be rather homogeneous across cities in Sweden, this was not the case for the cities in Italy and Spain.

While useful in many respects these comparative studies also have limitations, which are not shared with the present study.[44] First, this study is not limited to one city but encompasses the entire country. Second, the follow-up period for this study is much longer; in fact in some of the analyses we can follow new recipients as long as for one decade after initial entry. Third, we are able to study the process of entry as our samples are of all Swedish-born persons, not limited to recipients. Fourth, our data permits

[44] In addition to the comparative studies there are a few studies on patterns of social assistance receipt in Sweden. One example is Milton and Bergström (1998) who compared duration of receipt among persons in one city who had their applications processed at two different social welfare offices. Other examples are the works of Salonen referred to below.

showing how experience of social assistance receipt increases with an increase in the period of observation. When doing this we follow Salonen (1993) and Salonen (1997). However, while the first study mentioned is limited to a few municipalities and the last mentioned to birth cohorts entering work life, the present study covers the whole country, and all birth cohorts. Fifth, by taking samples of persons who entered receipt in different years, and comparing the results cross entry-cohorts we can investigate how the deteriorating labor market situation has affected patterns of receipt. While earlier studies (e.g., Gustafsson (1984), Stenberg (1998), Gustafsson (1998)) have shown a clear link between unemployment and social assistance receipt, they have not contrasted patterns of receipt for cohorts who have entered receipt during various macroeconomic situations.

Our strategy of working with large samples of recipients as well as non-recipients drawn from the total population, for analysing patterns of social assistance receipt is new for Sweden. However, the strategy has been adopted in studies for United States (e.g., Duncan (1984), Blank and Ruggles (1994) and Blank and Ruggles (1996)). Voges and Rohwer (1992) is the only study from a European country we are aware of. The authors analysed the German Socio-Economic-Panel for the period 1984 to 1989, reporting that dependency of social assistance is typically a relatively short-term situation in an individual's life cycle.

The rest of the paper is laid out as follows. In the next section we provide an institutional background and our research strategy is lined out in Section 3. The subsequent sections centre on results; Section 4 on entry into social assistance receipt, and Section 5 on undergoing and exiting receipt. Finally, we sum up the conclusions in Section 6.

2 Institutional background

Systems of social assistance vary widely across the different countries because of cross-country differences in the role played by markets, families and the public sector. Results from recent comparative studies (e.g., Earderly (1996a,b), Heikkilä and Keskitalo (2001) and Saraceno et al. (2002)) are helpful in pointing out what features distinguish the Swedish system.

The Swedish system of social assistance has a very long history. Actually through its precursor, poor relief, it is older than the welfare state. It has always been administered and financed by local governments, but works under a common legal framework for the entire country. There were only minor changes in the legislation for the system during the period studied here.[45]

Different from the case in several other countries, categories of people aged 18 years or older are entitled to social assistance. With a few exceptions, parents are not legally required to support their grown-up children, and vice versa. The requirement for social assistance is possession of a low income in combination with no opportunity of obtaining a living in another way. The latter means, for example, that a household should try to support itself by paid work or by drawing on savings. Only if such possibilities do not exist is one eligible for social assistance.

Social assistance receipt requires an application. It is generally perceived that many who are entitled refrain from applying meaning that take-up rates are low. The application is made at a social welfare office, where it is processed by a social worker that in many cases also is the decision maker. The application is typically made for a period of one month, and is often followed by later applications. There is no time limit for how long of a period one can receive social assistance.

When making decisions on an application for social assistance, certain guidelines are used. A simplified description of the decision making process is the following: People with a disposable income lower than income thresholds laid down in the social worker's guidelines who cannot make a living in any other way receive social assistance. The sum closes the gap between disposable income and the relevant threshold. Compared to its counterparts in several other countries, the thresholds in the Swedish system are relatively generous.

While all categories of people can receive social assistance in Sweden, receipt is more common among; young adults, single mothers and recent immigrants, for example. The number of immigrants receiving assistance has risen dramatically (but will not be studied in this paper). While half a century ago many elderly persons received poor-relief, pension payments expanded rapidly during the 60s, 70s and 80s

[45] However, in January 1998 a change in the legal framework came into effect. While earlier each local government decided on the thresholds for social assistance, binding levels were introduced in the legal framework. Probably this has led to larger homogeneity in decision-making across local governments.

and receipt of social assistance among the elderly has become rather uncommon during the 90s.

The Swedish economy experienced major changes in the beginning of the 90s. Following a large drop in demand, unemployment grew rapidly and the labor force shrank. For example, the unemployment rate, which had stood at 2.1% in 1987 and even went down to 1.5% in 1989 rose till 5.2% in 1992, and continued up to 8.2% the following year after which it slowly decreased.

Young adults in particular were affected by this as their possibilities for finding jobs deteriorated. This is illustrated in the unemployment rate among persons aged 16-24 expanding from 4.6% in 1987 to 11.4% in 1992 and continuing to increase the following year after which it started to decline.[46] Looking at employment rates by birth cohorts, as they became older shows that while the median age for being employed at the end of the 80s were 18 years, it jumped to 22 years in 1993 and started to slowly decrease at the end of the 90s (Börjesson, 2001).

Table 1 provides some key numbers on social assistance receipt in Sweden from the institution of the present legislation in 1983 until year 2000 as background for this study. The number of persons living in a household that at least once during a year received social assistance was (with some yearly changes) about half a million up until 1991, after which it increased until 1996 by as much as 41%. In 1996, 8.2% of the population were social assistance recipients, a percentage which decreased during the remaining part of the 90s as the situation in the labor market improved. There was also an increase in duration of receipt measured by how many months during a calendar year social assistance was received. This number went up from 4.1 months in 1988 to 5.8 months in 1997 and has remained at that level.[47]

[46] Ryan (2001) reports that while Sweden had the lowest unemployment rate among youth (people under 25 years of age) at the end of the 80s, by 1993 the youth unemployment rate was higher than in Germany, Japan, Netherlands, UK and USA. Only France was reported to have a higher youth unemployment rate.

[47] It should be understood that these numbers are no estimate of the length of social assistance episodes, as a household can continue to receive assistance during subsequent years. In addition receipt is accumulated for a calendar year without taking into account if receipt was continuous or not.

Table 1 Social assistance in Sweden 1983 - 2000, some key numbers

Year	Number of recipients (thousands)	Rate of recipients (percent)	Average number of months of receipt	Total sum Million SEK (Prices of year 2000)
1983	475	5.7	4.2	4 476
1984	524	6.3	4.2	5 019
1985	536	6.4	4.2	5 670
1986	564	6.8	4.3	6 454
1987	540	6.4	4.2	6 453
1988	524	6.2	4.1	6 149
1989	505	5.9	4.2	5 990
1990	492	5.7	4.3	5 922
1991	511	5.9	4.4	6 473
1992	560	6.5	4.6	7 866
1993	642	7.4	4.8	9 339
1994	696	7.9	5.1	10 790
1995	689	7.8	5.4	11 035
1996	722	8.2	5.8	12 102
1997	718	8.1	5.8	12 541
1998	660	7.5	5.8	11 589
1999	581	6.6	5.8	10 571
2000	522	5.9	5.8	9 532

Source: Socialstyrelsen (2001) Socialbidrag 2000, Stockholm. Note: There is a change in data collection starting in 1990. Previously people who received social assistance in more than one municipality (for example due to migration) were counted more than once.

3 Research strategy

For this study we use the sample of Swedish-born people available in the Swedish Income Panel (SWIP). This sample was drawn from the register of the total population (RTB) kept at Statistics Sweden of all registered persons residing in Sweden.[48] We concentrate here on adult persons. Information available in the panel is composed of a number of demographic variables and variables from income registers kept at Statistics Sweden. The information is annual covering various income sources.[49]

SWIP has income records for each person sampled, as well as for his or her spouse. The same income information exists for cohabiting partners if the partners have a child (under 18 years of age) in common. We use information in the income records of both partners to determine if a household and its members received social assistance during a particular year. In the panel analysis, we follow individuals and evaluate social assistance receipt by assessing the income records of both partners.

[48] The sample was originally drawn for people who resided in Sweden 1978, but has since then been updated with newborns as well as people who have immigrated.
[49] When this study was made, income information up to year 1997 was available.

Information on social assistance receipt is available in this data from 1983, the year when the present legislation came into effect. The information contains the amounts of receipt during the calendar year, and starting from 1985, the number of months during a calendar year social assistance was received. However, the timing of receipt during the calendar year is not known, therefore it is not possible to create spells of receipt on a monthly basis.

The main interest of this study is to describe and analyze the pattern of entry, going through and finally exiting social assistance receipt among new applicants. To qualify as a "new" recipient we require that the person investigated has not been a social assistance recipient during any of the preceding three years. In the main part of the analysis, we study persons who entered social assistance receipt for the first time in 1987, respectively 1992, and follow them up as long as possible. This means we have a follow-up period of ten years for the first entry-cohort, and respectively a five-year follow-up period for the second. Our approach means that we can observe persons and their households before the year of entry as well.

We first analyze how many persons enter receipt for the first time and which characteristics affect entry into social assistance receipt. The populations studied for these analyses are those who were not receiving social assistance 1, 2 and 3 years before the year of entry. Entry is defined as a binary variable and we estimate probit models with a variable taking the value of 1 if the person is a new recipient, while otherwise it takes the value of 0. Explanatory variables measure age, household type, location and various components of income of the adult household members as observed the year before entry.[50] Starting with 1990, we also have access to information on education of the person and can use this in the model specifications. Comparing estimates across entry-cohorts shows the importance of the changed macroeconomic climate. We also test if the characteristics of the new recipients differ from characteristics of others who received social assistance during the same year.

In the following analysis, we work with balanced panels of people who entered social assistance receipt for the first time in 1987, respectively 1992. That means we require that the persons resided in Sweden continuously during the follow up period. Grounds for attrition from the samples are emigration and death. We find in our sample

[50] "Adult member" is a person sampled who has reached the age of 18 as well as his or her partner. The partner is either husband or wife or a cohabiting person who is also the other parent to a child under 18.

that only a very few emigrated, while the number of deaths was larger. The latter prompts us to investigate if mortality among social assistance recipients differs from mortality of others in the population; we ask particularly whether or not mortality is higher.

Patterns of social assistance receipt are many-sided, and therefore, different approaches are used to show it. Such include computing the conditional probability of receiving social assistance after 1, 2, and more years of entry, as well as computing the number of years and the total months of receipt during the follow-up period. In addition, we compute survival curves for the state of being on social assistance receipt, and also for being in the state of non-receipt among people who have left receipt, recidivism.

We are also interested in what distinguishes people who continue to be receivers of social assistance from others who entered the system at the same point in time. Out of several different ways to make an analysis, we chose to identify persons in receipt five years after entry and estimate probit models. While the approach of following new recipients provides information on how patterns of social assistance are perceived from the perspective of the first time claimants, we also change perspective to that of the social worker. Therefore, we address the question of how the group of people receiving social assistance during one calendar year is distributed regarding the year of first entry?

Finally, we investigate experience of social assistance receipt by extending the follow-up period from one year to cover subsequently more years. This analysis is made for different birth cohorts with an emphasis on those who reach the age of 18 during the period studied.

4 Entering social assistance receipt

Table 2 shows the number of new adult recipients and results from estimating models for first entry into social assistance receipt during 1987 and 1992, and for comparison we also show results for people who entered receipt during 1997. First, it can be observed that the number of first time receivers increased rapidly between the first two years, to have decreased in 1997. The number of new recipients corresponds to the following numbers in the total population: 61000, 97200 and 80000 and entry probabilities (among non-receivers) of 1.0% in 1987, 1.6% in 1992, and 1.3% in 1997. These numbers are much lower that the rate of recipients reported in Table 1, which

means that many social assistance recipients must have entered receipt during previous years.[51]

Table 2 Estimates of probit-models relating first time social assistance receipt to variables at the household level

Year of entry	1987		1992		1997	
	Estimate	Std-error	Estimate	Std-error	Estimate	Std-error
Demographics						
Constant	-2.694*	0.011	-2.879*	0.114	-2.949*	0.137
Age (18-26)	0.689*	0.077	1.054*	0.084	0.942*	0.096
Age (27-39)	0.519*	0.083	0.955*	0.087	0.807*	0.1006
Age (40-49)	0.381*	0.089	0.836*	0.088	0.848*	0.102
Age (50-59)	0.345*	0.092	0.606*	0.089	0.591*	0.102
Age (60-69)	-0.014	0.092	0.164	0.091	0.292*	0.103
Single male no child	0.325*	0.078	0.044	0.061	0.229*	0.081
Single female no child	0.233*	0.078	-0.015	0.062	0.184*	0.082
Single with child	0.452*	0.094	0.353*	0.072	0.382*	0.093
Couples with child	0.441*	0.084	0.244*	0.063	0.476*	0.088
Big city region	0.047	0.044	0.096*	0.038	0.059	0.041
Small city region	-0.095*	0.042	-0.002	0.038	-0.007	0.041
Income source year t-1						
Capital	-0.023*	0.005	-0.0001	0.0001	-0.001*	0.0003
Pension	-0.001	0.001	-0.002*	0.0007	-0.003*	0.0007
Unemployment	0.009*	0.002	0.005*	0.0008	0.001	0.0006
Earnings	-0.003*	0.001	-0.003*	0.0002	-0.003*	0.0002
Education						
Primary school	-	-	0.377*	0.048	0.272*	0.050
Secondary school	-	-	0.257*	0.044	0.256*	0.044
Log likelihood		-3039.59	-4350.69			-3613.64
Number of observations		59638	61637			62257
Number of first time receivers		614	972			801
All social assistance receivers		2529	2820			3255

Note: The omitted age category is a person aged 70 +, a couple with no dependent children, living in rural regions, having post secondary education. * means significant at the 5% level.

We find that the new adult recipients are in minority among all adults who are receivers for a given year. They stand for 24% of all receivers in 1987, while the proportion is 34% in 1992 and 25% in 1997. Thus, the relative frequency of first time receiver to all receivers not surprisingly increased when the labor market situation deteriorated and decreased when it improved. Looking at the parameter estimates

[51] An additional reason for the difference is that while Table 2 refers to people born in Sweden, Table 1 also includes foreign-born persons. Also, note that our analyses concern adult receivers, while Table 1 also included children living in households receiving social assistance.

reported in Table 2 several comments can be made.[52] There is a very clear and expected age pattern, meaning that the risk of entry is highest for young adults and decreases with age. This pattern is less pronounced in 1987. The rapidly deteriorating labor market situation for young adults has thus, as expected, led to increased risk for being a new social assistance recipient. The estimated effects of variables measuring household type do not change across years in an equally easily interpretable manner. The estimates show that families with children (and in some cases also single persons) having higher risks of entry compared to couples without children. Further, on we can see in Table 2 that effects of region are all small, although some of them are estimated with relatively large t-values. Income components included in the specification measure earnings, capital income, pensions and unemployment compensation, all observed the year before entry. Not surprisingly, for all years of entry, the negative coefficient for earnings is estimated with a large t-value and the parameter estimate does not change much across years. Establishing oneself on the labor market is thus one strategy for avoiding subsequent entry into social assistance receipt. The results also point towards the amount of capital income as well as pension income during the preceding year negatively affecting entry into social assistance receipt. However, the relevant parameters are not always estimated with a high t-value. The positive parameter for the variable unemployment compensation has a high t-value for the entry year 1987 and 1992, but not for 1997. Finally, we find that primary and secondary education positively affects entry into social assistance.

To illustrate the magnitude of the various effects in Table 3 we predict probabilities of becoming a first time social assistance recipient in 1987, 1992 and 1997 for persons with selected characteristics based on the parameters reported in Table 2. First, all are considerably lower than 50%. This observation means that we as analysts have not had access to variables having strong predictive power. For example in this study, we have not observed unemployment experience at the household level. The relatively low probabilities can also be taken as supporting the view that rates of social assistance take-up are low in Sweden.

[52] As information on education is not available for year 1987, the specification reported in Table 2 does not include such variables for the first year of entry. Therefore when making comparisons across years of entry, we have also estimated the same specification excluding the education variables. It turns out that the coefficients for age, family type and income during year t-1 are similar for the two specifications.

Table 3 Predictions of first time social assistance recipient, year 1987, 1992, and 1997 (percentage)

Year of entry	1987		1992		1997	
Position in the earnings distribution	1st perc. earnings	10th perc. earnings	1st perc. earnings	10th perc. earnings	1st perc. earnings	10th perc. earnings
Person 1						
Single parent						
Age 18-26	5.89	1.01	14.31	1.49	8.91	0.17
Large city region						
Primary education						
Person 2						
Single male no child						
Age 18-26	4.54	0.71	8.46	0.65	6.69	0.10
Large city region						
Primary education						
Person 3						
Single male no child						
Age 70 –	0.86		0.76		0.73	
Large city region						
Primary education						

Note: Predictions based on estimates reported in Table 2. As it is very unlikely that a person aged 70 has earnings in the 10 percentile of the earnings distribution the corresponding cells are not shown.

A second observation from Table 3 is the very large variation in the probability of becoming a first time recipient across persons with different demographic characteristics. Take the numbers for 1992 as an example. For a person with the same earnings, education and location, the probability of becoming a first time recipient relates as 1 to 19, that is when one compares the probability for a single elderly person with a single parent.

A third observation is that being established in the labor market and thus earning income strongly reduces the risk of entering social assistance. Take again the risk for a single parent aged 18-26 as an example. The predicted probability of becoming a first time recipient falls by almost nine-tenths if the parent was at the top of the earnings-distribution in 1991 as compared to being at the bottom.

