IFAU – INSTITUTE FOR
LABOUR MARKET POLICY
EVALUATION

The effect of vocational employment training on the individual transition rate from unemployment to work

Katarina Richardson
Gerard J. van den Berg

WORKING PAPER 2002:8

The effect of vocational employment training on the individual transition rate from unemployment to work*

Katarina Richardson**
Gerard J. van den Berg***
15 April 2002

Abstract
Amongst the active labor market policy programs for the unemployed in Sweden, the vocational employment training program is the most ambitious and expensive. We analyze its effect on the individual transition rate from unemployment to employment using a unique set of administrative data and a novel empirical approach that exploits variation in the timing of training and exit to work. The approach involves the estimation of duration models, and it allows us to quantify the individual effect of training in the presence of selectivity on unobservables. The data contain the full population of unemployed in the period 1993-2000 and include multiple unemployment spells for many individuals. The results indicate a significantly positive effect on exit to work after exiting the program. Its magnitude is very large shortly after leaving the course but diminishes afterwards. If we also take account of the time spent in the program then the net effect of participation in the program on the mean unemployment duration is close to zero.

Keywords: vocational training, program evaluation, duration analysis, Swedish labor market, selectivity bias, treatment effect.

JEL code: J64, J24, I28, J68.

---

* We thank the Swedish National Labour Market Board (AMS) and Statistics Sweden (SCB) for their permission to use the data. We thank Richard Smith, Costas Meghir, Annette Bergemann, Håkan Regnér, Anders Harkman, participants at the Conference on Active Labor Market Policy (Stockholm, 2001), the Conference on Adult Education (Uppsala, 2001), and seminars at Aarhus, Bristol, and Mannheim, as well as our colleagues at IFAU, for helpful comments. In addition, we thank Helge Bennmarker for help with the data.
** Institute for Labour Market Policy Evaluation. Address for correspondence: Box 513, SE-751 20 Uppsala, Sweden. Phone +46 (0)18 471 7074, e-mail: Katarina.Richardson@ifau.uu.se.
*** Department of Economics, Free University Amsterdam; IFAU-Uppsala; Tinbergen Institute; CEPR.

Table of contents

1 Introduction

Training programs for the unemployed have been cornerstones of the Swedish labor market policy for many decades now. In fact, training programs have been used in Sweden since 1918. Nowadays, the so-called employment training program is the most prestigious program (here we simply call this the AMU program). It aims to improve the chances of unemployed job seekers to obtain a job, by way of courses that in which skills are substantially improved. In 1997, on average 37,000 individuals were participating in AMU per month, which corresponds to over 10% of total unemployment[1] AMU is the most expensive type of active labor market program in Sweden and as such it adds to the tax burden. Nevertheless, the number of evaluation studies is rather small, and most of these analyze the effect of AMU on the participants' earnings and/or use data from early eighties and/or data on special subgroups of unemployed workers, notably youths in Stockholm (see references below).

This paper empirically analyzes the effect of AMU on the individual transition rate from unemployment to employment. Note that the officially stated objective of AMU is to generate a positive effect. The results are of obvious importance for the evaluation of the AMU program and the underlying "Swedish model". In addition, they are of importance in the light of the recent policy shifts in other European countries towards an increased use of active measures of bringing the unemployed back to work, notably by way of reschooling unemployed workers with low skills or obsolete qualifications (see e.g. Fay, 1996).

We use a unique fresh set of longitudinal administrative data that contains the full population of individuals who were unemployed in Sweden within the period from January 1, 1993 until June 22, 2000. This dataset matches detailed records from employment offices to records from unemployment insurance agencies. The employment office data report detailed information on the types of training and the corresponding dates of entry and exit. Many individuals have experienced multiple unemployment spells within our observation

[1] In 2000, these figures are 30,000 and 9%, respectively (see AMS, 2001).

window.

The empirical analysis applies a methodology in which the information in the *timing* of events (like the moment at which the individual enrols in training and the moment at which he finds a job) is used to estimate the individual training effect. This method takes account of the fact that the training decision and the decision to accept a job are affected by unobserved factors. If individuals who receive training would have found a job relatively fast anyway, and if the analysis ignores this, then the training effect on the exit rate to work is over-estimated. Our method allows us to distinguish between such selection effects and the causal individual training effect. This is a major advantage over the use of methods that require the selection effects to be captured completely by observed variables (like the so-called matching approach), in particular if the set of observed variables only contains a small number of indicators of past individual labor market behavior, as is often the case. Another advantage of our approach is that it enables an examination of the way the individual training effect changes over time. This way we can learn something about the reasons for why training works or not.

The empirical approach involves estimation of models that simultaneously explain the duration until work and training participation. The approach has been used in a number of empirical studies, like Eberwein, Ham and LaLonde (1997) and Van den Berg, Holm and Van Ours (2002) (on training programs), and Abbring, Van den Berg and Van Ours (1997) (on punitive sanctions for unemployment insurance recipients; see also Gritz, 1993, and Bonnal, Fougère and Sérandon, 1997, for extensions focused on training programs, and Van den Berg, 2001, for a survey). Abbring and Van den Berg (2000) provide a statistical underpinning of this approach, and they provide a systematic account of the behavioral assumptions required for valid use.[2] For example, individuals are not allowed to anticipate

[2] Here we just provide some clarification on what drives the identification of the training effect. Consider individuals who enter training at time t. The natural control group consists of individuals who are unemployed for the same period of time at t but who have not yet received training. A necessary condition for a meaningful comparison of these groups is that there is *some* randomization in the training assignment at s. The duration model framework allows for such randomization because it specifies assignment by way of the

the moment at which they go into training a long time in advance, although they are allowed to know the distribution of this moment over time. We argue that AMU fits into this methodological framework rather well, whereas other active labor market programs in Sweden may not fit so well into this. To demonstrate this, we rely to some extent on evidence from in-depth interviews with caseworkers, and we rely to some extent on existing studies on unemployment, unemployment insurance, and active labor market programs in Sweden. These include Eriksson (1997a, 1997b), Zettermark et al. (2000), Carling and Richardson (2001), Dahlberg and Forslund (1999), Edin et al. (1998), and Carling et al. (1996). The latter two studies deal with the interaction between the inflow into active labor market programs in general on the one hand, and expiration of benefits entitlement on the other. We return to this in Sections 2 and 3.

To date, a few econometric studies have addressed the effect of AMU on unemployment duration. Harkman and Johansson (1999) and some replication studies examine individuals who finish a program in the final quarter of 1996. Harkman and Johansson (1999) use a subset of the data that we use and match it to data from a postal survey conducted in late 1997. They estimate a bivariate probit model on the employment probability at one year after the program, for different programs. They assume that the composition of programs within the employment office affects the individual probability of program participation but does not affect the individual probability of exit to work. Under this assumption, an observed relation between the composition of programs and the exit probability to work indicates a training effect. However, it is not clear whether the assumption is justifiable. Their results indicate that persons in AMU have a higher probability to get a job. Subjective responses on the perceived importance of program

---

rate of entering training. In addition, we have to deal with the selection issue that the unobserved heterogeneity distribution is different between the treatment and control groups at t. This can be corrected by exploiting the information in the data on what happened to individuals who received training and/or left unemployment before t. Another way to explain identification emphasizes the importance of the timing of events. If treatment and outcome are typically realized very quickly after each other, no matter how long the elapsed duration in the state of interest before the treatment, then this is evidence of a positive causal treatment effect. The selection effect does not give rise to the same type of quick succession of events.

participation agree to the estimation results.[3]

Our empirical analyses focuses on the way in which the effect on the exit rate to work varies over time. For example, we allow this effect to depend on the elapsed time in unemployment since exiting the course. We also estimate a model that takes account of the real time spent in training. The latter mitigates any positive effect of training, in the sense that time in training by itself increases the mean unemployment duration. In addition, we estimate models where the training effects depends on gender, nationality, and level of education.

The papper is organized as follows. Section 2 describes the AMU program. In Section 3 we briefly discuss the model framework and we argue that AMU fits into it. Technical details are relegated to the appendix of the paper. Section 4 describes the data. Section 5 contains the main estimation results. Section 6 concludes.

[3] Edin and Holmlund (1991) and Larsson (2000) examine the effect of AMU on the transition rate from unemployment to work for *young* individuals aged below 25. Edin and Holmlund (1991) use data from Stockholm from the early 1980s. They compare the unemployment spells of individuals who become unemployed and do not enter the program with the unemployment spells after exiting an AMU-program, and they attempt to deal with selectivity by adding many variables on the individual's unemployment history. They find a positive effect. Larsson (2000) also uses a matching approach, with data from the 1990s. Her results are mixed. We do not examine these studies further because in our empirical analyses we restrict attention to individuals aged over 25 (see Subsection 3.2). See Björklund (1993) for a survey of other studies based on data from the 1970s and 1980s. Regnér (2002) analyses earnings effects of AMU using data from around the 1980s. A matching approach is used to construct a comparison group. He concludes that on average there is no effect of AMU on earnings.

