Assessing the Employment Effects of Labor Market Training Programs in Sweden*

Daniela Andrén and Thomas Andrén[α]

Working Papers in Economics no 70
May 2002

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 Office 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 et al., 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).[1] 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;[2] 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

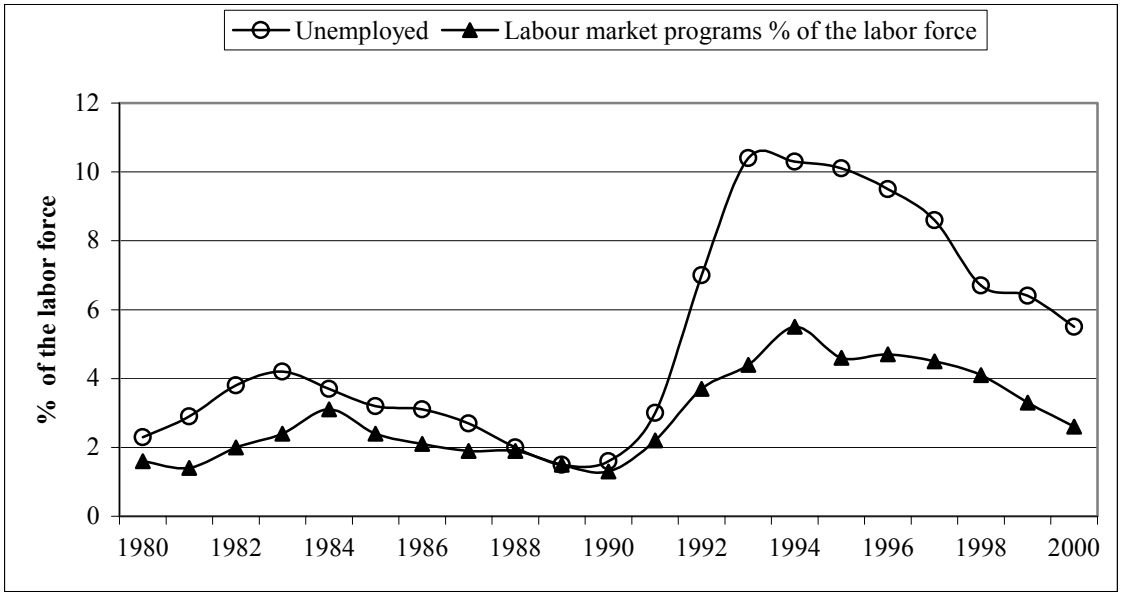
[1] We present the structure and rules of the system valid during 1993-1994, the period analyzed by this study.
[2] Until 1996, a 5-month participation in practically any labor market program would count as employment in allowing participants to become eligible for the first time.

believe that this is often the case.[3] 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.

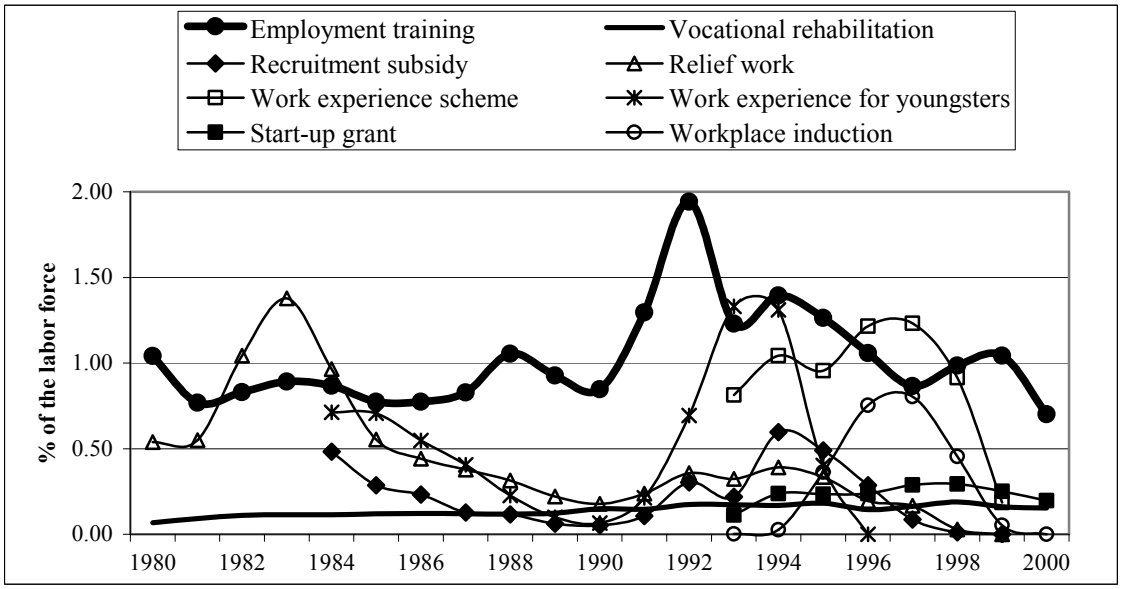
Originally, labor market training mainly consisted of vocational training programs, but over time, schemes comprised of more general educational training have grown more prevalent. During the 1990s, other education programs such as Swedish for immigrants, and computer training have been added to labor 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.