A fourth observation is that in some cases predictions change as suggested by the situation on the labor market across years studied, while this is not true in other cases. The changes are evident for young adults having earnings at the bottom of the earnings distribution for whom the risk of entering social assistance receipt is considerably higher in 1992 than in 1987, then falls to an intermediate level in 1997. On the other hand the predictions for elderly persons are similar for all years. The pattern of change

across cohorts is less clear for persons having earnings at the top of the earnings distribution.

Table 4 Tests of differences in the distribution of first time receivers and the stock of recipients in 1987, 1992 and 1997

Characteristic	1987			1992			1997		
	Stock	Debut	P < | t |	Stock	Debut	P < | t |	Stock	Debut	P < | t |
Age (18-26)	0.313	0.464	0.001	0.271	0.461	0.001	0.354	0.515	0.001
Age (27-39)	0.390	0.268	0.001	0.406	0.284	0.001	0.328	0.243	0.001
Age (40-49)	0.171	0.109	0.001	0.189	0.138	0.001	0.183	0.131	0.001
Age (50-59)	0.066	0.053	0.234	0.081	0.052	0.002	0.098	0.054	0.001
Age (60-69)	0.034	0.035	0.855	0.024	0.023	0.909	0.019	0.026	0.291
Age (70 -)	0.023	0.068	0.001	0.026	0.038	0.108	0.014	0.028	0.028
Male no child	0.373	0.394	0.358	0.406	0.374	0.092	0.415	0.428	0.532
Female no child	0.194	0.285	0.001	0.219	0.269	0.003	0.259	0.317	0.002
Single parent	0.210	0.089	0.001	0.194	0.109	0.001	0.193	0.099	0.001
Couple with child	0.184	0.180	0.828	0.145	0.179	0.021	0.101	0.124	0.081
Couple no child	0.037	0.050	0.181	0.034	0.067	0.001	0.028	0.029	0.881
Big City	0.364	0.363	0.945	0.441	0.400	0.036	0.456	0.425	0.123
Small City	0.381	0.359	0.343	0.363	0.396	0.087	0.326	0.370	0.023
Rural Region	0.209	0.234	0.202	0.195	0.203	0.622	0.296	0.203	0.420
Missing info.	0.044	0.042	0.818	0.000	0.000	0.000	0.000	0.000	0.000
Primary	-	-	-	0.488	0.352	0.001	0.483	0.305	0.001
Secondary	-	-	-	0.455	0.544	0.001	0.498	0.564	0.001
P. Secondary	-	-	-	0.056	0.102	0.001	0.098	0.129	0.019

Note: The test used is a simple t-test allowing for unequal variances between the distributions

We now turn to comparing how characteristics of new recipients relate to all who received social assistance during the same year (the stock). Such dissimilarities can suggest that the duration of assistance receipt differs by the characteristics investigated. Differences between new and all recipients can also be due to long-term changes in the composition of recipients, while duration for each category remains unchanged.

Table 4 reports the outcome of t-tests for differences across characteristics when splitting the sample according to household type, age, family type, location, education, and various income components the year before entrance. This is done for all three entry cohorts. Several comments can be made. Starting with age we find that in all three cohorts the new recipients contain a larger proportion of people aged 18-26 than among all who receive social assistance the same year. Actually, as many as half of the new recipients are found to be aged 18-26. In contrast, people aged 27-49 make up a smaller

proportion of new recipients than among all people who receive social assistance the same year.

Turning to types of family, we find single females without children (many young adults) make up a larger share of the new recipients than among all recipients for all cohorts, while the converse is the case for single parents. There is also a trend across the years studied of couples with children making up ever-smaller proportions of all recipients.

While there is not much of difference between the stock of recipients and new recipients, regarding distribution of recipients by region more is found when looking at education. The new recipients have longer educations compared to the stock of recipients. During the year of entry, new recipients receive much less social assistance than other recipients, which is the same difference seen regarding receipt of unemployment compensation. On the other hand, new recipients have higher earnings on average than all recipients.

5 Going through and exiting social assistance

In this section, we follow persons who entered social assistance receipt for the first time during 1987 respectively 1992. A few persons have left the sample due to emigration, a larger number (3,667) due to death. Is mortality higher among social assistance recipients than among others? It may be surmised that people with poor health are likely to have difficulties making a living and are therefore dependent on social assistance. In addition, it is known from surveys that social assistance recipients are more frequently exposed to violence than others (See for example Estrada & Nilsson 2001).

Table 5 A probit analysis of social assistance recipients and mortality during the period 1987 to 1992

Explanatory Variables	Parameter Estimates	Standard Errors
Constant	-1.981	0.033
If social assistance (yes = 1)	0.353	0.064
If newly introduced (yes = 1)	0.285	0.110
Male	0.283	0.020
Age (61-67)	0.853	0.031
Age (68-71)	1.112	0.035
Age (72-75)	1.366	0.034
Age (76-78)	1.548	0.037
Age (79-80)	1.833	0.045
Age (81-83)	2.023	0.040
Age (84-85)	2.229	0.053
Age (86-89)	2.421	0.048
Age (90-104)	2.952	0.063

Predictions

	Person aged 61-67	
	Male	Female
Reference	7.9	4.4
People who became first time recipients of social assistance 1987	12.9	7.9
Others Receiving Social Assistance 1987	14.4	8.9

Note: Number of observations 64,747 (deceased: 3,667)

We investigate mortality among social assistance recipients by following the sample of adults from 1987 to 1992 and observe mortality during this period. The sample for this analysis includes all adult persons living in Sweden in 1987. Thus, the sample includes in addition to new receivers, other persons on social assistance. A probit model relating mortality to variables measuring age, gender and social assistance receipt is estimated in the specification. We use one variable for first-time receivers another for others who received social assistance in 1987. The results are reported in Table 5. As is expected the results show mortality to increase with age and to be higher for men then for women. The results also show positive effects for both variables measuring social assistance receipt in 1987, and the parameters are both estimated with high t-values.

In the bottom of Table 5, we illustrate the magnitude of the effects by predicting the probability of death for a person aged 61-67. It shows that receivers of social assistance have a probability of death twice as great as for non-receivers. This difference is of the same magnitude as the well-known difference in mortality between

men and women, as well as the difference in mortality between white and blue-collar workers (for the latter see for example Kåreholt, 2000).

Table 6 Conditional probabilities of receiving social assistance a particular year after entry

	Year of entry	
	1987 (N=614)	1992 (N=972)
Years since entry	Percent	Percent
1	41.69	48.15
2	27.69	34.36
3	21.66	27.98
4	22.64	25.31
5	20.85	22.63
6	23.13	-
7	19.22	-
8	15.80	-
9	15.80	-
10	13.84	-

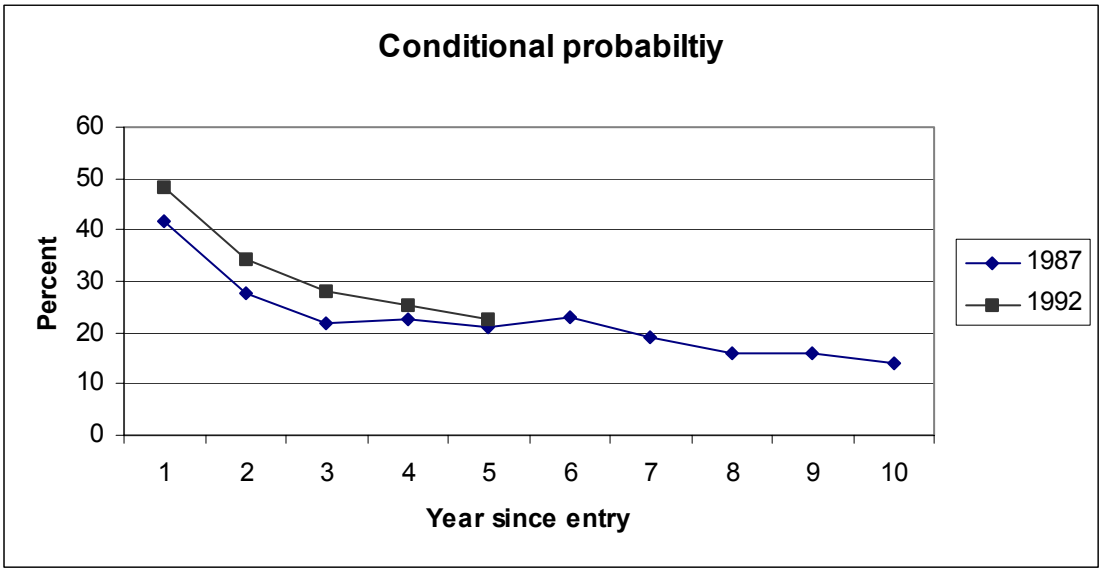


Figure 1 Conditional probabilities of receiving social assistance a particular year since entry.

We now follow the new recipients over time to describe their subsequent receipt of social assistance. Table 6 shows the probability of receipt 1, 2, and up to 10 years after entry for the first entry-cohort and up to 5 years for the second. When comparisons across cohorts can be made we find somewhat higher numbers for the second entry-cohort. Table 6 and Figure 1 show that more than a majority (the first entry cohort)

respectively about half (the second entry cohort) of those who entered receipt were not receivers the first year after entry.

While in the beginning of the follow-up period the proportion of recipients falls rapidly, this is not the case later on. Actually for the first entry cohort we can detect a plateau as three years after entry, 23% of the new receivers were receivers a number close to the percentage recorded at the time of seven years after entry. This plateau coincides with the deteriorating labor market situation and is most probably its cause. This interpretation is strengthened by the fact that we cannot observe a similar situation among those who entered receipt in 1992 and for whom the labor market improved starting two years after entry. Table 6 also shows that 10 years after entry one-sixth of the first cohort were receiving social assistance.

Table 7 First time social assistance recipients distributed by number of years on social assistance during the five-year and ten-year follow up period

	Year of entry		
	1987	1987	1992
Year	5 year	10 year	5 year
1	38.76	33.06	36.21
2	26.06	21.99	21.30
3	14.17	13.19	15.53
4	10.10	7.33	10.70
5	4.72	5.86	7.30
6	6.19	5.05	8.95
7	-	3.75	-
8	-	3.42	-
9	-	2.28	-
10	-	2.12	-
11	-	1.95	-
Mean (years)	2.39	3.33	2.64

The number of years of receipt is tabulated for the two entry cohorts in Table 7. Rather skewed distributions are reported. About one-third received social assistance during the year of entry only. Slightly more than half of the recipients received assistance not longer than two years. When looking at the first cohort during the year of entry and the preceding ten-year period it is found that about one-sixth received social assistance during most of the years. However, only two percent received social assistance during all of the eleven years. Comparison across the two entry-cohorts shows the mean years to be only slightly larger for the second entry cohort.

Table 8a First time social assistance receivers distributed by number of months of receipt during the year of entry plus a five-year respectively ten-year follow-up period

Year of entry	1987 (N = 614)				1992 (N = 972)	
	5 year period		10 year period		5 year period	
Month	Number	Percent	Number	Percent	Number	Percent
1	124	22.9	99	19.5	189	20.9
2	88	16.3	69	13.6	106	11.7
3	46	8.5	38	7.5	66	7.3
4	42	7.8	33	6.5	60	6.6
5	33	6.1	31	6.1	59	6.5
6	24	4.4	28	5.5	21	2.3
7	24	4.4	18	3.6	48	5.3
8	16	3.0	13	2.6	34	3.8
9	11	2.0	13	2.6	30	3.3
10	20	3.7	9	1.8	19	2.1
11	6	1.1	14	2.8	18	2.0
12	13	2.4	9	1.8	18	2.0
13	10	1.9	9	1.8	18	2.0
14	4	0.7	6	1.2	12	1.3
15	6	1.1	2	0.4	13	1.4
16	9	1.7	8	1.6	12	1.3
17	5	0.9	6	1.2	9	1.0
18	4	0.7	4	0.8	14	1.6
19 – 24	24	4.4	20	3.9	43	4.8
25 – 36	16	3.0	36	7.1	66	7.3
37 – 48	7	1.3	17	3.4	32	3.5
49 – 60	9	1.7	4	0.8	18	2.0
61 – 120	-	-	21	4.1	-	-
# of month the first year (mean)	2.41				2.49	
Missing	73	-	107	-	67	-
Mean	7.62	-	12.20	-	10.14	-
Median	4.0	-	5.0	-	5.0	-

By changing the unit of observation to months of receipt, we arrive at Table 8. The most frequent value is receipt during a one-month period, which is observed for about one out of five new receivers. The distribution is skewed with many receiving assistance for only a few months, and a few for many months. This also shows up in the mean being considerably larger than the median. While the median when analyzing the year of entry and a five-year follow up is found to be 4 months for the first cohort it had increased to 5 months for the second cohort.

Table 8b First time social assistance receipt distributed by the number of months of receipt during the year of entry

Year of entry	1987		1992	
Number of months	Number	Percent	Number	Percent
1	286	46.58	446	45.88
2	157	25.57	189	19.44
3	44	7.17	115	11.83
4	43	7.00	81	8.33
5	26	4.23	58	5.97
6	15	2.44	27	2.78
7	15	2.44	20	2.06
8	14	2.28	12	1.23
9	9	1.47	12	1.23
10	2	0.33	7	0.72
11	0	0.00	2	0.21
12	3	0.49	3	0.31
Mean	2.41		2.49	
N	614		972	

While extending the follow-up period by five years does not affect the median much, this is not the case for the mean. Table 8 reports that while average receipt was only 2 months during the year of entry, it had increased to 8 months when applying a five year follow-up period and is found to be as high as 12 months when applying the longest follow-up period.

Table 9a Survival in the state of social assistance receipt during subsequent number of years since entry

	Year of entry	
Consecutive years	1987	1992
0	100.0	100.0
1	41.69	48.15
2	18.24	26.54
3	10.42	17.08
4	7.98	11.83
5	6.19	8.95
6	5.05	-
7	4.39	-
8	3.26	-
9	2.61	-
10	1.95	-

Table 9b Survival in the state of non-receipt among those who left social assistance receipt year t+1 among first time receivers

	Year of entry	
Consecutive years	1987	1992
1	100.0	100.0
2	81.76	84.92
3	74.71	77.38
4	67.35	73.21
5	62.35	69.84
6	57.65	-
7	55.00	-
8	52.94	-
9	51.18	-
10	50.00	-

Returning to results obtained from yearly data in Tables 6 and 7 it can be understood from comparing them that some people who have left receipt re-enter at a later year. It also means that the proportion continuously receiving social assistance since entry decreases rather rapidly with time since entry. To show this we also report numbers showing survival in the state of non-receipt among those who exited receipt one year after entry.

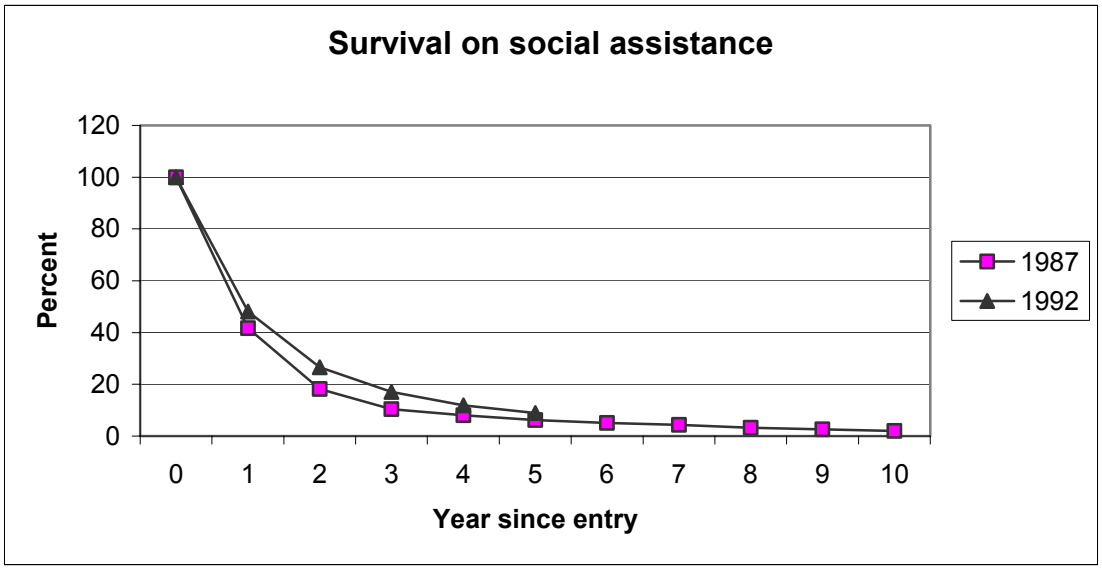


Figure 2 Survival on social assistance receipt during subsequent number of years since entry.

Looking at the first survival curve in Table 9 we find that in the first entry cohort as few as 18% had survived in the state of social assistance receipt during the first two years after entry. However, this proportion had increased to 27% for the second entry-cohort. Five years after entry the numbers were down to 6 respectively 9%.