2 Labor market training in Sweden

2.1 The AMU program

The purpose of the AMU program is to improve the chances of job seekers to obtain a job, and to make it easier for employers to find workers with suitable skills. This means that it aims to increase unemployed individuals' transition rate to work. The program attempts to achieve this by way of the participation of individuals in training and education courses.[4]

The program is targeted at unemployed individuals as well as employed individuals who are at risk of becoming unemployed. The individuals have to be registered at the local job center (which we shall call the (local) employment office) and must be actively searching for a job. The lower age limit is 20, although nowadays younger individuals are entitled to participate if they are disabled or receive unemployment insurance (UI) benefits.

During the 1980s, the yearly average number of individuals in AMU per month was about 40,000. During the big Swedish recession of the early 1990s, this number increased up to 85,000, with seasonal peaks of about 100,000 participants. After 1992, this number decreased again to about 30,000-40,000, which is about 1% of the total labor force (Dahlberg and Forslund, 1999; AMU, 2001). Nowadays, the annual inflow into AMU is about 80,000. The average duration of a course has fluctuated during the past decade and is now about six to seven months. In 1994, total expenditure on the AMU program amounted to about SEK 12 billion (US $ 1.2 billion), half of which was for training procurement and half for training grants. Per participant this equals about $ 10,000 for procurement and $ 10,000 for grants, on a yearly base (AMS, 1997).

There is strong evidence that in 1991 and 1992, participation in AMU was often used in order to extend benefits entitlement (Regnér, 2002, and Edin et al. 1998). This requires a brief exposition. A commonly recognized problem with Swedish labor market

[4] See e.g. AMS (1997). The formulation of the official aims of AMU has changed somewhat over time. For example, earlier formulations sometimes even refer to the prevention of cyclical inflationary wage increases. See e.g. Harkman and Johansson (1999) and Regnér (1997).

programs is that they are used to extend an individual's entitlement to unemployment benefits (which is 300 working days ($\approx$ 14 months) for those aged between 25 and 55). By participating in a program, the unemployed individual ensures that his benefits entitlement is extended until completion of the program; in fact, if the participation exceeds a few months then the new entitlement extends further into the future. Edin et al. (1998) examine this interaction between inflow into active labor market programs in general on the one hand, and expiration of benefits entitlement on the other. They do not consider differences across programs. They find that many unemployed workers move into programs shortly before expiration. Carling et al. (1996) use data from 1991-1992 to study these issues as well, and they reach similar conclusions.[5] In January 1993, a new large program called ALU ("work experience scheme") was introduced to end the abuse of AMU for benefits entitlement extension. ALU is specifically targeted towards individuals whose benefits entitlement expires. Participation usually amounts to performing tasks in the non-profit private that would otherwise not be carried out. Also, in 1993, the size of other non-AMU programs increased, and other new programs were designed. Again, these programs are much cheaper than AMU.

There are two types of AMU training: vocational and non-vocational. Vocational training courses are provided by education companies, universities, and municipal consultancy operations. The local employment office or the county employment board pay these organizations for the provision of courses. The contents of the courses should be directed towards the upgrading of skills or the acquisition of skills that are in short supply or that are expected to be in short supply. In recent years, most courses concerned computer skills, technical skills, manufacturing skills, and skills in services and medical health care. Vocational training is not supposed to involve the mastering a wholly different occupation

[5] Note that this also suggests that workers do not enjoy training very much, since otherwise they would have entered these programs earlier. Alternatively, caseworkers may have stimulated unemployed individuals to enter programs only shortly before the benefits expiration, or program participation was quantity constrained for individuals with low unemployment durations.

with a large set of new skills.

Non-vocational training (basic general training) concerns participation at courses within the regular educational system, i.e. at adult education centers and universities. Non-vocational training specifically intends to prepare the individual for other types of training (so that the aim of an increased transition rate to work is less direct here). Before 1997, a substantial part of AMU consisted of this non-vocational training. In 1997, a new program of adult education (called the Adult Education Initiative, or Knowledge Lift) has been introduced, and this program is, amongst other things, supposed to replace the non-vocational training part of AMU (see Brännäs, 2000). Nevertheless, for the period since January 1995, non-vocational training amounts to approximately 40% of all AMU courses followed. For 2000 this number is even higher (about 50%).

Concerning UI it should be mentioned that entitlement also requires registration at the employment office. In the mid-1990s, about 40% of the inflow into unemployment and about 65% of the stock of unemployed qualified for UI (Carling, Holmlund and Vejsiu, 2001). Part of the remaining 60% received “cash assistance” benefits, which are typically much lower than UI benefits. The average replacement rate for UI recipients is about 75% (Carling, Holmlund and Vejsiu, 2001).

During the training, the participants' income is called a training grant. Those who are entitled to UI receive a grant equal to their UI benefits level, with a minimum of SEK 240 per day (which is about $ 24). The other participants receive a grant of SEK 143 per day. These payments are made by the UI agency. In case of vocational training, the training organizations have to send in attendance reports, and the grant is withheld in case of non-attendance. In all cases, training is free of charge. In fact, additional benefits are available to cover costs of literature, technical equipment, travel, and hotel accommodation. In this sense, AMU training is far more attractive than regular education.

In Sweden there is a number of other active labor market programs (that is, apart from AMU and the above-mentioned ALU). Most of these concern subsidized employment. See AMS (1998) and Harkman and Johansson (1999) for descriptions of the programs and changes in program participation over time, respectively. In 1997, on average

191,000 individuals (4.5% of the total labor force) participated in one of the programs. The government's part of the total costs of this have amounted to over 3% of GDP (Dahlberg and Forslund, 1999, Regnér, 1997). In fact, Sweden has been the country with the highest percentage of GDP spending on active labor market policies in the world.

The benefits entitlement rules and programs for persons aged below 25 or over 55 differ from those aged between 25 and 55. Young persons must participate in a program after 100 days of unemployment, or otherwise they lose their unemployment benefits. They may use special programs that are not available for other age groups. Persons over 55 receive unemployment benefits for 450 days (instead of 300 days for those aged between 25 and 55).

Dahlberg and Forslund (1999) examine crowding out of non-participants by active labor market programs. They find no significant crowding out effects of AMU.

2.2 Decisions at the micro level

In this subsection we describe the process that leads to an individual's enrolment in AMU. The information is mostly obtained from documents of the Swedish National Labour Market Board (AMS) (see e.g. AMS, 1998) and from in-depth interviews with a number of individual caseworkers.[6] In addition, we rely on Zettermark et al. (2000), who provide a wealth of information on the day-to-day activities of employment offices and caseworkers. Most of that information confirms the interview outcomes.

Usually the employment office advertises, at the office and in the newspapers, about the availability of AMU courses. Most of the offices advertise one or two months before the scheduled starting date. In the advertisement they invite interested individuals to an information meeting. At this meeting individuals are informed about the content of the course and about the eligibility rules. The individuals can usually talk to their personal

[6] We did not use a formal sampling procedure to select caseworkers to be interviewed. Rather, we contacted a number of them to get detailed information concerning the actual decision process at the work floor of the employment offices.

caseworker at the meeting. Those who are interested can then apply to the course.

Enrolment requires approval from the caseworker. The eligibility rules usually include minimum requirements on the educational level upon inflow. However, most courses are at a fairly basic level, so that these requirements are not so restrictive. The caseworker also estimates the individual's “need” for AMU. In practice this means that he examines whether the individual's skills can be enhanced by the course. It is common that the applicants undergo a test in order to find out if they are able to benefit from the course. One may for example test the person's skills in mathematics or in the Swedish language. The test may also include some ability testing. Another way to address whether the individual's skills can be enhanced is by examining his “expected” unemployment duration. This expected duration is thought to be high in case of a low education or an obsolete type of education, or if the individual has an occupation in excess supply. This type of “profiling” is subjective. Sometimes the applicant should write a personal letter that explains why he wishes to participate in a specific AMU-course. If the person has work experience in his occupation, the caseworker might call employer references to ask if they would consider employing the person after the AMU. In general, caseworkers seem to be reluctant to offer an AMU course in a field that is completely different from the occupation of the individual. If an individual rejects a caseworker's offer of an AMU course then in principle the individual's unemployment benefits may be cut off completely, but this does not seem to happen in practice.

Occasionally, caseworkers may work closely with firms that demand certain skill categories. These firms may have an influence on who is accepted into the program. Training (of the unemployed individual) and job search effort (done by his caseworker) may go hand in hand, so the effect of AMU may consist of a skill enhancing effect as well as a search effort effect.

If the number of applicants is insufficient then the course may be canceled (i.e. may not be bought from the course provider). If there are more applicants than slots in a given course, then individuals with high elapsed durations and/or at risk of losing benefits (these are usually the same individuals) are often given priority. However, AMU is generally not offered to individuals if they are primarily concerned about the renewal of their unemployment benefits. It is commonly felt that such practices would not agree with the objective of AMU. Perhaps more importantly, there are in general cheaper alternative programs to deal with such cases, like workfare programs, and efforts are made to push the individual into those programs instead of AMU. Similarly, AMU is generally not offered to individuals who, in the opinion of the caseworker, need practical experience in order to be able to get a job, or just “need something to do” during daytime. In such cases the individual is offered another active labor market program, like a work experience program.