[3] 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[4]

[4] Data source: National Labor Market Board (Historisk statistik 1980-2000; AMS Statistikenhet; Arbetsmarknadsstyrelsen 2001).

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 3,000 and 7,000) 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.[5] 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.

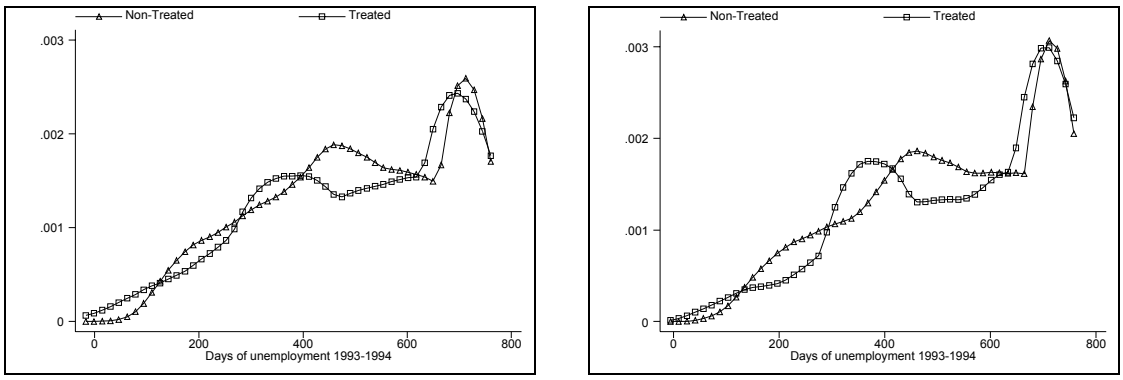
3.2 The construction of the comparison group

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;[6] 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 distribution of days of unemployment during the training period for trainees and non-

[5] 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.

[6] 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.

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),[7] 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.

[7] 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. [8] 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).[9] 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.

[8] 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.

[9] 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.[10] 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

[10] 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.[11] 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:

[11] 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:[12]

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

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.[13] 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:

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

For many questions, knowledge of distributional parameters is required. Heckman (1992), Heckman and Smith with Clements (1997) and Heckman and Smith (1998)

[12] $\text{TT}(X,Z) = \int \left[\Phi(Y_1 \mid X,\xi) - \Phi(Y_0 \mid X,\xi)\right] dF(\xi \mid D=1, X, Z)$ 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*).

[13] This estimator was first introduced into the evaluation literature by Björklund and Moffitt (1987) in the context of the Roy model.

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.[14] 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

[14] 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.[15]

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[16]	**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[17]	-	-	-	-	-	-	**-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[18]		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

[15] 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.

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

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

[18] 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.[19] 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

[19] 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.[20] 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.

[20] 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: >10 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.[21] 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

[21] 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[22]

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.

[22] 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.[23] The MTE

[23] 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 $\text{MTE}(X,u_D) = X(\beta_1 - \beta_0) - E(U_1 - U_0 \mid U_D = u_D)$ 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)| X=x, D=1]	0.405	0.399	0.492
E[1(Y_1=1,Y_0=1)| X=x, D=1]	0.308	0.294	0.238
E[1(Y_1=0,Y_0=0)| X=x, D=1]	0.164	0.203	0.194
E[1(Y_1=0,Y_0=1)| X=x, D=1]	0.123	0.104	0.076
Foreign-born			
E[1(Y_1=1,Y_0=0)| X=x, D=1]	0.111	0.249	0.233
E[1(Y_1=1,Y_0=1)| X=x, D=1]	0.342	0.213	0.236
E[1(Y_1=0,Y_0=0)| X=x, D=1]	0.168	0.335	0.299
E[1(Y_1=0,Y_0=1)| 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. 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.

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

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.

Calmfors L., A. Forslund and M. Hemström (2002), "Does active labour market policy work? Lessons from the Swedish experiences", IFAU Working Paper 2002: 4, Office of Labour Market Policy Evaluation, Uppsala.

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[24]	**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[25]	-	-	-	-	-		**-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[26]		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: **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.

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

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

[26] 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

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: >10 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: **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.

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[27]	**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[28]	-	-	-	-	-		**-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[29]		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: **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.

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

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

[29] 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: >10 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: **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.