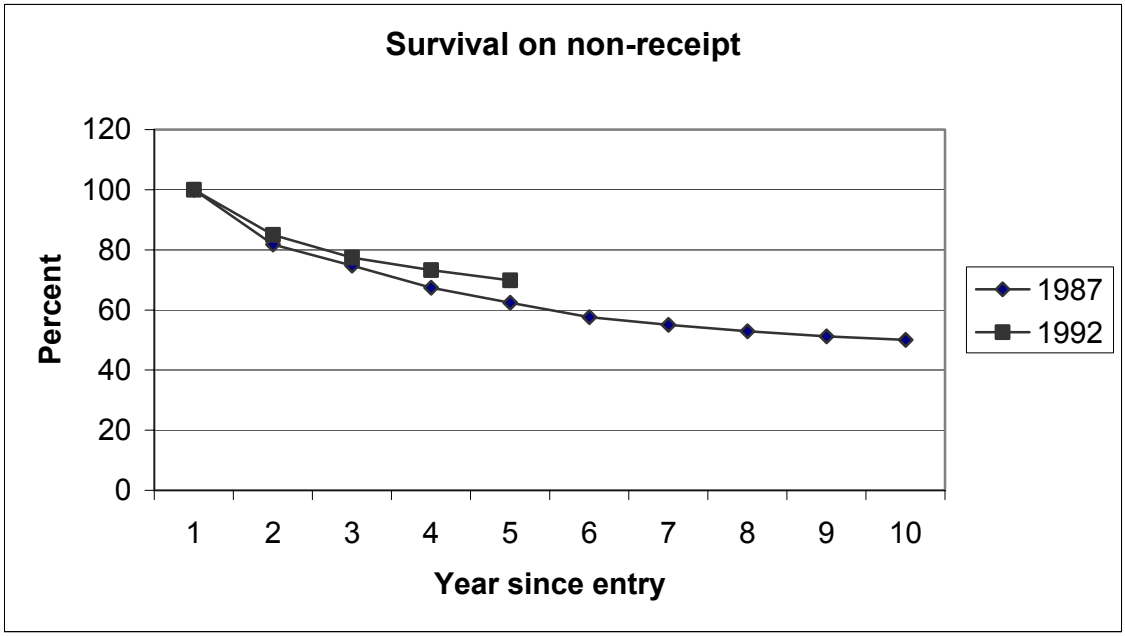


Figure 3 Survival on non-receipt among those who left social assistance receipt year t+1 among first time receivers by entry cohorts.

Turning to survival in the status of non-receipt we find that two years after entering non-receipt three-fourths had remained in the state of non-receipt. This proportion decreased to be two-thirds three years after entry into non-receipt and was one-half nine years after entry into non-receipt. These results are in agreement with results from comparative studies using monthly data referred to in the introduction as they indicate that recidivism is common among social assistance recipients in Sweden. Making comparisons across entry cohorts for the survival in non-receipt, more similarities than differences show up.

Why is recidivism in social assistance receipt high in Sweden? One, through not necessarily the only answer is probably the existence of the active labor market policy. Many social assistance recipients are job seekers and registered at the labor market authorities. After some time they receive offers to participate in labor market programs.

When involved in such programs they receive income. When the program ends, some succeed in finding a job, while others do not and have to rely on social assistance for a second period.

What distinguishes people who stay in the state of social assistance receipt for a long period from those who leave? We address this question by estimating a series of probit-models with the dependent variable taking a value of 1 if the person lives in a household receiving social assistance 1, 2, 3, 4 and 5 years after first-time entry. Explanatory variables include those used when analyzing entry into social assistance receipt in Section 4. In the specifications, we also add two dummy variables indicating if social assistance receipt makes up more than a marginal part of the disposable income for the household during the year of entry. We are interested in finding out if the degree of dependency when entering receipt affects the probability of receipt during the follow-up period.

In Table 10, we report results, which indicate that social assistance dependency during the year of entry clearly predicts subsequent receipt. Interesting enough when the variables measuring dependency are included in the specification, coefficients for variables indicating household type and region with only a few exceptions are not significant. Even more interesting is the fact that when we (in an alternative specification) add the variable measuring earnings during the year before entry, the coefficient for this variable is also not significant. However, in the specification for the 1992 entry cohort, which includes measures of education, we find additional variables affecting subsequent receipt. Young age means higher probabilities for subsequent receipt while long education have the reverse consequence.

Table 10 Estimates of probit-models relating subsequent social assistance receipt to variables at the household level, the entry cohort

a) Entry cohort 1987

	Year since entry				
Variable	T+1	T+2	T+3	T+4	T+5
Intercept	0.228	-0.329	-0.225	-0.218	-0.323
	(0.276)	(0.289)	(0.298)	(0.298)	(0.297)
An1	-0.924	-0.588	-0.514	-0.610	-0.407
	(0.139)	(0.146)	(0.153)	(0.146)	(0.149)
An2	-0.692	-0.223	-0.339	-0.497	-0.270
	(0.130)	(0.131)	(0.140)	(0.139)	(0.140)

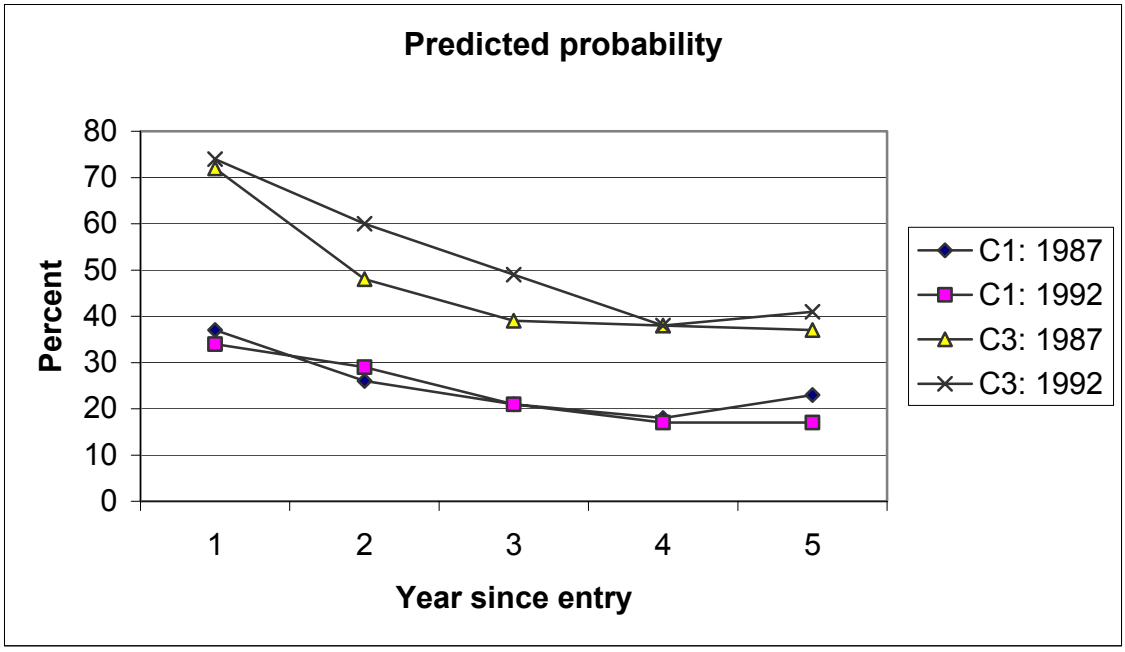
Note: The variable An1 assumes the value 1 if social assistance makes up at least 3.5 percent of disposable income but not more than 12 percent the year of entry. Variable An2 assumes a value of 1 if social assistance makes up at least 12 percent of disposable income. The specification also includes two dummy variables for age, four dummy variables for household characteristics and two dummy variables indicating region. Standard errors are reported within parentheses.

b) Entry cohort 1992

	Year since entry				
Variable	T+1	T+2	T+3	T+4	T+5
Intercept	0.540	-0.141	-0.158	-0.274	-0.517
	(0.211)	(0.229)	(0.234)	(0.242)	(0.248)
Age (18-26)	0.209	0.359	0.295	0.357	0.258
	(0.116)	(0.119)	(0.123)	(0.125)	(0.127)
Age (27 – 39)	-0.024	-0.157	-0.059	0.060	-0.035
	(0.121)	(0.129)	(0.134)	(0.135)	(0.139)
An1	-1.037	-0.775	-0.752	-0.628	-0.695
	(0.116)	(0.119)	(0.123)	(0.124)	(0.128)
An2	-0.553	-0.358	-0.455	-0.445	-0.381
	(0.104)	(0.104)	(0.107)	(0.110)	(0.111)
Secondary	-0.116	-0.147	-0.099	-0.266	-0.193
	(0.093)	(0.095)	(0.098)	(0.098)	(0.100)
Post Secondary	-0.384	-0.361	-0.457	-0.480	-0.348
	(0.153)	(0.163)	(0.179)	(0.178)	(0.176)

Note: The variable An1 assumes the value 1 if social assistance makes up at least 3.5 percent of disposable income but not more than 12 percent the year of entry. Variable An2 assumes a value of 1 if social assistance makes up at least 12 percent of disposable income. The specification also includes four dummy variables for household characteristics and two dummy variables indicating region. Standard errors are reported within parentheses.

In Figure 4 we show, based on the estimates, probabilities of subsequent receipt for a person with given characteristics with the exception of rather different degrees of dependency during the year of entry. The plots clearly show that if dependency is large this leads to a probability of receipt, which is about twice as large compared to only marginal dependency during the entry year.



Note: C1 corresponds to low social dependency, single men, aged 18-26, low educated, and living outside the metropolitan area; while C3 are high social dependency-peers;

Figure 4 Predicted probabilities for receiving social assistance for households with given characteristics but differing by the degree of social assistance dependency and year of entry (Predictions based on estimates reported in Table 10).

If dependency during year of entry is marginal, the curves are practically identical. However, for the household more dependent during the year of entry probabilities of receipt two and three years after entry are higher in the 1992 entry cohort.[53] Although receiving social assistance for many years is not typical for those entering receipt, those who stay on receipt make up a large proportion of those who are receivers at any given point in time. This is illustrated in Table 11 where we report year of first entry for people who received social assistance in 1992 respectively 1997. Among the group of receivers in 1992, only one-fourth were new receivers, while half had made their entry more than five years earlier. The picture in 1997 is somewhat different. The new receivers are an even smaller proportion of all receivers, while the fraction that had entered one to five years earlier had increased from 28 to 36%. From the finding that among many, social assistance is of short duration, it follows that the proportion of people that have experienced social assistance receipt increases relatively rapidly when the observation period is increased. Table 12 shows exactly how fast the

[53] Dependency is almost identical for the two cohorts. The distributions across the three classes of dependency were 35, 34 and 31% for the 1987 cohort while 32, 34 and 34% for the 1992 cohort.

numbers increase for various birth cohorts. Starting the comments with persons who were over 18 during the entire observation-period 1987-1997 (Figure 5), we find as expected rates of receipt vary by year of birth. At one extreme are people born in 1922 and earlier, that is people who were already over the general retirement age at the beginning of the observation period. Among them as few as 0.7% received social assistance in 1987, a number that can be compared to 8.9% among persons born 1963-1968.

Table 11 Year of entry among people receiving social assistance 1992 respectively 1997. Percent

	Year of receipt	
Year of entry	1992	1997
T	25.11	15.30
t-1	9.53	9.06
t-2	5.11	7.31
t-3	4.08	5.96
t-4	4.26	6.94
t-5	4.75	6.79
t- 6 and earlier	47.16	48.64
N	2820	3255

Table 12 Experience of social assistance receipt during cumulated number of years

Birth cohort	Year of observation											N row
	1987	87-88	87-89	87-90	87-91	87-92	87-93	87-94	87-95	87-96	87-97	
§1922	0.7	1.0	1.2	1.5	1.8	1.9	2.0	2.3	2.3	2.5	2.6	7285
1923-1932	1.1	1.4	1.7	2.0	2.3	2.5	2.7	2.9	3.0	3.1	3.3	6935
1933-1937	2.3	3.1	3.6	4.0	4.6	4.9	5.3	5.5	5.7	6.0	6,1	3692
1938-1942	2.6	3.6	4.0	4.7	5.1	5.7	6.3	6.7	7.2	7.4	7.5	4533
1943-1947	3.7	4.9	5.7	6.2	7.0	7.7	8.3	8.7	9.1	9.4	9.7	6183
1948-1952	5.2	6.9	8.0	9.0	10.0	11.0	12.0	12.7	13.2	13.6	13.9	5700
1953-1957	6.6	8.8	9.9	11.2	12.4	13.8	15.1	16.0	16.6	17.4	18.0	5394
1958-1962	7.4	10.0	11.5	12.9	14.6	16.0	17.3	18.0	18.8	19.5	20,0	5303
1963-1968	8.9	12.2	14.5	16.2	18.0	20.0	21.7	22.8	23.7	24.6	25,2	7242
1969	3.8	7.6	10.5	12.4	14.2	17.1	20.0	20.8	21.8	22.9	23,5	1089
1970	0.6	3.1	6.4	8.5	12.2	15.4	18.7	20.3	21.4	22.1	23.1	1086
1971	0.0	1.0	3.4	6.7	10.1	14.3	17.8	20.2	21.6	23.1	24,1	1155
1972	0.0	0.4	0.9	3.7	9.3	15.3	20.6	23.1	24.6	26.4	28.3	1141
1973	0.0	0.0	0.0	0.4	4.3	12.2	19.5	22.8	25.0	27.4	28.4	1125
1974	0.0	0.0	0.0	0.0	0.6	6.1	17.1	23.4	26.0	27.7	29.8	1130
1975	0.0	0.0	0.0	0.0	0.0	0.4	6.0	17.1	23.8	27.8	30.0	1055
1976	0.0	0.0	0.0	0.0	0.0	0.0	0.7	4.1	16.6	23.3	28.0	1015
1977	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	3.3	15.8	23.5	972
1978	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	2.4	9.1	933
1979	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.8	7.5	55
N column	2295	3175	3763	4334	5007	5797	6658	7287	7875	8446	8912	
N balanced panel												63023

Increasing the observation period increases the percentage of experience of receipt for the oldest group to be up to four times as large and for the youngest group, started from a much higher base, by almost three times. These numbers clearly show that experience of social assistance receipt is much more widespread when compared to annual figures of receipt from Table 1.

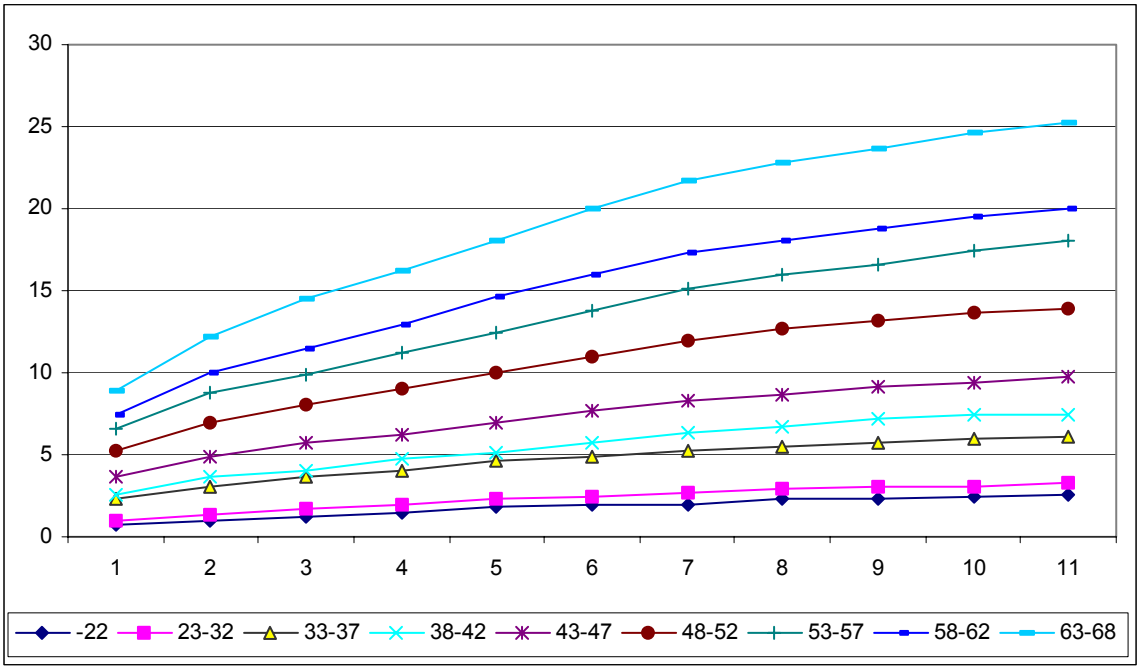


Figure 5 Experience of social assistance during observation periods lasting 1, 2, up to 11 years. Persons born during 1968 or earlier (Source Table 12)

Turning to social assistance experience among the cohorts who became 18 years of age during the observation period (Figure 6), surprisingly large differences due to year of birth can be seen. For a given age there is a very rapid increase in social assistance experience between on one hand those born 1970 and 1971, and those born 1974 and 1975. One way to show this is to make comparisons across cohorts keeping age constant. For example, at age 22, 15% of persons born 1970 experienced social assistance receipt, while the corresponding proportion was as high as 30% among those born in 1975.