It takes approximately one month from the first information meeting to the first day of the course. On average, the period from application to acceptance takes 2-3 weeks, while the period from acceptance to the start of the course takes 1-2 weeks. An individual may try the AMU-course before actually starting the course. For example, if he is interested in welding then he can make a one-week visit to the school that offers welding courses. Also, individuals may drop out of the course, because they find a job or for other reasons. In fact, in the first case, they are encouraged to do so, and they can come back later and complete the course. An AMU participant may also follow a sequence of courses, starting with basic vocational training and ending in a very narrow type of vocational training. Such a sequence may take 30-40 weeks. The participants do not receive grades or test-based certificates upon finishing a course.

We now show that the above information given by caseworkers on the process that leads to an individual's enrolment in AMU is confirmed by existing empirical studies. Eriksson (1997a, 1997b) analyzes choice and selection into different programs using the HÄNDEL data in combination with survey data on choice and selection by the unemployed as well as the caseworkers. The HÄNDEL data constitute the major administrative data set for our own analyses as well. It is shown that the personal characteristics that are observable in HÄNDEL are not able to give a very precise prediction of actual participation at AMU versus non-participation. The predictive performance can be substantially enhanced if one takes account of self-reported (by the unemployed) measures of the amount with which AMU is expected to have certain advantages for future labor market prospects.

These can be assumed to capture unobserved heterogeneity in the inflow rate into AMU and perhaps unobserved heterogeneity in the treatment effect. (Of course they may also reflect an ex-post rationalization of actual choices made in the past.) Eriksson (1997a) notes that informal interviews with caseworkers reveal that the motivation of the unemployed is a very important criterium for placing an unemployed individual into AMU.

Eriksson (1997b) exploits survey data obtained by letting caseworkers give AMU-advice on the basis of actual files of unemployed individuals that are supplied to them by the survey agency. The allocation of files to caseworkers is fully randomized. The data also allow for a comparison between the valuation of AMU as stated by the caseworkers and the actual (non-)participation of the individual. It turns out that heterogeneity of the caseworkers (which is typically unobserved but is here observed and used as an identifier) is a more important determinant of the caseworkers' stated decisions than the unobserved heterogeneity of the unemployed individuals as captured by fixed effects. So, there is a lot of variation in the caseworkers' decisions which can not be attributed to the unemployed individuals' identities but can be attributed to the caseworkers' identities. When selecting on the basis of observable personal characteristics, officials seem to use rules of thumb which are often not in accordance to the stated goals of AMU on priority groups. If the caseworkers think that an individual would benefit a lot from participation then the individual is also more likely to be an actual participant. But the actual participation also depends on the unemployed individual and on unexplained factors.

Carling and Richardson (2001) use the HÄNDEL data from 1995 onwards to study the choice of a particular type of training program conditional on going into one of these programs. They use a Multinomial Logit model for this. They find that employment agency identifiers have significant effects, and that these dominate the effects of characteristics of the unemployed individual.

According to Eriksson (1997b), caseworkers are reluctant to let current participants to non-AMU programs enter AMU. Also, work experience programs and public temporary employment are substitutes for each other but not for AMU. Caseworkers regard AMU to be a fundamentally different kind of program. So the variation in the caseworkers' behavior with respect to AMU mostly concerns the choice between AMU and no AMU, instead of the choice between AMU and another program. According to Dahlberg and Forslund (1999), nowadays, AMU is typically not used for UI entitlement extensions.

3 The model framework

3.1 A class of bivariate duration models for treatment evaluation

We normalize the point of time at which the individual enters unemployment to zero. The durations T_u and T_p measure the duration until employment and the duration until entry into the AMU training program, respectively. At this stage we assume that unemployment can only end in employment, and we take the period in AMU as part of the unemployment spell. Also, for the moment we ignore time spent in other training programs. As a result, T_u also denotes the duration of unemployment. The population that we consider concerns the inflow into unemployment.

We assume that the individual distribution of T_p can vary with observed and unobserved explanatory variables x and v_p, respectively. Similarly, we assume that the individual distribution of T_u can vary with observed and unobserved explanatory variables x and v_u and with the realized value of T_p of that individual. To construct a model, it is useful to focus on the *hazard rates* of T_p given x,v_p and T_u given x,v_u,T_p. The hazard rate of a duration variable is the rate at which the spell is completed at time t given that it has not been completed before, as a function of t. It provides a full characterization of the duration distribution (see Lancaster, 1990, and Van den Berg, 2001). Somewhat loosely, it is the speed at which the duration is realized.[7] We use notation $\theta_p(t|x,v_p)$ and $\theta_u(t|T_p,x,v_u)$,

[7] For a nonnegative random (duration) variable T, the hazard rate is defined as $\theta(t) = \lim_{dt \downarrow 0} \Pr(T \in [t,t+dt) |T \geq t)/dt$. Consider the distribution of a duration variable conditional on some other variables. It is customary to use a vertical "conditioning line" within the argument of a hazard rate in order to distinguish between (on the left-hand side) the value of the duration variable at which the hazard rate is evaluated, and (on the right-hand side) the variables that are conditioned upon.

respectively. We do not require that $v_p = v_u$ but we allow them to be dependent. The variable x may be different in θ_p and θ_u.

As noted in the introduction, we are interested in the causal effect of participation in AMU on the exit out of unemployment. The treatment and the exit are characterized by the *moments* at which they occur, so we are interested in the effect of the realization of T_p on the distribution of T_u. We assume that the realization t_p of T_p affects the shape of the hazard of T_u from t_p onwards, in a deterministic way. The assumption implies that the causal effect is captured by the effect of T_p on $\theta_u(t|T_p,x,v_u)$ for $t > T_p$. We adopt the following specification of the hazard rates $\theta_u(t|t_p,x,v_u)$ and $\theta_p(t|x,v_p)$,

$$\theta_p(t|x,v_p) \;=\; \lambda_p(t) \cdot \exp(x'\beta_p) \cdot v_p$$

$$\theta_u(t|T_p,x,v_u) \;=\; \lambda_u(t) \cdot \exp(x'\beta_u) \cdot \delta(t|T_p,x)^{I(t > T_p)}\, v_u$$

where I(.) denotes the indicator function, which is 1 if its argument is true and 0 otherwise.

Apart from the term involving $\delta(t|T_p,x)$, the above hazard rates have Mixed Proportional Hazard (MPH) specifications. The function $\lambda_i(t)$ is called the "baseline hazard" since it gives the shape of the hazard rate θ_i for any given individual. The hazard rate is said to be duration dependent if its value changes over t. Positive (negative) duration dependence means that $\lambda_i(t)$ increases (decreases). The term $\exp(x'\beta_i)$ is called the "systematic part" of the hazard. Finally, the term v_i is called the "unobserved heterogeneity term". MPH models are the universally most popular reduced-form duration models in econometrics (see Van den Berg, 2001, for a survey).

The term $\delta(t|T_p,x)^{I(t > T_p)}$ captures the AMU effect. Clearly, AMU has no effect if and only if $\delta(t|T_p,x) \equiv 1$. Now suppose $\delta(t|T_p,x)$ is equal to a constant larger than one. If T_p is realized then the level of the individual exit rate to employment increases by a fixed amount. This will reduce the remaining unemployment duration in comparison to the case where AMU is entered at a later point of time. More in general, we allow the effect of

AMU to vary with the moment T_p of entry into AMU and with x. Moreover, the individual effect may also vary over time, as we allow it to depend on the elapsed unemployment duration t. As a result, the individual effect may also vary with the time $t-T_p$ since entry into AMU (in fact in the empirical analysis we only consider variation of the effect with $t-T_p$ and not with t or T_p separately). The effect of $t-T_p$ may capture that the exit rate is low during the training course or high immediately after the end of participation. We return to the details of this below.

The AMU effect cannot be inferred from a direct comparison of realized unemployment durations of individuals with a given T_p to the realized unemployment durations of other individuals. If the individuals who enter AMU at t_p have relatively short unemployment durations then this can be for two reasons: *(1)* the individual causal AMU effect is positive, or *(2)* these individuals have relatively high values of v_u and would have found a job relatively fast anyway. The second relation is a spurious selection effect. If this is ignored then the estimate of the AMU effect may be inconsistent.

Recall that the same vector x may affect both hazards and that we allow for the possibility that v_u=v_p. This means that we allow individuals to be aware of the existence of the AMU, and we allow them to influence both the rate of entry into AMU and the rate of exit into employment. We do not assume that we observe determinants of training assignment that the individual does not use himself to update his strategy. Furthermore, it should be noted that we may allow x to vary over time.

The data provide observations on T_u and x. In addition, if T_p is completed before the realization of T_u then we also observe the realization of T_p, otherwise we merely observe that T_p exceeds T_u. In addition, the data provide multiple spells, i.e. for individuals in the sample we may observe more than one unemployment spell. We assume that an individual has a given time-invariant value of (v_u,v_p) and that, given these values and x, the spells of an individual are independent. Since v_u and v_p are unobserved, the duration variables given x are not independent across spells. Intuitively it is plausible that, the more individuals with multiple spells in the data, the less sensitive the results are with respect to the assumptions underlying the model framework. Basically, with multi-spell data, the empirical setting is similar to standard panel data analysis with fixed effects. It should be emphasized that this requires the assumption that the unobserved explanatory variables do not vary across spells. In reality these variables may change in between two consecutive unemployment spells, for example because of the accumulation of specific types of work experience.