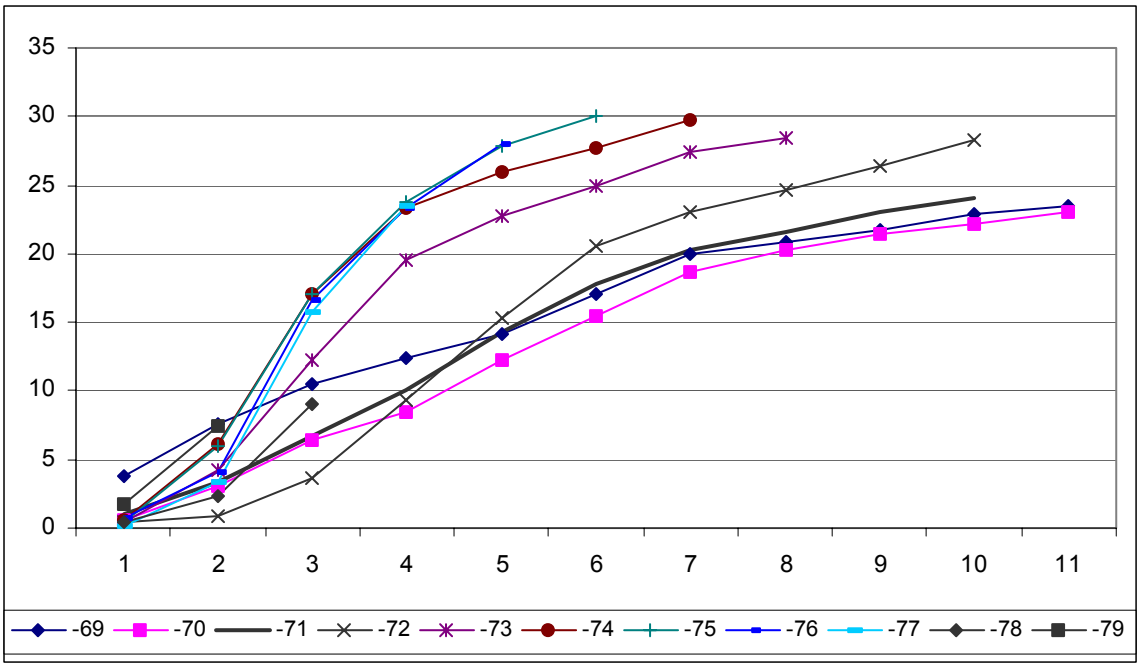


Figure 6 Experience of social assistance during observation periods lasting 1, 2, up to 11 years. Persons born 1969 and later (Source Table 12)

Alternatively, we can report that up to year 1993, 19% of the 1970 birth cohort (then aged 23) experienced social assistance receipt. This was very similar to the 1973 birth cohort although the later cohort for this year of observation was not more than 20 years of age. The table also shows that the increase in social assistance experience actually reached its peak for the 1975 cohort and did not continue to increase for the younger birth cohorts. These young persons turned 18 during a period when the labor market had started to improve.

6 Conclusions

In this paper, we have analyzed Swedish-born people who became first-time receivers of social assistance in 1987 and 1992 for a period of many years, actually starting before entry. The macroeconomic situation at the time of entry was rather different for the two cohorts. Comparisons across the two entry cohorts show several substantial differences. The number of new entrances increased from 1987 to 1992 by almost 50%. In addition, those who entered in 1992 had longer subsequent receipt.

Much of the increase consisted of young adults and experience of social assistance receipt has shown to be fairly widespread among people before they have reached the age of 25 years. We have documented a rather large increase between cohorts born in the beginning of the 70s. While about one out of six born in 1970 had experienced social assistance receipt at age 22, this was the case for as many as almost one-third of the persons born 1975.

Bridging the period for young people from being supported by their parents to being established as wage earners has become an important function of social assistance in Sweden. In this respect social assistance in Sweden differs from its counterparts in several European countries, where young adults are supposed to be maintained by their parents and are not entitled to social assistance.

We find that pattern of social assistance receipt is rather heterogeneous across new recipients. There is one group who received social assistance for a short period and did not return to the social assistance roles afterwards. We report that among new recipients about one-fifth received social assistance only during the month of entry. Further, one-third never received social assistance during the following ten years. In contrast, another group of receivers stayed on social assistance during most of the follow-up period. This group is not particularly large as we report that not more than 15% of the new recipients received social assistance during most of the eleven years following entry. The remaining group of people received assistance also after the year of entry but during less than half of the eleven years following entry. We find that patterns of receipt are complex also because half of the people who have exited receipt the year after entry return to the social assistance roles within the following decade.

The complex pattern of receipt means that due to choice of perspective, duration of social assistance can appear rather different. On one hand, we report that the median duration of social assistance receipt for a new receiver is as low as 5 months when an eleven-year follow-up period is applied. Consistent with this, experience of social assistance receipt during an eleven-year follow-up period is more widespread than suggested by yearly data. For example while 7% of people born 1957-58 in Sweden received assistance during 1987, evaluated over an eleven-year period the proportion was as high as 18%.

On the other hand, we report that if one focuses on people who were receiving social assistance during one particular year, as many as about half had entered receipt more than four years earlier. The composition of receivers according to year of first receipt changes according to the situation on the labor market. New receivers made up the largest proportion of all receivers when the labor market deteriorated.

We round off the paper by listing topics for future research. Quite in passing, we have found that social assistance recipients have higher mortality rates than other persons. This finding deserves more attention in future research. A natural extension of the present study would be to investigate patterns of social assistance receipt among the foreign-born, a category known to be over-represented among receivers of social assistance in Sweden. Finally, it should be most relevant and interesting to study how newer entry cohorts are faring. This issue is of considerable political concern as the Swedish government in 2001 adopted the goal of reducing the number of social assistance recipients by year 2004 to a level only half as high as in 1999.

References

Blank, R. and Ruggles, P (1994), "Short-Term Recidivism Among Public-Assistance Recipients", American Economic Review 65, 27 - 39.

Blank, R. and Ruggles, P. (1996), "When Do Women Use Aid to Families with Dependent Children and Food Stamps? The Dynamics of Eligibility versus Participation", Journal of Human Resources 31, 57 - 89.

Börjesson, M. (2001), "Vad innebar 1990-talet för ungdomars livsvilkor?", In Bergmark, Å (Editor) Ofärd i välfärden, Antologi från Kommittén Välfärdsbokslut, SOU 2001:54, Stockholm.

Duncan, G. (1984), Years of Poverty, Years of Plenty, Ann Arbor, The University of Michigan, Survey Research Centre.

Earderly, T., Bradshaw, J., Ditch, J, Gough, I. and Whiteford, P (1996a), Social Assistance in OECD Countries: Synthesis Report, London: HMSO, Department of Social Security Research Report No 46.

Earderly, T., Bradshaw, J., Ditch, J, Gough, I. And Whiteford, P (1996b), Social Assistance in OECD Countries: Country Reports, London: HMSO, Department of Social Security Research Report No 47.

Estrada, F. and Nilsson, A. (2001), "Brottslighet som välfärdsproblem - Utsatthet för brott och oro för brott 1988 – 1999" in Bergmark, Å (Editor) Ofärd i välfärden, Antologi från Kommittén Välfärdsbokslut, SOU 2001:54, Stockholm.

Gustafsson, B. (1984), "Macroeconomic Performance, Old Age Security and the Rate of Social Assistance Recipients in Sweden", European Economic Review, 319 - 338.

Gustafsson, B. (1998), "From the Employment Office to the Social Welfare Office: Social Assistance Recipiency Among First-time Unemployed in Sweden", European Journal of Social Work 1, 203 - 220.

Gustafsson, B. and Voges, W. (1998), "Contrasting Welfare Dynamics: Germany and Sweden compared", in Leisering, L. and Walker, R. (Editors), The Dynamics of Modern Society, Bristol: Policy Press.

Heikkilä, M and Keskitalo, E. (Editors) (2001), Social Assistance in Europe. A Comparative Study on Minimum Income in Seven European Countries, Helsinki: STAKES, National Research and Development Centre for Welfare and Health.

Kåreholt, I. (2000), Social Class and Mortality Risk, Stockholm: University of Stockholm Swedish Institute for Social Research - Dissertation Series 44.

Milton, P. and Bergström, R. (1998), Uppsalamodellen och socialbidragstagarna. En Effektutvärdering, Stockholm: Socialstyrelsen: Centrum för Utvärdering av socialt arbete, CuS skrift 1998:1.

Ryan, P. (2001), "The School-to-Work Transition: A Cross-National Perspective", Journal of Economic Literature 34, 34 - 92.

Salonen, T. (1993), Margins of Welfare, Torna Hällestad: Hällestad Press.

Salonen, T. (1997), "Socialbidraget - mellan personliga bekymmer och systembrister", in Socialstyrelsen, 11 röster om socialbidrag, Stockholm (SoS-rapport 1997:5).

Saraceno, C (Ed) (2002), Social Assistance Dynamics in Europe: National and Local Poverty Regimes, Bristol: Policy Press (In print)

Stenberg, S - Å. (1998), "Unemployment and Economic Hardship - A Combined Macro- and Micro- Analysis of the Relationship Between Unemployment and Means-Tested Social Assistance in Sweden", European Sociological Review 14, 1-13.

Voges, W and Rohwer, G. (1992), "Receiving Social Assistance in Germany: Risk and Duration", Journal of European Social Policy 2, 175 - 191.

A Structural Model of Childcare, Welfare, and the Labor Supply of Single Mothers*

Thomas Andrén[α]

Abstract
This paper considers the simultaneous relationship of the single mother's decision to choose paid childcare, welfare participation and labor supply, and estimates a structural model that allows for a free error covariance. We use a discrete approach to the choice of labor supply together with the discrete choices of utilized paid childcare and welfare participation, which allow formulating the model as a multiple-choice problem. The results show that there is an association between social assistance, paid childcare, and labor supply, but that the relationship is non-symmetric. An increase in the social assistance norms has a relatively small effect on paid childcare utilization, but a relatively larger effect on the mean labor supply. In contrast, a corresponding reduction in the childcare cost has a relatively large effect on the social assistance utilization but a relatively small effect on the mean labor supply. Our estimates suggest that a decrease in childcare cost increases the labor supply of those working rather than encourages non-workers to start work, which implies that childcare cost is foremost a barrier to fulltime work rather then a barrier to work at all.

Keywords: labor supply, paid childcare, welfare participation, structural model, simulated maximum likelihood, Halton draws.

JEL classification: J13, J22.

* I would like to thank Daniela Andrén, Lennart Flood, Henry Ohlsson and participants at the Labor and Public Economic Workshop at Department of Economics in Göteborg for useful suggestions and comments. The usual disclaimer applies.
α Göteborg University, Department of Economics. E-mail: Thomas.Andren@economics.gu.se.

1 Introduction

For single mothers of young children, participation in the labor market is strongly linked with the need for childcare. Childcare is often regarded as essential for full-time work and career development of single mothers, but sometime the cost of childcare can be a barrier for entering the labor market when municipal childcare is the only option.[54] This implies that some working mothers end up below the social assistance norm after paying for the childcare fee. The childcare fee per se can therefore be seen as a factor that induces a need for income support, such as social assistance. This is especially true for single mothers located at the lower end of the wage distribution. It is therefore reasonable to expect that there is a behavioral relationship among the choices of paid childcare utilization, welfare participation and labor supply.

In Sweden, large savings have been imposed on the municipal childcare system during the last decade, resulting in changes in the childcare fees. The fee levels have increased and the construction of the fees has changed. For many single parent households, income is too low to support the family after the childcare fee has been paid. In 1996, a single mother with a gross income of 13,200 Swedish crowns (SEK) per month and with 2 children on childcare 40 hours per week, had a consumption level below the social assistance norm in over 90% of the municipalities (Socialstyrelsen, 1997).[55, 56] Recently, public attitudes towards single mothers have changed dramatically, and policy makers have been under pressure to carry out reforms that reduce single mothers' welfare dependency by means of fee ceilings (Maxtaxa). It is therefore interesting to investigate the simultaneous effect of paid childcare utilization and welfare participation on the labor supply for single mothers. Would a lower childcare fee increase labor supply and/or decrease welfare participation? Would a change in the social assistance norm affect the single mother's preferences related to the labor market activity? Would a lower social assistance norm lower the demand for paid childcare and therefore the labor supply? These are the main questions to be addressed in this paper.

[54] Childcare activities are usually operated by the municipality and therefore subject to a childcare fee.
[55] In 1996 the grant for a PhD student (utbildningsbidrag) in economics was 12,000 SEK per month in Sweden.
[56] In March 2002, 1 SEK corresponds to 0.11 Euro.

The literature is rich with studies that analyze the childcare cost or welfare effects on the labor supply for single mothers. There is also a growing econometric literature relating childcare costs to female employment, though the vast majority focus on married mothers. Heckman (1974), Blau and Robins (1988), Ribar (1992, 1995), Connelly (1989, 1992), Averet et al. (1997) and Kimmel (1998) have explored the effect of childcare costs on married women's labor force participation in the United States. Kimmel (1998) compared married and unmarried women and found single women's employment elasticity to be lower than married women's. All found significant negative effects of childcare costs on women's employment.

There are also a few studies that use Swedish data. For example, Gustafsson and Stafford (1988) estimated the effect of such costs on the market work decision of women belonging to two-parent families. Flood and Wahlberg (2000) estimate the labor supply effects from introducing fee ceilings (Maxtaxa) on households with children. They found a significant negative effect of childcare costs on women's labor force participation, although the estimated mean childcare price elasticity of employment varied extensively. Flood and Wahlberg (2000) also found that the labor supply effect very much depends on where on the income distribution the household is located.

The empirical literature on the effect of transfer programs of labor supply on women is fairly large (see Danziger et al., (1981), Moffitt (1992) and Gustafsson et al. (1993) for surveys). The literature on the incentive effects of the U.S. welfare systems has shown unequivocal evidence of effect on labor supply and participation in the welfare system. These effects arise mostly for single women, which is a major recipient group. The econometric studies show that labor supply is reduced by welfare. Flood et al. (2001) analysed single mothers in Sweden and found rather small mean incentive effects. However, they found substantial effects for different income groups.

The literature on childcare costs and welfare participation contains only a few studies. Fore example, Connelly (1990) found small effects of childcare costs on welfare use. Kimmel (1995) using a low-income sub-sample found nearly zero elasticity. Crecelius and Lin (2000) found a one-percentage point reduction in the average probability of welfare receipt, if mothers received 20 hours of help weekly from relatives and friends. They also found that for each 10-cent reduction in childcare cost,

there were 0.15 to 0.21 more hours worked per week. Connelly (2001) found significant results indicating that subsidizing childcare reduces the welfare dependency of single mothers. To my knowledge there are no studies using Swedish data on the simultaneous effect of childcare cost, welfare participation and labor supply of single mothers.

There are at least two reasons for the lack of studies concerning the simultaneous effect. The first is the existence of self-selection into welfare programs and/or paid childcare utilization on the basis of unobserved heterogeneity. Some people are more likely to be eligible for social assistance than others, and the decision to utilize paid childcare is by no mean random. There are also individuals eligible for social assistance who choose not to participate. The latter are often referred to as welfare stigma (Moffitt 1983). This selection process, based on unobserved heterogeneity, requires that the welfare-participation equation and the paid childcare utilization equation be estimated jointly with the labor supply equation. This is a task that requires evaluation of high dimensional integrals if the choice set is large. Until recently, this has not been feasible since standard quadrature methods are very burdensome when integrating over high-dimensional density functions.[57] Instead, estimation methods based on simulation techniques may be used.

The second reason is the difficulty of deriving an analytical solution of the boundaries of the error space within which different choices are optimal.[58] This also makes it difficult to use high precision simulation methods such as the GHK–simulator since it requires that the residuals be expressed explicitly.[59] Instead, a frequency simulator (Lerman and Manski, 1981) may be used, but at a cost of many more random draws from the assumed distribution. The number of draws can be reduced dramatically using low discrepancy methods when evaluating the probabilities, but the method is still very computer intensive, which therefore only reduces the problem.

By assuming a discrete approximation of the hours of work distribution, we may formulate the model as a multiple-choice problem. We approximate the distribution of weekly hours of work by three distinct points: 0, 20, and 40, where H = 0, if the woman

[57] The number of evaluation points grow exponentially with the dimension which means that dimensions above 5 would be too burdensome since the integral has to be evaluated for each individual, several times during just one likelihood iteration.
[58] See Keane and Moffitt (1998) for a discussion.
[59] The GHK-simulator was developed by Geweke (1991), Hajivassiliou (1990), and Keane (1994), who gave the simulator its name.

does not work; H = 20, if she works part time, and H = 40, if she works full time. This approximating assumption simplifies the estimation dramatically and circumvents the problems with kinks and piecewise linear budget sections due to transfer programs, taxes and childcare fees. This is especially convenient since we are interested in investigating the factors behind the combined decision of labor supply, welfare participation and paid childcare utilization.

Our study differs from those published earlier in several respects. First, we estimate a structural labor supply model allowing for a free error structure among the involved equations, including paid childcare utilization and welfare participation. The model is structural in that we define an explicit expression for the preferences, which is used in the estimation of the parameters in the model. Second, depending on the hours of work, we allow for several different care modes. Third, we include a wage equation that is allowed to freely correlate with the other equations.[60] This is important since a substantial percentage of single mothers have unobserved wages, which therefore have to be simulated conditional on the wage equation. Fourth, fixed cost of work is an important factor for single mothers and is incorporated in the utility structure using a linear function to control for observed heterogeneity. Fifth, we perform a policy analysis using the full structural model in a micro simulation setting that provides responses on the choice variables given various changes in the budget set. Moreover, we focus on a group of individuals sensitive to changes in childcare fees and welfare norms, namely single mothers having at least 1 child aged 1-12 years.[61] The single mother's choice problem is to choose among the discrete hours, welfare participation, and paid childcare utilization, treating all choices as endogenous when the decision is taken.

The paper proceeds with Section 2, which presents the institutional background. The data is presented in Section 3. Section 4 presents the basic theoretical set-up, and Section 5 specifies and motivates the construction of the empirical model. Section 6 presents and discusses the results and Section 7 summarizes the main findings of the paper.

[60] To include a wage equation in a labor supply model is by no mean a novelty, but it is an important feature of this model, since we account for unobserved wages in a theoretically proper way.

[61] It is an empirical fact that almost all single mothers stay at home with the child during the first year. Hence we exclude households with children younger then 1 year, since paid childcare won't be an option for them.

2 Institutional background for paid childcare

According to the Education Act of 1995, municipalities are obliged to provide childcare in the form of pre-school activity and school childcare for children aged 1-12 of the extent required for parents to be able to work or study. The act states that placement ought to be provided "without unreasonable delay", i.e., normally within three to four months after the person having custody of the child has requested placement. However, if a parent is unemployed, the child usually loses the childcare placement at once or after a period of time depending on the municipality. According to a survey done by the Swedish Board of Education (Skolverket, 1998), almost half of the municipalities allowed the child retain the placement even though a parent became unemployed, while a child lost the placement in about 40 % of the municipalities. In only 28% of the municipalities were there no requirements put on the parents to be employed or have student status.

There are two types of municipal childcare in Sweden: pre-school childcare activity and school childcare. Pre-school childcare is intended for children who do not attend school and is carried out in the forms of pre-schools, family day-care homes and open pre-schools. School childcare is intended for children who attend school and is carried out in the form of after-school center, family day-care home and open after-school activities. In 1997 about 59 % of children aged 1-2 were enrolled in pre-school activity, while the corresponding number was 78% for children aged 3-6. The largest proportion of children (97%) participates the year before the children start school. The participation rate for children aged 7-9 was 56% and the number for children aged 10-12 was 6%.

In autumn 1999, the average hours per week at municipal pre-school for children aged 1-5 years were 31. Hours per week at private pre-school were on average an hour a week more, and at family day-care homes an hour a week less. Schoolchildren's average hours per week were just below 17 hours per week both in municipal after-school centers and family day care homes and an hour shorter in private leisure time centers. The spread is relatively great, as for example, the hours per week of a third of pre-school children were fewer than 20 or more than 40 hours per week. Hours per week are highest in big cities and suburban municipalities and lowest in industrial and rural municipalities.

The fees charged to parents finance an increased proportion of the gross expenditure on childcare. In 1999, approximately 18% of the costs for municipal childcare were met by parental payments, while in the early 1990s, this proportion was around 10%. The percentage of such financing is highest at the after-school centres, where charges to parents account for almost a quarter of the gross expenditure, compared with 16% at pre-school.

3 The Labor supply, childcare and welfare participation

In standard labor supply models, the utility of an individual is specified as a function of hours of leisure (L) and net income (Y), expressed as $U(L,Y)$. This function is equivalent to one with hours of work (H) as an argument, such as $U(H,Y)$. Expressing the utility function in terms of hours or work instead of leisure avoids the necessity of explicitly representing the total hours available to market and non-market activities. The hours of work is therefore defined as $T - L$, where T is the total time available, and L the non-market hours. The non-market hours are defined to include maternal childcare, other household production activity, and leisure. In these models, people are assumed to like income but to dislike working so that U increases in Y and decreases in H. In static labor supply models, people are assumed to maximize the current period utility subject to a current period budget constraint of the form $Y(H) = wH + N$, where w is the hourly wage rate and N the non-labour income. When paid childcare utilization and welfare participation is incorporated into the utility structure, the function is extended by a measure for paid childcare utilization (P_{CC}) and welfare participation (P_{SA}). The important question is how these two components work in the preferences of the individual; this is important when interpreting the parameters in the utility function. The factor that captures the utility of social assistance is assumed to be negative. [62]

We aim to analyze the structure and determinants of paid childcare utilization as oppose to non-paid childcare, and to what extent paid childcare constitutes a barrier to labor supply. We therefore follow and use Ribar (1995)'s conceptual framework regarding paid childcare, emphasizing the link between the quality of childcare and the choice of paid childcare utilization. He argues that it is the quality of childcare (Q)

[62] The reasons for this will be discussed and motivated when we discuss the empirical specification.

extended to the children that affects the preferences of the mother, and is therefore a factor in the utility function, $U(H,Y,Q(X_Q,P_{CC}),P_{SA})$. In this framework, paid childcare is only one of several components inside the production function of childcare quality. Childcare quality is assumed to increase with inputs of maternal care and market goods included in X_Q, but the contribution of paid childcare to the overall quality is ambiguous and depends on the quality of paid childcare relative to maternal and unpaid non-maternal childcare. Paid childcare enters the model as an input to the production of childcare quality but also as a cost component in the budget constraint. Unpaid childcare, on the other hand enters the model only as an input to the production of quality. The decision to use paid or unpaid childcare therefore depends only on the absolute cost and relative quality of paid childcare. It is therefore assumed that non-market care is available to all families at some indirect cost, which should be incorporated into the model. The indirect cost of unpaid care is determined in terms of the care provider's time in alternative activities. This provides a measure of the availability of unpaid care. The indirect cost is incorporated directly into the utility function by reformulating the utility function as

$$U = U(H,Y,Q,P_C,P_{SA}) \tag{1}$$

where utility increases with childcare quality (Q) and paid care utilization (P_C). The utility term (P_{CC}) acts as a flexible proxy for the indirect cost of unpaid care (Ribar, 1995).

For the labor supply of single mothers with young children, the budget constraint for the simple labor supply model must be elaborated. Most importantly, since a large proportion of single mothers have a strained economic situation, the budget constraint must include the available welfare benefits. The major welfare programs for single mothers are social assistance, housing allowance, child allowance, and alimony, which are all designed to reallocate economic resources to economically weak groups such as single mothers. This reallocation process affects the single mother's labor supply behavior in different ways, and needs to be included in the model in order to capture those effects. Furthermore, single mothers that participate in the labor market also need childcare, which often is associated with substantial costs. Therefore, the childcare cost

has to be included in the budget constraint in order to account for the importance of the cost of childcare on the single mother's preferences towards the endogenous variables in the model.

P_{SA} is an indicator that takes the value 1 if a single mother uses social assistance and 0 otherwise, while P_{CC} is the corresponding indicator for the decision to put her child in paid childcare. The budget constraint may then take the following form:

$$Y(H, P_{SA}, P_{CC}) = wH + N - P_{CC}B_{CC}(X_{CC}) + P_{SA}B_{SA}(X_{SA}) - T(X_T) \qquad (2)$$

$B_{CC}(X_{CC})$ represents the function for potential childcare cost with a vector of conditioning factors as argument (X_{CC}), where hours of work is one such factor. As the hours of work increase, the demand for more hours of childcare increase, causing potential childcare cost to increase. Hence, the childcare cost function increases in H.[63] $B_{SA}(X_{SA})$ is the corresponding function for potential social assistance with a vector of conditioning factors as arguments (X_{SA}), which also contains hours of work. The more she works, the higher the disposable income, which reduces the potential amount of welfare she could receive. Hence, potential welfare decreases with hours of work. $T(X_T)$ represents the tax function for the individual, which increases with H.[64,65] The other welfare programs are included in N as non-labor income.

4 The empirical specification

To make the model concrete we need to specify a functional form for the utility function. A convenient way to model preferences of an individual is to use the quadratic direct utility function with household net income (Y), individual labor supply (H), paid childcare utilization (P_{CC}), and welfare participation (P_{SA}) as arguments:

$$U(H, Y, P_{CC}, P_{SA}) = \beta_H H + \beta_{HH} H^2 + \beta_Y Y + \beta_{YY} Y^2 + \beta_{HY} HY + \beta_C P_{CC} - \beta_{SA} P_{SA} \qquad (3)$$

[63] B_{CC} is dependent on the age of the children, the number of children and the size of the gross household income. The system applied in this paper is described in the Appendix.
[64] B_{SA} is a function of several factors described in the Appendix, one of them being the level of housing allowance. If the housing allowance increases, the level of social assistance decreases with the same amount. Here we have a 100 % marginal effect.
[65] The income tax system used here is described in the appendix.

Y is defined by the budget constraint given by (2), and H is a discrete measure for hours of work. Specification (3) should be interpreted as a flexible approximation to (1), in which the coefficients represent combinations of utility and care quality production parameters (Ribar, 1995). The quadratic direct utility function has the disadvantage of being concave only when its arguments are below the amounts at which the utility peaks, but this is not a serious problem as long as the data falls into the concave range (Fraker and Moffitt, 1988). The quadratic utility function is simple and convenient to use, which is the main advantage for our purpose. It is also flexible enough to allow for backward bending labor supply behavior. The marginal utility of Y at $Y = H = 0$ is normalized to 1, i.e., β_Y=1, which means that the remaining parameters are expressed in terms of SEK. As mentioned above, having the utility being a function of hours of work rather then hours of leisure is arbitrary, but it affects the interpretation of the parameters, which here is a measure for hours of work rather then leisure. In the basic labor supply model it is assumed that utility decreases with H, which implies that β_H may be interpreted as a disutility factor in the preferences of the individual. No such assumption will be made here since it is more realistic to believe that utility increases in H with a decreasing pace. The sign of β_H will therefore be an empirical question. According to our earlier discussion the sign of the coefficient for the paid childcare indicator, β_C, is positive. Since quality of childcare is exogenously given, the parameter is a measure for paid childcare utilization compared to all other forms of unpaid childcare. If paid childcare in general offers higher relative quality compared to unpaid childcare, the sign of the coefficient will be positive.[66] However, the sign of the parameter is here also an empirical question. The interpretation of the parameter is in terms of utility and measures the marginal utility of paid childcare. Paid childcare also has an income effect, since childcare cost is a component in the budget constraint. The decision to use paid childcare is therefore in some sense restricted by the budget set. The coefficients β_{HH}, $\beta_{YY,}$ and β_{HY} have no clear interpretations, but play an important role in the determination of the labor supply elasticities.

[66] Many arguments about the importance of having the child on public childcare stress that this is not just a labor supply issue for the single parent but also an educational issue for the intellectual development of the child.

So far, we have not discussed the role of social assistance in the preferences. A problem that appears when estimating labor supply models that account for welfare participation is the well-known selection problem that appears when individuals who are eligible to participate choose not to. This welfare stigma must be accounted for if we are to receive consistent estimates. It is irrelevant whether the choice not to participate is truly stigmatic, or whether there are other reasons for it. The fact still exists that some people do not end up on the budget line, which means that those people prefer or choose a utility level lower than the maximum attainable. This must be taken into account if we assume that individuals are utility maximizers. This also makes it difficult to interpret parameter as a stigma effect, since it does not reflect truly stigmatic behavior, but rather a combination of non-stigmatic factors such as lack of information or too little gain from participation. It is therefore more reasonable to think of the parameter as representing a threshold cost of taking social assistance.

There are two different ways in which this disutility can manifest itself. First, there may be a flat component that arises from the mere fact of participation itself, but which does not vary with the size of the benefit received. Second, there may be a variable component that varies with the size of the benefit. In this paper we choose to model it as a flat component for social assistance. In order to adjust for the non-presence of eligible single mothers the flat component is incorporated additively as a negative component ($-\beta_{SA}P_{SA}$) in the utility function (3). Hence, if β_{SA} is sufficiently large, the program may not be chosen even thought participants increase the utility. β_{SA} is therefore a "cost" parameter or simply the marginal disutility of welfare, and P_{SA} the discrete indicator for whether or not a person participates in the welfare program. Maximizing the utility function (3) with respect to the endogenous variables in the model subject to the budget constraint gives us the functions that represent the single mother's equilibrium combination of hours of work, welfare participation and paid childcare utilization.

4.1 Preference Heterogeneity

Observed heterogeneity is introduced linearly through parameters $\alpha=\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$, that are allowed to vary in the population conditional on a set of observable socio-economic characteristics. In order to estimate the model, we must also specify a

stochastic structure. That is, since single mothers with identical observed characteristics make different decisions about labor supply, welfare and paid childcare, we must allow for random influences on these decisions. We therefore define the following equations:

$\beta_H = X_1\alpha_1 + \varepsilon_H$	Marginal utility of work	(4)
$\beta_{SA} = X_2\alpha_2 + \varepsilon_{SA}$	Marginal disutility of welfare	(5)
$\beta_C = X_3\alpha_3 + \varepsilon_C$	Marginal utility of paid childcare	(6)
$\text{Log}(w) = X_4\alpha_4 + \varepsilon_w$	The wage equation	(7)

where $\varepsilon = (\varepsilon_H, \varepsilon_{SA}, \varepsilon_C, \varepsilon_w)$ captures unobserved heterogeneity, and is assumed to be distributed jointly normal with mean zero and covariance Σ_ε.

4.2 Fixed entry cost of work

In general, it is reasonable to believe that there is an entry cost associated with work for any individual moving from the non-working state to the working state (Cogan, 1981). This is especially true for single mothers who need childcare in order to participate in the labor market. Entry cost of work is usually decomposed into two parts, namely: money cost, and the time cost of work. Both are of special importance when modeling single parent households with young children. Money cost is usually related to childcare costs, commuting costs or any other costs associated with work that are paid for by the individual. For single mothers, the ages of her children and the number of children are important factors that determine the major part of the money cost of work. Time cost of work is typically the time it takes to transport the child/children to a day-care center or the time it takes to commute to the working place, and therefore is directly related to the distance to the day-care center and work. A longer distance is associated with a higher cost since leisure time will be reduced by the same amount.

How does the fixed cost of work affect the single mother's labor supply? If money cost of work increases, the likelihood of an unemployed single mother waiting to supply labor decreases, but if she is employed, such an increase will increase labor supply and reduce leisure time. On the other hand, a rise in time cost to work has two consequences. First, it reduces "full income", that is, the maximum earning power of the

single mother. Second, the increase in time cost of work also reduces total time available for either leisure or work, as long as the single mother continues to work. If consumption and leisure are normal goods, and "full income" decreases, the single mother will reduce both consumption and leisure.

In this paper we express fixed cost of work by a linear function, allowing for observed heterogeneity: $Fc = X_5\alpha_5$, where X_5 is a matrix of explanatory variables and α_5 a corresponding vector of parameters. Fc will be used to reduce the net income for those who are working, by replacing Y by $(Y - Fc)$ in the utility function. The variables in such a function would be presence of children in different age categories, and the use of paid childcare, as well as the residence location. If a single mother lives in a city region, her transportation cost might be different from those who live in the country.

4.3 Estimation

In order to estimate the structural labor supply model we formulate a multiple-choice problem with 3 choices of hours of work, 2 welfare choices, and 2 childcare choices, which add up to a total choice set of 12 different alternatives. Letting j = 1, …, 12 indices the alternatives in the choice set, the problem is to choose the alternative generating the highest utility level. That is: an individual chooses alternative j if and only if $U_j > U_k \;\; for\ all\ k = 1, ..., 12, k \neq j$, where U_j denotes the evaluation of the stigma adjusted utility function for combination j obtained by inserting the budget constraint evaluated by setting H, P_{SA} and P_{CC} at their appropriate values for combination j.

To formulate the likelihood function, we need to determine the contributing probabilities that correspond to each alternative (utility level). Given the nature of the problem, we can only express differences between utility functions rather then being able to extract the residuals explicitly. This requires the use of a method that does not require such error bounds. The standard frequency simulator (Lerman and Manski, 1981) does not have such requirements and therefore makes the problem solvable. The standard frequency simulator has certain problems, which make it intractable. It is a step function, which therefore excludes the possibility of using gradient-based optimization methods. It may also happen that the estimated probabilities sum to a number different

from one. McFadden (1989) proposed a remedy for these two problems, suggesting a Kernel-Smoothed frequency simulator to handle them.[67] The basic idea behind this simulator is to add "noise" in a systematic way to the simulated choice probabilities, thereby making the probability space continuous. The method is based on the extreme value distribution function that is used as a kernel for the frequency simulator, which also by necessity imposes the sum-up criterion.

Let $P(j|\theta, X, w)$ be the likelihood contributing probability of the event that the individual chooses alternative j conditional on a parameter vector (θ), observed characteristics (X) and the wage rate (w), and ϕ being the normal p.d.f. If δ_{ij} is an indicator that is 1 if person i chooses alternative j and 0 otherwise, we have the following log likelihood function:

$$L(\theta, \Sigma_{\varepsilon} \mid X, w) = \sum_{i=1}^{n} \sum_{j=1}^{12} \delta_{ij} \ln\left[P(j \mid \theta, X_i, w_i)\phi(w_i \mid \theta, X_i)\right] \tag{8}$$

A critical assumption here concerns the observability of gross wages for all single mothers.[68] Since wage rate is unobserved for non-workers, it is important to specify an equation for the wage offer determination process so that the unobserved wages of non-workers can be inferred. According to economic theory, a person works as long as the wage offer exceeds the reservation wages of the individual. Hence, in all cases of non-workers, we know that the individual has been offered a wage rate that is below her reservation wage. By introducing a distribution for the wage offer we are able to average over the section of the distribution that is relevant for the unemployed single mother, and thereby use the resulting probability as the contribution to the likelihood in

[67] The kernel smoothed frequency simulator defines the choice probability as $P(choose\ k \mid \theta) = \exp(U_k(\theta)/\sigma) / \sum_{j=1}^{J} \exp(U_j(\theta)/\sigma)$, and the choice probability is received by averaging over repeated draws from the assumed distribution for the unobserved components of the utility function (McFadden, 1989). This simulator converges to the standard frequency simulator when σ goes to zero. Choosing a suitable σ is time-consuming. It should be as small as possible, but when choosing it too small the convergence property of the simulator gets worse, which makes it harder for the model to converge. We use $\sigma = 0.5$ in the estimation.

[68] See Van Soest (1995) for a discussion.

the estimation. We can do this by simulating wages for each individual over the truncated region of the wage offer distribution and then average out.

$$P(j \mid \theta, X) = \int P(j \mid \theta, X, w)\phi(w \mid \theta, X)dw \tag{9}$$

$P(j \mid \theta, X)$ is the probability of participation–hour combination j conditional on a vector of observed characteristics for the individual (X), and a vector of parameters of the model (θ), for those with unobserved wages. The likelihood function is adjusted and modified by replacing the likelihood contributing components in (8) by the logarithmic version of (9) for those with unobserved wages.