A number of assumptions are implicitly captured by the model specification. First of all, according to the model, the realization of entry into AMU training at say t_p does not have an effect on the individual's exit rate θ_u prior to that moment t_p. The individual's exit rate at t is the same irrespective of whether training will occur at t+1 or whether it will occur at t+100. This basically rules out anticipatory effects of the training. If an individual does anticipate participation in AMU at a particular future date t_p then he may want to wait for the treatment by reducing his search intensity for jobs, *i.e.* he may change his strategy and this may decrease the probability that T_u is quickly realized. If this is ignored in the empirical analysis then the training effect may be over-estimated. However, if the time span between the moment at which the anticipation occurs and the moment of the actual training is short relative to the durations T_p and $T_u – T_p$, and if the anticipatory effect is not very large, then estimation results may be rather insensitive to the assumption of no anticipation.

With well-established programs like AMU, it is plausible that *determinants* of the training assignment affect the individual's exit rate out of unemployment *before* the actual entry into training. For example, at any time before participation in AMU, the unemployed workers may search less because they know that there is a probability that their skills can be enhanced by AMU at some point during unemployment. In that case the program is said to have an *ex ante* effect on exit out of unemployment. Such an effect should not be confused with anticipation of the *realization* of entry into training, because in the letter case the individual knows the stochastic outcome rather than the determinants of the process. Likewise, absence of anticipation does not mean that ex ante effects of AMU are ruled out.

The ex ante effect can be contrasted to the *ex post* effect of training, which is the effect of actual training on the individual exit rate - this is of course the effect we focus on in this paper. The ex ante effect is an example of the macro effects that are present in a world in which a particular program is implemented. There may also be ex ante or macro effects on the magnitude and composition on the inflow into unemployment and on the behavior of employers. We do not estimate ex ante effects of AMU, but the model is compatible with ex ante effects. Individuals may know determinants of the process leading to training, including the probability distribution of the duration until training, but they do not know in advance the realizations of this process.

A different type of anticipation occurs if the future realization of the variable of interest T_u has an effect on the current level of θ_p. In reality, an individual may have private knowledge on a future job opportunity that is independent of whether the training will occur, and the individual may use this knowledge to avoid training. However, the model specification rules out that individuals anticipate the future outcome of T_u and use this to modify their strategy which would in turn affect the rate at which entry into training occurs. If something like does occur in reality and is ignored in the model then a positive effect of training on exit to employment is under-estimated. On the other hand, if the training course takes a long time, then such an effect may be unimportant, as employers may be unwilling to wait for a new employee for many months. Also, if the time span between the moment at which the anticipation occurs and the moment of the actual exit to work is relatively short, and if the anticipatory effect is not very large, then estimation results may be rather insensitive to this. Again, absence of anticipation does not rule out that individuals know the determinants of the process leading to employment and use these as inputs in their decision problem. For example, the individuals may know that $\lambda_u(t)$ increases in the near future, and modify their strategy accordingly, which may affect their θ_p. The latter can be captured in the model by way of $\lambda_p(t)$.

A more technical aspect of the model specification follows from the fact that we specify the assignment of training by way of specifying the hazard rate of a duration distribution. This approach implies that there is a random component in the assignment that is independent of all other variables (see e.g. Ridder, 1990, and Abbring and Van den Berg, 2000). This resembles the role of the error term in a regression equation. Intuitively it is clear that if there is not much variation in the moment of entry into AMU then it is difficult to address its effect. In the extreme case where individuals can only enter AMU at, say, exactly one year after flowing into unemployment, it is impossible to distinguish any effect of AMU from the duration dependence effect on the exit rate to work. In that case it is of course also impossible to justify that entry into AMU is not anticipated.

3.2 Applicability of the model framework to AMU

In this subsection we argue that the above model is particularly well suited for our study of the AMU program. We focus on the following issues: dependent (unobserved) heterogeneity, randomness in the (moment of) treatment assignment, absence of anticipatory effects, and absence of substitution with other programs.

From the information in Subsection 2.2 and from the studies by Eriksson (1997a, 1997b), it is obvious that unobserved (to us) heterogeneity of the unemployed individuals plays an important role in the assignment to AMU. The corresponding variables taken into account by the caseworker (like motivation, subjectively assessed expected unemployment duration, and subjective assessments of other aspects of the future career) are also indicative of unobserved determinants of the individual exit rate to work. The empirical analysis should therefore take account of potentially related unobserved heterogeneity terms in θ_u and θ_p.

If the individual knows that a variable is an important determinant of treatment assignment (like the amount and type of discretionary behavior of his caseworker), and the individual knows that he may be subject to treatment, then he has a strong incentive to inquire the actual value of the variable. Subsequently, he will take his value of the variable into account to determine his optimal strategy, and this strategy in turn affects the rate at which he moves to employment. The variables that are observed by us and that may have an effect on assignment to AMU are also observable to the individuals under consideration. Therefore we allow the same set of x variables to affect θ_p and θ_u.

Now let us consider the presence of randomness in the moment of entry into AMU. To some extent this may be generated by changes in the behavior of the caseworker or the employment agency that are beyond the observation window of the unemployed individual. More importantly, it is generated by the variation in the moment at which AMU courses start. In addition, admission to a course may depend on the extent to which other individuals apply to the course, which is random from the individual's point of view. Recall that Eriksson (1997b) finds residual variation in the AMU assignment process that can not be attributed to the individual or the caseworker.

We now turn to anticipation of the moment of entry into AMU. From Subsection 2.2, the time period between the moment at which the individual is informed about the possibility of enrolling into an AMU course and the moment at which the course starts is very short. (Of course, we allow individuals to be aware of the existence of the AMU *program*.) There are however two reasons for why individuals may anticipate the moment of entry, and both of these lead us to restrict the scope of the empirical analysis somewhat.

First, as shown in Section 2, in 1991 and 1992 AMU was often used to extend benefits entitlement. In that case, the date of inflow into AMU is mostly determined by the date of expiration of benefits entitlement. The latter date is known in advance by the unemployed individual and his caseworker (this date does not vary much across the unemployed; see the references). This allows for anticipation of the inflow into AMU, which violates a key assumption of our empirical methodology. Moreover, such self-selection into AMU is governed by different motives than self-selection in other years, so we may expect the unobserved heterogeneity distribution to be different across time. From January 1993 onwards, other programs took over its role as means to extend benefits entitlement. We therefore restrict attention to data from 1993 onwards.

Secondly, recall from Section 2 that part of AMU concerns non-vocational training (in particular before 1997). Non-vocational training primarily aims to prepare the individual for other types of training. It is often given within the regular school system. This implies that the starting date of the non-vocational training is often determined by institutional features of the school system, like the starting dates of the school seasons. As a result, it is easy for an unemployed individual to anticipate the date of inflow into such a program. We therefore restrict ourselves to vocational training. There are actually two other reasons to do so. First, vocational training is relatively expensive. Secondly, vocational training is difficult to get in other programs, whereas non-vocational training is easier to get elsewhere, so that in the latter case there are substitution possibilities.

Concerning substitution possibilities, recall from Subsection 2.2 that caseworkers regard vocational AMU training as a very different type of program than the other active labor market programs. The latter are regarded to be substitutable to a high degree. For persons under 25, there are programs that are more similar to AMU vocational training. Also, for these individuals, the similarity with courses and tracks in the regular school system may be important. For this reason we restrict attention to individuals aged over 25. Also, young individuals must enter a training course after 100 days of unemployment, which may generate anticipatory effects. We omit individuals over 55 because they face a different unemployment benefits system and because for them vocational AMU training seems to have relatively small advantages.

It follows from the above that our model framework may be less suited for the analysis of the effects of the other active labor market programs on unemployment duration. With other programs, individuals may anticipate their enrolment a long time in advance, because of their link to benefits entitlement expiration and/or because of their connection to the regular school system. Moreover, it is difficult to analyze them in isolation from each other because of the high degree of substitutability.

4 The data

The data are based on a combination of the administrative data sets called HÄNDEL (from the official employment offices) and AKSTAT (from the unemployment insurance fund). For the present project, the most important source concerns HÄNDEL, which contains information on unemployed individuals' training activities and work experience activities. These data cover all registered unemployed persons since August 1991 (approximately 2 million observations), and they contain detailed information on the types of training as well as the starting and ending dates of the participation in the program. According to Carling, Holmlund and Vejsiu (2001), more than 90% of the individuals who are ILO-unemployed[8]according to labor force surveys also register at the employment offices. The AKSTAT data are available from 1994 onwards and provide information on the wage level and working hours in the of job prior to the spell of unemployment, for individuals who are eligible for UI. The full HÄNDEL data are also informative on whether an individual in AMU gets vocational training or non-vocational training.

Our observation window runs from January 1, 1993 until June 22, 2000. The unit of observation is an individual. For each individual who is in HÄNDEL at least once during the observation window, we can construct an event history from the HÄNDEL data. For the spells of unemployment (to be defined below), the information in HÄNDEL and AKSTAT is used to make a list of characteristics at the beginning of the spell, and a list of dates at which changes occur, including the nature of the change. It is particularly important to include the information on participation in non-AMU programs, since such participation may rule out a transition to AMU, or may at least reduce the transition rate to AMU and/or work.