A way to reduce the computational burden is to use low discrepancy methods when simulating the probabilities. One such method is the so-called Halton sequence (Halton, 1960), which is a commonly used low discrepancy sequence that is much more efficient then standard random draws.[69]

4.4 Identification

When estimating the model we apply some normalizations. For example, the variances of the two choice parameters (β_{SA},β_C) are normalized to one. These normalizations are not used as identifying normalizations. We impose them to make the model more stable. The parameter for net income (β_Y) in the utility function is also normalized to one, which even here is by no mean necessary from an identification point of view. We impose it in the spirit of making the parameters more interpretable. The variance of β_H is identified since we have three H categories rather then two. We use exclusion constraints between the labor supply equation and the wage equation; these include variables in the labor supply equation that are not in the wage equation, and variables in the wage equation that are not in the labor supply equation. We use different sets of variable specifications for the welfare participation equation and the paid childcare equation to make the estimation more stable, but also since we believe that these two

[69] Train (2001) indicates that 100 Halton draws can outperform 1000 standard random draws. This implies that the computational burden can be reduced by a factor of 10 using low discrepancy methods.

choices to some extent are explained by different factors. No other restrictions or normalizations have been applied.

5 Data

The data used in the empirical analysis belong to the Swedish Household Income Survey (HINK) provided by Statistics Sweden. HINK provides information on labor market activities and incomes for a random sample of about 20,000 Swedish households per year, with information about each member of the household. It is comprised of two sections; a survey and a register. This construction provides us with relatively high quality income data, as well as with useful variables describing the household. However, the single mothers form a rather small group of the total Swedish population, and therefore the sample size of a cross-section is not big enough. To remedy this we decided to pool the 1997 and 1998 cross-sections of HINK. From each cross-section, we selected single women aged between 18 and 64 that have at least one child in the ages 1-12. We excluded women who were self-employed, early retired, and students. Having done this we ended up with a sample of 533 single mothers.

Hours of work and wage rate are important variables in any labor supply study. In the estimation, we use a discrete approximation for the continuous measure of hours of work. We use 3 discrete hour points $H = \{0, 20, 40\}$. In HINK the continuous variable for hours of work is derived with information from the survey section, which asked for the number of weeks of gainful employment (K_1) and for the number of worked hours per week with gainful employment (K_2). These two variables were used to construct the degree of employment as a percentage of normal working time. Normal full-time work is defined to be 1880 hours per year, assuming 5 weeks of vacation, and 40 hours of work per week (40*47=1880). This leads to the following formula $S = (K_1 * K_2) * 100 / 1880$, where S is used to construct the discrete time points used in the analysis.

$$H = \begin{cases} 0 & if \quad S = 0 \\ 20 & if \quad 0 < S \leq 88 \\ 40 & if \quad S > 88 \end{cases}$$

The wage rate is computed as the ratio of annual income from work, and annual hours of work.[70] This construction could lead to division bias in the case of measurement errors in annual hours of work. This is something that we do not consider and therefore implicitly assume that hours of work are measured with no errors.

The variable for paid childcare utilization is based on observed data. For each child in the household there is information about attendance of paid childcare during a year. If at least one child has been in paid childcare for at least one month, the household is registered as a paid childcare user. In general a single parent is not entitled to municipal childcare when unemployed, though, there are exceptions. If someone uses municipal childcare and becomes unemployed, it is sometimes possible to keep the child in the childcare center. However, since childcare cost is a function of hours of work, the cost for such children is 0. In the analyzed sample, there are a few cases were the wage rate is unobserved while the single mother uses paid childcare. For these households the childcare is free of charge.

Since the childcare fee varies by hours of work, it is necessary to specify a time use schedule for the households. This is important since we need to determine the potential childcare cost for different hours of work combinations during the estimation. Table 1 reports the assumed time use for the households in the model. It is assumed that a pre-school child stays at a childcare center as long as the mother is working. It is assumed that a school-aged child spends half the time at school, and the rest of the working day at the care center. This schedule is applied for those with observed childcare, and is considered a good approximation of the actual cases since we have information whether a particular child is within a childcare program.

[70] Each individual is asked how many hours she has been working each and every month a specific year. By summing the number of hours we receive a measure for total numbers of hours worked during the year. A variable from the tax register gives us the annual income from work. The ratio of these two gives us the wage rate.

Table 1 Time use of paid childcare in hours per week for the household [71]

Hours of work	Age of child		
	1 – 5	6 – 8	9 – 12
H = 0	0	0	0
H = 20	20	10	10
H = 40	40	20	20

The variable for social assistance utilization is a discrete indicator for those who received *any* social assistance during the year. Table 2 presents the distribution of the number of months with social assistance the single mothers received over the year, stratified over the discrete points of hours of work. We observe that the behavior differs, depending on whether the single mother is employed or not. If unemployed, the number of months of welfare participation is uniformly distributed, while if employed, just one or two months are used. In general, single mothers seem to take social assistance just a few months per year. This behavior is nothing that we control for since we define a welfare participant as someone who has received social assistance for at least one month during the year.

Table 2 Distribution of welfare dependency

Number of months	Hours of work					
	0		20		40	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
1	7	10.14	13	18.31	10	45.45
2	5	7.25	11	15.49	6	27.27
3	2	2.90	3	4.23	1	4.55
4	8	11.59	3	4.23	1	4.55
5	5	7.25	9	12.68	0	0
6	8	11.59	1	1.41	1	4.55
7	1	1.45	4	5.63	0	0
8	8	11.59	4	5.63	0	0
9	7	10.14	7	9.86	1	4.55
10	5	7.25	9	12.68	1	4.55
11	6	8.70	3	4.23	0	0
12	7	10.14	4	5.63	1	4.55
	69	100.00	71	100.00	22	100.00

[71] Lately, it has been common practice to link the childcare fee to the number of hours the child stays at the childcare centre. In 1999 almost all municipalities (90%) used a construction with a time-varying fee. In 1993 the corresponding number was 40% (Skolverket, 2000).

When choosing variables for the equations, the aim has been to have a specification as parsimonious as possible, yet still including what is relevant. The labor supply equation contains standard variables such as age, region, and education. Since we model paid childcare utilization, we also included dummies controlling for presence of children in different age intervals, which can affect the labor supply. The wage equation also contains standard choices namely region, education and years of work experience.

The social assistance equation should contain variables that are associated with social assistance utilization, such as age, region, education, and types of childcare. Younger people are more often exposed to social assistance then others. Educational level is typically related to how successful an individual is in the labor market, which in tern is associated with the need for social assistance. The link between paid childcare utilization, social assistance is explored in this paper, and we therefore include a dummy for this. We also include a corresponding dummy indicating if the child was at home with a parent as oppose to being in paid childcare.

The childcare equation measures the marginal utility of using paid childcare. It should therefore contain variables that capture such effects. Since the attendance duration differs across regions, regional dummies were included. The educational level of the parent is associated with the need for childcare. People with higher educations tend to work more than those with lower education. One could therefore argue that people with higher education potentially demand more childcare. The need for childcare also varies dependent on the age of the child. When children are younger, the need is greater, compared to the case when children are older and therefore more able to take care of themselves.

Table 3 presents sample means of variables used in the estimation, stratified into various groups dependent on hours of work, the choice of utilized paid childcare and welfare. The overall average age for a single mother is 35. Single mothers working part-time, using paid childcare and receiving social assistance are the youngest (about 30 years old), while single mothers working full-time using neither paid childcare nor social assistance have an average age of 39. This implies that social assistance is something that younger single mothers receive before they have established themselves on the labor market. On the other hand, we observe that older single mothers also have

higher educations and live in big city regions. The number of children per household is quite stable, ranging from 1.73 to 1.93. Younger mothers receiving social assistance and childcare more often have young children, while older mothers not using childcare or welfare more often have children aged 9-12.

Table 3 Variable means of household characteristics (standard deviations within parentheses for continuous variables)

Variable	*All*	H=0 SA=1	H = 0 SA=0	H = 20 CC = 1 SA = 1	H = 40 CC = 1 SA = 1	H = 20 CC = 1 SA = 0	H = 40 CC = 1 SA = 0	H = 20 CC = 0 SA = 0	H = 40 CC=0 SA=0
Age	35			30	33	35	37	36	39
	(7.4)	32	35	(6.3)	(7.2)	(6.6)	(6.8)	(6.5)	(6.1)
Age groups									
18-34 years	0.47	0.61	0,56	0.76	0.60	0.50	0.38	0.39	0.22
35-44 years	0.41	0.37	0,27	0.21	0.33	0.44	0.46	0.49	0.57
45-64 years	0.12	0.02	0,18	0.03	0.07	0.06	0.16	0.12	0.21
Experience	8.9	3.7	5,9	7.4	11.8	12.6	16.2	6.3	10.1
	(9.3)	(6.6)	(7,1)	(6.1)	(8.0)	(6.6)	(6.8)	(9.3)	(11.7)
Education									
Primary school	0.34	0.59	0,31	0.52	0.27	0.24	0.27	0.29	0.20
Secondary school	0.57	0.37	0,67	0.48	0.66	0.61	0.59	0.63	0.67
Post secondary	0.09	0.04	0,02	0.00	0.07	0.15	0.14	0.08	0.13
Number of children	1.78	1.90	1,56	1.88	1.93	1.88	1.80	1.73	1.79
	(0.84)	(1.07)	(0,78)	(0.94)	(0.96)	(0.85)	(0.79)	(0.81)	(0.71)
If children aged									
1- 5 years	0.43	0.54	0,42	0.76	0.67	0.52	0.49	0.24	0.14
6- 8 years	0.38	0.42	0,33	0.41	0.47	0.47	0.47	0.31	0.26
9-12 years	0.49	0.42	0,56	0.29	0.27	0.44	0.38	0.64	0.74
13-17 years	0.24	0.24	0,16	0.12	0.33	0.13	0.24	0.25	0.47
If paid childcare for children aged									
1- 5 years	0.28	0.31	0,24	0.74	0.67	0.52	0.49	-	-
6- 8 years	0.22	0.18	0,16	0.41	0.47	0.47	0.46	-	-
9-12 years	0.11	0.03	0,09	0.15	0.00	0.32	0.29	-	-
If home with parent for children aged									
1- 5 years	0.05	0.07	0,02	0.09	0.13	0.02	0.05	0.05	0.03
6- 8 years	0.03	0.03	0,04	0.03	0.00	0.05	0.01	0.02	0.01
9-12 years	0.25	0.18	0,31	0.18	0.27	0.18	0.15	0.35	0.43
Region groups									
Big city	0.43	0.31	0,36	0.30	0.40	0.32	0.47	0.53	0.54
Small city	0.30	0.32	0,27	0.41	0.20	0.37	0.30	0.20	0.30
Rural	0.27	0.37	0,38	0.29	0.40	0.31	0.23	0.27	0.16
N	533	71	45	34	15	62	100	83	76

Note: CC=paid childcare. SA=social assistance.

Table 4 reports mean statistics of variables that appear in the budget restriction of the household. Hours of work are expressed in annual numbers, and averages 1151 hours by individuals in the sample. The overall hourly wage rate was 73 SEK per hour. The wage rate for women who work part time, use paid childcare and no social

assistance, was 102, and for women who work full time, use social assistance and have children at paid childcare was 72. This latter group is very special, but also very small in our sample; only 15 individuals. Therefore, these numbers should be interpreted with certain reservations. This group also works most on average; 2153 working hours per year.

Table 4 Variable means of household budget components (standard deviation within parentheses)

Variable	*All*	H=0 SA=1	H = 0 SA=0	H = 20 CC = 1 SA = 1	H = 40 CC = 1 SA = 1	H = 20 CC = 1 SA = 0	H = 40 CC = 1 SA = 0	H = 20 CC = 0 SA = 0	H = 40 CC=0 SA=0
Hours of work, year	1151			600	2153	1051	2024	1068	2145
	(893)	-	-	(517)	(796)	(476)	(270)	(474)	(428)
Hourly wage rate (SEK)	73			82	72	102	97	92	96
	(47)	-	-	(51)	(29)	(44)	(26)	(22)	(15)
Monthly (potential) care	577			1243	1753	1196	1651		
expenditure (SEK)	(796)	-	-	(384)	(601)	(432)	(587)	-	-
Weekly potential social	318	1170		1054	663				
assistance (SEK)	(593)	(626)	-	(783)	(464)	-	-	-	-
Parental Allowance	5.61	0,63	0,55	6.29	12.49	7.98	6.68	6.39	5.45
(KSEK)	(13.26)	(1,60)	(1,42)	(7.33)	(21.05)	(17.32)	(9.58)	(17.41)	(14.99)
Child Allowance	14.87	16,55	12,79	16.10	15.68	16.21	14.75	14.81	14.19
(KSEK)	(8.62)	(11,12)	(7,45)	(12.23)	(8.34)	(8.70)	(8.32)	(7.80)	(6.97)
Housing allowance	16.76	21,85	19,61	20.40	15.79	18.85	11.13	17.82	11.36
(KSEK)	(11.59)	(11,29)	(12,38)	(10.83)	(9.88)	(11.28)	(10.11)	(10.90)	(11.09)
Alimony (KSEK)	16.51	20,11	16,86	16.85	18.71	18.23	14.94	16.59	14.08
	(14.75)	(17,95)	(14,14)	(14.79)	(16.14)	(13.88)	(13.54)	(15.53)	(13.99)
Sickness benefit (KSEK)	4.68	0,41	0,09	3.66	3.45	8.45	3.99	7.97	3.52
	(15.89)	(1,43)	(0,39)	(9.94)	(9.92)	(21.58)	(12.65)	(20.02)	(15.22)
Unemployment comp.	12.11	12,64	36,21	18.34	3.13	13.02	2.79	20.44	0.24
(KSEK)	(25.98)	(24,93)	(47,12)	(28.45)	(10.19)	(20.26)	(10.64)	(31.51)	(1.82)
Income from capital	0.53	0,02	0,63	0.01	0.02	0.45	1.74	0.21	0.42
(KSEK)	(3.90)	(0,07)	(1,88)	(0.02)	(0.07)	(1.53)	(8.63)	(0.88)	(1.42)
Housing rent per month	2.72	2,75	2,68	4.49	4.68	3.59	3.94	1.28	1.94
(KSEK)	(2.44)	(2,35)	(2,57)	(1.02)	(1.21)	(1.94)	(2.07)	(2.09)	(2.64)
N	533	71	45	34	15	62	100	83	76

Note: KSEK means thousands of SEK.

Table 5 reports how the endogenous variables in the model are distributed over the sample. Most single mothers (41%) work part-time, and 22% of them are unemployed. Across the choices, the smallest group of single mothers (1.5%) are full-time workers with no childcare but with social assistance, and the largest group (19%) are also full-time workers with childcare but with no social assistance.

Table 5 Distribution of women across welfare, childcare and hours of work (in %)

Welfare and childcare utilization		Hours of work			Row total
		0	20	40	
P_{CC}=1	P_{SA}=1	5.6	6.4	2.8	14.8
P_{CC}=0	P_{SA}=1	7.7	7.3	1.5	16.5
P_{CC}=1	P_{SA}=0	3.2	11.6	18.8	33.6
P_{CC}=0	P_{SA}=0	5.2	15.6	14.3	35.1
Column total		21.7	40.9	37.4	100.0

Table 6 reports the distribution of the choice combinations by education. Most single mothers have high school degrees and do not receive social assistance, while the less represented have post secondary education and receive social assistance. The distribution across the choice combinations is uniformly distributed among those with a primary education, and no clear pattern can be distinguished.

Table 6 Distribution of labor supply, welfare, and social assistance utilization by education

Welfare and childcare utilization		Education			Row total
		Primary	Secondary	Post Secondary	
P_{CC}=1	P_{SA}=1	7.3	6.9	0,5	14.8
P_{CC}=0	P_{SA}=1	9.0	7.1	0,4	16.5
P_{CC}=1	P_{SA}=0	8.4	20.8	4,3	33.6
P_{CC}=0	P_{SA}=0	9.4	22.3	3,4	35.1
Column total		34.1	57.2	8.6	100.0
H = 0		10.5	10.5	0.8	21.8
H = 20		14.6	23.1	3.1	40.9
H = 40		9.0	23.6	4.7	37.3
Column total		34.1	57.2	8.6	100.0

When it comes to the labor supply, single mothers with primary educations concentrate on no work or part-time work, while those with post secondary educations concentrate on full-time work. There is a clear pattern of labor supply increasing with level of education.

Table 7 presents the distribution of the choice variables by age groups. The largest group (17%) is that of single mothers aged 35-44, without paid childcare and with no social assistance. The next largest group is that of single mothers aged 18-34, with paid childcare and no social assistance. The smallest group (0.4%) is that of single mothers aged 45-64 with both paid childcare and social assistance. The probability of having social assistance is much lower for older mothers compared with younger mothers.

Using paid childcare is also less frequent for older mothers than with younger mothers. One reason is that older mothers have somewhat older children and therefore do not require childcare to the same extent.

The distribution of the labor supply by age groups shows that most single mothers are aged 18-34 and work part-time (21.4%), while the group less-represented is that of single mothers aged 45-64 who do not work at all.