We only use information on individuals who become unemployed at least once within the observation window. An individual becomes unemployed at the first date at which he registers at the employment office as being "openly" unemployed. This eliminates registration spells that start because the individual wants to change employer and also eliminates spells that start because the individual knows that he is going to be unemployed in the future (short term contract or notification of lay-off), at least until the individual does actually become unemployed. We also ignore unemployment spells that are already in progress at the beginning of the observation window, because using them would force us to make assumptions about the inflow rate before the beginning of the window. We thus obtain a so-called inflow sample of unemployment spells, and we follow the corresponding individual over time after this moment of inflow. (Note that we also use information

[8] The unemployment definition of the ILO (International Labour Organization) states that the individual must be without employment, actively searching for employment, and currently available for employment.

available on the period prior to such spells, notably on wages.) In addition, we exclude individuals who have experienced unemployment between August 1991 and January 1, 1993. August 1991 is the first month for which we have information on the individual's labor market state. The behavior of individuals who have been unemployed shortly before January 1993 may be different from that of those who have not.

For convenience, we use the term unemployment spell to include possible spells in AMU, relief work, ALU, etc. The spell ends if the individual leaves the employment office register or if he moves from the unemployment categories in the employment office register to a non-unemployment category in the register. If the exit destination is employment then we observe a realization of the duration variable of interest. If the exit destination is different (e.g. "regular education", or "other reason") then this duration variable is right-censored. The duration is independent right-censored if the spell is continuing at the end of the observation window. If exit occurs into "wage subsidy" or "(public) sheltered employment" then we remove the individual from the sample, since these programs are for handicapped people (who are typically are not in open unemployment anyway). As a result, our data set contains 500,960 individuals. Note that by following the individuals over time we may observe multiple unemployment spells per individual.

Occasionally, we observe coding errors in data at points of time at which individuals move between different categories in the register. Obvious typing errors are corrected, whereas otherwise we right-censor the duration variables at the moment at which such an error occurs.

If we would treat participation in other programs before participation in AMU as regular unemployment, then the transition rate from unemployment into AMU would be extremely low during the participation in the other programs. Participation in non-AMU programs most likely also reduces the transition rate into employment, so, during such a period of program participation, it may be preferable to halt the time clock of the duration until regular employment. So, as our baseline assumption, the time spent in training (in non-AMU programs as well as in AMU) does not contribute to the unemployment duration, and the time spent in other training programs does not contribute to the duration until

AMU. Note that this also means that time spent in non-AMU programs after AMU does not contribute to the unemployment duration. We address these assumptions in sensitivity analyses.

As mentioned in Subsection 3.2, we restrict attention to individuals who were at least 25 and below 55 at the moment they enter unemployment.

We distinguish between the following levels of education: junior high school or lower, short senior high school, long senior high school, short tertiary education, and long university degree or higher. These are roughly equivalent to ≤ 9, 10-11, 12-13, 14, and ≥ 15 years of education, respectively. Concerning nationality we also distinguish between three categories: Eastern Europe, Africa / Asia, and otherwise (including Sweden). Concerning the type of unemployment benefits received during unemployment we distinguish between three categories: UI, cash allowance, and neither. For UI recipients in 1994 and beyond, the AKSTAT data contain the hourly wage earned in the job that was held just before the onset of the spell of unemployment. This is almost linearly related to their UI level (see e.g. Carling, Holmlund and Vejsiu, 2001). For non-UI-recipients the wage variable is set to zero. This is also done for UI recipients who become unemployed and subsequently employed within 1993. However, if they move back to unemployment in 1994 we use the corresponding pre-unemployment wage as a proxy of the pre-unemployment wage for the unemployment spell in 1993.

The analyses are based on a 1% random subsample of the full data set at our disposal. For each individual we include at most 3 unemployment spells. This results in 5010 individuals with, in total, 8656 unemployment spells. We allow the x variables to differ across the spells of a given individual. For example, age and the pre-unemployment wage differ across different spells.

Of these spells, 656 contain a period of participation at an AMU course, and exactly 8000 do not. Some of the latter, of course, are right-censored due to the finiteness of the observation window, so in reality some of them may include AMU participation afterwards. The median T_p across the 656 spells that are observed to include participation is 161 days. Table 1 provides some summary statistics. The table takes T_p:=T_m if T_p is not realized.

Of the 656 spells that are observed to include AMU participation, 27% are also observed to include participation at another type of active labor market program before participation at the AMU program. Not surprisingly, this happens predominantly in long spells with a high realized value of T_p. Of the 328 spells with T_p smaller than its median of 161 days, only 12% are also observed to contain participation at another type of active labor market program before AMU participation. Of the 8000 spells that are not observed to include participation at AMU, 18% are observed to contain participation at another type of active labor market program. If we restrict focus to spells with T_u smaller than 161 days then this figure drops to 10%. Participation at other programs thus does not seem to be related to participation at AMU. The fact that spells with AMU participation relatively often also contain participation at other programs is because of the fact that by conditioning on AMU participation we condition on high realized durations.[9]

[9] More information on how the data set is constructed is available in an appendix available on request.

Table 1. Summary statistics for the 1 % sample. The unit of interval is one spell. Standard deviations in parentheses. The observed labor market outcomes are reported in fractions and in days.

	All spells	No AMU	With AMU
Individual characteristics			
log(age)	3.54 (0.23)	3.54 (0.23)	3.58 (0.22)
short senior high school	0.26	0.26	0.29
senior high school	0.20	0.20	0.20
short tertiary education	0.05	0.05	0.05
university	0.16	0.16	0.14
female	0.50	0.50	0.44
UI recipient	0.67	0.68	0.70
cash allowance recipient	0.07	0.07	0.09
from Eastern Europe	0.05	0.05	0.08
from Africa or Asia	0.05	0.05	0.05
log (hourly wage)	2.69 (2.31)	2.66 (2.32)	3.05 (2.14)
experience in occupation (dummy)	0.64	0.63	0.71
education in occupation (dummy)	0.62	0.61	0.65
professional and technical work	0.15	0.15	0.14
health, nursing and social work care	0.14	0.15	0.09
adm., managerial and clerical work etc.	0.13	0.12	0.18
sales	0.11	0.11	0.11
agriculture and mining	0.07	0.07	0.08
services (incl. not categorized occ.)	0.17	0.17	0.15
large city (dummy)	0.52	0.53	0.45
needs guidance (dummy)	0.08	0.07	0.14
willing to move (dummy)	0.15	0.15	0.18
accepts part time work (dummy)	0.05	0.06	0.03

The Table continues on next page.

Table 1 (Cont.) Summary statistics for the 1 % sample. The unit of interval is one spell. Standard deviations in parentheses. The observed labor market outcomes are reported in fractions and in days.

fraction of spells that starts in:			
1993	0.19	0.19	0.27
1994	0.16	0.16	0.21
1995	0.16	0.15	0.20
1996	0.13	0.13	0.10
1997	0.12	0.12	0.09
1998	0.11	0.11	0.09
1999	0.09	0.10	0.03
2000	0.04	0.05	0.01
Observed labor market outcomes			
spells contains AMU	0.08	0	1
spells ends in exit to work	0.57	0.58	0.53
realized t	170 (214)	153 (191)	372 (342)
realized t	158 (191)	153 (191)	210 (195)
time spent in AMU	9 (47)	0	124 (120)
time spent in other programs	38 (111)	38 (112)	46 (105)

5 The empirical analysis

5.1 Parameters

For the duration dependence functions and the bivariate unobserved heterogeneity distribution we take flexible specifications. We take both $\lambda_u(t)$ and $\lambda_p(t)$ to have a piecewise constant specification. This means that the value of λ_i is constant within duration intervals. In most of the empirical analyses we take 8 intervals for λ_u and 6 for λ_p. In both cases the length of an interval is 56 days, except for the last intervals which are unbounded from the right. The parameter λ_{ij} denotes the value of λ_i in the j^{th} interval. As an example, $\lambda_{u1} > \lambda_{u2}$ means that, everything else equal, the exit rate to work is higher during the first 56 days of the unemployment spell than between 56 and 112 days.

We assume that both v_u and v_p can take on two possible values, such that four combinations are possible. Each of these four possible outcomes for v_u,v_p has an associated probability. We estimate the possible values of the v_i as well as the probabilities. The joint unobserved heterogeneity distribution thus adds 7 unknown parameters to the model. This specification is popular, flexible, and computationally feasible (see Van den Berg, 2001, for an overview). In the Appendix we examine the specification in some more detail.

5.2 Estimation results

5.2.1 The basic model

We estimate the models using the method of Maximum Likelihood. We take the unit of time to be one day. The baseline set of parameter estimates is displayed in Table 2. These are obtained by estimation of the model under the following assumptions: δ is a constant, the lengths of the time intervals spent within AMU and within other programs are set to zero, and within a spell any subsequent participation in AMU after the first course is ignored. We include data on as many as three unemployment spells per individual, if available for the individual.

For the categorical variables in x we have the following baseline categories: education = less than short senior high school, gender = male, unemployment benefits type = none, nationality = not in Eastern Europe, Africa or Asia, and occupation type = manufacturing. Log age and log hourly wage in the previous job are measured in deviation from their mean across the 8656 spells. The "constant terms" in θ_u and θ_p are represented by the means of v_u and v_p, respectively. This is why we normalize λ_{u1}=λ_{p1}=0 and why x does not include a constant.