Table 7 Distribution of labor supply, welfare and social assistance utilization by age

Welfare and childcare utilization		Age			Row total
		18-34	35 - 44	45 - 64	
P_{CC}=1	P_{SA}=1	10.9	3.6	0.4	14.8
P_{CC}=0	P_{SA}=1	8.8	6.6	1.1	16.5
P_{CC}=1	P_{SA}=0	15.2	14.4	3.9	33.6
P_{CC}=0	P_{SA}=0	11.6	17.3	6.2	35.1
Column total		46.5	41.8	11.6	100.0
H = 0		12.8	7.1	1,9	21.8
H = 20		21.4	16.1	3,4	40.9
H = 40		12.4	18.6	6,4	37.3
Column total		46.5	41.8	11.6	100.0

6 Results

Tables 9 and 10 report the parameter estimates for the model containing 4 equations, i.e., the labor supply, the welfare participation, the paid childcare utilization, and the wage equation. Fixed cost of work is not included as a separate effect, and therefore this model will be used as a reference when we later analyze the model that includes fixed cost of work. Our estimates are based on a simulated log-likelihood function using 100 Halton draws per individual.[72] According to Train (2001), 100 Halton draws are a suitable number for our purpose. However, later we will investigate the sensitivity of the parameter estimates to the number of draws by comparing estimates from a model using twice as many draws.

The variable specification used in this paper is a result of extensive testing in the spirit of making a parsimonious variable specification. In general, structural models are

[72] This is a sufficient number based on studies in the literature where similar numbers are used. One should keep in mind that the simulated maximum likelihood is a biased estimator for the log likelihood for a finite number of Monte Carlo draws. The simulated log likelihood is only asymptotically unbiased as the number of draws used to simulate the choice probabilities grows large, and obtain consistency only as simulation size goes to infinity.

sensitive to the choice of variable specification and starting values. However, after testing a range of specifications we found that the significant parameters were quite stable, and the choice of different but reasonable starting values mainly affected the convergence time. The labor supply equation in the model is a measure for the marginal utility of work. The marginal utility of work increases with age at a decreasing pace, which is no surprise since the supply of labor is lower for younger mothers as well as for older mothers compared to the middle-aged group. Single mothers living in big city regions or having higher educations have higher marginal utility for work then others. The age of the children in the household is important for the single mother's labor supply: the older the children, the more likely the mother is working.[73]

The equation for social assistance should be thought of as a marginal disutility (cost) measure in the sense that a parameter with a positive sign indicates a reduced utility. Almost all parameters of the social assistance equation are significant. The parameter for younger mothers is negative which implies that the younger they are the more likely that they use social assistance, which might be related to difficulties of working or in finding a job. Those living in a big city region or having higher educations are less likely to participate, as these factors are associated with having higher earnings therefore being less likely to require additional income support. Having children aged 9-12 at a paid after-school center is also associated with a lower probability of receiving welfare, compared to having younger children at paid childcare. Having children in the same age group at home without using paid day-care is not associated with welfare participation. The need of paid childcare for children aged 9-12 is most likely less urgent, which implies that low-income groups choose not to utilize municipal childcare for these children. From official statistics, we also know that a rather small group have their children at after-school center when the children are of this age.

All estimated parameters of the equation associated with marginal utility for paid childcare utilization have the expected signs. Single mothers living in a big city region have lower utility from paid childcare and therefore use it to a lesser extent. From the

[73] Since the utility function is normalized (β_Y=1) by weakly income, Y, all parameters in the model are expressed in income units. As an example, at H=40 moving from a rural region to a big city region is roughly equivalent to an increase in weakly income of 27.8 SEK in utility terms, ignoring the quadratic income term (0.696*40=27.8).

labor supply equation we know that the marginal utility to work is higher for single mothers living in big-city regions as oppose to those living in other regions, and from official statistics we know that mothers in city regions use more hours of paid childcare compared to other single mothers. With that in mind, one would expect a greater need for paid childcare. Nevertheless, it seems as though these mothers use means other than municipal childcare. This might be an indication of a supply problem in that it is easier to find a childcare placement in regions where the concentration of people is lower. Education has also a significant effect on the childcare utilization, and the utility of paid childcare increases with the educational level of the single mother. This is associated with younger mothers being unemployed to a higher extent and therefore not being entitled to municipal childcare. To have younger children is obviously a reason for demanding childcare and the estimates suggest that the demand decreases with the age of the children.

Finally, the estimates for the wage equation are also, in line with what we would expect. Single mothers living in big-city regions have higher wages compared to those living in other regions. Single mothers with higher educations have higher wages compared to other single mothers. However, the return to years of experience is almost zero. We included a squared experience term initially and found the sign of the parameter to be negative but with no significance. We therefore decided to exclude the squared term but to keep years of experience as an indication of its lack of significance. This lack of effect is due to the system of equations since estimating the wage equation separately gives significant effects of years of experience on wages.

Table 10 shows the utility parameters and the error covariance estimates for the model. The utility parameters have no clear interpretation but they are important components in the expression for the labor supply elasticities. Except for the component related to social assistance all other covariances among the unobserved components are significant. The unobserved factor for the decision to receive social assistance seems to be unrelated with the other choices.

Table 9 Parameter estimates for single women with young children, 100 Halton draws, no fixed costs of work[74, 75]

	Behavioral equations			
Observed Characteristics	Marginal utility of work	Marginal disutility of SA	Marginal utility of paid childcare	Wage equation
Constant	2.369*	0.457	0.191	2.635*
	(0.844)	(0.059)	(0.183)	(0.199)
Age	0.057*	-	-	-
	(0.017)			
Age squared/100	-0.061*	-	-	-
	(0.021)			
Age (18-34)	-	-0.494	-	-
		(0.143)		
Big city region	0.696*	0.182	-0.408*	0.618*
	(0.051)	(0.087)	(0.087)	(0.091)
Small city region	0.399*	-	-	0.342*
	(0.074)			(0.096)
Primary school	-1.129*	-	-0.653*	-
	(0.265)		(0.124)	
Secondary school	-0.473*	0.651	-0.541*	0.708*
	(0.163)	(0.056)	(0.089)	(0.099)
P. Secondary school	-	1.115	-	1.205*
		(0.295)		(0.118)
If children aged 1 - 5	-	-	0.726*	-
			(0.119)	
If children aged 6 - 8	-	-	0.559*	-
			(0.087)	
If children aged 9 - 12	-0.058	-	-0.392*	-
	(0.044)		(0.088)	
If children aged 13 - 17	0.263*	-	-	-
	(0.038)			
If children aged 9-12 on paid childcare	-	0.821*	-	-
		(0.245)		
If children aged 9-12 with parent	-	0.072	-	-
		(0.137)		
Work experience	-	-	-	-0.003
				(0.009)
Mean Log-likelihood	-3.738			
N	533			

Note: * significant at the 10% level. Standard errors are in the parentheses.

We observe that the unobserved component for the social assistance equation is negatively correlated with hours of work yet with large standard errors. On the other hand, the corresponding component for childcare utilization is positively correlated with hours of work. That is intuitively appealing since more hours of work implies that someone else has to take care of the child to a greater extent, which therefore implies an increased demand of paid childcare. The unobserved components of social assistance

[74] The eigenvalues of the Hessian are all negative, ensuring the Hessian to be negative definite.
[75] The GFBS approximated Hessian is used in the optimization of the simulated log likelihood function and therefore used to determine the standard errors of the parameters.

and paid childcare utilization are negatively correlated, which could be interpreted as though the utility for paid childcare increases the utility of social assistance. One would expect the opposite sign since paid childcare increases with hours of work and therefore the earning level increases as well, which would reduce the potential amount of social assistance. Here the standard errors are also large and the parameter should therefore be interpreted as being zero.

Table 10 Utility and covariance parameters

a) Utility parameters

Parameter	β_{HH}	β_{YY}	β_{HY}
Estimate	-0.489*	-0.665*	-4.185*
Standard error	(0.136)	(0.081)	(0.307)

Note: β_{HH} is multiplied by 100, β_{YY} is multiplied by 1000000, and β_{HY} is multiplied by 10000.

b) Covariance matrix

	ε_H	ε_{SA}	ε_C	ε_w
ε_H	2.164*	-0.088	0.278*	2.712*
	(0.282)	(0.158)	(0.036)	(0.191)
ε_{SA}		1.000	-0.063	-0.238
			(0.095)	(0.215)
ε_C			1.000	0.263*
				(0.082)
ε_w				3.437*
				(0.098)

Table 11 reports the estimates of the model when including fixed cost of work. In general, fixed cost of work is important in any labor supply study, and is especially the case when dealing with single mothers. The parameter estimates are basically unchanged for those parameters, which previously were significant. However, after including fixed cost of work, the efficiency of the model decreased, even though the point estimates were basically unchanged compared to the reference model that excluded fixed cost of work. A city region dummy, and three dummies indicating the ages of the children in the household are assumed to capture observed heterogeneity in fixed cost of work, but only the child-age dummies capture this effect, having significant parameters. Unfortunately, including fixed cost of work made the childcare utilization equation lose all significant effects. The basic problem behind this is most likely the small sample size that simply cannot offer enough variation to separate the effect for fixed cost of work and marginal utility of paid childcare utilization in the same model, even though the functions have different variable specifications. Our conclusion is therefore to include only one of the equations, and the choice of equation must be based on what kind of model one would like to estimate. In our case we have a

multiple-choice model and the choice of paid childcare utilization is essential for the purpose of this study. It is therefore natural for us to exclude the fixed cost of work equation. After all, fixed cost is captured indirectly in the equation for paid childcare utilization. The rest of the analysis will therefore be based on the model excluding fixed cost of work.

Table 11 Parameter estimates for single women with young children, 100 Halton draws, with fixed cost of work

	Behavioural equations				
Observed characteristics	Marginal utility of work	Marginal disutility of SA	Marginal utility of paid childcare	Wage equation	Fixed Cost of work
Constant	3.397*	0.142	0.478	2.632*	2.470
	(1.387)	(0.194)	(0.477)	(0.145)	(1.940)
Age	0.058				
	(0.038)	-	-	-	-
Age squared/100	-0.055				
	(0.051)	-	-	-	-
Age (18 – 34)		-0.359*			
	-	(0.198)	-	-	-
Big city region	1.043	0.205	-0.517	0.619*	-0.281
	(0.674)	(0.155)	(0.347)	(0.124)	(0.383)
Small city region	0.468			0.347*	
	(0.248)	-	-	(0.154)	-
Primary School	-1.326*		-0.101		
	(0.514)	-	(0.192)	-	-
Secondary School	-0.558*	0.702*	-0.366	0.711*	
	(0.302)	(0.161)	(0.293)	(0.160)	-
P. Secondary School		1.104*		1.199*	
	-	(0.315)	-	(0.274)	-
If children aged 1 – 5			0.654		
	-	-	(0.693)	-	-
If children aged 6 – 8			0.481		0.227*
	-	-	(0.464)	-	(0.086)
If children aged 9 – 12	-0.508*		-0.491		0.781*
	(0.258)	-	(0.367)	-	(0.338)
If children aged 13-17	-0.402				1.048*
	(0.425)	-	-	-	(0.468)
If children aged 9 – 12 on paid childcare	-	1.463* (0.529)	-	-	-
If children aged 9 – 12 With parent	-	0.172 (0.189)	-	-	-
Work experience				-0.005	
	-	-		(0.006)	-
Mean Log-likelihood	-3.717				
N	533				

Note: * significant at the 10% level. Standard errors are in the parentheses.

Table 12 reports the corresponding utility and covariance parameters and less than half of them are significant at a conventional level. The utility parameters are all significant but larger in magnitude compared to the reference model. How this affects the labor supply elasticities is difficult to say and needs to be examined, which will be done later on.

Table 12 Utility and covariance parameters

a) Utility parameters

Parameter	β_{HH}	β_{YY}	β_{HY}
Estimate	-0.862*	-1.017*	-6.405*
Standard error	(0.435)	(0.159)	(2.749)

Note: β_{HH} is multiplied by 100, β_{YY} is multiplied by 1000000, and β_{HY} is multiplied by 10000.

b) Covariance matrix

	ε_H	ε_{SA}	ε_C	ε_w
ε_H	2.995*	0.045	-0.201	3.174*
	(1.279)	(0.136)	(0.144)	(1.208)
ε_{SA}		1.000	-0.951	-0.028
			(1.265)	(0.194)
ε_C			1.000	-0.161
				(0.215)
ε_w				3.431*
				(0.194)

Table 13 reports the actual and fitted probability values of the choices in the model with no fixed cost of work. The fit is quite good, but it is evident that the model has a tendency to over-predict those cells with small representations and under-predict others. The fitted probability values were computed using the parameter estimates in Tables 9 and 10 and using 100 Halton draws for each individual. The mean values were then calculated and reported as the choice probabilities in the table.

Table 13 Actual and fitted distributions of labor supply and program participation (in percent)

Welfare and childcare utilization		Hours of work			Row total
		0	20	40	
			Actual		
P_{CC}=1	P_{SA}=1	5.6	6.4	2.8	14.8
P_{CC}=0	P_{SA}=1	7.7	7.3	1.5	16.5
P_{CC}=1	P_{SA}=0	3.2	11.6	18.8	33.6
P_{CC}=0	P_{SA}=0	5.2	15.6	14.3	35.1
Column total		21.7	40.9	37.4	100.0
			Fitted		
P_{CC}=1	P_{SA}=1	5.97	5.53	3.94	15.44
P_{CC}=0	P_{SA}=1	8.28	5.83	3.39	17.50
P_{CC}=1	P_{SA}=0	3.03	14.22	15.61	32.89
P_{CC}=0	P_{SA}=0	4.03	16.33	13.77	34.17
Column total		21.4	41.9	36.7	100.0

Table 14 reports predicted probabilities conditional on the endogenous variables in the model. The relationship between social assistance and paid childcare utilization presents the most interest. Paid childcare utilization and welfare participation work in different directions in their effects on changes in employment. Welfare participation reduces the probability of working, while paid childcare utilization is associated with an increased probability to work. The probability of having social assistance is reduced when conditioning on childcare and the probability of using paid childcare is also reduced when conditioning on social assistance, which indicates that the two variables are stochastically related to each other.

Table 14 Conditional choice probabilities (in %)[76]

Labor force participation		Welfare participation		Paid Childcare	
P(H>0)	78.66	P(SA=1)	32.98	P(CC=1)	48.32
P(H>0|SA=1)	56.74	P(SA=1| H>0)	23.79	P(CC=1|H>0)	49.97
P(H>0|CC=1)	81.35	P(SA=1|CC=1)	31.97	P(CC=1|SA=1)	46.84
P(H>0|SA=1,CC=1)	61.34	P(SA=1| H>0,CC=1)	24.11	P(CC=1|SA=1,H>0)	50.64

[76] The conditional probabilities are computed using the mean choice probabilities in the model. The conditional probability of having social assistance given labor force participation was calculated in the following way: P(SA=1|H>0) = P(SA=1,H>0)/P(H>0). The other values were computed analogously.

6.1 Policy simulation

Table 15 presents the elasticities for single women with children. These numbers are based on comparative static analysis. The estimated choice probabilities have been evaluated before and after a specific change in the budget set. Based on these changes the corresponding change in labor supply has been calculated. The sign of the wage elasticity is ambiguous according to economic theory, but in the empirical literature, the sign is typically positive. One should hold in mind that the elasticities are just a measure of the mean effect, and different individuals may well have different wage elasticities. Even the same individual may have different wage elasticities both in sign and magnitude at different hour levels. Any inference drawn from single wage elasticities could therefore be dangerous. The sign of the income elasticity is unambiguously negative if leisure is a normal good. The signs of the elasticities in Table 15 are therefore all expected. Single mothers have a strong inclination for both net income and leisure (i.e., a sensitive trade-off), which might imply that they should have elasticities that differ greatly for different wages and hours of work combinations. The childcare cost elasticity on labor supply suggests that if childcare cost increases by 1%, the labor supply would be reduced by 0.16%. The welfare elasticity on labor supply measures the response on labor supply with respect to a change in the social assistance norm. If the social assistance norm increases by 1%, the elasticity suggests that the labor supply would decrease by 0.06%, which is a much lower figure compared to childcare cost elasticity.

The right side of Table 15 contains two measures for the elasticities between social assistance norm and childcare cost. The first one shows that if the social assistance norm increases by 10%, the probability of using paid childcare would decrease with 0.19%. Hence, the responsiveness in paid childcare utilization is very small from changes in the social assistance norm. The second elasticity, which measures the relationship in the opposite direction, shows that if the municipal childcare cost was reduced with 10% the probability of receiving social assistance would decrease by 1.6%. This shows the importance of the childcare cost for single mothers in relation to social assistance. The reason for this is the fact that a reduction in childcare cost directly affects the level of disposable income, which directly affects the level of social

assistance a single mother could receive. A change in the social assistance norm does not provide a direct link to the preferences to use paid childcare.

Table 15 Mean elasticities for single women with children

Labor supply elasticities	Mean	**Other elasticities**	Mean
Uncompensated wage elasticity	0.771	Social assistance norm effect (+10%)	-0.186
Income elasticity	-0.098	on paid childcare utilization	
Childcare cost elasticity on labor supply	-0.163	Childcare cost effect (-10%) on social	-1.607
Welfare elasticity on labor supply	-0.061	assistance participation	

Table 16 reports the simulated mean responses to changes in the budget set. The simulated responses are obtained by computing mean probabilities for each of the alternatives in the choice set across the individuals and for different alterations in the budget constraint using the SML estimates of Tables 9 and 10. The simulations are performed separately for each individual in the sample and then averaged across the individuals. The baseline represents the predicted probabilities of paid childcare utilization, welfare participation and hours of work as they are at the optimal point before any change in the budget set. These numbers are compared with the new numbers received after the change in the budget constraint. The simulation results show that a 1% increase in wages corresponds to a 0.77% increase in labor supply. This corresponds to the uncompensated wage elasticity given in Table 15. The mean labor supply is derived from the corresponding probabilities in Table 16 as the expected values of hours of work.[77] The values suggest that the wage effect on labor supply is non-linear and increases with a decreasing pace. There appear to be only a small impact on the part-time workers, the major change being on full-time work. The exact picture of the transitions would require a transition matrix. It is more reasonable to believe that non-workers move to part-time work and part-time workers to full-time work. What we see in the table is therefore only the net effect after the changes.