The main parameter of interest is δ, which represents the effect of training on the transition rate to work. The estimated value of δ is 0.83 and is significantly different from 0. Training thus raises this transition rate with slightly more than 100%, which means that it

Table 2. Estimation results for the baseline model. Standard errors in parentheses. The superindex * denotes significance at 5% level (only for elements in β_i and λ_i (with i = u, p) and δ).

	To work, θ_u		To AMU training, θ_p	
AMU effect				
δ	0.83	(0.10) *		
Individual characteristics				
log(age)	-0.59	(0.11) *	0.35	(0.29)
short senior high school	-0.01	(0.06)	0.20	(0.16)
senior high school	-0.10	(0.07)	0.05	(0.18)
short tertiary education	0.08	(0.12)	-0.13	(0.29)
university	0.23	(0.09) *	-0.14	(0.22)
female	0.01	(0.06)	-0.04	(0.14)
UI recipient	0.26	(0.07) *	0.04	(0.18)
cash allowance recipient	0.21	(0.10) *	0.43	(0.23) *
from Eastern Europe	-0.58	(0.13) *	0.35	(0.25)
from Africa or Asia	-0.85	(0.14) *	0.18	(0.28)
log(hourly wage)	0.00	(0.03)	0.16	(0.09)
experience in occupation (dummy)	0.17	(0.05) *	0.12	(0.14)
education in occupation (dummy)	0.23	(0.05) *	0.12	(0.13)

The Table continues on next page.

Table 2 (Cont.) Estimation results for the baseline model. Standard errors in parentheses. The superindex * denotes significance at 5% level (only for elements in β_i and λ_i (with i = u, p) and δ).

professional and technical work	-0.16	(0.09) *	0.08	(0.22)
health, nursing and social work	-0.08	(0.08)	-0.29	(0.24)
adm., managerial and clerical work	-0.39	(0.09) *	0.34	(0.22) *
sales	-0.29	(0.09) *	0.03	(0.22)
agriculture and mining	-0.08	(0.09)	0.09	(0.23)
services (incl. non categorized occ.)	-0.17	(0.09) *	-0.06	(0.19)
large city (dummy)	-0.08	(0.05)	-0.31	(0.12) *
needs guidance (dummy)	-0.48	(0.11) *	0.54	(0.20) *
willing to move (dummy)	0.04	(0.06)	0.18	(0.16)
accepts part time work (dummy)	0.11	(0.09)	-0.46	(0.32)
1994	0.24	(0.07) *	-0.07	(0.18)
1995	0.20	(0.07) *	-0.04	(0.17)
1996	0.24	(0.08) *	-0.48	(0.22) *
1997	0.53	(0.08) *	-0.44	(0.22) *
1998	0.58	(0.08) *	-0.37	(0.22) *
1999	0.75	(0.09) *	-1.07	(0.22) *
2000	0.75	(0.15) *	-0.69	(0.68)

The Table continues on next page.

Table 2 (Cont.) Estimation results for the baseline model. Standard errors in parentheses. The superindex * denotes significance at 5% level (only for elements in β_i and λ_i (with i = u, p) and δ).

	To work, θ_u		To AMU training, θ_p	
Duration dependence				
λ_{i2}	0.11	(0.05) *	-0.32	(0.15) *
λ_{i3}	0.05	(0.06)	-0.34	(0.17) *
λ_{i4}	0.10	(0.08)	-0.18	(0.18)
λ_{i5}	0.04	(0.09)	0.05	(0.18)
λ_{i6}	-0.17	(0.12)	-5.00	(0.19) *
λ_{i7}	-0.13	(0.13)		
λ_{i8}	-0.19	(0.11) *		
Unobserved heterogeneity				
log (v_1)	-5.82	(0.11)		
log (v_2)	-7.41	(0.16)		
log (v_3)			-8.02	(0.68)
log (v_4)			-7.00	(0.60)
q_{13}		-1.00	(1.19)	
q_{14}		0.94	(2.64)	
q_{23}		1.32	(3.81)	
log likelihood value		-22068.5		
number of individuals		5010		

more than doubles. The effect on the mean or median unemployment duration depends on the moment at which training occurs. If the training is given within the first month then the mean duration is more or less reduced by half. Similarly, training at a relatively early stage in an unemployment spell has a large effect on the probability of long-term unemployment. (Of course, such a policy can be costly if implemented on a wide scale.) Recall that (part of) the effect may be due to increased search effort on the part of the caseworker, especially when the individual's period of AMU participation comes to an end.

Now let us turn to the covariate effects on the transition rate to work. Not surprisingly, this rate is lower for older and non-Swedish individuals. It is higher for individuals with two years of high school and for university graduates. It is also higher for UI recipients, reflecting the stronger labor market attachment of these individuals. The disincentive effect of high UI benefits seems to be captured by the negative effect of a high previous wage on the exit rate to work, although this effect is insignificant. The interpretation of the calendar year effects is complicated by the fact that some vocational courses have been recoded as non-vocational courses during the time span of the data (Zettermark et al., 2000).

The estimated duration dependence of θ_u is such that the individual transition rate to work decreases as the duration increases. Apparently, stigmatization and discouraged worker effects play a significant role here. Also, some individuals may enter a loop of successive periods of unemployment and workfare.

Most observed individual characteristics have an insignificant effect on the rate θ_p at which the individual enters AMU training. If the individual receives cash allowance then this rate is higher. This may be due to the bad financial circumstances of such individuals. The rate at which individuals enter AMU fluctuates somewhat during the first 300 working days of unemployment. After that it is dramatically lower. Recall that UI recipients need to participate in some active labor market program after 300 working days of unemployment in order to extend their benefits entitlement, and that they do not use the AMU program for this.

As a first informal check on the robustness of the estimates, we compare them to those obtained from the misspecified model in which it is imposed that there is no unobserved heterogeneity. In that case the parameters of θ_u can be estimated in isolation from those in θ_p. The results are in Table A1 in Appendix. The constant term in θ_u is now represented by λ_{u1}, so that the estimates of the other λ_{ui} are now lower than in Table 2, with an order of magnitude equal to the estimated λ_{u1}.

There are no spectacular differences between the estimates of the θ_u parameters in Tables 2 and Table A1. Typically, when unobserved heterogeneity is ignored in duration analysis, the estimated duration dependence is more negative (i.e., θ_u decreases more over time), and the estimated covariate effects are smaller (see e.g. Lancaster, 1990, and Van den Berg, 2001).

5.2.2 Heterogeneous treatment effects on the individual transition rate to work

So far in this subsection we have assumed homogeneity of the treatment effect δ on the exit rate to work over individuals and over time. (Of course, the treatment effect on other outcomes of interest, like the mean duration or the fraction employed within a year is heterogeneous, due to the nonlinear way in which they depend on δ and x, v_u, v_p.) We now allow for heterogeneous treatment effects. First, we allow δ to be a non-constant function of the time $t - t_p$ that has elapsed since AMU participation. As we have seen in Subsection 2.2, there are reasons to suspect that the effect is smaller if this elapsed time is large. Also, the data show that many individuals move to employment the day they leave training.

To capture this, we take δ to be a piecewise constant function of $t - t_p$. Specifically, $\delta = \delta_1$ if $0 \leq t - t_p \leq 28$ days, and $\delta = \delta_2$ if $t - t_p > 28$ days. Alternatively, one could extend the model by incorporating real time spent in training and allowing for a time-dependent transition rate from training directly to employment. We return to this below.

Table 3 gives the estimates for δ_1 and δ_2. Clearly, the training effect is very large right after the training participation period. It is three times as likely to move to employment within a month after AMU training, in comparison to when the individual would not have participated in the training. After the first month, the effect is still positive, but it is much smaller in magnitude.

Table 3. Estimation results for the training effect on the transition rate to work when this is allowed to depend on the elapsed time since training. Standard errors in parentheses. The superindex * denotes significance at 5% level.

	To work, θ_u	
AMU effect		
δ_1 (i.e., ≤28 days)	1.24	(0.13) *
δ_2 (i.e., >28 days)	0.29	(0.13) *
log likelihood value	-22057.8	
number of individuals	5010	

For sake of brevity we do not report the other parameter estimates for this extended model. The estimates of the covariate effects β_u and β_p and their standard errors are the same as in Table 2. This is also true for the estimates of the duration dependence λ_p. The estimated duration dependence λ_u is slightly less negative, which is not surprising given that now $\delta(t-t_p)$ has become a source of negative duration dependence as well. The estimates of the unobserved heterogeneity distribution also change slightly.

The value of the test statistic of the likelihood ratio test of $\delta_1 = \delta_2$ equals 21.4 (see the log likelihood values reported in Tables 2 and 3). As this statistic has a chi-square distribution with one degree of freedom under the null hypothesis, we conclude that this null hypothesis is rejected. The training effect on the exit rate to work is mostly short-run. Note incidentally that this supports our assumption that training effects do not cross over to subsequent spells.

The results on δ_1 and δ_2 may indicate that job search effort by the caseworker is an important ingredient of the treatment. An alternative explanation is that trained workers who do not find a job within one month after finishing training become stigmatized, so that their chances to find a job decrease. Yet another explanation is that the individual treatment effects are heterogeneous across individuals, so that the decreasing shape of $\delta(t-t_p)$ reflects dynamic sorting. The individuals who benefit a lot from the course find a job quickly, and those who do not benefit from it remain unemployed longer. The heterogeneity may be due to heterogeneity of individual characteristics or due to heterogeneity of characteristics of the training course. The results suggest that human capital accumulation by itself is not a good explanation for the training effect. After all, it is unlikely that the human capital acquired in AMU becomes obsolete within one month.

To proceed, we examine individual heterogeneity in the treatment effect. Individuals with a certain level of education, or with a particular gender or nationality may not be able to benefit as much from training as other individuals. For example, individuals with a high level of education may not benefit simply because not many courses are available at an academic level. We investigate this by allowing δ to depend on the level of education, gender and nationality. Specifically, δ is allowed to have a different value if the individual has short tertiary or university education.

The main results are in Table 4. As expected, the estimated training effect is smaller for those with a high level of education. It is significantly different from zero. The estimated effects for women and immigrants are not significantly different from zero. Again, we do not report the other parameter estimates for this extended model, because they are the same as in Table 2 (even the β_u effects of level of education). The likelihood ratio test statistic of the null hypothesis that δ does not depend on the explanatory variables has the value 6.4. Under the null hypothesis, this test statistic has a chi-square distribution with three degrees of freedom. This leads to acceptance of this null hypothesis at the 5% level (though not at the 10% level).

Table 4. Estimation results for the training effect on the transition rate to work when this is allowed to depend on gender, educational level and nationality. Standard errors in parentheses. The superindex * denotes significance at 5% level.

	To work, θ_u	
AMU effect		
main effect	0.94	(0.13) *
woman	-0.00	(0.15)
high education	-0.40	(0.20) *
Immigrant from Africa, Asia, or Eastern Europe	0.11	(0.26)
log likelihood value	-22065.3	
number of individuals	5010	

5.2.3 Time in training and in other programs

We now replace the rule that the lengths of the time intervals spent within AMU and within other programs are set to zero by the rule that the time clock keeps on running during such periods. For T_u this is more appropriate if individuals move to employment at the same rate within such periods as they do when they are "openly" unemployed. For T_p this is more appropriate if individuals move into AMU at the same rate when they are in other programs as they do when they are "openly" unemployed. We also let the treatment effect work from the moment the individual *enters* AMU. We use the symbol Δ to denote the treatment effect parameter in the exit rate to work, and we assume that this parameter is constant over time. Somewhat loosely, Δ captures the average of the effect during participation and the effect after participation, where the latter was captured by the parameter δ in Subsection 5.2.1. Since exit to work during the first months in AMU is very rare, we expect the estimate of Δ to be smaller than the estimate of δ in Subsection 5.2.1.

The parameter estimates are reported in Table 5. Most are similar to those in Table 2. However, the estimate of the parameter of interest Δ is virtually equal to zero. The estimate of Δ is a compromise between the very low transition rate from AMU to work during the first months and the very high transition rate from AMU to work after that.

In this extension, T_u is real time spent out of work (i.e., in unemployment and in training programs and in other active labor market programs) since inflow into unemployment. Also, the treatment effect Δ summarizes the effect of entering AMU on the total time out of work. We conclude that the net effect on the individual's time spent out of work is about zero. Thus, from this point of view, the program does not appear to be cost-effective.

We now turn to the other estimates. The duration dependence of the inflow rate into AMU is more negative than before. This reflects the fact that individuals rarely move from other programs directly into AMU. The fact that the β_i parameter estimates are similar to before means that they are insensitive to whether we include time spent in other programs or not. Of course, other programs may have their own causal effect on θ_u, but an analysis of that raises new selection problems and would be beyond the scope of this paper.

Table 5. Estimation results for the training effect on the transition rate to work when real time in programs are included and the training effect works from the moment of entering (instead of leaving) AMU. Standard errors in parentheses. The superindex * denotes significance at 5% level (only for elements in β_i and λ_i (with i = u, p) and Δ).

	To work, θ_u		To AMU training, θ_p	
AMU effect				
Δ	0.11	(0.11)		
Individual characteristics				
log(age)	-0.71	(0.11) *	0.59	(0.30) *
short senior high school	0.04	(0.06)	0.18	(0.16)
senior high school	-0.08	(0.07)	0.04	(0.18)
short tertiary education	0.11	(0.12)	-0.12	(0.30)
university	0.22	(0.09) *	-0.14	(0.22)
female	0.03	(0.06)	-0.06	(0.14)
UI recipient	0.22	(0.07) *	0.02	(0.19)
cash allowance recipient	0.21	(0.11) *	0.34	(0.23)
from Eastern Europe	-0.75	(0.15) *	0.39	(0.26)
from Africa or Asia	-0.95	(0.15) *	0.19	(0.30)
log(hourly wage)	-0.01	(0.04)	0.17	(0.09)
experience in occupation (dummy)	0.20	(0.06) *	0.08	(0.14)
education in occupation (dummy)	0.22	(0.05) *	0.10	(0.14)

The Table continues on next page.

Table 5 (Cont.) Estimation results for the training effect on the transition rate to work when real time in programs are included and the training effect works from the moment of entering (instead of leaving) AMU. Standard errors in parentheses. The superindex * denotes significance at 5% level (only for elements in β_i and λ_i (with i = u, p) and Δ).

professional and technical work	-0.11	(0.09)	0.03	(0.22)
health, nursing and social work	0.07	(0.09)	-0.34	(0.25)
adm., managerial and clerical work	-0.38	(0.09) *	0.32	(0.23)
sales	-0.26	(0.09) *	0.01	(0.22)
agriculture and mining	-0.06	(0.10)	0.09	(0.24)
services (incl. non categorized occ.)	-0.17	(0.08) *	-0.09	(0.20)
large city (dummy)	-0.07	(0.05)	-0.26	(0.12) *
needs guidance (dummy)	-0.54	(0.11) *	0.51	(0.20) *
willing to move (dummy)	0.02	(0.07)	0.18	(0.16)
accepts part time work (dummy)	0.16	(0.10) *	-0.46	(0.33)
1994	0.22	(0.07) *	-0.15	(0.18)
1995	0.17	(0.07) *	-0.17	(0.18)
1996	0.21	(0.08) *	-0.63	(0.22) *
1997	0.46	(0.08) *	-0.63	(0.22) *
1998	0.52	(0.08) *	-0.59	(0.22) *
1999	0.81	(0.09) *	-1.39	(0.33) *
2000	0.84	(0.16) *	-1.03	(0.69) *

The Table continues on next page.

Table 5 (Cont.) Estimation results for the training effect on the transition rate to work when real time in programs are included and the training effect works from the moment of entering (instead of leaving) AMU. Standard errors in parentheses. The superindex * denotes significance at 5% level (only for elements in β_i and λ_i (with i = u, p) and Δ).

Duration dependence				
λ_{i2}	0.08	(0.05)	-0.62	(0.15) *
λ_{i3}	-0.05	(0.07)	-0.72	(0.17) *
λ_{i4}	0.03	(0.08)	-0.70	(0.19) *
λ_{i5}	0.11	(0.09)	-0.61	(0.20) *
λ_{i6}	-0.11	(0.11)	-5.52	(0.19) *
λ_{i7}	-0.06	(0.12)		
λ_{i8}	-0.19	(0.09) *		
Unobserved heterogeneity				
log (v_1)	-5.74	(0.12)		
log (v_2)	-7.09	(0.14)		
log (v_3)			-8.50	(0.97)
log (v_4)			-6.76	(0.40)
q_{13}		-0.01	(0.50)	
q_{14}		2.20	(2.06)	
q_{23}		3.63	(11.81)	
log likelihood value		-22781.8		
number of individuals		5010		

6 Conclusions

After participation at an AMU vocational training course, the individual's transition rate from unemployment to employment is significantly and substantially higher than it would have been if the individual had not participated. This effect is largest during the first few weeks after exiting the course. However, when we take the time spent *within* the program into account as well, then the net effect on the individual's unemployment duration is about zero. Thus, from this point of view, the program does not appear to be cost-effective.

The results are consistent with the view that AMU vocational training shifts (part of) the burden of skill improvement, screening effort and search effort from employers to the state. We find that this does not primarily affect the unemployed individuals' time out of work. It is an open question whether on an aggregate level this policy repairs market failures or reduces variation in individual outcomes, and whether this would makes the policy socially effective. A comprehensive cost-benefits analysis has to take account of effects on post-unemployment wages and the duration of subsequent employment, but the available evidence suggests that these effects are insignificant (see earlier references and Korpi, 1994).

There are some topics for further research. First, it would be interesting to shed more light on the reason for why the effect on the individual transition rate to work is mainly short-run. It may reflect extra search effort during the course or stigmatization of workers who do not find a job within one month after finishing training. Alternatively, individual treatment effects are heterogeneous. The heterogeneity may be due to heterogeneity of individual characteristics or due to heterogeneity of characteristics of the training course. In future work we aim to distinguish between these explanations by exploiting additional data information and estimating richer models.