[77] The expected value is given by E[H] = 0*P(H=0) + 20*P(H=20) + 40*P(H=40).

Table 16 Simulated responses to changes in the budget set. SML 100 Halton draws

	Participation (%)		Work hours distribution (%)			Mean hours worked	Hours change (%)
	P(SA=1)	P(CC=1)	P(H=0)	P(H=20)	P(H=40)		
Baseline	32.98	48.32	21.33	41.93	36.74	23.08	-
Wage change							
+ 1 %	32.63	48.50	20.93	41.85	37.22	23.26	0.771
+ 10 %	30.32	49.91	17.97	41.97	40.06	24.42	5.797
Change in the tax structure							
Municipality tax: - 1 % point	32.74	48.45	21.14	41.80	37.06	23.18	0.431
Basic deduction: 18kkr flat	32.54	48.74	21.22	41.30	37.48	23.25	0.745
Social assistance norm: +10 %	34.50	48.23	21.95	42.03	36.02	22.81	-1.151
Childcare cost: - 10 %	32.45	48.56	21.29	41.52	37.19	23.18	0.433

Note: SA=social assistance, CC=paid childcare utilization

A reduction in the tax level for low-income persons would most likely reduce welfare participation, and it would therefore be interesting to simulate how large such a response would be. Reducing the municipality tax by a 1-percentage point was found to have a positive effect on labor supply by a 0.43% increase and a reducing effect on welfare participation by 0.73%. The last effect is most likely the result of the increase in net income that has an immediate effect on the welfare amount received. The probability of using paid childcare also increases slightly. This effect comes from the increased labor supply inducing a need for childcare.

A change in the structure of the tax system is believed to have an effect on the labor supply and welfare behavior. When we simulate a system with a flat basic deduction of 18,000 SEK (modeled to be the same for all individuals independent of the income level) the labor supply increases by 0.75%, while the probability of using social assistance decreased and the probability of using paid childcare increased. This change affects mostly low-income groups, which might be due to the economic incentives.

Changes in the social assistance norm have important incentive effects on the labor supply. If the norm is increased by 10 %, a transfer from full-time work to part-time work takes place, which implies that the part-time work increases from the increased norm. There is also an increase in the probability of not working at all, which implies a decrease in the need for paid childcare.

Finally, we simulated the effect of changing the overall childcare fee for those with paid childcare. A 10% decrease in the total cost has several implications. It increases the income opportunities, which lead to a welfare reduction, which induces the

single mother to increase her labor supply. However, the transfer from non-work to work is very small; the major transfer instead being from part-time to full-time work.

6.2 Sensitivity analysis

Table 17 presents key components from different estimations of the model using different specifications. The figures to the left are considered to be the base model using 100 Halton draws, imposing no restrictions on the covariances, but excluding fixed cost of work as a separate effect. When using SML it is important to know how sensitive the estimates are to the number of draws used in the estimation. This is easily evaluated by comparing the results with estimates determined with twice as many draws. If the parameter estimates differ extensively, it is a sign of using too few draws since the estimated probabilities are far from convergence. In Table 17 we see that the parameter estimates hardly changed at all when using twice as many draws. This indicates that more draws add little to the precision of the point estimates. We therefore conclude that 100 Halton draws are a sufficient number when estimating the choice probabilities. This confirms the results from Train (2001), which claim that 100 draws are a sufficient number for stable and reliable estimates.

Table 17 Sensitivity analysis

	Cov free 100 Halton No fixed cost	Cov free 200 Halton No fixed cost	Cov free 100 Halton With fixed cost	Cov zero 100 Halton No fixed cost
β_{HH}	-0.492	-0.494	-0.862	-0.414
β_{YY}	-0.765	-0.768	-1.017	-0.664
β_{HY}	-4.235	-4.768	-6.405	-1.686
Simulated mean Log-Likelihood	-3.738	-3.739	-3.717	-3.942
Wage elasticity	0.771	0.773	0.878	0.548
Income elasticity	-0.098	-0.091	-0.141	-0.052

The next specification adds fixed cost to the model. As discussed above including fixed cost makes all parameters in the equation for paid childcare utilization lose their significance. We consider this to be a negative sign even though a likelihood ratio test would say that the model improved. The elasticities did not change greatly, even though the wage elasticity increased slightly and the income elasticity increased in magnitude.

Finally, we have the estimates from a model restricting all the covariances to zero, which show that the utility parameters differ slightly from the other models. A likelihood ratio test is easily rejected when testing if the reference model differs from one with covariances being zero.

7 Summary

In this paper, we have applied a simulation method to estimate a structural labor supply model incorporating welfare participation and paid childcare utilization for single mothers in Sweden. By approximating the hours of work for three discrete points (unemployed, part-time work, full-time work) and defining the choices of welfare and paid childcare as discrete alternatives, we were able to formulate the model as a multiple-choice problem, giving the single mothers a choice set of 12 alternatives. We estimated the full structural model with and without a separate effect of fixed cost of work, and found a conflict in having both fixed cost of work and paid childcare utilization incorporated in the model at the same time. When estimating the model including fixed cost of work, the parameters in the paid childcare equation all lost their significance while some of the parameters in the fixed cost equation were significant. The transport of significance into the fixed cost might be an indication that it is the fixed cost component that should stay in the model. However, if the model is design to analyze the relationship between the endogenous choices, the equation for paid childcare utilization has to stay.

The model excluding fixed cost of work was used in the remaining part of the study, and was used to perform simulations of the responses to changes in the budget constraints. Simulations of the labor supply elasticities showed inelastic values, with an uncompensated wage elasticity of 0.77 and an income elasticity of -0.1. The elasticity between childcare cost and welfare participation on labor supply is of special interest, and the results show a childcare cost elasticity of -0.16 and a welfare norm elasticity of -0.06. The relation between social assistance and childcare cost is non-symmetric, as a 10% increase in the social assistance norm reduces the probability of using paid childcare by 0.19% while a 10% reduction in childcare cost reduces the probability of having social assistance by 1.6%.

The performed micro simulations suggested that an increase in the social assistance norm has a relatively large mean labor supply effect where the major change is from full-time work to part-time and non-work. In comparison, a reduction in the childcare cost has a relatively small effect on the mean labor supply; the major change was within the group of those who already were working, by reducing part-time work with full-time work. The effect on overall employment was therefore very small. Childcare cost can therefore be seen as a barrier to full-time work rather than as an obstacle to employment.

References

Averet, S. L., H. Elizabeth Peters, and D. M. Waldman (1997). "Tax Credits, Labor Supply, and Child Care", Review of Economics and Statistics 79, 125-135.

Blau, D., and P. Robins (1988), "Child Care Costs and Family Labor Supply", Review of Economics and Statistics 70, 374-381.

Cogan, J. (1981), "Fixed Costs and Labor Supply", Econometrica 49, 945-964.

Connelly, R. (1989), "Determinants of Weekly Child Care Expenditures: A Comparison of Married and Unmarried Mothers", Bowdoin College Working Paper.

Connelly, R. (1990), "The Cost of Child Care for Single Mothers – Its Effect on LFP and AFDC Participation", Institute for Research on Poverty, Discussion Paper #920-90.

Connelly, R. (1992), "The Effect of Child Care Costs on Married Women's Labor Force Participation", The Review of Economics and Statistics 74, 83-90.

Connelly, R., and J. Kimmel (2001), "The Effect of Child Care Costs on the Labor Force Participation and Welfare Recipiency of Single Mothers: Implications for Welfare Reforms", Working paper W. E Upjohn Institute.

Crecelius, E, and E. Y. Lin (2000), "The Effects of Child Care Costs on Labor Supply and Welfare participation", Mimeo Upjohn Institute.

Danziger, S., R. Haveman and R. Plotnick (1981), "How Income Transfers Affect Work, Savings, and the Income Distribution: A Critical Review", Journal of Economic Literature 19, 975-1028.

Fraker, T., R. Moffitt (1988), "The Effect of Food Stamps on Labor Supply: A Bivariate Selection Model", Journal of Public Economics 35, 25-56.

Geweke, J. (1991), "Efficient Simulation from the Multivariate Normal and Student Distributions Subject to Linear Constraints", Computer Science and Statistics: Proceedings of the Twenty-Third Symposium on the Interface, 571-578.

Gustafsson, S., and F. Stafford (1992), "Child Care Subsidies and Labor Supply in Sweden", The Journal of Human Resources 27, 204-230.

Gustafsson, B., and A. Klevmarken (1993), "Taxes and Transfers in Sweden: incentive effects on labor supply", in A. B. Atkinsson and G. Viby Mogensen, editors, Welfare and Work Incentives. A North European Perspective, Oxford University Press.

Hajivassiliou, V. (1990), "Smooth Simulation Estimation of Panel Data LDV Models", Working paper, Department of Economics, Yale University.

Halton, K. H. (1960), "On the Efficiency of Certain Quasi-random Sequences of Points in Evaluating Multi-dimensional Integrals", Numer. Math. 2, 84-90.

Heckman, J. (1974), "Effects of Child-Care Programs on Women's Work Effort", Journal of Political Economy 82, Supplement, 153-161.

Keane, M. (1994), "A Computationally Practical Simulation Estimator for Panel Data with Applications to Estimating Temporal Dependence in Employment and Wages", Econometrica 62, 95-116.

Keane, M., and R. Moffitt (1998),"A Structural Model of Multiple Welfare Program Participation and Labor Supply", International Economic Review 39, 553-589.

Kimmel, J. (1995), "The Effectiveness of Child Care Subsidies in Encouraging the Welfare to Work Transition of Low Income Single Mothers", American Economic Review papers and proceedings 85, 271-275.

Kimmel, J. (1998), "Child Care Costs as a Barrier to Employment for Single and Married Mothers", The Review of Economics and Statistics 80, 287-299.

Lerman, S., and C. Manski (1981), "On the Use of Simulated Frequencies to Approximate Choice Probabilities", in C. Manski and D. McFadden, eds., Structural Analysis of Discrete Data with Econometric Applications (Cambridge, MA; MIT Press), 305-319.

McFadden, D. (1989), "A Method of Simulated Moments of Estimation of Discrete Response Models without Numerical Integration", Econometrica 57, 995-1026.

Moffit, R. (1983), "An Economic Model of Welfare Stigma", American Economic Review 73, 1023-35.

Moffitt, R. (1992), "Incentives Effects of the U.S. Welfare System: A Review", Journal of Economic Literature 30, 1-61.

Ribar, D. (1992), "Child Care and the Labor Supply of Married Women: Reduced Form Evidence", Journal of Human Resources 27, 134-165.

Ribar, D. (1995), "A Structural Model of Childcare and the Labor Supply of Married Women", Journal of Labor Economics13, 558-597.

Socialstyrelsen (1997), "Råd med barnomsorg? En kartläggning av avgifterna i dag- och fritidshem för 14 typfamiljer", Socialstyrelsens publikationsserie.

Skolverket (1998)," Barnomsorgen i siffror 1998: Barn och personal", Skolverkets rapport nr 152.

Skolverket (2000), "Descriptive Data on Child Care and Schools in Sweden in 2000", Report no. 192.

Train, K. (2001), “Discrete Choice Methods with Simulation”, Cambridge University Press, Draft version.

Van Soest, Arthur (1995), “Structural Models of Family Labor Supply: A Discrete Choice Approach”, The Journal of Human Resources 30, 63-88.

Appendix

A1 The Swedish tax system in 1997

The tax system applied in this study is based on several components, namely the municipality tax rate, the public tax rate, the basic deduction (grundavdrag), and the tax rate on capital. Municipality tax varies among the municipalities and has an average rate of 31.76% in 1997. We have access to each rate, and apply them here. The public tax rate has a basic flat rate of 200 SEK, and 25% of the taxable income exceeding 209,100 SEK. The tax rate on capital is 30%. Table A1 presents how the basic deduction works, while Figure A1 presents how the marginal and average taxes work.

Table A1 Basic deduction 1997

Income intervals	Basic deduction
8,700 – 67,800	8,700
67,900 – 104,600	8,712 + 25% of income > 67,518
104,700 – 110,900	18,000
111,000 – 202,900	18,059 – 10% of income > 110,352
203,000 –	8,700

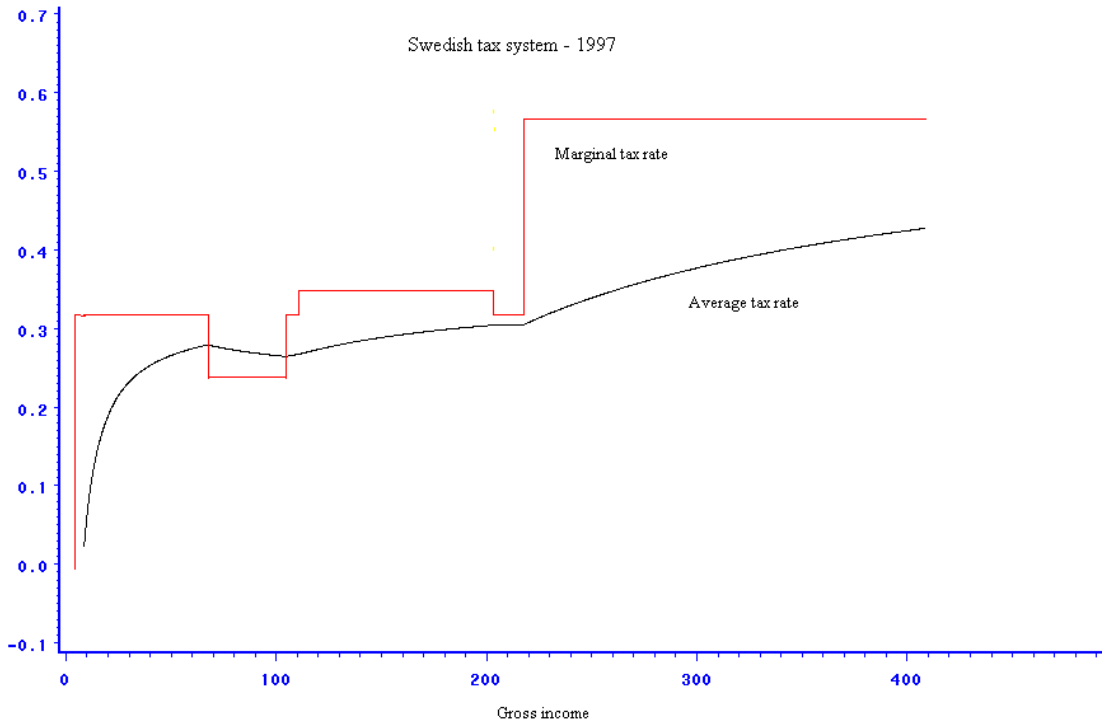


Figure A2 Marginal and average tax rate

A2 The social assistance norm

The level of social assistance an individual may receive is decided by each municipality, which has the right to pinpoint the exact amount. The National Board of Health and Welfare (Socialstyrelsen) gives guidelines to the municipalities so that the program will be equally defined throughout the country. Nevertheless, differences exist. This study uses the general guidelines proposed by the National Board of Health and Welfare and thereby assumes a unified system throughout the country. The assumed system consists of 4 parts. Individual parts for adults and children, a common part and the housing rent cost.

Table A2 The social assistance norm (SEK)

Single person component	2,320						
Child component	**Age of the child**						
	0	1-2	3	4-6	7-10	11-14	15-18
	1,230	1,440	1,120	1,410	1,530	1,830	2,070
Common component	**Individuals in the household**						
	1	2	3	4	5	6	7
	580	670	760	820	910	960	1020

For example: a single mother with two children aged 4-7 would receive:

MAX[0, 2320 + 1410 + 1530 + 760 + rent – disposable income] = *SA amount per month* .

A3 The potential cost of childcare

The cost of childcare differs extensively among the municipalities in Sweden. A single parent with an average income, and 1 child at a day care centre 26 hours a week, pays 200 SEK per month in the cheapest municipality and 1,600 SEK per month in the most expensive municipality. That is a difference of about 17,000 SEK per year. It is not feasible to try to model each and every system into the model. We therefore have to assume a united system, one that is the same for all individuals and represents an average situation. Most of the childcare cost systems in Sweden have several common components that are important to incorporate, such as the cost reduction for extra children on childcare, differentiated cost dependent on the age of the children, and a

fixed and variable cost component. Table A3 presents the unified cost of childcare system used in this study.

Table A3 The unified cost of childcare system (SEK)

	Children			
Base fee	1	2	3	4
Children aged 1 – 5	500	300	100	0.0
Children aged 6 – 8	450	270	90	0.0
Children aged 9 – 12	400	240	80	0.0
Variable fee per hour	1	2	3	4
Children aged 1 – 5	8.50	5.10	1.70	0.0
Children aged 6 – 8	7.65	4.59	1.53	0.0
Children aged 9 – 12	6.80	4.08	1.36	0.0

The cost of childcare is also a function of the household yearly income and is related to the base amount (basbelopp). If a household has an income lower than 20 % of the base amount, there is no charge for childcare. If the income is between 20-70 % of the base amount, the childcare cost is reduced linearly. When the household has an income higher then 70% of the base amount, the household pays the full childcare cost. The base amount was 36,300 SEK in 1997 and 36,400 SEK in 1998. We used the value for 1997 in this paper.