Secondly, in reality, participation in other programs may be affected by unobserved determinants that are related to the unobserved determinants of entry into AMU training and exit to work. In addition, one typically cannot participate in multiple programs at the same time. As a result, individuals who do not enter AMU training may flow into other programs at a relatively high rate. If participation in the latter programs has a positive effect on exit to work then our results under-estimate the effect of AMU training on exit to work. In such cases the analysis should include participation in other programs. This of course increases the complexity of the model and the estimation burden.

References

Abbring, J.H. and G.J. van den Berg (2000), The nonparametric identification of treatment effects in duration models, Working paper, Free University Amsterdam, Amsterdam.

Abbring, J.H., G.J. van den Berg, and J.C. van Ours (1997), The effect of unemployment insurance sanctions on the transition rate from unemployment to employment, Working paper, Tinbergen Institute, Amsterdam.

AMS (1997), Employment training, Working paper, AMS, Solna.

AMS (1998), Labour market policy programmes, Working paper, AMS, Solna.

AMS (2001), Arbetsmarknadspolitiska program. Årsrapport 2000 (English transl.: Yearly report on labour market policy programmes, Prora 2001:2, AMS, Solna (in Swedish).

Björklund, A. (1993), The Swedish experience, in: Measuring labour market measures, Ministry of Labour, Denmark, Copenhagen.

Bonnal, L., D. Fougère, and A. Sérandon (1997), Evaluating the impact of French employment policies on individual labour market histories, Review of Economic Studies, 64, 683-713.

Brännäs, K. (2000), Estimation in a duration model for evaluating educational programs, Working paper, Umeå University, Umeå.

Carling, K., P.A. Edin, A. Harkman, and B. Holmlund (1996), Unemployment duration, unemployment benefits and labor market programs in Sweden, Journal of Public Economics, 59, 313-334.

Carling, K., B. Holmlund, and A. Vejsiu (2001), Do benefit cuts boost job findings? Swedish evidence from the 1990s, Economic Journal, 111, forthcoming.

Carling, K. and K. Richardson (2001), The relative efficiency of labor market programs: Swedish experience from the 1990s, Working paper, IFAU, Uppsala.

Dahlberg, M. and A. Forslund (1999), Direct displacement effects of labour market programmes: The case of Sweden, Working paper, IFAU, Uppsala.

Eberwein, C., J.C. Ham, and R.J. LaLonde (1997), The impact of being offered and receiving classroom training on the employment histories of disadvantaged women: Evidence from experimental data, Review of Economic Studies, 64, 655-682.

Edin, P.A., A. Harkman, B. Holmlund, and H. Söderberg (1998), Escaping long-term unemployment in Sweden, Working paper, Uppsala University, Uppsala.

Edin, P.A. and B. Holmlund (1991), Unemployment, vacancies and labour market programmes: Swedish evidence, in: F. Padoa Schioppa, editor, Mismatch and Labour Mobility, Cambridge University Press, Cambridge.

Eriksson, M. (1997a), Comparison of compensatory and non-compensatory models for selection into labour market training, Umeå Economic Studies:437, Umeå University, Umeå.

Eriksson, M. (1997b), Placement of unemployed into labour market programs: a quasi-experimental study, Umeå Economic Studies:439, Umeå University, Umeå.

Fay, R.G. (1996), Enhancing the effectiveness of active labour market policies: Evidence from programme evaluations in OECD countries, Working paper, OECD, Paris.

Gritz, R.M. (1993), The impact of training on the frequency and duration of employment, Journal of Econometrics, 57, 21-51.

Harkman, A. and A. Johansson (1999), Training or subsidized jobs - what works?, Working paper, AMS, Solna.

Korpi, T. (1994), Escaping unemployment, Working paper, SOFI, Stockholm.

Lancaster, T. (1990), The Econometric Analysis of Transition Data, Cambridge University Press, Cambridge.

Larsson, L. (2000), Evaluation of Swedish youth labour market programmes, Working paper, IFAU, Uppsala.

Regnér, H. (2002), A nonexperimental evaluation of training programs for the unemployed in Sweden, forthcoming *i*n Labour Economics.

Regnér, H. (1997), *Training at the job and training for a new job: Two Swedish Studies*, Dissertation Series no 29, Swedish Institute for Social Research, Stockholm University, Stockholm.

Ridder, G. (1990), The non-parametric identification of generalized accelerated failure-time models, Review of Economic Studies, 57, 167-182.

Van den Berg, G.J. (2001), Duration models: Specification, identification, and multiple durations, in: J.J. Heckman and E. Leamer, editors, Handbook of Econometrics, Volume V, North Holland, Amsterdam.

Van den Berg, G.J., A. Holm, and J.C. van Ours (2002), Do stepping-stone jobs exist? Early career paths in the medical profession, Journal of Population Economics, forthcoming.

Zettermark, Å., H. Regnér, M. Aspgren, and J. Ericson (2000), Efficient training programmes for the unemployed, Ministry Publication Series 2000:38, Ministry of Industry, Elanders Gotab: Stockholm. (In Swedish, with an English summary.)

Appendix

We take the joint distribution of the unobserved heterogeneity terms v_u and v_p to be bivariate discrete with two unrestricted mass point locations for each term. Let v_1, v_2, v_3 and v_4 denote the points of support of v_u and v_p, respectively (note that v_u and v_p are random variables whereas $v_1,..,v_4$ are realizations). The associated probabilities are denoted as $p_{ij} :=$ $\Pr(v_u=v_i,v_p=v_j)$ with i=1,2 and j=3,4, and with $p_{24}=1-p_{13}-p_{14}-p_{23}$. The covariance of v_u and v_p equals

$$\mathrm{cov}\,(v_u,v_p)=(p_{13}p_{24}-p_{14}p_{23})\cdot(v_1-v_2)\cdot(v_3-v_4)$$

It is easy to show that v_u and v_p are independent if and only if $\mathrm{cov}(v_u,v_p)=0$.

In the estimation procedure we actually estimate the transformed probabilities q_{ij} which are implicitly defined as logistic versions of p_{ij} :

$$p_{ij} = \exp(q_{ij}) \,/\, \Sigma \exp(q_{i^* j^*})$$

Because the p_{ij} sum to one, we normalize by taking $q_{24}=0$. There is a one-to-one mapping between p_{13}, p_{14} and p_{23} on [0,1] and q_{13}, q_{14} and q_{23} on $(-\infty,\infty)$, so estimating the q_{ij} instead of the p_{ij} has the advantage that no boundary restrictions have to be imposed on the parameter space. Moreover, conditional on $v_1 \neq v_2$ and $v_3 \neq v_4$, there holds that $\mathrm{corr}(v_u,v_p)=0$ if and only if $q_{23}=q_{13}-q_{14}$.

Table A1. Estimation results for the training effect on the transition rate to work when we impose absence of unobserved heterogeneity. Standard errors in parentheses. The superindex * denotes significance at 5% level (only for elements in β_i and λ_i (with i = u, p) and δ).

	To work, θ_u	
AMU effect		
δ	0.57	(0.05) *
Individual characteristics		
log(age)	-0.50	(0.07) *
short senior high school	0.19	(0.04) *
senior high school	0.10	(0.05) *
short tertiary education	0.18	(0.07) *
university	0.38	(0.05) *
female	0.08	(0.04) *
UI recipient	0.65	(0.05) *
cash allowance recipient	0.61	(0.07) *
from Eastern Europe	-0.15	(0.09) *
from Africa or Asia	-0.39	(0.09) *
log(hourly wage)	0.08	(0.02) *
experience in occupation (dummy)	0.44	(0.04) *
education in occupation (dummy)	0.35	(0.04) *

The Table continues on next page.

Table A1 (Cont.) Estimation results for the training effect on the transition rate to work when we impose absence of unobserved heterogeneity. Standard errors in parentheses. The superindex * denotes significance at 5% level (only for elements in β_i and λ_i (with i = u, p) and δ).

professional and technical work	0.10	(0.06) *
health, nursing and social work	0.25	(0.06) *
adm., managerial and clerical work	-0.18	(0.06) *
sales	-0.01	(0.06)
agriculture and mining	0.22	(0.06) *
services (incl. non categorized occ.)	0.19	(0.05) *
large city (dummy)	0.04	(0.03)
needs guidance (dummy)	-0.20	(0.07) *
willing to move (dummy)	0.13	(0.04) *
accepts part time work (dummy)	0.23	(0.06) *
1994	0.52	(0.05) *
1995	0.47	(0.05) *
1996	0.51	(0.06) *
1997	0.70	(0.06) *
1998	0.71	(0.06) *
1999	0.91	(0.06) *
2000	0.96	(0.10) *

The Table continues on next page.

Table A1 (Cont.) Estimation results for the training effect on the transition rate to work when we impose absence of unobserved heterogeneity. Standard errors in parentheses. The superindex * denotes significance at 5% level (only for elements in β_i and λ_i (with i = u, p) and δ).

Duration dependence		
λ_{i1}	-7.46	(0.05) *
λ_{i2}	-7.42	(0.05) *
λ_{i3}	-7.56	(0.06) *
λ_{i4}	-7.58	(0.07) *
λ_{i5}	-7.72	(0.07) *
λ_{i6}	-8.00	(0.09) *
λ_{i7}	-8.02	(0.10) *
λ_{i8}	-7.99	(0.11) *
log likelihood value	-17413.1	
number of individuals	5010	